Plant Molecular Biology **27**: 853–862, 1995.

Analysis of a repetitive DNA family from *Arabidopsis arenosa* and relationships between *Arabidopsis* species

A. Kamm*, I. Galasso[1], T. Schmidt and J.S. Heslop-Harrison
*Karyobiology Group, Department of Cell Biology, John Innes Centre, Colney Lane, Norwich, NR4 7UH, UK (*author for correspondence); [1] Germplasm Institute, CNR, via G. Amendola 165/A, 70126 Bari, Italy*

Received 13 September 1994; accepted in revised form 6 December 1994

Key words: *Arabidopsis arenosa*, Cruciferae, genus *Arabidopsis*, *in situ* hybridization, molecular systematics, satellite DNA

Abstract

We have analysed a family of highly repetitive DNA from *Arabidopsis arenosa* (L.) Lawalrée [syn. *Cardaminopsis arenosa* (L.) Hayck] composed of AT-rich tandem repeats of 166–179 bp in head to tail organization. Sequence comparison between several repeat units revealed a high level of divergence of 4.5% to 25%. The sequence family shows more than 58% homology to satellite sequences described in *Arabidopsis thaliana* (L.) Heynh. but no homology to other satellite repeats in the Cruciferae. Within the genus *Arabidopsis* the satellite sequence was found to be present in *A. thaliana* and *Arabidopsis suecica* (Fries) Norrlin, but not in *Arabidopsis griffithiana* (Boiss.) N. Busch and *Arabidopsis pumila* (Stephan) N. Busch. *In situ* hybridization to metaphase chromosomes of *A. arenosa* ($2n = 4x = 32$) showed the sequence to be localized at the centromeres of all 32 chromosomes with substantial hybridization along the chromosome arms. Using Southern hybridization and *in situ* hybridization, we give evidence that *A. suecica* is a hybrid of *A. thaliana* and *A. arenosa*. A considerable reorganization of the *A. thaliana* satellite sequence pAL1 occurred in the hybrid genome while no molecular change of the *A. arenosa* repeat was observed in the hybrid. Analysis of related repeats enabled differentiation between closely related genomes and are useful for the investigation of hybrid genomes.

Introduction

Arabidopsis thaliana (L.) Heynh. is currently among the most extensively studied organisms because of its many advantages for molecular analysis including the small genome and plant size and rapid life cycle [21, 34]. However, there is relatively little information about the genomes and genetics of the closest relatives of *A. thaliana*. The present work was initiated to examine the genomes of close relatives of *A. thaliana* and, in particular, to investigate the type of changes which have occurred at the chromosomal and repetitive DNA sequence levels [17, 29]. It is possible that examination of such changes in genomes which are highly amenable to analysis may reveal test-

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers Z48158 (pAa214), Z48159 (pAa27), Z48160 (pAa271), Z48161 (pAa524), and Z48171 (pAa519).

able models for genome evolution in other species groups, and perhaps the functions of repetitive DNA.

Nuclear genomes of angiosperms vary in size more than 2400 times [5, 6], but a very large fraction of all plant genomes consists of repetitive DNA: sequences consisting of motifs between two and ten-thousand or more base pairs which are repeated hundreds or thousands of times, which may be at a single chromosomal site or many sites dispersed over the whole genome. Reassociation kinetic analysis showed that some 40% of the *A. thaliana* genome (100 Mb [18]) consists of repetitive DNA [25], while more than 95% of the DNA in species with large genomes, such as *Allium cepa*, may consist of repetitive DNA [13].

In previous work, we described the location and number of sites of the ribosomal DNA (rDNA), a sequence which comprises about 8% of the genome, in *A. thaliana* [28] and other *Arabidopsis* species [29]. Tandemly repeated DNA is another major class of repetitive DNA which has been analysed in *A. thaliana* [30, 42] and other crucifers [7, 8, 12, 13, 20]. Typically, these sequences consist of a monomeric unit of about 170–180 bp or 350–360, but they differ widely in copy number in a genome and in their nucleotide sequence. Some sequence motifs are conserved over quite large evolutionary distances, while others differ between closely related species [15]. Nevertheless, many studies of tandemly repeated sequences within groups of crucifers (*Raphanus* [12], *Brassica* [16, 24]) indicate that the distribution of these sequences along chromosomes and between species is of considerable evolutionary interest for understanding and reconstructing both phylogenetic and molecular aspects of evolution.

The taxonomy of the ephemeral, herbaceous annual species of the Cruciferae with rosettes currently in the tribes Arabideae and Sysimbrideae is extremely confused; the limits of both the tribes and genera are poorly defined and differ extensively between authorities [1]. There appears to be more agreement, from Linnaeus [27] onwards, about the constitution of species, although even here there are difficulties at least partly because of the plasticity of the phenotype typical of weedy ephemerals. While *Arabidopsis* has been considered a monotypic genus with only the species *A. thaliana* [37], most authors consider the genus much larger, with a centre of diversity in the Himalaya, with a few species extending into Eastern Europe and Scandinavia (including those studied here), and at least *A. thaliana* extending into North America. A much-needed revision of the group is currently underway (I.A. Al-Shebaz, personal communication). In view of the results presented here, involving both DNA sequence homology and natural hybrid species, we are referring to various species as members of the genus *Arabidopsis* (see Material and methods), although we recognize that the boundaries of the genus are subject to revision. The enormous amount of molecular data about one species, *A. thaliana*, in conjunction with morphological data, should be of value in defining the species and genera. Access to a wider pool of genetic variation and data about genomes of related species will be of considerable benefit to the analysis of evolution and genetics of *A. thaliana* [29].

In the present study, we investigated the molecular structure and chromosomal location of a major tandemly repeated satellite DNA family from *Arabidopsis arenosa* (L.) Lawalrée ($2n = 2x = 16$ or $2n = 4x = 32$) and compared them with repetitive sequences from *A. thaliana*. A second species, *A. suecica* (Fries) Norrlin ($2n = 26$), is considered to be of hybrid origin [44] with *A. arenosa* ($2n = 2x = 16$) and *A. thaliana* as parents. We therefore investigated the presence and distribution of tandemly repeated sequence families from *A. arenosa* and *A. thaliana* in the three species and other relatives to examine the evolution, diversification and genomic distributions of these sequences.

Material and methods

Plant material

The species used in this study were: *Arabidopsis arenosa* (L.) Lawalrée (syn. *Cardaminopsis*

arenosa (L.) Hayck) ($2n = 4x = 32$), *Arabidopsis thaliana* (L.) Heynh./wild-type Columbia ($2n = 10$), *Arabidopsis pumila* (Stephan) N. Busch ($2n = 32$), *Arabidopsis griffithiana* (Boiss.) N. Busch ($2n = 32$), *Arabidopsis suecica* (Fries) Norrlin (syn. Hylandra suecica (Fries) A. Löve) ($2n = 26$) kindly provided by Dr O. Savolainen, collected in Finland, and *Brassica nigra* (L.) Koch ($2n = 16$). The plants were grown in glasshouses and in controlled environment chambers.

Isolation, cloning and sequencing of plant DNA

Total genomic DNA was extracted from leaves (1 g fresh weight) following the protocol of Dellaporta *et al.* [11]. Genomic DNA from *A. arenosa* was digested with different restriction enzymes and fractionated on a 1.1% agarose gel. The restriction profile of *Hin*d III and *Bam* HI digestions showed a ladder of bands corresponding to multimeric DNA fragments. The basic *Hin*d III and *Bam* HI repeats were cloned in pUC18. The recombinant clones were screened with labelled genomic DNA from *A. arenosa*. Several clones with highly repeated inserts were obtained and tested by Southern hybridization. Five *Bam* HI clones and one *Hin*d III clone were sequenced on an automated sequencer (Pharmacia) using the dideoxy chain termination procedure.

Probe labelling and Southern hybridization

The ECL chemiluminescent system (Amersham) was used for labelling probes pAa214 (*Bam* HI satellite sequence from *A. arenosa*) and pAL1 (*Hin*d III satellite from *A. thaliana* [30]) by random priming. In general, all steps were carried out according to the protocol supplied by the manufacturer. For Southern blots, 3 μg of genomic DNA were digested, fractionated on a 1.1% agarose gel and transferred onto nylon membrane using standard protocols. Hybridization with repetitive sequences was carried out in 5× SSC, 0.1% (w/v) SDS, 5% (w/v) dextran sulphate, 20-fold dilution of liquid block (supplied) and 100 μg/ml denatured shared heterologous DNA at 60 °C overnight. Filters were washed twice for 20 min in 2× SSC/0.1% SDS at the same temperature.

Fluorescent in situ *hybridization*

Flower buds were used for chromosome preparations. After a treatment with 2 mM 8-hydroxyquinoline for 1 h the buds were fixed in methanol-acetic acid (3:1). The chromosome preparation and the *in situ* hybridization followed techniques described for *A. thaliana* [28]. pAa214 and pAL1 were used as probes for *in situ* hybridization and labelled with digoxigenin-11-dUTP and biotin-11-dUTP using the polymerase chain reaction. The *in situ* hybridization was performed with 84% stringency calculated according to Meinkoth and Wahl [33]. The probes were detected with FITC (fluorescein isothiocyanate) conjugated sheep anti-digoxigenin antibody and streptavidin-CY3 (Sigma) conjugate, respectively. The chromosomes were counterstained with DAPI (4′,6-diamidino-2-phenylindole).

Results and discussion

Genomic organization and sequence analysis

After digestion of genomic DNA of *A. arenosa* with different enzymes and electrophoretic separation, a ladder of fragments was visible in the *Bam* HI digestion indicating that a fraction of the genome consists of tandemly repeated elements. For detailed characterization, the highly abundant fraction of *Bam* HI fragments was isolated and cloned. Five clones of the *Bam* HI family were choosen for sequence analysis. Hybridization with one of the clones, pAa214, to *Bam* HI-digested genomic DNA of *A. arenosa* resulted in a ladder of increasing fragments typical of tandemly arranged repeats and caused by divergence or methylation of the restriction site (Fig. 1).

The nucleotide sequence of five clones of the

Bam HI satellite family and their consensus sequence are shown in Fig. 2. The length of the repeats (166–179 bp) corresponds to other satellite sequences [39, 40] but it does not seem to be highly conserved in length, as reflected by the varying repeat length. Simoens *et al.* [42] reported a *Hin*d III satellite sequence in *A. thaliana* differing also in length from 177-198 bp because of the presence of the so-called 'super-*Hin*d III' structure. The G + C content of the *Bam* HI monomer pAa214 (40.2%) is similar to the average G + C content (41.4%) of the *A. thaliana* genome [25].

All pAa repeats analysed show divergence due to randomly distributed point mutations, insertions and deletions resulting in a heterogeneity between monomers in the range from 4.5% to 25%. This unusually high divergence reflects the dynamic and more rapid divergence of tandemly repeated DNA sequences in comparison to structural genes [9]. A result of the divergence process is the mutation of *Hin*d III sites by deletion and AT→transversion in two out of five analysed repeats. These specific changes were not found in the remaining monomers, showing that a subset of the *Bam* HI sequence family is characterized by the occurrence of intact *Hin*d III sites. After digestion of genomic DNA with *Hin*d III the expected ladder pattern was detected, demonstrating that the *Bam* HI family can also be revealed as a *Hin*d III satellite of the same length. The *Hin*d III monomer was cloned and sequenced to analyse how adjacent *Bam* HI repeats are organized. These adjacent repeat units were found to be directed in head-to-tail orientation to each other as expected in clusters of tandem arrays [43]. A search of five clones representing the *Bam* HI sequence family from *A. arenosa* against the GenBank/EMBL DNA sequence database (release 80) shows 58% to 80% homology to a sequence family isolated from *A. thaliana* [30, 42]. No other significant matches were found.

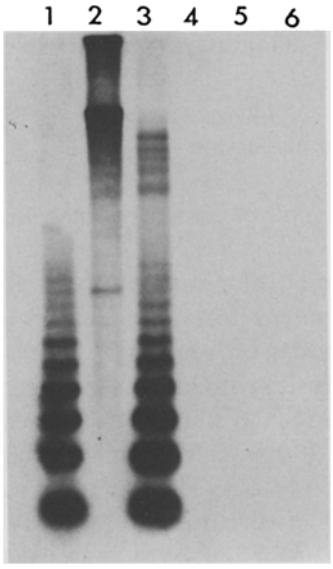


Fig. 1. Genomic organization of the *Bam* HI satellite DNA monomer in *A. arenosa* and distribution of this satellite sequence in different *Arabidopsis* species and *B. nigra*. DNA of *A. arenosa* (1), *A. thaliana* (2), *A. suecica* (3), *A. griffithiana* (4), *A. pumila* (5), *B. nigra* (6) was digested with *Bam* HI, electrophoresed and transferred onto a nylon membrane. The filter was hybridized with pAa214.

Chromosomal localization

In situ hybridization of repetitive DNA sequences has been used to study the organization of DNA sequences at the chromosomal level and demonstrated that satellite DNA may be clustered at heterochromatic regions at subtelomeric [4], centromeric [20] or more dispersed in intercalary chromosomal locations [3]. For physical mapping of the *A. arenosa* sequence, pAa214 was used as a probe for *in situ* hybridization (Fig. 3B). In a tetraploid line of *A. arenosa* ($2n = 4x = 32$), 32 sites of hybridization were detected in metaphase and most interphase nuclei. There was some variation in strength of hybridization; four weak sites and four less weak sites were detected reproducibly, while the remaining 24 strong sites showed limited variation. The differences probably resulted from unequal crossing-over that leads to the fixation and amplification of small sequence variants on different chromosomes [45]. In general, hybridization strength correlated with the brightness of the DAPI signal, although several major DAPI sites, presumably the condensed rDNA at the nucleolar organizing regions, were also observed. In metaphase chromosomes, hy-

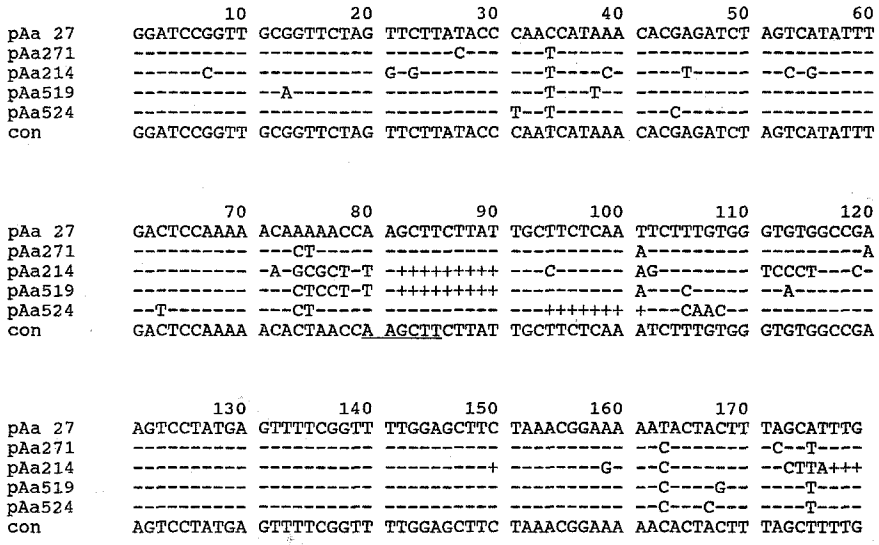


Fig. 2. Nucleotide sequence of five clones of the *Bam* HI monomer from *A. arenosa* and their consensus sequence. The *Hin*d III sites are underlined.

bridization to the two chromatids was clearly visible at the weaker pAa214 sites, while many of the major sites showed hybridization to both sides of the centromere with substantial hybridization along the length of the chromosome arms. In contrast, pAL1 was found to hybridize only at the centromeres of all five chromosome pairs of *A. thaliana* [28] and, moreover, this repetitive DNA family is thought to be present only near the centromeres in *A. thaliana* [34].

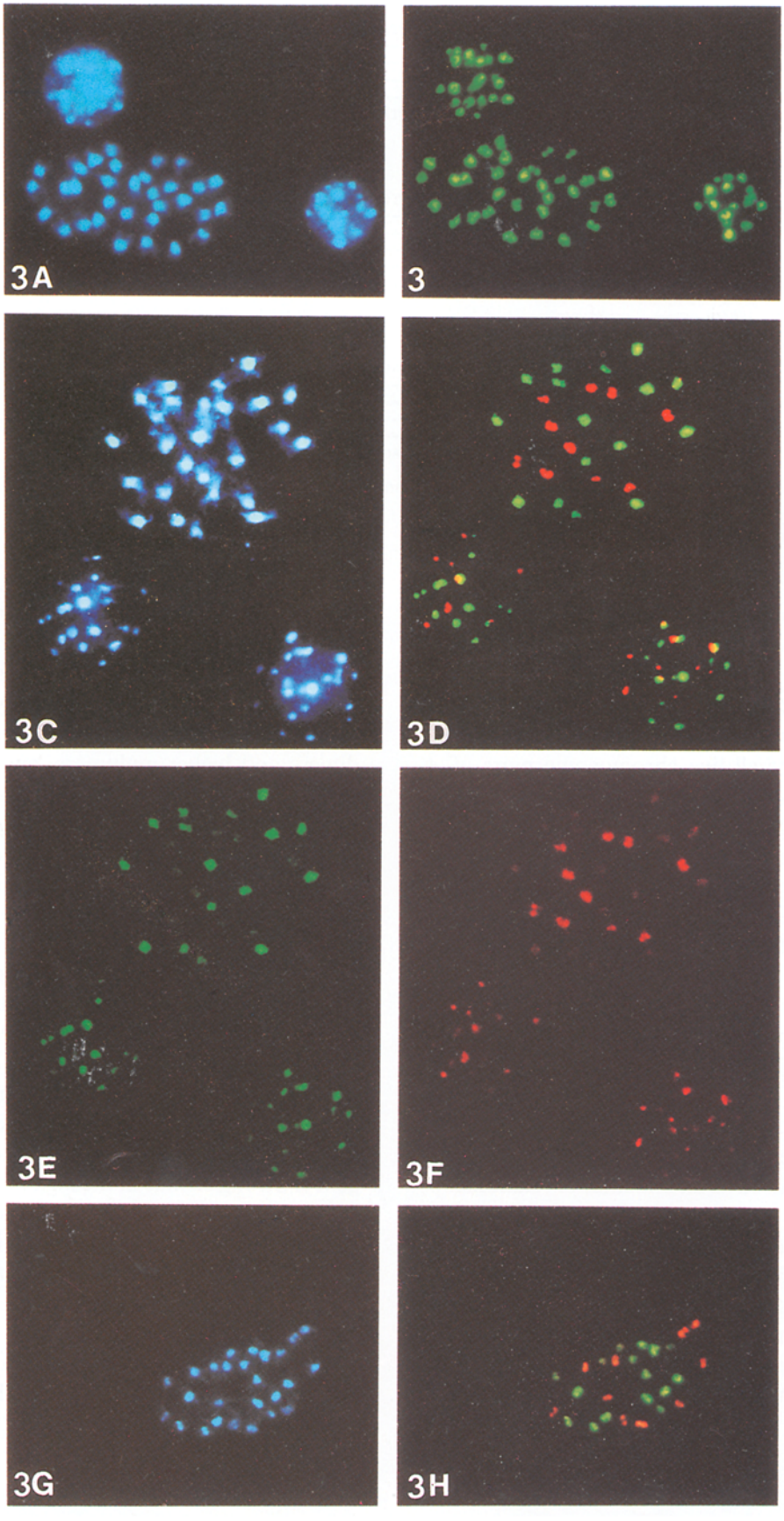
Distribution of the pAa214 family with the Cruciferae

Nuclear genomes of most plant species include satellite DNA sequence motifs, and there is often conservation of these motifs between different species [12, 15]. In some cases, the presence of tandemly repeated sequences from one species in others can be used as an indicator of relationships [2, 14].

To investigate the distribution of the *Bam* HI sequence family in related species, we used the monomer pAa214 to probe genomic Southern filters with *Bam* HI digested DNAs of *A. arenosa*, different *Arabidopsis* species and *Brassica nigra*. A ladder-like pattern was observed in *A. arenosa* and *A. suecica* (Fig. 1). Hybridization with pAa214 resulted in a similar pattern of monomers and multimers demonstrating that this sequence is present in clusters of tandem arrays in both species. However, the pattern of hybridization to *A. suecica* is more complex, showing monomers, dimers and higher multimers, and hybridization to fragments of higher molecular weight DNA, which can be reduced at higher stringency. This implies that the genome of *A. suecica* contains other DNA sequence families with different organization but homology to at least part of pAa214.

Hybridization to *A. thaliana* genomic DNA digests showed a very weak ladder-like pattern indicating that no sequence of the pAa214 family is present as a major class of tandem repeats. Other classes of fragments of the same size as in *A. suecica* were visible superimposed on a smear, which is weaker in hybridization at higher stringency. Strong signals in the high molecular weight range, indicating a relatively high copy number of this sequence, resulted from hybridization of pAa214 with DNA which lacks *Bam* HI sites. In addition, strong hybridization to a fragment of about 1.6 kb was detected.

These results suggest that sequence motifs present in the *A. arenosa Bam* HI family also


3A
3
3C
3D
3E
3F
3G
3H

occur in the genome of *A. thaliana* but in a different genomic organization and not as a *Bam* HI satellite family. These motifs exist in a relatively high copy number in the *A. thaliana* genome and have not been further characterized. The reduced cross-hybridization under conditions of higher stringency and the different genomic organization observed in the genome of *A. thaliana* confirm the substantial divergence between the *Bam* HI satellite monomer and satellite repeats of *A. thaliana* [30, 42], which is evident from the sequence. The satellite repeats pAa214 and pAL1 probably originate from a common ancestral sequence in the genome of a progenitor species of *A. arenosa* and *A. thaliana*.

No hybridization of pAa214 was detected to DNA from *A. pumila*, *A. griffithiana* and *B. nigra* suggesting that the pAa214 family, or repetitive DNA sequences including parts of it, is restricted to only some species of the genus *Arabidopsis sensu latu*, and indicates that the limits on some accepted genera in the Cruciferae are highly artificial. Although any generalization from data about a single sequence family must be qualified, it seems to us unlikely that a sequence would be present in two species, *A. thaliana* and *A. arenosa* (under the name *Cardaminopsis arenosa*), in different tribes but absent in other species in the genus *Arabidopsis*. Increasingly, molecular approaches are being used to examine taxonomic affinities, including chloroplast gene sequencing [10], ribosomal sequences [19], and total genomic DNA probes [36]. It is clear from work here and elsewhere [12, 38] that the sequencing and analysis of major repetitive DNA classes will also be useful in phylogenetic studies.

The hybrid species A. suecica

Two-colour *in situ* hybridization with pAL1 and pAa214, labelled with different fluorochromes, was carried out to examine the chromosomal distribution of the sequences (Fig. 3C-H). The stringency used for *in situ* hybridization was 84%, and was above the internal homology between pAL1 and pAa214. The repeat pAa214 hybridized preferentially and strongly to 16 chromosomes, presumably those originating from *A. arenosa* (Fig. 3E). The signals were centromeric, with extension along the chromosome arms, and similar to the *in situ* hybridization to *A. arenosa* chromosomes (Fig. 3B). Weaker hybridization was also observed in the centromeric region only of the remaining 10 chromosomes from *A. thaliana* with labelling of both chromatids. pAL1 showed a reciprocal hybridization pattern with 10 major centromeric sites and 16 very minor sites, also at centromeric positions (Fig. 3F). The strong signals on 10 major sites (corresponding to 10 *A. thaliana* chromosomes) confirm the results of Maluszynska and Heslop-Harrison [28]. The results support the view that *A. suecica* is a hybrid of *A. arenosa* and *A. thaliana*. Within the *Brassica* genus, we have examined satellite sequences with paracentromeric hybridization patterns and sites on only a few chromosomes in diploid species and the natural tetraploid derivatives. The tetraploids show different numbers of sites from the diploids [16], unlike *A. suecica* where the number of sites is the same as in the parents. The different behaviour may reflect differences in the behaviour of satellite DNA families, in age of the hybrid, or in meiotic pairing and recombination.

Fig. 3. Localization of centromeric satellite sequences on chromosomes of *A. arenosa* and *A. suecica*. A. DAPI staining of metaphase chromosomes and interphase nuclei of *A. arenosa* ($2n = 4x = 32$). B. The same metaphase and interphase nuclei after *in situ* hybridization with pAa214 (yellow-green signal). C. DAPI staining of prometaphase chromosomes and interphase nuclei of *A. suecica* ($2n = 26$). D. Simultaneous localization of pAa214 (digoxigenin-labelled) and pAL1 (biotin-labelled) in the same prometaphase and interphase nuclei. Exposure with a multiband passfilter shows 16 pAa214 sites (yellow-green fluorescence) and 10 pAL1 sites (red fluorescence). E. and F. The same prometaphase and interphase nuclei after *in situ* hybridization with pAa214 (yellow-green signal, E) and with pAL1 (red signal, F). G and H. DAPI staining of metaphase chromosomes of *A. suecica* (G) and simultaneous localization of pAa214 yellow-green signal) and pAL1 (red signal) (H).

Variants of satellite repeats (diverged subfamilies or subsets) are useful probes for the analysis of hybrid genomes. These probes may be a valuable complement to genomic *in situ* hybridization and enable differentiation even between closely related genomes under appropriate conditions.

Previously, it was shown that processes can lead to changes in the structural organization of repeated DNA sequences in sexual hybrids [22, 26]. Because of the conserved hybridization pattern of the *A. arenosa Bam* HI repeat (pAa214) in *A. arenosa*, *A. thaliana* and *A. suecica* (as described above, Fig. 1, lane 1, 2, 3) it is unlikely that this satellite family was reorganized in the hybrid genome. In addition, the *A. thaliana* repeat, pAL1, was hybridized to *Hin*d III-digested DNA of *A. arenosa*, *A. thaliana* and *A. suecica* (Fig. 4). The tandem arrangement of pAL1 in the genome of *A. thaliana*, as described by Martinez-Zapater *et al.* [30] was visible. A very weak ladder of hybridization was observed after probing pAL1 onto digested DNA from *A. arenosa*. The weak hybridization presumably reflects the high divergence of the majority of the pAa214 family repeats from the pAL1 family in *A. thaliana*. In digests of *A. suecica* DNA, a ladder of fragments in the low-molecular-weight ranges was detected, presumably from the *A. thaliana* origin DNA in the genome of the hybrid. The signal was weaker than that on the *A. thaliana* lane, reflecting the smaller amount of *A. thaliana* origin DNA in the equally loaded lanes.

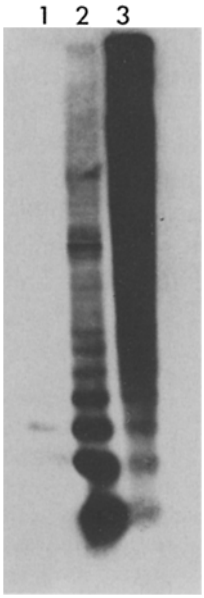


Fig. 4. Southern hybridization of pAL1 to *Hin*d III-digested DNA of *A. arenosa* (1), *A. thaliana* (2) and *A. suecica* (3).

pAL1, when probed to digests of *A. suecica* DNA, showed a strong and striking pattern of hybridization with a smear of fragments, thus indicating a considerable change in the organization of the pAL1 family from that found in the presumed parent *A. thaliana*. It might be that the *A. thaliana* satellite was rearranged leading to an increase in copy number and loss of most of the *Hin*d III sites. There is increasing evidence that rapid sequence rearrangement can occur in hybrids: McClintock [32] first suggested that remote hybridization is a form of stress, which places the parental genomes in an unusual genotypical environment. In cereals, changes in the copy number and genomic organization of tandem repeats have been shown in wheat-rye hybrids and backcrosses, where heterochromatin may be lost rapidly from the rye chromosomes [31, 41]. In hybrids between barley species, Vershinin *et al.* [46] showed that structural rearrangements of some highly repetitive DNA sequences may occur in hybrids only a few years old. However, the very high proportion of repetitive DNA in cereals makes analysis of even a small proportion difficult.

The detailed examination of genomic reorganization in natural hybrids of plants with small genomes, such as *A. suecica*, and in artificial hybrids which can be made between *Arabidopsis* species [23] and even *A. thaliana* and *Brassica* [35] may provide a useful route to study function and evolution of different types of repetitive DNA sequences.

Acknowledgements

A. K. was supported by grant PGIII/569 from BBSRC (former AFRC). I.G. acknowledges financial support from CNR. T.S. has been sup-

ported by a grant from Deutscher Akademischer Austauschdienst. We thank G. Harrison and Dr R. Doudrick for technical support in DNA sequencing, Professor N. Jacobsen for the valuable discussion about taxonomy of *Arabidopsis* and Professor J. Maluszynska for ongoing collaboration on molecular cytogenetics and provision of seeds.

References

1. Al-Shehbaz I: The genera of Sisymbrideae (Cruciferae, Brassicaceae) in the Southeastern United States. J Arnold Arboretum 69: 213–237 (1988).
2. Anamthawat-Jonsson K, Heslop-Harrison JS: Species-specific DNA sequences in the Triticeae. Hereditas 116: 49–54 (1992).
3. Anamthawat-Jonsson K, Heslop-Harrison JS: Isolation and characterization of genome-specific DNA sequences in Triticeae species. Mol Gen Genet 240: 151–158 (1993).
4. Bedbrook JR, Jones J, O'Dell M, Thomson RD, Flavell RB: A molecular description of telomeric heterochromatin in *Secale* species. Cell 19: 545–560 (1980).
5. Bennett MD, Smith JB: Nuclear DNA amounts in angiosperms. Phil Trans R Soc Lond B 274: 227–274 (1976).
6. Bennett MD, Smith JB: Nuclear DNA amounts in angiosperms. Phil Trans R Soc Lond B 334: 309–345 (1991).
7. Benslimane AA, Dron M, Hartmann C, Rode A: Small tandemly repeated DNA sequences of higher plants likely originate from a tRNA gene ancestor. Nucl Acids Res 14: 8111–8119 (1986).
8. Capesius I: Sequence of the cryptic satellite DNA from the plant *Sinapis alba*. Biochim Biophys Acta 739: 276–280 (1983).
9. Cavalier-Smith T: Eukaryote gene numbers, noncoding DNA and genome size. In: Cavalier-Smith T (ed) The Evolution of Genome Size, p. 112–115. John Wiley, Chichester (1985).
10. Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH and 42 others: Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene rbcL1. Ann Missouri Bot Gard 8: 528–580 (1993).
11. Dellaporta SL, Wood J, Hicks JB: A plant DNA minipreparation: version 2. Plant Mol Biol Rep 1: 19–22 (1983).
12. Grellet F, Delcasso D, Panabieres F, Delseny M: Organization and evolution of a higher plant alphoid-like satellite DNA sequence. J Mol Biol 187: 495–507 (1986).
13. Grellet F, Tremousaygue D, Delseny M: Isolation and characterization of a tandemly repeated DNA sequence from the crucifer *Sisymbrium irio*. Indian J Biochem Biophys 25: 483–487 (1988).
14. Gupta PK, Fedak G, Molnar SJ, Wheatcroft R: Distribution of a *Secale cereale* DNA repeat sequence among 25 *Hordeum* species. Genome 32: 383–388 (1989).
15. Hallden C, Bryngelsson T, Säll T, Gustafsson M: Distribution and evolution of a repeated DNA sequence in the family Brassicaceae. J Mol Evol 25: 318–323 (1987).
16. Harrison GE, Heslop-Harrison JS: Centromeric repetitive DNA in the genus *Brassica*. Theor Appl Genet 90: 157–165 (1994).
17. Heslop-Harrison JS, Maluszynska J: The molecular cytogenetics of *Arabidopsis*. In: Meyerowitz E, Somerville CJ (eds) *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, in press (1994).
18. Heslop-Harrison JS, Schwarzacher T: The ultrastructure of *Arabidopsis thaliana* chromosomes. In: Schweizer D, Peuker K, Loidl J (eds) Fourth International Conference on *Arabidopsis* Research, Vienna, June 2–5, Abstracts, p. 3 (1990).
19. Hori H, Lim BL, Osawa S: Evolution of green plants as deduced from 5S rRNA sequences. Proc Natl Acad Sci USA 82: 820–823 (1985).
20. Iwabuchi M, Itoh K, Shimamoto K: Molecular and cytological characterization of repetitive DNA sequences in *Brassica*. Theor Appl Genet 82: 349–355 (1991).
21. Koornneef M, van Eden J, Hanhart CJ, Braaksma FJ, Feenstra WJ: Linkage map of *Arabidopsis thaliana*. J Heredity 74: 265–272 (1983).
22. Kovtun YV, Korostash MA, Butsko YV, Gleba YY: Amplification of repetitive DNA from *Nicotiana plumbaginifolia* in asymmetric somatic hybrids between *Nicotiana sylvestris* and *Nicotiana plumbaginifolia*. Theor Appl Genet 8: 221–228 (1993).
23. Kranz AR, Kirchheim B: Genetic resources in *Arabidopsis*. Arabidopsis Inf Serv 24: 3.4.1–3.4.4 (1987).
24. Lakshmikumaran M, Ranade SA: Isolation and characterization of a highly repetitive DNA of *Brassica campestris*. Plant Mol Biol 14: 447–448 (1990).
25. Leutwiler LS, Hough-Evens BR, Meyerowitz EM: The DNA of *Arabidopsis thaliana*. Mol Gen Genet 194: 15–23 (1984).
26. Lin LS, Ho THD, Harlan JR: Rapid amplification and fixation of new restriction sites in the ribosomal DNA repeats in the derivatives of a cross between maize and *Trisacum dactyloides*. Devel Genet 6: 101–112 (1985).
27. Linnaeus C: Species Plantarum (1753).
28. Maluszynska J, Heslop-Harrison JS: Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. Plant J 1: 159–166 (1991).
29. Maluszynska J, Heslop-Harrison JS: Molecular cytogenetics of the genus *Arabidopsis*: *in situ* localization of rDNA sites, chromosome numbers and diversity in centromeric heterochromatin. Ann Bot 71: 479–484 (1993).
30. Martinez-Zapater JM, Estelle MA, Somerville CR: A highly repeated DNA sequence in *Arabidopsis thaliana*. Mol Gen Genet 204: 417–423 (1986).
31. May CE, Appels R: Rye chromosome translocations in hexaploid wheat: a re-evaluation of the loss of hetero-

chromatin from rye chromosomes. Theor Appl Genet 56: 17–23 (1980).
32. McClintock B: Mechanisms that rapidly reorganize the genome. Stadler Genet Sym 10: 25–48 (1978).
33. Meinkoth J, Wahl G: Hybridization of nucleic acids immobilized on solid support. Anal Biochem 138: 267–284 (1984).
34. Meyerowitz EM, Pruitt RE: *Arabidopsis thaliana* and plant molecular genetics. Science 229: 1214–1218 (1985).
35. Nagl W: 'Arabidobrassica': evidence for intergenic somatic hybrid nature from electron microscopic morphometry of chromatin. Eur J Cell Biol 21: 227–228 (1980).
36. Orgaard M, Heslop-Harrison JS: Relationships between species of *Leymus*, *Psathyrostachys*, and *Hordeum* (Poaceae, Trificeae) inferred from Southern hybridization of genomic and cloned DNA probes. Plant Syst Evol 189: 217–231 (1994).
37. Redei GP: A heuristic glance at the past of *Arabidopsis* genetics. In: Koncz C, Chua N-H, Schell J, (eds) Methods in *Arabidopsis* Research, pp. 1–15. World Scientific, Singapore (1992).
38. Schmidt T, Heslop-Harrison JS: Variability and evolution of highly repeated DNA sequences in the genus *Beta*. Genome 36: 1074–1079 (1993).
39. Schmidt T, Metzlaff M: Cloning and characterization of a *Beta vulgaris* satellite DNA family. Gene 101: 247–250 (1991).
40. Schweizer G, Ganal M, Ninnemann H, Hemleben V: Species-specific DNA sequences for identification of somatic hybrids between *Lycopersicon esculentum* and *Solanum acaule*. Theor Appl Genet 75: 679–684 (1988).
41. Seal AG, Bennett MD: The rye genome in winter hexaploid triticales: chromosome substitutions. Can J Cytol 23: 647–653 (1981).
42. Simoens CR, Gielen J, Van Montagu M, Inzé D: Characterization of highly repetitive sequences of *Arabidopsis thaliana*. Nucl Acids Res 16: 6753–6766 (1988).
43. Smith G: Evolution of repeated sequences by unequal crossover. Science 191: 528–535 (1975).
44. Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA: Flora Europaea vol. 1, 2nd ed., p. 352 (1993).
45. Willard HF, Waye JS: Hierarchical order in chromosome-specific alpha satellite DNA. Trends Genet 3: 192–198 (1987).
46. Vershinin AV, Salina A, Svitashev SK: Is there a connection between genomic changes and wide hybridization? Hereditas 116: 213–217 (1992).