The rapidly evolving field of plant centromeres

Anne E Hall[1], Kevin C Keith[1], Sarah E Hall[2], Gregory P Copenhaver[3] and Daphne Preuss[1,4]

Meiotic and mitotic chromosome segregation are highly conserved in eukaryotic organisms, yet centromeres — the chromosomal sites that mediate segregation — evolve extremely rapidly. Plant centromeres have DNA elements that are shared across species, yet they diverge rapidly through large- and small-scale changes. Over evolutionary time-scales, centromeres migrate to non-centromeric regions and, in plants, heterochromatic knobs can acquire centromere activity. Discerning the functional significance of these changes will require comparative analyses of closely related species. Combined with functional assays, continued efforts in plant genomics will uncover key DNA elements that allow centromeres to retain their role in chromosome segregation while allowing rapid evolution.

Addresses
[1]Department of Molecular Genetics and Cell Biology, The University of Chicago, Chicago, Illinois 60637, USA
[2]Committee on Genetics, The University of Chicago, Chicago, Illinois 60637, USA
[3]Department of Biology and the Carolina Center for Genome Sciences, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA
[4]Howard Hughes Medical Institute, The University of Chicago, Chicago, Illinois 60637, USA
[4]e-mail: dpreuss@midway.uchicago.edu

Current Opinion in Plant Biology 2004, **7**:108–114

This review comes from a themed issue on
Genome studies and molecular genetics
Edited by Joseph R Ecker and Doug Cook

1369-5266/$ – see front matter

DOI 10.1016/j.pbi.2004.01.008

Abbreviations

BAC	bacterial artificial chromosome
CEN4	centromere4
CENP-A	centromere protein-A
CRM	centromere-specific retrotransposon of maize
CRR	centromere-specific retrotransposon of rice
FISH	fluorescent *in situ* hybridization

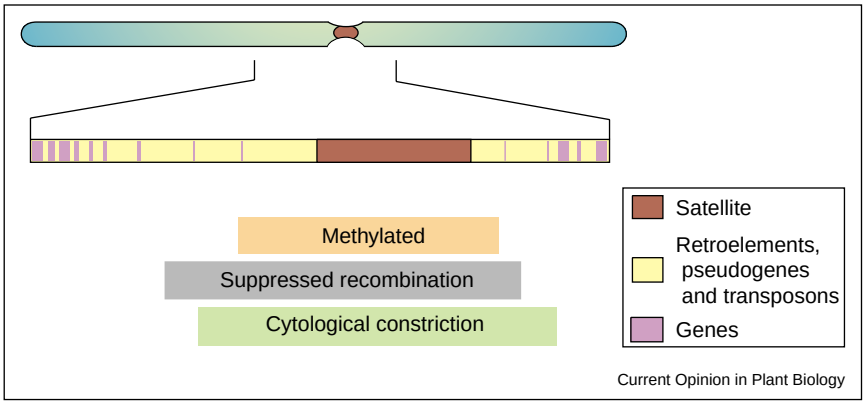
Introduction

Higher eukaryotic centromeres contain several megabases (Mb) of densely methylated, highly repetitive, heterochromatic DNA that nucleates kinetochores, promotes sister chromatid cohesion, and suppresses recombination. Although centromeric regions mediate these processes, the exact boundaries of each functional domain remain largely undefined. The innermost centromere region consists of tandemly repeated satellite DNA, flanked by DNA that is rich in middle repetitive elements, including transposons, retroelements and pseudogenes (Figure 1). The middle repetitive DNA is often referred to as the pericentromere, although domains in this region may contribute to centromere activity. Although centromere functions are highly conserved, there is a notable lack of sequence homology among the centromeres of distantly related species, highlighting the extremely rapid rate of centromere DNA evolution. Plants offer an exciting opportunity to discern the relationship between centromere DNA sequence and function. Genome projects have collected more centromere DNA sequence from plants than from other higher eukaryotes. In addition, plant genomes allow manipulation of chromosome number, permitting functional analysis with artificial chromosomes, and the abundance of closely related plant species facilitates the identification of evolutionarily conserved sequences. In this review, we examine the growing body of research on plant centromeres, discuss current models of centromere sequence evolution and consider assays to define functional components of centromere DNA.

Plant centromere DNA content

Satellite and middle repetitive centromere DNA has been characterized in several plant species (Table 1). These sequences provide probes for identifying centromere clones from large-insert genomic libraries. Such efforts, along with whole-genome shotgun sequencing, are generating an extensive database of centromere sequences, particularly from the grasses and crucifers. The centromeres of *Arabidopsis* are the most thoroughly characterized, with assembled sequences extending from the chromosome arms to the satellites of all five chromosomes [1,2]. *Arabidopsis* centromeres contain 2.8–4 Mb tracts of tandemly repeated 178 bp satellites, occasionally interrupted by insertions of *Athila*, a Ty3/*gypsy* retroelement [1–5]. The middle repetitive regions that flank the satellites also contain *Athila* and other retroelements, along with 5S ribosomal DNA (rDNA) tracts, transposable elements, and pseudogenes [1–5]. Despite the heterochromatic environment, the middle repetitive regions contain more than 200 expressed genes, at a density that is about one tenth of that on the chromosome arms [2]. Intriguingly, tracts of hemi-methylated cytosine residues are present throughout the *Arabidopsis* centromeres [6]. Whether these modifications contribute to centromere function or result from the impact of centromere structure on DNA-methylation mechanisms remains unresolved.

Figure 1



The higher eukaryotic centromere. Cytologically, centromeres are chromosomal constrictions (green) that nucleate the kinetochore, and mediate mitosis and meiosis. Centromeric regions encompass large arrays of satellite DNA (red) flanked by middle repetitive DNA (yellow); they often contain highly methylated heterochromatin (orange) that is recombinationally suppressed (gray). Although each of these features overlap, their precise boundaries vary from chromosome to chromosome and between species. At the edge of the centromere, middle repetitive DNA gradually transitions into the gene-rich chromosome arms (blue).

Although efforts to sequence the rice and maize centromeres are not complete, they have yielded intriguing data that are useful for comparative analysis. In rice, a contig of 1.16 Mb from centromere 4 (CEN4) revealed tracts of the CentO satellite (155 bp or 164 bp) interspersed with the Ty3/*gypsy* retroelement CRR (for centromere-specific retrotransposon of rice), as well as flanking regions that are rich in transposons, retroelements, and pseudogenes [7,8••]. In maize, two sequenced centromere bacterial artificial chromosomes (BACs) consist largely of CentC satellite (156 bp) and the centromere-specific retrotransposon of maize (CRM), a retroelement with homology to CRR [9•,10]. CentO and CentC are similar, yet neither shows homology to *Arabidopsis* satellites [8••]. In maize, heterochromatic knobs that can serve as alternatives to the conventional centromeres have also been characterized. These knobs contain two satellites, measuring 180 bp and 350 bp, that differ from CentC [11]. Supernumerary maize

Table 1

Current knowledge of centromere sequence, content and organization among several model organisms, species in the Brassicaceae family, and crops.

Organism	Sequence type[a]	Size of major satellite	Example of recent work[b]	Reference(s)
Human	AS	171 bp	365 kb of Chromosome X sequenced	[45]
Mouse	AS	Major 234 bp, minor 120 bp	Examined evolution of pericentric DNA	[41••]
Fruit fly	AS	5 bp	80 kb of CEN sequenced	[46]
Brassicaceae family				
Arabidopsis thaliana	AS	178 bp	Nearly complete sequence for all five CENS	[1–5]
Arabidopsis arenosa	US	178 bp	Satellites mapped by FISH	[47]
Arabidopsis suecica	US	178 bp	Satellites mapped by FISH	[47]
Olimarabidopsis pumila	US	168 bp, 178 bp	Satellites mapped by FISH	[19]
Brassica napus	US	176 bp	Satellites mapped by FISH	[48]
Crops				
Maize	AS	156 bp	Two CEN BACs sequenced	[9•]
Rice	AS	155 bp	1.16 Mb contig of CEN4 sequenced	[7]
Barley	US	Not known	18 kb of BAC clone sequenced	[49]
Beet	US	158–160 bp	Satellites/transposons mapped by FISH	[50]
Soybean	US	120 bp	Satellites mapped by FISH	[51]
Sorghum	US	140 bp	Satellites mapped by FISH	[52]
Radish	US	177 bp	Satellites mapped by FISH	[53]
Wheat	US	537 bp	Satellites/transposons mapped by FISH	[54]

[a]AS, assembled sequence contigs; US, unassembled sequence. These sequences were predominantly generated by restriction digestions or PCR with degenerate primers. [b]Examples of recent work indicate the methodology employed and the type of data generated for the listed organism.

B chromosomes have a third type of centromere, containing a 1.4 kb tandemly repeated sequence with homology to knobs and a region from CEN4 [12,13]. Discerning the role of the different satellite classes and the effects of interspersed DNA elements will require functional assays.

Centromere evolution

Cytological and sequence analyses of centromeres from different *Arabidopsis* ecotypes have illustrated their dynamic nature; significant large-scale changes that do not disrupt centromere function have occurred over timescales of just a few million years. For examples, several tandem copies of the mitochondrial genome have been integrated into CEN2 [14], an inversion involving CEN4 produced a heterochromatic knob and moved a gene-rich region into the heterochromatin [15,16], and a tract of 5S rDNA was inserted into CEN3 [17]. Whether such changes generally characterize the evolution of plant centromeres remains to be determined, but their occurrence on three of the five *Arabidopsis* centromeres warrants further analysis. Furthermore, although *Arabidopsis* chromosome arms have large tracts of DNA that reflect ancient genome duplications, similar segments have not been found in most of the centromeres, suggesting that evolutionary mechanisms differ between euchromatic regions and the centromere [1].

In addition to undergoing large-scale rearrangements, centromeres also evolve rapidly at the nucleotide level. Satellites undergo rapid change even within a species, and ecotype-specific satellite variants have been identified in *Arabidopsis thaliana* [18•]. Interestingly, the nucleotide substitutions that define variants are nonrandom, producing highly conserved and variable regions. Although such regions imply that the nucleotide substitutions have functional significance, investigations of satellites from the closely related species *Arabidopsis arenosa* and *Olimarabidopsis pumila* did not reveal conserved regions within these closely related genera [19]. By contrast, CentO and CentC maintain similar 5′ and 3′ ends despite a divergence of 50–70 million years [8••,20]. In maize, analysis of satellite polymorphisms made it possible to define at least 18 different satellite classes that are arranged in higher-order repeat arrays [9•]. Individual BACs contain a subset of the different satellite variants, indicating that satellites undergo local variation that is distinct from variation within the genome as a whole [9•,21]. This finding is consistent with observations of human centromeres, in which satellite variants tend to cluster non-randomly because of an intrachromosomal gene-conversion mechanism [21].

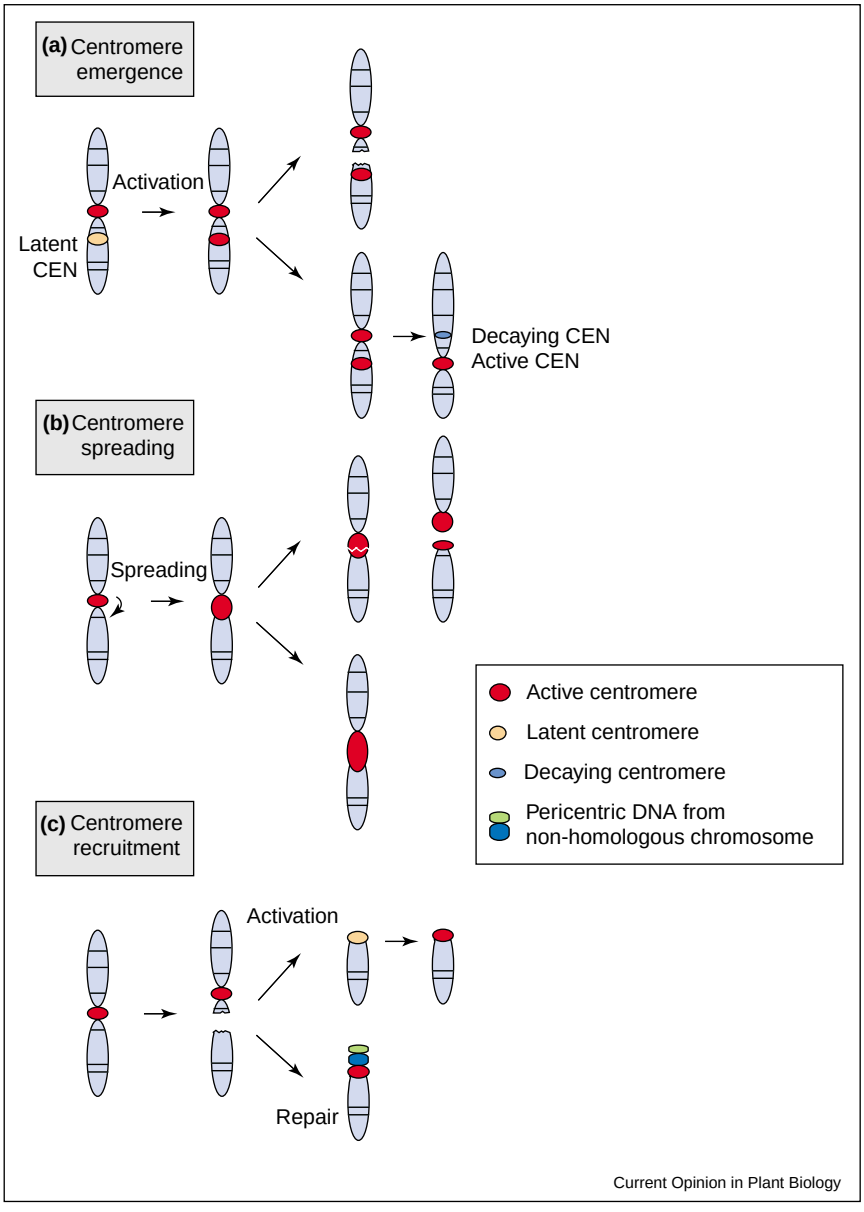
Several hypotheses have surfaced to explain how a genome can maintain satellite homogeneity while still allowing rapid evolution [22••]. The library hypothesis predicts that genomes contain sets of satellite variants in differing abundance. New satellites continually arise by mutation and are propagated within a genome by genetic drift. These satellites can be homogenized through unequal crossover, which would provide an explanation for chromosome-specific variation and the formation of higher-order satellite arrays [23–27]. Another model suggests that satellite evolution is driven by the selection and coevolution of satellites and centromere-binding proteins, rather than by random genetic drift. This requires that certain centromere satellite variants confer a selective advantage upon meiotic cells (i.e. eggs and sperm), a process known as meiotic drive. Such positive selection could result from a preferential interaction with centromere protein-A (CENP-A), a histone H3 variant that is incorporated into the nucleosomes that specifically bind to centromere DNA in plants, humans, yeast, and *Drosophila* [28]. If this is the case, even a slight advantage in satellite–CENP-A interactions could quickly result in genomic fixation of satellite arrays. The rapid adaptive evolution of CENP-A observed in both *Drosophila* and *Arabidopsis* is consistent with this interpretation [29,30•].

Meiotic recombination is dramatically suppressed in centromere regions; the effects of this suppression on centromere DNA content and evolution are not yet clear. Mathematical models predict that low levels of homologous recombination cause the accumulation of repetitive elements [31,32], but a study that considered only non-coding *Arabidopsis* DNA found no correlation between meiotic recombination frequencies and the distribution of most transposable elements [33]. Importantly, two superfamilies of *gypsy*-like elements (*Athila* and *CACTA*) were excluded from the study because of they are clustered within and around centromeres. On the basis of this work, Wright *et al*. [33] suggested that the apparent centromeric abundance of transposons is not due to suppressed recombination but rather results from ample non-coding DNA targets; insertions into chromosome arms, by contrast, are often deleterious because of the abundance of genes. Although there are few homologous recombination events in centromere regions, ectopic recombination may play a role in the evolution of these regions. For example, evidence from human and *Arabidopsis* centromere regions shows that ectopic recombination occurs at a relatively high frequency [34]. Interestingly, the centromere regions of wheat exhibit fewer local duplications than occur in the chromosome arms, resulting in a preservation of micro-synteny [35].

Acquisition of centromere function

Chromosome-wide studies of gene order have shown that non-centromeric regions can acquire centromere activity *de novo* (Figure 2a). Fluorescent *in situ* hybridization (FISH) mapping of closely related genomes shows that, in some cases, gene marker order remains constant while centromere positions migrate. This is the case when the X chromosomes of humans and lemur species are compared [36], and when the phylogeny of primate chromosome 6 is

Figure 2



Acquisition of centromere function. **(a)** Chromosomes can contain latent centromeres that become activated under some conditions (neocentromeres). If the original centromere also remains active, chromosome breakage can occur. Alternatively, as the endogenous centromere decays, centromere DNA structure becomes established at the new location. **(b)** When centromere protein components spread to linked chromosomal regions, they can form a larger centromere. If chromosome breakage occurs, two new centromeres may be formed. **(c)** Chromosome fission events can activate a latent centromere or can be repaired with ectopic centromeric or pericentromeric DNA from non-homologous chromosomes.

reconstructed [37••]. In both cases, marker order was conserved among multiple species, and neither transposition nor inversion events could account for centromere relocation. The migration of centromeres over evolutionary time scales could result from neocentromere formation — a process whereby ectopic centromeres are activated in previously non-centromeric locations (Figure 2a). Neocentromere formation in human cell lines is triggered by the loss of an endogenous centromere, coupled with a strong selection for inheritance of the DNA fragment (Figure 2a). These new centromeres often lack satellite DNA, suggesting that associated proteins and other epigenetic factors may be more important for centromere activity than the primary DNA sequence [38].

Although the topic of centromere emergence has not been addressed in most studies of plant genome synteny, satellite-rich heterochromatic knobs can often function as

neocentromeres. Maize knobs can promote the preferential inheritance (or meiotic drive) of particular chromosomes. The two different knob-specific maize satellites differ in their capacity to promote chromosome segregation [11], and several *trans*-acting factors that affect knob activation have been identified [39]. At least four distinct genetic functions mapping to Abnormal chromosome 10 (Ab10) play a role in the segregation of knobs; in some cases, these have been characterized through screens for the suppression of meiotic drive [39].

Two other models can account for the appearance of centromere functions in non-centromeric DNA. The first, centromere spreading, has been described in *Drosophila*. In this model, the migration of centromere functions to adjacent DNA is mediated by the deposition of centromere-binding factors (Figure 2b; [40]). Alternatively, centromeric and pericentromeric DNA can be recruited to repair the ends of broken chromosomes (Figure 2c). Support for this possibility is based on the discovery of large tracts of pericentromeric duplicated DNA (termed duplicons) that are thought to arise during the repair process. For example, mouse chromosomes 5 and 6 appear to be chromosome fission products, repaired with fragments of centromeric and pericentromeric DNA from other chromosomes [41••]. Intriguingly, the duplications produced by these repair processes often appear to become hot-spots for further duplications and rearrangements.

Studies of centromere function

Current approaches for delineating the functional components of centromeres include: co-immunoprecipitating centromere proteins with their associated DNA fragments, measuring transmission rates of non-essential chromosome fragments with truncated centromeres, and assessing transmission rates of *in-vitro*-assembled artificial chromosomes. In *Arabidopsis*, antibodies to the CENP-A homolog co-immunoprecipitate 178 bp satellite repeats [42]; whereas in maize, CenH3 binds to both CentC and CRM [43•]. These studies implicate satellites and retroelements as important DNA sequences for centromere function, but do not rule out a potential role for other DNA sequences in this function.

An important method for defining the functional domains of plant centromeres has relied on truncations of maize B chromosomes, which do not provide cellular functions and consequently can be lost without deleterious effects. Analysis of 25 B chromosomes that had truncated centromeres revealed a direct correlation between centromere size and meiotic transmission [13]. Centromere sizes of less than 1000 kb, or derivatives missing an internal 370 kb fragment, were lost at significantly higher rates. All of the derivatives were stable for multiple mitotic divisions, indicating that there may be genetic elements that distinguish meiosis and mitosis.

Assays that rely on artificial chromosome transmission would greatly facilitate the characterization of the functional components of plant centromeres. Such experiments have already been carried out using human artificial chromosomes: α-satellite arrays of 65 kb from chromosome 17 can confer mitotic centromere function, whereas chromosomes with similar arrays from the Y chromosome are lost during mitosis [44]. Interestingly, the Y chromosome arrays do not contain sites that bind to CENP-B, a centromere protein whose function is not completely understood. Future research using plant artificial chromosomes will enable analysis of the role of centromere length, satellite composition, and centromere-binding proteins during both meiosis and mitosis. Performing such experiments across different species will clarify the functional significance of the evolutionary changes that contribute to centromere diversity.

Conclusions

Plants hold great promise for clarifying the DNA elements that comprise centromere structure, determining the factors required for centromere function, and understanding the mechanisms that drive rapid centromere evolution. Limited comparisons of centromere DNA from species within the grass and crucifer families indicate that some DNA elements are conserved among closely related species. Further analysis of centromeres from many closely related plants will elucidate conserved functional elements. Plant artificial chromosomes will be useful for testing centromere function and identifying critical centromere elements. Functional assays and comparisons of centromere DNA sequence have far reaching implications, including improving our understanding of centromere evolution and its influence on speciation, allowing analysis of *cis*-and *trans*-acting factors that are important for proper inheritance of chromosomes, and permitting enhanced plant transformation using artificial chromosomes.

Acknowledgements

We thank members of the Preuss laboratory for helpful suggestions and critical reading of the manuscript. This work was supported in part by the National Science Foundation (grant to AEH and DP), a National Institute of Health National Research Service Award (to KCK), and the Howard Hughes Medical Institute.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. *Arabidopsis* Genome Initiative: **Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana***. *Nature* 2000, **408**:796-815.

2. 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.

3. Kumekawa N, Hosouchi T, Tsuruoka H, Kotani H: **The size and sequence organization of the centromeric region of *Arabidopsis thaliana* chromosome 5**. *DNA Res* 2000, **7**:315-321.

4. Kumekawa N, Hosouchi T, Tsuruoka H, Kotani H: **The size and sequence organization of the centromeric region of *Arabidopsis thaliana* chromosome 4**. *DNA Res* 2001, **8**:285-290.

5. Hosouchi T, Kumekawa N, Tsuruoka H, Kotani H: **Physical map-based sizes of the centromeric regions of *Arabidopsis thaliana* chromosomes 1, 2, and 3**. *DNA Res* 2002, **9**:117-121.

6. Luo S, Preuss D: **Strand-biased DNA methylation associated with centromeric regions in *Arabidopsis***. *Proc Natl Acad Sci USA* 2003, **100**:11133-11138.

7. Feng Q, Zhang Y, Hao P, Wang S, Fu G, Huang Y, Li Y, Zhu J, Liu Y, Hu X *et al.*: **Sequence and analysis of rice chromosome 4**. *Nature* 2002, **420**:316-320.

8. •• Cheng Z, Dong F, Langdon T, Ouyang S, Buell CR, Gu M, Blattner FR, Jiang J: **Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon**. *Plant Cell* 2002, **14**:1691-1704.
Using FISH analysis, the authors determine the centromeric distribution and organization of rice CentO satellites and CRR retroelements, including a quantification of the satellite arrays on all 12 chromosomes. Sequence alignments reveal homology between CentO and maize CentC, which is significant given the evolutionary separation of rice and maize. Finally, telotrisomic chromosomes derived from misdivisions indicate that chromosome breakpoints map within CentO arrays, providing strong evidence that CentO is a functional component of rice centromeres.

9. • Nagaki K, Song J, Stupar RM, Parokonny AS, Yuan Q, Ouyang S, Liu J, Hsiao J, Jones KM, Dawe RK *et al.*: **Molecular and cytological analyses of large tracks of centromeric DNA reveal the structure and evolutionary dynamics of maize centromeres**. *Genetics* 2003, **163**:759-770.
CentC and a centromere-specific retrotransposon were used to identify two centromere BAC clones from a maize genomic library. Sequencing indicated that both clones consist of short CentC arrays interspersed with retrotransposons. A detailed description of the CentC and retrotransposon sequences is provided, which includes alignments and phylogenetic trees generated with centromere-specific elements from other grasses.

10. Ananiev EV, Phillips RL, Rines HW: **Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions**. *Proc Natl Acad Sci USA* 1998, **95**:13073-13078.

11. Hiatt EN, Kentner EK, Dawe RK: **Independently regulated neocentromere activity of two classes of tandem repeat arrays**. *Plant Cell* 2002, **14**:407-420.

12. Page BT, Wanous MK, Birchler JA: **Characterization of a maize chromosome 4 centromeric sequence: evidence for an evolutionary relationship with the B chromosome centromere**. *Genetics* 2001, **159**:291-302.

13. Kaszas E, Birchler JA: **Meiotic transmission rates correlate with physical features of rearranged centromeres in maize**. *Genetics* 1998, **150**:1683-1692.

14. Stupar RM, Lilly JW, Town CD, Cheng Z, Kaul S, Buell CR, Jiang J: **Complex mtDNA constitutes an approximate 620-kb insertion on *Arabidopsis thaliana* chromosome 2: implication of potential sequencing errors caused by large-unit repeats**. *Proc Natl Acad Sci USA* 2001, **98**:5099-5103.

15. Fransz PF, Armstrong S, de Jong JH, Parnell LD, van Drunen C, Dean C, Zabel P, Bisseling T, Jones GH: **Integrated cytogenetic map of chromosome arm 4S of *A. thaliana*: structural organization of heterochromatic knob and centromere region**. *Cell* 2000, **100**:367-376.

16. *Arabidopsis* Sequencing Consortium: **The complete sequence of a heterochromatic island from a higher eukaryote**. *Cell* 2000, **100**:377-386.

17. Fransz P, Armstrong S, Alonso-Blanco C, Fischer TC, Torres-Ruiz RA, Jones G: **Cytogenetics for the model system *Arabidopsis thaliana***. *Plant J* 1998, **13**:867-876.

18. • Hall SE, Kettler G, Preuss D: **Centromere satellites from *Arabidopsis* populations: maintenance of conserved and variable domains**. *Genome Res* 2003, **13**:195-205.
The authors of this paper describe the evolution of centromere satellite repeats in *Arabidopsis thaliana* ecotypes. Consensus sequences were derived for each of the ecotypes and the species, and the conserved and variable regions determined. The results presented here are indicative of the functionality of satellites and their rapid evolution even within a species.

19. Heslop-Harrison JS, Brandes A, Schwarzacher T: **Tandemly repeated DNA sequences and centromeric chromosomal regions of *Arabidopsis* species**. *Chromosome Res* 2003, **11**:241-253.

20. Wolfe KH, Gouy M, Yang YW, Sharp PM, Li WH: **Date of the monocot–dicot divergence estimated from chloroplast DNA sequence data**. *Proc Natl Acad Sci USA* 1989, **86**:6201-6205.

21. Schindelhauer D, Schwarz T: **Evidence for a fast, intrachromosomal conversion mechanism from mapping of nucleotide variants within a homogeneous alpha-satellite DNA array**. *Genome Res* 2002, **12**:1815-1826.

22. •• Ugarkovic D, Plohl M: **Variation in satellite DNA profiles — causes and effects**. *EMBO J* 2002, **21**:5955-5959.
This review describes recent advances in studies of satellite sequence evolution and satellite copy number. The authors discuss current hypotheses regarding satellite evolution and their relevance to speciation events.

23. Smith GP: **Evolution of repeated DNA sequences by unequal crossover**. *Science* 1976, **191**:528-535.

24. Dover G: **Molecular drive: a cohesive mode of species evolution**. *Nature* 1982, **299**:111-117.

25. Ohta T, Dover GA: **The cohesive population genetics of molecular drive**. *Genetics* 1984, **108**:501-521.

26. Elder JF Jr, Turner BJ: **Concerted evolution of repetitive DNA sequences in eukaryotes**. *Q Rev Biol* 1995, **70**:297-320.

27. Nijman IJ, Lenstra JA: **Mutation and recombination in cattle satellite DNA: a feedback model for the evolution of satellite DNA repeats**. *J Mol Evol* 2001, **52**:361-371.

28. Sullivan KF: **A solid foundation: functional specialization of centromeric chromatin**. *Curr Opin Genet Dev* 2001, **11**:182-188.

29. Malik HS, Vermaak D, Henikoff S: **Recurrent evolution of DNA-binding motifs in the *Drosophila* centromeric histone**. *Proc Natl Acad Sci USA* 2002, **99**:1449-1454.

30. • Talbert PB, Masuelli R, Tyagi AP, Comai L, Henikoff S: **Centromeric localization and adaptive evolution of an *Arabidopsis* histone H3 variant**. *Plant Cell* 2002, **14**:1053-1066.
The authors of this paper identify and analyze the *A. thaliana* homolog of CENP-A (HTR12). Sequencing of this gene in other *Arabidopsis* species showed that although the protein core is conserved, the tail has undergone more nonsynonymous changes than expected. The authors present a model for how satellite evolution is driven by the evolution of CENP-A.

31. Charlesworth B, Sniegowski P, Stephan W: **The evolutionary dynamics of repetitive DNA in eukaryotes**. *Nature* 1994, **371**:215-220.

32. Stephan W: **Recombination and the evolution of satellite DNA**. *Genet Res* 1986, **47**:167-174.

33. Wright SI, Agrawal N, Bureau TE: **Effects of recombination rate and gene density on transposable element distributions in *Arabidopsis thaliana***. *Genome Res* 2003, **13**:1897-1903.

34. Eichler EE, Sankoff D: **Structural dynamics of eukaryotic chromosome evolution**. *Science* 2003, **301**:793-797.

35. Akhunov ED, Akhunova AR, Linkiewicz AM, Dubcovsky J, Hummel D, Lazo G, Chao S, Anderson OD, David J, Qi L *et al.*: **Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates**. *Proc Natl Acad Sci USA* 2003, **100**:10836-10841.

36. Ventura M, Archidiacono N, Rocchi M: **Centromere emergence in evolution**. *Genome Res* 2001, **11**:595-599.

37. •• Eder V, Ventura M, Ianigro M, Teti M, Rocchi M, Archidiacono N: **Chromosome 6 phylogeny in primates and centromere repositioning**. *Mol Biol Evol* 2003, **20**:1506-1512.
The authors use comparative FISH mapping in primates to document a centromere-repositioning event on chromosome 6. They propose this centromere-repositioning occurred because of an emergence event, rather than a transposition or inversion event, that relocated the endogenous centromere. They conclude that centromere emergence is

actually fairly frequent in primates and, surprisingly, find that the repositioning process does not affect neighboring chromosomal sequences.

38. du Sart D, Cancilla MR, Earle E, Mao JI, Saffery R, Tainton KM, Kalitsis P, Martyn J, Barry AE, Choo KH: **A functional neo-centromere formed through activation of a latent human centromere and consisting of non-alpha-satellite DNA**. *Nat Genet* 1997, **16**:144-153.

39. Hiatt EN, Dawe RK: **Four loci on abnormal chromosome 10 contribute to meiotic drive in maize**. *Genetics* 2003, **164**:699-709.

40. Maggert KA, Karpen GH: **The activation of a neocentromere in *Drosophila* requires proximity to an endogenous centromere**. *Genetics* 2001, **158**:1615-1628.

41. •• Thomas JW, Schueler MG, Summers TJ, Blakesley RW, McDowell JC, Thomas PJ, Idol JR, Maduro VV, Lee-Lin SQ, Touchman JW *et al.*: **Pericentromeric duplications in the laboratory mouse**. *Genome Res* 2003, **13**:55-63.

Comparative mapping in humans, mice and rats identifies regions on several mouse chromosomes that contain pericentromeric duplications (duplicons), which appear to flank an ancient chromosome breakpoint. This is significant because sequence comparisons indicate that the pericentromeric duplicons are hot-spots for rearrangements and duplications. The authors propose these regions are crucial for genome evolution.

42. 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.

43. • Zhong CX, Marshall JB, Topp C, Mroczek R, Kato A, Nagaki K, Birchler JA, Jiang J, Dawe RK: **Centromeric retroelements and satellites interact with maize kinetochore protein CENH3**. *Plant Cell* 2002, **14**:2825-2836.

This group demonstrates that antibodies to maize CENH3, a centromere-specific histone that marks functional centromeres, co-immunoprecipitate both CentC satellite and CRM retroelements. This work is significant because it suggests that, in addition to satellite DNA, a conserved retroelement may be a component of the functional maize centromere.

44. Grimes BR, Rhoades AA, Willard HF: **Alpha-satellite DNA and vector composition influence rates of human artificial chromosome formation**. *Mol Ther* 2002, **5**:798-805.

45. Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HF: **Genomic and genetic definition of a functional human centromere**. *Science* 2001, **294**:109-115.

46. Sun X, Le HD, Wahlstrom JM, Karpen GH: **Sequence analysis of a functional *Drosophila* centromere**. *Genome Res* 2003, **13**:182-194.

47. Comai L, Tyagi AP, Lysak MA: **FISH analysis of meiosis in *Arabidopsis* allopolyploids**. *Chromosome Res* 2003, **11**:217-226.

48. Xia X, Selvaraj G, Bertrand H: **Structure and evolution of a highly repetitive DNA sequence from *Brassica napus***. *Plant Mol Biol* 1993, **21**:213-224.

49. Hudakova S, Michalek W, Presting GG, ten Hoopen R, dos Santos K, Jasencakova Z, Schubert I: **Sequence organization of barley centromeres**. *Nucleic Acids Res* 2001, **29**:5029-5035.

50. Gindullis F, Desel C, Galasso I, Schmidt T: **The large-scale organization of the centromeric region in *Beta* species**. *Genome Res* 2001, **11**:253-265.

51. Morgante M, Jurman I, Shi L, Zhu T, Keim P, Rafalski JA: **The STR120 satellite DNA of soybean: organization, evolution and chromosomal specificity**. *Chromosome Res* 1997, **5**:363-373.

52. Zwick MS, Islam-Faridi MN, Zhang HB, Hodnett GL, Gomez MI, Kim JS, Price HJ, Stelly DM: **Distribution and sequence analysis of the centromere-associated repetitive element CEN38 of *Sorghum bicolor* (Poaceae)**. *Am J Bot* 2000, **87**:1757-1764.

53. Grellet F, Delcasso D, Panabieres F, Delseny M: **Organization and evolution of a higher plant alphoid-like satellite DNA sequence**. *J Mol Biol* 1986, **187**:495-507.

54. Kishii M, Nagaki K, Tsujimoto H: **A tandem repetitive sequence located in the centromeric region of common wheat (*Triticum aestivum*) chromosomes**. *Chromosome Res* 2001, **9**:417-428.