



**HAL**  
open science

## Characterization of minisatellites in *Arabidopsis thaliana* with sequence similarity to the human minisatellite core sequence

S Tourmente, J M Deragon, J Lafleuriel, S Tutois, T Pélissier, C Cuvillier, M C Espagnol, G. Picard

► **To cite this version:**

S Tourmente, J M Deragon, J Lafleuriel, S Tutois, T Pélissier, et al.. Characterization of minisatellites in *Arabidopsis thaliana* with sequence similarity to the human minisatellite core sequence. *Nucleic Acids Research*, 1994, 22 (16), pp.3317-3321. 10.1093/nar/22.16.3317 . hal-01920378

**HAL Id: hal-01920378**

**<https://uca.hal.science/hal-01920378v1>**

Submitted on 13 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Characterization of minisatellites in *Arabidopsis thaliana* with sequence similarity to the human minisatellite core sequence

S.Tourmente\*, J.M.Deragon, J.Lafleuriel, S.Tutois, T.Pélissier, C.Cuvillier, M.C.Espagnol and G.Picard

GDR 977 Biomove CNRS, Université Blaise Pascal, 24 Avenue des Landais, 63177 Aubiere Cedex, France

Received June 16, 1994; Revised and Accepted July 25, 1994

EMBL accession nos X79716–X79718 (incl.)

## ABSTRACT

**A strategy based on random PCR amplification was used to isolate new repetitive elements of *Arabidopsis thaliana*. One of the random PCR product analyzed by this approach contained a tandem repetitive minisatellite sequence composed of 33 bp repeated units. The genomic locus corresponding to this PCR product was isolated by screening a lambda genomic library. New related loci were also isolated from the genomic library by screening with a 14 mer oligonucleotide representing a region conserved among the different repeated units. Alignment of the consensus sequence for each minisatellite locus allowed the definition of an *Arabidopsis thaliana* core sequence that shows strong sequence similarities with the human core sequence and with the generalized recombination signal Chi of *Escherichia coli*. The minisatellites were tested for their ability to detect polymorphism, and their chromosomal position was established.**

## INTRODUCTION

Minisatellites, also called VNTR (Variable number of tandem repeats) are tandemly-repeated DNA sequences. They are ranging in size from a few hundred to several kilobase-pairs while their repeated units vary from 10 to 60 bp. A common feature of tandemly arranged repetitive sequences is their high degree of polymorphism, mostly due to variation in the number of tandem repeats but also to the internal structure of individual repeats (1). Minisatellites were first characterized in man (2,3) where they share a common motif known as the core sequence (4,5). These highly polymorphic sequences have been shown to be of fundamental importance to the study of human genetics (6–8).

Probes based on the human core sequence have been successfully used for DNA fingerprinting of a variety of species including mammals (9–11), birds (12,13), fungi (14,15) and plants (16–18). An internal repeat sequence from bacteriophage M13, which shares strong homologies to the human minisatellite

core sequence, has also been successfully used as a DNA fingerprinting tool in a variety of organisms including plants (19–21). Despite these results, minisatellite loci have not been yet characterized in plants. Broun and Tanksley (18) analysed recently DNA clones from tomato that share sequence similarity to human minisatellites but differ considerably in structure. We report the identification and characterization in *Arabidopsis thaliana* of a minisatellite family that shares strong structural and sequence similarity with human minisatellites. Three members of this family were placed on the *Arabidopsis* genetic map and their ability to detect polymorphisms among ecotypes was tested.

## MATERIAL AND METHODS

### Plant material

*Arabidopsis thaliana* DNA from 4 ecotypes C24, Columbia (Co), Landsberg (L) and Bensheim (B) was used.

### Southern blots

Genomic DNA was isolated from 3 to 4 weeks old plants (individually or in pools) with the CTAB method (22). The digestions were realized with 4  $\mu$ g of *Arabidopsis thaliana* genomic DNA and 15–20 units of the proper restriction enzyme in the recommended buffer (New England Biolabs) supplemented with 1 mM spermidine. For human, mouse, *Drosophila melanogaster*, *Brassica napus*, maize and wheat, 4, 10, 4, 5, 10 and 10  $\mu$ g of DNA were used respectively. Digested DNA were electrophoresed in 1% agarose gels overnight, depurinated in 0.25% HCl and transferred to Hybond membranes (Amersham).

### Probe labelling

DNA probes were PCR amplified inserts or plasmid fragments excised from a gel. They were labelled with  $\alpha^{32}$ P using a random hexamer priming method (Amersham). Oligonucleotides probes were labelled with  $\gamma^{32}$ P using the T4 polynucleotide kinase.

\*To whom correspondence should be addressed

**Genomic library screening, subcloning and sequencing**

A lambda EMBL3 genomic library of *Mbo*I partially digested *Arabidopsis thaliana* DNA was screened with the PCR probe or with the 14 mer oligonucleotide GGAGGTGGAGGCGG. For the DNA probe, hybridizations and washes (in 1×SSC, 0.1% SDS) were at 65°C. For the oligonucleotide probe, hybridization (in 5×SSPE, 0.05 M Tris HCl pH 8, EDTA 0.01 M, Denhardt 10×, SDS 0.5%, skimmed milk 0.5%) was realized at 37°C, and washes were performed as follows: Tetra methyl ammonium chloride 3 M, EDTA 0.002 M, Tris HCl pH 8 0.25 M, SDS 0.2%, 10 min 3 times at room temperature, and 5 min at 43°C (TM -5°C). Selected lambda clones were cut with a variety of restriction enzymes, and for each phage, the smallest fragment containing the whole minisatellite was subcloned into the pBluescript SK plasmid vector (Stratagene). For each phage, larger DNA fragments were also subcloned to recover flanking sequences. Sequencing was performed using the dideoxynucleotide chain termination method (23) using the T7 DNA polymerase (Pharmacia).

**Polymerase chain reaction (PCR)**

Random PCR were performed on C24 genomic DNA using a Techne PHC-3 apparatus. The amplifications were in a total volume of 25 µl containing 25 ng genomic DNA, 1 µM of a random primer (TAGGTGGTAGACTGGTC), 200 µM of each dNTP, 0.5 U of Vent polymerase (New England Biolabs), and the recommended buffer. Samples covered with mineral oil were processed as follows: one cycle: 95°C 1 min, 25°C 10 min; 45 cycles: 50°C 5 sec, 72°C 30 sec, 95°C 30 sec, 25°C 30 sec; 1 cycle: 50°C 5 sec, 72°C 5 min. The transition between 25°C and 50°C was realized at a rate of 1°C each 5 sec (slope: 12).

For standard PCR amplifications on loci 1 and 3, primers were selected in the regions flanking the tandem repeats. PCR amplifications were performed on the genomic DNA of the 4 ecotypes: Co, C24, L, B, and on the plasmid containing the

```

1   CAAAATTATAGAGGTGGCCATGGCGG 26
27  AGGATGGAAGGAAGGCGGTGGCC----AAGGTGG 56
57  AGGATGGAAGGAGGCGGTGGCC----AAGGTGG 85
86  AGGATGGAAGGAGGC-GTGGCC----AAGGTGG 114
115 AGGATGGAAGGAGGC-GTGGC-----AAGGC-- 140
141 AGGATGGAAGGAGGC-GT--C-----AAGGCG- 165
166 AGGATGGCTAGGAGGCGGGGGTGGATATGGCGG 198
199 AGGATGGCATGGAGGTGGAGGAGGACATGGCGG 231
232 AGGATGGACGGGAGGTGGAGGTGGACGAAGGAG 263
264 AGGATGGAAGGAGGTGGTGGACGTGGCGGAGG 296
297 ----ATGGCAGGAGGCGGTGGAC----ATGGCGG 323
324 AGGATGGCAGAGAGGCGGTGGAA----GAGGTGG 353
354 AGGATGGAAGCAGGCGGTGGCC----ATGGCGG 383
384 AGGATGGAGAGGAGGCGGTGGC-----GGAGG 410
411 AGGATGGAGAGGAGGCGGTGGCA-----GCGG 437
438 AGGATAGAG 446
447 TAATCGCCGTCTCTACTCATCT468

```

**Figure 1.** DNA sequence of the minisatellite 1. The minisatellite sequence, in bold, is presented with flanking regions (1–26; 447–468). The dashes (–) represent deletions in the tandem repeats. The DNA sequence of the PCR product is between positions 166 and 446.

minisatellite fragment. The amplifications were realized in a volume of 25 µl containing 100 ng of genomic DNA (or 0.1 ng plasmid DNA), 1 µM of each dNTP, 1 µM of each of the two primers, with the recommended buffer and 0.5 U of Vent polymerase. Samples were covered with mineral oil and processed through 42 temperature cycles: 1 cycle: 51°C 30 sec; 40 cycles: 72°C 1 min, 95°C 30 sec, 51°C 30 sec; 1 cycle, 72°C 5 min. PCR products of these amplifications were run on 2% agarose gels.

**Regeneration**

Independent regenerants were produced from roots as described in (24).

**Genetic mapping**

An F2 mapping population of 96 F8 recombinant inbred lines originating from the C.Dean laboratory and resulting from a cross between Columbia and Landsberg was used for mapping the cloned sequences. Hybridizations were performed with DNA probes in stringent conditions. Map positions were established using the Mapmaker software.

**RESULTS****Amplification of a minisatellite DNA sequence by random PCR**

Using random primers, different PCR products were obtained after amplification at low stringency. As revealed by hybridization (not shown), most of these products are members of the tandemly repeated 180bp family (25), the rDNA (26,27), or represent chloroplastic DNA. One of these fragments did not hybridize with control probes. Its nucleotide sequence, after subcloning in the pBluescript vector revealed a minisatellite DNA sequence (Figure 1). This minisatellite is very GC rich (around 65%), which is a characteristic of previously described minisatellites (2,9). Most of the repeated units are 33bp long, although some have short deletions. The PCR product does not contain the complete minisatellite since flanking regions are absent.

**Genomic library screening with the minisatellite PCR product**

To isolate the complete locus, we have screened a genomic library with the minisatellite PCR product. Among 50,000 plaques (4–5 equivalent genomes), one positive clone was found by high stringency hybridization of the minisatellite PCR product. This phage (phage 1) was purified and analyzed by restriction analysis, and the fragment containing the minisatellite was subcloned and sequenced. It contained a 622bp minisatellite (14 repeated units) and flanking sequences on each side. The central region of the sequence, between positions 166 and 438, corresponds to the PCR product used for the screening. The length of the repeated units varies from 33 to 27 bp, and the left parts of the units are more conserved than the right parts in which deletions can be observed (Figure 1).

**Isolation of other members of this minisatellite family**

When used at high stringency, human minisatellites are single-copy probes (3). However, each minisatellite of a given family share a common short 'core' sequence present in each repeated unit. This core sequence provides a powerful probe for detecting different members of a minisatellite family. Based on the first results of sequence of the minisatellite from phage 1 available at that moment, we have selected a 14 mer oligonucleotide

(GGAGGTGGAGGCGG) representing the most conserved sequence.

With this 'core' oligonucleotide we screened the genomic library to isolate other loci corresponding to the same minisatellite family. Among 50,000 plaques, 40 positive clones were found by hybridization of the oligonucleotide at high stringency. Among the 40 positive clones, 8 were initially selected for analysis. Restriction analysis and cross-hybridization between the phage inserts revealed only 3 different phages (phages 2, 3, 4) hybridizing with the oligonucleotide. The sequence of the fragment hybridizing with the oligonucleotide was determined for these 3 phages (not shown).

For the phage 2, a 182bp *AluI* fragment was sequenced. It includes 4 complete units of 33bp with a consensus sequence sharing 12 out of 14 bp homologies with the oligonucleotide (Table 1). On the phage 3 the minisatellite is 198bp long with 5 complete repeated units, without deletions. The repeated units are 33 bp long, and the central part of the repeat has 12 out of 14bp homology with the oligonucleotide used for the screening (Table 1). In the 994bp fragment of the phage 4 hybridizing with the oligonucleotide, we could surprisingly not identify a minisatellite. It contains however GC rich microsatellite motifs that can account for the cross hybridization with the oligonucleotide. As shown on Table 1, the oligonucleotide used for the screening is GGAGGTGGAGGCGG and the microsatellites motifs hybridizing with this oligonucleotide are multimers of the motif GGX (not shown). This type of microsatellite is generally poorly represented in higher plants (28,29), as this is the case in *Arabidopsis thaliana* (Depeiges *et al.*, in preparation).

#### Definition of a core motif with sequence homology with the human minisatellites

All the repeated units of the minisatellites characterized so far contain a GGAGG motif between positions 11 and 15, and share a core sequence composed of 11GGAGGXGGXGGXGG24. This sequence has a high homology with the human core sequence (2), and particularly with the myoglobin core sequence (3) (Table 1). The minisatellite sequence on phage 3 has the highest homology with the minisatellite core sequence present near the human myoglobin gene (Table 1). The complex pattern found

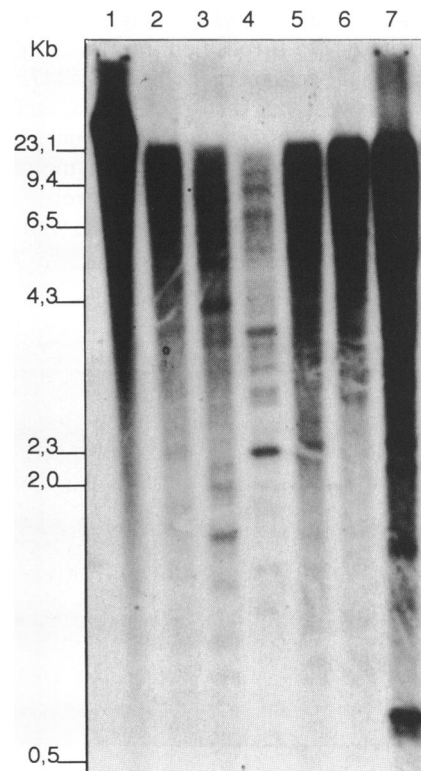
**Table 1.** A common core sequence is shared by the consensus sequence of the 3 minisatellites (ms1, ms2, ms3), and shows a high homology with the consensus sequence of the tandem repeats present at the human myoglobin locus

Oligonucleotide used for the screen	GGAGGTGGAGGCGG		
Ms1	AGGATGGAAA	GGAGGCGGTGGCGG	ACAAGGCGG
Ms2	TGGTGGTGGT	GGAGG <sup>A</sup> GGCGG <sup>C</sup> GG	TGGTGGTGG
		<sub>C</sub>	<sub>T</sub>
Ms3	GGGAGG <sup>G</sup> GGT	GGAGGTGGAGG <sup>A</sup> GA	<sup>A</sup> GGCGG <sup>G</sup> G <sup>G</sup>
	<sub>T</sub>	<sub>T</sub>	<sub>C</sub> <sub>A</sub> <sub>A</sub>
<i>Arabidopsis thaliana</i> . core	GGAGGXGGXGGXGG		
Human myoglobin 2 consensus sequence	GGAGGCTGGAGGAG		
Human core	GGAGGTGGGCAGGAXG		

after hybridization of our oligonucleotide on a transferred DNA from different species (man included) is a reflection of these homologies (Figure 2).

#### Polymorphism at *Arabidopsis thaliana* minisatellite loci

When used as probes against southern blots of *Arabidopsis* genomic DNA, minisatellite sequences detect only one band at high stringency. The polymorphism corresponds to different numbers of tandem repeats. It is revealed by the differences in fragment length from one ecotype to another, after hybridization of the minisatellite probe on genomic DNA of different ecotypes digested with the same enzyme. For minisatellite 1, polymorphism is detected among the 4 ecotypes Co, C24, B, L, with 2 restriction enzymes (Figure 3A). The differences in fragment length from one ecotype to another probably reflect different numbers of repeated units. It could also reflect RFLPs (restriction fragment length polymorphisms). To test these hypotheses, PCR amplifications were performed on genomic DNA with primers, the same 2 for all ecotypes, located in the very near flanking sequences of the locus. The results are presented on figure 3B and show that as expected the PCR products have different lengths in the 4 ecotypes, confirming that polymorphism is due to different numbers of tandem repeats. For minisatellite 2, polymorphisms were detected with 4 bp cutting site enzymes like *HhaI*. However, no polymorphism was detected among the 4 ecotypes with enzymes having a 6 bp cutting site. The fragments generated by these enzymes are probably too large to allow the detection by agarose electrophoresis of the



**Figure 2.** DNA fingerprints of different species probed with the *Arabidopsis thaliana* core oligonucleotide 1: human, 2: mouse, 3: *Drosophila melanogaster*, 4: *Arabidopsis thaliana*, 5: *Brassica napus*, 6: maize, 7: wheat.

variation in the number of repeats. For minisatellite 3, PCR products have the same length in the 4 ecotypes indicating that they have the same number of tandem repeats at this locus. This agrees with our inability to find polymorphism with 4bp or 6bp cutting sites enzymes.

These results show that 2 of the 3 minisatellites are polymorphic among the 4 ecotypes. We have also tested the polymorphism among individual plants within the C24 ecotype. No difference in pattern was observed after an *Eco*RI digestion of the genomic DNA from 15 independent plants was hybridized with the oligonucleotide or with the minisatellites 1 and 2 probes. The same results were obtained after hybridization of the oligonucleotide and the minisatellites 1 and 2 probes on *Eco*RI digested DNA of 12 plants regenerated from tissue culture (not shown).

### Estimation of the number of loci for this minisatellite family

The screening of the genomic library with the oligonucleotide at high stringency, revealed around 40 positive plaques in 4–5 equivalent genomes. Since among the 10 phages studied, 4 are different for the sequence hybridizing with the oligonucleotide, and 1 among them does not contain a minisatellite sequence, we can estimate that the minisatellite family is composed of around 10 members. This agrees with the number of bands revealed by DNA fingerprint with the oligonucleotide probe.

### Environment and location of the minisatellites in the *Arabidopsis thaliana* genome

Minisatellites 1 (ms1) and 2 (ms2) have revealed polymorphism between ecotypes Co and L and have been mapped using 96 recombinant inbred lines. Ms1 is located on chromosome 5 between markers g4715-b (position 27,2) and m247 (position 54,4) (30) at 11,3 centimorgans from m247. Ms2 is on chromosome 4 between markers g3843 (position 0) and g2616 (position 4,9) (30), at 2,8 centimorgans from g3843. For the minisatellite 3, we were not able to detect polymorphism. 6 DNA fragments flanking this minisatellite were assayed for polymorphism between the ecotypes Co and L with 12 restriction

enzymes. No polymorphism was detected and we could not map ms3. Ms4 which does not contain a minisatellite but microsatellites, is located on chromosome 3 at the same position as m583 (position 0) (30).

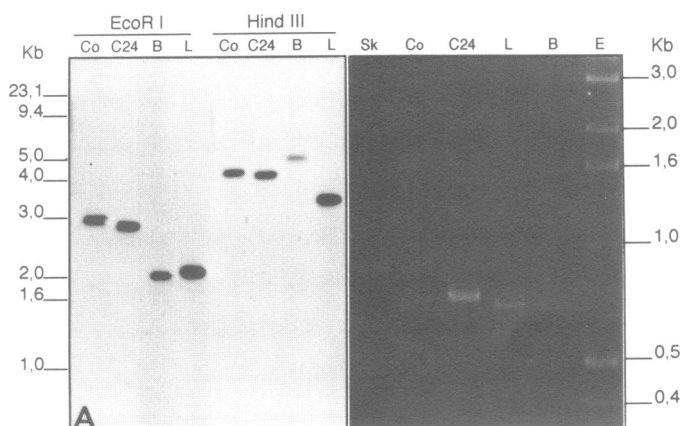
In addition to the localization of the minisatellite loci, we have analysed the DNA sequences flanking the minisatellites 1, 2 and 3. For each minisatellite, a random sample of adjacent sequences was hybridized on genomic southern blots. From a total of 15 clones, 10 were multicopy sequences and 5 corresponded to single sequence. The hybridization pattern, on genomic southern blots, of clones corresponding to multicopy sequences revealed that most of them were the 180bp satellite family. The 180bp units are adjacent to the minisatellites 1 and 2.

### DISCUSSION

We observed that low stringency PCR using random primers like RAPD (random amplified polymorphic DNA) (31) often amplifies tandemly or dispersed repeated sequences. One of the fragments amplified by this approach consists in a minisatellite DNA sequence with 33 bp repeated units and a high percentage in G+C. This PCR product and a 14 mer oligonucleotide based on the central part of the PCR minisatellite sequence were used as probes to isolate 4 clones, 3 of which contain a minisatellite. The repeated units for all minisatellites are 33bp long and present a 65% GC content. The minisatellite sequences can be in a completely different context since, for the minisatellite 3, the flanking sequences are 60% A+T rich.

In the human genome, minisatellites belonging to the same family share a central core sequence (2). The *Arabidopsis thaliana* minisatellites characterized in this work also share a central 14bp motif that is homologous with the human core sequence defined by Jeffreys *et al.* (2) (Table 1). The best homology (12 out of 14 bp) is observed between the *Arabidopsis* minisatellite 3 consensus sequence and the human myoglobin-2 consensus sequence (3) (Table 1). Strong cross-hybridization of the *Arabidopsis* core region on human DNA (Figure 2) confirms the high level of similarities between human and *Arabidopsis* core sequences. The *Arabidopsis* core region also cross-hybridized with mouse, *Drosophila melanogaster*, *Brassica napus*, maize and wheat DNA (Figure 2). This is not surprising since previous results have reported the detection of DNA fingerprints by hybridization of human minisatellite probes in various species (9–15) including plants, like *Populus deltoides* and *P. tremuloides* (16), rice (17) and tomato (18).

The presence of a core sequence in each repeat unit of a subset of minisatellites suggests that the core sequence may be directly involved in generating these hypervariable loci (9). Human hypervariable minisatellites act as hotspots of recombination (2,32,33), and show sequence similarity to the known recombination hotspot Chi in *E. coli*. In addition, a few groups have described DNA binding proteins that bind specifically to DNA containing a repeat sequence related to the human minisatellite core sequence. The protein described in (34) was purified from mouse tumor cells and was shown to enhance the ability of the minisatellite to undergo recombination. The function of the murine protein (35) is unknown but it is an ubiquitous minisatellite binding protein present in species ranging from *Drosophila* to man. Some authors (36) have also demonstrated the existence of two proteins, in nuclear extracts of human HeLa cells that bind specifically to hypervariable minisatellite DNA



**Figure 3.** Polymorphism of the locus 1 minisatellite. A: hybridization at high stringency of the minisatellite 1 probe on genomic DNA of four ecotypes; B: PCR amplifications with primers located in flanking regions. Sk: corresponding subclone in the pBluescript Sk vector, E: ladder.

sequences, but their role in recombination is still unproved. The identification of minisatellites in plants that share a core sequence highly similar to human suggests a conserved role for this region in various organisms. Based on these results, it is tempting to speculate that all the minisatellites sharing a similar core sequence could be recombinogenic and that ubiquitous proteins could be implicated in the recombination process.

To evaluate the potential of plant minisatellite as DNA markers, the different minisatellite loci were tested for their level of polymorphism among four ecotypes Co, C24, B, L. For the minisatellite 1, the length of the fragment is function of the number of copies of the tandem repeats present within the fragment (Figure 3B). This is probably true for minisatellite 2 since polymorphism could be revealed by 4bp cutting site enzymes. For the minisatellite 3, no polymorphism was observed, reflecting a same number of tandem repeats at this locus in the four ecotypes. The different probes tested (oligonucleotide, minisatellites 1 and 2 probes) did not detect polymorphism between individuals. Similar results were obtained in rice, where differences in hybridization patterns with a human probe were found only among cultivars, but not among individuals within a cultivar (17). On the contrary, a human probe revealed fragment differences between individuals from natural populations of the angiosperms *Populus deltoides* and *P.tremuloides* (16). However, in autogamous species, isolates tend very rapidly to homozygosity and this can explain the failure to observe polymorphism between individuals. We did not observe differences in hybridization pattern with the oligonucleotide and minisatellites 1 and 2 probes on regenerants DNA. The same result was obtained in rice where the hybridization pattern was the same before and after the regeneration of plants from tissue culture (17). However, changes affecting only a proportion of regenerants would probably have gone undetected in this experiment due to the low number of individuals analysed.

Ideally for mapping, minisatellites should be evenly dispersed throughout the linkage groups to provide general genome coverage. This is apparently the case in certain species such as mice and bovidae. However, in man, minisatellites are often clustered near telomeres. In *Arabidopsis*, 2 minisatellites have been mapped on chromosomes 4 and 5. Interestingly, the analysis of the DNA sequences flanking the minisatellites 1 and 2 revealed that they correspond to the 180bp satellite family. The 180bp satellite family is present at the centromeric region of all chromosomes (37,38). The minisatellite 1 map in the middle of chromosome 5 between markers m247 and g4715-b and is centromeric according to previous results (39,40). The centromere of chromosome 4 has not been mapped but this acrocentric chromosome possesses a large heterochromatic region that may include its entire short arm (40). This could explain the association of the minisatellite 2 with 180bp repeats despite its localization at the end of the short arm on chromosome 4. Thus, if the minisatellites 1 and 2 are adjacent to centromeric 180bp repeats, they would represent new markers for these two centromeres. They could also be useful to identify chromosome specific subfamilies of the 180 bp satellite. To determine the characteristics of the flanking regions of the minisatellites a random sample of clones carrying adjacent sequences was hybridized to southern blots of genomic DNA. 10 were multicopy sequences (like the 180bp satellite family) and 5 corresponded to single sequence. There is no enrichment for interspersed repeats around *Arabidopsis* minisatellites, in contrast to what was previously observed around human minisatellites (41).

## ACKNOWLEDGEMENTS

The authors thank A.Depeiges and S.Genestier for providing plants from the different ecotypes and regenerants. This work was supported by the CNRS (GDR 977) and by the Université Blaise Pascal.

## REFERENCES

1. Jeffreys A. J., Neumann R. and Wilson V. (1990) *Cell* **60**, 473–485.
2. Jeffreys A. J., Wilson V. and Thein S. L. (1985) *Nature* **314**, 67–73.
3. Nakamura Y., Leppert M., O'Connell P., Wolff R., Holm T., Culver M., Martin C., Fujimoto E., Hoff M., Krumlin E. and White R. (1987) *Science* **235**, 1616–1622.
4. Wong Z., Wilson V., Patel I., Povey S., Jeffreys A. J. (1987) *Ann. Hum. Genet.* **51**, 269–288.
5. Nakamura Y., Carlson M., Krapcho K., Kanamori M. and White R. (1988) *Am. J. Hum. Genet.* **43**, 854–859.
6. Jeffreys A. J., Wilson V. and Thein S. L. (1985) *Nature* **316**, 76–79.
7. Jeffreys A. J., Brookfield J. F. Y. and Semeonoff R. (1985) *Nature* **317**, 818–819.
8. Gil P., Jeffreys A.J. and Werret D. J. (1985) *Nature* **318**, 577–579.
9. Jeffreys A. J. (1987) *Biochem. Soc. Trans.* **15**, 309–317.
10. Jeffreys A. J. and Morton D. B. (1987) *Animal Genetics* **18**, 1–15.
11. Georges M., Gunawardana A., Threadgill D. W., Lathrop M. (1991) *Genomics* **11**, 24–32.
12. Burke T. and Bruford M. W. (1987) *Nature* **327**, 149–152.
13. Wetton J. H., Carter R. E., Parkin D. T. and Walters D. (1987) *Nature* **327**, 147–149.
14. Braithwaite K.S. and Manners J.M. (1989) *Curr. Genet.* **16**, 473–475.
15. Monastyrskii O. A., Ruban D. N., Tokarskaya O. N. and Ryskov A. P. (1990) *Genetika* **26**, 374–377.
16. Rogstad S. H., Patton II J. C., Schaal B. A. (1988) *Nucl. Acids Res.* **16**, 11378.
17. Dallas J. F. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6831–6835.
18. Broun P. and Tanksley S. D. (1993) *Plant Mol. Biol.* **23**, 231–242.
19. Vassart G., Georges M., Monsieur R., Brocas H., Lequarre A. S., Christophe D. (1987) *Science* **235**, 683–684.
20. Huey B. and Hall J. (1989) *J.Bacteriol.* **171**, 2528–2532.
21. Rogstad S. H., Patton II J. C., Schaal B. A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9176–9178.
22. Doyle J. J. and Doyle J. L. (1987) *Phytochem. Bull.* **19**, 11–15.
23. Sanger F., Nicklen S., Coulson A. R. (1977) *Proc. Natl. Acad. Sci. USA* **77**, 5463–5467.
24. Valvekens D., Van Montagu M. and Van Lijsebettens M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5536–5540.
25. Martinez-Zapater J. M., Estelle M. A. and Sommerville, C. R. (1986) *Mol. Gen. Genet.* **204**, 417–423.
26. Gruendler P., Unfried I., Pointner R. and Schweizer D. (1989) *Nucl. Acids Res.* **17**, 6395–6396.
27. Unfried I. and Gruendler P. (1990) *Nucl. Acids Res.* **18**, 4011.
28. Lagercrantz U., Ellegren H. and Andersson L. (1993) *Nucl. Acids Res.* **21**, 1111–1115.
29. Wang Z., Weber J.L., Zhong G. and Tanksley S.D. (1994) *Theor. Appl. Genet.*, **88**, 1–6.
30. Lister C; and Dean C. (1993) *The Plant J.* **4**, 745–750.
31. Williams J. G. K., Kubelik A. R., Livak K. J., Rafalski J. A. and Tingey S. V. (1990) *Nucl. Acids Res.* **18**, 6531–6535.
32. Wahls W.P., Wallace L.J. and Moore P. D. (1990) *Cell* **60**, 95–103.
33. White J.H., DiMartino J.F., Anderson R.W., Lusnak K., Hilbert D. and Fogel S. (1988) *Mol. Cell. Biol.* **8**, 1253–1258.
34. Yamazaki H., Nomoto S., Mishima Y. and Kominami R. (1992) *J. Biol. Chem.* **267**, 12311–12316.
35. Collick A. and Jeffreys A. J. (1990) *Nucl. Acids Res.* **18**, 625–629.
36. Wahls W. P., Swenson G. and Moore P. D. (1991) *Nucl. Acids Res.* **19**, 3269–3274.
37. Maluszynska J. and Heslop-Harrison J. S. (1991) *The plant J.* **1**(2), 159–166.
38. Bauwens S., Van Oostveldt P., Engler G. and Van Montagu M. (1991) *Chromosoma* **101**, 41–48.
39. Sears L.M.S. and Lee-Chen S. (1970) *Can. J. Genet. Cytol.* **12**, 217–223.
40. Koornneef M., Van Eden J., Hanhart C.J., Stam P., Braaksma F.J. and Feenstra W.J. (1983) *J. Heredity* **74**, 265–272.
41. Armour J.A., Wong Z., Wilson V., Royle N.J., Jeffreys A. (1989) *Nucl. Acids. Res* **17**, 4925–4935.