Chromosome Research 1997, **5**, 561–568

Two tandemly repeated telomere-associated sequences in *Nicotiana plumbaginifolia*

Chung-Mong Chen, Chi-Ting Wang, Chung-Ju Wang, Chia-Hsing Ho, Yen-Yu Kao & Chi-Chang Chen

Received 5 June 1997; received in revised form and accepted for publication by Pat Heslop-Harrison 17 September 1997

Two tandemly repeated telomere-associated sequences, NP3R and NP4R, have been isolated from *Nicotiana plumbaginifolia*. The length of a repeating unit for NP3R and NP4R is 165 and 180 nucleotides respectively. The abundance of NP3R, NP4R and telomeric repeats is, respectively, 8.4×10^4, 6×10^3 and 1.5×10^6 copies per haploid genome of *N. plumbaginifolia*. Fluorescence *in situ* hybridization revealed that NP3R is located at the ends and/or in interstitial regions of all 10 chromosomes and NP4R on the terminal regions of three chromosomes in the haploid genome of *N. plumbaginifolia*. Sequence homology search revealed that not only are NP3R and NP4R homologous to HRS60 and GRS, respectively, two tandem repeats isolated from *N. tabacum*, but that NP3R and NP4R are also related to each other, suggesting that they originated from a common ancestral sequence. The role of these repeated sequences in chromosome healing is discussed based on the observation that two to three copies of a telomere-similar sequence were present in each repeating unit of NP3R and NP4R.

Key words: fluorescence *in situ* hybridization, highly repeated tandem sequence, *Nicotiana plumbaginifolia*, telomere-associated sequence

Introduction

Although the telomeres of most plant species are composed of tandemly repeated sequence blocks of the oligonucleotide T_3AG_3 (Richards & Ausubel 1988, Fuchs *et al*. 1995), the sequences adjacent to the telomere repeats (telomere-associated sequences, TASs) are diverse. In *Arabidopsis*, while single-copy sequences are identified in subtelomeric regions of most chromosomes (Richards *et al*. 1992), tandemly repeated rDNA sequences are also found adjoining the telomeres of chromosomes 2 and 4 (Copenhaver & Pikaard 1996). In plants with larger genomes, such as *Aegilops speltoides* (Anamthawat-Jonsson & Heslop-Harrison 1993), *Aegilops squarrosa* (Rayburn & Gill 1987), *Allium cepa* (Barnes *et al*. 1985), *Hordeum vulgare* (Brandes *et al*. 1995), *Nicotiana tabacum* (Kenton *et al*. 1993) and *Secale cereale* (Vershinin *et al*. 1995), the subtelomeric regions are often composed of long arrays of tandemly repeated sequences. These tandemly repeated sequences often exhibit a high variation of copy number even between closely related species or lines (Ganal *et al*. 1988, Anamthawat-Jonsson & Heslop-Harrison 1993), making them good markers for genetic mapping (Burr *et al*. 1992, Gardiner *et al*. 1996). Furthermore, the spatial and quantitative variation of tandemly repeated TASs also occurs between chromosomes in the same genome, making it possible to distinguish individual chromosomes by *in situ* hybridization using these sequences as probes (Lapitan *et al*. 1989, Ganal *et al*. 1991, Castilho & Helsop-Harrison 1995).

Previously, we constructed a *Nicotiana plumbaginifolia* genomic library in a yeast artificial chromosome (YAC) (Chen *et al*. 1996). Several hundreds of restriction fragment-length polymorphic (RFLP) markers have been identified from the YAC, cDNA and random genomic libraries of *N. plumbaginifolia*, and construction of a genetic map based on these markers is under way (Chen *et al*. unpublished data). In order to isolate markers for mapping chromosome ends, we have cloned several TASs from *N. plumbaginifolia* using a modified Vectorette polymerase chain reaction (PCR) method (Riley *et al*. 1990, Kilian & Kleinhofs 1992). Two of them have been characterized and reported herein.

Materials and methods

Preparation of plant and plasmid DNA

Nuclear (Jofuku & Goldberg 1988) and total genomic (Suen *et al*. 1997) DNA was isolated from the leaves of *N. plumbaginifolia*. Plasmid DNA was isolated using an alkaline lysis method (Sambrook *et al*. 1989).

Isolation of TAS clones

To amplify TASs from *N. plumbaginifolia* genomic DNA using the polymerase chain reaction (PCR), oligonucleotides VT2

C.-M. Chen (corresponding author) and C.-T. Wang are at the Institute of Botany, Academia Sinica, Taipei, Taiwan, Republic of China, 115. Tel: (+886 2 7899 590 ext. 216; Fax: (+886 2 7827 954; Email: bocmchen@ccvax.sinica.edu.tw. C.-J. Wang, Y.-Y. Kao and C.-C. Chen are at the Department of Botany, National Taiwan University, Taipei, Taiwan, Republic of China. C.-H. Ho is at the Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan, Republic of China.

(5′-CTCTCCCTTCTCGAATCGTAACCGTTCGTACGAGAATC-GCTGTCGTCGACTTG-3′) and VB2 (5′-CAAGTCGACGA-CGCTGTCTGTCGAAGGTAAGGAACGGACGAGAGAAGGG-3′) were annealed to generate a vectorette (VT2/VB2) with a loop in the middle and a *Sal*I site (underlined) near its blunt end. *Hae*III-digested *N. plumbaginifolia* genomic DNA fragments rich in telomeric sequence (0.1 μg) were ligated to the blunt end of annealed vectorette (5 pmol), and TASs were PCR amplified in the presence of 33 pmol each of telomere primer (TP, 5′-CCGAATTCAACCCTAAACCCTAAACCCTAAACCC-3′) and vectorette primer (VP, 5′-CGAATCGTAACCGTTCG-TACGAGAATCGCT-3′) in a total volume of 100 μl (Kilian & Kleinhofs 1992). The *Eco*RI site in the TP (underlined) and the *Sal*I site (underlined) in the annealed vectorette were designed for subsequent cloning. The PCR product was digested with *Eco*RI and *Sal*I, ligated with pUC18, transformed into *Escherichia coli* HB101 and selected for colonies resistant to ampicillin. Amp^R colonies were screened by colony hybridization using $(T_3AG_3)_4$ as a probe. Inserts in the plasmids containing plant telomeric sequences were sequenced in both strands using an ABI 373 automatic sequencer. Based on the sequence information, the non-telomeric region of each clone was PCR amplified and subcloned into pUC18. The subcloned non-telomeric fragments were isolated and used as probes for genomic Southern hybridization and for screening a *Hae*III genomic library of *N. plumbaginifolia* with inserts ranging from 0.1 to 0.5 kb in length. Inserts of the plasmids isolated from positive clones were reconfirmed by Southern hybridization and then were sequenced on both strands. The sequences were subjected to a search for homologous sequences using the BLAST program at the GenomeNet in Tokyo Center.

Determination of the copy number of repeated sequences

Serial dilutions of plant nuclear DNA and plasmid containing the repeated sequence were dot blotted onto a Hybond-N membrane and hybridized with the corresponding probe (Koukalova *et al*. 1989). After exposure to the radiographic film, densities of the individual spots were measured and the amount of both plant nuclear and plasmid DNA giving the same density of hybridization were calculated. On the basis of these values, the fraction of nuclear DNA of *N. plumbaginifolia* homologous to the corresponding repeated sequence was estimated.

Colony, dot and Southern hybridization

Digestion of DNA, agarose gel electrophoresis, colony and Southern hybridization were performed as described by Sambrook *et al*. (1989). For preparation of probes, oligonucleotide $(T_3AG_3)_4$ was end-labelled with digoxigenin (DIG)-11-ddUTP using terminal transferase, and the fragments corresponding to the non-telomeric regions of TAS clones were labelled with DIG-11-dUTP using a random-primed DNA labelling kit (Boehringer Mannheim, Germany). For DNA/DNA hybridizations and washings, two different stringencies were used: 55°C, 1 × SSPE (0.18 M NaCl, 10 mM NaH_2PO_4, pH 7.4, and 1 mM EDTA), 0.1% sodium dodecyl sulphate (SDS) for $(T_3AG_3)_4$; and 65°C, 0.2 × SSPE, 0.1% SDS for probes derived from the non-telomeric regions of TAS clones. The hybridization signal was detected by chemiluminescent reaction using CSPD ($C_{18}H_{20}ClO_7PNa_2$) as substrate (Boehringer Mannheim, Germany).

In situ hybridization

Root tips excised from *in vitro*-cultured haploid *N. plumbaginifolia* plants (Chen *et al*. 1985) were treated with 2 mM 8-hydroxyquinoline at 18–20°C for 2.5 h, fixed in ethanol–glacial acetic acid (3:1) overnight and stored in 70% ethanol at −20°C. The root tips were digested in 2% (w/v) cellulase Onozuka R10 (Yakult Honsha, Japan) and 1% (w/v) macerozyme Onozuka R10 (Yakult Honsha, Japan) in citrate buffer (4 mM citric acid and 6 mM sodium citrate, pH 4.8) at 25°C for 1 h and squashed in a drop of 45% acetic acid on microscope slides pretreated with Vectabond (Vector Laboratories, UK).

Chromosome preparations were treated with 100 μg/ml DNAse-free RNAase for 1 h and post-fixed in 4% paraformaldehyde for 10 min. Chromosomal DNA was denatured in 70% formamide, 2 × standard saline citrate (SSC) at 70°C for 2.5 min and dehydrated through an ethanol series at 4°C. The inserts of clones NP3R.4 and NP4R.3, labelled with DIG-11-dUTP by PCR (Suen *et al*. 1997), were used as probes. The hybridization mixture consisted of 50% formamide, 2 × SSC, 10% dextran sulphate, 0.1% SDS, 5 ng/μl probe DNA and 1250 ng/μl herring sperm DNA. Hybridization was carried out at 37°C overnight. Slides were washed for 10 min in 20% formamide, 0.2 × SSC at 42°C, 10 min in 2 × SSC at 42°C and 3 × 5 min in 2 × SSC at room temperature. Probes were detected with fluorescein-conjugated anti-DIG antibody (Boehringer Mannheim, Germany) and the signals were amplified with fluorescein-conjugated anti-sheep IgG (Vector Laboratories, UK). Chromosomes were counterstained with 2 μg/ml propidium iodide.

Results

Cloning of TAS

The telomeres of *N. plumbaginifolia* chromosomes ranged from 60 to 160 kb in size and were resistant to the digestion of many restriction enzymes (Chen *et al*. unpublished data). To enrich telomeric DNA of *N. plumbaginifolia*, plant genomic DNA was digested with various restriction enzymes, fractionated by gel electrophoresis, Southern transferred and probed with the oligonucleotide $(T_3AG_3)_4$ (Figure 1). A fraction of plant DNA homologous to the telomere probe remained as uncleaved high-molecular-weight DNA in all enzyme digests (Figure 1b). However, in the *Hae*III digests, this high-molecular-weight DNA was well separated from the non-telomeric bulk DNA (Figure 1a). Therefore, this *Hae*III-digested high-molecular-weight plant DNA was recovered from the gel and used for cloning of TAS by vectorette PCR. Ninety putative NPTAS clones were identified after sequential colony hybridization and Southern hybridization of plasmid DNA using $(T_3AG_3)_4$ as a probe. The inserts of these clones ranged from 200 bp to 1 kb. All clones consisted of varying stretches of telomeric repeats $(T_3AG_3)_n$ and adjacent non-telomeric sequences. At least six classes of non-telomeric regions were identified based on the sequences (data not shown).

Identification of NP3R and NP4R repeated sequences

To characterize the putative NPTAS clones, DNA fragments corresponding to the non-telomeric region of

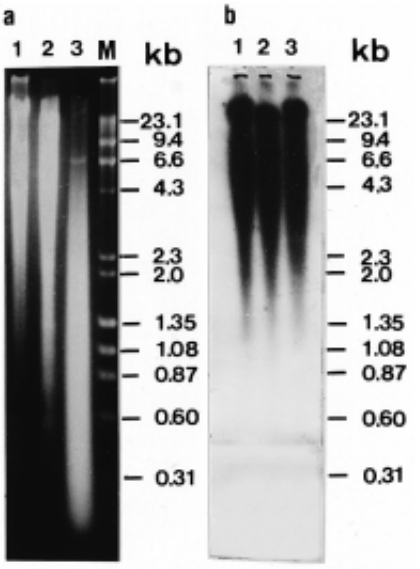

Figure 1. Southern hybridization of *N. plumbaginifolia* genomic DNA with telomeric probe. Genomic DNA of *N. plumbaginifolia* was digested with *Sal*I (1), *Bam*HI (2) or *Hae*III (3) and then fractionated by gel electrophoresis. The gel was stained with EtBr (**a**) and then Southern hybridized with $(T_3AG_3)_4$ probe (**b**). M represents the DNA size markers.

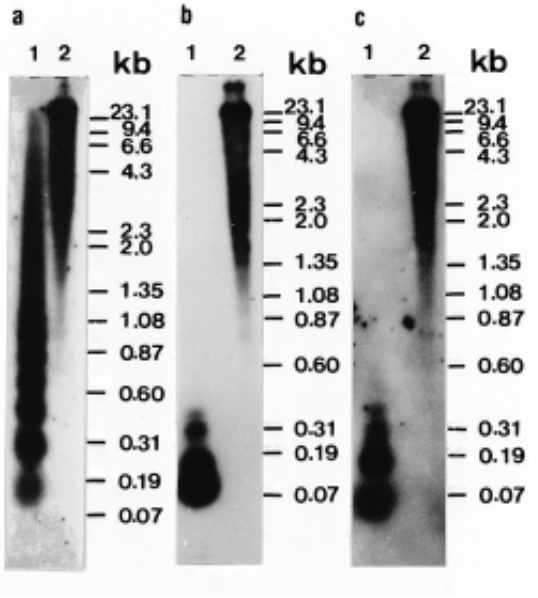

Figure 2. Southern Hybridization of *N. plumbaginifolia* genomic DNA with non-telomeric probes. Genomic DNA of *N. plumbaginifolia* was digested with *Hae*III (1) or *Bam*HI (2), fractionated by gel electrophoresis and Southern hybridized with probes corresponding to the non-telomeric regions of NPTAS3 (**a**), NPTAS4 (**b**), and fragment B of NP4R.3 (**c**).

each clone were used as probes for genomic Southern hybridization of *N. plumbaginifolia* DNA. Hybridization of *Hae*III-digested plant DNA with the non-telomeric region of clone NPTAS3 revealed a ladder of hybridization bands (Figure 2a, lane 1). Hybridization of *Bam*HI-digested plant DNA with the same probe showed that a large fraction of plant DNA homologous to the probe remained as uncleaved 'relic' DNA (Figure 2a, lane 2). This result suggests that NP3R.1, the non-telomeric region of clone NPTAS3, belongs to the family of a highly repeated tandem sequence called NP3R. Similarly, hybridization of *Hae*III-digested (Figure 2b, lane 1) and *Bam*HI-digested (Figure 2b, lane 2) plant DNA with the non-telomeric region of NPTAS4 indicates that NP4R.1, the non-telomeric region of clone NPTAS4, is a member of another tandemly repeated sequence designated as NP4R.

The sequences of clones NPTAS3 and NPTAS4 are shown in Figures 3a & 4a respectively. Degenerated telomeric sequences T_4AG_3 and T_3AG_2T are scattered in the telomeric regions of NPTAS3 and NPTAS4 respectively. Telomere-similar sequences T_5G_2 and T_5G_3 were observed in the telomeric region and at the telomere–subtelomere junction of NPTAS3; T_6G_2 was found near the telomere–subtelomere junction of NPTAS4. The size of NP3R.1 (117 bp) is smaller than that of the shortest fragment of *Hae*III-digested plant DNA detected by NP3R.1 (170 bp) (Figure 2a, lane 1), suggesting that NP3R.1 was only a portion of the monomeric unit of the repeated sequence NP3R. In order to obtain a full-length NP3R and to understand the genomic organization of the NP4R repeats, a *Hae*III library of plant DNA was screened with the probes NP3R.1 and NP4R.1. Five and six clones hybridized with NP3R.1 and NP4R.1, respectively, were isolated and sequenced.

Alignment of the sequences of clones hybridized with NP3R.1 (Figure 3b) revealed that a full-length repeating unit of NP3R is 165 bp; a repeating unit consists of three subfragments (I, II and III) and, depending on the clones, a *Hae*III site could occur at the junction of any two subfragments (Figure 3b) as the result of a single base change. Consequently, ladder-like bands appeared when *Hae*III-digested plant DNA was probed with NP3R.1 (Figure 2a, lane 1). A region near the end of subfragment III contained three copies of a TG-rich telomere-similar sequence. An imperfect inverted repeat, $T_4ACG_2TCATA_4$, was observed at the junction of subfragments III and I.

Comparing the sequences of clones hybridized with NP4R.1 (Figure 4b) suggested that a complete repeat unit of NP4R contains fragments A (113 bp) and B (67 bp) and that all the five clones containing fragment A have only a *Hae*III site at the junction of fragments A and B, whereas clone NP4R.3 lacks this restriction site. To test this prediction, the signal of Figure 2b was stripped off and the blot was reprobed with NP4R.3x corresponding to fragment B of NP4R.3 (Figure 4b). No difference was observed in the hybridization patterns for *Bam*HI-digested plant DNA when probed with NP4R.1 or NP4R.3x (Figure 2b & c, lane 2). However, in

a

CCGGGACGGAGCCGGGTCGGAATCGCCTAATATTTGTCCTGGACGTTAAATACGACATAAGAAGACATTGTCACCCTTCCGAAA
CGACAAAATTTCA***TTTTTGG***TA***TTTTTTGGG***CAttttagggttttagggtttagggtttagggttttagggtttagggttt
agggtttagggtttagggttttagggttttagggttttagggtttagggttttagggtttagggtttagggtttagggtttagg
gtttagggtttagggtttagggtttagggtttagggtttagggtttagggtttagggtttagggtttagggtttaggg***tttttggg***ttta
gggtttaggg***tttttggg***tttagggtttagggtttaggg

b

```
         |<------------- III -------------------
                                                               60
NP3R.3   CCCTAAGAAAACATTGTCACCCTTCCGAAATGACAAAATTACATTTTTGGTATTTTTTGG
NP3R.6    CCTAAGAAAATATTCTCACACTTCCGAAACGACAAAATTGCATTTTTGGCATTTTTTGG
NP3R.8    CCTAAGAGAACATTATCACCCTTCTGAAACGACAAAATTGCATTTTTAGCATTTTTTGG

   ------>|<---------- I ---------->|<----
                                                               120
NP3R.3   GCATTTTACAGCCATAAAAA-CAGAAATCCGCGCGAGTCGCAAATTTTCGGCAGGGACGGA
NP3R.4                                                       CCGGGACGGA
NP3R.6   GCATTTTACCGCCATAAAAC-CAATAATCCGCGCGAGTCGCAAGTTTTCGGCAGGGACGGA
NP3R.7                                                       CCGAAAAAGA
NP3R.8   GCATTTTATTGACATAAAAACCAGATATCCGCG--AGTCGTAAATTTTCGG
NP3R.1                                                       CCGGGACGGA

   ------------- II ------------->|<-------
                                                               180
NP3R.3   CCCGTCTCGAAATCGCCTAATATTTGTCCTGGACGTCAAGTACTGCCTAAGAAAACATTG
NP3R.4   GCCGGCTCGGAATCGCCTAATATTTGTCCTGGACGTTAAATACGACATAAGAAGACATTG
NP3R.6   CCCAGCTCGGAATCGCCTAATATTTGTCCTGGACGTCAAATATGG
NP3R.7   GCCGGCTCGGAATCGCCTAATATTTGTCCTGGACGTCAAATACGGCTTAAGTAGACATTG
NP3R.1   GCCGGGTCGGAATCGCCTAATATTTGTCCTGGACGTTAAATACGACATAAGAAGACATTG

   ------------- III ------------->|<--
                                                      235
NP3R.3   ACACTCTTTCGAAACGACAAAATTGCATATTTGGCATTTTTTGGGCGTTTTACGG
NP3R.4   TCACCCTTCCGAAATGACAAAATTGCATTTTTGATACTTTTTGGGTATTTTACGGTCATA
NP3R.7   CCACCCTTCCGAAATGACAAAATTGTATTTTTTGTATTTTTTGGGCATTTTACGGCATAA
NP3R.1   TCACCCTTCCGAAACGACAAAATTTCATTTTTGGTATTTTTTGGGCATTTTAGGG

   ------- I -------->|
NP3R.4   AAAGCTAGAAATCCGCACGAGTCCCAAATTGTCGG
NP3R.7   GAACCAAAAATT-CGCACGAGTCCCAAATTGTCGG
```

Figure 3. a Nucleotide sequence of NPTAS3. The sequence of the non-telomeric region is expressed as upper case and that of the telomeric region as lower case. The degenerated telomeric sequences are underlined. Plant telomere-similar sequences are bold and italicized. **b** The alignment of the nucleotide sequences of NP3R clones. Nucleotides not in common with NP3R clones are dotted underneath. Plant telomere-similar sequences are in boxes. The imperfect inverted repeats are underlined. Three subfragments in each repeating unit of NP3R were arranged in the order as shown. NP3R.1 was derived from the non-telomeric region of NPTAS3 shown in (**a**). The sequences have been submitted to the EMBL Data Library and have been assigned the following accession numbers. Y12621 (NPTAS3); Y12622 (NP3R.3); Y12623 (NP3R.4); Y12624 (NP3R.6); Y12625 (NP3R.7); Y12662 (NP3R.8).

a

CCCGAGCCCCCGGACCCGAATCGCCTAAAATTTTACCGGAACATCAAATACGCCCTAAGGAACCGTTGTCACCGCCCCGCCATG
ACCGAATTATA***TTTTTTGG***CAtttagggtttagggtttagggtttaggtttagggtttagggtttagggtttagggtttagg
gtttagggtttagggtttagggtttagggtttagggtttagggtttagggtttagggtttagggtttagggtttagg
gtttagggtttaggtttagggtttagggtttagggtttagggtttagggtttagggtttagggtttagggtttaggg

b

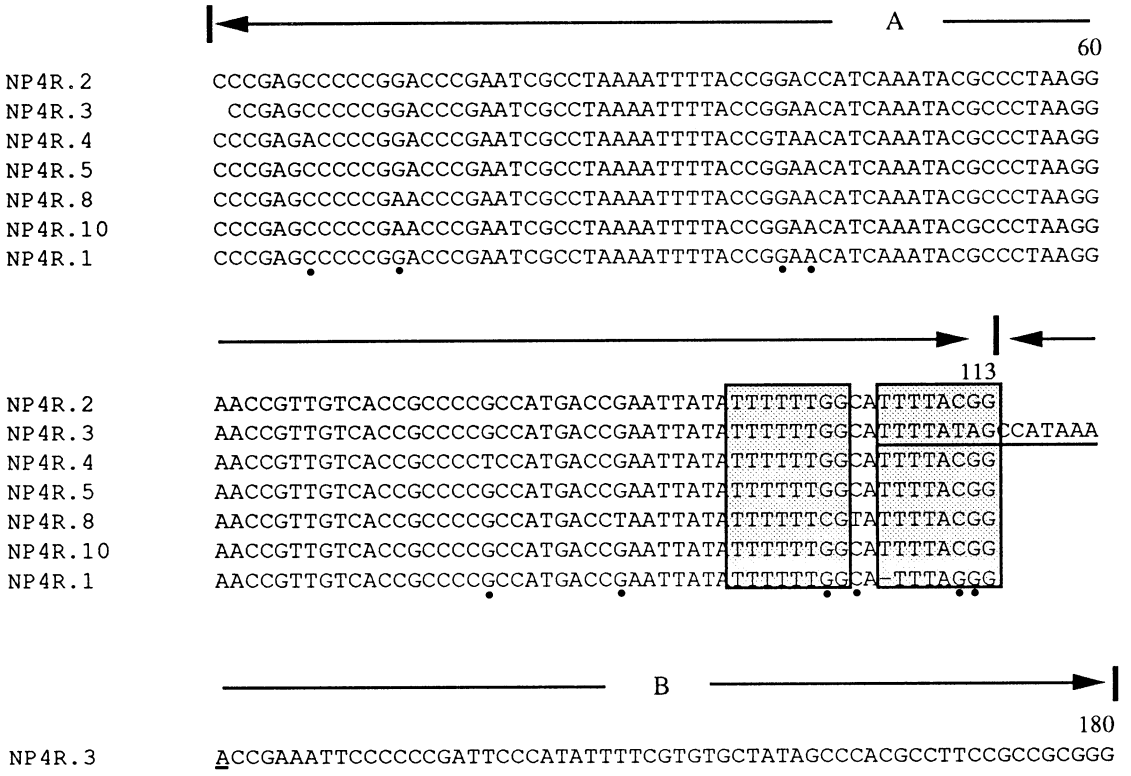

Figure 4. **a** Nucleotide sequence of NPTAS4. The sequence of the non-telomeric region is expressed as upper case and that of the telomeric region as lower case. The degenerated telomeric sequences are underlined. Plant telomere-similar sequences are bold and italicized. **b** The alignment of the nucleotide sequences of NP4R clones. Nucleotides not in common with NP4R clones are dotted underneath. Plant telomere-similar sequences are in boxes. The imperfect inverted repeat is underlined. NP4R.1 was derived from the non-telomeric region of NPTAS4 shown in (**a**). The sequences have been submitted to the EMBL Data Library and have been assigned the following accession numbers: Y12620 (NPTAS4); Y12626 (NP4R.10); Y12627 (NP4R.2); Y12628 (NP4R.3); Y12629 (NP4R.4); Y12630 (NP4R.5); Y12631 (NP4R.8).

the blot of *Hae*III-digested plant DNA, the major 120-bp band that appeared when probed with NP4R.1 was replaced by a major band of 70 bp when probed with NP4R.3x (Figure 2b & c lane 1). This result is consistent with our prediction. Sequence variation among NP4R repeats was not as frequent as that in NP3R repeats (Figure 3b & 4b). An imperfect inverted repeat, $T_4ATAGC_2ATA_4$, also appeared at the junction of fragments A and B of the NP4R repeat.

Copy number of NP3R, NP4R and telomere repeats

Using the procedure and calculations given in Materials and methods, it was found that the copy numbers of NP3R, NP4R and telomere repeats in the haploid genome of *N. plumbaginifolia* were 8.4×10^4, 6×10^3 and 1.5×10^6 respectively.

Chromosomal locations of NP3R and NP4R repeats

The somatic chromosomes of haploid *N. plumbaginifolia* could not be easily distinguished when stained with propidium iodide. *In situ* hybridization using NP3R.4 as a probe showed the presence of signals at the ends and/or in interstitial regions of all 10 chromosomes (Figure 5a). The exact number of hybridization sites was difficult to determine, because some signals were very weak and were not present consistently. The signals in some interstitial regions appeared to be somewhat scattered (Figure 5a), indicating the possibility of interspersion of

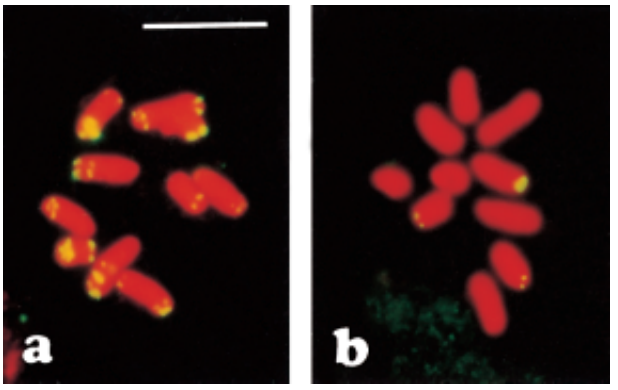

Figure 5. Fluorescence *in situ* hybridization of somatic chromosomes of haploid *N. plumbaginifolia* using clones NP3R.4 (**a**) and NP4R.3 (**b**) as probes. The hybridization sites fluoresce yellow while the other parts of the chromosomes fluoresce red. Scale bar = 10 μm.

NP3R.4 with other repeated and non-repeated sequences. In contrast to the distribution of NP3R.4, NP4R.3 sequences were confined to only three terminal or subterminal regions where the repeats were not seriously interrupted by other sequences, as revealed by the patterns of hybridization signals (Figure 5b).

Homology between NP3R and NP4R repeated sequences

A homology search from Genbank revealed that NP3R was 80.4% homologous to HRS60.1 (Figure 6a), a tandem repeat isolated from *N. tabacum* (Koukalova *et al*. 1989) and later identified to be specific for the genome of the parental species *N. sylvestris* (Kenton *et al*. 1993, Koukalova *et al*. 1993). The sequences most similar to NP4R were GRS (Gazdova *et al*. 1995) and ECO (Suter-Crazzolara *et al*. 1995) with a homology of 79.68% and 74.2% respectively (Figure 6b). Both GRS and ECO tandem repeats were isolated from *N. tabacum* and found to be members of the same family of repeated sequence specific for the genome of the parental species *N. tomentosiformis* (Gazdova *et al*. 1995, Suter-Crazzolara *et al*. 1995). Because there was a homology of 65% between HRS60 and ECO repeats (Suter-Crazzolara *et al*. 1995), the sequences of NP3R and NP4R repeats were also compared (Figure 6c). The homology between them was 74.67%. We also noted that the 3′ end of HRS60.1 resembled fragment B of the NP4R repeat both in sequence and in length (Figure 6a & c).

Discussion

The huge size and undigestibility by most restriction enzymes of the telomeres have made it difficult to isolate telomeres and TASs from the *N. plumbaginifolia* genome even by YAC cloning (Richards *et al*. 1992). Our method including enrichment of telomeric fragments followed by vectorette PCR amplification enabled us to isolate several TAS clones from a tiny amount of plant DNA. Without enrichment of telomeric fragments only false-positive clones of TAS were obtained (data not shown), suggesting that the telomere primer used in the vectorette PCR amplification may bind to the interstitial telomeric sequences in the chromosomes of *N. plumbaginifolia*. Interstitial telomeric sequences have been reported in some plant genomes (Fuchs *et al*. 1995, Presting *et al*. 1996). This method should be useful for characterizing the TAS and telomere regions from a genome containing interstitial telomeric sequences and huge telomeres.

The structure around the telomere–subtelomere junction in most plant species remained obscure. A spacer region of varying length has been proposed to exist between subtelomeric and telomeric repeat arrays in barley (Roder *et al*. 1993), rice (Wu & Tanksley 1993) and tomato (Broun *et al*. 1992), based on physical mapping of telomeres. However, sequencing of TAS clones from barley revealed that at least some barley telomeres contained subtelomeric repeats immediately after the telomere repeat arrays (Kilian & Kleinhofs 1992). In *N. tabacum*, the subtelomeric repeat HRS60 was found to attach directly to the telomeric sequence (Fajkus *et al*. 1995). Our results of sequencing TAS clones and *in situ* hybridization provided evidence that at least some chromosomes of *N. plumbaginifolia* contain no other sequence between telomeric repeat arrays and NP3R or NP4R repeat arrays. The result of *in situ* hybridization also showed that interstitial NP3R repeats may be interspersed with other sequences. Whether NP3R and NP4R repeats appear at the same subtelomeric region is not known.

Sequence homology search revealed homology for NP3R and HRS60, NP4R and GRS/ECO, and NP3R and NP4R (Figure 6), indicating that these repeats may have originated from a common ancestral sequence. The presence of NP3R and NP4R repeats in the genome of *N. plumbaginifolia* and the isolation of HRS60 and GRS or ECO repeats from the genome of *N. tabacum* suggested that both classes of repeats may co-exist in many *Nicotiana* species with variation in chromosomal location, copy number and sequence similarity. Experiments supporting this idea were the detection under low stringency of a HRS60-like sequence in the genome of *N. tomentosiformis* (Koukalova *et al*. 1993) and the observation of a weak signal in the genomic Southern blot of *N. sylvestris* hybridized with ECO repeat (Suter-Crazzolara *et al*. 1995).

Although the 5′ end of the HRS60 repeat is homologous to that of the NP3R repeat (Figure 6a), the 3′ end of HRS60 is more similar to that of the NP4R repeat both in sequence and in size (Figure 6a & c). This implies that HRS60 may be derived from the product of recombination between NP3R and NP4R repeats. Some short inverted repeats and stems are present in both NP3R and NP4R (Figure 3b & 4b). Such structures provide a potential source for homologous recombination, allowing creation of rearrangements. Perhaps, the genome of ancient *N. tabacum* may contain both NP3R-

a

```
NP3R.4          CCGGGACGGAGCCGGCTCGGAATCGCCTAATATTTGTCCTGGACGTTAAATACGACATAA
NTHRS60.1(-)              CGTCGTGGAATCGCCTAATATTTGTCCCGGACATCAAATACGGCATAA

NP3R.4          GAAGACATTGTCACCCTTCCGAAATGACAAAATTGCATTTTTGATACTTTTTGGGTATTT
NTHRS60.1(-)    GAAACAATTTGCACCCACTCGAAATGACCACATTATATTTTT---------TTGGCATTT

NP3R.4          TACGGTCATAAAAGCTAGAAATCCGCACGAGTCCCAAATTGTCGG
NTHRS60.1(-)    TACAGCCATAAAA-CTACAAATCCGCGCGAGTCCCAAATTTTTGTGTGCTATAGCCCATG

NTHRS60.1(-)    CCTTCCGCCTTGGGCCCGGATGGATC
```

b

```
NP4R.3                                                                CC-
NTPGRS61(-)     GAATTCGGCCCGATTTCTATATTTTTGTGTGCTCTAGCCCATGCCTTTCGAGTGGGCCCG
NTREECOA2(-)

NP4R.3          GAGCCCCCGGACCCGAATCGCCTAAAATTTTACCGGAACATCAAATACGCCCTAAGGAAC
NTPGRS61(-)     GGTCCCCCGAAACCGGATTGCCTAAAATTTTTCTAGGCCATCAAATACGACCTAAGGAAC
NTREECOA2(-)      AGACCCAGATCGCCTAAAAGTATTTTTGGACTTC-AATACAACTTAAGGAAC

NP4R.3          CGTTGTCACCGCCCCGCCATGACCG--AATTATATTTTTTGGCATTTTATAGCCATAAAAC
NTPGRS61(-)     CGTAGTCACCGCCCCGCCATGGCCGTTCCTTGTTGTTTTTAGCGTTTTT-AGCCAAAACAC
NTREECOA2(-)    CATATTCACCACCCCTCCATGACCGTTCCTCATT--TTTTGGTGTTTTTACAGCTATAAAGT

NP4R.3          CGAAATTCCCCCCGATTCCCATATTTTCGTGTGCTATAGCCCACGCCTTCCGCCGCGGG
NTPGRS61(-)     TAGAATTC
NTREECOA2(-)    TGGAATTCCGCTCGATTCCCCTATTTTCGTGTGCCATAGACCATGCCTTCTATCGTGGC

NTREECOA2(-)    CAAAGGCCTTC
```

c

```
NP3R.4          CCGGGACGGAGCCGGCTCGGAATCGCCTAATATTTGTCCTGGACGTTAAATACGACATAA
NP4R.3          CC--GA-GCCCCCGGACCCGAATCGCCTAAAATTTTACCGGAACATCAAATACGCCCTAA

NP3R.4          GAAGACATTGTCACCCTTCCGAAATGACAAAATTGCATTTTTGATACTTTTTGGGTATTT
NP4R.3          GGAACCGTTGTCACCGCCCCGCCATGACCGAATTATATTTTTT----------GGCATTT

NP3R.4          TACGGTCATAAAAGCTAGAAAT-CCGCACGAGTCCCAAATTGTCGG
NP4R.3          TATAGCCATAAAACC--GAAATTCCCCCCGATTCCCATATTTTCGTGTGCTATAGCCCACG

NP4R.3          CCTTCCGCCGCGGG
```

Figure 6. Alignment of nucleotide sequences. **a** NP3R.4 and NTHRS60.1 (Koukalova *et al.* 1989). The sequence in NTHRS60.1 resembling that of NP4R.3 is underlined. **b** NP4R.3, NTPGRS61 (Gazdova *et al.* 1995) and NTREECOA2 (Suter-Crazzolara *et al.* 1995). The sequences labelled with the (−) symbol are complementary to the published ones. **c** NP3R.4 and NP4R.3. Gaps have been introduced to maximize the homology. Nucleotides not identical between sequences are dotted underneath.

and NP4R-like repeats, which recombined during evolution to generate a HRS60-like sequence in the genome of modern tobacco.

In humans (Morin 1991) and *Saccharomyces cerevisiae* (Kramer and Haber 1993), new telomeres could be regenerated at the broken ends of chromosomes when the breakpoints were right next to G- and/or T-rich telomere-similar sequences. In this study, we observed that NP3R repeats appeared at the ends and/or interstitial regions of all *N. plumbaginifolia* chromosomes and that each repeating unit contained two to three copies of GT-rich telomere-similar sequence. *N. plumbaginifolia* belongs to the section *Alatae* in which all species are diploid, but they differ in basic chromosome number, karyotype symmetry and DNA content (Narayan 1987). Chromosome breakage and rearrangement after inter-

specific hybridization are thought to have played an important role in chromosome evolution in this section (Suen *et al*. 1997). We postulate that the telomere-similar sequences in interstitial TASs may serve as substrates for the enzyme telomerase in telomere healing at the broken chromosome ends.

Acknowledgements

We thank Y. C. Lee for a critical review of this article, F. M. Pan for sequencing cloned DNA and R. Wu for constructive suggestions. This work was supported by grant no. NSC86-2311-B-001-028-B05 from the National Science Council and Academia Sinica in the Republic of China.

References

Anamthawat-Jonsson K, Heslop-Harrison JS (1993) Isolation and characterization of genome-specific DNA sequences in *Triticeae* species. *Mol Gen Genet* **240**: 151–158.

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.

Brandes A, Roder MS, Ganal MW (1995) Barley telomeres are associated with two different types of satellite DNA sequences. *Chrom Res* **3**: 315–320.

Broun P, Ganal MW, Tanksley SD (1992) Telomeric arrays display high levels of heritable polymorphism among closely related plant varieties. *Proc Natl Acad Sci USA* **89**: 1354–1357.

Burr B, Burr FA, Matz EC, Romero-Severson J (1992) Pinning down loose ends: mapping telomeres and factors affecting their length. *Plant Cell* **4**: 953–960.

Castilho A, Heslop-Harrison JS (1995) Physical mapping of 5S and 18S-25S rDNA and repetitive DNA sequences in *Aegilops umbellulata*. *Genome* **38**: 91–96.

Chen CC, Huang CR, To CY (1985) Anther cultures of four diploid *Nicotiana* species and chromosome numbers of regenerated plants. *Bot Bull Acad Sin* **26**: 147–153.

Chen CM, Wang CT, Lee FM, Ho CH (1996) Construction and characterization of a *Nicotiana plumbaginifolia* genomic library in a yeast artificial chromosome. *Plant Sci* **114**: 159–169.

Copenhaver GP, Pikaard CS (1996) RFLP and physical mapping with an rDNA-specific endonuclease reveals that nucleolus organizer regions of *Arabidopsis thaliana* adjoin the telomeres on chromosomes 2 and 4. *Plant J* **9**: 259–272.

Fajkus J, Kralovics R, Kovarik A, Fajkusova L (1995) The telomeric sequence is directly attached to the HRS60 subtelomeric tandem repeat in tobacco chromosome. *FEBS Lett* **364**: 33–35.

Fuchs J, Brandes A, Schubert I (1995) Telomere sequence localization and karyotype evolution in higher plants. *Plant Syst Evol* **196**: 227–241.

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

Ganal MW, Lapitan NLV, Tanksley SD (1991) Macrostructure of the tomato telomeres. *Plant Cell* **3**: 87–94.

Gardiner JM, Coe EH, Chao S (1996) Cloning maize telomeres by complementation in *Saccharomyces cerevisiae*. *Genome* **39**: 736–748.

Gazdova B, Siroky J, Fajkus J, Brzobohaty B, Kenton A, Parokonny A, Heslop-Harrison JS, Palme K, Bezdek M (1995) Characterization of a new family of tobacco highly repetitive DNA, GRS, specific for the *Nicotiana tomentosiformis* genomic component. *Chrom Res* **3**: 245–254.

Jofuku K, Goldberg RB (1988) Analysis of plant gene structure. In: Shaw CH, ed. *Plant Molecular Biology. A Practical Approach*. Oxford: IRL Press, chapter 2.

Kenton A, Parokonny AS, Gleba YY, Bennett MD (1993) Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. *Mol Gen Genet* **240**: 159–169.

Kilian A, Kleinhofs A (1992) Cloning and mapping of telomere-associated sequences from *Hordeum vulgare* L. *Mol Gen Genet* **235**: 153–156.

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**: 77–80.

Koukalova B, Komarnitsky IK, Kuhrova V (1993) The distribution of tobacco HRS60 DNA repeated sequences in species of the genus *Nicotiana*. *Plant Sci* **88**: 39–44.

Kramer KM, Haber JE (1993) New telomeres in yeast are initiated with a highly selected subset of TG_{1-3} repeats. *Genes Dev* **7**: 2345–2356.

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

Morin GB (1991) Recognition of a chromosome truncation site associated with α-thalassaemia by human telomerase. *Nature* **353**: 454–456.

Narayan RKJ (1987) Nuclear DNA changes, genome differentiation and evolution in *Nicotiana (Solanaceae)*. *Plant Syst Evol* **157**: 161–180.

Presting GG, Frary A, Pillen K, Tanksley SD (1996) Telomere-homologous sequences occur near the centromeres of many tomato chromosomes. *Mol Gen Genet* **251**: 526–531.

Rayburn AL, Gill BS (1987) Molecular analysis of the D-genome of the *Triticeae*. *Theor Appl Genet* **73**: 385–388.

Richards EJ, Ausubel FM (1988) Isolation of higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* **53**: 127–136.

Richards EJ, Chao S, Vongs A, Yang J (1992) Characterization of *Arabidopsis thaliana* telomeres isolated in yeast. *Nucleic Acids Res* **20**: 4039–4046.

Riley J, Butler R, Ogilvie D, Finniear R, Jenner D, Powell S, Anand R, Smith JC, Markham AF (1990) A novel, repid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res* **18**: 2887–2890.

Roder MS, Lapitan NLV, Sorrells ME, Tanksley SD (1993) Genetic and physical mapping or barley telomeres. *Mol Gen Genet* **238**: 294–303.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Suen DF, Wang CK, Lin RF, Kao YY, Lee FM, Chen CC (1997) Assignment of DNA markers to *Nicotiana sylvestris* chromosomes using monosomic alien addition lines. *Theor Appl Genet* **94**: 331–337.

Suter-Crazzolara C, Brzobohaty B, Gazdova B, Schell J, Reiss B (1995) T-DNA integrations in a new family of repetitive elements of *Nicotiana tabacum*. *J Mol Evol* **41**: 498–504.

Vershinin AV, Schwarzacher T, Heslop-Harrison JS (1995) The large-scale genomic organization of repetitive DNA families at the telomeres of rye chromosomes. *Plant Cell* **7**: 1823–1833.

Wu K-S, Tanksley SD (1993) Genetic and physical mapping of telomeres and macrosatellites of rice. *Plant Mol Biol* **22**: 861–872.