Volume 16 Number 14 1988 Nucleic Acids Research

Characterization of highly repetitive sequences of *Arabidopsis thaliana*

C.R.Simoens, J.Gielen, M.Van Montagu* and D.Inzé

Laboratorium voor Genetica, Rijksuniversiteit Gent, B-9000 Gent, Belgium

Received May 4, 1988; Accepted June 15, 1988 Accession nos+

ABSTRACT

We have analyzed three classes of highly repetitive DNA sequences of Arabidopsis thaliana, composed of tamdemly repeated units of 180 bp, 500 bp, and 160 bp, respectively. The three families comprise approximately 2% of the Arabidopsis genome and are the major component of the highly repetitive DNA. The 500-bp element arose by duplication of one half of a 180-bp ancestor and insertion of a foreign segment between the two duplicated parts followed by amplification. The repeat elements contain occasionally palindromes and other motifs but none are significantly conserved. There is no significant similarity with previously published repetitive elements. Heterogeneity between monomers ranges from 6% to 17%. Monomers derived from different clusters in the genome are more diverged than monomers of the same array.

INTRODUCTION

Arabidopsis thaliana has the smallest known haploid genome size (7 x 10^7 bp) among higher plants [1]. Gene families in Arabidopsis are more simply organized compared to other plants [2,3]. The genes tend to contain less and smaller introns and to be more densely packed on the genome [4,5]. Furthermore, only a limited amount of repetitive DNA is present [1,6]. The latter makes Arabidopsis a useful system to study structure and function of repetitive DNA.

Repetitive sequences usually constitute more than 50% of plant genomes [7]. Some sequence elements are arranged as long tandem arrays in heterochromatic regions at telomeres, centromeres, or interspersed chromosomal locations whereas others are scattered through the genome as individual elements. The repeat elements can display a wide spectrum of repetition frequency up to more than 10^6 copies per genome, and the basic repeat unit can range from a simple dinucleotide pair to several kilobases in length [8].

Apart from a small fraction which encodes important functions such as ribosomal RNAs, the biological significance of the repetitive DNA remains

controversial. It has been proposed to be involved in chromosome-folding and pairing, determination of nuclear size, gene regulation or speciation processes [8-12]. Alternatively, repetitive DNA may have no function at all. It might reside and accumulate in the genome because there is no phenotypical disadvantage. From this point of view, the influence on processes such as chromosome-folding or speciation is considered to be accidental [13,14].

Arabidopsis could be regarded as a species in which selective forces acted to reduce the amount of repeated sequences, e.g. to allow its very short life cycle [15]. Nevertheless, it has retained some repetitive DNA. One family has already been described [16]. It comprises 1-1.5% of the Arabidopsis genome and consists of long tandem arrays of a 180-bp element possibly associated with heterochromatic regions. In this paper we give additional information on the 180-bp family and describe two more repeat families.

MATERIALS AND METHODS

Plasmid and Phage Clones

The plasmids and phage clones used in this work are listed in Table 1.

Plant Culture

Arabidopsis thaliana (collection number C24) seeds were a generous gift of Prof. M. Jacobs (Vrije Universiteit Brussel, Belgium). The plants were grown in a greenhouse under 16 hours light/8 hours dark illumination at

Table 1. Plasmids and phage clones

	Antibiotic resistance	Relevant characteristics	Origin
Plasmids			
pGem-2	Cb	RNA transcription vector	Promega Biotec
pC22	Cb Sp Sm	Binary vector containing the oriV of both pBR322 and the Ri plasmid pRiHR1	[17]
Phage clones			
λbAt002		A λ clone carrying the 9.9-kb repeat unit of the large ribosomal DNAs (18S and 26S) of A. thaliana	[6]

Abbreviations : Cb, carbenicillin; Sm, spectinomycin; Sp, streptomycin.

22–24°C and 70–80% humidity. Six weeks after germination mature plants were harvested.

Construction and Screening of an Arabidopsis thaliana Cosmid Library

The construction and characterization of the A. thaliana (variety Columbia) library in the cosmid vector pC22 has been described [17]. The library was used to construct a sublibrary of 1145 randomly selected clones, which were individually grown in wells of microtiter plates. All manipulations with the sublibrary were as described elsewhere [18].

Recombinant DNA Techniques and Hybridizations

For large- and small-scale plasmid DNA preparations the procedure of [19] was used. Total Arabidopsis DNA was prepared according to [6]. Chloroplast DNA of A. thaliana (variety Columbia) was a generous gift of Dr. G. Redei (Columbia, U.S.A.). Phage DNA was isolated essentially as described by [20]. All recombinant DNA manipulations were according to [21]. As cloning vector pGem-2 was used (Promega Biotec, U.S.A.).

^{32}P-labelled DNA probes were obtained by nick translation as described by [22]. The nylon filter (Hybond-N, Amersham) obtained upon Southern blotting was prehybridized for 2 hours in 3 x SSC, 0.25% nonfat dry milk. After denaturation (10 minutes boiling) the ^{32}P-labelled probe was added to the filter in hybridization mix (3 x SSC, 0.25% nonfat dry milk; approximately 10^6 cpm/ml). The hybridization was carried out for 16 hours at 68°C. The filters were washed in 3 x SSC with 0.1% SDS at 68°C (three quick rinses and twice for 30 minutes). DNA dot-blot filters were hybridized and washed under the same conditions.

Sequence Analysis

Sequencing was carried out on both strands according to the chemical method developed by [23]. Sequence comparisons were made using previously described computer programs [24,25]. The dot-matrix was obtained using a S.A.S.I.P. program [26].

DNA Dot-Blot

For DNA dot-blotting the procedure described by Amersham (U.K.) was followed using a Schleicher & Schüll (F.R.G.) dot-blotting apparatus. After hybridization the number of cpm bound at each spot on the filter was determined by liquid scintillation counting.

RESULTS

Isolation of Three Repeat Families

To identify clones containing repetitive sequences, a set of 1145 genomic

cosmid clones was replicated on nitrocellulose filters and subsequently hybridized with ^{32}P-labelled total Arabidopsis DNA. The degree of hybridization, as revealed by autoradiograms of these filters, reflects the repetitivity of the Arabidopsis insert DNA in the clones. In total, 204 (17.8%) of the analyzed clones, each with an average insert size of 23.6 kb, showed a positive hybridization signal with varying intensities [18]. As a background control a strain was used containing the cosmid vector pC22 without insert DNA.

The nature of the repetitive DNA was further analyzed by colony hybridization with ^{32}P-labelled A. thaliana chloroplast DNA and 18S/26S ribosomal DNA. The latter DNA was isolated from λbAt002, a λ clone carrying the 9.9-kb repeat unit of the large ribosomal DNAs of A. thaliana [6]. Fifty-four percent (110 out of 204) of the repetitive clones show homology with chloroplast DNA. Although some of these clones might contain homologous nuclear DNA, studies on promiscuous DNA in other plants suggest that these sorts of sequences should constitute only 1 to 2% of the haploid genome [27]. Mitochondrial DNA was not ruled out but can only comprise 1–2% of the total DNA [1]. This means that approximately 2.4 x 10^7 bp of nuclear DNA were analyzed. Two percent (21) of the remaining 1035 clones contain ribosomal DNA. This represents a 4-fold under-representation of rDNA in our library compared to the calculations of [6]. This difference is probably due to extensive rearrangements by homologous recombination of the tandemly repeated units [17]. The rearrangements could have been enhanced by the subrepeats present in the intergenic spacer which serve as hot spots for recombination [28]. Alternatively, the divergent copy numbers could at least partly be the result of between-individual and within-individual differences as frequently noticed in other plant species [28]. Of the remaining 73 repetitive clones, 47 hybridizing clones were not analyzed further. Presumably, some of these contain genes encoding for the 5S RNA, tRNAs or small nuclear RNAs.

Extensive restriction analysis showed that the 26 strongly hybridizing clones could be subdivided into three families. The first family contains 10 clones which in general gave the most intense signal in the colony hybridization. They are characterized by the presence of multiple copies of a 180-bp HindIII fragment. Four different repeats of this family were subcloned. The second family, consisting of 13 clones, contains a tandemly repeated HindIII fragment of about 500 bp. Two 500-bp repeats, derived from two clones, were subcloned. The three remaining clones constitute the

third family. A small repetitive fragment of about 160 bp could be obtained upon AluI digestion. One trimer was subcloned. Using the three repeats as a probe, colony hybridizations were performed which confirmed the composition of the different families. All the clones of the first family cross-hybridize weakly at high stringency with a 500-bp repeat probe. In all three groups a typical ladder consisting of integral multiples of the basic repeat unit can be obtained upon partial digestion, indicating that the fragments are tandemly organized (see further).

Sequence Analysis

180-bp Repeat. Four isolated 180-bp repeats were analyzed (Figure 1A). AR11 and AR12 were derived from pATC5D2 whereas AR13 and AR14 originated from pATC12G8. At the same time that the first sequencing data were obtained similar sequences were published [16]. One particular feature we observed is the presence of a so-called "super-HindIII" structure in AR11 (Figure 1A). It contains three HindIII sites (AAGCTT) separated from each other by a box of 14 nucleotides, one consisting mainly of AAG repeats and the other of repeated CTT. In the more common repeat of ± 180 bp length (in our case varying from 177 bp to 179 bp) only the CTT box is present bordered by two potential HindIII sites. In practice, only one of these has the correct recognition sequence, thus conserving the 180-bp repeat size [16]. However, AR11, which has three correct HindIII sites, indicates that at least some heterogeneity must occur, not only resulting in multimers, but also in shorter and longer repeats of 160 bp and 200 bp, respectively. It is difficult to imagine a selective force that maintains functional HindIII sites 180 bp apart. Although the presence of the "super-HindIII" structure is clearly not the rule, it could be a relic of the original repeat organization.

Sequence variation between class members is due to point mutations and to a lesser extent to single base pair insertions and deletions. Between monomers derived from an independent clone homology ranges from 85% to 87%. Comparison between AR11 and AR12, or AR13 and AR14 reveals a homology of 92% and 94%, respectively. The additional 20 bp present in AR11 was not considered when calculating the percentages.

The GC content of the repeats (35.5% to 38.2%) is slightly lower than the average GC content (41.4%) of Arabidopsis [1]. No internal subrepeat could be found. In AR13 and AR14 a stem-loop structure flanked by short, direct repeats (TTGT) could be localized between position 52 and 86. It consists of a 7-bp loop and a 10-bp stem with nine complementary base

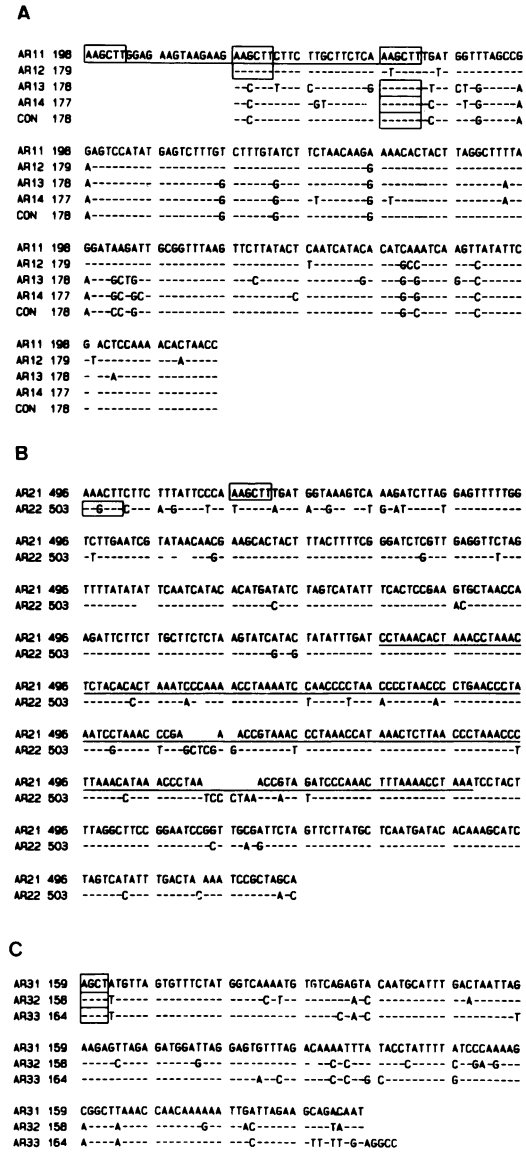

Figure 1. Nucleotide sequences of the different cloned repeat elements of Arabidopsis.
For each family only the first line shows the complete sequence written from 5' to 3'. Changes in the other sequences are noted. An empty space without hyphen refers to a deletion. Functional restriction sites (HindIII or AluI) which border the units are indicated by a box. A. 180-bp family; the "super-HindIII" structure is underlined; CON, designates the published consensus sequence [15]. B. 500-bp family; the telomere-like domain is underlined. C. 160-bp family.

pairs. Despite the low probability that this would have arisen by chance ($P = 7 \times 10^{-10}$), the structure is not conserved in AR11 and AR12 nor in the four repeats published by [16] : one to two base substitutions occur in the direct repeats and one to three substitutions in the stem part.

500-bp Repeat. Sequences of AR21 and AR22, subcloned from pATC3E8 and pATC6H7, respectively, are shown in Figure 1B. They have a 36.7% and 36% GC content, respectively, and are 88% homologous. A ± 190-bp domain can be found, mainly consisting of C and A, sometimes interrupted by T or less frequently by G. As an approximation a repeat unit can be written as $C_{1-4}T_{1-3}A_{4-2}T_{0-2}$ which resembles the telomeric repeats found in Tetrahymena (CCCCAA), Trypanosoma (CCCTAA), Stylonichia (CCCCAAAA), or Physarum ($CCCTA_n$) [29].

Sequence comparison revealed striking homology (70–80%) between some

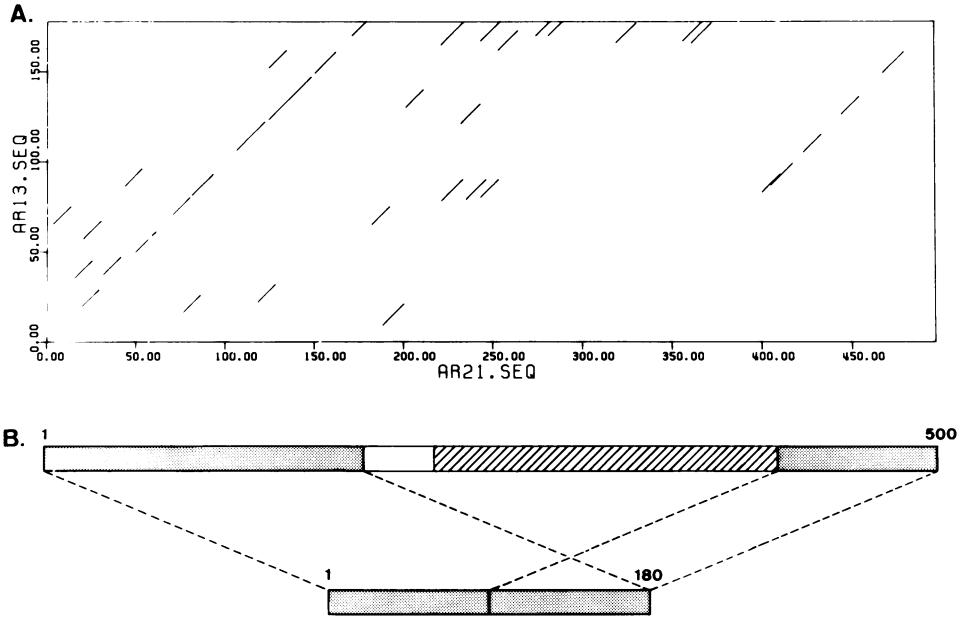

Figure 2.
A. Dot-matrix comparing a 500-bp element (AR21) with a 180-bp element (AR13). Each diagonal line represents a box of 8 base pairs of which six are homologous between the two compared sequences. The dot-matrix shows that the first part of AR21 is homologous to AR13 whereas the last part of AR21 is homologous to the second half of AR13.
B. Scheme of the homologies found between the 500-bp and 180-bp repeat elements. Grey areas indicate homologous regions and the hatched area, the telomere-like domain.

regions of the 500-bp repeat and the 180-bp unit. This is outlined in Figure 2. It seems that the 500-bp element arose by sequence rearrangement of a 180-bp fragment and subsequent reamplification. One half of the 180-bp fragment was duplicated and between the duplicated parts a foreign sequence (the telomere-like domain) was inserted.

160-bp Repeat. The 160-bp family presumably arose independently since no significant homology with the other families could be found. Three adjacent repeats, subcloned from pATC8G4, were sequenced (Figure 1C). Homology ranges from 83% to 86%. Especially close to the AluI site, some variation occurs which explains the differences in length (158, 159, 164 bp). The GC content lies between 32% and 36%. No special features were noticed concerning the internal structure (direct or inverted repeats, subrepeats).

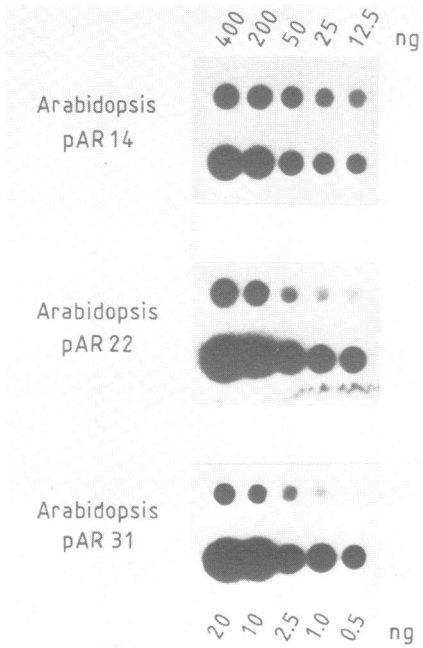

Figure 3. Dot hybridization to estimate the relative amount of the repetitive families in the genome.
Nylon filters were prepared with increasing amounts of plasmid DNA (upper numbers) and of total Arabidopsis DNA (bottom numbers). After hybridization to the corresponding ^{32}P-labelled plasmid, the number of counts per ng of repeat was calculated and compared with the radioactivity bound to the genomic DNA spots. The percentage of the different repetitive families has been deduced from linear regression analysis of five independent experiments.

Comparison with Repeat Sequences of Other Organisms

All repeats were compared with published repeated sequences of the following organisms : cauliflower [30], rice [31], radish [32], pumpkin [33], mustard [34], onion [35], maize [36], man [37], and the insect Chironomus [38]. In many cases one or more sequence similarities could be identified which could have arisen by chance alone at probabilities of 10^{-5} to 3×10^{-8} (data not shown). Even with the Chironomus and human repeats similarities could be registered with P varying between 10^{-6} and 5×10^{-6}. Comparison between the 180-bp repeat of Arabidopsis and a 177-bp HindIII repeat of cauliflower (belonging to the same family Cruciferae) could only reveal one similarity ($P = 3 \times 10^{-6}$). The observed sequence similarities are most likely not significant.

Genomic Content and Organization

The fraction of the Arabidopsis genome constituted by the different families was determined with a reconstruction experiment as indicated in Figure 3. The 180-bp family constitutes 0.8–1.4% of the total Arabidopsis genome which corresponds to 3100–5500 copies per haploid genome. This is in good agreement with the number obtained by [16]. The second family comprises 0.2–0.4% of the total DNA which means that 280–560 copies of the 500-bp repeat element exist per haploid genome. The 160-bp family also comprises 0.2–0.4% of the total DNA corresponding to a copy number of 875–1750 per haploid genome. These values presumably represent an underestimate since between 20% and 30% of the total cellular DNA can be chloroplastic DNA [1].

In general, these data agree with the results obtained by colony hybridization except for the second family of which 13 clones were isolated or 1.3% of the 1035 nuclear DNA-containing clones. This anomaly could be explained if the 500-bp repeats are more dispersed through the genome. The way the library was constructed could have offered a selection for interspersed repeated sequences. The 180-bp repeat clusters were estimated to be $\geq$ 50 kb [16]. Although in some of the 180-bp repeats a Sau3A site is present, it is possible that during the construction of the partial Sau3A cosmid library, cloning of these clusters was counterselected. Since the 500-bp elements have a normal distribution of Sau3A sites (2 and 4 sites, respectively) counterselection does not occur.

The ladders obtained upon partial digestion of the genomic clones indicate that the repetitive elements are tandemly organized. Hybridization to partially restricted Arabidopsis DNA with the different repeats

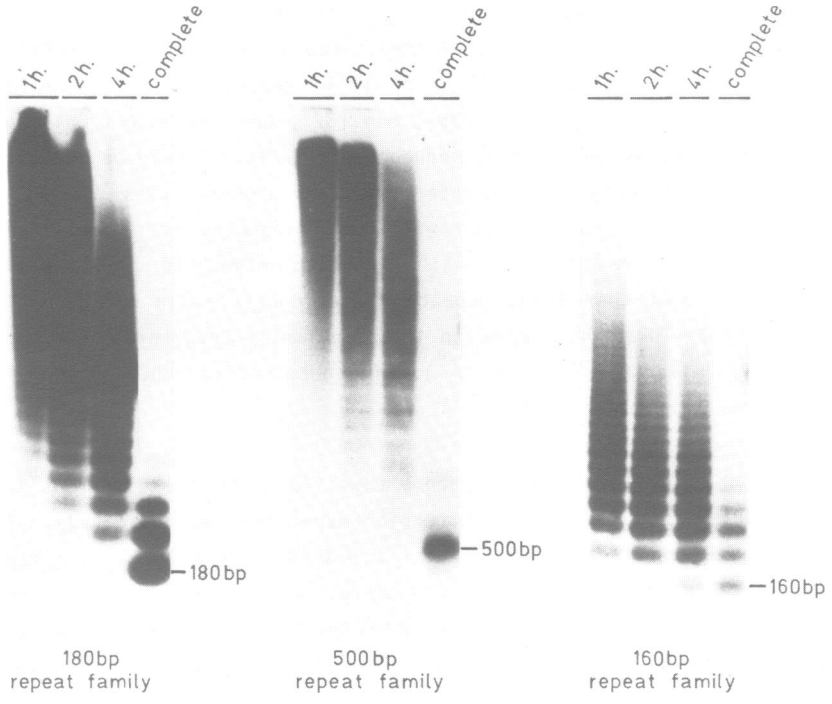

Figure 4. Tandem arrangement of the three repeat families.
One µg of Arabidopsis cellular DNA was incubated with one unit of restriction enzyme for different times (1 hour, 2 hours, 4 hours). In the fourth lane the Arabidopsis DNA was digested to completion by the addition of excess enzyme and a prolonged incubation. Electrophoresis was carried out on a 0.8% agarose gel. Labelled AR14 and AR22 were hybridized to HindIII-restricted Arabidopsis DNA to demonstrate the organization of the 180-bp and 500-bp family, respectively. The labelled 160-bp element AR31 was probed against AluI-digested Arabidopsis DNA.

also resulted in typical ladders consisting of integral multiples of the basic repeat sequence (Figure 4). In case of the 500-bp unit, a faint band of 300–400 bp appeared under the main band of 500 bp, presumably representing fragments which lost part of the sequence. In all three families some multimers are not digested even if excess enzyme is added (Figure 4). This was also noticed with several genomic clones which were originally isolated indicating some sequence heterogeneity.

By comparing the patterns obtained upon hybridization of MspI- or HpaII-digested Arabidopsis DNA with a 500-bp repeat probe, the inner cytosine was shown to be methylated in a majority of the CCGG sites situated in

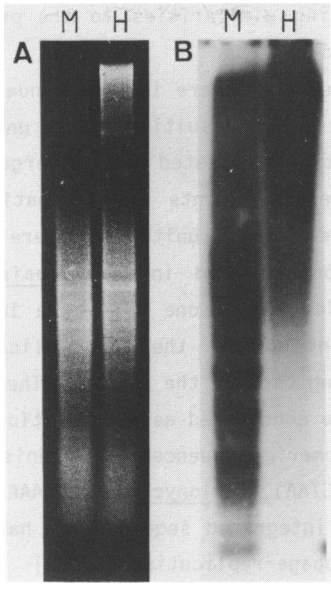

Figure 5. Restriction of total Arabidopsis DNA with the methylation-sensitive enzymes HpaII (H) and MspI (M).
A. Three µg of completely digested DNA were loaded on each lane, resolved on a 0.8% agarose gel and visualized by ethidium bromide staining.
B. Autoradiogram of the same lanes after Southern hybridization to the labelled 500-bp element AR22. The 500-bp family is only partially cleaved by HpaII.

the 500-bp elements (Figure 5). A similar conclusion was drawn concerning the 180-bp repeats which is in agreement with previously published results [16]. There is no HpaII/MspI site in the 160-bp repeat.

DISCUSSION

In this paper we describe three classes of highly repetitive DNA sequences of Arabidopsis thaliana. The high degree of methylation, the tandem arrangement, and the slightly lower GC content, compared to the average value for the Arabidopsis genome, suggest that the 500-bp and 180-bp family and possibly the 160-bp family can be considered as satellite-like DNA [8]. Together they comprise approximately 2% of the Arabidopsis nuclear genome which is consistent with its simple organization [1,6]. According to our colony hybridization experiment the described repeat families comprise the major fraction of highly repetitive DNA in Arabidopsis. The Arabidopsis

repetitive DNA has striking similarities to the properties of repetitive sequences found in other organisms.

It is generally agreed that there is a continual flux of DNA sequences in the higher eukaryotic genome resulting in new unique or repetitive DNA. By accumulation of mutations repeated DNA diverges to unique sequences whereas insertion of foreign elements or combination with adjacent, non-repeated DNA leads to new repeat units which are amplified [39,40]. A similar phenomenon can be observed in Arabidopsis. The 500-bp element presumably arose by duplication of one half of a 180-bp repeat and insertion of a foreign element between the two duplicated parts followed by amplification and spreading through the genome. The inserted element which can as an approximation be considered as a repetition of $[C_{1-4}T_{1-3}A_{4-2}T_{0-2}]$ blocks resembles the telomeric sequences of organisms such as Tetrahymena (CCCCAA), Trypanosoma (CCCTAA), Stylonychia (CCCCAAAA) or Physarum ($CCCTA_n$) [29]. Alternatively, the integrated sequence may have arisen independently by processes such as slippage-replication [41,42]. This mechanism, involving the slippage of short oligonucleotide motifs against each other with formation of single-stranded loops, may be an important source of variation in repetitive DNA [42]. The so-called "super-HindIII" structure present in one of the 180-bp repeats could also have arisen by DNA slippage.

All three families show a number of inverted and direct repeats but none of these were significantly conserved within the family. Although the observed motifs could be remnants of the original structure of the repeat element, no individual sequence is essential for the maintenance of the larger repeat [43,44].

Comparison of the isolated repeat sequences with previously published ones (see results) revealed several similar sequences that could occur by chance alone with a low probability (10^{-6} to 5×10^{-8}). However, no consistantly conserved region could be identified. Repetitive elements of the crucifers cauliflower, mustard and radish, all belonging to the same tribe Brassiceae, were shown to have homologies ranging from 75% to 80% [30,32]. However, the 180-bp elements of Arabidopsis, belonging to a different tribe Sysimbrieae [45], are no more similar to the sequences in these plants than to a human repetitive element [37].

The rapid divergence between different species in contrast to the considerable homogeneity within a species suggests that a family of repeated sequences evolves in concert. Several mechanisms such as unequal crossover and gene conversion may be involved [46,47]. Although only a limited

number of repeat units has been analyzed, the observation in the 180-bp family that monomers derived from different clusters in the genome are more diverged (13–15%) than monomers of the same array (6–8%) suggests that if concerted evolution is at work, then it is not occurring uniformly throughout the genome. The proposed mechanisms both require contact between different chromosomes. It is possible that not all arrays present in the genome interact equally [35]. On the other hand, the elements of one array could have arisen from a common ancestor, e.g. by a rolling circle mechanism [30].

Although many hypotheses on the function of satellite-like DNA have been postulated, their specific role in the genome remains a matter of debate. Their low occurrence in Arabidopsis will allow us to study if these repetitive DNAs are retained because of a specific function or if they are simply the last junk or parasitic relics in the economical Arabidopsis genome.

Acknowledgements

The authors wish to thank Drs. E. Krebbers and A. Caplan for critical reading of the manuscript, and V. Vermaercke, K. Spruyt and M. De Cock for help in preparing it. This work was supported by grants from the "A.S.L.K. Kankerfonds", the "Fonds voor Geneeskundig Wetenschappelijk Onderzoek" (FGWO #3.0001.82), and the Services of the Prime Minister (OOA #12.0561.84 and IUAP #12.0C.0187). CRS is a Research Assistant and DI a Senior Research Assistant of the National Fund for Scientific Research (Belgium).

*To whom correspondence should be addressed

+X06466, X06467, X06468, X06469, X06470, X06471, X06472

REFERENCES

1. Leutwiler, L.S., Hough-Evans, B.R. and Meyerowitz, E.M. (1984) Mol. Gen. Genet. 194, 15-23.
2. Meyerowitz, E.M. and Pruitt, R.E. (1985) Science 229, 1214-1218.
3. Chaboute, M.-E., Chaubet, N., Philipps, G., Ehling, M. and Gigot, C. (1987) Plant Mol. Biol. 8, 179-191.
4. Chang, C. and Meyerowitz, E.M. (1986) Proc. Natl. Acad. Sci. USA 83, 1408-1412.
5. Klee, H.J., Muskopf, Y.M., and Gasser, C.S. (1987) Mol. Gen. Genet. 210, 437-442.
6. Pruitt, R.E. and Meyerowitz, E.M. (1986) J. Mol. Biol. 187, 169-183.
7. Bennett, M.D. and Smith, J.B. (1976) Phil. Trans. R. Soc. London B 274, 227-274.
8. Singer, M.F. (1982) Int. Rev. Cytol. 76, 67-112.
9. Cavalier-Smith, T. (1978) J. Cell Sci. 34, 247-278.
10. John, B. and Miklos, G.L.G. (1979) Int. Rev. Cytol. 58, 1-114.
11. Davidson, E.H. and Britten, R.J. (1979) Science 204, 1052-1059.

12. Flavell, R.B. (1982) in Genome evolution, Dover, G.A. and Flavell, R.B. Eds., pp. 301-323. Academic Press, London.
13. Orgel, L.E. and Crick, F.H.C. (1980) Nature (London) 284, 604-607.
14. Doolittle, W.F. and Sapienza, C. (1980) Nature (London) 284, 601-603.
15. Bennett, M.D. (1972) Proc. Roy. Soc. (ser. B) 181, 109-135.
16. Martinez-Zapater, J.M., Estelle, M.A. and Somerville, C.R. (1986) Mol. Gen. Genet. 204, 417-423.
17. Simoens, C., Alliotte, Th., Mendel, R., Müller, A., Schiemann, J., Van Lijsebettens, M., Schell, J., Van Montagu, M. and Inzé, D. (1986) Nucl. Acids Res. 14, 8073-8090.
18. Simoens, C.R., Peleman, J., Valvekens, D., Van Montagu, M. and Inzé, D. (1988) Gene, in press.
19. Birnboim, H.C. and Doly, J. (1979) Nucl. Acids Res. 7, 1513-1523.
20. Kao, F.-T., Hartz, J.A., Law, M.L. and Davidson, J.N. (1982) Proc. Natl. Acad. Sci. USA 79, 865-869.
21. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, New York.
22. Rigby, P.W.W., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
23. Maxam, A.M. and Gilbert, W. (1980) in Nucleic Acids, Part I, (Methods in Enzymology, Vol. 65), Grossman, L. and Moldave K. Eds., pp. 499-559. Academic Press, New York.
24. Staden, R. (1980) Nucl. Acids Res. 8, 817-825.
25. Kanehisa, M.I. (1982) Nucl. Acids Res. 10, 183-196.
26. Clavérie, J.-M., Caudron, B. and Gérard, O. (1984) Biofutur 25, 35-37.
27. Timmis, J.N. and Scott, N.S. (1984) Trends Biochem. Sci. 9, 271-273.
28. Rogers, S.O. and Bendich, A.J. (1987) Plant Mol. Biol. 9, 509-520.
29. Blackburn, E.H. and Szostak, J.W. (1984) Ann. Rev. Biochem. 53, 163-194.
30. Benslimane, A.A., Dron, M., Hartmann, C. and Rode, A. (1986) Nucl. Acids Res. 14, 8111-8119.
31. Wu, T. and Wu, R. (1987). Nucl. Acids Res. 15, 5913-5923.
32. Grellet, F., Delcasso, D., Panabieres, F. and Delseny, M. (1986) J. Mol. Biol. 187, 495-507.
33. Leclerc, R.F. and Siegel, A. (1987) Plant Mol. Biol. 8, 497-507.
34. Capesius, I. (1983) Biochim. Biophys. Acta 739, 276-280.
35. Barnes, S.R., James, A.M. and Jamieson, G. (1985) Chromosoma 92, 185-192.
36. Peacock, W.J., Dennis, E.S., Rhoades, M.M. and Pryor, A.J. (1981) Proc. Natl. Acad. Sci. USA 78, 4490-4494.
37. Kiyama, R., Matsui, H., Okumura, K. and Oishi, M. (1987) J. Mol. Biol. 193, 591-597.
38. Saïga, H. and Edström, J.-E. (1985) EMBO J. 4, 799-804.
39. Flavell, R. (1980) Ann. Rev. Plant Physiol. 31, 569-596.
40. Roizès, G.P. and Pagès, M. (1982) in Genome evolution, Dover, G.A. and Flavell, R.B. Eds., pp. 95-111. Academic Press, London.
41. Hentschel, C.C. (1982) Nature (London) 295, 714-716.
42. Tautz, D., Trick, M. and Dover, G.A. (1986) Nature (London) 322, 652-656.
43. Wu, J.C. and Manuelidis, L. (1980) J. Mol. Biol. 142, 363-386.
44. Miklos, G.L.G. and Gill, A.C. (1981) Chromosoma 82, 409-427.
45. Hedge, J.C. (1976) in The biology and chemistry of Cruciferae, Vaughan, J.G., Jones, B.M.G. and Macleod A.J. Eds., pp. 1-45. Academic Press, New York.
46. Smith, G.P. (1976) Science 191, 528-535.
47. Dover, G. (1982) Nature (London) 299, 111-117.