Theor Appl Genet (1989) 78:77–80



# A BamHI family of highly repeated DNA sequences of *Nicotiana tabacum*

**B. Koukalová, J. Reich, R. Matyášek, V. Kuhrová and M. Bezděk**

Institute of Biophysics, Czechoslovak Academy of Sciences, CS-612 65 Brno, Czechoslovakia

Received December 12, 1988; Accepted March 17, 1989
Communicated by Yu. Gleba

**Summary.** HRS60.1, a monomer unit (184 bp) of a highly repeated nuclear DNA sequence of *Nicotiana tabacum*, has been cloned and sequenced. Following BamHI digestion of tobacco DNA, Southern hybridization with HRS60.1 revealed a ladder of hybridization bands corresponding to multiples of the basic monomer unit. If the tobacco DNA was digested with restriction endonucleases which have no target site in HRS60.1, the larger part of DNA homologous to HRS60.1 remained as uncleaved "relic" DNA. These results suggest a tandem arrangement of this DNA repeat unit. Four other clones of tobacco nuclear DNA cross-hybridized with HRS60.1, thus forming a "HRS60-family". Sequencing their inserts has shown their strong mutual homology. HRS60-family comprised about 2% of the nuclear genome of *N. tabacum*. Computer comparisons with other tandem plant-repeated DNA sequences could not detect any other homologous sequence.

**Key words:** Repeated DNA – Tandem repeat – Nucleotide sequence – *Nicotiana tabacum*

## Introduction

The presence of DNA-repeated sequences is a typical feature of most eukaryotic genoms. In plants, their sum often comprises a larger part of the genome, as reviewed by Walbot and Golberg (1979) and Flavell (1983). According to their organization in the genome, repeated DNA sequences can be either dispersed or clustered, forming a tandem array of closely related units. Generally, clustered DNA repeats can be detected in centromeric and telomeric heterochromatin and, as a rule, they are transcriptionally inactive. Some of the tandemly repeated DNA sequences can be separated by density gradient centrifugation as a satellite DNA. Others may exist as cryptic satellites.

In the last few years, several tandemly repeated DNA sequences of plants were analyzed with the aim to obtain more information on primary structure, molecular organization, evolution and eventually, the function of this genetic material (Bedbrook et al. 1980; Appels et al. 1981; Deumling 1981; Peacock et al. 1981; Brennicke and Hemleben 1983; Capesius 1983; Kato et al. 1984; Barnes et al. 1985; Benslimane et al. 1986; Grellet et al. 1986; Martinez-Zapater et al. 1986; Leclerk and Siegel 1987; Wu and Wu 1987).

Nevertheless, the number of these DNA repeats analyzed at the molecular level is not yet very large. Here we describe a family of tandemly repeated DNA sequences from *Nicotiana tabacum* nuclear DNA.

## Materials and methods

### *DNA extractions and plant material*

Plasmid DNA was extracted according to Birnboim and Doly (1979). Nuclear DNA of *N. tabacum* was isolated from leaves. These were homogenized in 1% citric acid by grinding in the mortar at 0 °C. The homogenate was filtered through a nylon cloth (ca. 50 μm mesh width) and centrifuged 3 min at 1,000 g. The pellet contained crude nuclei. Pelleting of the nuclei resuspended in 1% citric acid by low speed centrifugation was repeated twice. The nuclei were lysed by 2% SDS in 50 m*M* TRIS.Cl, pH 8, 250 m*M* NaCl at 70 °C, for 10 min. SDS was precipitated by KAc (final concentration of 0.5 *M*) at 0 °C. The supernatant was extracted twice with phenol and twice with chloroform and precipitated with 2.5 vol. of ethanol. DNA was spooled, washed by 70% ethanol and dissolved in 10 m*M* TRIS.Cl pH 7.6, 1 m*M* EDTA. This DNA migrates as a sharp single band on agarose gel electrophoresis. The total DNA of *N. tabacum* was isolated essentially according to Murray and Thompson (1980), but the CsCl gradient centrifugation was routinely omitted (Karlovská et al. 1987).

*Construction of genomic library and selection of recombinant plasmids*

A partial BamHI digest of the tobacco nuclear DNA was ligated, using T4 ligase, into BamHI site of the plasmid pUC19 (Yanisch-Perron et al. 1985). This vector was treated with bovine alkaline phosphatase before the ligation (Maniatis et al. 1982). *Escherichia coli* JM109 cells (Yanisch-Perron et al. 1985) were transformed (Hanahan 1983) with the ligation mixture. Bacterial clones bearing recombinant plasmids were selected on X-gal plates with ampicillin.

*Restriction endonuclease digestion and gel electrophoresis*

The tobacco DNA was digested with an excess of restriction enzymes in order to obtain complete digests. DNA fragments were fractionated by electrophoresis on agarose gel. After staining with ethidium bromide and photographing, DNA fragments were blot-transferred onto nitrocellulose membrane (Southern 1975).

*DNA/DNA hybridization*

After transfer, DNA was hybridized with heat-denatured probes according to Maniatis et al. (1982). For the washing procedure, the stringent conditions (0.1 × SSC, 65 °C) were used. In some experiments, DNA dotted on nitrocellulose membrane was hybridized with a probe (Thomas 1980).

*$^{32}$P-labelling of DNA probes*

The procedure according to Maniatis et al. (1982) was used.

*Determination of approximate number of copies of HRS60 repeat unit*

The procedure essentially according to Grellet et al. (1986) was used. Serial dilutions of tobacco nuclear DNA and those of HRS60.5 DNA were blotted onto nitrocellulose membrane using the Schleicher and Schuell Minifold II apparatus and hybridized with a $^{32}$P-labelled HRS60.5 probe. After autoradiography, densities of the individual bands were measured and the amounts of both tobacco and HRS60.5. DNAs giving the same density of hybridization were calculated. On the basis of these values, the fraction of tobacco DNA homologous to HRS60.5 was estimated.

*DNA sequence analysis*

The method of Maxam and Gilbert (1977) was used. Both DNA strands were sequenced.

## Results and discussion

*Isolation of HRS60-family of DNA repeated sequences*

Recombinant plasmids containing the *N. tabacum* nuclear DNA were dot-hybridized with the $^{32}$P-labelled total tobacco DNA. Plasmids which gave strong hybridization signals were expected to harbor repeated DNA sequences and served as sources of DNA probes in further hybridization experiments. BamHI digests of *N. tabacum* DNA, size-separated by electrophoresis on 0.7% agarose gel and blotted to nitrocellulose membrane, were then hybridized with the $^{32}$P-labelled DNA of selected recombinant plasmids. Various patterns of molecular hybridization were obtained with different

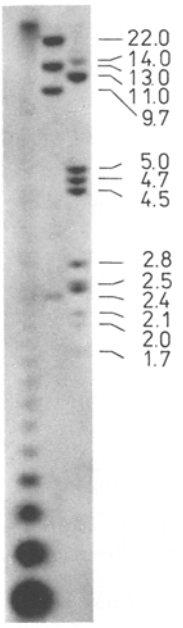


**Fig. 1.** Molecular arrangement of HRS60-family of DNA repeats in *N. tabacum*. *Lane 1:* BamHI digest of tobacco DNA probed with HRS60.1. *Lanes 2* and *3:* Lambda DNA digested with BglII and PstI, respectively. The *numbers* indicate the sizes of DNA fragments in kb

probes; one of them, the plasmid No 60, containing a 184-bp DNA insert (HRS60.1) of tobacco DNA, hybridized with multiple BamHI fragments of tobacco DNA, forming a ladder of bands. This hybridization pattern did not change with increasing amount of BamHI nuclease. The strongest hybridization signal corresponded to ca. 180-bp DNA fragment. Hybridization signals of longer DNA fragments were gradually less pronounced and their positions corresponded to exact multiples of the basic monomer unit (Figs. 1 and 2). HRS60.1 cross-hybridized with four other recombinant plasmids. Their inserts were denoted as HRS60.2, HRS60.3, etc. They represent, together with HRS60.1, the "HRS60-family".

*Members of HRS60-family are arranged in tandems*

The pattern of hybridization of HRS60.1 with the BamHI digest of tobacco DNA (Fig. 1) suggested that monomers of the HRS60-family were organized in tandem arrays. In some autoradiograms, bands of hybridization could be seen which corresponded up to the 20-mer of the basic monomer repeat unit. Tobacco DNA was also digested with several restriction endonucleases which have no target sites in HRS60.1, i.e. XbaI, BglII, or AluI. BamHI digestion of tobacco DNA was included for comparison. An excess of endonucleases was used to obtain a complete digest of tobacco DNA (Fig. 3A). Southern blot hybridization with $^{32}$P-labelled HRS60.1

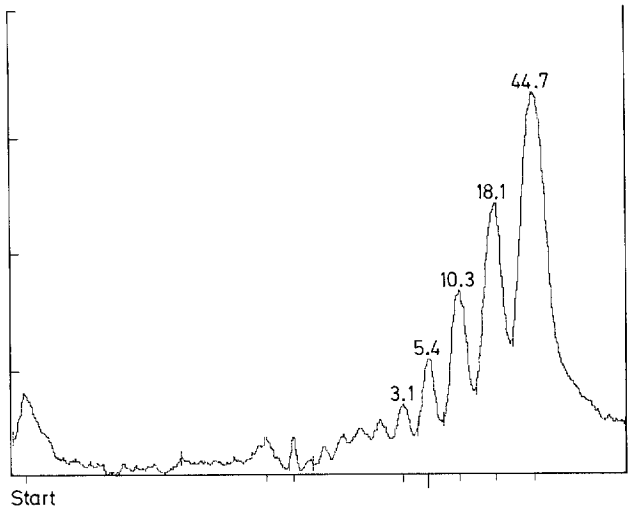


**Fig. 2.** Densitometric evaluation of hybridization bands of the BamHI digest of *N. tabacum* DNA probed with HRS60.1. The *numbers* above the peaks indicate the percentage of the total hybridization, the strongest hybridization (44.7%) being for the monomeric band

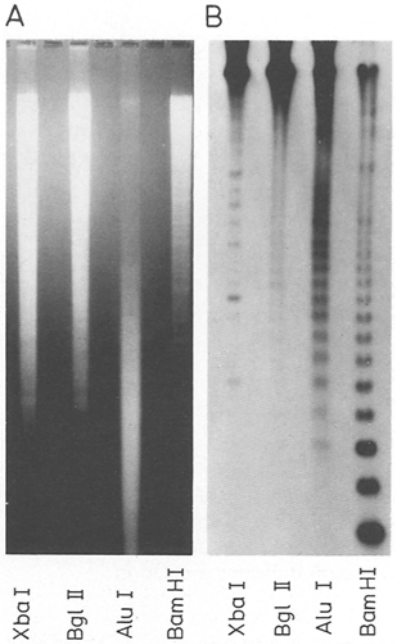


**Fig. 3 A and B.** Demonstration of "relic" DNA homologous with HRS60.1 as an undigested fraction of tobacco genome. **A** Tobacco DNA was digested with the restriction nucleases XbaI (*lane 1*), BglII (*lane 2*), AluI (*lane 3*), and BamHI (*lane 4*), electrophoresed on a 0.8% agarose gel and stained with ethidium bromide. **B** After blotting on nitrocellulose membrane, hybridization with the HRS60.1 DNA was carried out

showed that most of the DNA, homologous to HRS60.1, was localized (in the cases of XbaI, BglII and AluI digests) in undigested "relic" DNAs (Fig. 3 B). These results proved that the HRS60-family was organized in large clusters in the genome of *N. tabacum*. Fractions of DNA, homologous to HRS60.1, corresponding to XbaI, BglII, and AluI "relic" DNAs, were 0.77, 0.69, and 0.49, respectively. The smaller parts of tobacco DNAs homologous to HRS60.1 form ladders. Apparently, some members of the HRS60-family have, in contrast to HRS60.1, recognition sites for the above-mentioned restrictases (e.g., HRS60.5 has one AGCT recognition site for AluI; Fig. 5).

*Two percent of tobacco nuclear genome is formed by HRS60-family*

Two independent hybridization experiments were made in order to estimate the fraction of tobacco nuclear DNA consisting of HRS60-family. Using the procedure and calculations given in 'Materials and methods' it was found that the HRS60.5 probe hybridized 50 times more with HRS60.5 DNA than with the whole nuclear tobacco DNA (Fig. 4). Thus, HRS60-family comprises about 2% of the nuclear genome of *N. tabacum*, which represents about $1.64 \times 10^5$ copies of its monomeric unit per haploid genome. This latter value is based on the DNA content of $1.5 \times 10^9$ bp per haploid genome of *N. tabacum* (Walbot and Goldberg 1979).

*Sequence analysis of cloned repeats*

The five cloned DNA repeat units (HRS60.1 to HRS60.5) were sequenced (Fig. 5). Although they had

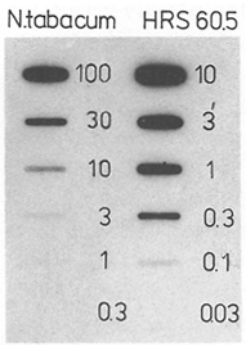


**Fig. 4.** Hybridization of HRS60.5 probe with nuclear tobacco DNA and HRS60.5 DNA. Serial dilutions of tobacco nuclear DNA (100, 30, 10, 3, 1, and 0.3 ng) and HRS60.5 DNA (10, 3, 1, 0.3, 0.1, and 0.03 ng) were blotted onto nitrocellulose membrane, hybridized with $^{32}$P-labelled probe and autoradiographed. Stringent hybridization conditions (0.1 × SSC, 65 °C) were used

been isolated from independent plasmid clones, sequencing showed that some of them formed a pair of identical "twins", i.e., HRS60.1 = HRS60.2 and HRS60.3 = HRS60.4. Therefore, only the repeat units HRS60.1, HRS60.3, and HRS60.5 were further considered. The repeat units HRS60.1, HRS60.3, and HRS60.5 differed slightly in their length, having 184 bp, 182 bp, and 183 bp, respectively. Homologies in nucleotides of

```
                10         20         30         40         50
HRS60.1  GATCCATCCG GGCCCAAGGC GGAAGGCATG GGCTATAGCA CACAAAAATT
HRS60.3  .......... .........T .......... .......... .........-
HRS60.5  .......... .......... .......... .......... ..........

                60         70         80         90        100
HRS60.1  TGGGACTCGC GCGGATTTGT AGTTTTATGG CTGTAAAATG CCAAAAAAAT
HRS60.3  .......... .........C .......... ..A......T ..-C......
HRS60.5  .......... .......... ..C....... .......... ..-...G...

               110        120        130        140        150
HRS60.1  ATAATGTGGT CATTTCGAGT GGGTGCAAAT TGTTTCTTAT GCCGTATTTG
HRS60.3  .......... .......G.. .....G.... .......... ..........
HRS60.5  ..T....... .......G.. .....G.... .........G ..........

               160        170        180
HRS60.1  ATGTCCGGGA CAAATATTAG GCGATTCCAC GACG
HRS60.3  .....G..T. .......... .......... ....
HRS60.5  .......A.. .......... .......... ....
```

**Fig. 5.** Nucleotide sequence organization of HRS60.1, HRS60.3, HRS60.5 DNA repeats. For HRS60.3 and HRS60.5, only the positions of nonhomologies are indicated. Identical nucleotides are represented by *dots*. A *dash* indicates a deletion

HRS60.1 with HRS60.3 and HRS60.3 with HRS60.5 were 94% and 95.7%, respectively. HRS60.3 had a 92.9% homology with HRS60.5. Non-homologies had the character of point mutations and deletions. They were dispersed randomly, perhaps with the exception of the "hotspot" region between the positions 90 and 97, where five non-homologies were accumulated. HRS60.1, HRS60.3, and HRS60.5 appeared to be basic repeat units with no long subrepeats; they contained only very short perfect inverted repeats (maximum 6 bp). Further experiments are in progress with the aim of isolating and characterizing some multimers of basic repeat units. Computer search for a homology of HRS60.1 with other sequenced plant tandem repeats (Appels et al. 1981; Deumling 1981; Peacock et al. 1981; Brennicke and Hemleben 1983; Capesius 1983; Kato et al. 1984; Barnes et al. 1985; Benslimane et al. 1986; Grellet et al. 1986; Martinez-Zapater et al. 1986; Leclerk and Siegel 1987; Wu and Wu 1987) was carried out. However, no long-range homology was found. Thus, HRS60.1 and the corresponding sequence family seem to be a DNA repeat not yet described. Nucleotide sequences of three members of HRS60-family are registrated in EMBL data library under the accession numbers X12489, X12490, and X12491.

*Acknowledgements.* The authors thank Dr. P. Karlovský for help in preparing the genomic DNA library, Mrs. L. Jedličková for excellent technical assistance, and Dr. J. Soška for critical reading of the manuscript.

## References

Appels R, Dennis ES, Smyth DR, Peacock WJ (1981) Two repeated DNA sequences from the heterochromatic region of rye (*Secale cereale*) chromosomes. Chromosoma 84:265–277

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

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

Benslimane AA, Dron M, Hartmann C, Rode A (1986) Small tandemly repeated sequences of higher plants likely originate from a tRNA gene. Nucleic Acids Res 14:8111–8119

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523

Brennicke A, Hemleben V (1983) Sequence analysis of the cloned *Cucumis melo* highly repetitive satellite DNA. Z Naturforsch Teil C 38:1062–1065

Capesius I (1983) Sequence of the cryptic satellite DNA from the plant *Sinapis alba*. Biochim Biophys Acta 739:276–280

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

Flavell RB (1983) Repeated sequences and genome architecture. In: Cifferi O, Dure III L (eds) Structure and functions of plant genomes. Plenum Press, New York, pp 1–14

Grellet F, Delcasso D, Panabieres F, Delseny M (1986) Organization and evolution of a higher plant alphoid-like satellite DNA sequence. J Mol Biol 187:495–507

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166:557–580

Karlovská L, Brzobohatý B, Reich J, Koukalová B, Bezděk M (1987) Transfer of *Drosophila melanogaster* transposable genetic element mdg-4 into plant cells. Folia Biol (Prague) 33:40–49

Kato A, Yakura K, Tanifuji S (1984) Sequence analysis of *Vicia faba* repeated DNA, the FokI repeat element. Nucleic Acids Res 12:6415–6426

Leclerk RF, Siegel A (1987) Characterization of repetitive elements in several *Cucurbita* species. Plant Mol Biol 8:497–507

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor/NY

Martinez-Zapater JM, Estelle MA, Somerville OR (1986) A highly repeated DNA sequence in *Arabidopsis thaliana*. Mol Gen Genet 204:417–423

Maxam AM, Gilbert W (1977) A new method for sequencing DNA. Proc Natl Acad Sci USA 74:560–564

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

Peacock W, 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

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

Thomas P (1980) Hybridization of denaturated DNA and small DNA fragments transferred to nitrocellulose. Proc Natl Acad Sci USA 77:5201–5205

Walbot V, Goldberg R (1979) Plant genome organization and its relationship to classical plant genetics. In: Hall TC, Davies J (eds) Nucleic acids in plants. CRC Press, Boca Raton/FL, pp 3–40

Wu T, Wu R (1987) A new rice repetitive DNA shows sequence homology to both 5S and tRNA. Nucleic Acids Res 15:5913–5923

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119