J Mol Evol (1995) 41:498–504

JOURNAL OF MOLECULAR EVOLUTION



# T-DNA Integrations in a New Family of Repetitive Elements of *Nicotiana tabacum*

**Clemens Suter-Crazzolara,[1]* Bretislav Brzobohaty,[1,2] Blanka Gazdova,[2] Jeff Schell,[1] Bernd Reiss[1]**

[1] Max-Planck-Institut für Züchtungsforschung, Carl-von-Linné-Weg 10, 50829 Köln, Germany
[2] Institute of Biophysics of the Academy of Sciences of the Czech Republic, Kralovopolska 135, 612 65 Brno, Czech Republic

Received: 26 August 1994 / Accepted: 5 January 1995

**Abstract.** A number of T-DNA insertions in the genome of *Nicotiana tabacum* were characterized. One class of integrations was found to have occurred in a new family of highly repetitive sequences. Three genomic regions (ecoA, ecoB, and ecoC) were isolated, all of which contain basic units of 180 bp, organized in direct tandem repeats. Several of the 180-bp elements contain an *Eco*RI recognition site within the repeating unit and are therefore named "eco repeats." All members of this family are weakly homologous in sequence to a previously described class of repeat elements which contained a *Bam*HI site (HRS60 repeat family), which suggests that both groups of sequences are of common evolutionary origin. The allotetraploid genome of *N. tabacum* is presumed to originate from the hybridization of two diploid genomes. The HRS60 elements previously described have been found exclusively in the genome of one of the ancestors, *N. sylvestris,* and in *N. tabacum* itself. Our DNA hybridization data suggest that the eco elements originate from the genome of the other ancestor, *N. tomentosiformis.* Whereas the HRS60 elements are transcriptionally silent, at least some eco elements appear to be transcribed.

**Key words:** T-DNA — Integration—Tobacco — Repetitive sequences — Transcription

## Introduction

Genomes of higher organisms generally contain different types of repetitive sequence elements. These elements may be present in several millions of copies per genome and can constitute a major portion of the total DNA of a species. Repetitive DNA is either found dispersed throughout the genome or arranged in tandem arrays. The latter type is higher repetitive and consists of clusters of relatively simple, short sequence units. Highly repetitive DNA or satellite DNA is frequently found in telomeric and centromeric heterochromatin and is usually transcriptionally inactive. However, exceptions have been reported (reviewed by Jelinek and Schmid 1982).

The presence of repetitive sequences in all eukaryotic organisms, similarities in repeat unit size, and chromosomal localization in different species suggest that these sequences serve a function (reviewed by Lapitan 1992). The association of highly repetitive DNA with heterochromatin is consistent with a role as a structural component of chromosomes. The length of the repeat units of satellites from different species coincides with the size of DNA occupied by a nucleosome. Based on this observation, Martinez-Zapater et al. (1986) suggested that these repeats might help to organize nucleosomes. Since repeats are often found at telomeric positions, a role in recombination between satellite arrays of nonhomologous chromosomes (Flavell 1986) or in stabilization of telomeres (reviewed by Blackburn 1991) has been proposed. In contrast, repetitive elements have also been regarded as having no functional relevance and to accumulate simply as selfish DNA (Von Sternberg et al. 1992).

Nucleotide sequence data reported here will appear in the EMBL, Genbank, and DDBJ Nucleotide Sequence Databases under the accession numbers ecoA: X78472; ecoB X78473; ecoC: X78474

* Present address: Universität Heidelberg, Institut für Anatomie und Zellbiologie, Im Neuenheimer Feld 307, 69120 Heidelberg, Germany
*Correspondence to:* B. Reiss

In plants, different copies of the repeating unit of satellite DNAs are highly conserved within one species (at least 90% identity). In contrast, these DNAs are often highly diverged across species within the same family. Although a number of these repeats have been well analyzed at the DNA sequence level, their overall organization is less well understood. The tomato TGR1 repeat was shown to be organized in 27 clusters, distributed over different chromosomes. The size of the clusters ranges from 150 copies to 6,000 copies (reviewed by Lapitan 1992).

The genome of tobacco, *Nicotiana tabacum,* also contains a number of repetitive sequences. One of them, the HRS60 family, has recently been identified and characterized (Koukalova et al. 1989). This repeat was isolated by *Bam*HI restriction endonuclease digestion and consists of a basic unit of 180 bp, organized in direct tandem repeats. The diploid genomes of *N. tomentosiformis* and *N. sylvestris* are assumed to have hybridized relatively recently in evolutionary terms to form the allotetraploid tobacco genome (Okamuro and Goldberg 1985). The HRS60 family, originally identified in *N. tabacum,* is present in *N. sylvestris* but absent from *N. tomentosiformis* (Koukalova et al. 1990). This indicates that this repetitive sequence was contributed to the tobacco genome by only one of the progenitors of tobacco. In this contribution, we describe a new repetitive sequence in tobacco, the "eco repeat." These elements are related in sequence to the HRS60 family and are present in *N. tabacum* and *N. tomentosiformis* but absent from *N. sylvestris.*

## Materials and Methods

Plant transformation experiments and the vectors used for *Agrobacterium*-mediated transformation were essentially as described by Koncz et al. (1989). The experimental setup which gave rise to the series of T-DNA insertions was described in Schell et al. (1993).

Enzymes were purchased from Gibco BRL, Boehringer Mannheim, USB, and AGS and were used according to the manufacturers' guidelines. Chemicals were purchased from Merck, Roth, and Sigma. Nucleotides were obtained from Pharmacia.

Genomic DNA was isolated from transformants of *N. tabacum* cv SR1, *N. sylvestris,* and *N. tomentosiformis* according to Murray and Thompson (1980). Total RNA was isolated by the method of Logemann et al. (1987). PolyA$^+$ RNA was purified by one passage through an oligo(dT)-cellulose column. All DNA and RNA handling was carried out according to Sambrook et al. (1989).

For the isolation of genomic sequences flanking T-DNA integrations by plasmid rescue (plasmid pEcoC), total genomic DNA was digested with *Eco*RI, religated in a large volume (Koncz et al. 1989), and transformed into *E. coli* MC1061.

The plasmid pEcoA insert was obtained by direct PCR (Saiki 1990). Oligonucleotides "a" (5′ CGG GAT CCA TAT TGA CCA TCA TAC TCA TTG) and "b" (5′ GCA GGG TAC CCG GGG ATC AGA TTG TCG) are complementary to sequences within the T-DNA close to the left and right borders and were thus suitable to amplify sequences between tandem integrations (Fig. 1). These oligonucleotides contain additional recognition sites for either *Kpn*I or *Bam*HI (for "b" and "a") at their 5' ends, which were used for cloning.

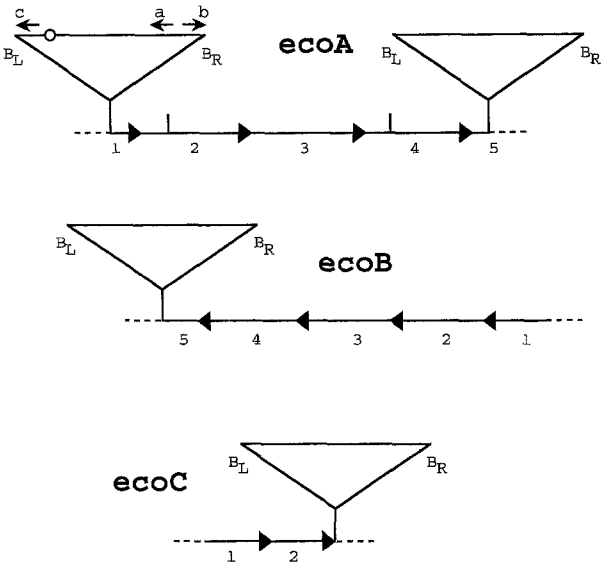


**Fig. 1.** Diagrammatic summary of eco elements and sites of T-DNA integration. *Large triangles* represent T-DNA inserts. The position of the right ($B_R$) and left ($B_L$) T-DNA borders is given. *Solid horizontal arrows* denote position and orientation of repetitive elements. *Vertical lines* mark the *Eco*RI restriction sites in ecoA2 and ecoA4. All other elements contain no intact *Eco*RI site. Unknown genomic DNA sequences are shown as *dashed lines.* T-DNA insertions were isolated by PCR with the help of primers (oligonucleotides *a, b,* and *c*) and plasmid rescue taking advantage of a pBR322 replicon (*small circle*) present close to the left border.

The plasmid pEcoB insert was obtained by inverse PCR (Triglia et al. 1988). Genomic DNA was digested with *Hae*III, religated in a large reaction volume, and digested with *Eco*RV. *Eco*RV cuts at the right border of the T-DNA between primer pair "b" and "c" (5′ CGG AAT TCG GCG AGT TCT GTT AGG TCC TCT A) (Fig. 1). Oligonucleotide "c" contains an *Eco*RI site at its 5′ end and the corresponding sequence is located 500 bp upstream of oligonucleotide "b," which anneals to the opposite strand. The DNA was concentrated by ethanol precipitation and amplified using primer pair b and c. After purification, the PCR products were digested with *Kpn*I and *Eco*RI.

For all PCR reactions, the equivalent of 100 ng genomic DNA in a 20-µl reaction volume was used. The reactions were carried out in 2.5 mM $MgCl_2$, 0.01% Triton X-100, 10 mM Tris-HCl, pH = 8.3, 50 µM of each dNTP, and 2 µM of each oligonucleotide. The reaction conditions were: 30 cycles; 1 min at 93°C, 1 min at 65°C, and 2 min at 72°C. Purified PCR products were digested with the restriction enzymes indicated above and ligated to appropriately cut pUC19 (Yanisch-Perron et al. 1985). Ligation mixtures were transformed in MC1061, plasmid DNA was purified, and DNA sequences were determined by the chain termination method as described (Sambrook et al. 1989). Sequences were analyzed with the GCG programs (Devereux et al. 1984).

For Southern analysis, DNA was separated by agarose gel electrophoresis, transferred to Zetaprobe membranes (Biorad), and hybridized for 16 h to the radioactively labeled probe in 50% formamide, 0.12 M $Na_2HPO_4$ pH 7.2, 0.25 M NaCl, and 7% SDS at 42°C. To prepare radioactive probes, the inserts of pEcoA, pEcoB, pEcoC, and pHRS60.1 were purified and labeled by random priming (Feinberg and Vogelstein 1983). These probes had specific activities of $10^9$ cpm/µg DNA. After hybridization, the filters were washed for 30 min at 55°C in 0.5 × SSPE and 1% SDS and exposed to Kodak X-AR film. The exposure time was 16 h unless otherwise indicated. Radioactive probes were removed from the filters by washing in 0.01 × SSPE and 1% SDS at 90°C for 2 h. The filters were reexposed for 1 week to X-ray film to ensure that no residual signal was left.

For Southern blots with genomic DNA, 10 µg of DNA was digested

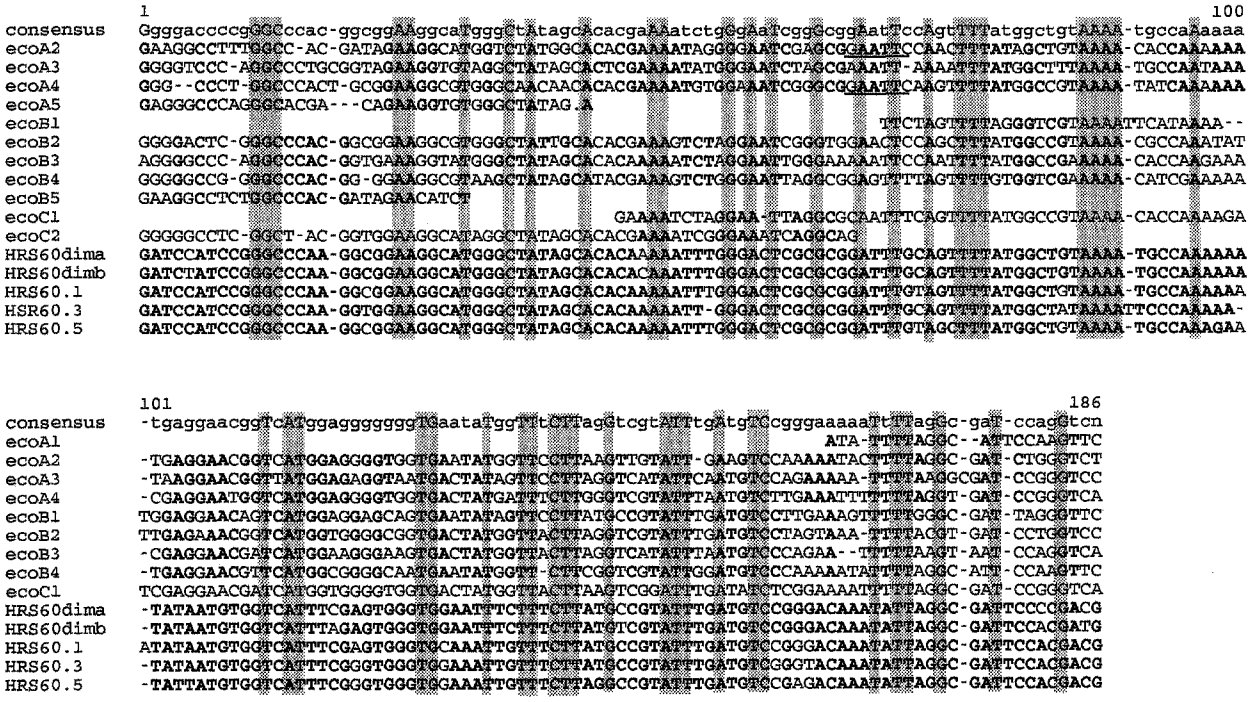


**Fig. 2.** Sequence comparison of different members of the hrs family. Sequences were aligned to the HRS60 sequences (Koukalova et al. 1989; Mataysak et al. 1989). The *Bam*HI site is located at position *1* and the *Eco*RI site around position *64* (*underlined*). These sites are cryptic in most elements. A consensus sequence of all repeats is shown in the *first lines.* Residues which are conserved in all sequences are indicated by *gray boxes* and *capital characters* in the consensus sequence. *Lowercase* characters in the consensus sequence denote residues with the largest frequency. Residues conserved within groups are shown in *bold* characters. *Hyphens* indicate deletions. Due to the procedure used to isolate the repeats, only partial sequences for some elements were obtained (as indicated in Fig. 1).

with *Eco*RI, *Bam*HI, and *Hind*III, respectively, and the fragments were separated by 0.8% agarose gel electrophoresis. For the Southern blots with plasmid DNA, pEcoA, pEcoB, pEcoC, and pHRS60.1 were digested with *Eco*RI, *KpnI*/*Bam*HI, *Eco*RI/*Sau*3AI, and *Bam*HI, respectively, to release fragments of approximately the size of the monomeric unit. The fragments were separated by 2.5% agarose gel electrophoresis and blotted to Zetaprobe. Hybridization and washing conditions were identical for genomic and plasmid blots.

For Northern analysis, poly$A^+$ RNA (10 μg per lane) was separated by 0.8% agarose–formaldehyde gel electrophoresis according to Sambrook et al. (1989) and transferred to Hybond filter (Amersham) according to the manufacturer's guidelines. Hybridization, washing, exposure, and removal of residual radioactive probe were as described for Southern hybridizations.

## Results

### *Identification of the Eco Repeat Family of* N. tabacum

As part of an ongoing project, we systematically characterized T-DNA integration sites in the genome of *N. tabacum,* transformed by cocultivation of protoplasts with *Agrobacterium tumefaciens.* Tagged genomic sequences were isolated by three different methods. Direct PCR (Saiki 1990) using primers specific for the right and left borders of the T-DNA were used to amplify a genomic region localized between two T-DNA inserts integrated in the same relative orientation (ecoA, Fig. 1). A genomic sequence flanking the right border of another T-DNA integration site was obtained by inverse PCR (Triglia et al. 1988) using two primers specific for sequences at the right border of the T-DNA (ecoB, Fig. 1). The third integration site was obtained by plasmid rescue (Koncz et al. 1989) taking advantage of an *E. coli* replicon located close to the left border of the T-DNA (ecoC, Fig. 1).

Characterization of the T-DNA integration sites by nucleotide sequence analysis and Southern blotting hybridization revealed that in several instances the T-DNA had inserted into repetitive sequences. In particular, the Southern analysis demonstrated the repetitive character of this DNA and showed that several of these repeats contain a recognition site for *Eco*RI within the repeating unit. The tagged genomic regions were therefore named ecoA, ecoB, and ecoC.

### *Sequence Analysis*

Inspection of the DNA sequences showed the tandem arrangement of basic elements 180 bp long (Figs. 1 and 2). For ecoA, the isolated genomic region consists of a stretch of five such repeated elements (ecoA1 to ecoA5). As shown in Fig. 1, two copies of the T-DNA are integrated into this genomic region in close proximity. The

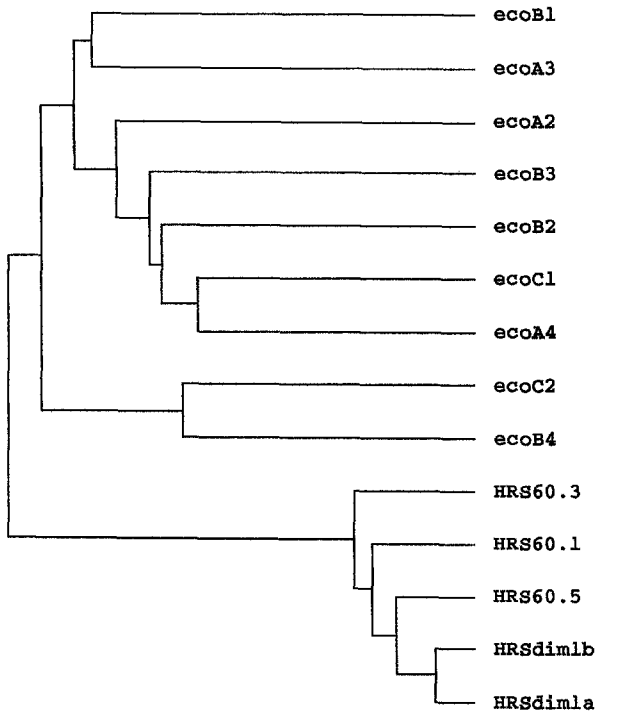


**Fig. 3.** Sequence similarity of the repetitive elements. Sequences were aligned with the GCG Pileup program. Sequences shorter than 50 bp (ecoA1, ecoA5, and ecoB5) were not considered in this alignment and only partial sequences of ecoB1, ecoC1, and ecoC2 were available for this analysis. The degree to which one family member is similar to others is indicated by the number of branch points between them. Members more closely related are separated by fewer branches.

fragment carrying the ecoB genomic region consists of a cluster of five repetitive elements which were named ecoB1 to ecoB5. Only two adjacent elements, ecoC1 and ecoC2, were isolated from the ecoC repeat. The T-DNAs are integrated into the ecoB repeat cluster in the opposite relative orientation compared to ecoA and ecoC. Only two of the repeat elements, ecoA2 and ecoA4, actually contain an *Eco*RI recognition site. The other units presumably lost the recognition site by sequence drift.

By comparing the sequences of these elements with each other and with the previously described HRS60 elements (Fig. 2), strict conservation of 54 out of 180 residues in all repeats was observed. These residues are separated by less-conserved or nonhomologous regions. The least-conserved region is found around the position of the *Bam*HI site present in the HRS60 element. The majority of base changes between repeats can be explained by transitions indicating that these repeats may have originated from a single progenitor by simple mutations. Cytosines appear to be underrepresented in this family, whereas AT residues are relatively frequent (49–59%).

Individual elements within one cluster are only 76–80% homologous to one another. Elements of the ecoA cluster are also moderately homologous to members of the ecoB and ecoC cluster (between 70% and 81%). Homology between different members of ecoB and ecoC is in the same range. In contrast, HRS60 elements are more than 90% homologous to one another and all eco elements are clearly less homologous to the HRS60 repeats (about 65%) than to one another. A graphic display of these relationships is shown in Fig. 3.

### *Genomic Organization of the Repeats*

Sequence analysis of the ecoA genomic region revealed the presence of two *Eco*RI sites (Fig. 2) while no recognition sequence for *Hind*III or *Bam*HI was found. In order to analyze the genomic organization of the eco repeats in more detail, genomic tobacco DNA was digested with either *Bam*HI or *Eco*RI or *Hind*III. The fragments generated were separated and hybridized to the radioactively labeled ecoA, ecoB, or ecoC probe by Southern blotting (Fig. 1). Results are shown in Fig. 4A (panels a, b, c). In all lanes containing *Eco*RI-digested tobacco DNA, the ladderlike pattern characteristic for tandem repeats was obtained. The weakest hybridization signal was obtained with the ecoB probe. These results indicate that these classes of repeats are present in clusters in the genome and indeed contain an *Eco*RI restriction site within some of the repeat units. However, the monomeric unit of 180 bp is strongly underrepresented in all three cases and the most abundant species appears to be a trimer. Most of the genomic DNA hybridizing to the probes is found in the form of repeat structures. Only rather weak hybridization is found to the bulk of fragments produced in the digest. In contrast, most of the signal is found in this fraction when the DNA is digested with *Bam*HI or *Hind*III. Therefore, the majority of the repeat units contain no restriction sites for either enzyme. However, a minor fraction of *Bam*HI-digested DNA yields a ladderlike pattern, similar to that seen in the *Eco*RI digest. This pattern is shifted to higher multimer forms. For comparison, the same filter was probed with the HRS60.1 element (Koukalova et al. 1989). With this probe, the ladderlike pattern is most clearly visible in *Bam*HI-digested genomic DNA (Fig. 4A, panel d). However, it is also weakly present in the *Eco*RI digest (not clearly visible in Fig. 4A, panel d, due to overexposure).

In order to determine the degree of cross-hybridization between the individual repetitive elements in these Southern blots, plasmids carrying the ecoA, ecoB, ecoC, and HRS60.1 elements were digested with restriction enzymes to obtain fragments of approximately the size of a monomeric unit. A Southern filter prepared with these digests was hybridized with the ecoA probe under the same conditions as the genomic Southern blots (Fig. 4B). No cross-hybridization with the HSR60.1 repeat

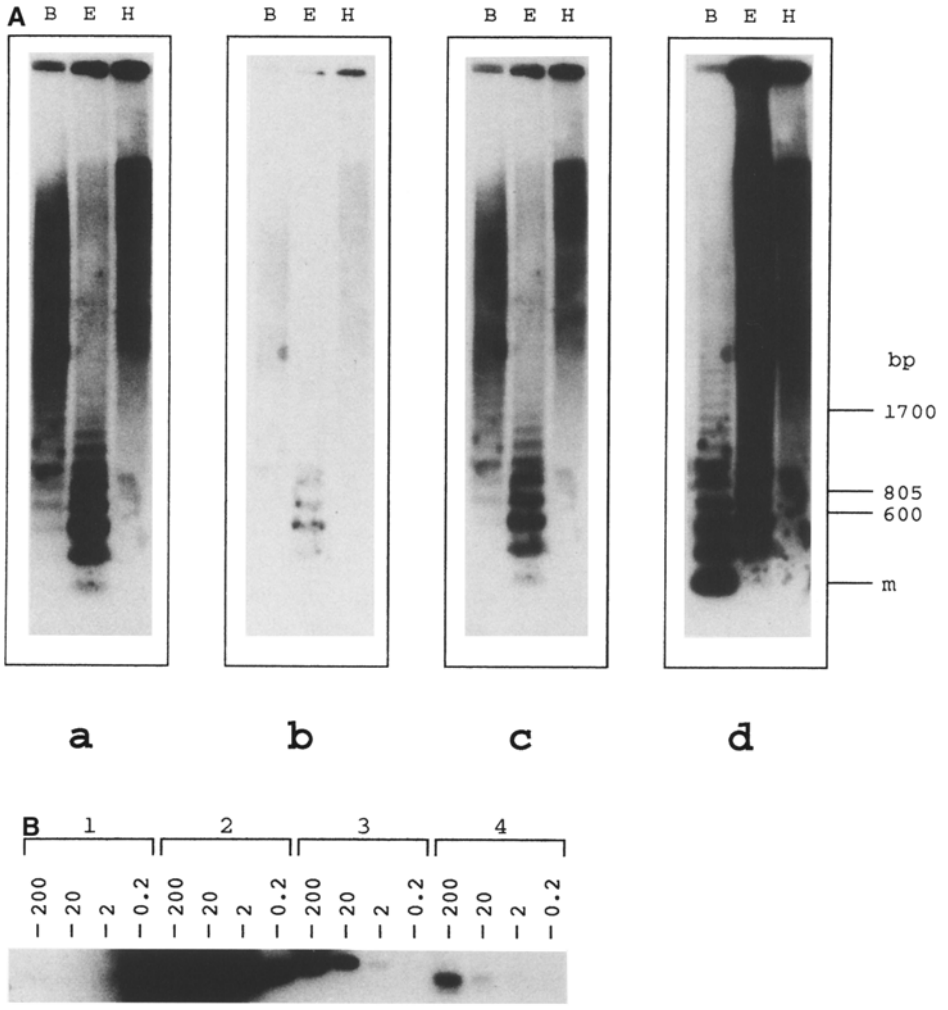


**Fig. 4.** Southern analysis. **A** Genomic RNA was isolated from tobacco SR1 and digested with either *Bam*HI (*B*), *Eco*RI (*E*), or HindIII (*H*). Fragments were separated by gel electrophoresis and blotted to a Nylon membrane. This filter was hybridized to the following radioactively labeled probes: panel a: ecoA, *b:* ecoB, *c:* ecoC, and *d:* HRS60.1. Positions of size markers are indicated on the right. *"m"* marks the position of the monomer. **B** Plasmids pHRS60.1 (panel *1*), pEcoA (panel *2*), pEcoB (panel *3*), and pEcoC (panel *4*) were digested with restriction enzymes to release fragments of approximately the size of the monomeric unit of each family. These fragments were serially diluted, separated by gel electrophoresis, and blotted to a Nylon membrane. The filter was hybridized with the radioactively labeled ecoA probe. The amount of fragment loaded per lane is indicated (200, 20, 2, 0.2 pg).

was detected under these conditions. The ecoB and ecoC repeats gave a 100- and 1,000-times-weaker signal than the ecoA repeat, respectively. No cross-hybridization was observed with the ecoA, ecoB, or ecoC elements if the HRS60.1 element was used as probe (data not shown).

### *Origin of the eco Repetitive Elements*

To determine the evolutionary origin of the eco repeat family, genomic DNA of *N. sylvestris* and *N. tomentosiformis* was digested with either *Bam*HI or *Eco*RI and hybridized after blotting with the ecoA probe (Fig. 5). A strong hybridization signal, similar to the one observed with the same probe in tobacco, was obtained with *N. tomentosiformis* DNA. In contrast, only a very weak hybridization signal was observed with *N. sylvestris* DNA. This indicates that the eco elements are present in the genomes of *N. tomentosiformis* and *N. tabacum,* but absent from the genome of *N. sylvestris.* A similar result was obtained when ecoB was used as a probe (data not shown).

### *Eco Repeats Are Transcribed in the Tobacco Genome*

In order to determine whether eco repeats are represented in transcribed RNA, Northern blotting experiments with polyA$^+$ RNA from tobacco leaves were performed. Using a labeled probe consisting of the entire ecoA repeat, two prominent bands corresponding to RNA sizes of approximately 5,000 and 1,500 bases and some minor bands were detected on this blot (Fig. 6, panel 1). In

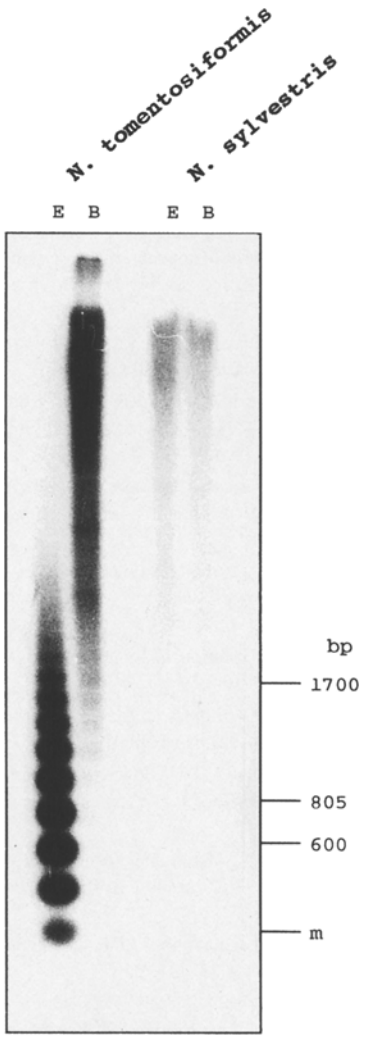


**Fig. 5.** Southern blot to analyze the evolutionary origin of the eco repeat. Genomic DNA of *N. sylvestris* and *N. tomentosiformis* was digested either with *Bam*HI (*B*) or *Eco*RI (*E*). Fragments were separated by gel electrophoresis and blotted to a Nylon membrane. The filter was hybridized with the ecoA probe. Positions of a DNA size marker are indicated on the right; *"m"* denotes the position of the monomer.

contrast, only background signals were observed when the same filter was probed with the HRS60.1 element (Fig. 6, panel 2). The latter signals correspond to the position of rRNAs, as was confirmed in an independent hybridization (not shown). To estimate the relative abundance of the transcripts hybridizing with the ecoA repeat, the same filter was hybridized with additional probes. Genes for actin, the proteasome-subunit C3, and glyceraldehyde-3-phosphate dehydrogenase (GaPDH) are transcribed at comparable levels in most tissues of many organisms. Tobacco cDNA clones (unpublished) for these genes were used as probes and detected the expected transcripts (Fig. 6, panels 3–5). Based on these hybridizations, the transcripts detected with the ecoA repeat probe appear to be quite abundant, but less so than actin and GaPDH mRNAs.

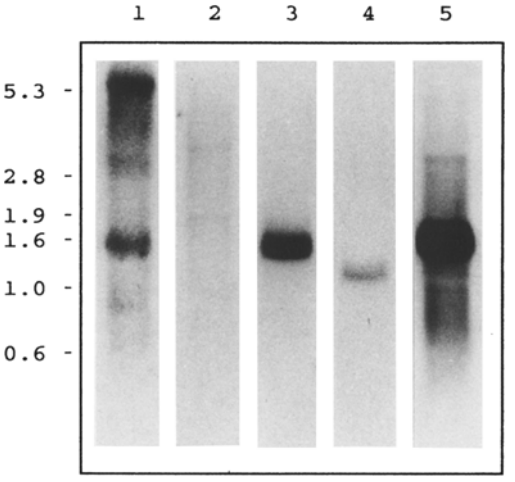


**Fig. 6.** Northern blot hybridizations to analyze transcriptional activity of the eco repeats. Tobacco-leaf polyA$^+$ RNA was separated by gel electrophoresis and transferred to a Nylon filter. This filter was hybridized to the following probes and exposed for the time indicated in parentheses: panel *1:* ecoA (4 days); *2:* HRS60.1 (3 days); *3:* tobacco actin cDNA (2 days); *4:* tobacco proteasome subunit C3 (3 days); and *5:* tobacco GaPDH (6 days). Positions of a RNA size marker are shown on the left side of the blot.

## Discussion

A new family of repetitive elements of *N. tabacum* has been identified in the course of an analysis of T-DNA integrations. Three genomic regions containing eco repeats, ecoA, ecoB, and ecoC, with 12 different members in total, were characterized. The repeats are organized in tandem arrays of 180-bp monomeric units. These elements are highly repetitive, as suggested by signal strengths obtained in Southern blotting experiments. The eco repeats are moderately conserved in sequence to each other, irrespective of which cluster they were derived from. In addition, only a few *Eco*RI restriction sites were actually found in the sequence of individual repeats, and in Southern blots, the most prominent species in the *Eco*RI digest was the trimer. Therefore, these repeats appear to be rather degenerate. Based on DNA sequence analysis, the family of eco repeats is related to, but clearly distinct from, the previously identified HRS60 family. Although according to DNA sequence homology comparison the eco repeats fall in one group, this family appears to consist of at least three distinct classes which can be discriminated on Southern blots.

The relatedness of the eco and HRS60 repeats suggests that both families evolved from an archetypal repeat which originated in an ancestor of the Nicotianaceae. The presence of *Bam*HI sites in a minority of eco elements and *Eco*RI sites in some HRS60 repeats indicates that this ancestor may have had both sites. Since eco and HRS60 repeats are found in different progenitors of tobacco, both families of repeats evolved independently in *N. sylvestris* and in *N. tomentosiformis*. The

hybridization of these two species which led to the formation of tobacco presumably combined these repeats in the same genome.

Tandemly repeated sequences are usually transcriptionally silent (Lapitan 1992). However, two ecoA-repeat-specific RNAs were found in tobacco-leaf RNA. These RNAs were larger in size than the corresponding, isolated, genomic region. In addition, this region did not contain any significant open reading frames. Therefore these RNAs might not be derived directly from ecoA, but instead might be detected due to cross-hybridization with a rather abundant mRNA. Also, ecoA sequences might be present in active genes and could be cotranscribed accidentally. At present, the exact origin of these transcripts and their function remain to be determined—for example, by isolation and sequencing of corresponding cDNAs.

Whereas repetitive elements are usually detected via a diagnostic restriction-enzyme recognition site, the eco repeats were found by T-DNA tagging. T-DNA in general integrates preferentially into transcribed regions of the genome (Koncz et al. 1989). Therefore, transcriptional activity of the eco repeats may have contributed to T-DNA integration at these sites. It should be noted that only a fraction of the integration sites analyzed were insertions in eco repeats, the other integrations being in single-copy DNA. The ecoB and ecoC sequences most likely arise from the same T-DNA integration event (data not shown).

While this work was ongoing, six members of another family of repetitive elements in tobacco, GRS, were identified by restriction-enzyme digestion (Gazdova et al. 1995). The sequence identities of members of GRS with members of the eco repeats are in the same range (74–79%) as the identities between elements of the eco family (70–81%). Therefore, it is likely that elements of eco and GRS both represent the same family of repeats.

*Acknowledgments.* We thank Dr. Josefa-Maria Balsalobre-Balibrea for providing the Northern blot filter, Dr. Klaus Palme and Stefanie Rechmann for DNA sequencing, Astrid Blau, Petra Krieger, Gisela Simons, and Edelgard Wendeler for excellent technical assistance, and the BMFT for financial support of this project (0316101A). Clemens Suter-Crazzolara acknowledges support by fellowships from the EC (B/BIOT-913008) and the Max-Planck-Society.

## References

Blackburn E (1991) Structure and function of telomeres. Nature 350: 569–573

Devereux J, Haeberli L, Smithies T (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12(1): 387–395

Feinberg AP, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Flavell, R. (1986) Repetitive DNA and chromosome evolution in plants. Philos Trans R Soc Lond Biol 312:227–298

Gazdova B, Siroky J, Fajkus J, Brzobohaty B, Kenton A, Bezdek M. (1995) Characterisation of a new family of tobacco highly repetitive DNA, GRS, specific for the N. tomentosiformis genomic component. Chromosome Res 3:245–254

Jelinek WR, Schmid CW (1982) Repetitive sequences in eukaryotic DNA and their expression. Annu Rev Biochem 51:813–844

Koncz C, Martini N, Mayerhofer R, Koncz-Kalman Z, Körber H, Redei PG, Schell J (1989) High-frequency T-DNA-mediated gene tagging in plants. Proc Natl Acad Sci U S A 86:8467–8471

Koukalova B, Reich J, Matyasek R, Kuhrova V, Bezdek M (1989) A BamHI family of highly repeated DNA sequences of *Nicotiana tabacum*. Theor Appl Genet 78:77–88

Koukalova B, Reich J, Bezdek M (1990) A *Bam*HI family of tobacco highly repeated DNA: a study of its species specificity. Biol Plant (Praha) 32:445–449

Lapitan NLV (1992) Organization and evolution of higher plant nuclear genomes. Genome 35:171–181

Logemann J, Schell J, Willmitzer L (1987) Improved method for the isolation of RNA from plant tissues. Anal Biochem 163:16–20

Martinez-Zapater J, Estelle M, Sommerville C (1986) A highly repeated DNA sequence in Arabidopsis thaliana. Mol Gen Genet 204:417–423

Mataysak R, Koukalova B, Reich J (1989) Isolation and sequencing of HRS60diml, a dimeric member of the HRS60-family of a tobacco DNA repeat. Nucleic Acids Res 17:4377

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

Okamuro JK, Goldberg RB (1985) Tobacco single-copy DNA is highly homologous to sequences present in the genomes of its diploid progenitors. Mol Gen Genet 198:290–298

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Saiki RK (1990) Amplification of genomic DNA. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols: a guide to methods and applications. Academic Press, San Diego, pp 13–20

Schell J, Palme K, Reiss B, Walden R (1993) Recent advances in the search for genes involved in the mechanisms of action of auxin and cytokinins. J Plant Res Special Issue 3:221–227

Triglia T, Peterson MG, Kemp DJ (1988) A procedure for in vitro amplification of RNA segments that lie outside the boundaries of known sequences. Nucleic Acids Res 16:8186

Von Sternberg RM, Novick GE, Gao G-P, Herrera RJ (1992) Genome canalization: the co-evolution of transposable and interspersed repetitive elements with single copy DNA. Genetica 86:215–246

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