Theor Appl Genet (1992) 85:303–308

THEORETICAL AND APPLIED GENETICS

Species-specific *Eco*RI repetitive elements of at least 16 kb in length are present in *Lupinus luteus*

T. Sakowicz

Department of Plant Cytology and Cytochemistry, University of Lodz, 90-237 Lodz, Banacha 12/16, Poland

Received February 29, 1991; Accepted March 10, 1992
Communicated by G. Wenzel

Summary. A genomic DNA library of *Lupinus luteus* cv. Ventus was constructed in the phage vector EMBL3 using *Mbo*I-digested DNA. Screening with a 1070 bp labelled repetitive unit from *L. luteus* yielded several DNA clones. The repetitive family is composed of elements whose length is at least 16 kb. The average copy number of the cloned fragments is 5.0×10^4 per haploid genome and constitutes approximately 3% of the total *L. luteus* genome. The homologous repeats were found in all ten cultivars of *L. luteus* tested but were not detected in two cultivars each of the closely related species *Lupinus albus* and *Lupinus angustifolius*. The *Eco*RI family fragments could thus be considered as species-specific DNA elements. These fragments may be useful as molecular markers in the genetic manipulation of *L. luteus*.

Key words: *Lupinus luteus* – Repetitive sequences – Species–specific probes – Molecular marker

Introduction

Eucaryotic genomes are characterized by repetitive DNA sequence families which constitute a major part of the genome and have been widely investigated. These repetitive sequences may be interspersed with single-copy DNA sequences and there is a great diversity in the size, organization and copy number of this DNA. Moderately and highly repetitive DNA sequences often represent non-coding regions of the genome that may undergo evolutionary changes at a relatively rapid rate. Analysis of such repetitive families has revealed differences between closely related species and has contributed to the determination of evolutionary relationships.

Specific repeated sequences have been reported in several higher plants, including *Triticum* (Metzlaff et al. 1986), *Secale* (Appels et al. 1986), *Aegilops* (Rayburn and Gill 1986), *Lycopersicon* (Ganal et al. 1988), *Actinidia* (Crowhurst and Gardner 1991), and *Oryza* (Zhao et al. 1989; Wu et al. 1991).

We have previously decribed a new family of middle repetitive DNA sequences in *L. luteus* denoted as the *Eco*RI family. DNA/DNA hybridization experiments have revealed that the 1070-bp long fragments representing this family which we cloned and sequenced are in fact internal parts of much longer units (Sakowicz et al. 1986).

The present paper reports mapping and characterisation of a 16-kb element of *L. luteus* DNA containing sequences homologous to *Eco*RI-family fragments. We show that the occurrence of such fragments are restricted to cultivars of *L. luteus* and are not found in those of related species. We conclude that repetitive *Eco*RI-family fragments from *L. luteus* are species-specific sequences.

Materials and methods

Plant material

A list of the plant material used in this study is given in Table 1. All the material was obtained from the collection held at the Plant Breeding Station, Wiatrowo, Poland.

Isolation of DNA

The isolation of high molecular-mass DNA from plants was carried out according to the procedure developed by Murray and Thompson (1980) using 0.5 cm-long root tips grown in the dark for about 2–3 days. The isolated DNA was then purified on cesium chloride – ethidium bromide gradients. Phage and plasmid DNA were prepared according to the standard procedures of Maniatis et al. (1982).

Enzymes

Restriction endonucleases, T4 DNA ligase, and the DNA polymerase I large fragment were all obtained from Bethesda Re-

Table 1. Plant material used in this study

Species	Cultivar	Feature
L. luteus	Ventus	Sweet
	Topaz	Sweet
	Orbit	Sweet
	Juno	Sweet
	Jantar	Sweet
	Manru	Sweet
	Cyt	Sweet
	Alb	Sweet
	Palestyna 2	Bitter
	Palestyna 1300	Bitter
L. albus	Wat	Sweet
	Bac	Bitter
L.angustifolius	Turkus	Sweet
	Mirela	Bitter
Pisum sativum		
Phaseolus vulgaris		
Vicia faba		

search Laboratories, GIBCO, and Boehringer Mannheim. [α^{32}P] dNTP and [γ-^{32}P] ATP were purchased from Amersham. The chemicals used for DNA sequencing and modifications were from the Sigma Chemical Company Ltd.

Construction and screening of a genomic DNA library

Total *L. luteus* cv. Ventus DNA was partially digested with *Mbo*I and 15–20 kb of the DNA fraction was ligated into the *Bam*HI site of the EMBL3 phage vector (Frischauf et al. 1983). The library was constructed by Dr. A. Konieczny (Konieczny et al. 1987).

Screening was carried out as described (Sakowicz et al. 1986) with a purified 1070 bp insert from plasmid pKoK5 labelled by nick-translation. Phages showing a positive signal with the 1070 bp fragment were further characterized and were designated as λ2, λ12, and λ22.

Subcloning of the phage sequences

Southern hybridization with λ2 digestion products was used to identify appropriate fragments of λ2 for subcloning into pUC18. The newly created plasmids were named:

p2 = pUC18/*Eco*RI + *Eco*RI–*Eco*RI 4.6 kb
p4 = pUC18/*Hin*dIII/*Sal*I + *Hin*dIII–SalI 2.7 kb
p5 = pUC 18/*Eco*RI + *Eco*RI–*Eco*RI 2.1 kb
p8 = pUC18/*Hin*dIII + *Hin*dIII–*Hin*dIII 7.0 kb
p12 = pUC18/*Sal*I/*Bgl*II + *Sal*I–*Bgl*II 3.1 kb
p24 = pUC18/*Hin*dIII + *Hin*dIII–*Hin*dIII 3.3 kb

DNA hybridizations

The DNA probes used for hybridizations were labelled by the nick-translation method (Rigby et al. 1977). Specific activities of the probes used were approximately 4×10^7 cpm/μg of DNA. Southern hybridization was carried out according to the procedure of Southern (1975). After 1% agarose-gel electrophoresis DNA was denatured and transferred to a nitrocellulose filter paper (supplied by Schleicher and Schüll). Plaque hybridization was performed according to Benton and Davis (1977) while filters for dot blot hybridization were prepared as described by Katafos et al. (1979). The stringency of washing was as recommended by Meinkoth and Wahl (1984).

Determination of the copy number of the subcloned repeated DNA sequences in L. luteus

The frequency of the *Eco*RI-family members in the *L. luteus* genome was determined using the dot-blot technique. Appropriate amounts of total *L. luteus* DNA and recombinant plasmids containing one copy of the fragment belonging to the *Eco*RI family were fixed on nitrocellulose filters. Hybridization with nick-translated DNA probes was carried out at 68 °C for 18 h in a solution containing 0.1% BSA, 0.1% ficoll, 0.1% PVP, 0.5% SDS and 200 μg/ml of denatured *E. coli* DNA. A comparison of the radioactivity bound within the linear range of response was used to determine the copy number in the *L. luteus* genome.

DNA sequence determination

The sequence of end-labeled DNA restriction fragments was determined according to the base-specific chemical degradation protocol of Maxam and Gilbert (1980). Cleavage products were displayed on either 6,8 or 20% polyacrylamide 8 M urea gels and exposed to X-ray film (Kodak XAR5) with intensifying screen (Lighting Plus, Du Pont Co.) at −70 °C.

Isolated DNA fragments were either labelled at their 3′ end using the Klenow fragment of *E. coli* DNA *pol*I with the appropriate [α-^{32}P] dNTP (Amersham), following the protocol of Drouin (1980), or were labelled at their 5′ end with [γ-^{32}P] ATP using T4 polynucleotide kinase (Boehringer Mannheim), as described by Maxam and Gilbert (1980). The ends of labelled fragments were separated by secondary cleavage with appropriate restriction endonucleases. Labelled fragments were purified from agarose or polyacrylamide gels by electroelution into dialysis bags according to McDonnell et al. (1977).

Results

Mapping of the 16 kb element containing sequences homologous to EcoRI – family fragments

Previous analysis of dispersed *Eco*RI-family fragments in the *L. luteus* L. genome indicated that 1070 bp fragments of that family were internal parts of much longer repetitive units (Sakowicz et al. 1986).

Labelled 1070 bp fragments (cloned and sequenced in plasmid pKoK5) were used to determine phages hybridizing to it from the *L. luteus* genomic library. Of the 2000 phages screened 20 provided distinct signals. We selected three of these called λ2, λ12, and λ22, for analysis.

After digestion with *Eco*RI, *Hin*dIII, *Sal*I, *Kpn*I and *Hae*III the DNA of these phages was separated on gels, transferred onto nitrocellulose and hybridized with the 1070 bp fragment as a probe. The restriction patterns of the tested phages showed distinct differences (see Fig. 2a), but all of them revealed fragments homologous to the probe.

A DNA fragment, 16 kb in length, was cloned in λ2 phage. *Eco*RI and *Hin*dIII restrictases were mostly used to construct a map of the 16 kb fragment. Mapping started with finding out which λ2/*Hin*dIII fragments hybridized to the 1070 bp fragment. Subsequently one of three such fragments was cloned into pUC18/*Hin*dIII.

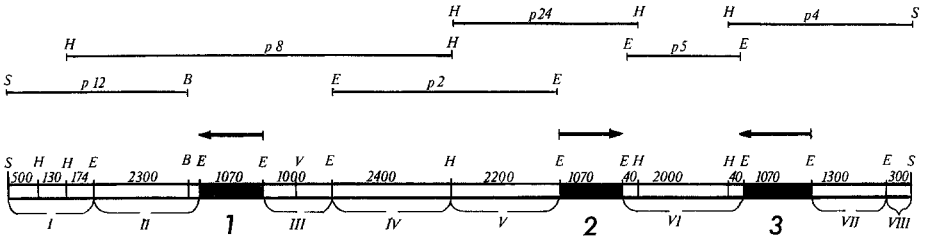

Fig. 1. Restriction map of the 16 kb *L. luteus* fragment cloned in λ2 phage. The following cleavage sites are shown: *E*, *Eco*RI; *H*, *Hin*dIII; *S*, *Sal*I; *B*, *Bgl*II; *V*, *Eco*RV. *Black boxes* indicate 1070 bp fragments denoted as *1*, *2*, *3*. The relative directions of the 1070 bp fragments are indicated by *arrows*. Restriction fragments of λ2 used for subcloning in pUC18 are shown as *horizontal bars*. Fragments I–VIII were used as hybridization probes to determine internal sequence organization

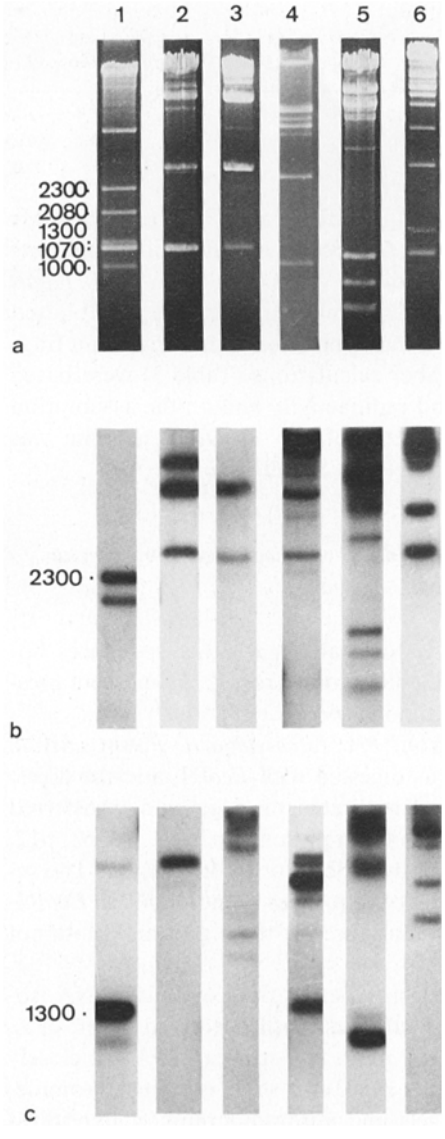

Fig. 2. **a** Restriction patterns of λ2, λ12, λ22 DNA digested with *Eco*RI (*lanes* 1, *2*, *3*) and *Hin*dIII (*lanes* 4, *5*, *6*). Digestion products revealed essential differences between phages. **b, c** autoradiograms of the gel in **a** after blotting and hybridizing with the 2,300 pb (**b**) and 1,300 bp (**c**) fragments from λ2

The insert was 7000 bp in length. The resulting plasmid was called p8. A restriction map of the p8 insert was constructed on the basis of the results of the digestion of the end-labeled 7000 bp fragment. To identify neighbours of the 7000 bp fragment on λ2 we used flanks of 7000 bp as probes in Southern hybridization. A series of similar experiments (involving "walks" along λ2 molecules) was repeated until the arms of the phage were reached.

The resulting experimental data made possible the creation of a restriction map of the 16 kb λ2 insert (Fig. 1). This map indicated that the 1070 bp fragment was repeated three times in the 16 kb region. The surroundings of each of these three repeats were sequenced to determine their relative direction to one another on λ2.

Internal organization of the 16 kb sequences of λ2 and their homology with λ12 and λ22

Cloned inserts isolated from appropriate plasmids were used as hybridization probes to analyse the internal organization of λ2 DNA. Using fragments I–VIII (Fig. 1) as probes several cross hybridizations were made. The *Eco*RI and *Hin*dIII digestion products of λ2, λ12, λ22 phages and p2, p4, p5, p8, p12, p24 plasmids were blotted onto membranes. Selected autoradiograms are shown in Fig. 2b, c. The 2300 bp *Eco*RI–*Eco*RI fragment (left neighbour of 1070 bp, 1 on λ2 in Fig. 1) and the 1300 bp *Eco*RI-*Eco*RI (right neighbour of 1070, 3 on λ2 in Fig. 1) were used as probes.

The main conclusion resulting from studies on three independent λ2, λ12, λ22 phage clones was that the repetitive elements in *L. luteus* were at least 16 kb in length and were homologous to one another. Each of probes I–VIII, derived from λ2, hybridized to *L. luteus* DNA fragments in λ12 and λ22, through the electrophoretic pattern of the λ12/*Eco*RI, λ22/*Eco*RI, λ2/*Eco*RI, λ12/*Hin*dIII, and λ2/*Hin*dIII digestion products revealed essential differences between the phages (Fig. 2a). The above data suggest that the *L. luteus* 16 kb λ2, λ12, λ22 inserts belong to the same family of dispersed sequences.

Table 2. A comparison of sequence homology of the λ2 fragments. On the basis of the densitometric tracing of autoradiograms (Fig. 2b, c) the percentage of relative homologies of λ2 fragments was calculated (+++, 100% homology; ++, 70–80% homology; +, 20–30% homology)

	S-E 800	E-E 2300	E-E 1070 1	E-E 1000	E-H 2400	H-E 2200	E-E 1070 2	E-E 2080	E-E 1070 3	E-E 1300	E-S 300
S-E 800	+++										
E-E 2300		+++						++			
E-E 1070 1			+++				+++		+++		
E-E 1000				+++		+				+	
E-H 2400					+++						
H-E 2200				+		+++				+	
E-E 1070 2			+++				+++		+++		
E-E 2080		++						+++			
E-E 1070 3			+++				+++		+++		
E-E 1300				+		+				+++	
E-S 300											+++

Table 3. Copy number of *Eco*RI-family fragments in the *L. luteus* genome

Fragments of λ2 (see Fig. 2)	Copy number
I	0.9×10^4
II	5.4×10^4
1070	6.6×10^4
III	3.3×10^4
IV	1.4×10^4
V	3.3×10^4
VI	5.4×10^4
VII	3.3×10^4
VIII	0.9×10^4

The following conclusions apply to the internal organization of *Eco*RI-family fragment representatives (Table 2):

(1) Some of the DNA fragments were repeated several times in the 16 kb region (three times in the case of the 1070 bp fragment).

(2) The 2300 bp fragment, being the left neighbour of the 1070 bp fragment (shown as 1 in Fig. 1) on λ2 molecules, showed high homology (70–80%) to the 2000 bp fragment from the area between the 1070 fragments (shown as 2 and 3 in Fig. 1), the distance between the homologous fragments being several thousand bp;

(3) A lower homology existed in the 16 kb of λ2. For example, the 1300 bp *Eco*RI–*Eco*RI right neighbour of 1070 (shown as 3 in Fig. 1) revealed only slight (20–30%) homology to the 2200 bp *Hin*dIII–*Eco*RI fragment and to the 1000 bp *Eco*RI–*Eco*RI fragment which is the right neighbour of the 1070 fragment (shown as 1 in Fig. 1).

(4) Unique sequences were also found in the λ2 DNA fragment, i.e., in the fragments flanking it (I, VIII in Fig. 1) and in the central 2400 bp *Eco*RI–*Hin*dIII sequence.

Copy number

Using the protocol of Katafos et al. (1979) it was possible to establish the copy number of *Eco*RI-family fragments in the *L. luteus* genome. A series of dilutions of *L. luteus* nuclear DNA and of recombinant plasmids were blotted onto nitrocellulose and hybridized with the relevant fragments. Copy number calculations (Table 3) were based on counting bound radioactivity and on the assumption that the DNA content of the *L. luteus* genome was 1.7×10^9 bp. (Bennett and Smith 1976).

EcoRI-family fragments reveal sequence homology only with the genome of L. luteus

Efforts were made to establish whether sequences homologous to the *Eco*RI family from *L. luteus* were present in other leguminous species.

Total DNA from *Phaseolus vulgaris*, *Pisum sativum* and *Vicia faba* was digested with *Eco*RI, and the digestion products were separated on a gel and transferred onto a filter. Inserts from plasmids p2, p4, p5, p8, p12, and p24 were used as probes for hybridization. The results showed a lack of sequences homologous to *Eco*RI-family fragments in all three of these genomes (data not shown).

Similar experiments were carried out with DNA isolated from several cultivars, both sweet and bitter, of *L. angustifolius* and *L. albus*, two species which are closely related to *L. luteus*. Fig. 3a, b show ethidium bromide-stained agarose gels and autoradiograms of hybridizations with the 1070 bp *Eco*RI-*Eco*RI fragment used as a probe. From these autoradiograms it is clear that fragments of the *Eco*RI family are species-specific and are present exclusively in all the tested cultivars of *L. luteus*.

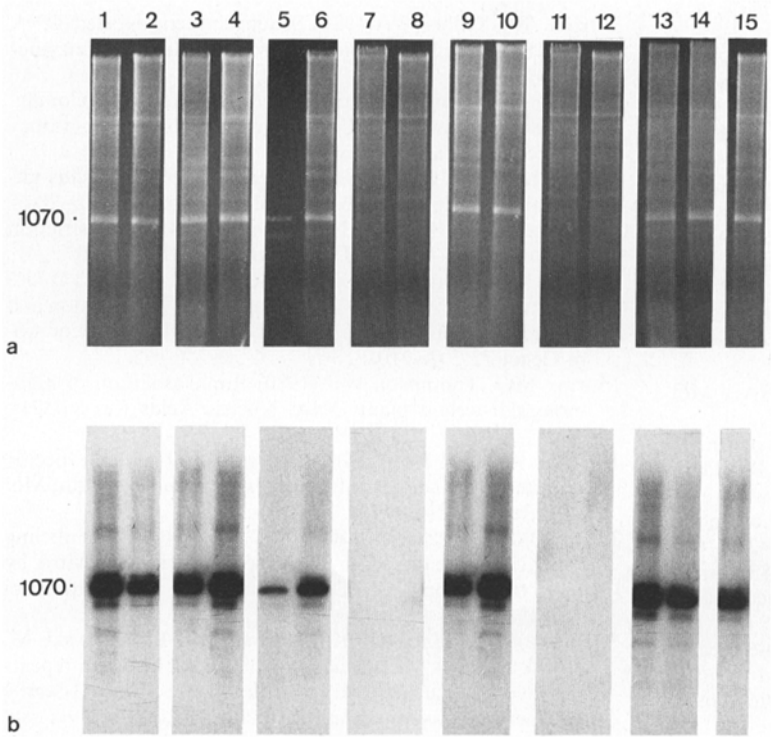

Fig. 3 a, b. Species-specificity of the *Eco*RI-family fragments. **a** Ethidium bromide-stained agarose gel containing different *Lupinus* DNA digested with *Eco*RI: *L. luteus* cvs Ventus, Topaz, Orbit, Juno, Jantar, Palestyna 2, Palestyna 1,300, Manru, Cyt, Alb (*lanes 1, 2, 3, 4, 6, 9,* 10, *13, 14, 15* respectively); *L. angustifolius* cvs Turkus, Mirela (*lanes 11, 12*); *L. albus* cvs Wat, Bac (*lanes 7, 8*); 1,070 bp fragment (*lane 5*). **b** Autoradiogram of the **a** gel after blotting and probing with a labelled 1,070 bp fragment from $\lambda 2$

```
                    20                  40
5' ...TCTTTTTTTACCCCACTAGTTTCTCCCCTTATCTTTTTAATTTTTTTTTTAATTTTT
      60                  80                  100
     TTTCATTTTTTTCATTTTTTTCTTCTTTTTCTTTCTCATTCTTTT...3'
```

Fig. 4. The nucleotide sequence located on the right flank of the *Eco*RI-*Eco*RI 1,300 bp fragment of $\lambda 2$ contains a long polypurine/polypyrimidine region

Discussion

Comparisons between the repeated sequences of closely related species which have diverged from a common ancestor have shown that many of the repeats found in more than one species are present in different copy number and are organized differently in different species (Flavell 1986). These changes are brought about by amplification, by dispersion of the repeats around the genome, by the deletion of sequences, and by the new synthesis of repetitive DNA (Flavell 1980).

Some middle repetitive sequences may be retained in a number of plant species or else completely eliminated from others and thus become species-specific. An example of the latter type is provided by the *Eco*RI family of fragments, which are specific to the *L. luteus* genome. No members of this sequence family were present in *L. angustifolius* and *L. albus*. The internal organization of the members of the dispersed *Eco*RI fragments from *L. luteus*, i.e., the very high number of direct and inverted repeats, reflects the divergence of the fragments analysed. Because the copy number of *Eco*RI-family fragments show distinct differences from one another it appears that the intensity of the amplification process may differ for different fragments coming from the 16 kb region and may result in a type of arrangement in which repeats are interspersed with one another.

Partial sequencing (9.0 kb) of a $\lambda 2$ genomic clone showed:

(1) A very high number of direct and inverted repeats (4–7 bp long).

(2) A very high AT content (approaching 70%), and

(3) The presence of at least four 50–100 bp-long regions that contain almost exclusively pyrimidines (one such a sequence is presented in Fig. 4).

Long polypurine/polypyrimidine sequences have been observed in the upstream regions of eukaryotic genes. It has been suggested that such sequences may affect gene expression by adopting an altered chromatin conformation (Hoffman-Liebermann 1986). In *L. luteus* blocks in question show sensitivity to either S1 nuclease or to diethyl pyrocarbonate (data not shown), which may suggest the presence of unusual DNA structures in these regions (Wells 1980; Galazka et al. 1986).

Acknowledgements. The author acknowledges the help of J. Klysik at the initial stage of this work. Additionally, he thanks A. Konieczny for access to the *L. luteus* genome library, M. J. Olszewska for valuable suggestions and a careful review of the manuscript, and W. Swiecicki (Plant Breeding Station in Wiatrowo Poland) for supplying the plant material.

References

Appels R, Moran LB, Gustafson JP (1986) Rye heterochromatin I. Studies on clusters of the major repeating sequence and the identification of a new dispersed repetitive sequence element. Can J Genet Cytol 28:645–657

Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. Phil Trans R Soc Lond Ser B 274:227–274

Benton WD, Davis RW (1977) Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196:180–182

Crowhurst RN, Gardner RC (1991) A genome-specific repeat sequence from kiwifruit (*Actinidia deliciosa* var. *deliciosa*. Theor Appl Genet 81:71–78

Drouin J (1980) Cloning of human mitochondrial DNA. J Mol Biol 140:15–34

Flavell RB (1980) The molecular characterization and organization of plant chromosomal DNA sequences. Annu Rev Plant Physiol 31:569–596

Flavell RB (1986) Repetitive DNA and chromosome evolution in plants. Phil Trans R Soc Lond Ser B 312:227–242

Frischauf AM, Lehrach H, Poustka A, Murray N (1983) Lambda replacement vectors carrying polylinker sequences. J Mol Biol 170:827–842

Galazka G, Palecek E, Wells RD, Klysik J (1986) Site-specific OsO_4 modification of the B-Z junctions formed at the $(dA\text{-}dC)_{32}$ region in supercoiled DNA. J Biol Chem 261:7093–7098

Ganal MW, Lapitan NLV, Tanksley SD (1988) Molecular and cytogenetic survey of major repeated DNA sequences in tomato (*Lycopersicon esculentum*). Mol Gen Genet 213:262–268

Hoffman-Liebermann B, Liebermann D, Troutt A, Kedes LH, Cohen SN (1986) Human homologs of TU transposon sequences: polypurine/polypyrimidine sequence elements that can alter DNA conformation in vitro and in vivo. Mol Cell Biol 6:3632–3642

Katafos FC, Jones CW, Efstratiadis A (1979) Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. Nucleic Acids Res 7:1541–1552

Konieczny A, Szczygloski K, Madrzak C, Legocki AB (1987) Cloning and identification of lupin nodule-specific cDNA sequences. Acta Biochim Polon 34:21–27

Maxam AM, Gilbert W (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol 65:499–560

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbour Laboratory Cold Spring Harbor, New York

Meinkoth J, Wahl G (1984) Hybridization of nucleic acids immobilized on solid supports. Anal Biochem 138:267–284

McDonnell MW, Simon MN, Studier FW (1977) Restriction fragments of T7 DNA. J Mol Biol 110:119–149

Metzlaff M, Troebner W, Baldauf F, Schlegel R, Cullum J (1986) Wheat-specific repetitive DNA sequences – construction and characterization of four different genomic clones. Theor Appl Genet 72:207–210

Murray MG, Thompson WF (1980) Rapid isolation of high-molecular-weight plant DNA. Nucleic Acids Res 8:4321–4325

Rayburn AL, Gill BS (1986) Isolation of a D-genome-specific repeated DNA sequence from *Aegilops squarrosa*. Plant Mol Biol Rep 4:102–109

Rigby PWJ, Dieckman M, Rhodes C, Berg P (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J Mol Biol 113:237–251

Sakowicz T, Galazka G, Konarzewska A, Kwinkowski M, Klysik J (1986) An unusually high number of direct repeats detected by sequence analysis of the dispersed *Eco*RI-family fragments in *Lupinus luteus* L. Planta 168:207–213

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

Wells RD (1980) The chemistry and biology of unusual DNA structures adopted by oligopurine /oligopyrimidine sequences. FASEB J 2:2939–2946

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

Zhao X, Wu T, Xie Y, Wu R (1989) Genome-specific repetitive sequences in the genus *Oryza*. Theor Appl Genet 78:201–209