In plants, as in all eukaryotes, centromeres are chromatin domains that govern the transmission of nuclear chromosomes to the next generation of cells/individuals. The DNA composition and sequence organization of centromeres has recently been elucidated for a few plant species. Although there is little sequence conservation among centromeres, they usually contain tandem repeats and retroelements. The occurrence of neocentromeres reinforces the idea that the positions of centromeres are determined epigenetically. In contrast to centromeric DNA, structural and transient kinetochoric proteins are highly conserved among eukaryotes. Candidate sequences have been identified for a dozen putative kinetochore protein homologues, and some have been localized to plant centromeres. The kinetochore protein CENH3, which substitutes histone H3 within centromeric nucleosomes, co-immunoprecipitates preferentially with centromeric sequences. The mechanism(s) of centromere assembly and the functional implication of (peri-)centromeric modifications of chromatin remain to be elucidated.

**DNA composition of plant centromeres**

In contrast to telomeres, centromeres are not specified by highly conserved DNA sequences. Centromeric sequences have been described for several eukaryotes. Except for the approximately 125-bp centromeres of budding yeast [5], whose functional importance was verified by mutation analyses, the functional importance of centromeric sequences is at least controversial. In *Vicia faba* and *Tradescantia paludosa* it has not yet been possible to detect centromere-specific repeats [6], whereas such sequences have been found in other plants. For instance, a 178-bp tandem repeat [7,8] together with intermingled portions of *Athila* retroelements [9] (mainly part of this retroelement’s long terminal repeats [LTRs] [10]), forms the centromere core (~3 Mbp) in *Arabidopsis thaliana* [11,12,13]. This tandem repeat differs significantly in sequence but not in length (178 ± 1 bp) between *A. thaliana* accessions [14]. A similar organization (i.e. retroelements embedded within short tandem repeat arrays) has been reported for *Drosophila* centromeres [15]. For cereals, two conserved centromere-specific repeats (CEREAL CENTROMERIC SEQUENCE1 [CCS1] [16] and Sau3A9 [17]) were reported and later found to represent parts of a Ty3/gypsy-like retroelement [18,19]. The sequence organization of barley centromeres has been elucidated by analysis of the 23-kb insert of a centromere-specific bacterial artificial chromosome (BAC) clone (BAC 7), which contained three apparently complete copies of the gypsy-like retroelement ‘*cereba*’ (~7 kb). Parts of the LTRs of *cereba* (~1 kb) correspond to CCS1 and parts of the integrase region of its polygene to Sau3A9. *Cereba*-elements, together with a G+C-rich satellite sequence (AGGGAG)_n, constitute the major DNA components of barley centromeres [19]. Each barley centromere contains around 200 *cereba*-elements, comprising at least 1.4 Mbp. *Cereba*-related centromeric retroelement (CR) sequences (including complete and truncated elements or solo-LTRs) have also been found within the centromeres of maize, sorghum, wheat, rice, rye, oats and *Aegilops squarrosa* [20–26,27,28], and even in...
dicotyledonous *Beta* species [29]. In most cases, these retroelements were not detectable along chromosome arms. They have probably occurred within the centromeres of all grass species since their phylogenetic separation from other monocots about 60 million years ago, are probably derived from a single transposon family, are interspersed irregularly between centromeric satellite sequences and are apparently still able to transpose [25,27,28].

In addition, *Ty1/copia*-like elements have been found in the centromeres of rye (in the polycene region *Bilby* [30]) and *Zingeria biebersteiniana* [31]. The satellite sequences that occur within the centromeres of most eukaryotes are usually species-specific and possibly represent structural homologues of the alphoid satellite of human centromeres. Nevertheless, some contain CENTROMERIC PROTEIN-B (CENP-B)-box-like sequences [8,32], which bind the conserved transposase-like protein CENP-B in human centromeres. These sequences show similarity in rice and maize [27*], although no conserved CENP-B protein has been found outside of mammals and fission yeast. The centromere of maize chromosome 4 also contains repeats that are present in the centromeres of B chromosomes and that are similar to parts of repeats from potentially neocentromeric knobs [33]. Centromeric satellites and retroelements are preferentially immuno-precipitated with a maize homologue of the human centromeric histone H3 variant CENP-A [34*]. All of the breakpoints from 14 centromere-misdivision events in rice were mapped within the centromeric satellite arrays [27*]. Robertsonian wheat-rye translocation chromosomes that had centromeres composed of similar proportions of wheat and rye centromeric DNA behave normally during mitosis and meiosis [35]. Maize mini-B-chromosomes that contained centromeric repeats of less than 1000 kb showed reduced meiotic transmission, those with centromeric repeats of less than 500 kb also showed reduced mitotic transmission [36]. All of these observations suggest that centromeric satellites and retroelements are functionally important. Barley telosomes and isochromosomes that lack *cereha* and centromeric satellite sequences are mitotically and meiotically stable [37], and possibly indicate the formation of neocentromeres that do not require the presence of such sequences.

Large-scale sequence organization of centromeres has been reported for *A. thaliana* (reviewed in [13*]), rice [24,27*], maize [28], barley [19] and *Beta* [38].

**Protein composition of plant kinetochores**

Contrary to centromeric DNA, structural and functional kinetochore proteins are highly conserved between yeast and metazoa. Almost fifty proteins are constitutively or transiently associated with *S. cerevisiae* centromeres (see also [39]), and more than 20 kinetochore proteins (KPs) are known to occur in human centromeres and neocentromeres [40]. The KPs of human neocentromeres are free of detectable centromeric alphoid satellite sequences [40]. The occurrence of neocentromeres (about 50 are known in humans [41]) supports the assumption that centromere formation is regulated epigenetically and, once established, remains stable [42,43].

Human CREST (calcinosisc, Raynaud’s phenomenon, esophageal dismotility, sclerodactyly, telangectasia) antisera from patients with a variant of sclerodermia have previously been used to label animal kinetochores [44]. Two CREST sera were the first reagents used to identify plant kinetochores [45,46]. Currently, the most promising method for the identification of plant KPs is to carry out similarity searches within plant genomic and expressed sequence tag (EST) databases for sequences that correspond to non-plant KPs. The plant KPs identified to date by sequence similarity/immunolocalization are listed in Table 1.

It has been suggested that the evolutionarily conserved centromere-specific histone-H3-like protein CENH3 (the human variant is CENP-A) determines the chromosomal position of kinetochore assembly and forms a link between the centromeric DNA and the proteinaceous kinetochore [47]. Indeed, histone H3 is replaced by CENH3 within the nucleosomes of active centromeres in several non-plant species [48], and all of the KPs examined were mislocated when CENH3 was disrupted [49,50]. CENH3 homologues of *A. thaliana* (designated HTR12 [51]) and maize [34*] have been identified through database searches and were found at centromeres throughout the cell cycle. CENH3 evolves rapidly; the known CENH3 proteins share a common histone H3 core sequence but diverge in their amino-terminal tails and in the internal loop 1 region [52]. The co-adaptive evolution of CENP-A-like KPs and centromeric repeats was reported to have taken place in *Drosophila* and *Arabidopsis* [53]. CENH3 of maize was found to co-localize with CENPC in the inner kinetochore, in a region beneath an outer regulatory domain that contains the cell-cycle checkpoint protein MAD2 and the unidentified 3F3/2 antigen [34*,54]. Because of its close association with DNA, CENH3 of maize has been used to identify the centromere sequences that interact with the kinetochore. Chromatin immunoprecipitation revealed that the *Arabidopsis* 178-bp satellite repeat [55], the maize centromeric CentC tandem repeat arrays and the centromeric retrotransposon CRM preferentially interact with CENH3, whereas non-centromeric sequences do not [34*].

The DNA-binding protein CENP-C is a fundamental component of the centromere, is highly conserved among species and is necessary for the proper assembly of the kinetochore structure and for the metaphase→anaphase transition [56]. A CENP-C-like KP of maize [57] might represent a hybrid protein that is partially homologous to human CENP-C and partially homologous to another protein apparently located within human centrosomes.
CENP-E is one of the KPs that are involved in chromosome motility and spindle checkpoint control. When this kinesin-like protein is depleted from mammalian kinetochores, incomplete chromosome alignment during metaphase and subsequent mitotic arrest result [59]. The cross-reactivity of antibodies against human CENP-E with kinetochores from field bean and barley indicated that KPs similar to CENP-E might occur in plants [58]. Indeed, two putative CENP-E-like KPs of barley (Cpel1 and Cpel2) have been identified. Cpel2 is the most likely candidate for a plant CENP-E homologue; Cpel1 shows similarity to the coiled-coil domain of CENP-E but does not possess a kinesin motor domain [60].

Putative homologues of the centromere-binding factor CBF5p and the yeast KP SUPPRESSOR OF KINETOCHORE PROTEIN1 (SKP1), which is part of the CBF2 complex and interacts with the centromeric DNA element II (CDEII) of yeast centromeres, are apparently present in plant kinetochores. Their metazoic homologues are located within centromeres (SKP1) or nucleoli (the human CBF5 homologue, dyskerin) [58].

The disruption of putative KP genes by T-DNA or their silencing using RNAi approaches is expected to elucidate the functional importance of plant KPs in vivo.

### Pericentromere-specific chromatin modifications

Specific post-translational modifications of histone H3 at pericentromeres are known to occur in metazoa and plants. In plants, serine 10 (H3S10) and 28 (H3S28) of H3 are strongly phosphorylated at pericentromeres during mitosis and during the second meiotic division from prophase until telophase. During the first meiotic division, this modification occurs along entire chromosomes (as is the case for monocentric animal chromosomes during all nuclear divisions). Single chromatids that result from the equational division of univalents at anaphase I, have almost no pericentromeric phosphorylation during the second meiotic division. It has been assumed that H3 phosphorylation is required for the cohesion of sister chromatids during metaphase I and of sister chromatid centromeres during mitosis and metaphase II [63,64]. This assumption was further supported by observations of a maize mutant that was defective in sister chromatid cohesion. In this mutant, the univalents at metaphase I were strongly phosphorylated only in the pericentromeric regions [63]. Similarly, a ‘semi-dicentric’ barley chromosome revealed hyperphosphorylated H3 only at the functional centromere [61], whereas the polycentric chromosomes of Luzula luculoides were hyperphosphorylated along the entire chromosomes during mitosis [62].

Pericentromeres of *A. thaliana* are characterized by heterochromatin-specific dimethylation of histone H3 at lysine 9 and by strong DNA methylation, both of which are considered to be involved in heterochromatin assembly [65]. On the basis of studies carried out in fission yeast [66–68], Dawe [69] discussed the possibility that RNA interference (using small transcripts from centromeric retroelement-derived sequences) and histone H3 lysine

### Table 1: Known plant kinetochore proteins

<table>
<thead>
<tr>
<th>Type</th>
<th>Yeast</th>
<th>Fruit fly</th>
<th>Human/mouse</th>
<th>Plant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural proteins</td>
<td>SKP1</td>
<td>Cse4</td>
<td>p19Skp1b</td>
<td>ASK1–9b, Skp1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CENP-A</td>
<td>HTR12, CENH13</td>
</tr>
<tr>
<td>Mad, mitotic arrest deficient</td>
<td>Mit2</td>
<td>CBF5</td>
<td>CENP-C</td>
<td>ZmCENP-C</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dykerin&lt;sup&gt;a&lt;/sup&gt;</td>
<td>CBF5</td>
</tr>
<tr>
<td>Motor proteins</td>
<td>Kar3p</td>
<td>CENPmeta</td>
<td>CENP-E</td>
<td>Cpel1, Cpel2</td>
</tr>
<tr>
<td>Passenger proteins</td>
<td></td>
<td></td>
<td>CENP-F</td>
<td>CENP-F antigen&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Checkpoint proteins</td>
<td></td>
<td></td>
<td>hMad2</td>
<td>Mad2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hZw10&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Zw10&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hBub1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Bub1-like&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3F3/2 antigen</td>
<td>3F3/2 antigen&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bub3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Bub3-like&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Spindle-associated proteins</td>
<td></td>
<td></td>
<td>γ-tubulin antigen&lt;sup&gt;a&lt;/sup&gt;</td>
<td>γ-tubulin antigen&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>Modified according to R ten Hoopen, unpublished data.
<sup>b</sup>Centrosomal localization
<sup>c</sup>Only DNA sequence data
<sup>d</sup>Nucleolar localization
<sup>e</sup>Antigen only detected by immunostaining
<sup>f</sup>R ten Hoopen, unpublished data

Abbreviations: ASK, associated with spindles and kinetochores; Bub1, budding unsensitive to benomyl1; CID, centromere identifier; Cse4, defects in mitotic chromosome segregation; Mit2, affecting mitotic fidelity; Nop60Bs, nucleolar protein that maps to chromosomal band 60B13-C1s; Zwr10, protein product of the *Drosophila melanogaster* gene I zw10.
9 methylation (recruiting HETEROCHROMATIN PROTEIN1 [HP1]) is used as an epigenetic mechanism in eukaryotes to establish the specific chromatin organization of centromeres that is required for chromosome segregation. Centromere dysfunction has not, however, been reported for Arabidopsis mutants that are characterized by a reduced degree of DNA methylation and H3 lysine 9 dimethylation at pericentromeric chromocenters (e.g. DNA methylation1 [ddm1] and methyl transferase1 [met1]), nor for the kryptonite (kyp) mutant, which has a high level of DNA methylation but low H3 lysine 9 dimethylation at the chromocenters.

Neocentromeres
Most plant neocentromeres differ from those described for humans [70] and Drosophila melanogaster [71] because they occur on chromosomes that have a normal centromere. Plant neocentromeres, which are best known in maize and rye, appear during meiosis rather than during mitosis. Maize neocentromeres occur in terminal heterochromatic domains called knobs, which are composed mainly of tandem repeats that differ from those of regular centromeres [20,72]. Neocentric activity results in preferential migration (i.e. meiotic drive) and causes chromosomes with such knobs to accumulate in the progeny [73,74]. The neocentromeres become active in the presence of an abnormal chromosome 10 (Ab10), which has an extra segment containing at least two genes that control the neocentric activity [75]. Thus, the presence of heterochromatin is necessary but not sufficient for neocentromere activity at interstitial or terminal chromosomal positions. CENP-C, a constitutive component of the maize kinetochore, is missing at the neocentromeres and so it has been suggested that maize neocentromeres possess a simplified type of kinetochore [57]. In rye, no centromere-specific repeats (i.e. CCS1 [16] and Bilby [30]) have been found, but non-centromeric tandem repeats were detected by in-situ hybridization within the terminal regions that show potential neocentric activity [76,77]. It remains to be seen whether the functional centromeres of barley telosomes that lack centromere-specific satellite and cereba sequences [37] harbor known plant KPs and represent neocentromeres that were formed after the loss of regular centromeric sequences.

Conclusions
Centromeric repetitive sequences are not highly conserved and may not be sufficient (e.g. the inactive centromeres of some human dicentric chromosomes) or necessary (e.g. neocentromeres) for the assembly of functional kinetochores. Nevertheless, clusters of tandem repeats that are interspersed with retroelements are typical features of the regular centromeres of most studied plants. Whether RNA interference [69] and/or stretching of centromeric chromatin, caused by the bi-orientation of sister centromeres during nuclear division [39], generate the landmarks for the initiation/maintenance of kinetochore assembly (i.e. loading of CENP-A-like...
Histon H3 variants into centromeric nucleosomes) remains to be elucidated. Future work will reveal the extent to which homologues of KPs from non-plant species are present, replaced or substituted in plant centromeres, and which modifications/interactions are necessary to ensure proper centromere function.

Despite the generally high variability of centromeric DNA and some differences in protein composition, it is clear that the basic features of plant kinetochore assembly and function are conserved with other eukaryotes. Understanding the minimum requirements for functional centromeres will be important in understanding what is needed for the generation of artificial plant chromosomes as vectors for gene transfer.

Acknowledgements
We thank K Dawe, TR Endo, R ten Hoopen, E Schroeder-Reiter and R Rieger for helpful discussion, and G Wanner for providing the electron micrograph. This work was supported by grants from the Deutsche Forschungsgemeinschaft to AH and IS.

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
28. The authors present a compilation of the sequence and sequence organization of all of the rice chromosomes.
32. Saunders VA, Houben A: The pericentromeric heterochromatin of the grass Zingeria biebersteiniana (2n = 4) is composed of Zbcen1-type tandem repeats that are intermingled with accumulated dispersedly organized sequences. Genome 2001, 44:952-961.


The authors of this paper show the preferential co-immunoprecipitation of the CENP-A-like KP with plant centromeric sequences. This affinity indicates a possible importance for this KP in centromere assembly.


The author presents an interesting hypothesis on the epigenetic determination of centromere position by the conjunction of RNAi and H3K9 methylation processes.


