

# Analysis of 5S rDNA Arrays in *Arabidopsis thaliana*: Physical Mapping and Chromosome-Specific Polymorphisms

C. Cloix,<sup>1</sup> S. Tutois,<sup>1</sup> O. Mathieu, C. Cuvillier, M.C. Espagnol, G. Picard, and S. Tourmente<sup>2</sup>

Unité Mixte de Recherche, 6547 BIOMOVE, Université Blaise Pascal, 63177 Aubière Cedex, France

A physical map of a pericentromeric region of chromosome 5 containing a 5S rDNA locus and spanning ~1000 kb was established using the CIC YAC clones. Three 5S rDNA arrays were resolved in this YAC contig by PFGE analysis and we have mapped different types of sequences between these three blocks. 5S rDNA units from each of these three arrays of chromosome 5, and from chromosomes 3 and 4, were isolated by PCR. A total of 38 new DNA sequences were obtained. Two types of 5S rDNA repeated units exist: the major variant with 0.5-kb repeats and one with short repeats (251 bp) only detected on YAC IIA3 from chromosome 3. Although the 38 sequences displayed noticeable heterogeneity, we were able to group them according to their 5S array origin. The presence of 5S array-specific variants was confirmed with the restriction polymorphism study of all the YACs carrying 5S units.

[The sequence data described in this paper have been submitted to the EMBL data library under accession nos. AFI98182–AFI98224.]

The 5S ribosomal DNA in higher eukaryotes is generally organized in tandem arrays, the repeated unit of which contains the transcription unit and a spacer sequence. These tandem arrays may be localized on either a single or several chromosomes and are separated from the genes encoding the large rRNAs (Appels et al. 1980; Long and Dawid 1980; Ellis et al. 1988).

*Arabidopsis thaliana* is known to contain ~1000 copies of 5S rDNA per haploid genome and these occur in tandem arrays (Campbell et al. 1992). Within the 497-bp unit lies a 121-bp fragment that shows >95% homology to the angiosperm consensus 5S rDNA transcribed sequence. There is, however, considerable sequence and length variation in the intergenic spacer regions. Nonetheless, within a species the 5S rDNA units are virtually identical in sequence, indicating that 5S rRNA repeats evolve as a group, a phenomenon known as concerted evolution.

5S rDNA FISH experiments have been reported on both mitotic (Murata et al. 1997) and meiotic chromosomes (Fransz et al. 1998). Using a 5S rDNA probe, these investigators were able to distinguish individual chromosomes and show the polymorphism between different chromosomes. All ecotypes tested possess one major 5S rDNA locus in the short arm of chromosome 4 and the largest one in the upper arm of chromosome

5 [Columbia (Col), C24, Cape Verde Islands (Cvi), Kashmir 1 (Kas), Landsberg erecta (Ler), Wassilekija (WS)]. Both loci map to the pericentromeric heterochromatin regions of these chromosomes. A minor 5S rDNA locus in the lower arm of chromosome 5 was detected in Col, C24, Ler, and WS, and a third major locus in chromosome 3 was detected in some ecotypes (Col, Cvi, Kas, and Ler). In Ler, this locus is found in the middle of the long arm (3L), whereas it maps to the pericentromeric heterochromatin region of chromosome 3 in Col, Cvi, and Kas. A study of the molecular organization of pericentromeric repeats with extended DNA fibers has revealed long tandem arrays of 5S rDNA genes (Fransz et al. 2000). The 5S rDNA repeat appears to hybridize in the pericentromeric heterochromatin of the chromosome 4 short arm, adjacent to the central pAL1 (180-bp satellite) domain. In all chromosomes, 106B, which has DNA sequence homology with the long terminal repeat (LTR) of the *Athila* retroelement, colocalizes with pAL1 in the center of the heterochromatic regions. The 17A20 repeats map to the heterochromatic blocks and colocalize partly with 5S rDNA in chromosomes 3, 4, and 5.

Mozo et al. (1999) have established a complete BAC-based physical map of the *Arabidopsis* genome, but gaps remain at centromeres and 5S locus. Previously, we (Tutois et al. 1999) have established an ~1000-kb physical map of the chromosome 5 pericentromeric region anchored to the left arm YAC contig generated by Schmidt et al. (1997) and Kotani et al. (1997). One 5S block was present on the right side of

<sup>1</sup>These authors contributed equally to this work.

<sup>2</sup>Corresponding author.

E-MAIL Sylvette.TOURMENTE@geem.univ-bpclermont.fr; FAX (33) 4 73 40 77 77.

the contig (centromere-proximal side). In this paper, we present a YAC contig going through this 5S locus localized in the pericentromeric region of the left arm of chromosome 5 in the Columbia ecotype. Three 5S rDNA arrays were resolved in this locus and their length analyzed by pulsed-field experiments. Several 5S rDNA units from each array were subcloned and sequenced, two 5S length-variants were identified, and sequences were grouped according to their 5S array origin. Restriction site polymorphism was analyzed with 14 enzymes on all the YACs carrying 5S units (5S YACs) from the CIC library.

**RESULTS**

5S rDNA units occur in pericentromeric regions (Thompson et al. 1996; Tutois et al. 1999; this paper), which are largely composed of repetitive elements, thus greatly complicating the construction of contigs. Nevertheless, using many probes, including some that occur in tandem arrays, we have been able to construct a YAC contig going through the 5S locus present in the pericentromeric region of chromosome 5 left arm. We report the analysis of 5S rDNA units from the different chromosomal locations (chromosomes 3, 4, and 5) at both sequence and restriction polymorphism levels.

**YAC Contig Establishment: Three 5S rDNA Arrays in the Contig**

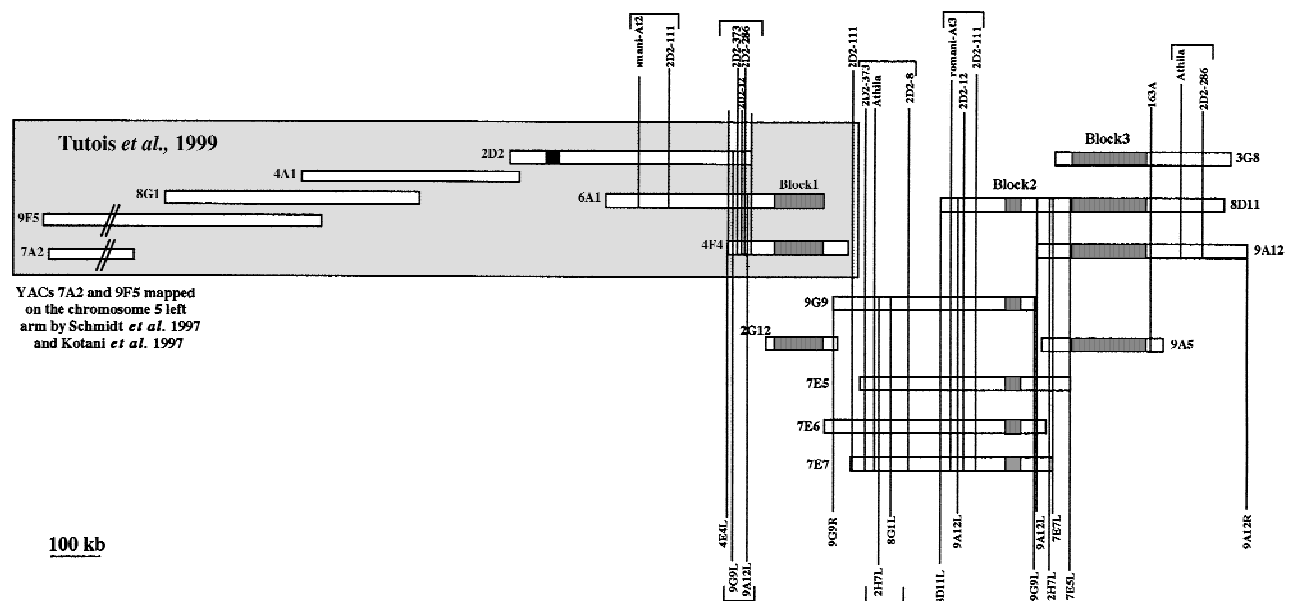
A total of 35 YAC clones from the CIC library (Creusot et al. 1995) carrying 5S rDNA units have been identified. Figure 1 shows both our preceding YAC contig (Tutois et al. 1999) and that described here. Because

our preceding YAC contig ended on its centromere-proximal side with a 5S rDNA block (Fig. 1, YACs 4E4, 6A1), DNA from 30 5S CIC YAC clones was purified and *EcoRI*-digested. Southern blots were hybridized with numerous probes to generate a YAC contig. The probes, all corresponding to repeated sequences, were *EcoRI* subclones of YAC 2D2 obtained previously (8, 12, 104, 111, 286, 373, r-At3; Fig. 1; Table 1) and YAC end probes, sequenced and quantified, of different 5S YACs (Table 1). Overlapping YACs share common restriction fragment(s) revealed by a probe and each overlap was confirmed with other probes.

The resulting YAC contig spanning ~1000 kb with three 5S rDNA arrays labeled 1, 2, 3 is shown in Figure 1. We have placed different non-5S sequences relative to the three 5S rDNA blocks and infer the nature of the sequences present between the 5S rDNA blocks (Table 1). We were not able to order some sequences and these are presented between brackets in Figure 1.

**Pulsed Field Electrophoresis: Size of the Three 5S rDNA Clusters**

To assess the size of the 5S rDNA clusters, we tested enzymes *SacI*, *SacII*, *XhoI*, and *XbaI* based on the absence of their recognition sites in the 5S rDNA published sequence (Campell et al. 1992). If virtually identical units are arranged in simple tandem arrays, enzymes that do not cleave the 5S rDNA should yield fragments containing unit 5S rDNA blocks. Recombinant yeast DNA was prepared, embedded in agarose, and cut with *SacI*, *SacII*, *XhoI*, and *XbaI*, and the resulting DNA fragments were resolved by PFGE. 5S rDNA



**Figure 1** Graphic representation of the YACs contig, including contig published previously (Tutois et al. 1999). Shaded boxes represent 5S blocks. The black box represents a 180-bp satellite block. Markers above the contig represent 2D2 subclones (Tutois et al. 1999). Markers below the contig represent YACs end probes left (L) or right (R). Markers between brackets are not ordered.

**Table 1.** Probes Used for the Physical Mapping of the 5S rDNA Locus of Chromosome 5

Probe	Copy no.	Hybridization pattern	Reference
2D2-8	70	tandem repeat	Tutois et al. (1999)
2D2-12	22	tandem repeat	
2D2-286	5	dispersed	
2D2-373	110	dispersed	
8G1-L	N.D.	dispersed	
2D2-111	5	transposable element	
romani-At2	5	transposable element	
romani-At3	27	transposable element	
2D2-104	90	centromeric sequence = 163A	
2H7-L	20	centromeric sequence = 278A	
4H7-R	150	transposable element <i>Athila</i>	Thompson et al. (1996)
4E4-L	160	tandem repeat	Péliissier et al. (1995)
7E5-L	N.D.	dispersed	this paper
7E7-L	N.D.	dispersed	
8D11-L	280	dispersed	
9A12-L	20	dispersed	
9A12-R	50	dispersed	
9G9-L	200	dispersed	

was detected by Southern blotting followed by hybridization using a radioactive 5S rDNA probe.

The restriction endonuclease *SacI* confirmed that YACs 2G12, 4E4, and 6A1 from block 1 were carrying the same 5S rDNA block spanning around 100 kb (Fig. 2). The same results were obtained with the enzymes *XhoI* and *XbaI*. Surprisingly, no hybridization was ob-

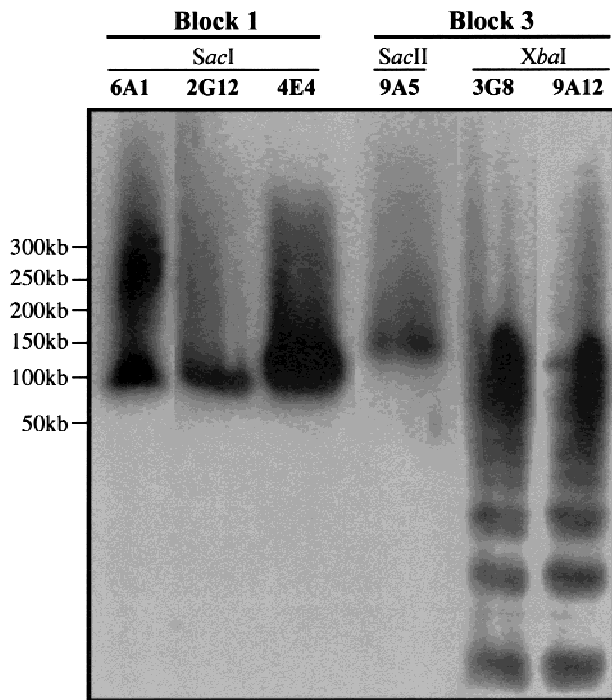
tained for either block 2 (YACs 7E5-E7, 9G9) or block 3 (YACs 3G8, 8D11, 9A5, 9A12) after *SacI* digestion, suggesting a digestion of these 5S rDNA blocks, leading to monomer (497 bp) or lower multimer units, not visualized by PFGE. With *SacII*, a 5S rDNA block of around 30 kb was observed for YACs 7E6 and 7E7 (block 2) using standard electrophoresis with a 0.4% agarose gel and with *XbaI* this 30-kb 5S array was confirmed. With *SacII*, a 150-kb 5S array was obtained for YACs from block 3 (9A5, Fig. 2) although with *XbaI* a ladder of bands was obtained suggesting that *XbaI* digested some 5S rDNA units of block 3 (3G8, 9A12, Fig. 2).

These results demonstrated the presence of three 5S rDNA blocks, 100-, 30-, and 150-kb long, arranged from euchromatin towards centromere. A restriction polymorphism between these three blocks was revealed. The presence of the same 5S rDNA block, at both length and restriction polymorphism levels, on the different overlapping YACs confirmed the contig.

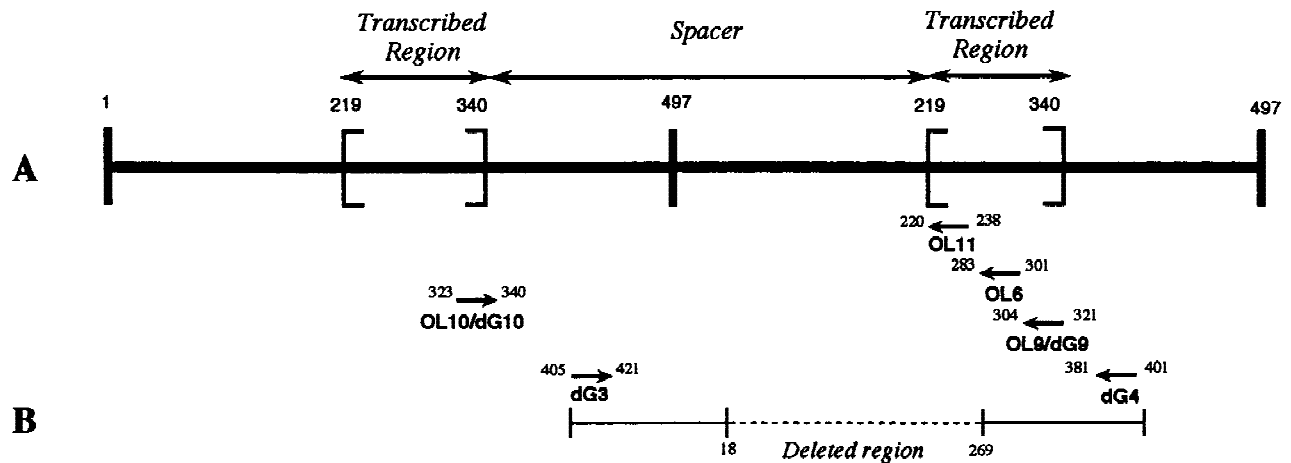
**PCR-Based Isolation of Members of the Different 5S rDNA Arrays**

To isolate 5S rDNA sequences from the different 5S loci of the *Arabidopsis thaliana* genome, CIC YACs were used as template DNA for PCR. We used 5S YACs mapped on our contig of chromosome 5, that is, 6A1, 4E4, block 1; 7E6, 7E7, block 2; 9A5, 9A12, block 3; 11A3 mapped on chromosome 3 (Camilleri et al. 1998); and 9D3 mapped on chromosome 4 (Schmidt et al. 1995).

Different sets of primers were used, based on the consensus sequence of Campell et al. (1992), and the positions of the primers are given in Figure 3A. To avoid a selective amplification of one population of 5S



**Figure 2** Identification of the size of the three 5S arrays of chromosome 5 by pulsed field electrophoresis. Prior to PFGE, YACs were digested with *SacI* (2G12, 4E4, 6A1, block1), with *SacII* (9A5, block3), and *XbaI* (3G8, 9A12, block 3).



**Figure 3** Positions of the primers used for PCR amplifications. Two consecutive 5S rDNA units are represented. (A) The PCR products are expected to be 412 bp with OL10–OL11; 497 bp with OL10–OL9 and dG10–dG9; 476 bp with OL10–OL6, and 495 bp with dG3–dG4. (B) The short 251-bp variant was amplified with dG3–dG4.

units, two sets of degenerate primers (dG) were also chosen:

1. OL10–OL9(dG10–dG9) and OL10–OL6—these sets were expected to amplify nearly complete units of 497 and 476 bp between the 3' end of the transcribed region of two adjacent 5S units. Primers are situated in the transcribed region.
2. OL10–OL11—this primer set (with primers localized respectively at the end and the beginning of the transcribed region of two adjacent 5S rDNA units) flanks the whole spacer region of 412 bp.
3. dG3–dG4—this degenerate primer set is expected to amplify nearly complete units of 495 bp. The primers are localized in the spacer region 3' from the transcribed region of two adjacent units.

These PCR-amplified fragments were sequenced giving a total of 38 sequences for the different 5S rDNA arrays; 4 for chromosome 3, 4 for chromosome 4, 9 for chromosome 5 block 1, 11 for chromosome 5 block 2, and 10 for chromosome 5 block 3.

The first conclusion is that the same types of 5S units were obtained with the different sets of primers confirming the absence of selective amplification. Monomers within a group are more closely related to each other than to monomers of other groups. Alignment of sequences from each 5S rDNA block resulted in a consensus sequence presented in Figure 4A. It appears that the consensus sequence of Campell et al. (1992) is much more similar to 5S units from chromosome 5 block 1 (6A1, 4E4) and was possibly derived from this array. However, the Campell sequence differs from all our sequences by a 5-bp deletion located in position 477–483. Therefore, the lengths of the *Arabidopsis* major 5S units vary between 498 bp (chromo-

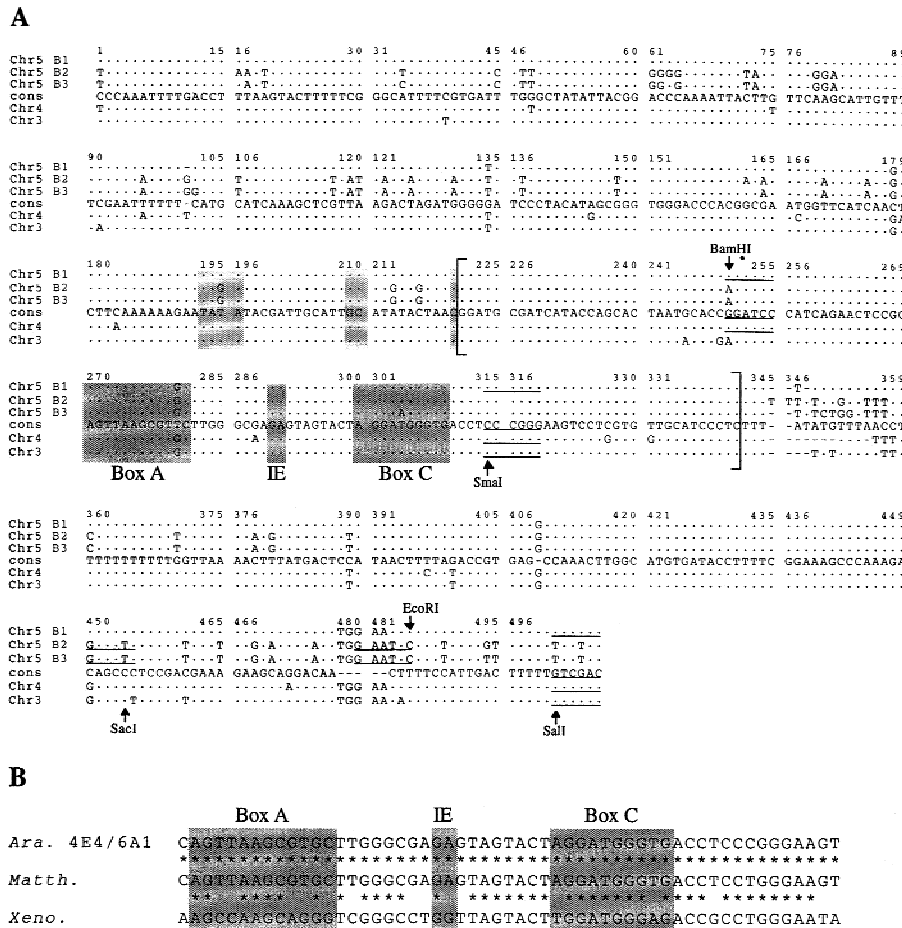
some 4) and 506 bp (chromosome 5, block 2). We have compared our 5S sequences with other published 5S rDNA sequences. The comparison reveals three conserved sequence elements 5' to the transcribed sequence and a higher homology in the region of the intragenic promoter (Fig. 4B) was found with plants like rice, *Vigna radiata*, tomato, sugar beet (*Beta vulgaris*), and the Brassicaceae *Brassica napus* and *Matthiola incana* than with the well-characterized *Xenopus* model (Bogenhagen et al. 1980; Sakonju et al. 1980; Pieler et al. 1987).

#### Intra-5S Block Sequence Variation

Sequence heterogeneity inside each 5S array was calculated. Nucleotide diversity ( $N$ ) is numerically equivalent to the mean number of nucleotide differences per site between all pairs of sequences. Table 2 shows nucleotide diversity expressed as a percentage of all sites, within 5S arrays for 5S transcribed sequence, spacers, and entire repeats. There is an overall equivalence of diversity in the transcribed and spacer regions within a 5S array, but there is a considerable variance from one 5S block to another. 5S sequences are more homogenous in two blocks: on chromosome 4 ( $N=1\%$ ) and in block 1 of chromosome 5 ( $N=1.9\%$ ).

#### Inter-5S Blocks Sequence Variation

Nucleotide diversity is the percentage of the sites showing nucleotide difference between two consensus sequences. Alignment of the consensus sequence from each 5S block illustrates the specific differences for each block resulting occasionally in specific restriction sites (Fig. 4A). Numerous mutations are shared by the consensus sequence from blocks 2 and 3 of chromosome 5 showing that sequences from chromosome 5 blocks 2 and 3 are the most homologous, although



**Figure 4** Alignment of the consensus sequences. (A) Alignment of the consensus sequences derived from the different 5S arrays and the consensus sequence of Campell et al. (1992). Restriction sites are indicated by arrows; conserved sequence elements upstream from the transcribed region (193–196, 209–210) and in the transcribed region BoxA, IE (Intermediate Element), and BoxC are indicated. The length of each consensus sequence is 501 bp, chromosome 3; 498 bp, chromosome 4; 502, 506, and 504 bp for chromosome 5 blocks 1, 2, and 3, respectively. The transcribed region is between brackets. (The *SmaI* site is not present in blocks 2 and 3, but contained in the primer sequence.) Dots (.) indicate identical residue; dashes (-) indicate a gap. (B) Homology between the internal promoter of *Arabidopsis* 5S rDNA (*Ara* 4E4/6A1) and the respective region of *Matthiola incana* (*Matth*) and *Xenopus* 5S rDNA (*Xeno*).

some differences are specific to one of the two blocks (16, 34, 63, 102, 162, etc.). For spacer regions, the diversity between 5S arrays is often greater than intra-5S array diversity (Tables 2 and 3). In fact, fixed specific mutations present in the spacer do not contribute to inter-repeat polymorphism within a 5S block. In contrast, inter-5S block nucleotide diversity in the transcribed region is usually lower than intra-block diversity. Rare fixed mutations were observed in the transcribed region. These mutations are too rare to appear in the consensus sequence but do create some intra-block diversity.

Distance analyses were conducted using all entire monomer sequences of the five 5S arrays. They revealed the existence of five monomer subtypes, although sequences from chromosome 5 blocks 2 and 3

are more homologous leading to two less divergent groups (not shown).

**PCR-Based Isolation of Members of the 5S rDNA Short Repeat Multigene Family**

The primer set dG3–dG4 is expected to amplify almost complete units of 500 bp. These degenerate primers are located in the spacer region, 3' from the transcribed region of two adjacent 5S units (Fig 3A). With YAC 11A3, this primer set indeed produced a PCR product of the expected size as well as an additional unexpected PCR product of 251 bp.

Analysis of this 251-bp variant revealed that it starts at the positions expected with primers dG3–dG4, but lacks a 251-bp region situated between nucleotides 18 and 268 of the consensus sequence. This deletion covers part of the spacer and part of the transcribed region. This deletion explains the absence of amplification of this 251-bp 5S variant with the primer set OL10–OL11, as primer OL11 is situated within the deletion (Fig. 3B). The short variants are interspersed with the 0.5-kb variants on chromosome 3 because PCR products corresponding to two adjacent units of 251 bp and 501 bp were amplified several times with primers OL10 and OL11, with a

proportion of about three 500-bp PCR products for one 750-bp PCR product containing two consecutive units of 251 bp and 501 bp.

**The Restriction Polymorphism Is 5S-Block Specific**

Some restriction polymorphism was observed with several restriction enzymes in pulsed field electrophoresis experiments and confirmed by DNA sequence analysis. This restriction polymorphism was further studied with 14 enzymes (*Bgl*II, *Bam*HI, *Dra*I, *Eco*RI, *Eco*RV, *Hind*III, *Nco*I, *Pst*I, *Sac*I, *Sac*II, *Sal*I, *Sma*I, *Xba*I, *Xho*I) on standard electrophoresis, to perform a more thorough analysis for the different 5S blocks of the genome. Only *Bam*HI and *Sma*I sites are present in the 5S rDNA published sequence (Campell et al. 1992). However, only single point mutations are needed, relative to this pub-

**Table 2. Intra-5S Block Nucleotide Diversity Values (N) for Transcribed Region, Spacer, and Entire Unit Sequences**

5S Block	Transcribed sequence (N)	Spacer (N)	Entire unit (N)
Chromosome 3	6.40%	5.60%	5.70%
Chromosome 4	1.00%	0.96%	1.00%
Chromosome 5 block 1	1.70%	1.96%	1.90%
Chromosome 5 block 2	6.70%	6.30%	6.40%
Chromosome 5 block 3	5.40%	5.98%	5.80%

lished sequence, to generate the sites for the other restriction enzymes used in this study. Restriction polymorphism was studied for all the YACs of the chromosome 5 contig (Fig. 1), for YAC 11A3 mapped on chromosome 3 (Camilleri et al. 1998), for YACs 5C3, 7H3, 8B12, 9D3, mapped on chromosome 4 (Schmidt et al. 1995), and for 5S YACs not yet anchored on the physical map of *Arabidopsis thaliana*: 2A6, 2B6, 2D4, 2E11, 2H7, 3D6, 4B12, 4D12, 4H7, 7G3, 8C4, 9B1, 10G10, 11G4, and 12F1.

Three types of profiles were obtained:

1. Following digestion with an enzyme that does not cut in a 5S rDNA block, all of the 5S rDNA hybridizing with the probe migrated at the position of uncut DNA at the top of the gel (Fig. 5, 9A5 *EcoRV*).
2. Digestion with an enzyme that cuts in almost all of the 5S rDNA units of a block yielded a strong band at 0.5 kb corresponding to a monomeric unit, as well as a band of lower intensity corresponding to dimers of the repeat (Fig. 5, 4E4 *Sall*).
3. With some enzymes, a ladder was obtained corresponding to monomers and or multiples (dimer, trimer, tetramer, etc.) of the 0.5-kb basic repeat. The ladder can be irregular depending on the number of adjacent units that lack the restriction site. Some uncut DNA can be present (Fig. 5, 7E7 *NcoI*) at the top of the gel. The results are summarized in Table 4.

#### 5S Blocks 1, 2, 3 on Chromosome 5

5S units from 6A1, 4E4, and 2G12 (block 1, chromosome 5) are digested by *BamHI* and *Sall* (Fig. 5, 4E4 *Sall*) almost exclusively into unit length 0.5-kb fragments (a faint 1-kb band is observed), indicating that *BamHI* and *Sall*-bearing units comprise nearly all of this 5S block 1. A large proportion of 5S units contain the *SmaI* site as judged by Southern blot. The other enzymes did not cut the 5S units of these YACs.

The situation is more heterogenous for the second 5S block of chromosome 5, with the YACs 7E5-E7 and

9G9. A nearly total digestion was only obtained for *SacI* (only 0.5-kb and 1-kb bands were observed) and most of the units were digested by *EcoRI*. The enzymes *BamHI*, *DraI*, *NcoI* (Fig. 5, 7E7 *NcoI*), *PstI*, and *XhoI* digest a lower proportion of the 5S units. The situation is comparable for the third 5S block on chromosome 5, with strong digestion by *SacI* and *EcoRI*: Only bands between 0.5 kb and 2 kb were obtained, indicating the presence of no more than three adjacent 5S units without the restriction site. Some units were digested with *BamHI*, *BglII*, *DraI*, *PstI*, and *XbaI*. YAC 8D11, which contains the 5S blocks 2 and 3, presents an intermediate restriction polymorphism with some enzymes. With *XbaI*, the bands obtained are similar to those obtained with 9A5 and 9A12 (block 3) in addition to uncut DNA (like for 7E7 block 2). *BamHI* cut twice in a small proportion of the 5S units of chromosome 5 (blocks 1, 2, 3), giving bands <0.5 kb.

#### 5S Block on Chromosome 4

In the YACs 5C3, 7H3, 8B12, and 9D3, localized on chromosome 4, all of the 5S units were digested with *BamHI* and *SmaI*, except an 8-kb band presumably consisting of 16 units with *SmaI*-bearing repeats only at the ends. Most of the units were digested with *Sall* and the 11 other enzymes did not cut the 5S units of these YACs.

#### 5S Block on Chromosome 3

YAC 11A3 from chromosome 3 was found to be more heterogenous. As shown by sequence analysis, 5S units of 0.25 and 0.5 kb are present in this 5S array. Using *SmaI*, two major bands were obtained at 0.5 and 0.75 kb as well as a minor band at 1 kb (Fig. 5, 11A3 *SmaI*). Sequence analysis showed that the 251-bp units frequently have the *SmaI* site (4/4) and PCR products containing one small unit followed by a large unit were obtained. These results lead us to propose the hypothesis that the block contains both tandem array(s) of the 0.5 kb unit and tandem array(s) of a higher-order repeat that consists of a 0.5-kb linked to a 0.25-kb unit. A large proportion of both units contain the *Sall* site as only bands between 0.25 and 1.25 kb are observed with *Sall*.

#### Unmapped 5S YACs

Some YACs not yet mapped in the genome present the same polymorphism as one of these blocks and probably map at the same 5S locus (Table 4). YACs 2B6, 2D4, 3D6, 9B1, and 10G10 are digested by *BamHI*, *Sall*, and *SmaI* and could map either on chromosome 4 or on chromosome 5 block 1. YACs 2A6, 7G3, and 8C4 gave restriction results different from all above YACs and could be candidates for the small 5S block on the right arm of chromosome 5.

All of these results agree with the organization of the YACs on the contig, as overlapping YACs present

**Table 3.** Inter-5S Block Nucleotide Diversity for Spacer and Transcribed Region

		Chromosome 5				
		Chromosome 3 (%)	Chromosome 4 (%)	block 1 (%)	block 2 (%)	block 3 (%)
Chromosome 5	Chromosome 3		5.50	3.40	16.50	16.20
	Chromosome 4	5.70		5.80	17	16.20
	block 1	2.50	3.30		16	15.40
	block 2	2.50	4.10	0.82		2.60
	block 3	2.50	4.90	1.60	0.82	

Nucleotide diversity is the percentage of the sites showing nucleotide difference between two consensus sequences. Spacer is above the diagonal; transcribed region is below the diagonal.

the same restriction polymorphism, and with the sequence results (Table 4). This restriction analysis confirms the presence of numerous 5S rDNA variants at both length (units of 0.25 and 0.5 kb) and sequence level. This polymorphism is both inter- and intra-block.

The results obtained exclude the presence of non-

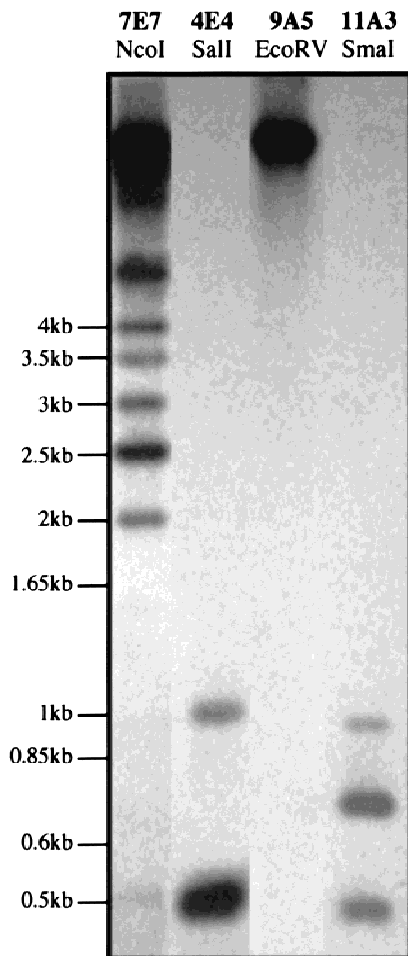
rDNA sequences interspersed with rRNA units inside the 5S arrays. First, non-rDNA interspersed with the 5S RNA units should cause larger than unit-length fragments to be generated by single-cutter enzymes. Furthermore, enzymes that do not cut within individual 5S RNA units apparently did not cut within a 5S block at all. Non-rDNA interspersed among the rDNA units might be expected to harbor sites for one or more of these enzymes. We could occasionally observe one or two faint bands whose lengths were not multiples of 0.5 kb. These minor fragments could represent junction fragments containing part of a 5S rDNA unit and non-5S flanking sequence. Our finding that some enzymes cleave the 5S rDNA to unit length strongly suggests that most of the rDNA units are arranged in simple tandem arrays.

**DISCUSSION**

**YAC Contig Establishment: Three 5S rDNA Arrays in the Contig**

Starting from YACs 6A1 and 4E4, carrying a 5S block and anchored in the pericentromeric region of chromosome 5 left arm, we have aligned 11 YACs for a total of 1300 kb and defined three 5S rDNA blocks. We have now covered 2 Mb in this pericentromeric region of chromosome 5 (Tutois et al. 1999; this paper). Our data confirm the idea that the heterochromatic region occupies a larger area than the ~200 kb proposed by Thompson et al. (1996). On chromosome 5, the genetically defined functional centromere is already 1070 kb long (Copenhaver et al. 1999).

Numerous probes were necessary to generate this contig as these YACs contain, probably exclusively, repeated sequences. Between the three 5S rDNA arrays we mapped DNA sequences such as the retroelements *Athila* (Pélissier et al. 1995); Romani At2 and 3; the three small satellite families 8, 12 (Tutois et al. 1999), and 4E4L; the 163A and 278A centromeric repeats (Thompson et al. 1996); and several other new dispersed-repeat DNA sequences. The 17A20 repeat sequences, not yet characterized, but used



**Figure 5** Southern blot of yeast recombinant DNA, probed with 5S rDNA.

**Table 4.** Comparison of Restriction Polymorphism Results

5S block	Size and unit no.	Nucleotide diversity (%)	Sequence analysis from consensus sequence	Restriction polymorphism		
				frequent digestion	Rare digestion	Unmapped YACs
Chromosome 3L 11A3	N.D.	5.70	<i>SmaI, Sall</i>	<i>SmaI &gt; Sall</i>		2H7,11G4
Chromosome 4 L 5C3, 7H3, 8B12, 9D3	200 kb (400 units)	1	<i>BamHI, Sall, SmaI</i>	<i>BamHI, SmaI &gt; Sall</i>	] <i>BamHI, DraI, NcoI,</i> <i>PstI, XhoI</i> <i>BamHI, BglII, DraI,</i> <i>PstI, XbaI</i>	2B6, 2D4, 3D6, 9B1, 10G10
Chromosome 5 L block 1	100 kb	1.90	<i>BamHI, Sall, SmaI</i>	<i>BamHI, Sall &gt; SmaI</i>		
2G12, 4E4, 6A1	(200 units)					
block 2	30 kb	6.40	<i>SacI, EcoRI</i>	<i>SacI &gt; EcoRI</i>		4D12, 4H7, 2E11
7E5-E7, 9G9	(60 units)	5.80	<i>SacI, EcoRI</i>	<i>SacI = EcoRI</i>		4B12, 12F1
block 3	150 kb					
3G8, 9A5, 9A12	(300 units)					

in FISH experiments by Fransz et al. (1998) are probably present in this contig and could be one of the sequences identified in this paper. FISH experiments with 17A20 and 5S rDNA revealed local regions on the DNA with alternating hybridization spots of 5S rDNA and 17A20 suggesting that these repeats form a partly interspersed pattern. In chromosome 5, the 5S rDNA blocks in both left and right arms would be at the euchromatin–heterochromatin borders (Fransz et al. 1998).

In all chromosomes the centromere accommodates the 180-bp tandem repeat and dispersed repeats such as 106B. According to Fransz et al. (2000), all pericentromeric regions also contain the same repeats. This is in agreement with the sequence results of Lin et al. (1999) and Mayer et al. (1999) who reported that chromosomes 2 and 4 share both similar types and distributions of repeats in the heterochromatic region. As seen above, on chromosome 5 left arm we have established a YAC contig carrying a small pAL1 block (30 kb) next to three 5S rDNA arrays adjacent to the centromere (Fig. 1). This situation is comparable to that observed in chromosome 4 where the YAC contig contains two pAL1 blocks, each 50 kb long, and one 5S array, 200 kb long (Mayer et al. 1999), adjacent to the centromere. The difference resides in the number of the 5S arrays.

In *Arabidopsis*, all of the 5S rDNA loci are separated from the 18S 5.8S 25S rDNA gene clusters. In other plants, the 5S rDNA is clustered at one (tomato, Lapitan et al. 1991; rice, Kamisugi et al. 1994; sugar beet, Schmidt et al. 1994; maize, Mascia et al. 1981) or several sites (pea, Ellis et al. 1988; barley, Leitch and Heslop-Harrison 1993; wheat, Appels et al. 1980; rye, Reddy and Appels 1988; Dvorak et al. 1989; flax, Schneeberger et al. 1989), all of which are separated from the other rRNA genes. A telomeric–subtelomeric (pea, barley, wheat, rye, maize; see references above)

or pericentromeric location (tomato, rice; see references above) for 5S rDNA genes is usually observed, although some 5S rDNA arrays have an interstitial position (one 5S array in pea and 5S arrays from *Helianthus annuus* and *Helianthus argophyllus*, Cuellar et al. 1999).

#### Two Length 5S rDNA Variants Were Identified

In *Arabidopsis*, the major variant is 0.5 kb long and we have been able to amplify a second 5S rDNA variant of 251 bp interspersed with the 0.5-kb variant, from a YAC mapped on chromosome 3. The short variant results from a deletion in the 0.5-kb variant and is probably functionally inactive because a deletion in the transcribed region is observed. 5S rDNA length variants have been reported for several plant species such as sugar beet (Schmidt et al. 1994), sea barley (Baum and Johnson 1998), wheat (Campenhout et al. 1998), pea (Ellis et al. 1988), and flax (Schneeberger et al. 1989; Agarwal et al. 1992) where different length classes of the 5S rDNA units have been identified. The different repeats are organized in separate clusters with each cluster present on a separate chromosome in wheat and rye, although they are interspersed with each other in pea. Because no band at 0.25 kb was obtained using *SmaI*, although the 251-bp units frequently have the *SmaI* site, tandem arrays of 251-bp variants are probably absent. One hypothesis to explain this would be the presence of a higher-order repeat unit composed of a 0.5-kb unit linked to a 251-bp variant. Such higher-order repeat units have been reported for the human  $\alpha$  satellite DNA where many chromosomes are characterized by specific subsets of  $\alpha$  satellite (Willard and Wayne 1987).

#### The Sequences Group According to their 5S Array Origin

DNA sequence heterogeneity present throughout the



amplified sequences enabled their classification in five groups, although sequences from blocks 2 and 3 of chromosome 5 are very similar. Sequence heterogeneity inside each 5S rDNA block differs from one block to another but is equivalent in the transcribed and the spacer regions. The more homogeneous 5S blocks are those on chromosome 4 and block 1 of chromosome 5. Despite similar levels of transcribed-region and spacer-intrablock diversity, spacer sequences are more divergent than transcribed sequences from one block to another. Fixed specific mutations could explain this situation. Numerous specific fixed mutations are observed in the spacer sequences and constitute differences between the blocks, but not within a block. Fixed differences in the transcribed region are nearly nonexistent. Consequently, diversity is equivalent for transcribed region and spacer inside a block, but larger between spacers of different blocks. The overall sequence divergence of *Arabidopsis* 5S rDNA (up to 16%) is higher than that found for the 180-bp Atcon satellite (5%) (Heslop-Harrison et al. 1999). Specific mutations, leading sometimes to new specific restriction sites, are found in each 5S block. The presence of 5S array-specific variants was confirmed with the restriction polymorphism study, as we observed a restriction polymorphism between each major 5S rDNA array of chromosomes 3, 4, and 5. In *Arabidopsis*, Heslop-Harrison et al. (1999) have amplified different variants of Atcon (the 180-bp satellite) and could correlate sequence variations with chromosome specificity. A chromosome specificity has also been observed in the soybean STR 120 satellite DNA (Morgante et al. 1997) and the human centromeric  $\alpha$  satellite DNA (Willard 1985). Using the 5S array restriction polymorphism specificity, nonanchored 5S YACs were predicted to colocalize with one of the 5S rDNA blocks (Table 4).

A comparison of the upstream sequences in our 5S sequences as well as from the other published 5S rDNA sequences (Venkateswarlu et al. 1991) reveals three conserved sequence elements (Fig. 4A): 1) a universally conserved C nucleotide one base before the initiation site that very likely plays a role in the initiation of transcription; 2) a GC dinucleotide at -11, which can be a GC-rich element in other plants, 3) a TATA sequence is observed from -28 to -25. Both 1) and 2) were observed in all our sequences, however the TATA sequence is less conserved in sequences from blocks 2 and 3 from chromosome 5. Comparing the *Arabidopsis* 5S rDNA transcribed region with other transcribed regions, a higher homology was found with the plants than with the well characterized *Xenopus* model (Bogenhagen et al. 1980; Sakonju et al. 1980; Pieler et al. 1987), in the region of the intragenic promoter defined by three elements named box A, Intermediate Element (IE), and box C (Fig. 4B). The *Arabidopsis* 5S rDNA transcribed region starts 5' with GGA and ends with CTC,

differing from all *Fabaceae* investigated, which start with AGG and from plants like flax, sugar beet, and *Brassicaceae* starting with GGG (Hemleben and Werts 1988). T-stretches occur in the spacer shortly after the 3' end of the transcribed region, and could be involved in correct termination of transcription. *Arabidopsis* spacer regions do not contain small stretches of tandemly duplicated sequences.

### Concerted Evolution

Our results show that *Arabidopsis thaliana* 5S rDNA units evolve together as a group. This phenomenon, known as concerted evolution, has been documented for a number of tandemly repeated gene families. A combination of two mechanisms could be involved in this concerted evolution. Unequal crossing-over, with its potential for allowing exchanges between rDNA gene clusters on the same or on different chromosomes, coupled with some mechanism for intrachromosomal homogenization, particularly gene conversion, has been proposed (Dover 1982; Dover et al. 1982; Williams et al. 1989), and there is strong evidence for gene conversion in 5S and tRNA genes (Klein and Petes 1981; Morzycka-Wroblewska et al. 1985). As expected, we found that rDNA units within a cluster are more similar to one another than they are to the rDNA units on another cluster or chromosome. Interlocus interactions have maintained a moderate interlocus heterogeneity both in the spacer and in the transcribed region. In human, ~90 5S rDNA units, tandemly repeated within a 200-kb fragment, are inferred to be nearly identical. The presence of highly repetitive elements in the spacer affect the conformation of the DNA and could promote recombination, helping to homogenize the repeating units (Little and Braaten 1989). Recombination frequencies could vary between our three 5S blocks adjacent to the centromere, because of different DNA modifications or chromatin states, leading to a variable homogenization of the repeats. We report the absence of fixed mutations in the transcribed region. According to Cronn et al. (1996), as variant nucleotides in the transcribed region become disadvantageous owing to their effect on 5S transcription or 5S RNA function, fixed differences fail to accumulate. In this respect, our results are consistent with those observed previously in 5S rDNA from diploid *Triticaceae* (Kellog and Appels 1995) and ITS sequences from *Drosophila* (Schlotterer and Tautz 1994).

Although 5S sequences from blocks 2 and 3 of chromosome 5 are very similar, they are distinct. Specific nucleotides are present for each block and sequences separating the two blocks are clearly not 5S sequences. A single block could have been interrupted by an insertion of non-5S sequences leading to the two blocks 2 and 3 of chromosome 5. A recent insertion

could explain the small divergence between these two blocks.

## METHODS

The CIC YAC library (12 microtiter plates) (Creusot et al. 1995) was used and propagated in yeast strain AB 1380 (Burke et al. 1987) grown in selective medium (yeast nitrogen base 0.75%, glucose 2%, casamino acids 0.75%, adenine 15 mM).

### Preparation of Yeast Chromosomal DNA in Agarose Microplugs and PFGE Separation

Selective medium (10 ml) was inoculated with a single yeast colony and cultured at 30°C for 18–24 hr. Yeast cells were harvested by centrifugation and embedded in 1 volume of 1.2% Sea Plaque GTG agarose (FMC Inc). Plugs were then incubated in 5 ml of SCE medium (1 M Sorbitol, 100 mM sodium citrate; 60 mM EDTA) containing 30 mM DTT and lyticase (0.4 mg/ml) at 37°C for 2 hr. Plugs were further incubated in a solution (5 ml) containing 1% lauryl sarkosyl, 0.5 M EDTA, 20 mM Tris (pH 7.5), 1 mg/ml proteinase K at 50°C for 16 hr. Plugs were stored in 5 ml of 5 mM EDTA. Yeast chromosomes in agarose microplugs were separated by PFGE on a CHEF-DRII apparatus (BioRad). Plugs were placed in a trough prepared in a 1% low-melting-point agarose gel (BRL) in 0.5X TBE. The gel was subjected to electrophoresis at 200 V using a 20 or 50 sec switch time for 15 hr at 14°C. The YAC sizes were estimated by comparison with the chromosome sizes of the yeast strain AB 1380.

### Preparation of Yeast Transformant DNA and Preparation of YAC End Probes

Selective medium (20 ml) was inoculated with a single yeast colony and cultured at 30°C for 36 hr. Yeast cells were harvested by centrifugation (10 min, 2000 rpm) and resuspended in 1 ml of TES (100 mM Tris HCl at pH 8.0, 100 mM EDTA, 0.9 M Sorbitol, 0.2%  $\beta$ -mercaptoethanol, 2 mg/ml lyticase) and incubated for 1 hr at 37°C. Spheroplasts were pelleted by centrifugation (5 min, 1200 rpm) and resuspended in 2 ml of 50 mM Tris, 20 mM EDTA, then lysed by addition of 100  $\mu$ l 20% SDS and incubation for 30 min at 60°C. Potassium acetate 5 M (600  $\mu$ l) was added. After centrifugation (1200 rpm, 15 min), the supernatant was recovered and 5 ml of 100% ethanol added. This was centrifuged and the pellet was resuspended in 500  $\mu$ l of TE containing Ribonuclease A (0.01 mg/ml). The preparation was incubated for 30 min at 37°C, and the DNA was precipitated and resuspended in 100  $\mu$ l of TE.

YAC end probes from CIC clones were prepared using vectorette-PCR as described by Matallana et al. (1992).

### Determination of Copy Number

Copy number of each subclone in the *Arabidopsis* genome was determined by slot blot hybridization. Amounts of 50, 150, 250, 500, and 1000 ng of total *Arabidopsis* genomic DNA were blotted. Dilutions of each fragment were performed and 50, 150, 250, and 500 pg was blotted. Both membranes (genomic DNA and fragment dilutions) were hybridized with one of the following probes: 2H7L, 4E4L, 4H7R, 7E5L, 7E7L, 8D11L, 9A12L and R, 9G9L, and 2D2 107 as a six-copy control (Tutois et al. 1999). After hybridization, slots were counted and copy number of each fragment in the genome was determined using both countings obtained with the fragments' dilutions,

and countings obtained by hybridization of 2D2 107 (6 copies) on genome dilutions.

### Southern Blots, Genetic Mapping, and Probe Labeling

*Arabidopsis thaliana* genomic DNA (ecotype Columbia) was isolated from 3- to 4-week-old plants with the CTAB method (Doyle and Doyle 1987). The digestions were realized with 4  $\mu$ g of *Arabidopsis thaliana* genomic DNA and 15–20 units of restriction enzyme in the recommended buffer (New England Biolabs) supplemented with 1 mM spermidine. Digested DNA was electrophoresed in 0.8% agarose gels overnight, depurinated in 0.25 M HCl, and transferred to hybrid membranes (Amersham). (A total of 15  $\mu$ l of yeast transformant DNA was used per digestion in the same conditions as genomic DNA.)

DNA probes were prepared from plasmid fragments excised from gels. They were labeled with  $\alpha$ -<sup>32</sup>P using a random hexamer priming method (Amersham).

### PCR Amplification and Primers

The primer sets used to isolate 5S rDNA units were derived from the 5S consensus sequence of Campell et al. (1992). The primer sequences are as follows: OL10: CCTCGTGTGCATC-CCTC; OL11: GCT-GGTATGATCGCATCC; OL9: CTCCC-GGGAGGTCACCC; OL6: CCTAGTACTACTCTCGC-CC; dG10: CCTNGTGNTGNANCCCTC; dG9: CTNCCNG-GNAGNTCACCC; dG3: GNCAAANTTNGCAT-GTG; dG4: CNGTNTANAAGTTATNGAGTC.

5S rDNA units were amplified by PCR using YACs as template DNA. The positions of the primers in the 5S rDNA repeating units are indicated in Figure 3A. The amplification reactions were carried out in 25  $\mu$ l volume. After initial denaturation at 95°C for 5 min, 30 amplification cycles were performed, each consisting of denaturation at 95°C for 30 sec, annealing at  $T_m$  for 30 sec, and primer extension at 72°C for 30 sec. PCR products were purified with the PCR purification kit (Qiagen).

### Subcloning and Sequencing

PCR products were subcloned in the pGem-T plasmid using the pGem-T vector system I (Promega). Sequencing was performed using the dideoxynucleotide chain termination method (Sanger et al. 1977), using the T7 DNA polymerase (Pharmacia). Computer sequence analysis was performed with the Blast Search program (Altschul et al. 1996).

### Stability and Chimerism of 5S YACs

5S CIC YACs are much more stable than pAL1 and 18S–25S YACs, as observed by J. Lafleuriel (pers. comm.) during sizing and characterization of the CIC library. The mean insert size is around 340 kb for 5S YACs compared with 240 and 160 kb for pAL1 and 18S–25S YACs, respectively (Creusot et al. 1995). Using PFGE we have shown that all our overlapping 5S YACs do contain a 5S block with the same size. In the CIC library the proportion of chimeric clones is <10% (Creusot et al. 1995). During the course of the experiment, many YAC fragments were used as probes without any conflicting results. Therefore we can consider that these YACs are not chimeric.

### ACKNOWLEDGMENTS

We thank C. White and J.C. Côté for critical reading of the manuscript. This work was supported by the CNRS and by the Université Blaise Pascal. C. Cloix and O. Mathieu are recipi-

ents of a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche.

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.

## REFERENCES

- Agarwal, M.L., J. Aldrich, A. Agarwal, and C.A. Cullis. 1992. The flax ribosomal RNA-encoding genes are arranged in tandem at a single locus interspersed by "non-rDNA" sequences. *Gene* **120**: 151-156.
- Altschul, S.F. and W. Gish. 1990. Local alignment statistics. *Methods Enzymol.* **266**: 460-480.
- Appels, R., W.L. Gerlach, E.S. Dennis, H. Swift, and W.J. Peacock. 1980. Molecular and chromosomal organization of DNA sequences coding for the ribosomal RNAs in cereals. *Chromosoma* **78**: 293-311.
- Baum, B.R. and D.A. Johnson. 1998. The 5S rRNA gene in sea barley (*Hordeum marinum hudson sensu lato*): Sequence variation among repeat units and relationship to the X haplome in barley (*hordeum*). *Genome* **41**: 652-661.
- Bogenhagen, D.F., S. Sakonju, and D.D. Brown. 1980. A control region in the center of the 5S RNA gene directs specific initiation of transcription: II. The 3' border of the region. *Cell* **19**: 27-35.
- Burke, D.T., G.F. Carle, and M.V. Olson. 1987. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**: 806-812.
- Camilleri, C., J. Lafleurriel, C. Macadre, F. Varoquaux, Y. Parmentier, G. Picard, M. Caboche, and D. Bouchez. 1998. A YAC physical map of *Arabidopsis thaliana* chromosome 3. *Plant J.* **14**: 633-642.
- Campbell, B.R., Y. Soung, T.E. Posch, C.A. Cullis, and C.D. Town. 1992. Sequence and organization of 5S ribosomal RNA-encoding genes of *Arabidopsis thaliana*. *Gene* **112**: 225-228.
- Campanhout, S.V., R. Aert, and G. Volckaert. 1998. Orthologous sequence variation among 5S ribosomal RNA gene spacer sequences on homoeologous chromosomes 1B, 1D and 1R of wheat and rye. *Genome* **41**: 244-255.
- Copenhaver, G.P., K. Nickel, T. Kuromori, M.I. Benito, S. Kaul, X. Lin, M. Bevan, G. Murphy, B. Harris, L.D. Parnell, et al. 1999. Genetic definition and sequence analysis of *Arabidopsis* centromeres. *Science* **286**: 2468-2474.
- Creusot, F., E. Fouilloux, M. Dron, J. Lafleurriel, G. Picard, A. Billault, E. Le Paslier, D. Cohen, M.E. Chabouté, A. Durr, J. Fleck, et al. 1995. The CIC library: A large insert YAC library for genome mapping in *Arabidopsis thaliana*. *Plant J.* **8**: 763-770.
- Cronn, R.C., X. Zhao, A.H. Paterson, and J.F. Wendel. 1996. Polymorphism and concerted evolution in a tandemly repeated gene family: 5S ribosomal DNA in diploid and allopolyploid cottons. *J. Mol. Evol.* **42**: 685-705.
- Cuellar, T., J. Orellana, E. Belhassen, and J.L. Bella. 1999. Chromosomal characterization and physical mapping of the 5S and the 18S-5.8S-25S ribosomal DNA in *Helianthus annuus*. *Genome* **42**: 110-115.
- Dover, G.A. 1982. Molecular drive, a cohesive model of species evolution. *Nature* **299**: 111-117.
- Dover, G., S. Brown, E. Cohen J., Dallas, T. Strachan, and M. Trick. 1982. The dynamics of genome evolution and species differentiation. In *Genome evolution* (ed. G.A. Dover and R.B. Flavell), pp. 343-372. Academic Press, London, UK.
- Doyle, J.J. and J.L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh tissue. *Phytochem. Bull.* **19**: 11-15.
- Dvorak, J., H.B. Zhang, R.S. Kota, and M. Lassner. 1989. Organization and evolution of the 5S ribosomal RNA gene family in wheat and related species. *Genome* **32**: 1003-1016.
- Ellis, T.N., D. Lee, C.M. Thomas, P.R. Simpson, W.G. Cleary, M.A. Newman, and K.W.G. Burcham. 1988. 5S rRNA genes in *pisum*: Sequence, long range and chromosomal organization. *Mol. Gen. Genet.* **214**: 333-342.
- Fransz, P., S. Armstrong, C. Alonso-Blanco, T.C. Fisher, R.A. Torres-Ruiz, and G. Jones. 1998. Cytogenetics for the model system *Arabidopsis thaliana*. *Plant J.* **13**: 867-876.
- Fransz, P., S. Armstrong, J. Hans De Jong, L.D. Parnell, C. Van Drunen, C. Dean, P. Zabel, T. Bisseling, and G.H. Jones. 2000. Integrated cytogenetic map of the short arm of chromosome 4 of *Arabidopsis thaliana*: Structural organization of a heterochromatic knob and pericentromeric heterochromatin. *Cell* **100**: 367-376.
- Hemleben, V. and D. Werts. 1988. Sequence organization and putative regulatory elements in the 5S rRNA genes of two higher plants (*Vigna radiata* and *Matthiola incana*). *Gene* **62**: 165-169.
- Heslop-Harrison, J.S., M. Murata, Y. Ogura, T. Schwarzacher, and F. Motoyoshi. 1999. Polymorphisms and genomic organization of repetitive DNA from centromeric regions of *Arabidopsis thaliana* chromosomes. *Plant Cell* **11**: 31-42.
- Kamisugi, Y., S. Nakayama, R. Nakajima, H. Ohtsubo, E. Ohtsubo, and K. Fukui. 1994. Physical mapping of the 5S ribosomal RNA genes on rice chromosome 11. *Mol. Gen. Genet.* **245**: 133-138.
- Kellogg, E.A. and R. Appels. 1995. Intraspecific and interspecific variation in 5S RNA genes are decoupled in diploid wheat relatives. *Genetics* **140**: 325-343.
- Klein, H.L. and T.D. Petes. 1981. Intrachromosomal gene conversion in yeast. *Nature* **289**: 144-147.
- Kotani, H., S. Sato, M. Fukami, T. Hosouchi, N. Nakazaki, S. Okumura, T. Wada, Y.G. Liu, D. Shibata, and S. Tabata. 1997. A fine physical map of *Arabidopsis thaliana* chromosome 5: Construction of a sequence-ready contig map. *DNA Research* **4**: 371-378.
- Lapitan, L.N., M.W. Ganai, and S.D. Tanksley. 1991. Organization of the 5S ribosomal RNA genes in the genome of tomato. *Genome* **34**: 509-514.
- Leitch, I.J. and J.S. Heslop-Harrison. 1993. Physical mapping of four sites of 5S rDNA sequences and one site of the  $\alpha$ -amylase-2 gene in barley (*Hordeum vulgare*). *Genome* **36**: 517-522.
- Lin, X., S. Kaul, S. Rounsley, T.P. Shea, M.I. Benito, C.D. Town, C.Y. Fujii, T. Mason, C.L. Bowman, M. Barnstead et al. 1999. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature*. **402**: 761-768.
- Little, R.D. and D.C. Braaten. 1989. Genomic organization of Human 5S rDNA and sequence of one tandem repeat. *Genomics* **4**: 376-383.
- Long, E.O. and I.B. Dawid. 1980. Repeated genes in eukaryotes. *Annu. Rev. Biochem.* **49**: 727-764.
- Mascia, P.N., I. Rubenstein, R.L. Phillips, A.S. Wang, and L. Xiang. 1981. Localization of the 5S rRNA genes and evidence for diversity in the 5S rDNA region of maize. *Gene* **15**: 7-20.
- Matallana, E., C.J. Bell, P.J. Dunn, M. Lu, and J.R. Ecker. 1992. Genetic and physical linkage of the *Arabidopsis* genome: Methods for anchoring yeast artificial chromosomes. In *Methods in Arabidopsis research* (ed. C. Koncz, N.H. Chua, and J. Schell), pp. 144-169. World Scientific, Singapore.
- Mayer, K., C. Schüller, R. Wanbutt, G. Murphy, G. Volckaert, T. Pohl, A. Düsterhöft, W. Stiekema, K.D. Entian, N. Terryn et al. The European Union Arabidopsis Genome Sequencing Consortium, and The Cold Spring Harbor, Washington University in St Louis, and PE Biosystems Arabidopsis Sequencing Consortium. 1999. Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*. *Nature* **402**: 769-777.
- Morgante, M., I. Jurman, L. Shi, T. Zhu, P. Keim, and J.A. Rafalski. 1997. The STR 120 satellite DNA of soybean: Organization, evolution and chromosomal specificity. *Chrom. Res.* **5**: 363-373.
- Morzycka-Wroblewska, E., E.U. Selker, J.N. Stevens, and R.L. Metzenberg. 1985. Concerted evolution of dispersed *Neurospora crassa* 5S RNA genes: Pattern of sequence conservation between allelic and nonallelic genes. *Mol. Cell. Biol.* **5**: 46-51.
- Mozo, T., K. Dewar, P. Dunn, J.R. Ecker, S. Fischer, S. Kloska, H. Lehrach, M. Marra, M. Martienssen, S. Meier-Ewert, and T. Altmann. 1999. A complete BAC-based physical map of the *Arabidopsis thaliana* genome. *Nat. Genet.* **22**: 271-275.
- 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.
- Pélissier, T., S. Tutois, J.M. Deragon, S. Tourmente, S. Genestier, and G. Picard. 1995. *Athila*, a new retroelement from *Arabidopsis thaliana*. *Plant Mol. Biol.* **29**: 441–452.
- Pieler, T., J. Hamm, and R.G. Roeder. 1987. The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. *Cell* **48**: 91–100.
- Reddy, P. and R. Appels. 1988. A second locus for the 5S multigene family in *Secale L*: Sequence divergence in two lineages of the family. *Genome* **32**: 456–467.
- Sakonju, S., D.F. Bogenhagen, and D.D. Brown. 1980. A control region in the center of the 5S RNA gene directs specific initiation of transcription: I. The 5' border of the region. *Cell* **19**: 13–25.
- Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci.* **77**: 5463–5467.
- Schlotterer, C. and D. Tautz. 1994. Chromosomal homogeneity of *Drosophila* ribosomal DNA arrays suggests intrachromosomal exchanges drive concerted evolution. *Curr. Biol.* **4**: 777–783.
- Schmidt, T., T. Schwarzacher, and J.S. Heslop-Harrison. 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., 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.
- Schmidt, R., K. Love, J. West, Z. Lenehan, and C. Dean. 1997. Description of 31 YAC contig spanning the majority of *Arabidopsis thaliana* chromosome 5. *Plant J.* **11**: 563–572.
- Schneeberger, R.G., G.P. Creissen, and C.A. Cullis. 1989. Chromosomal and molecular analysis of 5S RNA gene organization in the flax, *Linum usitatissimum*. *Gene* **83**: 75–84.
- Thompson, H.L., R. Schmidt, and C. Dean. 1996. Identification and distribution of seven classes of middle-repetitive DNA in the *Arabidopsis thaliana* genome. *Nucleic Acids Res.* **24**: 3017–3022.
- Tutois, S., C. Cloix, C. Cuvillier, M.C. Espagnol, J. Lafleurriel, G. Picard, and S. Tourmente. 1999. Structural analysis and physical mapping of a pericentromeric region of chromosome 5 of *Arabidopsis thaliana*. *Chrom. Res.* **6**: 1–14.
- Venkateswarlu, K., S.W. Lee, and R.N. Nazar. 1991. Conserved upstream sequence elements in plant 5S ribosomal RNA-encoding genes. *Gene* **105**: 249–253.
- Willard, H.F. 1985. Chromosome-specific organization of human alpha satellite DNA. *Am. J. Hum. Genet.* **37**: 524–532.
- Willard, H.F. and J. S. Waye. 1987. Hierarchical order in chromosome-specific human alpha satellite DNA. *Trends Genet.* **3**: 192–198.
- Williams, S.M., J.A. Kennison, L.G. Robbins, and C. Strobeck. 1989. Reciprocal recombination and the evolution of the ribosomal gene family of *Drosophila melanogaster*. *Genetics* **122**: 617–624.

Received November 8, 1999 ; accepted in revised form March 8, 2000.