

Arabidopsis thaliana Centromere Regions: Genetic Map Positions and Repetitive DNA Structure

Elaine K. Round,¹ Susan K. Flowers, and Eric J. Richards²

Department of Biology, Washington University, St. Louis, Missouri 63130

The genetic positions of the five *Arabidopsis thaliana* centromere regions have been identified by mapping size polymorphisms in the centromeric 180-bp repeat arrays. Structural and genetic analysis indicates that 180-bp repeat arrays of up to 1000 kb are found in the centromere region of each chromosome. The genetic behavior of the centromeric arrays suggests that recombination within the arrays is suppressed. These results indicate that the centromere regions of *A. thaliana* resemble human centromeres in size and genomic organization.

Genetic mapping of centromeres is essential for the integration of cytological and genetic maps, and marks an important step toward the molecular characterization of centromeric DNA. Although the centromere is one of the most conspicuous markers on the cytological map, determination of the location of centromeres on genetic maps is frequently difficult to achieve. This is especially true in higher animals and plants where the genetic tools for centromere mapping are limited.

We have been pursuing the characterization of centromere regions of the model angiosperm *Arabidopsis thaliana*. This plant's small genome (~100 Mb/haploid) (Meyerowitz 1994) and relatively low abundance of repetitive DNA [~10% of total (Leutwiler et al. 1984)] make it well suited for molecular chromosome studies. Aiding in this analysis is the availability of dense genetic maps that exist for each of *A. thaliana*'s five chromosomes (e.g., Hauge et al. 1993; Lister and Dean 1993). In addition, physical maps are being developed for all *A. thaliana* chromosomes in the form of overlapping cloned genomic fragments (Schmidt et al. 1995; Zachgo et al. 1996).

Considering the generally advanced molecular characterization of the *A. thaliana* genome, the genomic organization and genetic location of centromeres in this species remain poorly characterized. The *A. thaliana* centromere regions are heterochromatic (Schweizer et al. 1987) and contain tandem arrays of related repeats (exhibiting $\geq 80\%$ similar-

ity) that are ~180 bp in length (Martinez-Zapater et al. 1986; Simoens et al. 1988; Maluszynska and Heslop-Harrison 1991), a genomic organization that resembles the ~170-bp aliphoid repeat arrays at primate centromeres (Willard 1990; Pluta et al. 1995). It is not clear whether the 180-bp repeat arrays flank or span *A. thaliana* centromeres, but data from mammalian systems suggest that the aliphoid repeats are intimately associated with the centromere and are likely to play a role in centromere function (Heartlein et al. 1988; Haaf et al. 1992; Tyler-Smith et al. 1993; Brown et al. 1994; Larin et al. 1994; Harrington et al. 1997). A number of other middle-repetitive sequence elements have been found to be associated with the 180-bp repeats in genomic clones (Richards et al. 1991; Schmidt et al. 1995; Pelissier et al. 1996; Thompson et al. 1996a,b), suggesting that islands of more complex sequence arrangement are located in the *A. thaliana* centromeric regions.

An approximate genetic location of three of the five *A. thaliana* centromeres (on chromosomes 1, 3, and 5) was achieved by Koornneef and coworkers using telocentric mapping techniques (Koornneef 1983; Koornneef and Van der Veen 1983). More recently, physical mapping projects have localized the centromeres on chromosomes 2 and 4 as gaps in the YAC contig maps that are bordered by genomic clones containing 180-bp repeats (Schmidt et al. 1995; Zachgo et al. 1996). We report here the mapping of all five *A. thaliana* centromere regions, identified by large polymorphic 180-bp repeat arrays, on the standard recombinant inbred genetic map. The organization and genetic behavior of the 180-bp arrays will also be discussed.

¹Present address: Department of Genetics, University of Washington, Seattle, Washington.

²Corresponding author.

E-MAIL richards@biodec.wustl.edu; FAX (314) 935-4432.

RESULTS

General Organization of Centromeric Arrays of the 180-bp *HindIII* Repeat

To assess the size and organization of the 180-bp centromeric arrays, we performed genomic Southern analysis using techniques that allow large restriction fragments (up to 2 Mbp) to be examined. We digested megabase-sized *A. thaliana* genomic DNA with various restriction endonucleases and separated the resulting fragments by pulsed-field gel electrophoresis. Restriction fragments containing the 180-bp repeat were detected by Southern blot hybridization.

A number of endonucleases that recognize six base-pair sites yielded large (>400 kb), discrete restriction fragments containing 180-bp repeat arrays. We focused our attention on the 180-bp repeat hybridization patterns generated after cleavage with three endonucleases, *AseI* (ATTAAT), *BclI* (TGATCA), and *NcoI* (CCATGG) (see Fig. 1). These endonucleases were chosen in part because they are insensitive to canonical cytosine methylation (^mCpG & ^mCpNpG) (McClelland et al. 1994), which is abundant in the 180-bp repeat arrays (Martinez-Zapater et al. 1986). Although most of the hybridization signals correspond to fragments <400 kb in size, a number of >400-kb restriction fragments were detected. Restriction sites for the three endonucleases are absent from members of the 180-bp repeat family present in the nucleotide databases (data not shown) suggesting that the large restriction fragments are composed of long, uninterrupted arrays of 180-bp repeats or arrays containing a mixture of 180-bp repeats and other low-complexity repeats. We detected only one of the >400-kb fragments seen in Figure 1 after hybridization with probes corresponding to the two other highly reiterated satellite-type repeats abundant in *A. thaliana*, the so-called 500-bp and 160-bp families (Simoens et al. 1988; Richards et al. 1991) (see Fig. 1), suggesting that most of the large restriction fragments hybridizing with the 180-bp repeat probe are composed primarily of uninterrupted 180-bp repeat arrays.

To investigate this point further, we performed a two-dimensional gel analysis to examine the composition of the >400-kb fragments hybridizing to the 180-bp repeat probe. Figure 2 demonstrates that the large Landsberg *NcoI* 180-bp repeat hybridizing fragments can be cleaved completely to small fragments corresponding to the various multimers of the 180-bp repeat seen in the one-dimensional Landsberg/*HindIII* sample. The variation in the lad-

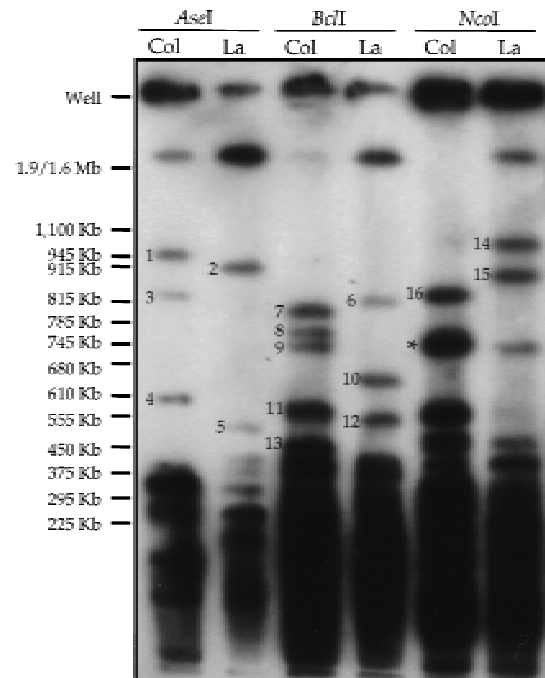


Figure 1 Identification of large polymorphic restriction fragments containing 180-bp centromeric repeat arrays. High-molecular-weight genomic DNA samples from *A. thaliana* strain Columbia (Col) or Landsberg erecta (La) were digested with the restriction enzymes indicated and size-fractionated by pulsed-field gel electrophoresis. After electrophoresis, the DNA was transferred to a nylon membrane and hybridized with a radiolabeled 180-bp repeat probe. The polymorphic fragments used for genetic mapping are numbered 1–16 (see Table 1). The fragment indicated by the asterisk cross-hybridized with 160-bp and 500-bp (telomere-similar) repeat probes. The sizes of yeast chromosomal molecular standards are shown at the left.

der pattern resulting from digestion of different *NcoI* fragments demonstrates a nonrandom distribution of 180-bp repeat variants.

A significant portion of the *A. thaliana* genome is accounted for by the long 180-bp arrays. For example, the hybridization patterns seen in Figure 1 indicate that the large (>400 kb) *BclI* restriction fragments containing 180-bp repeat arrays total over 3 Mbp in length in strain Columbia (see Table 1), corresponding to ~3% of the ~100-Mbp haploid *A. thaliana* genome.

Genetic Mapping of *A. thaliana* Centromere Regions

We took advantage of polymorphisms in the size of large 180-bp repeat array restriction fragments to map the genetic location of the five *A. thaliana* cen-

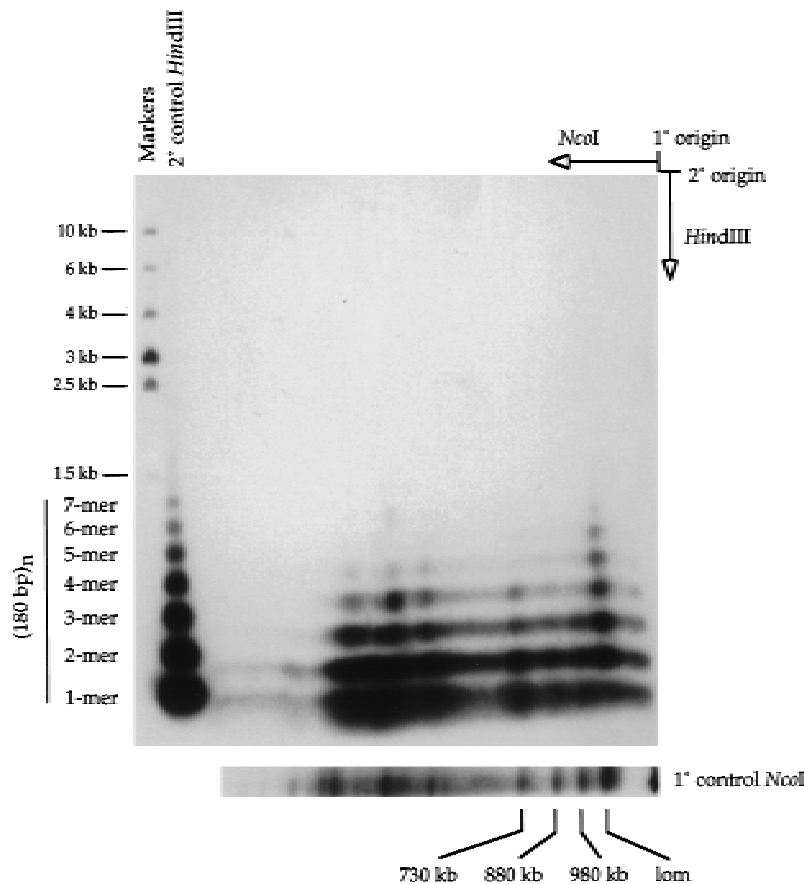


Figure 2 Two-dimensional gel analysis of large *NcoI* fragments containing 180-bp centromere repeats. A high-molecular-weight genomic DNA sample from *A. thaliana* strain Landsberg erecta was digested with *NcoI* and size-fractionated by pulsed-field gel electrophoresis in the first dimension. An excised gel slice containing the separated *NcoI* fragments was then incubated with *HindIII* and placed horizontally at the origin of a 1% agarose gel. After conventional electrophoresis in the second dimension, the fragments were transferred to a nylon membrane and detected by hybridization with a radiolabeled 180-bp repeat probe. *HindIII*-digested Landsberg erecta genomic DNA was run in the second dimension as a control (shown at left with molecular weight markers). Repeat multimers up to 7-mers were detected. The origin of the second-dimension *HindIII* control lane was offset slightly from the first-dimension gel slice, accounting for the faster apparent migration of the sample vs. control multimers. A separate first-dimension *NcoI* control lane is shown at the bottom of the figure. (lom) Limit of mobility.

tromere regions (Willard et al. 1986; Kipling et al. 1994). We first identified size polymorphisms in these large fragments between the two common laboratory strains of *A. thaliana*, Columbia and Landsberg erecta, allowing us to use the Norwich Columbia/Landsberg recombinant inbred (RI) population (Lister and Dean 1993) to map the genetic position of the centromeric repeat arrays. We scored between 29 and 49 independent RI lines for

the large polymorphic fragments seen in Figure 1. A total of 16 polymorphic fragments were scored (Table 1; Fig. 3). Five different segregation patterns were discerned, and, in most cases, candidate allelic pairs could be identified. Genotypes for the five patterns were added to the RI mapping database and a genetic map position for each centromere repeat marker, designated *RCEN#*, was determined by a maximum-likelihood algorithm (Lister and Dean 1993; Arnold and Anderson 1996) (see Fig. 4). The reliability of the *RCEN* marker locations is reflected by the high logarithm of odds (lod) scores, shown in Table 2, which express the relative probability of linkage to a reference marker versus placement at an unlinked location.

The 180-bp array polymorphisms map to five positions, one on each of the five *A. thaliana* chromosomes (Fig. 4) as expected for localized centromere markers. Furthermore, the map positions of the 180-bp array *RCEN* markers are consistent with the available *A. thaliana* centromere mapping information. The map positions of *RCEN2* and *RCEN4* established here correspond well to the centromere locations deduced from the current physical maps of chromosomes 2 and 4. The *RCEN2* marker maps between RFLP markers mi310 and mi421. These markers border gap 1 in the chromosome 2 physical map, which is thought to be the likely position of the centromere (Zachgo et al. 1996). Similarly, the *RCEN4* marker is closely linked to the RFLP marker mi87, which is present on candidate centromeric YAC clones containing 180-bp repeats at the end of contig 1 (Schmidt et al. 1995). The map positions of the 180-bp array *RCEN1* and *RCEN3* markers fall within the genetic intervals containing the centromeres on chromosomes 1 and 3 determined by reference to telocentric chromosome breakpoints. Koornneef and colleagues (Koornneef 1983; Koornneef and Van der Veen 1983) used the telocentric mapping technique to place the centromere of chromosome 1 between the morphological markers *tt-1* and *ch-1*, an ~5-cM interval that contains the mi133 marker (E.J. Richards,

Table 1. Large Polymorphic Restriction Fragments Containing 180-bp Centromeric Repeat Arrays

Fragment no.	Enzyme	Size	Strain origin	Location	Possible alleles (fragment no.)
1	<i>Asel</i>	950 kb	Columbia	<i>RCEN4</i>	
2	<i>Asel</i>	900 kb	Landsberg	<i>RCEN2</i>	3,9,16
3	<i>Asel</i>	820 kb	Columbia	<i>RCEN2</i>	2,10,12,15
4	<i>Asel</i>	590 kb	Columbia	<i>RCEN3</i>	5,6
5	<i>Asel</i>	500 kb	Landsberg	<i>RCEN3</i>	4,8
6	<i>BclI</i>	815 kb	Landsberg	<i>RCEN3</i>	4,8
7	<i>BclI</i>	810 kb	Columbia	<i>RCEN1</i>	14
8	<i>BclI</i>	780 kb	Columbia	<i>RCEN3</i>	5,6
9	<i>BclI</i>	730 kb	Columbia	<i>RCEN2</i>	2,10,12,15
10	<i>BclI</i>	650 kb	Landsberg	<i>RCEN2</i>	3,9,16
11	<i>BclI</i>	570 kb	Columbia	<i>RCEN1</i>	14
12	<i>BclI</i>	550 kb	Landsberg	<i>RCEN2</i>	3,9,16
13	<i>BclI</i>	450 kb	Columbia	<i>RCEN5</i>	
14	<i>NcoI</i>	980 kb	Landsberg	<i>RCEN1</i>	7,11
15	<i>NcoI</i>	880 kb	Landsberg	<i>RCEN2</i>	3,9,16
16	<i>NcoI</i>	830 kb	Columbia	<i>RCEN2</i>	2,10,12,15

unpubl.). A much larger genetic window for centromere 3, extending ~40 cM south of *gl-1* to *tt-5*, was defined by the telocentric strategy (Koornneef 1983; Hauge et al. 1993). The 180-bp array *RCEN3* marker falls within this interval, ~5.5 cM south of the *GL1* genetic marker. Finally, the telocentric strategy indicated that centromere 5 is located in an ~11-cM window between morphological markers *ga-3* and *tt-2* (Koornneef 1983; Hauge et al. 1993). Neither of these markers is present on the current recombinant inbred genetic map, yet the localization of the 180-bp array *RCEN5* marker in the center of the chromosome is consistent with the position of these morphological markers in the center of the integrated genetic map of *A. thaliana* chromosome 5. Moreover, *RCEN5* corresponds well to the centromere location estimated by the position of repetitive DNA containing clones on the emerging physical map of chromosome 5 (Murata et al. 1997; S. Tabata, pers. comm.; www.kazusa.or.jp/arabi).

DISCUSSION

The identification of large polymorphic restriction fragments containing 180-bp repeats allowed us to determine the genetic locations of the centromeric repeat arrays in *A. thaliana*. A number of considerations indicate that the mapped centromeric repeats accurately predict the location of the five *A. thaliana* centromeres. Chromosome in situ hybridization ex-

periments localized the 180-bp repeat exclusively to the primary constrictions and failed to detect interstitial 180-bp repeat loci in the chromosome arms (Maluszynska and Heslop-Harrison 1991). In addition, the 180-bp repeat RFLP markers map to locations consistent with centromere mapping data developed using telocentric approaches or correlations with the gaps in the current physical maps as described in the Results. Moreover, the centromere map positions defined here fall within the genetic intervals containing the meiotic *A. thaliana* centromeres that have been defined recently by tetrad analysis (G. Copenhaver and D. Preuss, pers. comm.).

Beyond defining useful cytogenetic landmarks on the *A. thaliana* map, the data presented here provide information on the genomic organization of a centromere from a eukaryotic model organism. The *A. thaliana* centromere regions are rich in repetitive DNAs, which is a hallmark of the molecularly characterized centromere regions from humans (Willard and Wayne 1987), mice (Kipling et al. 1991; Narayanswami et al. 1992), maize (Alfenito and Birchler 1993; Kaszás and Birchler 1996), *Drosophila* (Le et al. 1995; Murphy and Karpen 1995), *Neurospora* (Centola and Carbon 1994), and fission yeast (Nakaseko et al. 1986; Clarke 1990). *A. thaliana* centromeres contain long arrays of 180-bp repeats with little or no interspersions, in addition to the more complex arrays of 180-bp repeats mixed with other repeat

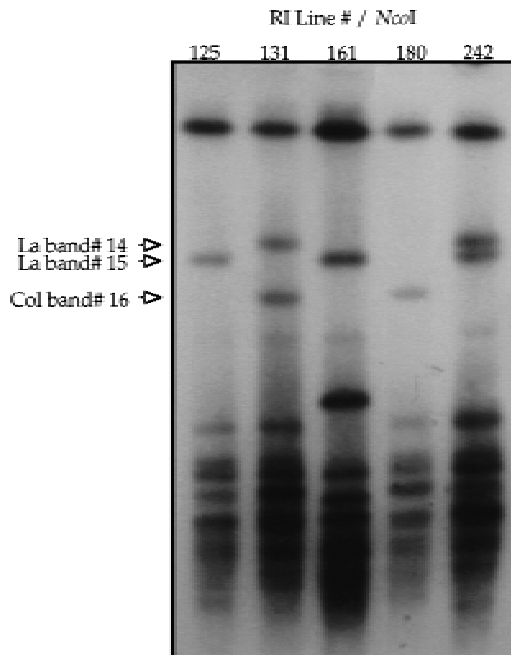


Figure 3 Segregation of large polymorphic restriction fragments containing 180-bp centromeric repeat arrays in a recombinant inbred mapping population. High-molecular-weight genomic DNA samples from five *A. thaliana* Columbia (Col)/Landsberg (La) RI lines were digested with *NcoI* and size-fractionated by pulsed-field gel electrophoresis. After electrophoresis, the DNA was transferred to a nylon membrane and hybridized with a radiolabeled 180-bp repeat probe. Dominant markers corresponding to fragments 14 and 15 segregated independently. The inheritance of the markers represented by fragments 15 and 16 is mutually exclusive as expected for allelic markers in RI populations (Lister and Dean 1993).

elements characterized previously (Richards et al. 1991; Pelissier et al. 1996; Thompson et al. 1996a,b). A conservative estimate of the size of the uninterrupted 180-bp repeat arrays is ~0.6 Mbp per centromere assuming an equal distribution of the ~3 Mbp of uninterrupted 180-bp repeat arrays (summing the *BclI* fragment lengths; Table 1) among the five centromeres. However, in some cases we can assign >1 Mbp of uninterrupted arrays to a single centromere; for example, centromere 1 (810 + 570 kb in Columbia). Our estimates of the size of the *A. thaliana* centromere regions are comparable to the size of the alphoid repetitive arrays at human centromeres (Tyler-Smith and Brown 1987; Willard and Wayne 1987; Wevrick and Willard 1989, 1991; Cooper et al. 1993; Jackson et al. 1993).

The size and composition of the 180-bp repeat arrays make it unlikely that a complete *A. thaliana*

centromere can be isolated in a single genomic clone or accurately represented by overlapping clones, using available molecular cloning technology (Neil et al. 1990). The current physical maps of the *A. thaliana* genome based on overlapping genomic clones do not represent the large 180-bp arrays described here, and we believe that the centromeres are missing from these maps (Schmidt et al. 1995; Zachgo et al. 1996). It may be necessary to analyze *A. thaliana* centromeres that have been pared down in subchromosomal derivatives rather than rely on assembly of centromeric regions from constituent genomic clones (Tyler-Smith et al. 1993; Brown et al. 1994; Le et al. 1995; Murphy and Karpen 1995; Kaszás and Birchler 1996).

Our analysis of the large polymorphic restriction fragments carrying the 180-bp repeat arrays in the recombinant inbred lines provides insight into the genetic behavior of *A. thaliana* centromere regions. No nonparental fragments were detected in our analysis, which scored 677 large 180-bp array fragments distributed over 54 RI lines. The absence of nonparental fragments indicates that the centromere repeat arrays are more stable than the long arrays of ribosomal RNA genes [nucleolus organizing regions (NOR)], which have been observed to undergo genomic rearrangements in the recombinant inbred mapping material used here (Copenhaver et al. 1995). Nonparental fragments could also arise from homologous recombination events, including unequal exchanges leading to expansion or contraction of direct repeat arrays, intrachromosomal exchanges leading to deletions, or reciprocal crossover events within the polymorphic arrays. The paucity of homologous recombination events involving the centromeric 180-bp repeat arrays suggests that recombination is suppressed within *A. thaliana* centromere repeat arrays. Similar conclusions were drawn from the meiotic and mitotic stability of centromeric alphoid repeat arrays in humans (Wevrick and Willard 1989). Recombinational suppression in centromere regions has been observed in other organisms, including mice (Kipling et al. 1994), *Neurospora* (Centola and Carbon 1994), and fission yeast, *Schizosaccharomyces pombe* (Nakaseko et al. 1986). The *A. thaliana* centromere repeats arrays cannot be completely genetically inert, however, as evidenced by the nonrandom distribution of 180-bp variants seen in Figure 2. Similar regional clustering of repeat variants is found in centromeric repetitive DNA arrays in humans (Warburton and Willard 1990) and is thought to reflect intermediates in a recombination-driven homogenization process.

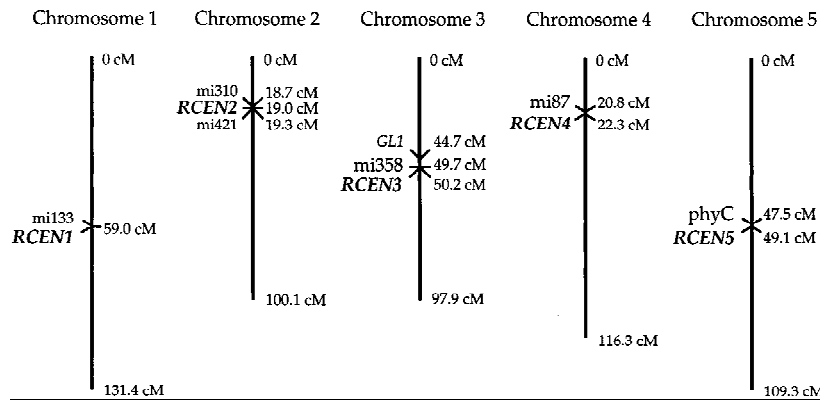


Figure 4 Genetic location of the *A. thaliana* centromere regions. The locations (in cM) of the centromeric repeat (RCEN) array polymorphisms on the Columbia/Landsberg recombinant inbred map (March 1997 version; <http://nasc.nott.ac.uk/>) are shown along with relevant reference markers (see text). The mi# markers are as described (Liu et al. 1996). The RCEN markers may be identified on earlier versions of the RI map by their provisional names, EKR#.

It is clear from our initial characterization of the centromere repeat arrays that *A. thaliana* centromeres are large, complex regions that resemble their mammalian counterparts. The results described here lay a foundation for further characterization of the centromeres of a model higher eukaryotic organism.

METHODS

Preparation of Large Genomic DNA

Protoplasts from axenic seedling cultures (~2 weeks old) were prepared and embedded in agarose molds as described (Copenhaver et al. 1995). The embedded protoplasts were lysed and digested at 55°C in two overnight incubations in a solution of 0.5 M EDTA, 1% N-laurylsarcosine, 1 mM Tris at pH 9.5, 1 mg/ml of Pronase (Sigma). The embedded genomic DNA was stored at 4°C in lysis solution minus pronase.

Restriction Endonuclease Digestion and Size Separations

Agarose-embedded DNA samples were treated at 55°C with 100 µg/ml of phenylmethylsulfonyl fluoride for 2 hr to inactivate the protease. The treated samples were then washed in 10 mM Tris-Cl, 10 mM EDTA at pH 8.0 before restriction endonuclease digestion. Agarose-embedded DNA samples were digested with 50–100 units of restriction endonuclease (New England Biolabs) in 300 µl of the appropriate buffer + 100 µg/ml of bovine serum albumin + 2.5 mM spermidine. Digested DNA was size-fractionated through a 1% agarose gel (FastLane, FMC) in 0.5× TBE buffer (89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA) using a CHEF (Chu et al. 1986) pulsed-field gel apparatus (CHEF-DRII, Bio-Rad) and the fol-

lowing conditions: 60 sec switch time, 12 hr, 200 V → 90 sec switch time, 9 hr, 200 V at -15°C.

Two-Dimensional Gel Electrophoresis

Landsberg genomic DNA was digested with *NcoI* (New England Biolabs) and size-fractionated on a 1% low-melting point agarose (SeaPlaque, FMC) pulsed-field gel as described above. An excised lane from the first-dimension gel was equilibrated in restriction buffer and subsequently incubated overnight at 37°C after addition of two aliquots of 1000 units of *HindIII* (New England Biolabs). The gel slice was then equilibrated in 0.5× TBE, cast at the origin of a 1% agarose gel (SeaKem, FMC), and subjected to conventional gel electrophoresis in 0.5× TBE.

Southern Hybridization/Probes

Size-fractionated DNA was depurinated in 0.25 N HCl for 10–15 min, denatured in 0.4 N NaOH, 3 M NaCl and equilibrated in 8 mM NaOH, 3 M NaCl. The nucleic acid was then transferred using a downward capillary apparatus (Scheicher & Schuell) to an uncharged nylon membrane (GeneScreen, DuPont/NEN). After transfer, the membrane was neutralized in 50 mM Na phosphate buffer (pH 6.5) and the nucleic acids were immobilized by UV cross-linking. Hybridizations were carried out in the high-SDS buffer as described (Church and Gilbert 1984). The 180-bp centromere repeat hybridization probe was generated by random-priming DNA synthesis (Ausubel et al. 1987) using the cloned 180-bp repeat monomer from pARR20-1 (E.J. Richards, unpubl.) as a template. The 500-bp repeat hybridization probe was generated by random-priming DNA synthesis using the cloned telomeric insert from pAtT4.1 (Richards and Ausubel 1988; Richards et al. 1991) as a template. The 160-bp repeat hybridization probe was generated by 5'-endlabeling (Ausubel et al. 1987) an oligonucleotide of the following sequence: 5'-TGTTAGTGTTCATGGTCA-3' (Simoens et al. 1988). High-

Table 2. Iod Scores Assessing the Significance of Linkage of the RCEN Markers to Reference Markers

RCEN marker	Reference marker	Iod score
RCEN1	mi133	11.4
RCEN2	mi310	14.1
RCEN2	mi421	14.1
RCEN3	mi358	14.4
RCEN4	mi87	13.2
RCEN5	phyC	4.5

The scores were calculated from two-point mapping data and using a linked recombination frequency of 0.04 cM.

stringency washes in $0.2 \times$ SSC, 0.1% SDS at 60°C were used for the 180-bp repeat hybridizations. Even under high-stringency conditions, the 180-bp repeat probe does not exhibit chromosome-specificity, which is characteristic of some human alphoid repeat probes (Willard and Wayne 1987). Lower stringency wash conditions were used to detect the 500-bp repeat ($2 \times$ SSC, 0.1% SDS at 60°C) and the 160-bp repeat ($2 \times$ SSC, 0.1% SDS at 45°C). Hybridization signals were detected by autoradiography or PhosphorImager (Molecular Dynamics).

ACKNOWLEDGMENTS

This work was supported by a grant from the United States Department of Agriculture (CRG 94-37300-0295). We thank G. Copenhaver for DNA preparations and technical advice. We thank M. Anderson, M. Arnold, C. Dean, and C. Lister for RI mapping resource management (see <http://nasc.nott.ac.uk/>) and the Nottingham *Arabidopsis* Stock Centre for recombinant inbred stocks.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

NOTE ADDED IN PROOF

We wish to cite recently published cytogenetic analysis of centromeric repetitive DNA organization in *A. thaliana* by Brandes et al. (1997), as well as previous molecular analysis of *Arabidopsis* centromeric 180-bp arrays published by Murata et al. (1994).

REFERENCES

- Alfenito, M.R. and J.A. Birchler. 1993. Molecular characterization of a maize B chromosome centric sequence. *Genetics* 135: 589–597.
- Arnold, M. and M. Anderson. 1996. Mapping and map interpretation—A learning experience. *Weeds World* 3iii: 7–8.
- Asubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1987. *Current protocols in molecular biology*. John Wiley/Greene, New York, NY.
- Brandes, A., H. Thompson, C. Dean, and J.S. Heslop-Harrison. 1997. Multiple repetitive DNA sequences in the paracentromeric regions of *Arabidopsis thaliana* L. *Chrom. Res.* 5: 238–246.
- Brown, K.E., M.A. Barnett, C. Burgtorf, P. Shaw, V.J. Buckle, and W.R.A. Brown. 1994. Dissecting the centromere of the human Y chromosome with cloned telomeric DNA. *Hum. Mol. Genet.* 3: 1227–1237.
- Centola, M. and J. Carbon. 1994. Cloning and characterization of centromeric DNA from *Neurospora crassa*. *Mol. Cell. Biol.* 14: 1510–1519.
- Chu, G., D. Vollrath, and R.W. Davis. 1986. Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science* 234: 1582–1585.
- Church, G.M. and W. Gilbert. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci.* 81: 1991–1995.
- Clarke, L. 1990. Centromeres of budding and fission yeasts. *Trends Genet.* 6: 150–154.
- Cooper, K.F., R.B. Fisher, and C. Tyler-Smith. 1993. Structure of the sequences adjacent to the centromeric alphoid satellite DNA array on the human Y chromosome. *J. Mol. Biol.* 230: 787–799.
- Copenhaver, G.P., J.H. Doelling, J.S. Gens, and C.S. Pikaard. 1995. Use of RFLPs larger than 100 kbp to map the position and internal organization of the nucleolus organizer region on chromosome 2 in *Arabidopsis thaliana*. *Plant J.* 7: 273–286.
- Haaf, T., P.E. Warburton, and H.F. Willard. 1992. Integration of human α -satellite DNA into simian chromosomes: Centromere protein binding and disruption of normal chromosome segregation. *Cell* 70: 681–696.
- Harrington, J.J., G.V. Bokkelen, R.W. Mays, K. Gustashaw, and H.F. Willard. 1997. Formation of de novo centromeres and construction of first-generation human artificial chromosomes. *Nature Genet.* 15: 345–355.
- Hauge, B., S.M. Hanley, S. Cartinhour, J.M. Cherry, H.M. Goodman, M. Koornneef, P. Stam, C. Chang, S. Kempin, L. Medrano, and E.M. Meyerowitz. 1993. An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. *Plant J.* 3: 745–754.
- Heartlein, M.W., J.H.M. Knoll, and S.A. Latt. 1988. Chromosome instability associated with human alphoid DNA transfected into the Chinese hamster genome. *Mol. Cell. Biol.* 8: 3611–3618.
- Jackson, M.S., P. Slijepcevic, and B.A. Ponder. 1993. The organisation of repetitive sequences in the pericentromeric region of human chromosome 10. *Nucleic Acids Res.* 21: 5865–5874.
- Kaszás, É. and J.A. Birchler. 1996. Misdivision analysis of centromere structure in maize. *EMBO J.* 15: 5246–5255.
- Kipling, D., H.E. Ackford, B.A. Taylor, and H.J. Cooke. 1991. Mouse minor satellite DNA genetically maps to the centromere and is physically linked to the proximal telomere. *Genomics* 11: 235–241.
- Kipling, D., H.E. Wilson, A.R. Mitchell, B.A. Taylor, and H.J. Cooke. 1994. Mouse centromere mapping using oligonucleotide probes that detect variant of the minor satellite. *Chromosoma* 103: 46–55.
- Koornneef, M. 1983. The use of telotrisomics for centromere mapping in *Arabidopsis thaliana* (L.) Heynh. *Genetica* 62: 33–40.

- Koornneef, M. and J.H. Van der Veen. 1983. Trisomics in *Arabidopsis thaliana* and the location of linkage groups. *Genetica* 61: 41–46.
- Larin, Z., M.D. Fricker, and C. Tyler-Smith. 1994. *De novo* formation of several features of a centromere by introduction of a Y alphoid YAC into mammalian cells. *Hum. Mol. Genet.* 3: 689–695.
- Le, M.-H., D. Duricka, and G.H. Karpen. 1995. Islands of complex DNA are widespread in *Drosophila* centric heterochromatin. *Genetics* 141: 283–303.
- Leutwiler, L.S., B.R. Hough-Evans, and E.M. Meyerowitz. 1984. The DNA of *Arabidopsis thaliana*. *Mol. Gen. Genet.* 194: 15–23.
- Lister, C. and C. Dean. 1993. Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. *Plant J.* 4: 745–750.
- Liu, Y.G., N. Mitsukawa, C. Lister, C. Dean, and R.F. Whittier. 1996. Isolation and mapping of a new set of 129 RFLP markers in *Arabidopsis thaliana* using recombinant inbred lines. *Plant J.* 10: 733–736.
- Maluszynska, J. and J.S. Heslop-Harrison. 1991. Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. *Plant J.* 1: 159–166.
- Martinez-Zapater, J., M.A. Estelle, and C.R. Somerville. 1986. A highly repeated DNA sequence in *Arabidopsis thaliana*. *Mol. Gen. Genet.* 204: 417–423.
- McClelland, M., M. Nelson, and E. Raschke. 1994. Effect of site-specific modification on restriction endonuclease and DNA modification methyltransferases. *Nucleic Acids Res.* 22: 3640–3659.
- Meyerowitz, E.M. 1994. Structure and organization of the *Arabidopsis thaliana* nuclear genome. In *Arabidopsis* (ed. E.M. Meyerowitz and C.R. Somerville), pp. 21–36. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Murata, M., Y. Ogura, and F. Motoyoshi. 1994. Centromeric repetitive sequence in *Arabidopsis thaliana*. *Jpn. J. Genet.* 69: 361–370.
- Murata, M., J.S. Heslop-Harrison, and F. Motoyoshi. 1997. Physical mapping of the 5S ribosomal RNA genes in *Arabidopsis thaliana* by multi-color fluorescence in situ hybridization with cosmid clones. *Plant J.* 12: 31–37.
- Murphy, T.D. and G.H. Karpen. 1995. Localization of centromere function in a *Drosophila* minichromosome. *Cell* 82: 599–609.
- Nakaseko, Y., Y. Adachi, S.-I. Funahashi, O. Niwa, and M. Yanagida. 1986. Chromosome walking shows a highly homologous repetitive sequence present in all the centromere regions of fission yeast. *EMBO J.* 5: 1011–1021.
- Narayanswami, S., N.A. Doggett, L.M. Clark, C.E. Hildebrand, H.-U. Weier, and B.A. Hamkalo. 1992. Cytological and molecular characterization of centromeres in *Mus domesticus* and *Mus spretus*. *Mamm. Genome* 2: 186–194.
- Neil, D., A. Villasante, R. Fisher, D. Vetrie, B. Cox, and C. Tyler-Smith. 1990. Structural instability of human tandemly repeated DNA sequences cloned in yeast artificial chromosome vectors. *Nucleic Acids Res.* 18: 1421–1428.
- Pelissier, T., S. Tutois, S. Tourmente, J.M. Deragon, and G. Picard. 1996. DNA regions flanking the major *Arabidopsis thaliana* satellite are principally enriched in *Athila* retroelement sequences. *Genetica* 97: 141–151.
- Pluta, A.F., A.M. Mackay, A.M. Ainsztein, I.G. Goldberg, and W.C. Earnshaw. 1995. The centromere: Hub of chromosomal activities. *Science* 270: 1591–1594.
- Richards, E.J. and F.M. Ausubel. 1988. Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* 53: 127–136.
- Richards, E.J., H.M. Goodman, and F.M. Ausubel. 1991. The centromere region of *Arabidopsis thaliana* chromosome 1 contains telomere-similar sequences. *Nucleic Acids Res.* 19: 3351–3357.
- Schmidt, R., J. West, K. Love, Z. Lenehan, C. Lister, H. Thompson, D. Bouchez, and C. Dean. 1995. Physical map and organization of *Arabidopsis thaliana* chromosome 4. *Science* 270: 480–483.
- Schweizer, D., P. Ambros, P. Grundler, and F. Varga. 1987. Attempts to relate cytological and molecular chromosome data of *Arabidopsis thaliana* to its genetic linkage map. *Arabidopsis Inf. Serv.* 25: 27–34.
- Simoens, C.R., J. Gielen, M.V. Montagu, and D. Inze. 1988. Characterization of highly repetitive sequences of *Arabidopsis thaliana*. *Nucleic Acids Res.* 16: 6753–6766.
- Thompson, H., R. Schmidt, A. Brandes, J.S. Heslop-Harrison, and C. Dean. 1996a. A novel repetitive sequence associated with the centromeric regions of *Arabidopsis thaliana* chromosomes. *Mol. Gen. Genet.* 253: 247–252.
- Thompson, H.L., R. Schmidt, and C. Dean. 1996b. Identification and distribution of seven classes of middle-repetitive DNA in the *Arabidopsis thaliana* genome. *Nucleic Acids Res.* 24: 3017–3022.
- Tyler-Smith, C. and W.R.A. Brown. 1987. Structure of the major block of alphoid satellite DNA on the human Y chromosome. *J. Mol. Biol.* 195: 457–470.
- Tyler-Smith, C., R.J. Oakey, Z. Larin, R.B. Fisher, M. Crocker, N.A. Affara, M.A. Ferguson-Smith, M. Muenke, O. Zuffardi, and M.A. Jobling. 1993. Localization of DNA sequences required for human centromere function through an analysis of rearranged Y chromosomes. *Nature Genet.* 5: 368–375.
- Warburton, P.E. and H.F. Willard. 1990. Genomic analysis of sequence variation in tandemly repeated DNA: Evidence

for localized homogeneous sequence domains within arrays of α -satellite DNA. *J. Mol. Biol.* 216: 3–16.

Wevrick, R. and H.F. Willard. 1989. Long-range organization of tandem arrays of α satellite DNA at the centromeres of human chromosomes: High-frequency array-length polymorphism and meiotic stability. *Proc. Natl. Acad. Sci.* 86: 9394–9398.

———. 1991. Physical map of the centromeric region of human chromosome 7: Relationship between two distinct alpha satellite arrays. *Nucleic Acids Res.* 19: 2295–2301.

Willard, H.F. 1990. Centromeres of mammalian chromosomes. *Trends Genet.* 6: 410–416.

Willard, H.F. and J.S. Wayne. 1987. Hierarchical order in chromosomal-specific human alpha satellite DNA. *Trends Genet.* 3: 192–198.

Willard, H.F., J.S. Waye, M.H. Skolnick, C.E. Schwartz, V.E. Powers, and S.B. England. 1986. Detection of restriction fragment length polymorphisms at the centromeres of human chromosomes by using chromosome-specific α satellite DNA probes: Implication for development of centromere-based genetic linkage maps. *Proc. Natl. Acad. Sci.* 83: 5611–5615.

Zachgo, E.A., M.L. Wang, J. Dewdney, D. Bouchez, C. Camilleri, S. Belmonte, L. Huang, M. Dolan, and H.M. Goodman. 1996. A physical map of chromosome 2 of *Arabidopsis thaliana*. *Genome Res.* 6: 19–25.

Received July 11, 1997; accepted in revised form September 25, 1997.