Theor Appl Genet (1999) 98: 387–395

K. D. Kenward · D. Bai · M. R. Ban · J. E. Brandle

Isolation and characterization of Tnd-1, a retrotransposon marker linked to black root rot resistance in tobacco

Received: 3 March 1998 / Accepted: 18 August 1998

Abstract In *Nicotiana debneyi*, resistance to a wide range of black root rot (*Chalara elegans*) isolates is conferred by a single dominant gene. This gene has been transferred to cultivated tobacco (*Nicotiana tabacum*) and was recently discovered to be linked in coupling to a 1050-bp random amplified polymorphic DNA (RAPD) marker generated with the UBC 418 primer. We have cloned and sequenced the $UBC418_{1050}$ marker and found it to be part of a retrotransposon. This retrotransposon is a remnant of the *N. debneyi* genome and is the first to be isolated from this species. Transposon *N. debneyi* (Tnd)-1 is present in the tobacco genome as two directly repeated copies, and in multiple copies in the donor species *N. debneyi* and in a number of related *Nicotiana* species. The retrotransposon appears to have been introduced into the *Nicotiana* genome after the development of the Suavolentes progenitors. The gene associated with black root rot resistance co-segregates with the retrotransposon in tobacco and is thought to be contained within the introgressed fragment marked by Tnd-1. The retrotransposon will therefore be a useful species-specific landmark that can be used for future cloning of the resistance gene.

Key words Retrotransposon · Long-range and inverse PCR · Black root rot · Resistance gene

The complete Tnd-1 nucleotide sequence has been deposited into the GenBank database, accession number AF059674

Communicated by M. A. Saghai Maroof

K. D. Kenward · D. Bai · M. R. Ban · J. E. Brandle (✉)
Southern Crop Protection and Food Research Centre, Agriculture and Agri-Food Canada, 1391 Sandford St., London, ON, Canada N5V 4T3
E-mail: brandleje@em.agr.ca

Introduction

Complete resistance to a wide range of black root rot *Chalara elegans* isolates has been transferred from *Nicotiana debneyi* Domin to cultivated tobacco, *N. tabacum*, by interspecific hybridization and backcrossing (Hoffbeck et al. 1965; Clayton 1969). This resistance is not widely used, however, due to associated negative agronomic and chemical traits such as later maturity and reduced yield (Legg et al. 1981). These negative traits may be due to linked deleterious genes carried over with the resistance gene during introgression. Isolation of the resistance gene and re-introduction via transformation would reduce linkage drag and permit transfer of the resistance to non-related plant species that are also affected by *C. elegans*.

Random amplified polymorphic DNA (RAPD) screening has identified two markers associated with the *N. debneyi* resistance gene (Bai et al. 1995). Of particular interest is a 1050-bp polymerase chain reaction (PCR) product generated with the UBC 418 primer from lines carrying the *N. debneyi* resistance gene and found to be linked in coupling to the resistance gene. $UBC418_{1050}$ is a primary PCR product in *N. debneyi* and appears to be a repetitive element within this genome, whereas only a single copy of the marker co-segregates with black root rot resistance in cultivated tobacco. Such species-specific markers associated with introgressed DNA segments were essential for the cloning of a gene associated with cyst nematode resistance in sugar beet (*Beta vulgaris*) (Cai et al. 1997) and warrant further investigation.

We have cloned the $UBC418_{1050}$ marker and found that it represents the 5′ end of a retrotransposon. As mobile genetic elements, retrotransposons are thought to be important forces in genome evolution (Kidwell and Lisch 1997). Insertion events have been associated with changes in gene regulation (Kloeckener-Gruissem and Freeling 1995; Marillonnet and Wessler 1997),

gene duplication (White et al. 1994), introduction of introns (Giroux et al. 1994), and DNA repair (Moor and Haber 1996). Their utility as molecular markers has been demonstrated in the development of genetic maps (Kumar et al. 1997) and transposon tagging experiments for the identification and isolation of functional genes (Osborne and Baker, 1995). The full-length retrotransposon present in a *N. tabacum* cultivar resistant to black root rot was isolated using long range PCR, and its organization in the plant genome was characterized. No open reading frame was found within the retrotransposon which would indicate a direct role in disease resistance. The isolated retrotransposon does not appear to be mobile, although related forms were found in a number of *Nicotiana* species suggesting previous activity. The implications for finding the resistance gene in the flanking genomic DNA are discussed.

Materials and methods

Plant materials

The black root rot resistance phenotype was originally introgressed from *Nicotiana debneyi* to tobacco cv 'Burley 49' (Clayton 1969; Hoffbeck et al. 1965). The trait was transferred to tobacco cv 'Delgold' via the breeding line 'PB19' using five backcrosses to cv 'Delgold', followed by eight generations of selfing. The resulting near-isogenic line was designated RM12-3. Sequences associated with the $UBC418_{1050}$ marker were cloned from RM12-3 genomic DNA.

Preparation and Southern blot analysis of genomic DNA

Genomic DNA was prepared according to the method of Palmiter (1974) . For Southern blotting, DNA (25 μg) was digested with 20 units of restriction enzyme for 3–6 h, supplemented with the same amount of enzyme, digested overnight, precipitated, and subjected to agarose gel electrophoresis. DNA was transferred to Hybond-N + (Amersham) nylon membrane by capillary transfer and fixed to the membrane by UV cross-linking. Blots were pre-hybridized for 3 h in 5 × Denhardt's/5 × SSC/0.5% SDS with herring DNA at 65°C and then hybridized with the probe overnight at the same temperature. Probes were generated by PCR amplification of RM12-3 genomic DNA with *Taq* DNA Polymerase (Gibco/BRL) and labelled with [^{32}P]-dCTP using a random primer labelling kit (Gibco/BRL).

Generation and cloning of PCR products

For inverse PCR, genomic DNA (3 μg) was digested as described above for Southern blotting. Reactions were incubated at 95°C for 10 min to heat inactivate the remaining enzyme, their volumes increased to 800 μl at 1 × ligase buffer concentration, and the DNA ligated overnight at 16°C with 5 units of T4 DNA ligase (Gibco/BRL). Each reaction was divided into 4 aliquots, precipitated, and stored as a dried pellet at −20°C.

PCR amplifications were performed with 0.75–1 μg genomic DNA template. The $UBC418_{1050}$ PCR product was generated with *Taq* DNA polymerase as previously described (Bai et al. 1995), gel-purified, and cloned into pGEM-T (Promega). The Pfu1.8 PCR product was amplified with Pfu DNA polymerase (Stratagene) using 0.5 m*M* dNTP and 0.4 μ*M* of each primer over 30 cycles of 95°C denaturation for 1 min, 60°C annealing for 1 min, and 72°C extension for 4 min, followed by a 20-min incubation at 72°C. When long PCR products were anticipated, amplifications were done with the Expand™ Long Template PCR system (Boehringer Mannheim) according to the manufacturer's instructions and using a 15-min extension time. For nested PCR, initial amplification products were subjected to QIAquick PCR purification (Qiagen) for nucleotide and primer removal prior to a second round of amplification as described above but with a different primer set. Reactions using Pfu DNA polymerase and the Expand™ system were relatively specific and resulted in only 1–2 major PCR products. DNA fragments generated with these polymerases were purified (Qiagen), cloned into pCR-Script™ SK + (Stratagene) or pGEM-T_{easy} (Promega), and the desired clones selected by size and restriction mapping.

DNA sequencing and analysis

Clones were sequenced with the ABI Prism Sequencer 377. The full-length RT clone was sequenced completely in both directions: all other isolates were sequenced in one direction. Assembled sequences were analyzed using the National Center for Biotechnology Information (NCBI) Gapped Blast Search, DNAstar Genequest, Baylor College of Medicine (BCM) Gene Finder, and Genie programs to identify potential open reading frames, promoter elements, transcription/translation signals and introns, as well as to search for homology with known nucleotide sequences, proteins, and expressed sequence tags.

Results

Cloning of the transposon *N. debneyi* (Tnd)-1 retrotransposon

The $UBC418_{1050}$ marker was amplified from RM12-3, then cloned and sequenced (Fig. 1). The sequence was 1067 bp long, flanked by inverted repeats of the UBC 418 primer. A 193-amino acid open reading frame (ORF) was present that started at nucleotide 491. A search of the non-redundant protein database at NCBI using the Gapped BlastX algorithm, showed that the ORF was homologous (score = 62, $E = 4 \times 10^{-9}$) to the gag polyprotein of the SIRE-1 retrotransposon from soybean (*Glycine max*) (Laten and Morris 1993; Bi and Laten 1996) (Genbank accession number U22103). Weaker homology was also exhibited with the gag genes of the Opie-2 (San Miguel et al. 1996) (Genbank accession number U68408) (score = 57, $E = 1 \times 10^{-7}$) and PREM-2 (Turcich et al. 1996) (Genbank accession number U41000) (score = 46, $E = 4 \times 10^{-4}$) retrotransposons from maize (*Zea mays)*. Upstream of the gag-homologous region, a 327 bp sequence was defined by a 6-bp invert repeat sequence at either end. Within these borders, two areas of homology to the SV40 core enhancer sequence and a TATA box were identified. Immediately following the 3′ border was a 12-bp sequence, of which 11 nucleotides were an exact complement of the 3′ end of the wheat

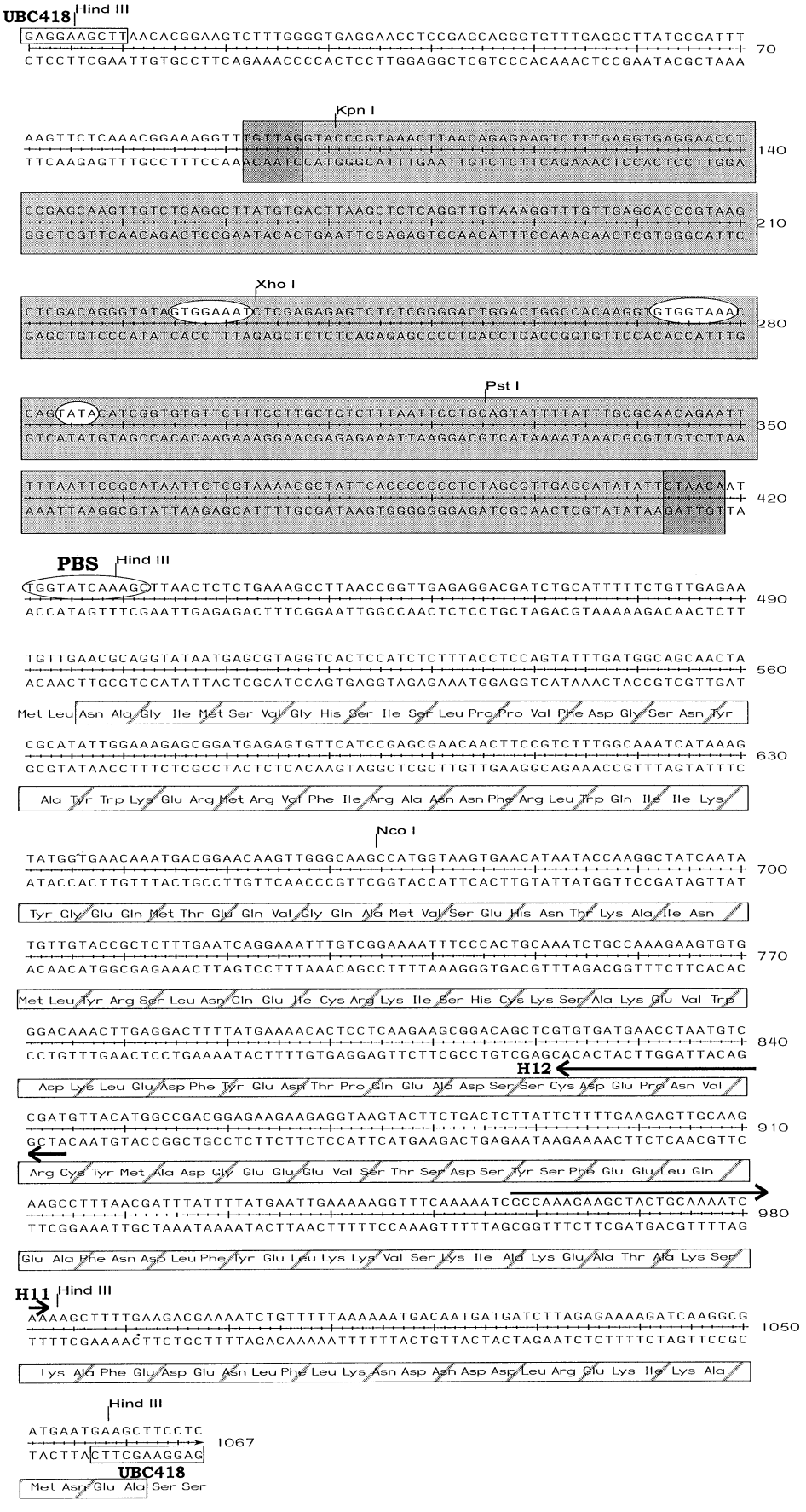
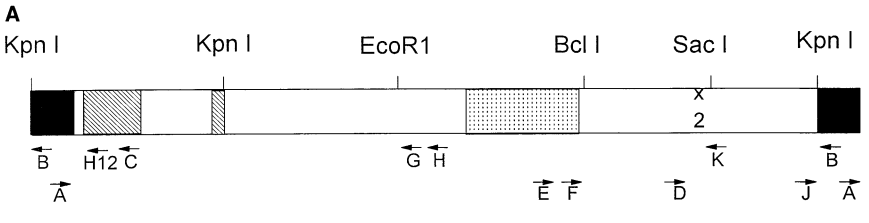


Fig. 1 Sequence of the UBC418$_{1050}$ RAPD marker clone. The putative LTR region and defining invert repeat borders are indicated in *light* and dark *grey*, respectively. Amino acid sequences demonstrating homology to gag protein genes are *outlined* by the *striped box*. The sequence homologous to the SV40 core enhancer sequence, the TATA box, and primer binding site (PBS) are identified *within elipses*. UBC 418 primer sequences are indicated *within rectangles*, and the H11 and H12 primers used for generation of the Pfu1.8 clone are marked by arrows



B

Primer	Orientation	Sequence (5'-3')	Position in retrotransposon
A	Sense	AAGTTGTCTGAGGCTTATGTGACTT	56-80, 6912-6936
B	Antisense	TTCCTCACCTCAAAGACTTCTCTG	46-23, 6902-6879
C	Antisense	TGCAAAGATTCACACGATAAACAAGTAAGACA	1018-987
D	Sense	AAGGACAGTACAAGTTCAAGAATCGCAAATG	5687-5717
E	Sense	GATCGTTCCCTACTTTACTTCACTGCT	4225-4251
F	Sense	AACGTGTTAATGTCGCTTTGTC	4510-4531
G	Antisense	TCAAGTGACACAATACAGATAGGAA	3663-3639
H	Antisense	TCCTTAAAGAAAATTCCCCCTCAA	3800-3777
H12	Antisense	ACATCGGACATTAGGTTCATCACAC	755-731
J	Sense	TTGAGAAGTCTTAGGGATAGGTT	6676-6698
K	Antisense	AGAAACTTACTGCCTCGTGACA	5980-5959

Fig. 2A,B Primer strategy for amplification of the full-length retrotransposon and associated sequences. **A** Relative position of primers within the full-length Tnd-1 retrotransposon. LTR, gag-, and pol-homologous regions are indicated by *black*, *striped*, and *dotted boxes*, respectively. The position and copy number of the 38-bp repeat sequence (*X*) are indicated. **B** Identity and sequence of indicated primers, and their actual position within the full-length retrotransposon

methionine initiator tRNA (Ghosh et al. 1982). These structural elements and the proximity of the 327-bp sequence to the gag-homologous region are characteristic of the long terminal repeats (LTR) which define the ends of retrotransposons (Boeke et al. 1989), indicating that the UBC418$_{1050}$ marker was part of the 5′ end of an LTR-type retrotransposon present within the *N. debneyi* genome.

Repeated attempts to build a RM12-3 genomic library failed to generate clones carrying the UBC418$_{1050}$ sequence, although other low-copy-number genes were present. Instead, the sequence flanking the UBC418$_{1050}$ marker was obtained by inverse PCR conducted on *Kpn*I-digested and ligated RM12-3 DNA using divergent primers, H11 and H12, in the gag-homologous region of the UBC418$_{1050}$ sequence (Fig. 1). A single *Kpn*I restriction site at the 5′ end of the putative LTR had been identified within the UBC418$_{1050}$ clone, and it was anticipated that the product of this inverse PCR would include sequences at the 3′ end of the UBC418$_{1050}$ clone from the interior of the retrotransposon. A 1.65-kb amplification product was obtained and cloned (Pfu1.8). Sequencing of the Pfu1.8 clone confirmed much of the existing UBC418$_{1050}$ sequence and yielded an additional 800 bp of downstream sequence to a second *Kpn*I site approximately 1.7 kb from the site present in the LTR (Fig. 2A). Gag-homology was present as initially observed in the UBC418$_{1050}$ clone and for an additional 42 amino acids later in the same ORF just upstream of the second *Kpn*I restriction site.

Assuming that a full-length retrotransposon did exist, the LTR sequence represented in the UBC 418$_{1050}$ clone was predicted to be repeated in the same orientation at the 3′ end of the retrotransposon. If a second LTR copy was located downstream of the cloned sequence, divergent primers homologous to the LTR sequences would amplify a product encompassing the complete interior of the retrotransposon. Based on this strategy, long-range PCR amplification was performed using RM12-3 genomic DNA and primers A and B (Fig. 2), resulting in the amplification of a 6.8 kb fragment designated RT. Sequencing confirmed the presence of sequences matching the UBC418$_{1050}$ LTR at either end of the cloned RT fragment. Amalgamation of the sequences for the UBC418$_{1050}$, Pfu1.8, and RT clones generated the full-length 7.182 kb Tnd-1 retrotransposon sequence (Fig. 2A). Analysis of this sequence showed that the *Xho*I and *Pst*I restriction sites previously identified in the LTR were not present in the remainder of the retrotransposon and that they were unique to the LTR. *Kpn*I restriction sites were present at the 5′ ends of both LTRs, and at one interior site 1.75 kb from the 5′ LTR *Kpn*I site, which was consistent with the position predicted from the Pfu1.8 clone. Unique *Eco*RI, *Bcl*I, and *Sac*I sites were found in the complete retrotransposon sequence. Immediately upstream of the *Sac*I site were two direct repeats of the 38-bp sequence "GTTGAGTCGTCTTTATTTTGTAATCTTGACTCAACGACT".

A number of potential ORFs were found throughout the retrotransposon, but these were interrupted by numerous stop codons in all reading frames. The longest ORF was the gag-homologous region, which terminated immediately downstream of the internal *Kpn*I site. A second area of interest started at nucleotide 3929 and terminated at nucleotide 4850, just prior to the *Bcl*I site. This 300-amino acid sequence was divided over three reading frames and interrupted by six termination codons, and had significant homology to the pol regions of the Opie-2 (San Miguel et al. 1996) (Genbank accession number U68408 (Score = 121, $E = 1 \times 10^{-49}$) and Hopscotch (White et al. 1994)

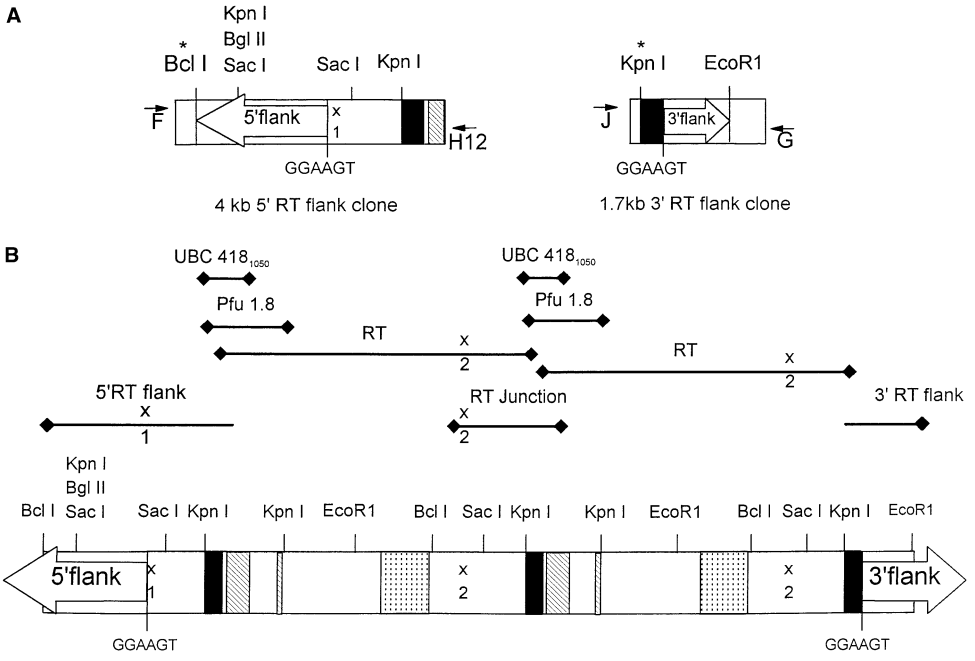


Fig. 3 **A** Structure of the nested inverse PCR clones carrying sequences flanking the 5′ and 3′ ends of Tnd-1. LTR, gag-, and pol-homologous regions are indicated by *black*, *striped*, and *dotted boxes*, respectively. The position and copy number of the 38-bp repeat sequence (*X*) are indicated. Primers used to generate each clone are represented by *arrows* and identified by *letter* as described in Fig. 2. Restriction/ligation sites used to generate DNA template for inverse PCR are marked (*). **B** Structure of the Tnd-1 insert in the introgressed *N. debneyi* DNA segment present in RM12-3. Isolated clones and their relative position within the genome insert are indicated. Each line (◆–◆) represents a full length clone except the 5′ RT and 3′ RT flank clones (–◆) where only retrotransposon sequences directly linked to the flanking sequence are shown

(Genbank accession number U12626) (Score = 66, $E = 5 \times 10^{-14}$) retrotransposons from maize and retrotransposons from *Volvox carteri* (Genbank accession number AF015276) (Score = 57, $E = 3 \times 10^{-13}$), *Arabidopsis thaliana* (Konieczny et al. 1991) (Genbank accession number X53973) (Score = 61, $E = 6 \times 10^{-13}$), and *Nicotiana tabacum* (Grandbastien et al. 1989) (Genbank accession number X13777) (Score = 50, $E = 6 \times 10^{-09}$). Homology was also found with several *Drosophila* copia-type retrotransposons.

Isolation of retrotransposon flanking regions

Sequences flanking Tnd-1 were isolated using inverse PCR and *Bam*H1- or *Sma*I-digested and ligated DNA from RM12-3. Neither restriction enzyme cut within the retrotransposon. Two divergent primers were designed that were specific to opposite ends of the retrotransposon: primer D would hybridize in sense orientation near the 38-bp repeat sequence 1.17 kb upstream from the 3′ LTR, and primer C would hybridize in antisense orientation 0.7 kb downstream from the 5′ LTR copy (Fig. 2). Surprisingly, inverse PCR from either template preparation produced identical 2.2-kb fragments. The same fragment was also amplified from uncut RM12-3 DNA. Sequencing of these putative junction clones showed that they included 1.17 kb of sequence identical to the 3′ end of the retrotransposon, 0.32 kb of LTR sequence, and 0.7 kb of sequence identical to the 5′ end of the retrotransposon. This result indicated that at least one additional copy of the Tnd-1 retrotransposon, either partial or complete, was present within the plant genome and directly linked to the full-length sequence previously isolated.

Based on the possibility of multiple-linked copies of the retrotransposon within the plant genome, an alternative PCR strategy was used to specifically amplify sequences associated with either the 5′ or 3′ end of the retrotransposon. Template DNA was prepared for inverse PCR by digesting RM12-3 genomic DNA with either *Eco*RI or *Bcl*I restriction enzymes and ligating. To amplify sequences associated with the 5′ end of the retrotransposon, we carried out inverse PCR using *Bcl*I-treated DNA template and divergent primers C and E (Fig. 2). The products of this PCR were then subjected to a second, nested PCR using primers F and H12, which resulted in amplification of two fragments of approximately 2.8 and 4 kb. The sequence associated with the 3′ end of the retrotransposon was amplified from *Eco*RI-treated DNA template using the divergent primers D and H, followed by primers G and J. Two DNA fragments, of approximately 1.7 and 3.5 kb, were generated.

Sequencing of the nested PCR clones showed that both the 2.8-kb fragment generated from *Bcl*I-treated-DNA and the 3.5 kb fragment generated from *Eco*RI-treated DNA included only retrotransposon sequences. The 4-kb *Bcl*I-nested PCR clone contained 364 bp of retrotransposon sequence from primer F to the *Bcl*I ligation site, 2246 bp of 5′ flanking sequence, 1231 bp of sequence corresponding to the 3′ end of the retrotransposon, one copy of the LTR, and 428 bp of sequence identical to the 5′ end of the retrotransposon to primer

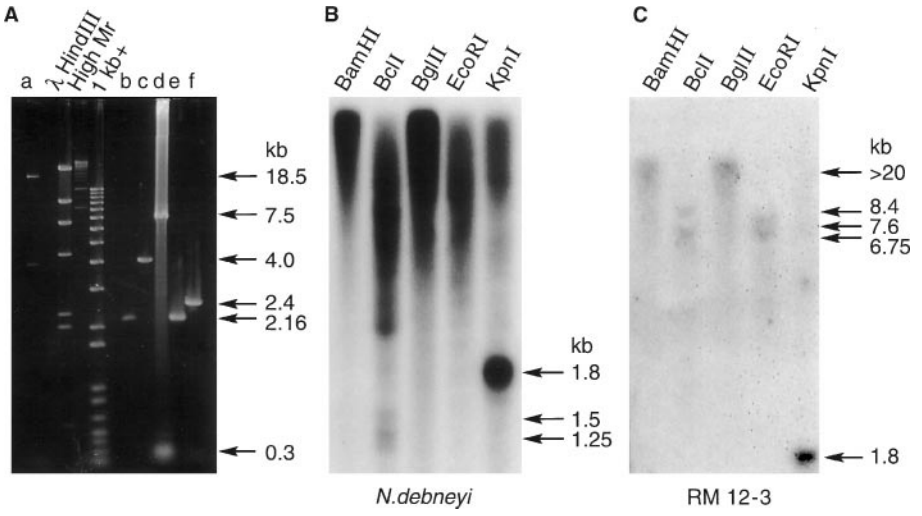


Fig. 4A–C Structural analysis of the genomic Tnd-1 retrotransposon insert. **A** PCR analysis of RM12–3. Retrotransposon fragments were amplified using (*a*) primers to the 5′ and 3′ flanking regions, (*b*) a primer to the 5′ flanking region and primer K to the 3′ end of the retrotransposon, (*c*) a primer to the 5′ flanking region and primer C to the 5′ end of the retrotransposon, (*d*) primers D and K to either side of the 38-bp repeat sequence, (*e*) primers C and D to either end of the retrotransposon, and (*f*) primer D to the 3′ end of the retrotransposon and a primer to the 3′ flanking region. Primers to the retrotransposon are described in Fig. 2. The 4-kb fragment in *lane a* does not consistently amplify: due to the length and repetitive nature of the insert complex, this fragment is believed to be a recombinant artifact of the PCR. **B** Southern blot analysis of *N. debneyi*. **C** Southern blot analysis of RM12-3. Blots were probed with the 5′ end of the Tnd-1 retrotransposon

H12 (Fig. 3A). This clone was similar to the RT junction clone isolated previously but was distinguished by the presence of only one copy of the 38-bp repeat sequence in the region equivalent to the 3′ end of the retrotransposon. The 1.7-kb clone amplified from *Eco*RI-treated DNA contained primer J, 180 bp of sequence identical to the 3′ end of the retrotransposon, one copy of the LTR, 1068 bp of 3′ flanking sequence to the *Eco*RI ligation site, and 472 bp of internal retrotransposon sequence up to primer G (Fig. 3A). Alignment of the various cloned sequences suggests a genomic organization that corresponds to one partial copy of the 3′ end of the retrotransposon linked to two directly repeated copies of the retrotransposon (Fig. 3B). This organization provides for two potential amplification sites within the insert for each of the UBC418$_{1050}$, Pfu 1.8, and RT sequences. Due to the strong sequence identity between these sites, however, the actual site from which the clones were generated cannot be determined. A 6-bp sequence "GGAAGT" is present at either end of the retrotransposon insert and is consistent with a target site duplication formed on insertion of the retrotransposon element into the plant genome.

Mapping the Tnd-1 genomic insert

PCR analysis of uncut RM12-3 genomic DNA was conducted to confirm the structure of the retrotransposon insert. Amplification reactions using a sense primer specific to the 5′ flanking region and an antisense primer specific to the 3′ flanking region yielded a product of approximately 18.5 kb (Fig. 4A, lane a). This size is comparable, within an acceptable margin of error, to the 17.7-kb size predicted for an insert complex consisting of one partial and two complete copies of the retrotransposon. A primer to the 5′ flanking region in combination with primer K (Fig. 2) to the 3′ end of the retrotransposon yielded a PCR product of 2.16 kb (Fig. 4A, lane b); the same 5′ flanking primer in combination with primer C (Fig. 2) to the 5′ end of the retrotransposon generated a 4-kb fragment (Fig. 4A, lane c). The sizes of these products confirmed the presence of a partial copy of the 3′ end of the retrotransposon at the 5′ end of the insertion complex. PCR products of 7.5 and 0.3 kb (Fig. 4A, lane d) were generated using primers D and K (Fig. 2). The larger fragment is equivalent to an internal copy of the retrotransposon, while the smaller fragment represents the region containing the 38-bp repeat sequence. Amplification of the 7.5-kb product indicates the duplication of sequences homologous to primer D and is further evidence of more than one copy of the retrotransposon. Multiple retrotransposon copies were further confirmed by the amplification of a 2.16-kb fragment (Fig. 4A, lane e) using primers C and D (Fig. 2). This size is comparable to the predicted 2.18 kb fragment that represents the junction between two copies of the retrotransposon. Continuity between the 3′ end of the retrotransposon and the 3′ flanking region was confirmed using primer D (Fig. 2) and an antisense primer specific to the 3′ flanking region. This amplification produced a 2.4-kb fragment (Fig. 4A, lane f) which approximated a predicted size of 2.55 kb.

Distribution of Tnd-1 in the genomes of '*N. debneyi*, 'Delgold', and RM12-3 was assessed by Southern blot analysis. A probe consisting of 90 bp of the LTR and 650 bp of the gag-homologous region did not hybridize to the 'Delgold' DNA (data not shown). It did, however, hybridize strongly to multiple restriction fragments in the *N. debneyi* DNA (Fig. 4B), indicating that a large number of retrotransposon copies were present in the *N. debneyi* genome. Three low-molecular-weight fragments could be clearly resolved from the *N. debneyi* blot: two *Bcl*I restriction fragments of 1.5 and 1.25 kb and a 1.8-kb *Kpn*I restriction fragment. Hybridization of the probe to RM12-3 DNA (Fig. 4C) gave results consistent with the retrotransposon insertion structure shown in Fig. 3. *Bam*HI and *Bgl*II digests each yielded a single fragment of 20 kb or larger. Since neither restriction enzyme cuts within the retrotransposon, these fragments were consistent with the expectation that each enzyme would generate restriction fragments in excess of 18.6 kb and encompassing the entire retrotransposon insertion complex. *Bcl*I digestion resulted in two restriction fragments of 8.4 and 6.75 kb, and *Eco*RI digestion gave two fragments of 7.6 and 6.75 kb. The 6.75 kb fragment common to both these digests is consistent with the expectation of a 6.85-kb internal copy of the retrotransposon. The 8.4-kb *Bcl*I restriction fragment is consistent with the predicted 8.35-kb fragment present between the 5′ flanking region and the closest internal *Bcl*I site. The 7.6 kb *Eco*RI restriction fragment suggests that an *Eco*R1 site is located approximately 1 kb upstream of the *Bcl*I site. *Kpn*I digestion generated a single 1.8-kb fragment which correlates with the 1.75-kb fragment encompassing the $UBC418_{1050}$ and Pfu1.8 clones. The greater intensity of probe hybridization to this restriction fragment relative to the *Bam*HI, *Bcl*I, *Bgl*II, and *Eco*RI fragments is also consistent with the expectation that this fragment is present at a two-fold higher copy number within the retrotransposon insertion.

Distribution of the Tnd-1 retrotransposon in *Nicotiana* species

The presence of the Tnd-1 retrotransposon within a species genome may serve as an indication of the relatedness of *Nicotiana* species. To test this hypothesis we subjected *Kpn*I-digested genomic DNA from a selection of *Nicotiana* species, varying in chromosome number and degree of relatedness, to Southern blot analysis using the retrotransposon LTR as a probe. The retrotransposon, or related homologues, was present in all representatives of the Sauveolentes (*N. sauveolens*, *N. exigua*, *N. maritima*, *N. velutina*, *N. benthaminana*, *N. goodspeedii*, *N. ingulba*, *N. rotundifolia*, *N. occidentalis*, *N. rotundifolia*, and *N. debneyi*) but was not present in representatives of the Alatae

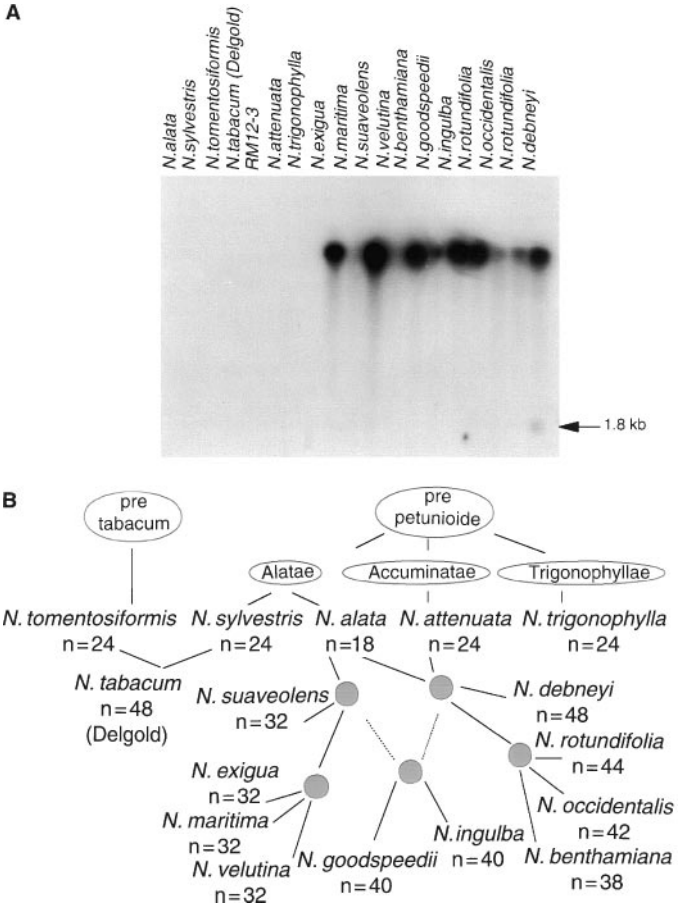


Fig. 5A, B Tnd-1 in *Nicotiana* species. **A** Southern blot analysis of *Kpn*I-digested genomic DNA probed with the Tnd-1 LTR sequence. *B* Phylogenetic relationship of *Nicotiana* species (Goodspeed 1954). Subgenus and section of each species are indicated in *open elipses*, while unknown progenitors presumed to have sequences homologous to Tnd-1 are represented by the *shaded elipses*. Unconfirmed relationships between species are indicated by a *dotted line*, whereas established relationships are indicated by a *solid line*. The somatic chromosome number (n) of each species is indicated

(*N. sylvestris* and *N. alata*) or Accuminatae (*N. attenuata*), the progenitor families of the Suaveolentes (Fig. 5).

Discussion

RM12-3 is a near isogenic line derived from the *N. tabacum* cultivar 'Delgold', it carries an introgressed DNA segment from *N. debneyi* which confers resistance to the fungal disease black root rot. The resistance phenotype is closely linked to the $UBC418_{1050}$ RAPD marker: no recombination event was observed in $48spF_2s$ that separates the resistance gene from this genetic marker (Bai et al. 1995). However, the exact size of the alien transfer DNA is difficult to estimate because recombination frequency is often lower between interspecific DNA segments (Chunvongse et al. 1994).

We have determined that the UBC418$_{1050}$ marker is part of a retrotransposon, designated Tnd-1. Both UBC418$_{1050}$ (Bai et al. 1995) and the retrotransposon (Fig. 5A) are absent in black root rot susceptible *N. tabacum* and are therefore part of the *N. debneyi* genomic component of RM12-3. Tnd-1 is the first retrotransposon to be isolated from *N. debneyi*.

Tnd-1 is 7183 bp long and has features typical of an LTR-type retrotransposon (Boeke and Corces 1989). The 327-bp LTR sequence is perfectly repeated at either end of the retrotransposon, contains areas of homology with the SV40 core enhancer sequence and a TATA box, and its borders are defined by the palindromic "TGTTAG … CTAACA" sequence. Immediately downstream of the 5′ LTR is a 12-bp primer binding site of which 11 nucleotides are complementary to the 3′ end of the wheat initiator tRNA^{met}. A polypurine tract, "AAACGGAAAGG", is located just upstream of the 3′ LTR. Within the retrotransposon, a large 457-amino acid ORF occurs that has homology to the gag protein, and another 300-amino acid sequence has homology to the pol genes of several copia-type retrotransposons. A concatomer of one partial copy and two complete Tnd-1 copies make up the retrotransposon insert in the black-root resistant RM12-3 genome. The insert itself is flanked by a 6-bp target site duplication of the sequence "GGAAGT".

Two types of LTR retrotransposons are defined by the arrangement of the functional domains within the pol region: in copia-type retrotransposons these domains are in the order protease-integrase-reverse transcriptase-RNase H, whereas in gypsy-type retrotransposons the integrase follows RNase H. Copia-type retrotransposons are considered to be relatively common among plants (Voytas et al. 1992). Database searches have revealed that the Tnd-1 sequences have homology to copia-type retrotransposons. The first 13 amino acids of the pol region were similar to conserved amino acids at the C-terminal end of the reverse transcriptases of the *Drosophila* copia, tobacco Tnt-1, and maize Hopscotch (White et al. 1994) and Opie-2 (San Miguel et al. 1996) retrotransposons. The remaining pol region homology was to RNase H, and there were no sequences homologous to protease or integrase. The absence of these homologies leaves domain ordering in Tnd-1 ambiguous, and Tnd-1 cannot be absolutely defined as either a copia- or gypsy-type retrotransposon. Stop codons appear frequently throughout the Tnd-1 pol region in all reading frames. A minimum of three frame shifts and suppression of five stop codons would be required to translate the RNase H domain. The two regions of gag-protein homology are present within a single ORF and might produce a functional protein, but reverse transcription PCR experiments (data not shown) using primers against the gag sequence were unable to find a Tnd-1 transcript. This failure, together with the apparent lack of pol-region functional domains, makes it unlikely that the retrotransposon isolated from RM12-3 would be mobile within the plant genome.

A mobile form of the Tnd-1 might persist in one of the related *Nicotiana* plants. In our survey of *Nicotiana* species, all representatives of the Suaveolentes hybridized to the Tnd-1 LTR probe, suggesting the existence of sequences homologous to Tnd-1. The Suaveolentes originate from Accuminatae and Alatae progenitors (Fig. 5B). The presence of Tnd-1 in all of the *N. suaveolens* and *N. debneyi* derivatives, and its absence from the members of the Alatae (*N. sylvestris* and *N. alata*), which is the purported source of the progenitor to both *N. debneyi* and *N. suaveolans*, suggests that Tnd-1 became part of the Suaveolentes genome after the development of the common Alatae progenitor. Polymorphisms were observed in the restriction map of the Tnd-1-homologous sequences resident in these plants: the distinctive 1.8-kb *Kpn*I restriction fragment was present only in *N. debneyi* and RM12-3 (Figs. 3 and 4), and two small *Bcl*I restriction fragments were detected in *N. debneyi* (Fig. 3) that are not present in the RM12-3 retrotransposon insert. Some polymorphisms are anticipated because of flanking restriction sites, but the *Kpn*I sites are associated with the LTR and gag-homologous region, and would be expected to be well conserved. Tnd-1 may therefore be one of a family of retrotransposons present within the Sauveolentes.

In RM12-3, resistance to black root rot has been linked to the single insertion complex of Tnd-1 retrotransposon described here. No ORF was found within the retrotransposon that could be related to the resistance. Similarly, the Tgm*r* retrotransposon which is linked to the *Rps1*-k allele for soybean resistance to *Phytophthora sojae* was found to be inactive and not to contain any apparent resistance gene (Bhattacharyya et al. 1997). Other such degenerate retrotransposon insertions have been found in close proximity to functional plant genes and have been suggested to be involved in gene duplication and regulation of gene expression (White et al. 1991; Marillonnet and Wessler 1997). Analysis of the immediate sequence flanking Tnd-1 has identified several potential intron/exon boundaries, but numerous stop codons present in all of the reading frames have precluded assembly of a recognizable ORF. Database searches have also failed to find any significant homology to known gene sequences. It therefore seems unlikely that insertion of the Tnd-1 retrotransposon into the genome has disrupted a protein function. The promoter elements present in the Tnd-1 LTRs might be able to activate or enhance expression of neighboring gene sequences, however we have not identified any potential gene(s) adjacent to the retrotransposon. Although Tnd-1 may not be a direct participant in the resistance phenotype, it is a significant species-specific marker for the resistance gene.

Acknowledgments We would like to thank Dr. Katherine Dobinson for her advice and critical reading of the manuscript.

References

Bai D, Reeleder R, Brandle JE (1995) Identification of two RAPD markers tightly linked with the *Nicotiana debneyi* gene for resistance to black root rot of tobacco. Theor Appl Genet 91: 1184–1185

Bhattacharyya MK, Gonzales RA, Kraft M, Buzzell RI (1997) A copia-like retrotransposon Tgm*r* closely linked to the Rps1-k allele that confers race-specific resistance of soybean to *Phytophthora sojae*. Plant Mol Biol 34: 255–264

Bi YA, Laten HM (1996) Sequence analysis of a cDNA containing the gag and prot regions of the soybean retrovirus-like element SIRE-1. Plant Mol Biol 30: 1315–1319

Boeke JD, Corces VG (1989) Transcription and reverse transcription of retrotransposons. Annu Rev Microbiol 43: 403–434

Cai D, Kleine M, Kifle S, Harloff HJ, Sandal NN, Marcker KA, Stiekema WJ, Wyss U, Grundler FMW, Jung C (1997) Positional cloning of a gene for nematode resistance in sugar beet. Science 275: 832–834

Chunwongse J, Bunn TB, Crossman C, Jiang J, Tanksley SD (1994) Chromosomal localization and molecular-marker tagging of the powdery mildew resistance gene (*Lv*) in tomato Theor Appl Genet 89: 76–79

Clayton EE (1969) The study of resistance to the black root rot disease of tobacco. Tob Sci 13: 30–37

Ghosh HP, Ghosh K, Simsek M, RajBhandary UL (1982) Nucleotide sequence of the wheat germ cytoplasmic initiator methionine transfer ribonucleic acid. NAR 10: 3241–3247

Giroux MJ, Clancy M, Baier J, Ingham L, McCarty D, Hannah LC (1994) *De novo* synthesis of an intron by the maize transposable element *Dissociation*. Proc Natl Acad Sci USA 91: 12150–12154

Goodspeed TH (1954) The genus *Nicotiana*. Chron Botanica Co., p 536

Grandbastien MA, Speilmann A, Caboch M (1989) Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature 337: 376–380

Hoffbeck LJ, Neas MO, Heggestad HE, Skoog HA (1965) 'Burley 49', a new disease-resistance burley tobacco. Tenn Agric Exp Stn Bull 395

Kidwell MG, Lisch D (1997) Transposable elements as sources of variation in animals and plants. Proc Natl Acad Sci USA 94: 7704–7711

Kloeckner-Gruissem B, Freeling M (1995) Transposon-induced promoter scrambling: a mechanism for the evolution of new allelles. Proc Natl Acad Sci USA 92: 12836–1840

Konieczny A, Boytas DF, Cummings MP, Ausubel FM (1991) A superfamily of *Arabidopsis thaliana* retrotransposons. Genetics 127: 801–809

Kumar A, Pearce SR, McLean K, Harrison G, Heslop-Harrison JS, Waugh R, Flavell AJ (1997) The Ty-1 *copia* group of retrotransposons in plants: genomic organisation, evolution and use as molecular markers. Genetica 100: 205–217

Laten HM, Morris RO (1993) SIRE-1, a long interspersed repetitive DNA element from soybean with weak sequence similarity to retrotransposons: initial characterization and partial sequence. Gene 134: 153–159

Legg PD, Litton CC, Collins GB (1981) Effects of the *Nicotiana debneyi* black root rot resistance factor on agronomic and chemical traits in burley tobacco. Theor Appl Genet 60: 365–368

Marillonnet S, Wessler SR (1997) Retrotransposon insertion into the maize *waxy* gene results in tissue-specific RNA processing. Plant Cell 9: 967–978

Moor JK, Haber JE (1996) Capture of retrotransposon DNA at sites of chromosomal double-strand breaks. Nature 383: 644–646

Osborne BI, Baker B (1995) Movers and shakers: maize transposons as tools for analyzing other plant genomes. Curr Opin Cell Biol 7: 406–413

Palmiter RD (1974) Magnesium precipitation of ribonucleoprotein complexes: expedient techniques for the isolation of undegraded polysomes and messenger ribonucleic acid. Biochemistry 13: 3606–3615

SanMiguel PJ, Tiknonov A, Jin Y-K, Motchoulskaia N, Zakharov D, Melake-Berham A, Springer PS, Edwards KJ, Lee M, Bennetzen JL (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274: 765–768

Turcich MP, Bokhari-Riza A, Hamilton DA, He C, Messier W, Stewart CB, Mascarenhas JP (1996) PREM-2, a copia-type retroelement in maize is expressed preferentially in early microspores. Sex Plant Reprod 9: 65–74

Voytas DF, Cummings MP, Konieczny A, Ausubel FM, Rodermel SR (1992) Copia-like retrotransposons are ubiquitous among plants. Proc Natl Acad Sci USA 89: 7124–7128

White SE, Habera L, Wessler SR (1994) Retrotransposons in the flanking regions of normal plant genes: a role for copia-like elements in the evolution of gene structure and expression. Proc Natl Acad Sci USA 91: 11792–11796