TECHNICAL ADVANCE

Display and isolation of transposon-flanking sequences starting from genomic DNA or RNA

Alexander Yephremov* and Heinz Saedler
Max-Planck-Institut für Züchtungsforschung, Carl von Linné Weg 10, 50829 Cologne, Germany

Received 8 September 1999; revised 26 January 2000; accepted 1 February 2000.
*For correspondence (fax +49 221 5062 113; e-mail efremov@mpiz-koeln.mpg.de)

Summary

Insertion mutagenesis using transposons has become a powerful tool for the isolation of genes involved in any given biochemical or developmental pathway. We describe here ligation-mediated PCR techniques for the isolation of sequences flanking the transposable elements En/Spm, Mu1 and Cin4 in Zea mays and Arabidopsis thaliana. Two versions of this transposon insertion display (TID) method use biotinylated linkers or biotinylated primers to rapidly isolate transposon-flanking sequences starting from digested genomic DNA. TID protocols have been employed to clone several genes from En/Spm insertion mutants of Arabidopsis. A novel procedure, expression TID (ETID), is also introduced, which provides a direct approach for the isolation of transposon insertions that tag transcribed portions of genes. ETID uses RNA as a starting material and exploits 5’ RACE PCR to identify transposon copies that form parts of gene transcripts. The detection of several En/Spm insertion mutations in Arabidopsis illustrates the power of this method. ETID offers important advantages for the isolation of mutant alleles of novel genes that are expressed in specific tissues in plants and animals.

Introduction

Identification of a transposon insertion in a gene of interest makes it possible to clone it directly by using the known DNA sequence of the transposon as a molecular tag. The cloning strategy relies on the demonstration of coincident segregation of the tag with a mutant phenotype in a progeny population. In practice, two types of techniques have been used to follow the segregation of transposon insertions. Although these approaches are both based on the RFLPs that result from transposon insertions, and on segregation analysis of DNA fragments in a gel, they use different techniques to reveal tag-specific DNA fragments. The first procedure utilizes Southern hybridization with a transposon-derived molecular probe (Walbot, 1992), while the second approach exploits PCR to amplify DNA sequences flanking transposons (Earp et al., 1990; Frey et al., 1998; Van et al., 1998; Yephremov and Saedler, 1995). As with other applications of PCR aimed at the amplification of DNA fragments outside the boundaries of known sequences, in the latter procedure the PCR template must first be constructed. In principle, this can be achieved either before PCR – by circularization of DNA fragments, tailing reactions or ligation of DNA fragments with defined sequences (Ochman et al., 1993) – or directly during PCR by using a degenerate oligonucleotide as a primer for flanking sequences (Liu et al., 1995).

Ligation of linkers to digested genomic DNA is used in one PCR-based technique, transposon insertion display (TID), to synthesize the PCR templates (Yephremov and Saedler, 1995). TID protocols have been developed for three classes of maize transposable elements – Cin4 (Schwarz-Sommer et al., 1987), Mu1 (Bennetzen et al., 1984) and En/Spm (McClintock, 1954; Pereira et al., 1986; Peterson, 1953) – and used to examine the segregation of transposons and isolate their flanking sequences in maize, teosinte and Arabidopsis. Using the TID protocol, several transposon-tagged genes have been cloned from Arabidopsis populations mutagenized with En/Spm transposons.

This paper describes details of the TID techniques and introduces expression transposon insertion display (ETID), a novel approach for cloning genes that are tagged by transposons. Unlike other methods for the isolation of

© 2000 Blackwell Science Ltd

495
transposon-flanking sequences, the ETID technique is not based on RFLPs in genomic DNA but on the fact that part of a transposon inserted in a gene may be co-transcribed with that gene, from its normal promoter, to form chimeric RNA. Chimeric transcripts composed of gene and transposon segments are isolated by 5' RACE PCR (Frohman et al., 1988) using essentially the same transposon primers as for TID. Segregation of fragments is analysed by gel electrophoresis or by using them to probe Southern blots for RFLPs.

Results and Discussion

Rationale for amplification of flanking sequences from genomic templates

The TID approach uses a type of ligation-mediated PCR (LM-PCR) (Mueller and Wold, 1989) on genomic DNA that relies on RFLPs generated by digestion of DNA with a frequently cutting enzyme. In order to amplify sequences flanking transposon insertions, all TID protocols use PCR with nested primers that anneal to the 5' or 3' end of a transposon.

However, our experience with TID has shown that the use of nested PCR alone is not sufficient to ensure that essentially all resulting fragments are derived from transposon-flanking sequences. In order to improve the selectivity of amplification reactions, we took advantage of the high affinity of biotinylated DNA for streptavidin-coated magnetic beads. Two strategies were applied in TID, one using biotinylated linkers (Figure 1a,b) and the other using biotinylated primers in PCR (Figure 1c). Both versions of TID were successful, underlining the importance of using a biotin-capture step in the protocols.

In TID version I, the use of biotinylated linkers allows one to discard most genomic DNA fragments before amplification starts. Higher selectivity for transposon fragments is achieved by choosing a hapaxotermistic restriction enzyme that cuts close to the 3' or the 5' end of the

![Diagram](image)

**Figure 1.** Two strategies for amplification of transposon flanking sequences in TID.

(a) Amplification of flanking sequences at the 3' end of Cin4 retrotransposon copies using TID version I. Digestion with the hapaxotermistic restriction enzyme BsaI produces Cin4 fragments that can be selectively ligated to a biotinylated linker.

(b) Use of the hapaxotermistic restriction enzyme BstXI allows selection for Mu1 transposon fragments, as in (a), in Mu1 TID version I. BstXI produces 3' protruding cohesive ends upon digestion; therefore, in contrast to (a), synthesis of linker templates is necessary.

(c) Amplification of sequences flanking EnSpm transposons of maize using TID version II. Sequences around hapaxotermistic enzyme recognition sites in transposable elements are circled; light grey bars are transposon sequences; white bars are flanking genomic DNA sequences; dark grey bars are linkers and primers; *, labelling with 32P; O, labelling with biotin.
transposon. This type of restriction enzyme (Berger, 1994) does not cut its recognition sequence but rather an adjacent sequence, and is thus able to produce unique cohesive ends. Since digestion with a hapaxterministic enzyme can produce up to 256 types of four-base cohesive ends, it provides up to 256-fold better selectivity in LM-PCR compared to a type II restriction enzyme with an equally long recognition sequence.

A biotinylated oligonucleotide is used as a primer to start synthesis from a transposon end into flanking genome sequences in TID version II. This strategy allows greater flexibility in primer design and enzyme selection, since only one restriction enzyme, a frequent cutter, is required. Biotinylated DNA is synthesized using PCR for only 12–15 cycles. This often helps to reduce the background. After purification using streptavidin-coated magnetic particles, the products are amplified by PCR and labelled with a nested primer. Version II TID is similar to the recently published AIMS protocol for Mu transposable elements in maize (Frey et al., 1998).

In both versions of TID, segregation analysis of labelled fragments is done under denaturing conditions in a polyacrylamide gel. After identification of fragments of interest, a preparative gel separation is performed; the fragments are then isolated and amplified for sequencing.

**TID procedures for Cin4, Mu1 and En/Spm transposable elements**

Sequences flanking transposon insertions produced by three classes of maize transposable elements were analysed by TID techniques. DNA sequencing showed that the major fragments generated by TID include portions of transposons.

Both versions of TID allowed display of flanking sequences at the 3' end of the non-LTR retrotransposon Cin4 (Schwarz-Sommer et al., 1987), which can accumulate in tens of copies in maize and teosinte genomes. Capture with biotinylated selective linkers in TID version I dramatically improves the specificity of the subsequent PCR, as illustrated by the use of different nested primers. This is seen as a shift in the fragment pattern, as shown by the amplification of Cin4-flanking sequences in Mexican teosinte races (Figure 2a).

Digestion with BstXI within Mu1 transposons in TID version I allows selection for members of this subfamily and visualization of their flanking sequences in the gel. These amplification reactions were remarkably consistent over a 7°C range in annealing temperatures, as demonstrated in Figure 2(b).

TID version II, based on the use of biotinylated primers for amplification, allowed display of sequences of En/Spm transposable elements from maize in Arabidopsis (Figure 2c). This protocol is especially practical considering the widespread usage of En/Spm mutagenesis in Arabidopsis (Aarts et al., 1995; Pereira and Aarts, 1998; Tissier et al., 1999; Wisman et al., 1998a). The Arabidopsis genes isolated using TID include FIDDLEHEAD (Yephremov et al., 1999), PIN1 (Gaelweiler et al., 1998), SUCROSE UNCOUPLED6 and MYB2 (C. Huijser, personal communication). TID of En/Spm was useful in the isolation of flanking sequences for mapping of insertion sites (Wisman et al., 1998b; E. Wisman, personal communication) and for the estimation of numbers of insertions (A. Pereira, personal communication). In all these instances, a set of nested primers at the 5' end of En/Spm was used, since TID with primers at the 3' end appeared to be less reliable.

**Isolation of a DWF4 insertion allele from high-copy transposon material by 3' end En/Spm TID**

In order to improve 3' end En/Spm TID, we used the recessive dwarf mutant 2AA145, which originated from an Arabidopsis population (ecotype Columbia) mutagenized with En/Spm (Wisman et al., 1998a). Mutant plants were essentially sterile and had dark green, rounded leaves, resembling the phenotype of steroid-deficient dwarfs (Chory et al., 1991; Szekeres et al., 1996). Dwarf plants responded to the application of 0.1% brassinolide by elongation of organs and change in colour (data not shown). Putative revertant R0 shoots (with wild-type morphology) that set seeds were observed on some plants of mutant 2AA145. R1 offspring of three independent reverting shoots segregated wild-types and dwarfs in a 3:1 ratio (data not shown), suggesting that the revertant phenotype is due to transposon excision from the dwarf locus. None of the 96 offspring of one reverting shoot, however, had a wild-type phenotype. Chimerism of this shoot is the only explanation for the exclusive appearance of dwarfs in this R1.

Wild-type R1 plants of three reverting shoots were fertilized with pollen from the Columbia parent, giving rise to BC1R1. The BC1R1 plants were again back-crossed to Columbia. BC1R1S (selfed) progeny of BC1R1 were tested for segregation of the mutants, and leaves of BC1R1 plants of positive families were used for molecular analysis. Half of these BC1R1 plants are expected to be heterozygous at the 2AA145 locus and to contain the En/Spm transposon insertion, while the other half are expected to be homozygous wild-type.

Southern analysis of BC1R1 has shown that on average 16 En/Spm copies segregate in these plants (data not shown). DNA of three heterozygotes and three BC1R1 homozygotes at the 2AA145 dwarf locus were analysed by TID version II using the new primer combination at the 3' end of En/Spm. The amplification resulted in an average of 16 prominent fragments on the autoradiogram (Figure 2c).
Three fragments – F212, F257 and F313 – which distinguish heterozygotes from homozygotes at the 2AA45 locus were cut out and re-amplified. Sequencing showed that all the fragments had the expected portion of the En/Spm transposon. Sequences flanking two of the fragments did not show similarity to any sequences in the database; however, the F212 fragment appeared to contain a portion of the DWF4 gene of Arabidopsis (GenBank accession number AF044216). The product of DWF4 is a cytochrome P450 mono-oxygenase involved in hydroxylation during brassinsteroid biosynthesis, mutation of which is known to result in a dwarf phenotype (Choe et al., 1998). Further analysis of this insertion allele, designated dwf4-En1240, using PCR and sequencing revealed that the En/Spm transposon is located in exon 1 at nt 1240 of the DWF4 gene. PCR analysis of two other insertions in the BC2 R1 plants did not reveal co-segregation with the dwarf phenotype.

The principle of ETID

Transposon insertions located in the transcribed portion of a gene should generally abolish gene function and therefore are of great interest. In this kind of mutation, the gene promoter is unaffected by the insertion and hence should still be transcriptionally active. Expression of the mutant allele should result in a chimeric RNA, because transcription proceeds from the 5′ end of the gene into the transposon itself.

The ETID technique is based on the specific amplification of 5′ end portions of these chimeric RNAs, followed by gel analysis (Figure 3). Although a number of protocols have been described for the amplification of sequences at the 5′ ends of transcripts (Schafer, 1995), the SMART technique (Chenchik et al., 1998), which utilizes oligo-mediated template switch at the 5′ end of the RNA during the reverse transcription (RT) reaction, is the method of choice for our purposes.

Provided that cDNA synthesis starts with a transposon-specific primer and continues into the transcribed portion of the tagged gene and into the SMART oligonucleotide template, the resulting fragments will acquire defined sequences at their ends and can be readily amplified by PCR. ETID may employ essentially the same nested transposon primers as TID, which are chosen to lie close to the ends of transposons. The first transposon-specific primer, which is used to start cDNA synthesis in ETID, is biotinylated. Also, as in TID, selection for biotinylated nucleic acids is carried out as early as possible in the protocol. To accomplish this, a biotinylated transposon-specific oligonucleotide is bound to streptavidin-coated particles and hybridized to total RNA to permit direct capture of hybrids. DNA–RNA hybrids are washed to remove unhybridized RNA and inhibitors (van Doorn et al., 1992). Subsequently, a reverse transcription reaction is performed on the matrix-bound hybrids, followed by nested PCR and gel electrophoresis.

Identification of En/Spm transposon insertions in transcribed portions of genes by ETID

In order to investigate how ETID could be applied to the isolation of genes tagged by transposons, we analysed BC2 S (selfed) progeny of BC2 plants that were used in the cloning of DWF4. Vegetative and floral tissues of about 20 apparently wild-type plants from each BC2 S family were combined, at the onset of bolting, for total RNA isolation. After ETID, some PCR products were visible as prominent

Figure 2. Amplification of transposon flanking sequences in teosinte, maize and Arabidopsis by TID. (a) Stability of TID fingerprint patterns generated by two different Cin4 retrotransposon primers in Mexican races of teosinte (Zea mays ssp. parviglumis). Genomic DNA was digested with MseI and BstU1. Enrichment for transposon-specific fragments was performed by linker ligation to GCCT-cohesive ends. PCR was performed with primers 7116 and LR26 on DNA fragments bound to streptavidin particles. The next amplification was done with the labelled primers 7116 or 7127 (as indicated at the bottom), and the products were separately loaded on a gel. Note that the 3′ end Cin4 7127 primer is located downstream of the 7116 primer and produces fragments that are 11 bp shorter. L, labelled Sequmark ladder (Research Genetics) loaded together with 7116 TID products displays bands at 10 bp intervals up to 500 bp. BA, teosinte race (accession 11407) collected in Michoacan; GU, teosinte race (accession 11402) collected in Guerrero; MC, teosinte race (accession 11403) collected in Michoacan. Accession numbers were assigned by the Maize Germplasm Bank of the International Maize and Wheat Improvement Center (CIMMYT). (b) Stability of amplification of Mu1 transposon insertions in offspring of branching mutants of maize (Zea mays ssp. mays). Genomic DNA was digested with MseI and BstU1. Enrichment for transposon-specific fragments was performed by linker ligation to CGCG-cohesive ends. PCR on DNA fragments bound to streptavidin particles was accomplished with Mu26 transposon and LR26 linker primers at two annealing temperatures as indicated. TID patterns of three plants representing the parental bx2-mu2 ‘big spot’ line – a source of active Mu1 of transposons – are shown next to that of the out-cross offspring (CR1 and CR2) of two independent insertion mutants. This progeny is expected to harbour a transposon insertion allele of a branching locus. DNAs of 16 out-cross offspring were mixed in equal amounts for TID. ‘Big spot’ line samples were amplified using annealing at 63°C, and the out-cross offspring as indicated. The fragments depicted in squares were isolated and proven to contain Mu1 flanking sequences. L, end-labelled ladder (PCR molecular weight marker, USB). (c) Analysis of co-segregation of En/Spm transposon-flanking sequences with the dwarf phenotype in the back-cross progeny of revertants in Arabidopsis thaliana. The homozygous wild-type (DWF/dWF) or heterozygous (DWF/DWF) allelic state of the dwarf locus is indicated. Lanes 1, 4, 5, and 6 represent progeny of revertant 106, while lanes 2 and 3 represent progeny of revertant 107. The positions of three bands that distinguish heterozygotes from wild-type homozygotes are indicated by arrows. All fragments contain a portion of En/Spm, with the F212 fragment corresponding to the dwf4-1240 insertion allele. The sequencing gel allowed resolution of individual bands (asterisks) that differ from F212 and F257 only by two nucleotides. L, the labelled SequaMark ladder (Research Genetics) displays bands at 10 bp intervals up to 500 bp.
bands on an agarose gel (Figure 4a,b). These were cut out, and sequencing of three fragments has shown that all include transposon-flanking sequences. One of them, the C5 fragment (360 bp), represented a portion of a novel receptor kinase gene with highest sequence similarity to GenBank accession AC005957, which encodes a putative disease resistance protein of Arabidopsis. This result was confirmed by sequencing of a 1.2 kb DNA fragment flanking En/Spm at the opposite side of this insertion (data not shown).

The second fragment, C78, of 278 bp could be aligned with a sequence from chromosome 4 in the BAC clone F23E12 (GenBank accession AL022604) at nucleotide positions 72480–72751. Interestingly, according to the database, computer analysis did not predict that this region is part of a transcript. The ETID results strongly suggest that RNA transcription starts on the opposite strand at the position corresponding to nucleotide 72751 of this clone.

A very short flanking sequence of 31 bp in the third fragment, designated C6, did not allow reliable homology searches in databases. A PCR analysis demonstrated that the corresponding En/Spm insertion is present in the BC2 parental plant. In order to isolate a longer portion of the C6 fragment, BC2 genomic DNA was subjected to inverse PCR (Earp et al., 1990) with nested primers, taking advantage of
the known sequence of the C6 fragment (see Experimental procedures). The addition of only three C6-specific nucleotides at the 3' end of the EnC6-3 primer resulted in preferential amplification of the 2.4 kb fragment as compared to the combination of nested transposon primers (Figure 5). The fragment was isolated and sequenced, and this allowed us to identify the corresponding gene as a novel homologue of late embryogenesis abundant (LEA) proteins, particularly of Arabidopsis ECP63 (Yang et al., 1997).

Critical points of TID and ETID

TID is reminiscent of the AFLP fingerprinting technique (Vos et al., 1995), in which distinct adaptor sequences attached to the ends of restriction fragments permit the use of selective primers for amplification. Combinations of primers, which are complementary to adaptors but extend beyond these, generate multiple gel patterns for each DNA digestion reaction in AFLP. AFLP allows one to amplify restriction fragments reproducibly without the use of biotinylated oligonucleotides. In contrast to this, in TID the use of biotinylated oligonucleotides was found to be necessary to ensure specific amplification of transposon-flanking sequences. This difference may be attributed to the use of only one segment of the fragments – that which corresponds to the transposon portion – as a template for extended or nested PCR primers. TID is designed to result in a single gel pattern for each DNA digestion reaction and to display only transposon-specific fragments.

We have found that the quality of biotinylated oligonucleotides is the factor that has the greatest influence on the outcome of TID. This can be inferred from the observation that different batches of biotinylated oligonucleotides obtained from various suppliers perform differently in TID.
Like any analysis based on RFLP, both versions of TID depend on the availability of restriction sites at the ends of transposable elements and in flanking sequences. In addition, the occurrence of methylation at these sites affects digestion and therefore amplification in TID. However, although sequences corresponding to BstXI sites in Mut1 transposons and Bsal sites in Cin4 retrotransposons are often methylated (data not shown), we nevertheless could readily amplify flanking sequences using TID version I. Since methylation of DNA is usually partial, it appears to have a similar effect on TID as do impurities in DNA preparations that inhibit enzymatic reactions. In our hands, however, DNA purity was found to be less important for TID than for Southern hybridization.

The main TID genomic linkers are constructed similarly to Splinkerettes (Devon et al., 1995; Qureshi et al., 1994) and have a Y-like shape (Figure 1). The Y-shaped linker structure prevents DNA synthesis at the linker template during heating prior to the first DNA denaturation. The TID linkers comprise a 32 bp linker oligo as a PCR template and a short subsidiary oligo of 16-20 bp, which bears mismatch base pairs at its AT-rich 3’ end. Unlike the case with Splinkerettes, the 3’ end of the subsidiary oligo does not have a hairpin structure. In Cin4 and En/Spm TID, the subsidiary oligos were modified during synthesis by introducing an amino group at the 3’ end and a phosphate group at the 5’ end. The 3’ amino group blocks DNA synthesis and mismatch repair in the linker that could result from residual exonuclease activity in DNA mini-preps. The 5’ phosphate group allows ligation of the subsidiary oligo to a genomic DNA strand to block it in PCR. This linker design also allows one to avoid DNA synthesis when restriction fragments that include DNA repeats re-anneal during PCR.

The extreme care needed in the design of linkers for TID version I in order to reduce background becomes superfluous in TID version II, since effective elimination of background is achieved by using a single biotinylated transposon primer for pre-amplification, followed by affinity purification on streptavidin. We have found that Y-shaped linkers do not provide notable advantages in TID version II compared to a standard l-shaped design. ‘Hot-start’ PCR was also not critical for segregation analysis of En/Spm transposons in Arabidopsis (data not shown).

Reducing background in TID has another implication for isolating fragments of interest from a preparative gel. Spurious products that are not visible in autoradiographs can complicate the purification of selected fragments after re-amplification. This problem arises in differential RNA display as well. In our experience, 90–95% of fragments identified in TID can be isolated and sequenced without cloning.

We have shown here that ETID allows isolation of transposon insertions that tag the transcribed region of a gene. Hybridization of a biotinylated oligonucleotide to total RNA at the beginning of the ETID procedure is equivalent to the early use of biotin selection in TID procedures. Direct hybrid capture, similar to that used for the isolation of mRNA on oligo(dT) matrices, was incorporated in ETID. An indirect method involving hybridization of a biotinylated oligonucleotide in solution followed by addition of streptavidin-coated particles may be a useful alternative. Depending on the abundance of chimeric transcripts, conventional 5’ RACE PCR on total RNA may be sufficient in some cases for their amplification. However, hybrid selection for transposon sequences is the method of choice for ETID, since it allows one to begin with a large amount of RNA and to perform an additional purification step before the reverse transcription reaction. We believe that further experience with ETID will allow us to improve some of the steps in the protocol. For instance, the use of concentrated guanidine hydrochloride and guanidine thiocyanate solutions as hybridization buffers is conceivable. In addition, end labelling and separation of ETID products in polyacrylamide (as in TID) rather than agarose gels, should facilitate the identification of minor bands.

Applications of TID and ETID procedures

The TID techniques were designed for analysis of the distribution of transposon insertions and the characterization of mutants obtained by conventional gene tagging. Existing TID protocols make it possible to perform the analysis with the En/Spm and Mut1 transposons of Zea mays, which can effectively produce insertion mutants. Considering that the technique was successfully applied to the Cin4 retrotransposon as well, it may be anticipated that insertions of other transposons can be analysed by TID provided corresponding transposon primers are made. Although a universal AIMS protocol for Mut transposable elements has been published (Frey et al., 1998), for some applications it would be advantageous to restrict amplification to one Mut1 subfamily only. The use of other hapaxterministic enzymes to select specific Mut transposons from within the range available is also possible. For instance, digestion with Esp3I in terminal inverted repeats would allow display of sequences flanking Mut1, Mut2, Mut4, Mut5 and Mut8 elements.

Beside gene-tagging applications, the TID techniques could be of use in revealing relationships between species, varieties, breeds and strains. For instance, since stable retrotransposon insertions provide landmarks for chromosomal fragments, the corresponding flanking sequences used as sequence tag markers may be used to uncover lineage relationships. Flanking transposon
sequences amplified by TID can be analysed by spotting them onto filters (S. Steiner-Lange, personal communication), followed by hybridization with probes similar to those described for P-elements in Drosophila (Eggett et al., 1998). This reverse genetics approach allows one to identify new insertions during transposon mutagenesis.

ETID is a direct approach for the isolation of transposon insertions that reside in transcribed portions of genes. This should effectively limit the number of candidate transposon insertions that have to be used for analysis of co-segregation with a mutant phenotype. If RNA is isolated from a particular organ or tissue that is affected in a mutant, this number could be reduced further. ETID enables one to identify knock-out insertion alleles in the heterozygous state. Heterozygous wild-type plants can provide tissues that do not develop properly in mutant plants. Therefore, even lethal insertion alleles may be recovered by ETID. Indeed, ETID may result in the identification of knock-out insertion alleles of genes, which do not cause a phenotype that is easily distinguishable from that of the wild-type. Such alleles may also be found in reverse genetic screens; however, in contrast to this approach, ETID does not require prior knowledge of gene sequences.

Although ETID is very effective in Arabidopsis, we anticipate that its advantages over DNA-based techniques will become particularly obvious in species that form large organs, and whose genomes are too large to allow complete sequencing. Some such species – maize, snapdragon (Antirrhinum majus) and petunia (Petunia hybrida) – harbour numerous active transposons, while others may have to be mutagenized with already characterized transposable elements. While TID and ETID have thus been shown to be very successful in plants, ETID should also be practicable with animal species, where the applicability of genetic methods of segregation analysis may be limited.

Experimental procedures

**Plant material**

Seeds of the following races of Mexican teosinte (Zea mays ssp. parviglumis) were obtained from Maize Germplasm Bank of the International Maize and Wheat Improvement Center (CIMMYT, Mexico); Balsas accession 11407 (collected in Michoacan), Balsas accession 11402 (collected in Guerrero, Central Plateau), and accession 11403 (collected in Michoacan). Seeds of an out-cross progeny of maize (Zea mays ssp. mays) branching mutants and a bz2-mu2 ‘big spot’ line (developed by V. Walbot, Stanford University, Stanford, CA, USA) were obtained from P. Peterson (Iowa State University, Ames, IA, USA). Seeds of the Arabidopsis thaliana dwarf mutant 2AAL45 were obtained from E. Wisman (Max-Planck-Institut für Züchtungsforschung). Plants were grown in greenhouses and growth chambers.

**DNA and RNA isolation**

Genomic DNA was isolated from maize and teosinte plants with 5–6 leaves as described by Dellaporta et al. (1983). Arabidopsis genomic DNA was isolated with the Nucleo Prep Plant DNA extraction kit (Amersham, Braunschweig, Germany) according to the manufacturer’s protocol, with modifications that improved purity. Specifically, the DNA-containing supernatant was extracted with a phenol:chloroform mixture (1:1) and with chloroform after the first centrifugation before removing polysaccharides with the Nucleon matrix. A nucleic acid pellet obtained with sodium acetate and isopropanol was dissolved in TE (0.2 ml per g starting tissue) and the solution was treated with RNAse before the final precipitation in the protocol. Rapid DNA isolation for PCR analysis was performed as described by Edwards et al. (1991).

RNA was isolated from aerial parts of plants with total RNA isolation reagent (Bioun, Hamburg, Germany) and dissolved in formamide (0.5 ml per g of fresh tissue) (Chomczynski, 1992).

**Oligonucleotides and linkers**

Most of the oligonucleotides were synthesized on a Model 392 DNA/RNA synthesizer (Applied Biosystems, Foster City, CA, USA), eluted with ammonia and purified by gel filtration through Sephadex G-25 NAP10 columns (Pharmacia Biotech). Some of the biotinylated oligonucleotides were synthesized by Life Technologies (Neu-Isenburg, Germany), and some of the unlabelled oligonucleotides by MWG-Biotech (Ebersberg, Germany).

**Genomic linkers**

The oligonucleotide LR32 (5’-ACTGATTCTCAACCCGAAAGTATGAGTCCA-3’) constituted the invariable part of several genomic linkers, and the corresponding linker primer, LR26 (5’-ACTCCTACTTCAACCCGAAAGTATGAGTCCA-3’), was used in all TID protocols. In order to produce linkers with Msel-compatible ends, LR32 was annealed either with L16 (5’-TATGGAATCAGATGAGTCCA-3’), giving L1632, or with APL16, giving APL1632. APL16 is a modified L16 with phosphate and amino groups at the 5’ and 3’ ends, respectively.

Cin4 TID

A Cin4-selective linker Bio3’Select used in TID version I was composed of the oligonucleotides 3’ Select (5’-AGCAATCCTGTTAGTCTGAA-3’) and BioLS27 (5’-bio-ATCCTAAATCCCCGCAGTATGAGTCCA-3’). The nested Cin4 primers 7116 (5’-TATGGAAGACGTCTTGGCTCGG-3’) and 7127 (5’-GGAAAGGCTCTTGGCTCGG-3’) were used, respectively, in the first PCR and for radioactive amplification, together with the LR26 linker primer.

Alternatively, in TID version II the biotinylated oligonucleotide Bio8681 (5’-bio-CTACGCAAGGGCGTCTGGTGTTA-3’) was used prior to 7116 and 7127 primers on genomic DNA digested with Msel or Csp6l.

**Mu1 TID version I**

The Mu1-selective linker BioMu1Select was made up of Bio5’Mu (5’-bio-ATCCTAAATCCCCGCAGTATGAGTCCA-3’) and Bst19 (5’-ATCCTAGTCTGGGGGATTTT-3’). A genomic MsTL2032 linker giving TA cohesive ends was obtained by annealing LR32 to MsTL20 (5’-TATGGAATCAGATGAGTCCA-3’). This linker combin-
ation requires that DNA synthesis is performed on the linker template, therefore PCR begins with an extension step at 77°C. The primers Mu26 (5’-GAAGCCACCGCCAACGCTCATTTT-3’) and Mu152 (5’-ATCCCTTTTCTCCTTGATATG-3’), complementary to the 5′ end of Mu1, were used for PCR and for radioactive amplification, respectively.

En/Spm TID version II
The En/Spm transposon primers were En205R (5’-GAAGCAGCAGTGTAAGAAGGA-3’), BioEn205R (5’-bio-GAAGCAGCAGTGTAAGAAGGA-3’), En203R (5’-GAAGCAGCAGGGTCAAACTTTGTAATTTT-3’) and En48R (5’-TTGCTTCTTGAAAAAGGCTTGGC-3’) at the 5′ end; En8130 (5’-GAGCCTGCTTCCCACACTTCTTAC-3’), BioEn8130 (5’-bio-GAGCCTGCTTCCCACACTTCTTAC-3’), En8113 (5’-CAGGTCCTCCACACTTCTTACAGAT-3’), En8136 (5’-CGGCCCACAAGGTGCTCGGAAC-3’) and En8179 (5’-TGACGGCTAAGGTGCTCGGAAC-3’) at the 3′ end.

En/Spm ETID
The SMART II oligonucleotide (5’-AAGCAGTGGTAACAAACGG-AAGGTACAGCAGG-3’) and a PCR primer (5’-AAGCAGTGGA-TAACAGCCAGGT-3’) were components of the SMART PCR cDNA synthesis kit (Clontech, Palo Alto, CA, USA).

PCR components
Taq polymerase and 10× PCR buffer were from Roche Molecular Biochemicals (Mannheim, Germany) and dNTPs from Pharmacia Biotech. PCR reactions were performed in UNOBLOCK (Biometra, Gottingen, Germany) or PCT-200 (MJ Research, Watertown, MA, USA) thermocyclers.

TID
Digestion and ligation of linkers
A typical digestion reaction in 50–100 µl of corresponding buffer with 100 µg/ml of acetylated bovine serum albumin (NEB, Beverly, MA, USA) contained 1–2 µg of miniprep genomic DNA and 8–10 units of restriction enzymes. Digestion with BbsI and Msel (NEB) for Cin4 TID was performed simultaneously in NEB buffer 2, digestion with BsaDI (NEB) for Mu1 TID was performed at 55°C in NEB buffer 3, after which DNA was precipitated and digested with Msel. Digestion with Csp6I (MBI Fermentas) and enzymes from Roche Molecular Biochemicals was carried out in buffers provided. Ligation of linkers was performed overnight at 12–16°C directly in the digestion reaction as in AFLP (Vos et al., 1995) or after purification with QIAquick PCR purification kit (Qiagen, Hilden, Germany). For 1 µg of digested DNA, 50 pmol of genomic linkers, and, if necessary, 12.5 pmol of selective linkers (TID version I) were used.

The ligation mixture was then optionally (TID version I) or obligatorily (TID version II) purified with QIAquick PCR purification kit and eluted with 50 µl of a provided buffer containing 10 mM Tris–HCl, pH 8.5.

Binding to streptavidin-coated particles in TID version I
Dynabeads M-280 Streptavidin (Dynal, Oslo, Norway) paramagnetic particles were used in TID and ETID protocols. Binding and washing steps were conducted mainly according to Dynal protocols after addition of an equal volume of 4 M NaCl to the reaction mix. Inclusion of 0.01% Tween-20 into TE washing buffer was found to be necessary to avoid Dynabeads sticking to tube walls, which readily occurs without salt. Finally beads were suspended in 200 µl of TE containing 0.01% Tween-20.

Cycling biotinylated primer extension in TID version II
Each 50 µl reaction contained 50 ng of genomic DNA with linkers (typically 12.5 µl), 10 pmol of corresponding biotinylated primer, 200 nm dNTPs, 5 µl 10× PCR buffer and 2.5 units of Taq polymerase (Roche Molecular Biochemicals). The reaction programme was: 90°C for 30 sec, 12 × (94.5°C for 35 sec, 64°C for 1 min 30 sec, 73°C for 1 min 30 sec), then 73°C for 3 min. The reduced number of cycles was chosen to avoid random priming PCR. Oil overlay improved the consistency of results. Extension products were purified out of primers using QIAquick PCR purification kit, eluted with 50 µl of 10 mM Tris–HCl, pH 8.5 and bound to Dynabeads as in TID version I.

PCR amplification
The reaction in 25 µl contained appropriate PCR components together with 25 pmol of transposon and linker primers and a pellet of DNA on Dynabeads corresponding to a one-fifth/one-third part of the starting material. In Cin4 and En/Spm TID, a ‘hot start’ was performed during initial incubation at 82°C. The reaction programme was: 82°C pause, 25–30 × (94°C for 35 sec, 64°C for 1 min, 73°C for 1 min), the 73°C for 3 min. In contrast, Mu1 TID was performed without a ‘hot start’, and the initial incubation temperature was 77°C to allow strand displacement during DNA synthesis on linker templates.

Labelling extension reaction
A corresponding transposon primer was labelled with [γ-32P]ATP (Amersham, Braunschweig, Germany) and 0.75 pmol was taken for each reaction. Appropriate PCR components and labelled primer in 8.4 µl were added to 1.6 µl of the PCR amplification reaction and overlaid by oil. The reaction programme was as for PCR amplification, except that the number of cycles was reduced to 15–25. Labelled products were treated similarly to the sequencing reactions, and a one-eighth/one-quarter part of them was loaded on a 6% sequencing polyacrylamide gel. [γ-32P]ATP-labelled PCR molecular weight markers (USB, Cleveland, Ohio, USA) or cycling labelled SegueMark™ (Research Genetics, Huntsville, AL, USA) were used as reference ladders.

Gene electrophoresis and fragment elution
Sequencing gels were bound to supporting glass plates with Silane A174 (Merck, Darmstadt, Germany) while notched glass plates were treated with dichlordimethyliosilane (Merck) and removed after gel electrophoresis. Gels were fixed with 15% acetic acid for 40 min, dried at 80°C and exposed to X-ray films. A preparative gel having wider slots was covered with plastic foil without fixation and drying and was exposed at −70°C to a properly marked and oriented X-ray film. Bands of interest were cut as corresponding gel slices and eluted overnight in 50–100 µl TE. DNA was precipitated using 10 µg of linear acrylamide as carrier (Gaillard and Strauss, 1990) and dissolved in TE. Elution steps were monitored with a counter.

Re-amplification of fragments
Fragments were re-amplified under conditions similar to that of the last TID or ETID PCR using one-fifth to one-third of the elution volume and purified in a 2.5% agarose gel (see below).
ETID

Binding the bait oligonucleotides Bio-En205 or Bio-En8130 and hybridization in ETID

Biotinylated oligonucleotide (200 pmol, 2 μl) was bound to 1 mg of Dynabeads M-280 for 15 min at room temperature in 22 μl of BAITBIND buffer (20 mM Tris–HCl, pH 7.5, 1 mM NaCl), prepared for RNA work according to the manufacturer’s instructions, and resuspended in 65 μl of the BAITBIND buffer. The resulting beads are hereafter called Dynabeads-BBO (biotinylated bait oligonucleotides).

Total RNA (80 μg) was precipitated from formamide with 4 vol EtOH (Chomczynski, 1992), solubilized in 80 μl of a buffer containing 75 μl H2O, 4 μl 1 M Tris–HCl, pH 7.5 and 1 μl RNAsin (40 units μl⁻¹), and 2 μl of 20% SDS were added after solubilization. A 60 μl aliquot of Dynabeads-BBO in the BAITBIND buffer was combined with 60 μl 20× SSPE and transferred to the RNA solution, resulting in a hybridization buffer of 20 mM Tris–HCl, pH 7.5, 6× SSPE, 0.2% SDS, 0.2 units μl⁻¹ RNAsin. Tubes were heated for 3 min at 65°C and incubated for 3 h at 30°C with occasional shaking.

Washing was performed twice with the WASH buffer (20 mM Tris–HCl, pH 7.5, 0.15 M LiCl, 2 mM EDTA, 0.2 units μl⁻¹ RNAsin and 2× Reverse SuperscriptII 1st strand buffer (Life Technologies, Rockville, MD, USA).

Reverse transcription reaction and PCR

RT reactions were done on re-suspended beads in a 20 μl volume with SuperscriptII enzyme (Life Technologies) and SMART II oligonucleotide at conditions close to that specified in the SMART PCR cDNA Synthesis Kit (Clontech, Palo Alto, CA, USA). First PCR was performed in 50 μl using EXPAND High Fidelity PCR System components (Roche Molecular Biochemicals), the SMART II primer and a corresponding En/Spm primer at concentrations of 0.2 μM each and 5 μl of the RT reaction. The first PCR programme for PCT-200 (MJ Research) was: 85°C for 2 min, 20× (94.5°C for 35 sec, 60°C for 35 sec, 68°C for 3 min), the 72°C for 3 min. The second PCR was performed in 50 μl using Taq polymerase (Roche Molecular Biochemicals), the SMART II PCR primer and a corresponding En/Spm primer, each at 0.4 μM concentration, and 2 μl of the first PCR reaction. The second PCR programme was: 85°C for 2 min, 20× (94.5°C for 33 sec, 60°C for 1 min, 72°C for 3 min), 72°C for 3 min. Transposon primers for the 5′ end and the 3′ end were En205R and En8130 in the first PCR and En203R and En8133 in the second PCR, respectively. The third PCR before loading on an agarose gel was done with the En91R or En8136 primers using 2–5 μl of the second PCR under conditions identical to the second PCR. Radioactive products were obtained in the third PCR with the En48R and En8179 labelled primers for analysis on a sequencing gel similarly to TID or on an 2.5% small mol wt agarose (Biozym, Hameln, Germany) gel. Fragments were purified for sequencing with QIAquick gel extraction kit (Qiagen).

Inverse PCR for En/Spm transposon insertions

Aliquots of corresponding genomic DNA (300 ng) were digested in 50 μl with Apal, Clal, MluI, Nael, NciI and XhoI which do not cut in the transposon, combined, extracted with phenol: chloroform (1:1) and precipitated with isopropanol using 20 μg of linear acrylamide as a carrier (Gaillard and Strauss, 1990). The DNA fragments were then separated on 0.6% agarose gel and a portion of the gel corresponding to fragments from 8 to 14–20 kb was cut out. After purification from the gel with the QIAprep spin gel purification kit (Qiagen), the fragments were circularized overnight at 12°C with T4 ligase in 500 μl and concentrated with a Microcon 100 ultrafiltration device (Amicon, Beverly, MA, USA). Aliquots of these were digested separately in 20 μl volumes with BamHI, BglII, BstXI and HpaI restriction enzymes, which cut in the transposon portion. Aliquots (1 μl) of each reaction were taken and subjected to 20 cycles of long PCR (EXPAND High Fidelity PCR system, Roche Molecular Biochemicals) with the En205R and En8130 transposon primers. The amplification profile of long PCR in 25 μl was as follows: 85°C for 2 min, 8× (94°C for 30 sec, 56°C for 30 sec, 68°C for 4 min), 12× (94°C for 30 sec, 56°C for 30 sec, 68°C for 4 min plus time extension for 20 sec at each cycle), 80°C for 10 min. Primers were used at 0.3 μM as recommended by the manufacturer.

Aliquots of 2 μl were taken for the second long PCR of 20 cycles at the same conditions except that the En203R and En8136 transposon nested primers or the En203R and En66-C3 (5'-CTCCCTACCTTTTTCTTGTAGGGCT-3') nested primers were used. The sequence of the En66-C3 primer was derived from that of a junction between the 3′ end of the En/Spm insertion (outlined) and genomic DNA at the C6 locus. The combination of transposon nested primers allowed amplification of multiple En/Spm flanking sequences, while the En203R and En66-C3 primers selectively amplified the C6 insertion fragment. The PCR products were separated in a 0.8% agarose gel. The same fragment of 2.4 kb was apparently amplified in BamHI, BglII and HpaI samples with the En203R and En66-C3 primers, suggesting that it contains the C6 insertion (Figure 5).

Acknowledgements

We are grateful to CIMMYT’s Maize Germplasm Bank (Mexico) for providing seed material of teosinte and maize races, to Peter Peterson (Iowa State University, Ames, IA, USA) for providing seeds of branching mutants of maize, to Ellen Wisman (Max-Planck-Institut für Züchtungsforschung, current address: Michigan State University, East Lansing, MI, USA) for giving us the Arabidopsis dwarf mutant as a material for TID. We thank Hans Sommer for the cDNA library of A. thaliana. We especially appreciate the criticisms of this manuscript made by Paul Hardy.

References


