Theor Appl Genet (1995) 90:312–316

K. Irifune · K. Hirai · J. Zheng · R. Tanaka
H. Morikawa

Nucleotide sequence of a highly repeated DNA sequence and its chromosomal localization in *Allium fistulosum*

Received: 5 July 1994 / Accepted: 18 August 1994

Abstract A highly repeated DNA sequence with a repeating unit of approximately 380 bp was found in *Eco*RV digests of the total genomic DNA of *Allium fistulosum*. Three independent clones containing this unit were isolated, and their repeating units sequenced. These units showed more than 94% sequence homology, and the copy number was estimated to be about 2.8×10^6 per haploid genome. In situ hybridization, with the repeating unit as a probe, and C-banding analyses indicated that the repeated DNA sequence of *A. fistulosum* is closely associated with the major C-heterochromatin in the terminal regions of all 16 chromosomes at mitotic metaphase. The characters of the repeating unit are similar to those of the *A. cepa* unit, which is taxonomically closely related to *A. fistulosum*.

Key words *Allium fistulosum* · Repeated DNA sequence · In situ hybridization · Nucleotide sequence · C-banding

Introduction

Eukaryotic genomes, including those of higher plants, generally contain highly repeated DNA sequences. The study of these sequences is important in relation to chromosome organization because many previously reported lines of evidence (Appels et al. 1978; Yakura et al. 1987; Bedbrook et al. 1980; Deumling 1981; Peacock et al. 1981; Iwabuchi et al. 1991; Wu et al. 1991) show that these sequences are closely associated with constitutive heterochromatin.

Communicated by K. Tsunewaki

K. Irifune · K. Hirai[1] · J. Zheng[2] · R. Tanaka · H. Morikawa (✉)
Graduate Department of Gene Science, Faculty of Science, Hiroshima University, Higashi-Hiroshima, 724, Japan

Present addresses:
[1] Shobara Kakuchi High School, Shobara, Hiroshima 727, Japan
[2] Department Biology, Nankai University, Tianjin 300071, China

Plants of the genus *Allium*, including *A. fistulosum*, are good material for sequence studies because they have large chromosomes and heterochromatin blocks. Only a single study of the nucleotide sequence of repeated DNA sequences in *Allium cepa* (Barnes et al. 1985), however, has been published. We here report on the highly repeated DNA sequence of *A. fistulosum* which has a repeating unit of approximately 380 bp and is closely associated with the C-heterochromatin located in the terminal regions of the chromosomes.

Materials and methods

Plant material, chromosome preparations and C-banding

Allium fistulosum cv 'Kujho' seeds obtained from a commercial source were soaked in water, then placed on a moist filter paper in Petri dishes and incubated at 20 °C. After 3–5 days, the root tips that had formed were excised and used for chromosome preparation according to the method of Hizume et al. (1980). C-banding was done as described in Tanaka and Taniguchi (1975).

Extraction of total DNA

Five grams of 3–5-day-old seedlings of *A. fistulosum* were used for DNA extraction by the method of Peacock et al. (1981).

Restriction endonuclease digestion and gel electrophoresis

A 10-µl reaction mixture containing 3.6 µg of total DNA, 10–15 units of restriction enzyme [*Eco*RI, *Eco*RV, *Bam*HI, *Sph*I, *Sal*I, *Sca*I, *Pst*I, *Pvu*II, or *Hin*dIII (Takara Shuzo Co., Ltd.)], in the appropriate reaction buffer recommended by the distributor, was incubated at 37 °C for 4 h, the reaction being terminated by the addition of 2 µl of 0.25% bromophenol blue-50% glycerol solution. The DNA digests were electrophoresed in a 1.2% agarose gel in TBE running buffer (45 mM Tris-borate, 1 mM EDTA pH 8.0) for 5 h at 4 V/cm.

Cloning of repeated DNA sequences

After fractionation of the *Eco*RV digest of *A. fistulosum* total DNA, fragments of approximately 380 bp were extracted from the agarose gel using Geneclean II (Bio 101, Inc.). They were then ligated into the

*Sma*I site of the polylinker region of pUC119 plasmid using T4 ligase (Takara Shuzo Co. Ltd), after which they were transformed into *E. coli* strain MV1184 or DH5α. Recombinant clones were selected as white colonies growing on ampicillin plates containing 5-bromo–4–chloro–3–indolyl–β–galactopyranoside (X-gal). Three recombinant plasmids (pAfi100, pAfi105, pAfi106) with 380-bp fragments were obtained.

Dideoxy sequencing

Double-stranded DNAs of the pAfi 100, pAfi 105 and pAfi 106 plasmids were sequenced with ^{32}P-deoxycytidine 5′-triphosphate (3 000 Ci/mmol; Du Pont Inc.) using the dideoxynucleotide chain-terminator method (Sanger et al. 1977). Sequence reactions were carried out according to the instructions for Sequenase version 2.0 (USB Co.). Reaction products were fractionated by 6% polyacrylamide gel electrophoresis and made visible by autoradiography. DNA sequences were analyzed using the Genetyx program (Software Development).

Southern hybridization

DNAs digested with the nine restriction enzymes were electrophoresed in a 1.2% agarose gel, then stained with ethidium bromide, after which they were photographed and blotted onto Gene Screen Plus nylon membranes (Du Pont, Inc.) for hybridization analysis as described by Southern (1975). In the case of the probe, the 380-bp insert DNA was extracted from pAfi 100 by double digestion with *Hin*dIII and *Eco*RI then labelled with ^{32}P-dCTP (3 000 Ci/mmol; Du Pont Inc.) by the random primer extension method using a DNA labelling kit (Nippon Gene Co.).

After hybridization, the membrane was washed three times in 2 × SSC, 0.5% SDS at 65 °C for 5 min; twice in 2 × SSC, 0.1% SDS at 65 °C for 15 min; and once in 0.1 × SSC, 0.5% SDS at 65 °C for 60 min, after which it was exposed to X-ray film (Fuji XR) at − 80 °C for about 2 days.

Copy-number estimation

A 100-ng sample of the genomic DNA digested with *Eco*RV, and a 2.4-ng portion of the 380-bp insert DNA that had been cut from pAfi-100, were electrophoresed in the same agarose gel. The DNA from this gel was transferred to a nylon membrane for hybridization analysis. The conditions for hybridization, washing, and the preparation of the radioactive probe, were as described above. Immediately after washing the blotted membrane, the radioactivity of the hybridized DNA sequence was measured directly with an Ambis radioanalytic imaging system (Ambis systems, Co.). The copy number of the repeated DNA sequence was estimated from a comparison of the relative radioactivity, within the linear range of cpm/ng DNA, of the genomic DNA and the insert DNA from pAfi100.

In situ hybridization

Forty microliters of hybridization solution, containing 150 ng of biotin-labelled probe DNA, 10% dextran sulphate in 2 × SSC, and 0.1 M sodium phosphate (pH 5.0), were dropped onto a slide and covered with a coverslip. The prepared slides were incubated for 15 min at 78 °C to denature both the probe and the target DNA, after which they were incubated in a humid chamber for 8–12 h at 42 °C for hybridization. They were then washed three times in a 2 × SSC solution at room temperature, and soaked twice in TBS solution (150 mM NaCl, 100 mM Tris-HCl pH 7.5) for 15 min at room temperature. Linearized pAfi100 DNA labelled with biotin-14-dATP, using a nick translation kit (BRL Inc.), served as the probe.

Hybridization signals were detected using the streptavidin-alkaline phosphatase conjugate with nitroblue tetrazolium-4-bromo-5-chloro-3-indolylphosphate (NBT-BCIP) as the substrate (BRL Inc.). Detection was as described by Iwabuchi et al. (1991).

Results and discussion

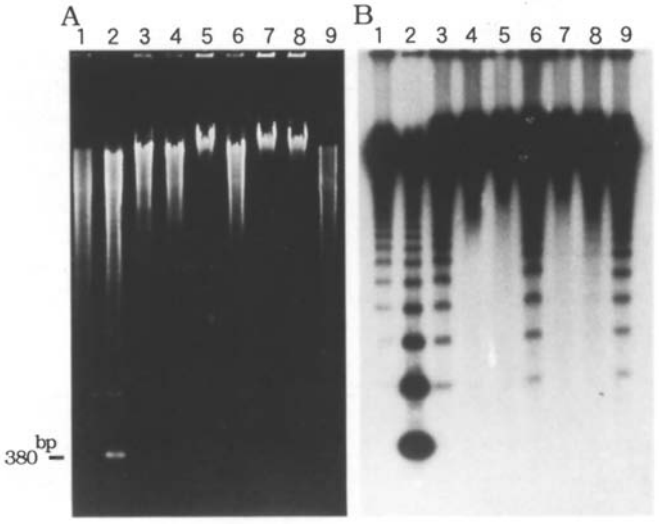
Fractionation and identification of highly repeated DNA fragments

Figure 1 A shows total *A. fistulosum* DNA that had been digested with *Eco*RI, *Eco*RV, *Bam*HI, *Sph*I, *Sal*I, *Sca*I, *Pst*I, *Pvu*II and *Hin*dIII, then electrophoresed in an agarose gel and stained with ethidium bromide. Bands were detected at about 380 bp and, although less clear, at 760 bp, in the *Eco*RV digest, indicative that the *A. fistulosum* genome has a repeated DNA sequence with a repeating unit of 380 base pairs. Interestingly, its size is very close to that of the *Bam*HI repeating unit (375 bp) of the repeated DNA sequence reported in *A. cepa* (Barnes et al. 1985). No clear bands are present in the digests made with the any of the other restriction enzymes (see Fig. 1 A).

The DNA of the 380-bp band from the gel shown in Fig. 1 A was recovered and cloned into *Sma*I-digested pUC119. Of the three independent recombinant clones, pAfi100, pAfi105 and pAfi106, a fragment of genomic DNA with the DNA of the 380-bp band was excised from clone pAfi100 by digestion with *Hin*dIII and *Eco*RI. Using this fragment as a probe, we carried out a Southern-blot analysis of the total *A. fistulosum* DNA that had been digested with the nine restriction enzymes. The results are shown in Fig. 1B.

Typical ladder patterns were obtained for the digests with *Eco*RV, *Eco*RI, *Bam*HI, *Sca*I, *Hin*dIII and *Pvu*II,

Fig. 1 Electrophoretic fractionation (**A**) and Southern-blot analysis (**B**) of *A. fistulosum* total DNA. **A** Total DNA (3.6 μg) was digested with *Eco*RI (*lane 1*), *Eco*RV (*lane 2*), *Bam*HI (*lane 3*), *Sph*I (*lane 4*), *Sal*I (*lane 5*), *Sca*I (*lane 6*), *Pst*I (*lane 7*), *Pvu*II (*lane 8*) and *Hin*dIII (*lane 9*). The digests were electrophoresed in a 1.2% agarose gel and then stained with ethidium bromide. Note the clear bands at approximately 380 and 760 bp only in the *Eco*RV digest. **B** The fractionated DNAs in the gel shown in **A** were transferred to a nylon membrane and subjected to Southern hybridization. A genomic DNA fragment (*Hin*dIII/*Eco*RI fragment of clone pAfi100) that corresponded to the 380-bp band in panel **A** was the probe employed.



but not for the digests with *Sph*I, *Sal*I and *Pst*I. The molecular size of each of the ladder bands roughly corresponded to multiples of 380 bp, the strongest band in the *Eco*RV digest. This agrees with the hypothesis that the *A. fistulosum* genome has a repetitive DNA sequence whose repeating unit is approximately 380 bp.

In the *Eco*RV digest, two additional and much fainter bands were present at approximately 620 and 530 bp (Fig. 1 B). Bands were also detected at approximately 290 and 240 bp when electrophoresis was done under conditions for low-molecular-weight separation (data not shown). This suggests that the *A. fistulosum* genome has minor repeated DNA sequences that have repeating units of less than 380 bp.

Hybridization signals were seen in the high-molecular-weight portion of each digest (see Fig. 1 B). These bands did not disappear even when the total DNA had been extensively digested with the restriction enzymes (data not shown). These signals are probably not caused by partial digestion of the total DNA but by sequence alteration in the repetitive DNA sequences (Singer 1982; Beridze 1986; and see below).

Quantification of the repeated DNA sequence

To determine the number of copies of the 380-bp repeat unit in the genome of *A. fistulosum*, we made a quantitative analysis of the hybridization bands in the Southern blot of the *Eco*RV digest using the Ambis radioanalytic imaging system. As shown in Fig. 2, the relative radioactivity detected in the 380-bp band of the *Eco*RV digest of 100 ng of total DNA (Fig. 2 A) was almost the same as that of the 2.4 ng of the 380-bp fragment isolated from pAfi100 that had been blotted on the same membrane (Fig. 2 B), evidence that the 380-bp fragments make up approximately 2.4% of the total DNA of this cultivar of leek. The copy number of the 380-bp fragment in the haploid genome excised by *Eco*RV was calculated to be 1.5×10^6, based on a haploid genome size of 26.3 pg for *A. fistulosum* (Jones and Rees 1968).

The radioactivity of the 380-bp band was estimated to be about 53% of the total radioactivity of all the bands in the *Eco*RV digest (see Fig. 2 A). The total copy number of 380-bp repeating units was therefore about 2.8×10^6 per haploid genome, which corresponds to approximately 4.5% of the total haploid DNA. On the basis of the results of slot-blot hybridization analysis, the content of the repeated DNA sequence was estimated to be about 4% (data not shown). This value is very close to that of the repeated DNA sequence (4%, 375 bp long) reported for *A. cepa* (Barnes et al. 1985).

Nucleotide sequences of the three repeated DNA sequences

Figure 3 shows the nucleotide sequences of the 380-bp repeat units cut from the independent clones, pAfi100, pAfi105 and pAfi106. The numbers of the nucleotide base pairs of the respective units are 378, 376 and 379. These sequences are not identical, but the homology between any of their two pairings is approximately 94%.

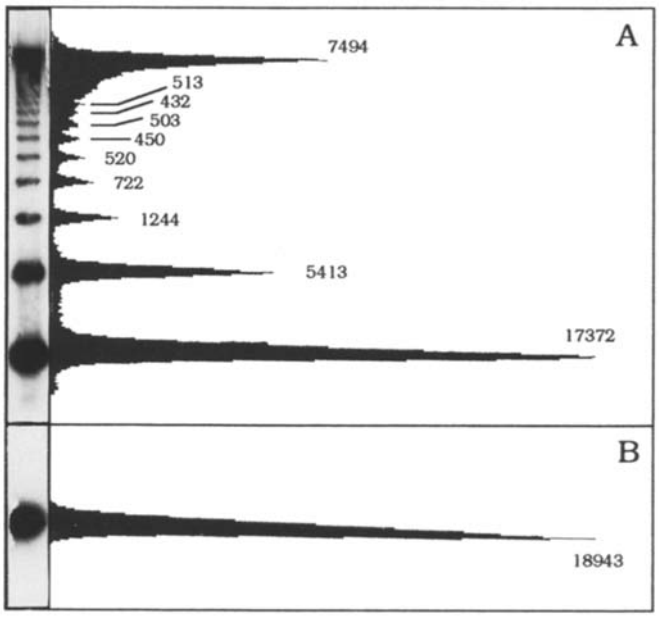


Fig. 2A,B Quantitative Southern-blot analysis of the *Eco*RV digest of *A. fistulosum* total DNA. A 100-ng sample of total DNA was digested with *Eco*RV. The digest and 2.4 ng of the *Hin*dIII/*Eco*RI fragment of clone pAfi100, which has a 380-bp fragment of genomic DNA, were subjected to Southern-blot analysis. The Southern bands are shown to the left of panels **A** and **B**. The relative radioactivities of the bands [ten bands in the *Eco*RV digest (**A**) and one in the genomic fragment (**B**)] were determined with an Ambis radioanalytic imaging system. Their values are shown to the right of the peaks in panels **A** and **B**

There are base substitutions (e.g., at positions 5, 34 and 53) and small insertions/deletions (at positions between 241 and 248) (see Fig. 3). The repeating unit contains 46.2% $G + C$, based on the three averages, and eight AT clusters (underlined in Fig. 3). A 5′-GGTGCA-3′ sequence is repeated five times tandemly in the first 120 bp of the repeating unit (arrows in Fig. 3). No recognition sites for *Eco*RI, *Bam*HI, *Hin*dIII, *Pvu*II or *Sca*I are present in any of the sequences of the three clones. This agrees with the results of the hybridization analysis of the digests, showing that these five enzymes do not have the 380-bp monomer band (see Fig. 1 B).

The 5′-GATAGC-3′ sequence at position 235 may be a potential *Eco*RV restriction site because substitution of a single base (from *G* to *T*) at this position forms a new *Eco*RV site, yielding two bands of 235 and 145 bp on digestion with the enzyme. This may account for the presence of minor repetitive DNA sequences that have repeating units of less than 380 bp derived from the 380-bp unit in the genome of *A. fistulosum*.

Our *Eco*RV repeating unit had 82% homology with the 375-bp *Bam*HI repeating unit in the genomic DNA of *A. cepa* reported by Barnes et al. (1985). The presence, though to different extents, of AT clusters and 5′-

Fig. 3 Nucleotide sequences of three repeating units from *Eco*RV-digested *Allium* DNA. *Asterisks* indicate matching nucleotides in the three repeats. Eight AT clusters are *underlined*. *Arrows* indicate five 5′-GGTGCA-3′ internal repeated sequences

```
                                                                      60
pAfi100   ATCCGCAGGG TGCAACATCT GCGGTGCAAG GTGCAACATT CGTCCAAAAA GACGAAACAG
pAfi105   *****A**** ********** ********** ***G****** ********** **********
pAfi106   *****A**** ********** ********** ********** ********** **T*******
                                                                     120
pAfi100   CGCCGTGCGA CTTTCAGAAT TGGGTGCATG TATTTGAAAT TGACCGGTAC CATGGTGAAC
pAfi105   ********** ****T****A ********G* ********** G********* *******C**
pAfi106   ********** ****T****A ********G* G********* C****A**** *******C**
                                                                     180
pAfi100   CACTCTCCAG GGATGGTAAA TCAAAGACCA GCCACGGTGT TCCCTTTTTT ACGTAAAAAT
pAfi105   ***CG***** ********** **GT****** *******C** ********** **********
pAfi106   ********** ********** **G******* **T******* ***A*A**** **********
                                                                     240
pAfi100   TCATGTGATG GAAAAACGAA GGCCAAAAAA CTTCAAAGAC CCGATTGTGG TTCGGATAGC
pAfi105   ********C* ********** ********** ********** *******A** **********
pAfi106   **G*****C* *G******** ********** ********** *******A** **********
                                                                     300
pAfi100   --TTTTTTGG GTATGAAACT ATCCGTTGAC GGCTTACAGC GGTCAATGCC GTAGAAATCA
pAfi105   TT******** ********** ********** ********** ******A*** ******C***
pAfi106   TTT******** ********** ********** ********** ******A*** **********
                                                                     360
pAfi100   CTGAGACCTT CGTTTTGACT GGTTATGGCT CCCGTAACTC TAAACGAGTC AACTGTTATG
pAfi105   ********** ********** ********** ********** ********G* ***A******
pAfi106   ********** *A******** ********** ********** ********G* ***A******
                                    378
pAfi100   GCCGTCCGAA GAATCGAT
pAfi105   ********** ********
pAfi106   ********** ********
```

GGTGCA-3′ sequence repeats in the 375-bp repeating unit of *A. cepa* has been noted by Barnes et al. (1985). The results of a homology search made at the EMBL data bank show that our *Eco*RV repeating unit has no appreciable homology with any other reported repeated DNA sequences.

Digestion of *A. fistulosum* total DNA by *Bam*HI produced a ladder pattern similar to that of the DNA digested with *Eco*RV, but no clear band was detected at 380 bp in the digest stained with ethidium bromide (see Fig. 1 A, B). The genomes of both *A. fistulosum* and *A. cepa* probably have a repetitive DNA sequence composed of repeating units similar in sequence and length but different in respect of their major restriction sites.

Distribution of the repeated DNA sequence on interphase and metaphase chromosomes

We carried out in situ hybridization with the linearized pAfi100 plasmid, labelled with nonradioactive biotin-dATP, as a probe, and made C-banding analyses of the interphase and metaphase chromosomes to investigate the distribution of the repeated DNA sequence, and more especially the relation between the sequence and the occurrence of C-band heterochromatin in the *A. fistulosum* chromosomes. The results are shown in Fig. 4.

At interphase, hybridization signals appeared as reddish dots (Fig. 4 A) on large condensed chromatin blocks (20–30 per nucleus). In some daughter nuclei, in what corresponded to the G_1 phase, the signals were clustered on one side of the nucleus (data not shown). A similar C-heterochromatin distribution appeared in the C-banded interphase nucleus (brown dots in Fig. 4 C). At metaphase, hybridization signals were present on the distal ends of all 16 *A. fistulosum* chromosomes and on the two satellites (Fig. 4 B). Similarly, large bands of C-heterochromatin were present in the terminal regions

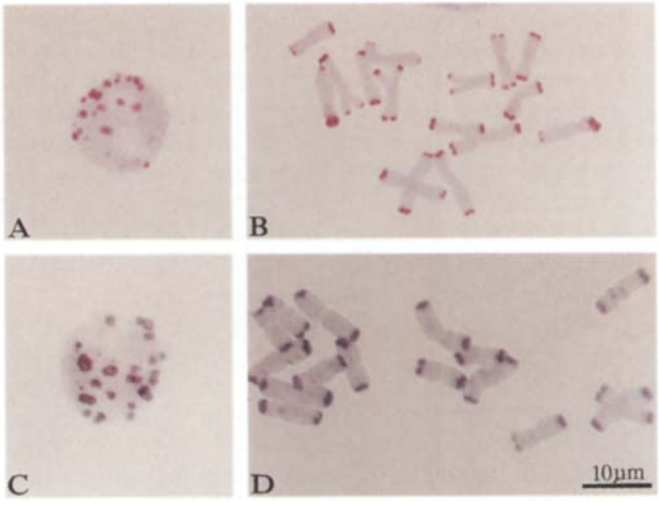


Fig. 4A–D In situ hybridization and C-banding for *A. fistulosum* chromosomes at interphase and metaphase. Clone pAfi100 labelled with nonradioactive biotin-dATP was the probe for in situ hybridization (**A** and **B**). C-banding (**C** and **D**) was done as described in Materials and methods. Note the many dots in the nucleus at interphase (**A** and **C**) and the distinct bands in the chromosomes at metaphase (**B** and **D**)

of all the chromosomes and the two satellites. In contrast, small bands or dots of C-heterochromatin were present in the centromere region in all the chromosomes, and in the interstitial region in some of them (Fig. 4 D), but there were no hybridization signals in these regions (Fig. 4 B).

The results of the in situ hybridization and C-banding analyses indicate that the 380-bp repeated DNA sequence clusters in the terminal region of the chromosomes and is associated with major heterochromatin blocks (but probably not with the minor ones in the interstitial and centromere regions) in the nucleus of *A. fistulosum*.

Barnes et al. (1985) reported that the *Bam*HI repeated DNA sequence is localized in the terminal regions of all but two of the 16 chromosomes and in the two satellites of *A. cepa*, which agrees well with the localization of the *Eco*RV repeated DNA sequence in *A. fistulosum* described above. Similar locations of repeated DNA sequences in terminal regions of chromosomes have been reported in the plant species *Secale cereale* (Appels et al. 1978), *Scilla siberica* (Deumling 1981) and *Lycopersicon esculentum* (Lapitan et al. 1989). It should be noted that the C-banding pattern of *A. fistulosum* shown in Fig. 4 is in good agreement with that reported for this leek species by El-Gadi and Elkington (1975) and by Vosa (1976).

The C-heterochromatin content in *Allium* species, including *A. fistulosum*, is reported to involve 10–12% of the genome (Vosa 1976). In contrast, our 380-bp sequence is estimated to make up approximately 4.5% of the *A. fistulosum* genome based on quantitative results of a Southern-blot analysis. Therefore, this 380-bp sequence is estimated to make up almost half of the C-heterochromatin in this leek species. This value is reasonable considering the results obtained for in situ hybridization and C-banding analyses (see Fig. 4).

Interestingly, the 380-bp repeated DNA sequence in *A. fistulosum* and the 375-bp one reported in *A. cepa* (Barnes et al. 1985) have many similarities including sequence length, nucleotide composition, genomic content, and chromosomal localization. These similarities suggest that *Allium* species may have a common ancestral sequence of repeated DNA that is closely associated with the chromosomal C-heterochromatin. Hybridization analyses, using the *A. fistulosum* repeated DNA sequence, to test this hypothesis, should provide information on the divergence of repeated DNA in the genus *Allium*.

References

Appels R, Driscoll C, Peacock WJ (1978) Heterochromatin and highly repeated DNA sequences in rye (*Secale cereale*). Chromosoma 70:67–89

Barnes SR, James AM, Jamieson G (1985) The organization, nucleotide sequence, and chromosomal distribution of a satellite DNA from *Allium cepa*. Chromosoma 92:185–192

Bedbrook JR, Jones J, O'Dell M, Thompson RD, Flavell RB (1980) A molecular description of telomeric heterochromatin in *Secale* species. Cell 19:545–560

Beridze T (1986) Satellite DNA. Springer, Berlin Heidelberg New York, pp 30–97

Deumling B (1981) Sequence arrangement of highly methylated satellite DNA of plant, *Scilla*: a tandemly repeated inverted repeat. Proc Natl Acad Sci USA 78:338–342

El-Gadi A, Elkington T (1975) Comparison of the Giemsa C-band karyotypes and the relationships of *Allium cepa*, *A. fistulosum*, and *A. galanthum*. Chromosoma 51:19–23

Hizume M, Tanaka A, Yonezawa Y, Tanaka R (1980) A technique for C banding in *Vicia faba* chromosomes. Jpn J Genet 55:301–305

Iwabuchi M, Itoh K, Shimamoto K (1991) Molecular and cytological chracterization of repetitive DNA sequences in *Brassica*. Theor Appl Genet 81:349–355

Jones RN, Rees H (1968) Nuclear DNA variation in *Allium*. Heredity 23:591–605

Lapitan NLV, Ganal MW, Tanksley SD (1989) Somatic chromosome karyotype of tomato based on in situ hybridization of TGRI satellite repeat. Genome 32:992–998

Peacock WJ, Dennis ES, Rhoades MM, Pryor AJ (1981) Highly repeated DNA sequence limited to knob heterochromatin in maize. Proc Natl Acad Sci USA 78:4490–4494

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:83–103

Singer MF (1982) Highly repeated sequences in mammalian genomes. Int Rev Cytol 76:67–112

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503–517

Tanaka R, Taniguchi K (1975) A banding method for plant chromosomes. Jpn J Genet 50:163–167

Vosa CG (1976) Heterochromatic patterns in *Allium*. 1. The relationship between the species of the *cepa* group and its allies. Heredity 36:383–392

Wu HK, Chung MC, Wu T, Ning CN, Wu R (1991) Localization of specific repetitive DNA sequences in individual rice chromosomes. Chromosoma 100:330–338

Yakura K, Kato A, Tanifuji S (1987) Cytological localization of highly repeated sequences, the *Fok*I sequence family and *Bam*HI sequence families, in *Vicia faba*. Jpn J Genet 62:325–332