Mol Gen Genet (1998) 258: 449–456

ORIGINAL PAPER

W.-Y. Song · L.-Y. Pi · T. E. Bureau · P. C. Ronald

Identification and characterization of 14 transposon-like elements in the noncoding regions of members of the *Xa21* family of disease resistance genes in rice

Received: 17 October 1997 / Accepted: 3 February 1998

Abstract The rice disease resistance gene *Xa21*, which encodes a receptor-like kinase, is a member of a multigene family. Based on comparisons of genomic sequences of seven family members, seventeen transposon-like elements were identified in the 5′ and 3′ flanking regions and introns of these genes. Sequence characterization revealed that these elements are diverse, showing similarity to maize *Ds*, *CACTA* and miniature inverted repeat-like elements, as well as novel elements. Only two elements were located in presumed coding regions, indicating that integration of transposable elements at the *Xa21* disease resistance locus occurred preferentially in noncoding regions.

Key words Rice · *Xa21* · *Xanthomonas oryzae* · Disease resistance · Transposable elements

Introduction

Transposable elements (TEs) are nucleic acid sequences that can move (transpose) from one chromosomal site into a new site (McClintock 1947; Starlinger and Saedler 1972). Based on the mechanism of transposition, TEs can be grouped into two classes: DNA elements which transpose via DNA intermediates (Gierl and Saedler 1992) and retroelements, such as retrotransposons, which move via RNA intermediates (Hull and Will 1989). The first DNA elements to be characterized, the maize *Activator*/*Dissociation* (*Ac*/*Ds*), were described by Barbara McClintock in 1947. Subsequent molecular analysis of these elements and their sites of integration indicated that *Ac* and *Ds* carry terminal inverted repeats (TIRs) at both ends of the elements and generate direct repeats (DRs) of target sites upon insertion (Fedoroff et al. 1983). Transposition of the *Ds* element requires the presence of an active transposase encoded by *Ac*. The *Ac*/*Ds* elements have only been found in maize, although they have been demonstrated to function in diverse plant species as transgenes (Baker et al. 1986; Izawa et al. 1991).

A second group of plant TEs belongs to the *CACTA* class. The elements of this group are large (up to 15.2 kb), encode two proteins thought to be important in transposition, and carry conserved TIRs that have the sequence CACTA... (Nacken et al. 1991). Mobile CACTA elements have been found in soybean, *Antirrhinum*, Japanese morning glory, and maize (Vodkin et al. 1983; Pereira et al. 1985; Nacken et al. 1991; Hoshino et al. 1995).

The largest and most recently characterized class of plant transposons is called the MITE (Miniature Inverted Repeat Transposable Elements) group (Wessler et al. 1995; Bureau et al. 1996). MITEs appear to be ubiquitous in most plant genomes and are characterized by their small size (100 ∼ 300 bp), the presence of TIRs, and the generation of 2- to 3-bp DRs upon insertion. MITEs have been found to insert preferentially in noncoding regions of genes. To date, the identified MITEs can be further classified into seven subclasses: *Tourist*, *Stowaway*, *Gaijin*, *Castaway*, *Ditto*, *Wanderer*, and *Explorer*. Members within each subfamily share structural and sequence similarity, whereas members of different subfamilies only resemble each other structurally. Some MITEs, such as *Tourist* and *Stowaway*, can form secondary structures.

In addition to the DNA elements, a large number of retrotransposons, which transpose via an RNA intermediate, have been identified in plant genomes. Retrotransposons are flanked by two long terminal repeats (LTRs) and encode enzymes required for transposition to occur. The first complete retrotransposon identified in

Communicated by H. Saedler

W.-Y. Song · L.-Y. Pi · P. C. Ronald (✉)
Department of Plant Pathology, University of California, Davis, CA 95616, USA
Fax: +1-916-752-5674

T. E. Bureau
Department of Biology, McGill University, Montreal, Quebec H3A 1B1, Canada

plants is the tobacco element *Tnt1*, whose transposition was detected by the recovery of an insertion mutation in the nitrate reductase gene (Grandbastien et al. 1989). *Tnt1* has LTRs at each end and encodes a single large open reading frame (ORF) with five domains: nucleic acid binding, protease, endonuclease, reverse transcriptase, and RNase H domains. Genomic Southern analysis revealed that there are at least 100 copies of *Tnt1* in the tobacco genome. Other identified plant retrotransposons include: *Bs1*, *Stonor*, *Hopscotch*, and *Magellan* in maize (Johns et al. 1985; Varagona et al. 1992; White et al. 1994; Purugganan and Wessler 1994); *Tto1* and *Tto2* in tobacco (Hirochika 1993); and *Tos10*, *Tos17*, *Tos19*, and *Retrofit* in rice (Hirochika et al. 1996; Song et al. 1997).

The existence, conservation and movement of such a large number of TEs in genomes suggests that these elements have a function. For instance, it has long been hypothesized that TEs play a role in the reconstruction of genomes in response to environmental stresses such as tissue culture, irradiation or pathogen infection (McClintock 1984; Wessler et al. 1995). Pouteau et al. (1994) demonstrated that the transcription of the tobacco element *Tnt1* is induced by a broad spectrum of microbial and fungal elicitors, which is compatible with this hypothesis. In rice, transposition of *Tos10*, *Tos17*, *Tos19* can be activated under tissue culture conditions (Hirochika et al. 1996). Such activation results in an increase in copy number of these elements in the rice genome and causes insertional mutations in diverse rice genes. In addition, evidence has suggested that TE insertion into, and excision from, regulatory and coding regions can change the coding capacity and expression patterns of a gene (McDonald 1995; Wessler et al. 1995; Marillonnet and Wessler 1997). Finally, recent studies in maize have led to the identification, in the spacer region of the *Adh1*-F and *u22* loci, of ten retrotransposon families which account for 50% of the nuclear DNA. These results suggest that TEs play an important role in the enlargement of complex genomes (SanMiguel et al. 1996).

The presence of disease resistance genes provides plants with defense systems with which to recognize and respond to pathogen attack. Given the fact that pathogens can mutate and evolve more quickly than plants due to their shorter life cycle, it has been hypothesized that plants possess genetic mechanisms to generate distinct resistance specificities and thus combat the changing virulence patterns of pathogens (Pryor 1987). One possible mechanism for resistance gene diversification is TE-induced gene alteration (Wise and Ellingboe 1985; Michelmore 1995). Molecular evidence supporting this hypothesis is lacking because few TEs have been identified at resistance loci. The only exception is the maize fungal resistance gene *Hm1*, which confers resistance to *Cochliobolus carbonum* race 1 (Johal and Briggs 1992). In this case, a 315-bp insertion (designated *dHBr*) was found in a mutant allele of this gene. Moreover, recent studies have led to the identification of a transposon, *Drone*, which inserted into the *Hm1* gene in a widely planted maize inbred line, disrupting the resistance function and leading to a rapid outbreak of leaf spot and ear mold disease in 1938 (Multani et al. 1998). Since both TEs at the *Hm1* locus result in loss of function, the question of whether TEs play a functional role in the evolution of resistance genes is still open.

Xa21, a rice gene which confers resistance to *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), was introgressed from the wild rice species, *O. longistaminata*, and was cloned from the introgression line IRBB21 using a positional cloning strategy (Song et al. 1995). *Xa21* encodes a receptor-like kinase protein and belongs to a multigene family (Ronald et al. 1992; Song et al. 1995). Pulsed-field gel electrophoresis and genetic analysis demonstrated that most of members of the *Xa21* gene family are located in a 230-kb genomic region on chromosome 11 (Ronald et al. 1992; Williams et al. 1996). We have cloned and sequenced seven *Xa21* family members designated A1, A2, B (*Xa21*), C, D, E, and F (Song et al. 1995, 1997; Wang et al. 1995). Based on sequence comparisons, two subfamilies were identified: the *Xa21* subfamily consisting of *Xa21*, D and F, and the A2 subfamily consisting of A1, A2, C and E. Within a subfamily, members show 95–98% identity. Two transposon-like elements were identified in the presumed coding regions of the *Xa21* family members D and E, creating ORFs that encode truncated proteins (Song et al. 1997). In addition, *Wanderer*, a MITE element, was found in an *Xa21*-linked noncoding region (Bureau et al. 1996). Based on these results, we undertook a detailed analysis of the *Xa21* family members to identify additional transposon-like elements in the *Xa21* locus and assay their contribution to variability of the family.

In this paper, we describe fourteen previously undescribed transposon-like elements in the noncoding regions of the *Xa21* locus. This result, together with our previous observation that *Retrofit* and *Truncator* are located in coding regions of two *Xa21* family members, indicates that *Xa21* is a transposon-rich locus, and that transposition events contribute to sequence diversity of the *Xa21* gene family members.

Materials and methods

Cloning and sequencing

Cloning and sequencing has been described previously (Song et al. 1995, 1997; Wang et al. 1995). Previously reported Genbank accession numbers for the sequenced regions are as follows: A1, U72725; A2, U72727; Xa21/C, U72723; D, U72726; E, U72724; F, U72728; 3′ flanking region of F, U72729; pTA 818, AF019881.

Sequence analysis

The GCG sequence analysis programs GAP and Pileup were used to calculate percentage identity and to carry out multiple alignments of DNA and protein sequences, respectively. The Genbank searches were performed with the BLAST program of the National Center for Biotechnology Information, Bethesda, Md.

Results

MITEs in the flanking regions of *Xa21* family members

The locations and characteristics of seventeen TE-like sequences located at the *Xa21* locus are summarized in Fig. 1 and Table 1. Five of these elements belong to the previously described MITE families of transposons (Bureau et al. 1996). These include two *Gaijin*-related elements (named *Gaijin-Ol1* and *Gaijin-Ol2*) and two *Tourist*-related elements (named *Tourist-Ol1* and *Tourist-Ol2*) and one *Wanderer* element that was described previously (Bureau et al. 1996). *Gaijin-Ol1* and *Gaijin-Ol2* are positioned in tandem in the 5′ flanking region of member F and share 67.4% identity with each other (Fig. 1, Table 1). Multiple alignments of these two elements with *Gaijin-Os1*, a *Gaijin* element identified in the first intron of the 4-coumarate-CoA ligase gene of *O. sativa* (Bureau et al. 1996), revealed the characteristic terminal motifs of *Gaijin* elements (Fig. 2).

Tourist-Ol1 and *Tourist-Ol2* are located in the 5′ flanking region of member F and the 3′ flanking region of member A2, respectively (Fig. 1, Table 1). *Tourist-Ol1* and *Tourist-Ol2* show a relatively low degree of sequence identity (44.2%) but share the defining characteristics of *Tourist* members including conserved terminal repeats and 3-bp target site duplications (Fig. 3) (Bureau and Wessler 1994a).

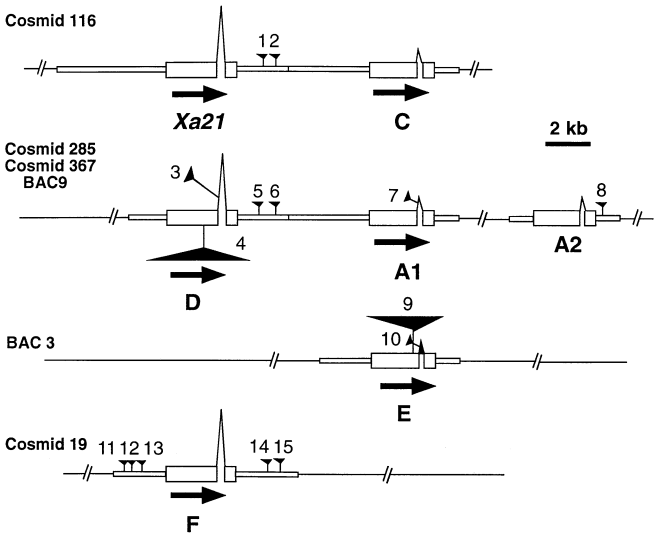

Fig. 1 Genomic organization of 15 transposon-like elements at the *Xa21* locus. Cosmid and BAC clones carrying the family members are designated. *Large open boxes* represent predicted coding regions, *small open boxes* represent non-coding regions, introns are indicated by *angled lines*, and the non-sequenced regions are shown by the *straight lines*. Distance between A1 and A2 is not drawn to scale. *Letters* refer to names of *Xa21* gene family members and *arrows* indicate the orientation of ORFs. The 15 transposon-like elements are numbered and represented by *filled triangles*. The Figure is derived from Fig. 1 of Song et al. (1997)

Snap, *Crackle* and *Pop*: novel miniature TEs in the noncoding regions of *Xa21* family members A1, E, D and F

In addition to the previously described classes of MITEs, we have identified three new classes of miniature elements named *Snap*, *Crackle* and *Pop*, based on comparisons of DNA sequences of closely related *Xa21* family members (Fig. 4). Two nearly identical 183-bp elements were named *Snap-Ol1* and *Snap-Ol2*. *Snap-Ol1* and *Snap-Ol2* insert into the same location in the intron regions of the highly conserved members A1 and E, respectively (Song et al. 1997) (Table 1, Figs. 1, 4). *Snap* elements carry perfect 10-bp TIRs immediately followed by 7-bp DRs. The *Pop* class of miniature elements is

Table 1 Summary of *Xa21*-associated transposable elements

Name	Class	Characteristics	Location (nt)[a]	Reference
Gaijin-Ol1	MITE	TIRs, 3-bp DRs	12 (404–532)	Present work
Gaijin-Ol2	MITE	TIRs, 3-bp DRs	11 (233–384)	Present work
Tourist-Ol1	MITE	TIRs, 3-bp DRs	13 (1077–1239)	Present work
Tourist-Ol2	MITE	TIRs, 3-bp DRs	8 (5453–5697)	Present work
Wanderer-Ol1	MITE		pTA8100 (410–642)[b]	Ronald et al. (1992); Bureau et al. (1996)
Ds-rice1	Ac/Ds	TIR	2 (12313–12521)	Present work
Ds-rice2	Ac/Ds	TIR	6 (13040–13248)	Present work
Ds-rice3	Ac/Ds	TIR	15 (6750–6960)	Present work
Xa21-CACTA	CACTA	TNP2/TnpD	pTA818[b]	Present work; Ronald et al. (1992)
Krispie	Novel	TIRs, 6-bp DRs	4 (10020–10975)	Present work
Snap-Ol1	Novel	TIRs, 7-bp DRs	7 (7432–7614)	Present work
Snap-Ol2	Novel	TIRs, 7-bp DRs	10 (8402–8583)	Present work
Crackle	Novel	SIRs, 9-bp DRs	14 (6202–6585)	Present work
Pop-Ol1	Novel	SIRs, 8-bp DRs	1 (11899–12023)	Present work
Pop-Ol2	Novel	SIRs, 8-bp DRs	5 (12625–12750)	Present work
Retrofit	Retroelement	LTRs, 5-bp DRs	3 (4201–9071)	Song et al. (1996)
Truncator	Novel	TIRs, 5-bp DRs	9 (5211–8128)	Song et al. (1996)

[a] The first number in this column corresponds to the numbered elements in Fig. 1
[b] *Xa21*-linked markers
Abbreviations: TIRs, Terminal inverted repeats; DRs, direct repeats; SIR, subterminal inverted repeat

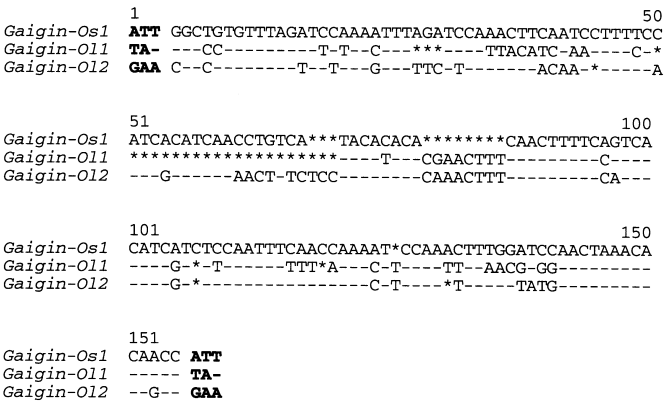

Fig. 2 Multiple alignment of *Gaijin-Ol1*, *Gaijin-Ol2* and *Gaijin-Os1*. Only nucleotides which differ from the *Gaijin-Os1* sequence are indicated. *Asterisks* represent gaps introduced to optimize alignment. The 3-bp target site sequences are indicated in *bold*

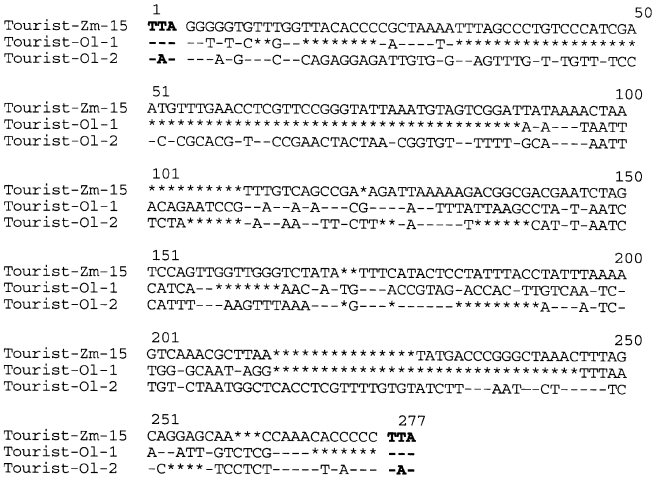

Fig. 3 Multiple alignment of *Tourist-Ol1*, *Tourist-Ol2* and *Tourist-Zm15*. *Asterisk*s represent gaps introduced to optimize alignment. Only nucleotides which differ from the *Tourist-Zm15* sequence are indicated. The 3-bp target site sequences are indicated in *bold*

represented by the two elements *Pop-Ol1* and *Pop-Ol2* present in the 3′ flanking regions of the highly conserved member D and in *Xa21*, respectively. These 125-bp elements generate nearly identical 8-bp DRs. Although no TIR is present following the DR, three nearly perfect subterminal inverted repeats (SIRs) were identified (Fig. 4). Finally, *Crackle* (385 bp), the third new class of miniature elements, inserts into the 3′ flanking region of F. *Crackle* carries four SIRs and generated 9-bp perfect DRs upon integration.

Maize *Ds*-like sequences are located in the 3′ flanking regions of *Xa21* and members F and D

Sequences showing similarity to maize *Ds2* were found in the 3′ regions of *Xa21*, D and F (Fig. 1, Table 1). These sequences, designated *Ds-rice1*, *2*, and *3*, span a 209-bp region and are nearly identical between family members (data not shown). Interestingly, the 209-bp sequence of *Ds-rice1* not only shows 65% identity to the terminus of the maize *Ds2* (Merckelbach 1992), but also carries the 11-bp TIR characteristic of the *Ac*/*Ds* class of maize transposons (Fig. 5) (Fedoroff et al. 1983). Since our data indicate that the genomic regions containing *Xa21*, D and F probably derived from a common progenitor (Song et al. 1997), the three *Ds*-like elements presumably originated through duplication rather than from independent insertion events (Song et al. 1997).

CACTA-like sequence encoded on the *Xa21*-linked marker pTA818

Translation of the 1158-bp sequence of pTA818 revealed an ORF showing similarity to the C-terminal region of the TNP2 gene product of the *Antirrhinum* TE *Tam1* (57.4% identity) and the TnpD product of the maize element *Spm1* (38.3% identity) (Pereira et al. 1985; Nacken et al. 1991; Ronald et al. 1992; Wang et al. 1995; Table 1; Fig. 6). *Tam1* and *Spm1* belong to the CACTA class of TEs. *tnp2* and *tnpD* encode proteins that are thought to interact with the *Spm1* TIRs to facilitate transposition (Masson et al. 1991).

Novel transposon-like element in the noncoding region of *Xa21* family member D

Compared with the MITEs and the maize elements, another previously undescribed 956-bp element (named *Krispie*) was found in the intron region of member D (Fig. 1, Table 1). This element shows no significant similarity to any sequences in the database. *Krispie* displays the characteristics of a transposon-like sequence, including 10-bp TIRs and 6-bp DRs (Fig. 7). No significant ORF was observed in this element.

Discussion

The sequencing of seven *Xa21* family members and their surrounding regions has led to the identification of seventeen transposon-like elements. The elements characterized in this study and those of Song et al. (1997) and Bureau et al. (1996) can be grouped into eleven families including three families of MITEs (*Gaijin*-related elements: *Gaijin-Ol1* and *Gaijin-Ol2*; *Tourist-related* elements: *Tourist-Ol1* and *Tourist-Ol2*, and *Wanderer*), five novel elements (*Truncator*, *Krispie*, *Snap*, *Crackle*, *Pop*), *Ds*-like elements (*Ds-rice1*, *2*, *3*), a *CACTA*-like element (encoded by pTA818) and a retrotransposable element (*Retrofit*) (Song et al. 1997). With the exception of *Retrofit* and possibly the *CACTA*-like element encoded by pTA818, all the identified *Xa21*-associated transposon-like elements appear to be

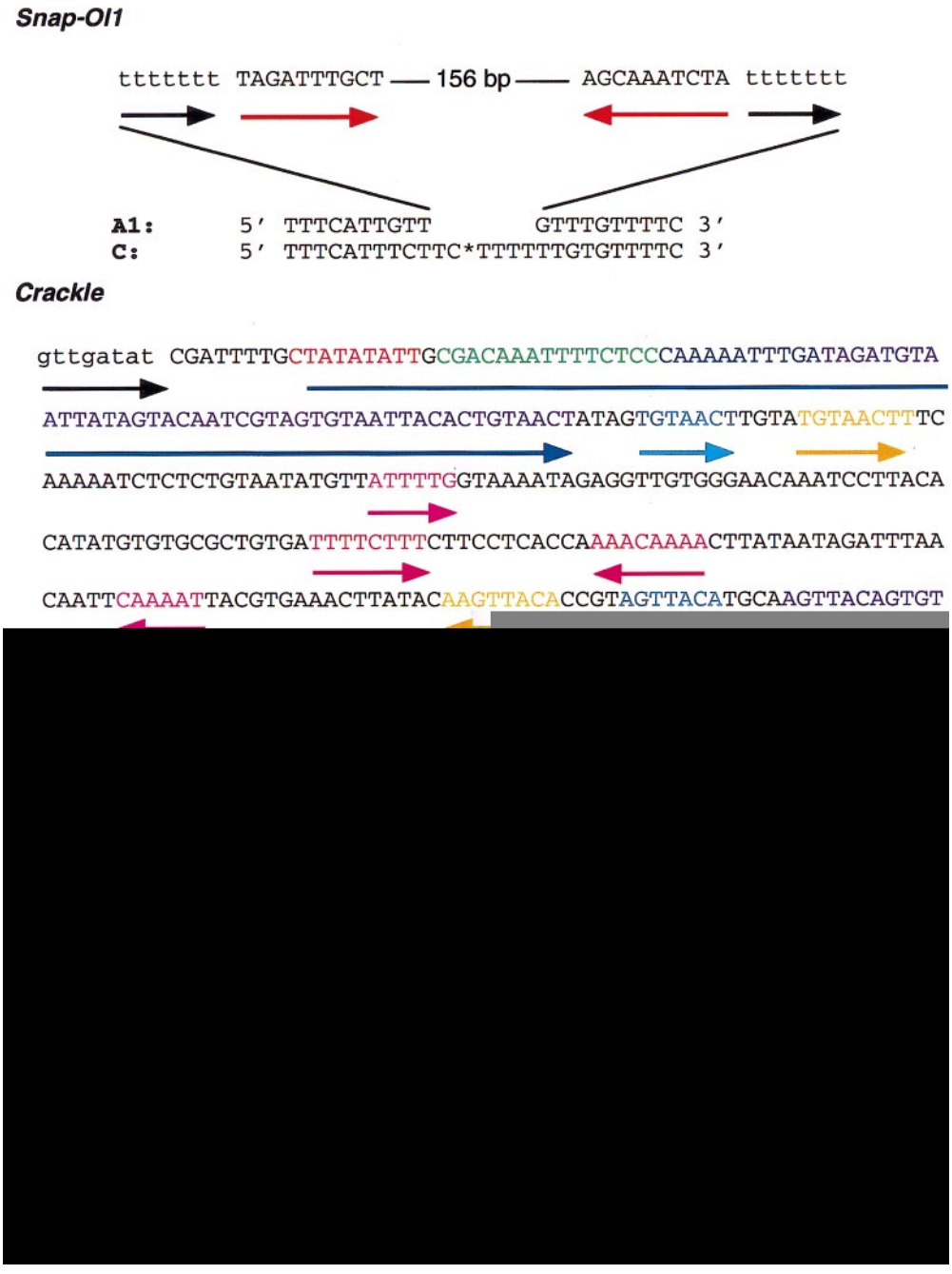

Fig. 4 Structures of the novel miniature elements *Snap*, *Crackle* and *Pop*. Target site sequences are shown in *lower case* and highlighted by *arrows*. Terminal and subterminal inverted repeats are indicated by *colored arrows*. For *Snap*, the number of nucleotides between the TIRs is indicated. Sequences flanking the presumed transposon insertion sites are shown *below* each diagram and are aligned with the corresponding regions of closely related family members lacking the elements. The *asterisk* represents a gap introduced to maximize the alignment

nonmobile or nonautonomous since no obvious ORFs were observed within them.

Like the previously described MITEs, *Snap*, *Crackle* and *Pop* are short (< 390 bp) and the *Pop* and *Crackle* elements have subterminal inverted repeats (SIRs) that could presumably form secondary structures, as has been proposed for *Tourist*, and *Stowaway* (Bureau and Wessler 1992, 1994a, b). However, unlike previously described MITEs, *Crackle* and *Pop* do not have recognizable TIRs. In addition, 7- to 9-bp DRs of target sequences were identified flanking *Snap*, *Crackle* and *Pop*, rather than the 3-bp DRs generated by MITES (Bureau et al. 1996). Based on these differences, these elements appear to represent a new class of miniature transposable elements.

The high level of identity of the TIR of Ds-*rice-1* to the terminal sequence of the maize *Ds2* element (Merckelbach 1992) suggests that intact *Ac* or *Ds*-like elements may have existed in rice at one time. This prediction raises the possibility that other regions of this element might be present in the rice genome; however, no such sequences have yet been identified in the 12 312-bp sequenced region flanking the rice *Ds*-like element (Song et al. 1997). Thus, the rice *Ds*-like sequences may represent disabled *Ds* elements, in which one terminus has been lost due to mutation or recombination.

```
          1                                                   50
Maize Ds2 TTTTATGGTTTGTTTTTACCGAACAAAATACCGGTTCCCGTCCGATTTCG
Ds-rice1  --C---A----AA-G--T-T-T-TTTGT--------TT--AT-TG-AC--

          51                                                  100
Maize Ds2 GCTTTCACCCGACCGGATCGTATCGGTTTTCGATTCCCGTATTTATCCCG
Ds-rice1  A-A-ATTT--AT-A-T--TA-TC-AT--CCG-T--T---ATA--TC-GAT

          101                                                 150
Maize Ds2 TTCGTTTTCGTTACCGATATATCCCGTTTTCGTTTTCGTCCCGCAAGTTA
Ds-rice1  A----------T----*C-T-A---------A----A-TT-*-G--AA-

          151                                                 200
Maize Ds2 AATATGAAAATGAAAACGGTAGAGGTATTTTACCGACCGTTCCCGACCGT
Ds-rice1  -------TT---G---T---C----CTG---T----T----T--------

          201      209
Maize Ds2 TTTCATCCCTA
Ds-rice1  -----------
```

Fig. 5 Alignment of *Ds-rice1* and Maize *Ds2*. Only nucleotides which differ from the *Xa21-Ds* sequence are indicated. *Asterisks* represent gaps for maximum alignment. The 11-bp TIRs are indicated in *bold*

```
          1                                                   50
  TNP2  LLRHNFDVMH IEKNVCESII GTLLNLEGRT KDHENSRLDL KDMGIRSELH
  TNPD  LLPHNIDLMH QERNVAESII SMCFDFTGQT KDNMNARRDL AELCDRPHLE
pTA818  LLRHNLDIMH IEKNVCDNIC GTLLGLEGKS EDNLQARLDL QDMNIRSELH

          51                                                  100
  TNP2  PISLESGKHY LPAACYSMSK KEKEIVFEIL KTVKVPDGYA SNISRRVQLK
  TNPD  LRKNPSGSES RPQAPYCLKR QEREEIFQWL KKLRFPDRYA ANIKRAVNLD
pTA818  PQRRANDKYY LPPASYTLSK KEKQQFCKVL HDIKVPDGYV GNISRCVNVE

          101                                                 150
  TNP2  PNKISGLKSH DHHILMQQLL PIALRKVLPK HVRTPLIKLC TFFRELCSKV
  TNPD  TGKLVGLKSH DYHILIERLV PVMFRGYFSP DVWKIFAELS YFYKQICAKE
pTA818  QGKISGLKSH DCHILMQELL PLALRGVLPD NVTAVLFDLC GHFRELNAKV

          151                                                 200
  TNP2  LNPQDLVRMG KDIAKTLCDL EKIFPPSFFD IMMHLPIHLA YEAQIAGPVQ
  TNPD  ISKKLMLRFE KEIVVLVCKM EKVFPPGFFN CMQHLLVHLP WEALVGGPAQ
pTA818  LYIDELKKLD ERIKLTLCRM EMIFPPGFFT IMVHLVSHLA TEALLGGPAC

          250
  TNP2  YRWMYPIER
  TNPD  FRWMYSQERE LKKLRGMVRN KARVEGCIAE AFAAREITLF SSK ...
pTA818  FCTMYFVERY FSLSNIRMYS LFYSNIFNTI FSLCCPDFL
```

Fig. 6 Multiple alignment of the homologous amino acid sequences predicted from pTA818, TNP2, and TNPD. Conserved residues are indicated in *bold*

In the *Xa21*-linked marker pTA818, we identified sequences showing similarity to those encoding the C-terminal region of TNP2, a protein that is specified by the *Antirrhinum* CACTA element *Tam1* (Nacken et al. 1991). We do not know if the lines used in these studies contain additional sequence homology to the *CACTA* elements. It is not known whether the rice *CACTA*-like element is autonomous, since the whole element has not been sequenced or identified. A small deletion in a complete element is sufficient to form a non-autonomous element (Gierl and Saedler 1989). Interestingly, TIR-like sequences belonging to the *CACTA* class of elements have been found in rice (Motohashi et al. 1996), suggesting that the entire element may be present in some varieties and be capable of transposition in rice. This hypothesis is supported by the observation that a single copy of the DNA sequence carrying the *CACTA*-like element encoded on pTA818 is found on chromosome 2 of *O. sativa*, whereas in the line IRBB21, which contains an *Xa21* locus that originated from the wild species *O. longistaminata*, a second unlinked copy of the *CACTA*-like element is found at the *Xa21* locus, suggesting that a duplication and/or transposition event occurred in the wild species (Jiang et al. 1995; J. Xiao, personal communication).

In addition to demonstrating the existence and revealing the distribution of the *Xa21*-associated TEs, the sequence analysis also provides information concerning the timing of transposon integration. Compared with other, more ancient, evolutionary events such as duplication and recombination (Song et al. 997), many of these elements seem to have been active quite recently. For example, the sequence of a 14 742-bp region spanning the *Xa21*/C cluster shows 97.7% identity to the corresponding sequence (14 871 bp) of the D/A1 cluster, suggesting these regions evolved through sequence duplication (Song et al. 1997). The major difference between these two regions is the presence of the transposon-like elements *Retrofit*, *Krispie* and *Snap-Ol1*, which are located in the intronic and coding regions of members D and A1. These observations suggest that integration of *Retrofit*, *Krispie* and *Snap-Ol1* occurred after duplication. Similarly, insertions of *Crackle*, *Pop-Ol1*, *Tourist-Ol2*, *Truncator*, *Snap-Ol1* and *Snap-Ol2* appear to have occurred after duplication events, because the presence of these elements accounts for the major differences observed among closely related family members (*Xa21* and member F; members A1, A2, C and E) (Song et al. 1997). In contrast, movement of the three *Ds*-like elements may have occurred before duplication of the *Xa21* subfamily (including *Xa21*, and members D and F), because these elements have almost identical DNA sequences and are present at the same positions of the 3′ flanking regions of all the *Xa21* subfamily members, which are believed to have arisen by duplication from a common progenitor (Song et al. 1997). Thus, our sequence data suggest that TEs have been active over the entire evolutionary history of the *Xa21* gene family.

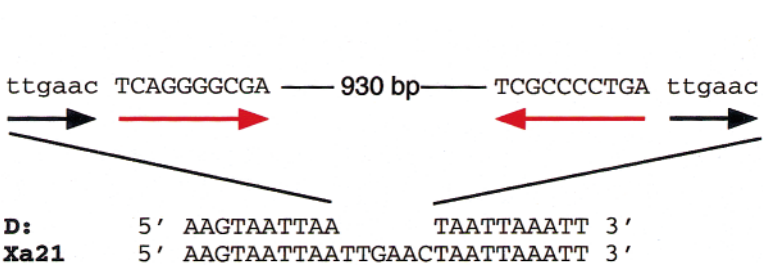

Fig. 7 Structure of the novel transposon-like element *Krispie*. Sequences of this element are indicated in the same format as those of other elements in Fig. 4

Most of the transposon-like elements identified at the *Xa21* locus are located in noncoding regions and appear to be functionally neutral. Since the *Xa21* locus has not been completely sequenced, there may be additional TEs in the noncoding regions. In rice, most MITEs characterized to date insert preferentially into noncoding regions (Bureau et al. 1996; Chen and Bennetzen 1996). A similar observation has been made in yeast and maize (Voytas 1996; SanMiguel et al. 1996). These results support the hypothesis that genomes have developed mechanisms to modulate the integration of TEs to minimize deleterious mutations caused by transposition (Voytas 1996).

The *Xa21*-associated transposable-like elements appear to be a major source of variability among the *Xa21*-gene family members and may affect resistance gene function. For instance, we have demonstrated that integration of the DNA element, *Retrofit*, into *Xa21* family member D leads to the generation of a novel molecule conferring an altered resistance phenotype (Song et al., 1997; Wang et al., 1998). Similarly, the integration of the retrotransposable element, Truncator, into family member E generated a receptor-like ORF structurally similar to the tomato fungal resistance gene *Cf-9* (Song et al., 1997). Furthermore, three MITEs, *Tourist-Ol1 Gaigin-Ol1* and *Gaigin-Ol2*, were found in the 5′ flanking region of the resistance gene family member F suggesting that these elements could alter the expression of member F as has been proposed for other MITES (Wessler et al., 1995). Movement of these elements in response to pathogen induced stress would provide genetic plasticity for evolution of the disease resistance locus. However, direct evidence demonstrating a clear selective advantage resulting from the movement of the TEs remains to be shown.

Acknowledgements We thank Dean Lavelle for technical assistance. We are grateful to R. Adamchak of suggesting the names *Snap*, *Crackle* and *Pop*. This work was supported by NIH and the Rockefeller Foundation.

References

Baker B, Schell J, Lorz H, Fedoroff N (1986) Transposition of the maize controlling element "*Activator*" in tobacco. Proc Natl Acad Sci USA 83:4844–4848

Bureau TE, Wessler SR (1992) *Tourist*: A large family of small inverted repeat elements frequently associated with maize genes. Plant Cell 4:1283–1294

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

Bureau TE, Ronald PC, Wessler SR (1996) A computer-based systematic survey reveals the predominance of small inverted-repeat elements in wild-type rice genes. Proc Natl Acad Sci USA 93:8524–8529

Chen MS, Bennetzen JL (1996) Sequence composition and organization in the Sh2/A1-homologous region of rice. Plant Mol Biol 32:999–1001

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

Gierl A, Saedler H (1989) Maize transposable elements. Annu Rev Genet 23:71–85

Gierl A, Saedler H (1992) Plant transposable elements and gene tagging. Plant Mol Biol 19:39–49

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

Hirochika H (1993) Activation of tobacco retrotransposons during tissue culture. EMBO J 12:2521–2528

Hirochika H, Sugimoto K, Otsuki Y, Tsugawa H, Kanda M (1996) Retrotransposons of rice involved in mutations induced by tissue culture. Proc Natl Acad Sci USA 93:7783–7788

Hoshino A, Inagaki Y, Iida S (1995) Structural analysis of *Tpn1*, a transposable element isolated from apanese morning glory bearing variegated flowers. Mol Gen Genet 247:114–117

Hull R, Will H (1989) Molecular biology of viral and nonviral retroelements. Trends Genet 5:357–360

Izawa T, Miyazaki C, Yamamoto M, Terada R, Iida S, Shimamoto K (1991) Introduction and transposition of the maize transposable element *Ac* in rice (*Oryza sativa* L.). Mol Gen Genet 227:391–396

Jiang J, Gill BS, Wang GL, Ronald PC, Ward DC (1995) Metaphase and interphase fluorescence in situ hybridization mapping of the rice genome with bacterial artificial chromosomes. Proc Natl Acad Sci USA 92:4487–4491

Johal GS, Briggs SP (1992) Reductase activity encoded by the *Hm1* disease resistance gene in maize. Science 258:985–987

Johns MA, Mottinger J, Freeling M (1985) A low copy number, *copia*-like transposon in maize. EMBO J 4:1093–1102

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

Masson P, Banks JA, Fedoroff N (1991) Structure and function of the maize *Spm* transposable element. Biochimie 73:5–8

McClintock B (1947) Cytogenetic studies of maize and *Neurospora*. Carnegie Inst (Wash) Yrbk 46:146–152

McClintock B (1984) The significance of responses of the genome to challenge. Science 226:792–801

McDonald JF (1995) Transposable elements – possible catalysts of organismic evolution. Trends Ecol Evol 10:123–126

Merckelbach A (1992) Genbank Accession Number X65746

Michelmore R (1995) Molecular approaches to manipulation of disease resistance genes. Annu Rev Phytopathol 33:393–427

Motohashi R, Ohtsubo E, Ohtsubo H (1996) Identification of *Tnr3*, a suppressor-mutator/enhancer-like transposable element from rice. Mol Gen Genet 250:148–152

Multani AS, Meeley RB, Paterson AM, Grey S, Briggs SP, Johal GS (1998) Plant pathogen microevolution, molecular basis for the origin of a fungal disease in maize. PNAS 95:1686–1691

Nacken WKF, Piotrowiak R, Saedler H, Sommer H (1991) The transposable element *Tam1* from *Antirrhinum majus* shows structural homology to the maize transposon En/Spm and has no sequence specificity of insertion. Mol Gen Genet. 228:201–208

Pereira A, Cuypers H, Gierl A, Schwarz-Sommer Zs, Saedler H (1985) Molecular analysis of the *En*/*Spm* transposable element system of *Zea mays*. EMBO J 5:835–841

Pouteau S, Grandbastien MA, Boccara M (1994) Microbial elicitors of plant defence responses activate transcription of a retrotransposon. Plant J 5:535–542

Pryor T (1987) The origin and structure of fungal disease resistance genes in plants. Trends Genet 3:157–161

Purugganan MD, Wessler SR (1994) Molecular evolution of magellan, a maize Ty3/gypsy-like retrotransposon. Proc Natl Acad Sci USA 91:11674–11678

Ronald PC, Albano B, Tabien R, Abenes L, Wu KS, McCouch S, Tanksley SD (1992) Genetic and physical analysis of the rice bacterial blight disease resistance locus, *Xa21*. Mol Gen Genet 236:113–120

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, Bennetzen JL (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274:765–768

Song WY, Wang GL, Chen LL, Kim HS, Pi LY, Gardner J, Wang B, Holsten TE, Zhai WX, Zhu LH, Fauquet C, Ronald PC (1995) A receptor kinase-like protein encoded by the rice disease resistance gene *Xa21*. Science 270:1804–1806

Song WY, Pi LY, Wang GL, Gardner J, Holsten TE, Ronald PC (1997) Evolution of the rice *Xa21* disease resistance gene family. Plant Cell 9:1279–1287

Starlinger P, Saedler H (1972) Insertion mutations in microorganisms. Biochimie 54:177–185

Varagona MJ, Purugganan M, Wessler SR (1992) Alternative splicing induced by insertion of retrotransposons into the maize *waxy* gene. Plant Cell 4:811–820

Vodkin LO, Rhodes PR, Goldberg RB (1983) A lectin gene insertion has the structural features of a transposable element. Cell 34:1023–1031

Voytas DF (1996) Retroelements in genome organization. Science 274:737–738

Wang GL, Holsten TE, Song WY, Wang HP, Ronald PC (1995) Construction of a rice bacterial artificial chromosome library and identification of clones linked to the *Xa21* disease resistance locus. Plant J 7:525–533

Wang GL, Ruan DL, Song WY, Sideris S, Chen L-L, Pi L-Y, Zhang S, Zhang Z, Fauquet C, Gaut B, Ronald P (1998) Xa21 D encodes a recepter like molecule with a leucine rich repeat domain that determines race specific recognition and is subject to adaptive evolution, Plant Cell 10:1–15

Wessler SR, Bureau TE, White SE (1995) LTR-retrotransposons and MITEs – important players in the evolution of plant genomes. Curr Opin Genet Dev 5:814–821

White SE, Habera LF, 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

Williams CE, Wang B, Holsten TE, Scambray J, da Silva F, Ronald PC (1996) Markers for selection of the rice *Xa21* disease resistance gene. Theor Appl Genet 93:1119–1122

Wise RP, Ellingboe AH (1985) Fine structure and instability of the *Ml-a* locus in barley. Genetics 111:113–130