

Multiple repetitive DNA sequences in the paracentromeric regions of *Arabidopsis thaliana* L.

A. Brandes, H. Thompson, C. Dean & J. S. Heslop-Harrison

Received 13 December 1996; received in revised form 24 February 1997; accepted for publication by S. Mizuno 6 March 1997

Nine repetitive DNA sequences, present in the haploid *Arabidopsis thaliana* genome in 7–300 copies, were hybridized *in situ* to metaphase and interphase chromosomes. Every sequence was detected on all five chromosome pairs, but was not evenly dispersed over the genome. Clusters of signals were found in particular regions of the centromeric heterochromatin, and each sequence showed a characteristic distribution pattern. Some sequences hybridized more strongly on different chromosomes, reflecting chromosome-specific amplification or the presence of homologous sequences. No hybridization signals could be detected on euchromatic regions. *In situ* hybridization on extended chromatin fibres showed that the pAL1 repeats are interrupted by another repetitive DNA sequence. A cosmid subclone (74A) contained a (GA)₃₈ microsatellite motif, and hybridization with a (GA) oligonucleotide revealed that most of the hybridization sites of 74A correspond to the distribution of this microsatellite motif. The results show that the paracentromeric heterochromatin of *A. thaliana* chromosomes is composed not only of the tandemly arranged 180-bp repeat family pAL1/pAtMr, but also of some other repetitive sequences, thus giving a better understanding of the organization of sequences at the centromeres of *A. thaliana*.

Key words: centromere, fibre DNA, genome organization, *in situ* hybridization, microsatellite

Introduction

The *Arabidopsis thaliana* L. genome is among the smallest plant genomes known (around 100–50 Mbp; Meyerowitz 1994) and, in contrast to species with larger genomes, the amount of repetitive DNA is low (Meyerowitz 1994). The 18S–25S ribosomal RNA-encoding genes and the intergenic spacers represent the most abundant repetitive sequences, accounting for 6–8% of the genome (Meyerowitz 1994), while the 5S rDNA makes up 0.7% (Campbell *et al.* 1992) and the telomeric sequences about 0.3% of the genome (Richards & Ausubel 1988). The only major satellite sequence, the pAL1/pAtMr family (Martinez-Zapater

et al. 1986, Murata *et al.* 1994), constitutes, together with a related 500-bp HindIII repeat (Simoens *et al.* 1988), around 3.5% of the genome (Murata *et al.* 1994). A 160-bp AluI repeat contributes to about 0.3% of the genome (Simoens *et al.* 1988). Microsatellites, although less abundant than in mammals (Bell & Ecker 1994, Depeiges *et al.* 1995), as well as minisatellite sequences (Tourmente *et al.* 1994), are scattered throughout the genome. At least 2–3% of the nuclear DNA corresponds to dispersed repetitive sequences (Pruitt & Meyerowitz 1986); most dispersed repetitive sequences in plants have been identified as mobile genetic elements or their remnants (Flavell 1986). In *A. thaliana*, long terminal repeat (LTR) retrotransposons (Voytas & Ausubel 1988, Konieczny *et al.* 1991) represent some 1% of the genome (Hirochika *et al.* 1992, Brandes *et al.* 1997). Non-LTR retrotransposons (Wright *et al.* 1996) and an unusual retroelement, 'Athila' (0.3% of the genome; Pélissier *et al.* 1995, 1996), have also been found.

In situ hybridization is a reliable method for the localization of sequences to chromosome regions. Tandemly arranged repetitive sequences are often associated with constitutive heterochromatic regions of chromosomes, which remain condensed at interphase and are localized at telomeric and subtelomeric positions (Bedbrook *et al.* 1980), pericentromeric regions (Harrison & Heslop-Harrison 1995) or interstitial blocks (Fuchs *et al.* 1994). In *A. thaliana*, heterochromatin is largely restricted to regions at or around the centromeres of all five chromosome pairs (paracentromeric heterochromatin) and at the nucleolus organizing regions (NORs), accounting for some 12% of the chromosomal target (Ambros & Schweizer 1976). The pAL1/pAtMr family was found at the centromeric regions of all chromosomes by *in situ* hybridization (Maluszynska & Heslop-Harrison 1991, Murata *et al.* 1994). The related 500-bp and the 160-bp repeated sequences also occur at the centromeres (Bauwens *et al.* 1991). 5S rRNA loci on chromosome pairs 4 and 5 (Schmidt *et al.* 1995, Murata *et al.* 1997) and telomere-

A. Brandes (corresponding author) and J. S. Heslop-Harrison are members of the Karyobiology Group, Department of Cell Biology, and H. Thompson and C. Dean are at the Molecular Genetics Department, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK. Tel: (+44) 1603 452571; Fax: (+44) 1603 501771; Email: brandes@bbsrc.ac.uk.

related repeat sequences on chromosome 1 (Richards *et al.* 1991) have been assigned to the centromeric regions by physical mapping and *in situ* hybridization. Recently, dispersed repetitive sequences, including the reverse transcriptase domain of Ty1-*copia*-like retrotransposons and mi167, a repetitive sequence associated with a restriction fragment length polymorphism (RFLP) marker, were found in the paracentromeric regions of all chromosomes (Thompson *et al.* 1996a, Brandes *et al.* 1997). The Ty1-*copia* elements hybridized in clusters at the paracentromeric heterochromatin, an organization that contrasts with the dispersed distribution of Ty1-*copia*-like elements over the chromosomes of most plant species (Brandes *et al.* 1997). Pélissier *et al.* (1995, 1996) described the association of *Athila* retrotransposable elements and pAL1 in stretches of cloned DNA.

Thompson *et al.* (1996a-c) isolated several cosmid subclones and end clones from yeast artificial chromosomes (YACs) and mapped them on the YAC contig map of chromosome 4. We were interested in the physical location of these sequences and examined their chromosomal and genomic distribution to gain

insights into the arrangement of repetitive DNA sequences in the *Arabidopsis* genome.

Materials and methods

Isolation and molecular characterization of *A. thaliana* plasmid probes are described in Thompson *et al.* (1996a-c), and clones and their copy numbers are shown in Table 1. The 5S rDNA probe pXV1 was isolated from sugar beet (Schmidt *et al.* 1994). It includes the 5S rRNA gene and the adjacent intergenic spacer. pAt18S, containing a part of the 18S rRNA gene, was isolated by polymerase chain reaction (PCR) from genomic *A. thaliana* 'Columbia' DNA and cloned into the pCRII vector of the TA cloning kit (Invitrogen).

Cloned probes were labelled by PCR partially replacing dTTP with biotin-16-dUTP (Boehringer) or with digoxigenin-11-dUTP (Boehringer). The 18S-5.8S-25S rDNA probe pTa71 (Gerlach & Bedbrook 1979) was labelled by nick translation and the microsatellite (GA)₁₂ was labelled by oligonucleotide tailing (Boehringer) with digoxigenin-11-dUTP.

Chromosome preparations were made from flower buds of *A. thaliana* 'Columbia'. Extended DNA fibres were prepared from nuclei by treating them with the STE2 lysis buffer, following the protocol of Parra & Windle (1993) and Fransz

Table 1 Size, copy number and chromosomal localization of clones isolated by Thompson *et al.* (1996a-c)

Name	Size (bp)	Copy number estimated by Southern hybridization	Origins	Localization by <i>in situ</i> hybridization	Figure
mi167 ¹	980 bp	~90	RFLP marker	Paracentromeric; extending beyond the pAL1 region, not on the short arm of chromosome 4	2a-g
164A ²	900 bp	~150	Cosmid subclone	Small clusters at the paracentromeric heterochromatin of all chromosome pairs	2h,i
163A ²	450 bp	~90	Cosmid subclone	Small clusters at the paracentromeric heterochromatin of all chromosome pairs; two chromosome pairs with weaker signals	2j,k
164B ⁴	1100 bp	ND	Cosmid subclone	Small clusters at the paracentromeric heterochromatin; weaker signals on two or three chromosome pairs, one of which is chromosome pair 4	2l-u
278A ²	1300 bp	~20	Cosmid subclone	Small clusters of unequal strength at the paracentromeric heterochromatin	3a-c
106B ²	400 bp	~300	Cosmid subclone	Dispersed over the paracentromeric heterochromatin of all chromosomes; highly amplified at the centromeres of four chromosome pairs	3d,e
15C8LE ³	1005 bp	At least 5	Left end from YAC	Small clusters at the paracentromeric heterochromatin of all chromosomes; highly amplified on the short arm of chromosome pairs 2 and 4	3f-h
11B7RE ³	349 bp	~16	Right end from YAC	Small clusters at the paracentromeric heterochromatin of all chromosome pairs; more distal to the centromere	3i-k
74A ²	700 bp	~7*	Cosmid subclone	Small clusters at the paracentromeric heterochromatin of all chromosomes; highly amplified on the short arm of chromosome pairs 2 and 4	3l-o

All clones are not organized in long tandem arrays and no continuous open reading frames were found.

1 =Thompson *et al.* (1996a).

2 =Thompson *et al.* (1996b).

3 =Thompson *et al.* (1996c).

4 =This paper.

ND, not determined.

*For the MYB homology.

et al. (1996). *In situ* hybridization was as described by Maluszynska & Heslop-Harrison (1991) and Brandes *et al.* (1997). Washing was carried out in several changes of standard saline citrate (SSC; 20 ×SSC: 3 M NaCl, 0.3 M sodium citrate) with a most stringent wash in 20% formamide, 0.1 ×SSC, which is equivalent to 84% stringency. For *in situ* hybridization with the microsatellite (GA)₁₂, the conditions followed Schmidt & Heslop-Harrison (1996a). Biotinylated probes were detected by streptavidin-conjugated Cy3 (Sigma), and digoxigenin-labelled probes by anti-digoxigenin-fluorescein (Boehringer). Chromosomes were counterstained with DAPI (4',6 diamidino-2-phenylindole, 6 µg/ml in McIlvaine's citric buffer). Photographs were taken on a Fujicolour 400 colour print film; negatives were digitized with a Nikon film scanner LS 10E and printed after contrast optimization of the whole image with Adobe Photoshop 3.0.

The preparation of nuclei for electron microscopy followed standard techniques (Heslop-Harrison 1996).

Results

Figure 1 shows an electron micrograph of an interphase nuclei of *A. thaliana*. Most chromatin is decondensed and only visible as diffuse threads, but some chromatin segments remain condensed. They are localized against the nucleolus (N) and the nuclear envelope (arrowhead). The former most probably correspond to unexpressed rRNA genes (Maluszynska & Heslop-Harrison 1993) and the latter to the pAL1/pAtMr repeat family (Martinez-Zapater *et al.* 1986, Murata *et al.* 1994) and other repetitive sequences, presumably included in the heterochromatin around the centromeres. Our *in situ* hybridization reveals that

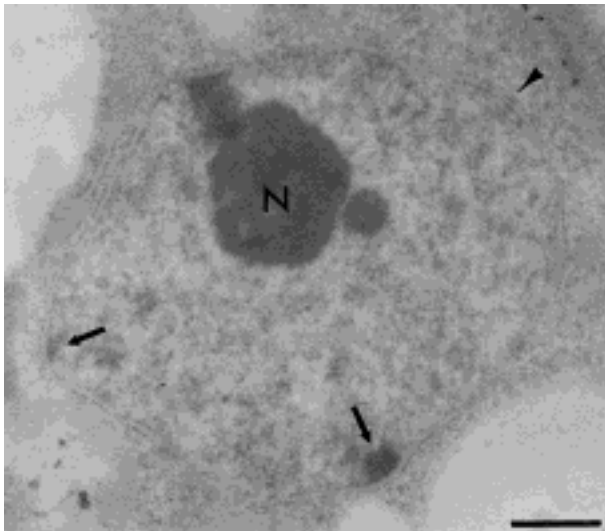


Figure 1. Electron micrograph of an ovule wall interphase nucleus of *A. thaliana*. The nuclear envelope (arrowhead) surrounds the nuclei. The dark-stained nucleolus (N) and dark condensed chromatin representing unexpressed rDNA sequences (adjacent to the nucleolus) and paracentromeric heterochromatin (peripheral, arrows) are visible. Decondensed chromatin is detected as diffuse darker regions. Bar = 500 nm.

several repetitive sequences cluster at paracentromeric regions of all five chromosome pairs (Figures 2 & 3), partially colocalizing with the 180-bp repeat family pAL1 (Maluszynska & Heslop-Harrison 1991).

mi167, a repetitive sequence of a genomic clone used as a RFLP marker and not found on YACs mapping on the short arm of chromosome 4 (Schmidt *et al.* 1995), colocalizes over most parts of the paracentromeric regions of metaphase chromosomes with the pAL1 sequences (Figure 2a-d). On at least one chromosome pair, carrying the 18S, 5.8S and 25S rRNA genes and the intergenic spacer (rDNA) and visible in the DAPI-stained metaphase by chromatid morphology (arrows), mi167 is restricted to the arm without rDNA, although the extension of the small metaphase chromosomes is not high enough to detect short clusters very close to the centromere on the short arm. Decondensed interphase chromosomes allow a more precise location, but some chromosome regions remain condensed and are visible as DAPI-positive sites (Figure 2e). Most of them hybridize with the pAL1 sequence (Figure 2f) and hence represent at least a part of the paracentromeric heterochromatin. The remaining DAPI-positive chromosome regions correspond to some of the 18S-25S rDNA or the 5S rDNA, which may stay condensed during interphase (Maluszynska & Heslop-Harrison 1993 and Figure 1). mi167 hybridizes to discrete regions on interphase chromosomes (Figure 2g), colocalizing with the pAL1 sequences as on the metaphase chromosomes. However, additional sites not hybridizing with the 180-bp repeats but with the mi167 sequences are present (arrows).

The repetitive clones 164A, 164B, 163A and 278A hybridize to all five chromosome pairs in varying strengths, reflecting different copy numbers and/or the existence of closely related sequences on the chromosomes (Figures 2h-u & 3a-c). Unlike mi167, the 164A, 164B and 163A repeats cluster at the pericentromeric regions in addition to some dispersed hybridization. 164A shows major signals on every chromosome (Figure 2h & i), whereas 163A hybridized only on six chromosomes with multiple strong sites (Figure 2j & k). Four to six chromosomes show stronger signals after hybridizing with 164B (Figure 2l - u). On two chromosomes, weaker signals on only one arm appeared (arrowed). These chromosomes hybridized on the opposite arm with 5S rDNA and with 18S rDNA sequences as shown by the yellow-green fluorescence in Figure 2n-p & s-u. As chromosome 4 is the only chromosome harbouring genes of both rDNA sequences (Schmidt *et al.* 1995, Murata *et al.* 1997), it can be concluded that one of the chromosome pairs with the weaker 164B hybridization signals is chromosome pair 4.

To define the arrangement of sequences in the centromeric regions, double hybridization with pAL1 and one of the repetitive sequences, clone 164B, was carried out on extended DNA fibres. This high-resolution mapping method revealed that tandem arrays of pAL1

(yellow-green fluorescence) are interrupted by sequences hybridizing with 164B (red fluorescence, Figure 2w).

In agreement with the low copy number estimate (Table 1), clone 278A shows fewer and weaker hybridization signals (Figure 3a-c) than the sequences described above. Double hybridization with the 5S rDNA probe pXV1 (yellow-green fluorescence in Figure 3c) marks the chromosomes 3, 4 and 5 (Murata *et al.* 1997).

Clone 106B, a diverged part of the long terminal repeat of the *Athila* element (Pélissier *et al.* 1995), present in the genome in about 300 copies, is dispersed over the paracentromeric regions of all chromosomes but amplified in regions most proximal to the centromere (Figure 3d & e). This is consistent with the close association of *Athila* elements and the 180-bp repeats (Pélissier *et al.* 1996).

15C8LE sequences also localize throughout the DAPI-positive chromosome regions and show discrete hybridization clusters (Figure 3f-h). Different copy numbers or different homology on individual chromosomes is evident from the strength of hybridization. This sequence is most amplified on chromosome pairs 2 and 4 (arrows), which are identified by the 18S, 5.8S and 25S rRNA genes (Figure 3h).

The right end sequence of YAC 11B7, 11B7RE, hybridizes to the paracentromeric heterochromatic regions of all five chromosome pairs, showing discrete clusters and some dispersion (Figure 3i-k). In most cases, the clustered signals appeared at the transition between the heterochromatin and the euchromatin.

74A, a cosmid subclone, which includes a (GA)₃₈ microsatellite repeat in addition to a region showing homology to the *Arabidopsis* MYB gene *GL1*, is amplified in approximately seven copies per haploid genome (Table 1). 74A shows bright signals on the two NOR chromosome pairs, as revealed by double *in situ* hybridization with the 18S-5.8S-25S rRNA genes (Figure 3l-o, arrowed). On all other chromosomes, only faint signals were visible at the heterochromatic regions. Thompson *et al.* (1996b) demonstrated that the microsatellite motif in 74A accounts for the highly

repetitive pattern on Southern blots. Hybridization with a microsatellite probe, (GA)₁₂, reveals four chromosomes with strong hybridization sites (Figure 3p & q, arrowed). The prometaphase shows two heterochromatic spots, probably the NORs, hybridizing with the microsatellite (Figure 3r & s, arrowed). The elongated chromosomes enabled the assignment of the microsatellite to the paracentromeric heterochromatin of all chromosomes, with some chromosomes showing a gap between the signals, most probably the primary constriction at the centromere.

Discussion

The results demonstrate the complex genomic organization of repetitive DNA sequences around, and perhaps at, the centromeres of *A. thaliana* chromosomes. They show that, in addition to the pAL1/pAtMr family, multiple middle repetitive DNA sequences are amplified at the paracentromeric heterochromatin of all chromosomes. Each repeat has a different and characteristic distribution over the genome, reflecting chromosome- and sequence-specific amplification (Table 1, Figures 2 & 3). Unlike many other plant species (Ganal *et al.* 1988, Schmidt & Heslop-Harrison 1996b), *A. thaliana* shows no uniform distribution of middle repetitive sequences over the chromosomes. Thus, the majority of the chromosome arms are composed predominantly of low- and single-copy DNA sequences with only a few repetitive DNA sequence motifs, as is also clear from the well-developed contigs of chromosomes 2 (Zachgo *et al.* 1996) and 4 (Schmidt *et al.* 1995). It agrees with the calculation of Pruitt & Meyerowitz (1986) that single-copy DNA is present in extremely long contiguous blocks in the *Arabidopsis* genome with an average length of single-copy sequence regions of 125 kb. The small proportion of repetitive sequences on the chromosome arms is also obvious from the extreme decondensation of chromatin in interphase nuclei, assuming that repetitive sequences remain condensed at interphase (Figure 1).

Figure 2. a Metaphase, after hybridization with pAL1 (b) and with mi167(c & d). Heterochromatic regions show strong DAPI staining. Arrows point to two NOR chromosomes, recognizable by their typical shape in the microscope (clearer in I). mi167 sequences hybridize to these two chromosomes only on the arm not carrying the genes for the 18S-25S rDNA. e Interphase, showing DAPI-positive regions of condensed chromatin. Hybridization with pAL1 (f) revealed that most of the DAPI-positive regions collocate with the paracentromeric heterochromatin. The other DAPI-positive sites correspond to condensed sites of 5S rDNA or 18S-25S rDNA. g The same interphase after hybridization with mi167, showing colocalization of pAL1 and mi167 on most of the paracentromeric heterochromatin. Arrows point to some chromosome sites only hybridizing with mi167. h & j DAPI-stained metaphases. i & k The same metaphases after hybridization with 164A (i) and 163A (k); both probes show distinct clusters at the paracentromeric heterochromatin of all chromosomes with some dispersed hybridization. 163A sequences hybridize only on six chromosomes with stronger sites. l Metaphase; two NOR chromosomes, identified by their chromosome morphology, are arrowed. m 164B hybridized to all chromosomes; the two NOR chromosomes arrowed in I show weaker signals. n The same metaphase after detection of the 5S rDNA sites reveals the two NOR chromosomes as chromosome pair 4. o Overlay of m and n. p Overlay of l, m and n. q-u As(l-p), but 18S rDNA sites on chromosome pairs 2 and 4 were detected with probe pAt18s. v and w Double hybridization of pAL1 (green fluorescence) and 164B (red fluorescence) to extended DNA fibres, showing interruption of pAL1 repeats with 164B sequences. Scale bars=10 μm.

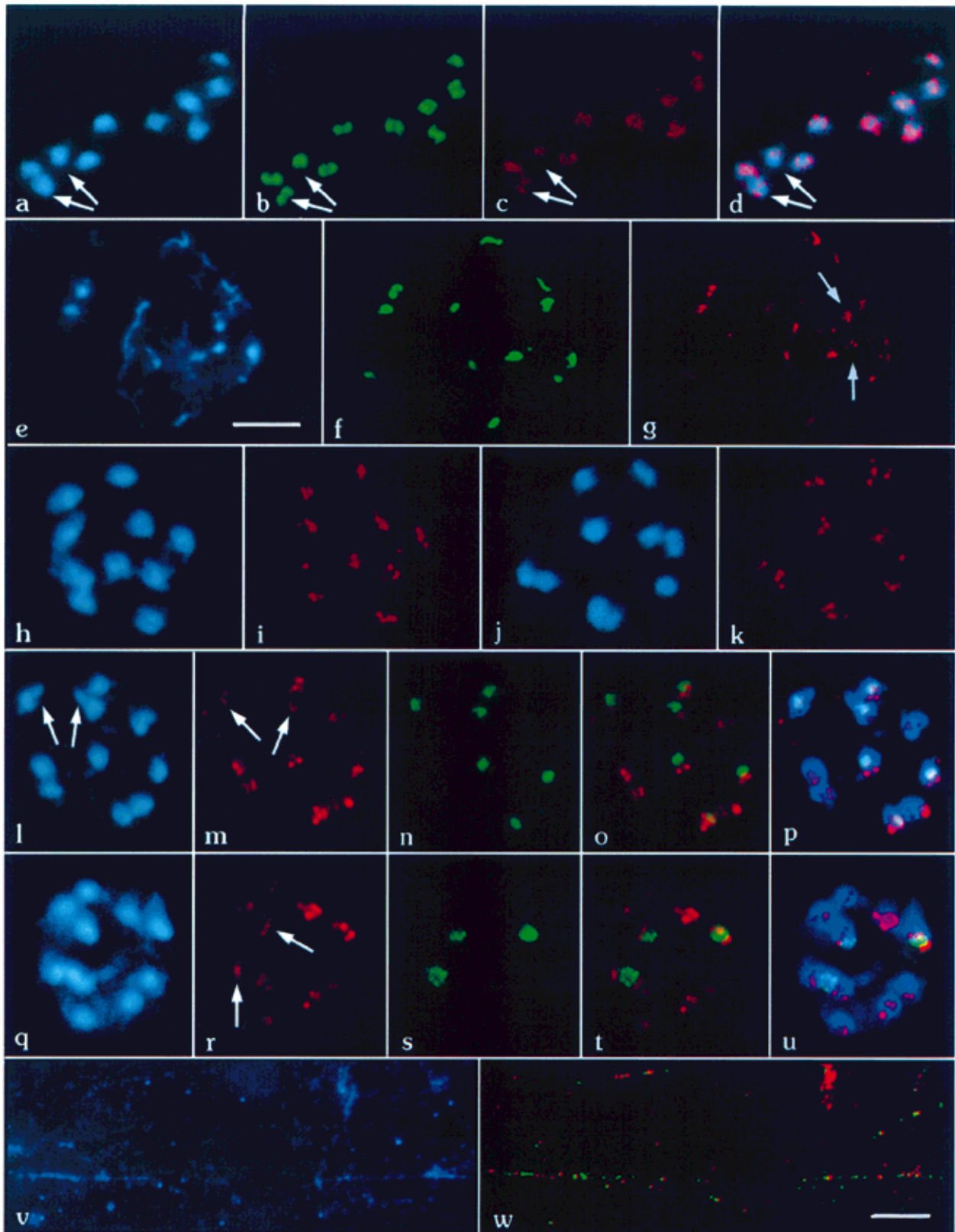


Figure 2.

Figures 2 & 3. *In situ* hybridization of repetitive sequences to metaphase and interphase chromosomes of *A. thaliana* ($2n=10$). Chromosomes were stained with DAPI (blue fluorescence), and the sites of the probe were visualized by red or yellow-green fluorescence. Magnification in Figures 2a-d and h-u and 3a-o, $\times 2880$. Bar = $10 \mu\text{m}$.

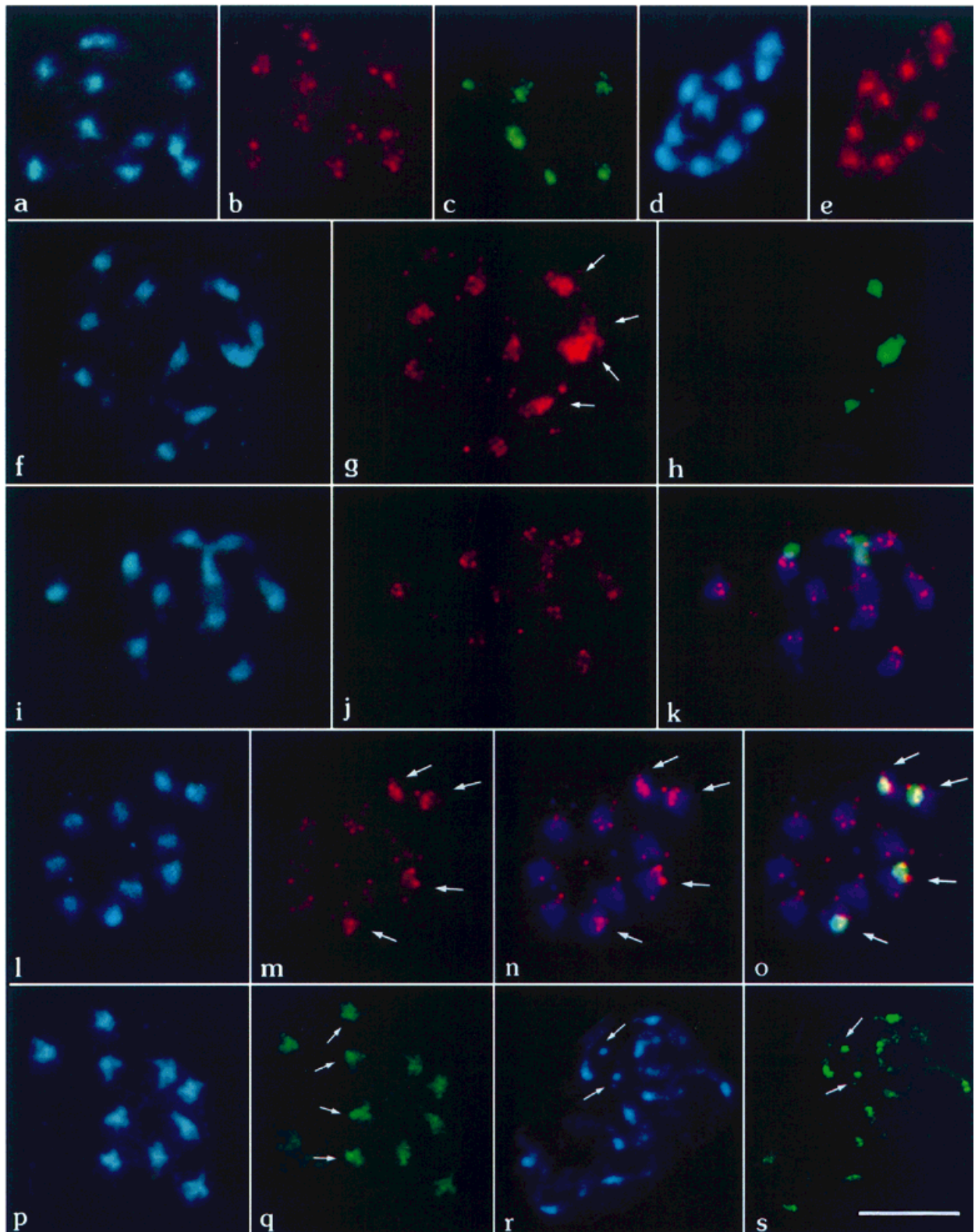


Figure 3.

The arrangement and genomic location of the repetitive sequences used was investigated by Southern hybridization onto YAC and cosmid clones (Schmidt *et al.* 1995, Thompson *et al.* 1996a–c). These hybridizations showed association between some sequences, supporting the *in situ* hybridization results (Figures 2 & 3) showing the colocalization of the sequences at all paracentromeric heterochromatic regions. However, there were some differences between chromosomes. The short arm of chromosome 4 is mainly composed of sequences coding for the 5S (Murata *et al.* 1997) and the 18S–5.8S–25S rRNA, and sequences homologous to mi167 and 163A were not detected on this arm by either *in situ* hybridization or genetic mapping approaches (Schmidt *et al.* 1995, Thompson *et al.* 1996a,b). 164B sequences were mapped to only a very small region of the short arm of chromosome 4, but no signals were visible after *in situ* hybridization. The hybridization of 278A to chromosome 4, although not found by Thompson *et al.* (1996b) on the YAC contigs available for chromosome 4, may be caused by stringency differences, hybridization kinetics or sequences not represented in the YAC library.

Because *in situ* hybridization shows the genomic distribution of sequences, it may facilitate the generation of YAC contigs for further chromosomes. For instance, 15C8LE hybridized to YAC libraries available for chromosomes 2, 3, 4 and 5 (Thompson *et al.* 1996c), but we also found 15C8LE sequences on chromosome 1. Subclone 11B7RE hybridized to YAC clones positioned on chromosomes 4 and 5 (Thompson *et al.* 1996c). *In situ* hybridization revealed that chromosomes 1, 2 and 3 also contain sequences homologous to 11B7RE. Furthermore, the localization and interspersed data may be of value in the interpretation of genomic sequence data, in which large arrays of tandem repeats are difficult to align.

The resolution of *in situ* hybridization signals on metaphase and interphase chromosomes is not high enough for the investigation of the precise arrangement of different sequences on a chromosome. Tandem repeated elements are arranged in continuous blocks like the TGRI repeat of tomato (Ganal *et al.* 1988) or are interspersed with other sequences like the type IV

repeat in *Cucumis sativus* (Ganal & Hemleben 1988). The long-range organization of the pAL1/pAtMr sequences, which are present in arrays of 50 kb (Martinez-Zapater *et al.* 1986, Murata *et al.* 1994), is not well characterized, but there is some evidence for interspersed with other sequences. Telomere-like repeated sequences (Simoens *et al.* 1988) and *Athila* retroelements (Pélissier *et al.* 1996) are inserted in pAL1 repeats, and mi167 sequences have been detected in YAC clones containing pAL1 repeats (Thompson *et al.* 1996a). We showed that both sequences, mi167 and the subclone 106B, which has homology to the LTR of the *Athila* retrotransposon, hybridized to the same heterochromatic regions of *A. thaliana* chromosomes as pAL1.

Further evidence for the interspersed of pAL1 repeat arrays comes from double hybridization with 164B sequences on extended DNA fibres, which give extremely high resolution of hybridization signals (Figure 2v & w). Short stretches of the DNA fibre hybridized only with pAL1 sequences and were interrupted by signals of 164B sequences. The gaps between the signals seen in Figure 2w often occur after the hybridization of repetitive sequences on DNA fibres and their interpretation is still a matter of discussion (Gerdes *et al.* 1994, Houseal & Klinger 1994). However, DNA fibres have been used for ordering YAC contigs, cosmid and lambda clones and repetitive sequences (Haaf & Ward 1994, Fransz *et al.* 1996). Here, we have shown that this technique may be especially useful for the mapping of large chromosomal arrays composed of different repetitive DNA sequences and for investigating the arrangement of these sequences within the arrays. YACs containing tandemly repeated sequences are often unstable because of recombination events (Schmidt *et al.* 1995), leaving a gap in their grouping into contigs and their positioning on a physical map.

A. thaliana chromosomes show a remarkable amount and concentration of heterochromatin around the centromeres. Sequences may have accumulated and homogenized by exchange events between different centromeres. It is noteworthy that most of the sequences revealed no homology to any known nucleotide or protein sequence in the database (Thompson *et al.*

Figure 3. **a** DAPI-stained metaphase with 278A **(b)** hybridizing in small clusters to each chromosome. **c** 5S rDNA sites are detected by pXV1 on three chromosome pairs. **d** Metaphase with the distribution of 106B **(e)**, a sequence with homology to the LTR of the *Athila* retroelement, hybridizing to all five chromosome pairs with concentration near four pairs of centromeres. **f** Metaphase. **g** Clone 15C8LE is distributed over the centromeric regions of all chromosomes but amplified on four chromosomes (arrows), identified as chromosomes 2 and 4 by the presence of the 18S–25S rDNA **(h)**. **i** and **j** Hybridization of 11B7RE to metaphase chromosomes (red fluorescence in **j**) results in small clusters on all chromosomes, mostly between centromeric heterochromatin and euchromatin of the chromosome arms. **k** The same metaphase after hybridization with the 18S–25S rDNA genes (clone pTa71), overlaid with the metaphases in **i** and **j**. **l** Metaphase with 74A **(m** and **n)**, showing faint signals on three and stronger signals on two chromosome pairs (arrows), identified as the NOR chromosome pairs 2 and 4 by the 18S–25S rRNA genes **(o)**. **p** and **r** DAPI-stained metaphase **(p)** and prometaphase **(r)** chromosomes and the appropriate metaphase **(q)** and prometaphase **(s)** after detection of the microsatellite motif (GA)₁₂. **q** Four chromosomes show stronger signals (arrows). **q** The prometaphase reveals amplification of the microsatellite in the paracentromeric regions of all chromosomes and near two NOR regions, which stayed condensed during prometaphase (arrows). Scale bar = 10 μm.

1996a–c). The functional relevance of repetitive sequences (without those coding for rRNA) is still speculative and controversial, but the accumulation of different types of repetitive sequences in the centromeric regions of many eukaryotes, including plants, indicates an involvement in centromeric activity, possibly as a stereophysical factor. For instance, the tandemly arranged 171-bp α -satellite repeats found at all human centromeres have been shown to be important for centromeric function (Haaf *et al.* 1992). Some plant species also have tandemly arranged satellite repeats at their centromeric regions, such as *Pennisetum glaucum* (Kamm *et al.* 1994), *Brassica* species (Harrison & Heslop-Harrison 1995), *Vigna unguiculata* (Galasso *et al.* 1995) and sugar beet (Schmidt & Heslop-Harrison 1996a). Chromosome-specific subrepeats of the pAL1/pAtMr family have been isolated in *A. thaliana* (F. Motoyoshi, M. Murata & J.S. Heslop-Harrison, unpublished results), thus resembling the organization of the α -satellite repeats in humans (Willard & Wayne 1987). Nevertheless, in plant species like *Vicia faba*, rye and wheat, no major, highly abundant centromere-associated satellite sequences have been found. Other plant species, such as barley and sugar beet, show amplification of microsatellite sequences around their centromeric regions (Pedersen & Linde-Laursen 1994, Schmidt & Heslop-Harrison 1996a). We have shown clustering of a (GA) microsatellite motif at the paracentromeric regions of prolonged prometaphase chromosomes of *A. thaliana*, but the centromeric constrictions themselves were free of hybridization signals.

Sequences and the arrangement of sequences making up a functional centromere in higher eukaryotes are still unknown. We have demonstrated that the paracentromeric regions of *A. thaliana* chromosomes are composed of many different repetitive sequences, each with a characteristic distribution between chromosomes, ranging from tandem repeated sequences to microsatellites and several middle repetitive sequences and even a single-copy sequence (Schmidt *et al.* 1995). The arrangement of different sequences around the centromere resembles the heterogeneous DNA composition of fission yeast centromeres (Baum *et al.* 1994) and of *Drosophila melanogaster* centromeres (Murphy & Karpen 1995), in which long stretches of complex repeated DNA sequences are required for centromere function. Thus, *A. thaliana* centromeres may be used as a model for the composition of plant centromeres.

Acknowledgements

The authors would like to thank Dr R. Whittier for kindly providing the marker mi167. We thank Dr T. Schmidt (Institute of Crop Science and Plant Breeding, Christian-Albrechts University, Kiel, Germany) for critical reading of the manuscript. A.B. was supported by a grant from BBSRC PAG/01502 to J.S.H.-H., C.D. by BBSRC PG208/0606 and H.T. by a BBSRC studentship.

References

- Ambros P, Schweizer D (1976) The Giemsa C-banded karyotype of *Arabidopsis thaliana*. *Arabidopsis Info Service* **13**: 167–171.
- Baum M, Ngan VK, Clarke L (1994) The centromeric K-type repeat and the central core are together sufficient to establish a functional *Schizosaccharomyces pombe* centromere. *Mol Biol Cell* **5**: 747–761.
- Bauwens S, Van Oostveldt P, Engler G, Van Montagu M (1991) Distribution of the rDNA and three classes of highly repetitive DNA in the chromatin of interphase nuclei of *Arabidopsis thaliana*. *Chromosoma* **101**: 41–48.
- Bedbrook JR, Jones J, O'Neil M, Thompson RD, Flavell RB (1980) Molecular characterization of telomeric heterochromatin in *Secale* species. *Cell* **19**: 545–560.
- Bell CJ, Ecker JR (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics* **19**: 137–144.
- Brandes A, Heslop-Harrison JS, Kamm A, Kubis S, Doudrick RL, Schmidt T (1997) Comparative analysis of the chromosomal and genomic organization of Ty1-*copia*-like retrotransposons in pteridophytes, gymnosperms and angiosperms. *Plant Mol Biol* **33**:11–21.
- Campbell BR, Song Y, Posch TE, Cullis CA, Town CD (1992) Sequence and organization of 5S ribosomal RNA-encoding genes of *Arabidopsis thaliana*. *Gene* **112**: 225–228.
- Depeiges A, Goubely C, Lenoir A *et al.* (1995) Identification of the most represented repeated motifs in *Arabidopsis thaliana* microsatellite loci. *Theor Appl Genet* **91**: 160–168.
- Flavell RB (1986) Repetitive DNA and chromosome evolution in plants. *Phil Trans R Soc Lond B* **312**: 227–242.
- Fransz PF, Alonso-Blanco C, Liharska TB, Peeters AJM, Zabel P, De Jong JH (1996) High resolution physical mapping in *Arabidopsis thaliana* and tomato by fluorescence *in situ* hybridization to extended DNA fibres. *Plant J* **9**: 421–430.
- Fuchs J, Pich U, Meister A, Schubert I (1994) Differentiation of field bean heterochromatin by *in situ* hybridization with a repeated *FokI* sequence. *Chrom Res* **2**: 25–28.
- Galasso I, Schmidt T, Pignone D, Heslop-Harrison JS (1995) The molecular cytogenetics of *Vigna unguiculata* (L.) Walp: the physical organization and characterization of 18s–5.8s–25s rRNA genes, 5s rRNA genes, telomere-like sequences, and a family of centromeric repetitive DNA sequences. *Theor Appl Genet* **91**: 928–935.
- Ganal M, Hemleben V (1988) Insertion and amplification of a DNA sequence in satellite DNA of *Cucumis sativus* L. (cucumber). *Theor Appl Genet* **75**: 357–361.
- Ganal M, Lapitan NLV, Tanksley S (1988) A molecular and cytogenetic survey of major repeated DNA sequences in tomato. *Mol Gen Genet* **213**: 262–268.
- Gerdes MG, Carter KC, Moen Jr PT, Lawrence JB (1994) Dynamic changes in the higher-level chromatin organization of specific sequences revealed by *in situ* hybridization to nuclear halos. *J Cell Biol* **126**: 289–304.
- Gerlach WL, Bedbrook JR (1979) Cloning and characterization of ribosomal RNA genes from wheat and rye. *Nucleic Acids Res* **7**: 1869–1885.
- Haaf T, Ward DC (1994) Structural analysis of α -satellite DNA and centromere proteins using extended chromatin and chromosomes. *Hum Mol Genet* **3**: 697–709.
- Haaf T, Warburton PE, Willard HF (1992) Integration of human α -satellite DNA into simian chromosomes: centromere protein binding and disruption of normal chromosome segregation. *Cell* **70**: 681–696.

A. Brandes et al.

- Harrison GE, Heslop-Harrison JS (1995) Centromeric repetitive DNA sequences in the genus *Brassica*. *Theor Appl Genet* **90**: 157–165.
- Heslop-Harrison JS (1996) Electron microscopy in plant chromosomes. In: Fukui K, Nakayama S, eds. *Plant Chromosomes: Laboratory Methods*. Baton Rouge: CRC Press, pp. 225–245.
- Hirochika H, Fukuchi A, Kikuchi F (1992) Retrotransposon families in rice. *Mol Gen Genet* **233**: 209–216.
- Houseal TM, Klinger KW (1994) Commentary: What's in a spot? *Hum Mol Genet* **3**: 1215–1216.
- Kamm A, Schmidt T, Heslop-Harrison JS (1994) Molecular and physical organization of highly repetitive, undermethylated DNA from *Pennisetum glaucum*. *Mol Gen Genet* **244**: 420–425.
- Konieczny A, Voytas DF, Cummings MP, Ausubel FM (1991) A superfamily of *Arabidopsis thaliana* retrotransposons. *Genetics* **127**: 303–308.
- Maluszynska J, Heslop-Harrison JS (1991) Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. *Plant J* **1**: 159–166.
- Maluszynska J, Heslop-Harrison JS (1993) Molecular cytogenetics of the genus *Arabidopsis*: *in situ* localization of rDNA sites, chromosome numbers and diversity in centromeric heterochromatin. *Ann Bot* **71**: 479–484.
- Martinez-Zapater JM, Estelle MA, Somerville CR (1986) A highly repeated DNA sequence in *Arabidopsis thaliana*. *Mol Gen Genet* **204**: 417–423.
- Meyerowitz EM (1994) Structure and organization of the *Arabidopsis thaliana* nuclear genome. In: Meyerowitz EM, Somerville CR, eds. *Arabidopsis*. New York: Cold Spring Harbor Laboratory Press, pp. 21–36.
- Murata M, Ogura Y, Motoyoshi F (1994) Centromeric repetitive sequences in *Arabidopsis thaliana*. *Jpn J Genet* **69**: 361–370.
- Murata M, Heslop-Harrison JS, Motoyoshi F (1997) Physical mapping of the 5S ribosomal RNA genes in *Arabidopsis thaliana* by multicolour fluorescence *in situ* hybridization with cosmid clones. *Plant J* (in press).
- Murphy TD, Karpen GH (1995) Localization of centromere function in a *Drosophila* minichromosome. *Cell* **82**: 599–609.
- Parra I, Windle B (1993) High resolution visual mapping of stretched DNA by fluorescent hybridization. *Nature Genet* **5**: 17–21.
- Pedersen C, Linde-Laursen I (1994) Chromosomal locations of four minor rDNA loci and a marker microsatellite sequence in barley. *Chrom Res* **2**: 65–71.
- Pélissier T, Tutois S, Deragon JM, Tourmente S, Genestier S, Picard G (1995) *Athila*, a new retroelement from *Arabidopsis thaliana*. *Plant Mol Biol* **29**: 441–452.
- Pélissier T, Tutois S, Tourmente S, Deragon JM, Picard G (1996) DNA regions flanking the major *Arabidopsis thaliana* satellite are principally enriched in *Athila* retroelement sequences. *Genetica* **97**: 141–151.
- Pruitt RE, Meyerowitz EM (1986) Characterization of the genome of *Arabidopsis thaliana*. *J Mol Biol* **187**: 169–183.
- Richards EJ, Ausubel FM (1988) Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* **53**: 127–136.
- Richards EJ, Goodman HM, Ausubel FM (1991) The centromere region of *Arabidopsis thaliana* chromosome 1 contains telomere similar sequences. *Nucleic Acids Res* **19**: 3351–3357.
- Schmidt T, Heslop-Harrison JS (1996a) The physical and genomic organization of microsatellites in sugar beet. *Proc Natl Acad Sci USA* **93**: 8761–8765.
- Schmidt T, Heslop-Harrison JS (1996b) High-resolution mapping of repetitive DNA by *in situ* hybridization: molecular and chromosomal features of prominent dispersed and discretely localized DNA families from the wild beet species *Beta procumbens*. *Plant Mol Biol* **30**: 1099–1114.
- Schmidt T, Schwarzacher T, Heslop-Harrison JS (1994) Physical mapping of rRNA genes by fluorescent *in situ* hybridization and structural analysis of 5S rRNA genes and intergenic spacer sequences in sugar beet (*Beta vulgaris*). *Theor Appl Genet* **88**: 629–636.
- Schmidt R, West J, Love K et al. (1995) Physical map and organization of *Arabidopsis thaliana* chromosome 4. *Science* **270**: 480–483.
- Simoens CR, Gielen J, Van Montagu M, Inzé D (1988) Characterization of highly repetitive sequences of *Arabidopsis thaliana*. *Nucleic Acids Res* **16**: 6753–6766.
- Thompson H, Schmidt R, Brandes A, Heslop-Harrison JS, Dean C (1996a) A novel repetitive sequence associated with the centromeric regions of *Arabidopsis thaliana*. *Mol Gen Genet* **253**: 247–252.
- Thompson H, Schmidt R, Dean C (1996b) Identification and distribution of seven classes of middle-repetitive DNA in the *Arabidopsis thaliana* genome. *Nucleic Acids Res* **24**: 3017–3022.
- Thompson H, Schmidt R, Dean C (1996c) Analysis of the occurrence and nature of repeated DNA in a 850 kb region of *Arabidopsis thaliana* chromosome 4. *Plant Mol Biol* **32**: 553–557.
- Tourmente S, Deragon JM, Lafleuril J et al. (1994) Characterization of minisatellites in *Arabidopsis thaliana* with sequence similarity to the human minisatellite core sequence. *Nucleic Acids Res* **22**: 3317–3321.
- Voytas DF, Ausubel FM (1988) A *copia*-like transposable element family in *Arabidopsis thaliana*. *Nature* **336**: 242–244.
- Willard HF, Wayne JS (1987) Chromosome-specific subsets of human alpha satellite DNA: analysis of sequence divergence within and between chromosomal subsets and evidence for an ancestral pentameric repeat. *J Mol Evol* **25**: 207–214.
- Wright DA, Ke N, Smalle J, Hauge BM, Goodman HM, Voytas DF (1996) Multiple non-LTR retrotransposons in the genome of *Arabidopsis thaliana*. *Genetics* **142**: 569–578.
- Zachgo EA, Wang ML, Dewdney J et al. (1996) A physical map of chromosome 2 of *Arabidopsis thaliana*. *Genome Res* **6**: 19–25.