J Mol Evol (1997) 45:137–144

JOURNAL OF MOLECULAR EVOLUTION

Characterization of the *Sol3* Family of Nonautonomous Transposable Elements in Tomato and Potato

Teruko Oosumi, William R. Belknap

United States Department of Agriculture, Agricultural Research Service, Western Regional Research Center, 800 Buchanan Street, Albany, CA 94710, USA

Received: 18 September 1996 / Accepted: 11 March 1997

Abstract. *Sol3* transposons are mobile elements defined by long terminal inverted repeats which are found in tomato and potato. Members of the *Sol3* family have been isolated from a variety of solanaceous species including *Solanum tuberosum* (potato), *S. demissum, S. chacoense, Lycopersicon esculentum* (tomato), and *L. hirsutum.* While highly conserved elements are found within different species, *Sol3* terminal inverted repeats can also flank unrelated sequences. Southern blot analysis indicates that *Sol3* elements are less prevalent in the potato (approximately 50 copies) than in the tomato (>100 copies) genome. No *Sol3*-hybridizing sequences were observed in tobacco. While a number of *Sol3* elements ranging in size from 500 bp to 2 kbp were sequenced, no transposase coding domains could be identified within the internal regions of the elements. The data suggest that the *Sol3* represent a heterogeneous family of nonautonomous transposable elements associated with an as-yet-unidentified autonomous transposon.

Key words: Transposon — Solanaceous — Plant — Inverted repeat

Introduction

Mobile DNA elements are classified into major divisions based upon transposition via RNA or DNA intermediates (Berg and Howe 1989). Transposable elements with RNA intermediates, retroelements, are further divided into the viral and nonviral superfamilies. Members of the viral superfamily, including retroviruses and retrotransposons, are characterized by long terminal direct repeats (LTRs) flanking reverse transcriptase coding domains. Members of the nonviral or retroposon superfamily, including SINEs and LINEs, lack LTRs (Deininger 1989; Hutchison III et al. 1989).

Mobile elements which transpose through DNA intermediates are generally flanked by short terminal inverted repeats (TIRs), which in the autonomous forms flank a transposase encoding domain. This transposase mediates the DNA excision/recombination reaction of the elements (Lampe et al. 1996). The short TIR class of elements includes *Tn3* of bacteria (Heffron et al. 1979), the *P* (O'Hare and Rubin 1983) and the *Hobo* (Streck et al. 1986) elements of *Drosophila,* the *Ac* (Fedoroff et al. 1983; Pohlman et al. 1984), and the *Spm* elements of maize (Masson et al. 1987; Pereira et al. 1986), the Tc2 (Ruvolo et al. 1992) elements of *Caenorhabditis elegans,* and the Tc1-*mariner* superfamily widespread in animals (Oosumi et al. 1995a; Plasterk 1996; Rosenzweig et al. 1983). The transposase encoded by the autonomous elements can act in *trans* to catalyze the transposition of related, nonautonomous, elements with conserved TIRs. The transposase recognizes sequences within the TIRs or repeat motifs within subterminal regions of the transposable element (Gierl et al. 1988; Kunze and Starlinger 1989; New et al. 1988; Vos et al. 1993). Some members of second category, including the foldback (FB) elements of *Drosophila* (Potter et al. 1980), the TU1 elements of

Abbreviation: TIR, terminal inverted repeat

Correspondence to: W. Belknap; *e-mail:* wrb@pw.usda.gov

sea urchin (Liebermann et al. 1983), the Tc4 (Yuan et al. 1991) and the Tc5 (Collins and Anderson 1994) elements of *C. elegans,* and the *Mu* elements of maize (Barker et al. 1984), are characterized by long TIRs. In the case of FB, TU1, and Tc4, there are short sequence repeats in the outer domain of the TIRs (Liebermann et al. 1983; Truett et al. 1981; Yuan et al. 1991).

This paper presents the characterization of a family of transposable elements, *Sol3,* in solanaceous plants. The original *Sol3* element was identified by its inverted repeated structure in the promoter of tomato polygalacturonase gene (Tom*PG*) (Bird et al. 1988; Oosumi et al. 1995b). The *Sol3* element within this promoter is of particular interest because previous experiments (Montgomery et al. 1993) showed that both a positive regulatory region and multiple nuclear protein binding domains are contained within the transposable element. This transposable element, therefore, represents a potential mobile enhancer element functioning within the plant genome. In a number of other cases transposable elements have been shown to contribute important nuclear protein binding domains involved in transcriptional regulation of eukaryotic genes (Britten 1996). These transposable elements provide a potential molecular mechanism for "enhancer shuffling" (Reue et al. 1988).

The *Sol3* element within the tomato polygalacturonase promoter consists of 250-bp TIRs flanking a 335-bp internal sequence. We have previously identified a similar element from potato (Oosumi et al. 1995b) as well as two smaller elements, from potato and tomato, which appear to be severely deleted members of this transposon family. Sequence similarity among all four loci is limited to the inverted repeated domain.

Here we report further characterization of the *Sol3* family. Southern blot analysis indicates that these elements are present in high copy number in tomato with a lower frequency in potato. Southern hybridization data indicates that no members of this family are present in another solanaceous species, tobacco. The *Sol3* elements do not share internal sequence similarity with any known autonomous transposon nor do they contain coding sequences for a transposase enzyme. This suggests that the *Sol3* elements represent a family of heterogeneous nonautonomous elements mobilized by a transposase encoded by an unknown autonomous transposon.

Materials and Methods

Plant Materials. Genomic DNA preparations from *Solanum chacoense, S. brevidans,* and *S. demissum* were a kind gift from D. J. Hannapel (Iowa State University). Genomic DNA from *S. tuberosum* cvs. Lemhi Russet and Lenape, *Lycopersicon esculentum* cvs. T27 and VF36, and *L. hirsutum* was prepared as described by Draper and Scott (1988). Tobacco genomic DNA preparations from *Nicotiana. sylvestris* and *N. tabacum* cv. SR1 were a kind gift from B. Baker (USDA/ARS, University of California, Berkeley).

Southern Blot Analysis. Genomic DNA from *L. esculentum* cv. VF36, *L. hirsutum,* and *N. tabacum* cv. SR1 was digested with *Eco*RI. Genomic DNA from *S. tuberosum* cv. Lemhi was digested with *Eco*RI and *Hinc*II. Fragments were separated on 0.8% agarose gels and transferred to GeneScreenPlus membranes (Du Pont-NEN) according to manufacturer's recommendations.

Two probes were generated by PCR amplification from the Pot*ten1* clone. A 343-bp TIR probe was amplified using two primers (Potten5′-4, 5′-TTAGAAACACTCATTGGCCG-3′; IROE3′, 5′-GGAAAATGGCCTAAAATATCC-3′). A 562-bp internal probe was produced by the single primer TIR3 (5′-CCACCCAATTAATAAACC-3′) (Fig. 1). The PCR mixture consisted of 10 μM Tris-HCl (pH 8.3), 50 μM KCl, 2 mM $MgCl_2$, all four dNTPs (each at 200 μM), 0.2 μM each primers (or 0.4 μM single primer), 5 units of Taq DNA polymerase (Perkin Elmer), and lambda DNA in a total volume of 100 μl. A step program of 1 min at 94°C, 30 s at an annealing temperature (49°C for TIR3, 55°C for Potten5′-4 and IROE3′), and 1 min at 72°C was repeated 30 times. Probes were labeled with [α-^{32}P]dCTP, using oligolabeling kit (Pharmacia) according to manufacture's recommendations.

The filters were hybridized at 65°C for 17 h in a solution of 1 M NaCl, 1% SDS, 10% dextran sulfate, 200 μg/ml salmon sperm DNA with the ^{32}P-labeled internal probe. The filters were washed at 65°C two times in 2 × SSC, 1% SDS for 40 min. Following the autoradiography, the filters were stripped and then rehybridized with the ^{32}P-labeled terminal repeat probe as described above.

Amplification of Sol3 *Elements from Solanaceous Genome.* The primers used to amplify internal regions of *Sol3* elements were designed using the internal region of the 5′ TIR of Pot*ten1* (Fig. 1): Primer TIR1, 5′-CACTACCCACCCATTACT-3′; Primer TIR2, 5′-CACTACCCACCCATTACTAATTAAA-3′; and Primer TIR3, CCACCCAATTATAAACC. The PCR mixture was the same as above, except that 40 ng of genomic DNA was used as template. A step program of 1 min at 94°C, 30 s at Tm of primer, 3 min at 72°C was repeated 31 times followed by a final extension at 72°C for 2 min. The products from the first round PCR using primer TIR1 were diluted to 2.5-fold, and 1-μl aliquots of the dilution were used for a second-round PCR using primer TIR3.

The TIR domain of the *Sch2* element was amplified from genomic DNA of *S. chacoense* and *S. tuberosum* cv. Lemhi Russet using the TAIL-PCR method (Liu et al. 1995). Two specific primers (2kSc1, 5′-CGCAAAGTCACCTAGAATTCCCCTC-3′; 2kSc2, 5′-CCCTCTAAATTGAAGACATCATC-3′), corresponding to the internal sequences, and a mixed primer, TIR4, 5′-T(C/T)AAAATTGACCA(C/T)TA-3′, from the TIRs of Tom*PG* and Pot*ten1* elements, were synthesized. The PCR mixture was the same as above. Thermal cycling was performed as described by Liu et al. (1995), except that an annealing temperature of 40°C was used in the reduced stringency cycles. The tertiary reaction was omitted.

Cloning of PCR Products. PCR products were cloned into pCRII or pCR2.1 using the TA cloning kit (Invitrogen) as directed by the manufacture.

DNA Sequence Analysis. Plasmid DNA isolation and sequencing were performed using a plasmid kit (QIAGEN) and fmol DNA sequencing system (Promega), respectively. DNA sequences were analyzed using MACVECTOR (Eastman Kodak). Sequence alignment was performed using ASSEMBLYLIGN (Eastman Kodak). Database search using the BLAST algorithm were performed via electric mail to the National Center for Biotechnology Information at the National Library of Medicine (Altschul et al. 1990).

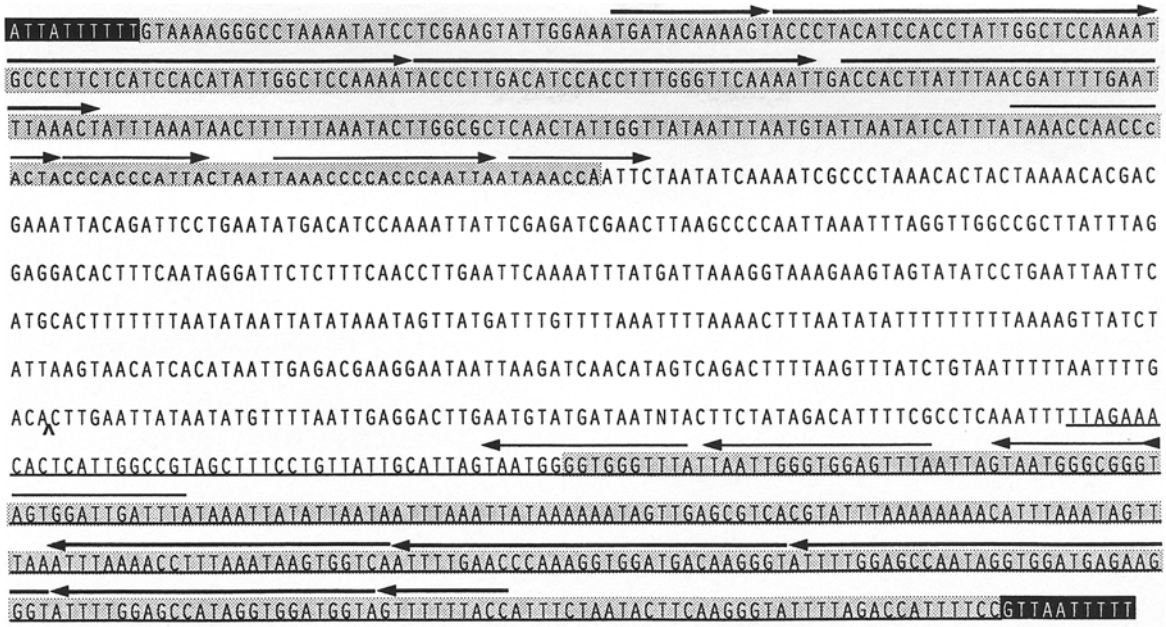


Fig. 1. Nucleotide sequence of the Pot*ten1* element. The TIRs are *shaded* and the flanking 11-bp sequences are indicated with the *black background.* The 31-bp repeat sequences are indicated by *thick arrows* and 15-bp repeat sequences are indicated by *thin arrows.* The TIR probe for Southern blot analysis (Figs. 3 and 4) is *underlined.* Nucleotides in the 5′ TIR in *small print* indicate source sequence source for the three PCR primers used for amplification of *Sol3* family members. The position of the transposon-like insertion within *Lhi1.4* (Fig. 5) is indicated by (^).

Results

Sol3 *Inverted Repeated Elements*

Four *Sol3* elements have been previously identified (Oosumi et al. 1995b) in the tomato polygalacturonase (Tom*PG*) and RuBP Carboxylase small subunit (LERBCS2) genes, a potato patatin psuedogene (STPATP1), and the Pot*ten1* element. The sequence of the Pot*ten1* element is similar to the element in tomato polygalacturonase (Tom*PG*) except that Pot*ten1* element has an additional 187 bp immediately 5′ to the 3′ TIR. The sequence of the Pot*ten1* element is shown in Fig. 1. The inverted repeats of the element are approximately 295 bp in length and consist of three regions: an outer region composed of tandem repeats of 31-bp sequence (consensus ACCCTTNNCATCCACCTATTGGCTCCAAAAT) and an inner region composed of tandem repeats of a 15-bp sequence (consensus TAAACCCACCCATTA) separated by an AT-rich domain.

The smaller elements in STPATP1 and LERBCS2 have shorter TIRs (115 and 108 bp, respectively). The element in LERBCS2 appears to be a deleted form of the STPATP1 element (Fig. 2A). While the internal domain of the STPATP1 element has no obvious similarity to the same region of Pot*ten1,* the repeats of these smaller elements are clearly related to the distal regions of the TIRs of the larger *Sol3* transposons (Fig. 2B).

The Pot*ten1* element is flanked by an 11-nucleotide imperfect direct repeat apparently generated by duplication at the recombination target site (Oosumi et al. 1995b). The *Sol3* elements in the Tom*PG* and STPATP1 loci are similarly flanked by imperfect direct repeats of 11 and 10 bp, respectively (Fig. 2A). All three insertion site sequences have the structure (A/T)(AT)NN(A/T)(A/T)(A/T)(A/T)TTT.

The structural similarity among these four elements suggests that they are members of the same family of transposable elements.

Characterization of Sol3 *Copy Number in Solanaceous Species*

Given the limited amount of sequence from solanaceous plants currently available in the database, the computational identification of three *Sol3* family members suggested that these elements are common in potato and tomato. In order to determine the relative copy numbers, Southern hybridization to DNA from four solanaceous species, *S. tuberosum* (potato), *L. esculentum* (tomato), *L. hirsutum,* and *Nicotiana tabacum* (tobacco), were carried out. Hybridization using the internal region of the Pot*ten1* transposon (Fig. 1) indicates that there are approximately 10 copies of this element in the potato genome (Fig. 3A), with many more copies in the *Lycopersicon* species (Fig. 4A). No detectable hybridizing bands were observed when tobacco DNA was hybridized to this probe (data not shown).

A

```
STPATP1 GTATCATAAAACTATTTTTGGGAAAAA-----GGATAAATATACCCCT-GAACTATCGTAAATGGTATGTAAATACCCTTT
LERBCS2 TCACTAGTCTGTATAAAAA        AAAGAA             T A   -    A   G       C G   -T   -

STPATP1 CG-TCATACTTTTGGG-CATTGTTGCCCCTGCCGCCGTCCAAAAACTAGAGCATATATACCCTTTAAACTAACGGACATAC
LERBCS2   G     G       A     G A TT     ---              A G TG  ----------------------

STPATP1 ACGTGTCATAATATTATCCGCCGATTGAATATCGAATCGACAGATATGATTGTGTCACGTGTCCCGATTTAAACTTCCGTT
LERBCS2 ------------------------------------------------------------------------------

STPATP1 AAAGTGAAGGGCATATATGCTCTAGTTTTTAAACGGCAGGGACATCAATGTCCCAATAGTATGACGGATGGTATCTGCATA
LERBCS2 -----------       T G        GG       AAG              A A     T A AAA      T

STPATP1 CCATTTACGATAATTCGGGAATATATTTGTCC-TTTTTTCCAATATTTTTTCGTGTATGTATGTAAGTGTGACATTACTAT
LERBCS2  T          G     --           C       T  GT    A AAA  TAAA A  T AAA TTTGAAT CGC
```

B

```
POTTEN1 (5')   -GTAAAA----GGGCCTAAA-ATATCCTCGAAGTATTGG-AAATGATA--CAAAAGTACCCT-ACATC
        (3')   -GGAAAA----TGGTCTAAA-ATACCCTTGAAGTATTAG-AAATGGTA--AAAAACTACC-----ATC
STPATP1 (5')   GGGAAAA----AGGA-TAAATATACCCCTGAACTAT-CGTAAATGGTATGTAAA--TACCCTTTCGTC
        (3')   GGAAAAA----AGGA-CAAATATATTCCCGAATTAT-CGTAAATGGTATGCAGA--TACCAT-CCGTC
LERBCS2 (5')   GGGAAAAAAAAGAAGATAAATATACCCTTAGAATAT-CATAAGTGGTATGCAGA--TATCTT-CGGTC
        (3')   GAAAAAA----GGGA-CAAATATATCC--GAACTAT-CGTAAATAGTATACAGA--TATTTT-TCATC
```

Fig. 2. Alignment of the *Sol3* transposable elements in STPATP1 and LERBCS2 genes. **A** GenBank loci STPATP1 (positions 1667–2062) and LERBCS2 (positions 2035–2315) are aligned with only the mismatched bases shown and gaps are indicated by *dashes*. The TIRs are indicated by *shading*. **B** Alignment of the *Sol3* TIRs from STPATP1, LERBCS2, and Pot*en1*. Gaps are indicated by *dashes*.

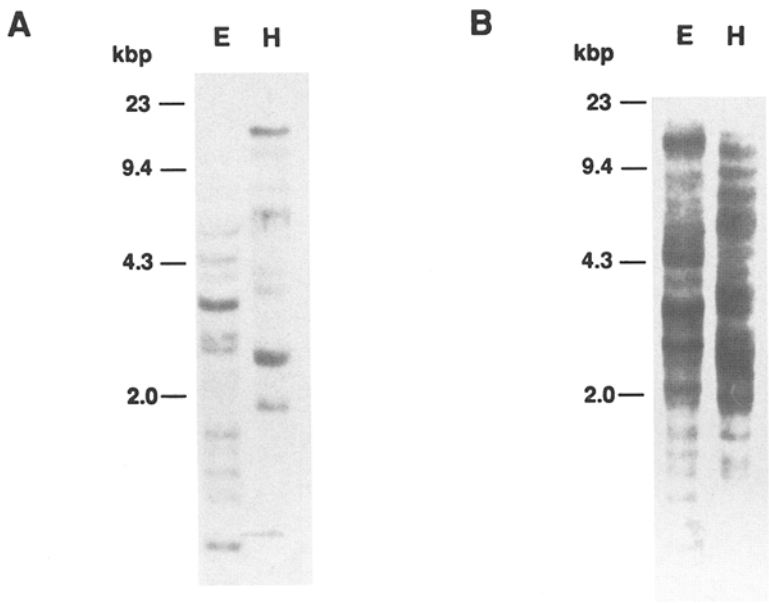


Fig. 3. Southern blot analysis of *Sol3* sequences in potato genomic DNA; 10 μg of genomic DNA from *Solanum tuberosum* cv. Lemhi Russet was digested with *Eco*RI (E) and *Hinc*II (H). **A** The genomic DNA was hybridized with the Pot*en1* internal probe generated by PCR amplification using a primer TIR3 (see Materials and Methods section). **B** Genomic DNA rehybridized with the Pot*en1* TIR probe indicated in Fig. 1.

In order to detect severely deleted *Sol3* elements (Fig. 2), these DNA samples were also probed with *Sol3* TIR sequences from Pot*en1* (Figs. 3B and 4B). As expected, more hybridizing bands were detected in *S. tuberosum* and the *Lycopersicon* species using the TIR probe relative to internal transposon sequences. Once again, no detectable bands were observed when the TIR probe was hybridized to tobacco DNA (data not shown).

These data indicate that the potato genome contains approximately 50 *Sol3* elements, while the *Lycopersicon* species contain hundreds of copies of members from this transposon family. The specificity of the probes and hybridization conditions employed is indi-

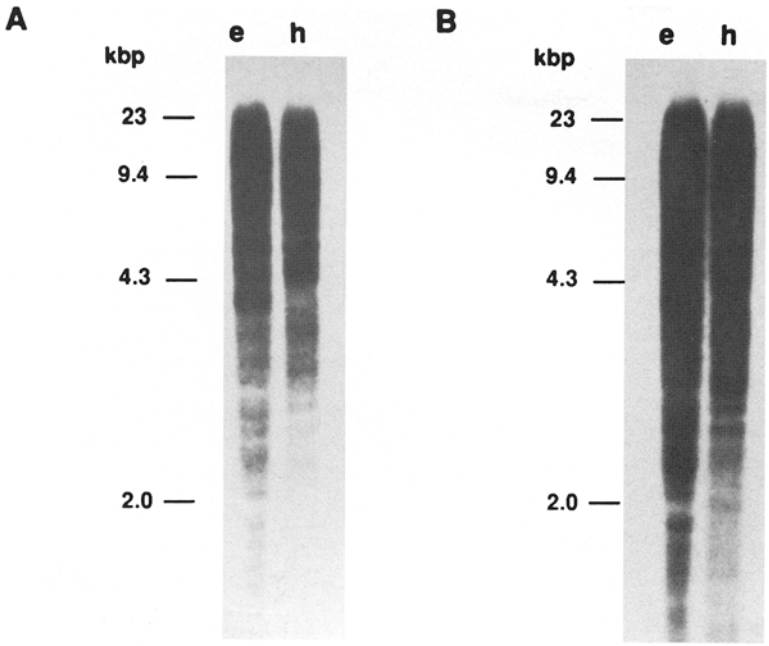


Fig. 4. Southern blot analysis of *Sol3* sequences in genomic DNA. of *Lycopersicon* species; 10 μg of genomic DNA from *Lycopersicon esculentum* cv. VF36 *(e)* and *Lycopersicon hirsutum (h)* was digested with *Eco*RI. **A** The genomic DNA was hybridized with the Pot*ten1* internal probe (see Fig. 3), and then **B** rehybridized with the Pot*ten1* TIR probe.

cated by the absence of hybridization to tobacco DNA.

PCR Amplification and Cloning

PCR analysis was carried out to amplify the internal sequences of *Sol3* elements using single primers corresponding to the TIRs (Oosumi et al. 1995a). DNA templates were derived from a variety of solanaceous species. Given the observed divergence in the TIRs from these elements, designation of a consensus sequence for primer design was difficult. Therefore, three oligonucleotide primers were employed which corresponded to the conserved inner domain in the inverted repeats of the Tom*PG* and Pot*ten1* transposons (Fig. 1). Each of the three primes was used individually in the PCR reactions.

When Primer TIR1 (Fig. 1) was used, multiple products ranging from 500 to 700 bp were amplified from the *Lycopersicon* species, while the *Solanum* species *(tuberosum, brevidans, chacoense,* and *demissum)* produced a limited number of 550-bp products (data not shown). In addition, *L. hirsutum, L. esculentum* cv. T27, *S. chacoense,* and *S. demissum* produced single larger products of 1.4, 2.0, 2.0, and 2.5 kbp, respectively. The products from the first-round PCR were subjected to a second round of amplification using Primer TIR3 internal to the Primer TIR1 (Fig. 1). The 500–600-bp products from the *Lycopersicon* species and *S. tuberosum,* as well as the 1.4-kbp product from *L. hirsutum* were reamplified. When Primer TIR2 (Fig. 1) was used for first-round PCR, once again, multiple small (500–600 bp) products were generated from the *Lycopersicon* species. However, in the case of the *Solanum* species, products of this size were amplified only from *S. tuberosum* and *S. brevidans*). The 2.0-kbp PCR product from *S. chacoense* was once again amplified using Primer TIR2.

When tobacco DNA was used as a template, Primers TIR1 and TIR3 yielded no amplification products, while Primer TIR2 amplified a single 550-bp band. The Southern hybridization data discussed earlier suggests that this PCR product is an artifact.

Three 500–600-bp products of *L. esculentum (Les0.5-1, Les0.5-2,* and *Les0.6* in Fig. 5), the 1.4-kbp product of *L. hirsutum* (Fig. 5, *Lhi1.4*), the 2.0-kbp product of *S. chacoense* (*Sch2*), and the 500-bp product of *S. demissum* (Fig. 5, *Sde0.5*) were cloned and subjected to sequence analysis.

Sequence Analysis of Sol3 *Family Members*

The structures of the five PCR clones that share internal similarity with the Tom*PG* and Pot*ten1* elements are shown in Fig. 5. The 1.4-kbp PCR product generated from *L. hirsutum* (*Lhi1.4*) has a 924-bp domain not found in the related elements (Figs. 5 and 6). Three lines of evidence suggest that this domain represents another transposable element which has recombined into the *Sol3* element in this locus. First, the sequence similarity between regions flanking the putative element and Pot*ten1* suggests an insertion event. Second, the putative element is bordered by 8 bp of imperfect direct repeats consistent with a target site duplication (Fig. 6). Finally, the termini of the 38 bp of imperfect TIRs which define the putative element have similarity to previously described plant transposons (Fig. 6).

The internal sequences of the seven elements (including Tom*PG* and Pot*ten1*) are highly conserved, except that the Tom*PG* element has a deletion of 187 bp

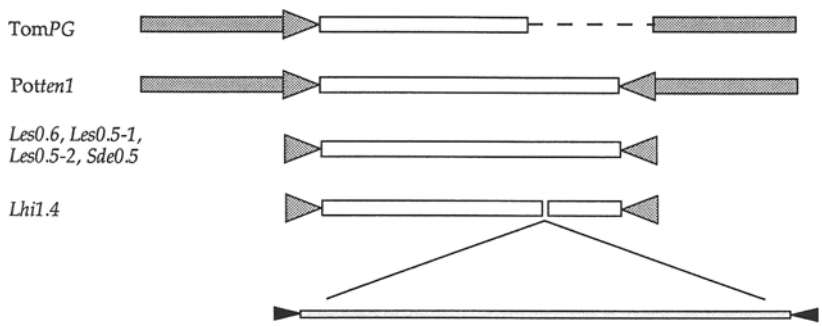
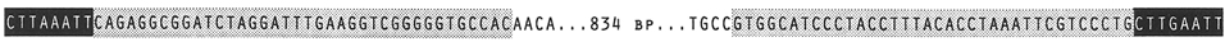


Fig. 5. Structure of the Pot*ten1*-type *Sol3* elements. Alignment of elements from potato (Pot*ten1*, GenBank locus U91987 positions 361–1469), tomato (Tom*PG*, GenBank locus TOMPGAAA positions 147–1000; *Les0.6*, GenBank locus U91991; *Les0.5–1*, GenBank locus U91989; *Les0.5–2*, GenBank locus U91990), *Solanum demissum* (*Sde0.5*, GenBank locus U91992), and *Lycopersicon hirsutum* (*Lhi1.4*, GenBank locus U91988). *Sol3* TIR domains are indicated by the *shaded arrows;* a gap is indicated by *dashed line.* The position of the transpson-like insertion within the *Lhi1.4 Sol3* element is indicted (Fig. 1), as defined by 38-bp imperfect inverted repeats shown as *black arrows.*

A



B

Fig. 6. Terminal inverted repeat domains of the insertion in the internal region of the *Lhi1.4* element. **A** Nucleotide sequence of the terminal inverted repeats and the flanking 8-bp directed repeats and putative TIRS (indicated as in Fig. 1) of the insertion found in the *Lhi1.4* internal region (Fig. 5). **B** Alignment of the terminal inverted repeats of the insertion in the *Lhi1.4* element with *Ac* and *dTph1* transposons.

(Oosumi et al. 1995b). The deletion in the Tom*PG* element includes the junction of the internal region and the 3′ TIR. Therefore, the Box C positive regulatory domain identified by Montgomery et al. (1993) is composed of the two sequences flanking this deletion. Three clones from *L. esculentum* and one from *Solanum demissum* show different patterns of smaller (<30 bp) deletions and insertions relative to each other and to the Tom*PG* and Pot*ten1* elements (Fig. 5).

The final PCR clone, *Sch2* (GenBank locus U91993), shares only inverted repeat sequence homology with the Pot*ten1*-type elements (data not shown). To confirm that the inverted repeats of *Sch2* are not artificial, one arm of this element was amplified by TAIL-PCR (Liu et al. 1995). For this reaction, two specific oligonucleotides were designed using the internal sequence of the element and a shorter oligonucleotide was designed using an outer sequence of Tom*PG* and Pot*ten1* TIRs. PCR using genomic DNA from either *S. chacoense* or *S. tuberosum* as template resulted in amplification of a 300-bp product. Sequence analysis of the product revealed similarity to the internal sequences of *Sch2*. The remainder of the 300-bp product shares sequence similarity with the TIR of Pot*ten1*, indicating that the *Sch2* is a member of the *Sol3* family.

When the internal region of *Sch2* was used as a database query (Altschul et al. 1990), two domains associated with genes from solanaceous plants were identified. A 107-bp domain at immediately 3′ to the 5′ TIR is similar to a sequence in intron 5 of a patatin gene (STPATG). In addition, a 125-bp sequence 1,412 bp 3′ to the 107-bp domain shares sequence similarity with five sequences that are located in the 5′ or 3′ flanking regions or introns of solanaceous genes (*L. esculentum* vacuolar invertase, *L. pimpinellfolium* vacuolar invertase, *S. tuberosum* proteinase inhibitor II, *S. tuberosum* patatin, and *S. tuberosum* sucrose synthase). The data indicate that *Sol3* TIRs define a heterogeneous family of nonautonomous transposable elements.

Discussion

We have previously described the use of computational analysis of the available database to identify transposable elements in solanaceous plants (Oosumi et al. 1995b), *C. elegans* (Oosumi et al. 1995b, 1996), and humans (Oosumi et al. 1995a). Of the elements identified in solanaceous plants, the *Sol3* family was of particular interest because the original family member identified in the tomato polygalacturonase gene contributed key regulatory elements to its promoter domain (Oosumi et al. 1995b; Montgomery et al. 1993). The initial characterization of this transposon family was limited to cloning of an additional *Sol3* element from potato and identifi-

cation of recombination target sequences. Here we present a more thorough characterization of this family of mobile elements.

Structural features of the *Sol3* family, in particular the long TIRs, suggest that they belong to the class of mobile elements which transpose via DNA intermediates. The *Sol3* elements are defined by long TIRs. Similar to the FB elements from *Drosophila* (Truett et al. 1981), the TU1 elements of sea urchin (Liebermann et al. 1983), and the Tc4 transposon from *C. elegans* (Yuan et al. 1991), the long TIRs which define the *Sol3* elements are composed of short tandem repeats. The TIRs of Pot*en1*-type elements also contain second repeating inner domain flanking the internal region.

In addition to the Pot*en1*-type *Sol3* elements, similar TIRs define other members of this family with unrelated internal sequences—*Sch2* for example, or severely deleted elements composed solely of the defining repeats (Fig. 2). This suggests that the family of elements defined by *Sol3* TIRs is heterogeneous both in size and sequence. This interpretation is supported by the Southern blot data presented in Figs. 4 and 5, where TIR-hybridizing sequences are shown to be significantly more frequent than Pot*en1*-type internal sequences in the genomes of both *Solanum tuberosum* and *Lycopersicon* species. Conserved TIRs flanking internally heterogeneous sequences are a common feature in eukaryotic DNA-based transposon families; examples include the *Ac/Ds* (Fedoroff 1989) and *Mu* (Talbert et al. 1989) transposons in maize and *Tc2* from *C. elegans* (Oosumi et al. 1996).

Similar to a number of other recently described plant transposable elements (Kikuchi et al. 1991; Meyer et al. 1994; Pozueta-Romero et al. 1995), none of the internal regions of the *Sol3* elements described here contain sequences encoding a tranposase. This suggests that these elements represent nonautonomous forms of an as-yet-unidentified autonomous transposable elements. While we have previously used a PCR-based protocol with TIR-specific primers to isolate autonomous transposons associated with nonautonomous forms in humans (Oosumi et al. 1995a), attempts to similarly identify an autonomous transposable element with *Sol3* TIRs have been unsuccessful to date.

Plant transposon families are often dispersed through multiple genera (Kikuchi et al. 1991; Bureau and Wessler 1994a) and in some cases are ubiquitous throughout the kingdom (Bureau and Wessler 1994b; Pozueta-Romero et al. 1996). In other cases, however, a more specific distribution similar to *Sol3* is observed (Meyer et al. 1994). The absence of *Sol3* hybridizing sequences in tobacco, using probes from either the TIR or internal sequences of Pot*en1*, suggests that the family members characterized to date represent a recently generated series of nonautonomous elements. This interpretation is consistent with evidence indicating that *Solanum* and *Lycopersicon* are congeneric (Spooner et al. 1993).

Mobile DNAs have the potential to contribute regulatory domains and alter gene regulation (Britten 1996). The *Sol3* element in the tomato polygalacturonase promoter (Oosumi et al. 1995b) provides an example of the introduction of both positive and negative regulatory elements (Montgomery et al. 1993) into a promoter by transposon insertion. A more thorough understanding of the role of mobile DNAs in the evolution of the eukaryotes is clearly dependent upon the identification and characterization of the transposon families which make up a considerable percentage of their genomes.

Acknowledgments. The authors wish to express their gratitude to D. Rockhold for expert technical assistance. References to a company and/or product by the USDA is only for purposes of information and does not imply approval or recommendation of the product to the exclusion of others that may also be suitable.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Barker RF, Thompson DV, Talbot DR, Swanson J, Bennetzen JL (1984) Nucleotide sequence of the maize transposable element *Mu1*. Nucleic Acids Res 12:5955–5967

Berg DE, Howe MM (eds) (1989) *Mobile DNA*. American Society for Microbiology, Washington, DC, 972 pp

Britten RJ (1996) DNA sequence insertion and evolutionary variation in gene regulation. Proc Natl Acad Sci USA 93:9374–9377

Bird CR, Smith CJS, Ray JA, Moureau P, Bevan MW, Bird AS, Hughes S, Morris PC, Grierson D, Schuch W (1988) The tomato polygalacturonase gene and ripening-specific expression in transgenic plants. Plant Mol Biol 11:651–662

Bureau TE, Wessler SR (1994a) Mobile inverted-repeat elements of the Tourist family are associated with the genes of many cereal grasses. Proc Natl Acad Sci USA 91:1411–1415

Bureau TE, Wessler SR (1994b) Stowaway: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. Plant Cell 6:907–916

Collins JJ, Anderson P (1994) The Tc5 family of transposable elements in *Caenorhabditis elegans*. Genetics 137:771–781

Deininger PL (1989) SINEs; short interspersed repeated DNA elements in higher eucaryotes. In: Berg DE, Howe MM (eds) Mobile DNA. American Society for Microbiology, Washington, DC, p 619

Draper J, Scott R (1988) The isolation of plant nucleic acids. In: Draper R, Scott R, Armitage P, Walden P (eds) Plant genetic transformation and gene expression: a laboratory manual. Blackstone Scientific, Oxford, p 199

Fedoroff N (1989) Maize transposable elements. In: Berg DE, Howe MM (eds) Mobile DNA. American Society for Microbiology, Washington, DC, p 375

Fedoroff N, Wessler S, Shure M (1983) Isolation of the transposable maize controlling elements *Ac* and *Ds*. Cell 35:235–242

Gierl A, Lutticke S, Saedler H (1988) TnpA product encoded by the transposable element En-1 of *Zea mays* is a DNA binding protein. EMBO J 7:4045–4053

Heffron F, McCarthy BJ, Ohtsobu H, Ohtsubo E (1979) DNA sequence analysis of the transposon Tn3: three genes and three sites involved in transposition of Tn3. Cell 18:1153–1163

Hutchison III CA, Hardies C, Loeb DD, Shehee WR, Edgell M (1989) LINEs and related retroposons; long intersersed repeated sequences

in the eucaryotic genome. In: Berg DE, Howe MM (eds) Mobile DNA. American Society for Microbiology, Washington, DC, p 593

Kikuchi S, Xiangjun L, Frommer WB, Koster-Topfer M, Willmitzer L (1991) Identification and structural characterization of further DNA elements in the potato and pepper genomes homologous to the transposable element-like insertion *Tst1*. Mol Gen Genet 230:494–498

Kunze R, Starlinger P (1989) The putative transposase of transposable element *Ac* from *Zea mays* L interacts with subterminal sequences of *Ac*. EMBO J 11:3177–3185

Lampe DJ, Churchill MEA, Robertson HM (1996) A purified mariner transposase is sufficient to mediate transposition in vitro. EMBO J 15:5470–5479

Liebermann D, Hoffman-Lieberman B, Weinthal J, Childs G, Maxon R, Mauron A, Cohen SN, Kedes L (1983) An unusual transposon with long terminal inverted repeats in the sea urchin *Strongylocentrotus purpuratus*. Nature 306:342–347

Liu Y-G, Mitsukawa N, Oosumi T, Whittier RF (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J 8:457–467

Masson P, Surosky RT, Kingsbury JA, Fedoroff NV (1987) Genetic and molecular analysis of the Spm-dependent a-m2 alleles of the maize a locus. Genetics 177:117–137

Meyer C, Pouteau S, Rouze P, Caboche M (1994) Isolation and molecular characterization of *dTnp1*, a mobile and defective transposable element of *Nicotiana plumbaginifolia*. Mol Gen Genet 242: 194–200

Montgomery J, Pollard V, Deikman J, Fischer RL (1993) Positive and negative regulatory regions control the spatial distribution of polygalactronase transcription in tomato fruit pericap. Plant Cell 5: 1049–1062

New JH, Eggleston AK, Fennewald M (1988) Binding of the *Tn3* transposase to the inverted repeats of *Tn3*. J Mol Biol 201:589–599

O'Hare K, Rubin GM (1983) Structures of P transposable elements and their sites of insertion and exision in the *Drosophila melanogaster* genome. Cell 34:25–35

Oosumi T, Belknap WR, Garlick B (1995a) *Mariner* transposons in humans. Nature 378:672

Oosumi T, Garlick B, Belknap WR (1995b) Identification and characterization of putative transposable DNA elements in solanaceous plants and *Caenorhabditis elegans*. Proc Natl Acad Sci USA 92: 8886–8890

Oosumi T, Garlick B, Belknap WR (1996) Identification of putative nonautonomous transposable elements associated with several transposon families in *Caenorhabditis elegans*. J Mol Evol 43:11–18

Pereira A, Cuypers H, Schwarz-Sommer Z, Saedler H (1986) Molecular analysis of the *En/Spm* transposable element system in *Zea Mayz*. EMBO J 5:835–841

Plasterk RH (1996) The Tc1/mariner transposon family. Curr Top Microbiol Immunol 204:125–143

Pohlman RF, Fedoroff NV, Messing J (1984) The nucleotide sequence of the maize controlling element *Activator*. Cell 37:635–643

Potter S, Truett M, Phillips M, Maher A (1980) Eucaryotic transposable genetic elements with inverted terminal repeats. Cell 20:639–647

Pozueta-Romero J, Klerin M, Houlne G, Schantz M-L, Meyer B, Schantz R (1995) Characterization of a family of genes encoding a fruit-specific wound-stimulated protein of bell pepper *(Capsicum annuum)*: identification of a new family of transposable elements. Plant Mol Biol 28:1011–1025

Pozueta-Romero J, Houlne G, Schantz R (1996) Nonautonomous inverted repeat *Alien* transposable elements are associated with genes of both monocotyledonous and dicotyledonous plants. Gene 171: 147–153

Reue K, Leff T, Breslow JL (1988) Human apolipoprotein CIII gene expression is regulated by positive and negative cis-acting elements and tissue-specific protein factors. J Biol Chem 263:6857–6864

Rosenzweig B, Liao LW, Hirsh D (1983) Sequence of the *C. elegans* transposable element Tc1. Nucleic Acids Res 11:4201–4209

Ruvolo V, Hill JE, Levitt A (1992) The Tc2 transposon of *Caenorhabditis elegans* has the structure of a self-regulated element. DNA Cell Biol 11:111–122

Spooner DM, Anderson GJ, Jansen RK (1993) Chloroplast DNA evidence for the interelationships of tomatoes, potatoes and pepinos (Solanaceae). Am J Bot 80:676–688

Streck RD, MacGaffey JE, Beckendorf SK (1986) The structure of hobo transposable elements and their insertion sites. EMBO J 5: 3615–3623

Talbert LE, Patterson GI, Chandler VL (1989) *Mu* transposable elements are structurally diverse and distributed throughout the genus *Zea*. J Mol Evol 29:28–39

Truett MA, Jones RS, Potter SS (1981) Unusual structure of the FB family of transposable elements in *Drosophila*. Cell 24:753–763

Vos JC, van Luenen HGAM, Plasterk RHA (1993) Characterization of the *Caenorhabditis elegans* Ts1 transposase *in vivo* and *in vitro*. Genes Dev 7:1244–1253

Yuan J, Finney M, Tsung N, Horvitz HR (1991) Tc4, a *Caenorhabditis elegans* transposable element with an unusual fold-back structure. Proc Natl Acad Sci USA 88:3334–3338