A Hyperactive Transposase of the Maize Transposable Element *Activator* (*Ac*)

Katina Lazarow,* My-Linh Du,* Ruth Weimer,†,[1] and Reinhard Kunze*,[2]
*Dahlem Centre of Plant Sciences, Freie Universität Berlin, 14195 Berlin, Germany, and †Institute of Genetics and Microbiology, Ludwig-Maximilians-Universität München, 82152 Martinsried, Germany

ABSTRACT *Activator/Dissociation* (*Ac/Ds*) transposable elements from maize are widely used as insertional mutagenesis and gene isolation tools in plants and more recently also in medaka and zebrafish. They are particularly valuable for plant species that are transformation-recalcitrant and have long generation cycles or large genomes with low gene densities. *Ac/Ds* transposition frequencies vary widely, however, and in some species they are too low for large-scale mutagenesis. We discovered a hyperactive *Ac* transposase derivative, *Ac*TPase$_{4x}$, that catalyzes in the yeast *Saccharomyces cerevisiae* 100-fold more frequent *Ds* excisions than the wild-type transposase, whereas the reintegration frequency of excised *Ds* elements is unchanged (57%). Comparable to the wild-type transposase in plants, *Ac*TPase$_{4x}$ catalyzes *Ds* insertion preferentially into coding regions and to genetically linked sites, but the mutant protein apparently has lost the weak bias of the wild-type protein for insertion sites with elevated guanine–cytosine content and nonrandom protein-DNA twist. *Ac*TPase$_{4x}$ exhibits hyperactivity also in *Arabidopsis thaliana* where it effects a more than sixfold increase in *Ds* excision relative to wild-type *Ac*TPase and thus may be useful to facilitate *Ac/Ds*-based insertion mutagenesis approaches.

DNA transposons are widely used in plants and animals as functional genomics tools and gene transfer vehicles. In plants, transposable elements are particularly valuable when large-scale T-DNA insertion mutagenesis is not feasible, *e.g.*, for transformation-recalcitrant species or plants with long generation cycles. The success of transposon insertion mutagenesis strategies depends on high forward mutagenesis rates and a favorable distribution of novel insertions. As forward mutagenesis rates are frequently limited by transposase activity, attempts were made to find hyperactive transposase mutants. For some transposons, such mutants were fortuitously found; in other instances, systematic screening approaches and molecular evolution were successful (Goryshin and Reznikoff 1998; Beall *et al.* 2002; Baus *et al.* 2005; Keravala *et al.* 2006; Mates *et al.* 2009).

The maize *Activator/Dissociation* (*Ac/Ds*) transposable elements have been widely used in plants for gene tagging and functional genomics approaches because they are active in numerous plant species, integrate preferentially into or near to coding regions, and frequently transpose to genetically linked sites, enabling local saturation mutagenesis approaches (reviewed in Kunze and Weil 2002). The successful introduction of *Ac/Ds* elements into yeast revealed that application of these elements is not restricted to plants (Weil and Kunze 2000). Their recent adoption as tools for transgenesis and the generation of gene trap lines in the teleost fishes zebrafish and medaka further emphasized the wide range functionality of *Ac/Ds* elements (Emelyanov *et al.* 2006; Boon Ng and Gong 2011; Froschauer *et al.* 2012).

An impediment for a universal application is the variability of *Ac/Ds* transposition frequency in different plant species. In tobacco, *Ac* transposition frequency is similar as that in maize, with the majority of plants showing between 1 and 5% germinal excisions (Jones *et al.* 1990). In tomato and the monocot crops rice and barley, transposition frequencies range from 2 to 40% with a strong variance in individual lines (Belzile *et al.* 1989; Enoki *et al.* 1999; Koprek *et al.*

doi: 10.1534/genetics.112.139642
Manuscript received February 14, 2012; accepted for publication April 25, 2012
Supporting information is available online at http://www.genetics.org/content/suppl/2012/05/04/genetics.112.139642.DC1.
Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. Ac, P08770.2; Hermes, AAC37217.1; Hobo, PIR: A39652, GI 85002; Tag2, AAD24567.1; Tam3, CAA38906.1; Herves, AAS21248.1; Restless, CAA93759.1; and Tol2, BAA87039.
[1]Present address: Lehárstrasse 3, 81243 München, Germany.
[2]Corresponding author: Dahlem Centre of Plant Sciences, Freie Universität Berlin, Albrecht-Thaer-Weg 6, 14195 Berlin, Germany. E-mail: rkunze@zedat.fu-berlin.de

2000; Lazarow and Lütticke 2009). In contrast, germinal transposition frequencies in *Arabidopsis* typically are low, ranging from 0.07 to 5.7% (Schmidt and Willmitzer 1989; Dean *et al.* 1992). The varying transposition frequencies result from the complex regulation of *Ac* activity, including epigenetic inactivation by the host plant. The DNA methylation state affects the *Ac* promoter activity (Kunze *et al.* 1988) and the binding of the *Ac*TPase to the subterminal binding sites (Wang *et al.* 1996; Wang and Kunze 1998; Ros and Kunze 2001). Moreover, there is no linear relationship between *Ac*TPase levels and transposition frequency. McClintock (1951) has noted already that an increase of *Ac* copy number can result in developmentally delayed transposition and an overall decrease in transposition frequency ("inverse dose effect"). The *Ac*TPase is active only in a limited concentration range because with increasing expression the protein aggregates into large, filament-like structures that are transpositionally inactive (Heinlein *et al.* 1994). The tight regulation of *Ac* transposition may be an evolutionary beneficial mechanism to protect the host against harmful transposition frequencies; however, for transposon tagging, it is a drawback.

Ac belongs to the eukaryotic *hAT* transposon superfamily (Kempken and Windhofer 2001; Kunze and Weil 2002) that had long been thought to be unique (Capy *et al.* 1997). Recently, it was suggested that the transposase catalytic centers of all eukaryotic transposon superfamilies, including the *hAT* elements, share a conserved aspartate-aspartate-glutamate (DDE) catalytic triad in an RNaseH-like fold with prokaryotic transposons like Tn*5* and retroviral integrases (Kulkosky *et al.* 1992; Yuan and Wessler 2011). To this end, however, among the *hAT* elements, the basic necessity of the DDE motif was proven only for *Hermes* (Zhou *et al.* 2004).

In this work, we studied several mutant derivatives of the maize *Activator* transposase. We showed that the putative *Ac*TPase DDE motif residues are essential for transposase activity. We found four amino acid substitutions that result in enhanced *Ac*TPase activity. Their combination in the quadruple mutant *Ac*TPase$_{4x}$ leads in yeast cells to a 100-fold increase in *Ds* element excision compared to the unaltered *Ac*TPase. In stably transformed *Arabidopsis* plants, *Ac*TPase$_{4x}$ also triggers an elevated *Ds* excision activity, suggesting that this protein might be suitable to improve *Ac/Ds* transposon-based insertion mutagenesis approaches.

Materials and Methods

Generation of AcTPase mutants

Amino acid substitutions were introduced into the *Ac*TPase ORF by site-directed mutagenesis of plasmid pWL80 (Weil and Kunze 2000). Mutagenesis primer sequences are listed in the Supporting Information, Table S3. Double mutants were generated from the single mutants by inserting a *Spe*I/*Pfl*MI fragment from pWL80-E249A into pWL80-E336A and an *Eco*RI fragment from pWL80-D545A into pWL80-D459A, respectively. For construction of the *Ac*TPase$_{4X}$ expression vector, a *Nar*I/*Nhe*I *Ac*TPase fragment from pWL80-E249A/E336A was inserted into pWL80-D459A/D545A.

Transposition assay in yeast

CWY1 yeast cells (*ade2:Ds1*; Weil and Kunze 2000) were transformed with the wild-type and mutant *Ac*TPase expression vectors. Growth of yeast cells, selection of Ade+ revertants, and amplification of *Ds* excision sites were performed according to Weil and Kunze (2000).

DNA gel-blot analyses of Ds transposition in yeast

Genomic DNA from yeast was prepared with the E.Z.N.A. yeast DNA kit (Omega Bio-tek) according to the manufacturer's instructions. For DNA gel-blot analysis, 1 μg yeast genomic DNA was digested with *Pvu*I, fractionated by gel electrophoresis, and transferred to Amersham Hybond-NX membrane (GE Healthcare). *Ds* elements were detected by hybridization with a 404-bp DNA probe labeled with DIG-11-dUTP (Roche Diagnostics GmbH).

Analysis of Ds insertion sites in yeast

Splinkerette-PCR was performed according to Uren *et al.* (2009) with minor modifications. The first base (cytosine) of the Splink 1 primer and the long-strand adaptor were omitted. Gene-specific primer sequences are listed in Table S3. For amplification of *Ds* insertion sites, 300 ng yeast genomic DNA was digested with *Sau*3AI. Following adaptor ligation, the DNA was purified and used in the primary PCR.

Plant material and transformation

Arabidopsis thaliana Col-0 plants were transformed with the *Ds* donor construct by the floral-dip method (Clough and Bent 1998). Transgenic plants were selected on solid half-strength MS medium supplemented with 50 μg/ml kanamycin. Two independent *Ds*-transgenic lines, Ds5 and Ds24, were transformed with constructs carrying *Ac*TPase and *Ac*TPase$_{4X}$ expression cassettes. *Ac*TPase and *Ds* transgenic plants were selected on half-strength MS medium supplemented with 50 μg/ml kanamycin and 25 μg/ml hygromycin B.

Construction of binary vectors

The 3-kb *Ds* element from pD16 was excised with *Pst*I and *Pau*I, blunted with T4-DNA-Polymerase (Fermentas), and inserted into the *Sma*I restriction site of pPGTKan3 (Kasaras and Kunze 2010). pD16 was derived from pKU4 (Baker *et al.* 1987) by insertion of the *Ds* element into the *Bam*HI restriction site of pUC19. *Ac*TPase expression constructs were generated by inserting the transposase expression cassette containing the *Ac*TPase$_{103–807}$ cDNA from pcATG10 (Kunze *et al.* 1993) into pCAMBIA1200, resulting in pCAM10ATG. Mutant *Ac*TPase$_{4X}$ expression constructs were generated by exchanging an *Nsi*I fragment from pcATG10 with the respective fragment region from pWL80-E249A/E336A/D459A/

D545A and subsequent transfer as a *Bam*HI fragment into pCAM10ATG.

Analysis of Ds excision sites in Arabidopsis

For analysis of *Ds* excision sites, 10 ng genomic DNA were amplified with primers P1 and P2, followed by a nested PCR with primers P3 and P2 (Figure 6A; Table S3). PCR products were fractionated by gel electrophoresis, and amplificates approximating the size of empty *Ds* donor sites (687 bp) were purified and sequenced either directly or after subcloning.

Results

Substitution of the putative DDE motif residues abolishes AcTPase activity

The amino acid sequence alignment of the *Ac*TPase with the *Hermes* transposase and six other *hAT* element transposases from plants, insects, fungi, and fish suggests that the highly conserved *Ac*TPase residues D301, D367, and E719, located in the conserved *hAT*1 and *hAT*3 regions, form the DDE motif in the catalytic center of the protein (Figure 1 and Figure 2A). We substituted each of these residues by alanine in the N-terminally truncated *Ac*TPase$_{103–807}$ (subsequently referred to as *Ac*TPase) (Kunze *et al.* 1993) and expressed the mutant proteins in the yeast (*Saccharomyces cerevisiae*) strain CWY1, which carries a nonautonomous *Ds* element in the yeast *ADE2* gene (Weil and Kunze 2000). The activity of *Ac*TPase derivatives was measured by the number of Ade$^+$ revertants. No *Ds* excision events were detected among 10^{11} cells for any of the three transposase mutants *Ac*TPase$_{D301A}$, *Ac*TPase$_{D367A}$, and *Ac*TPase$_{E719A}$ (Figure 2B). Proper expression and integrity of the *Ac*TPase derivatives was verified by protein blot analysis (Figure 2C). The complete loss of *Ac*TPase activity demonstrates that the residues D301, D367, and E719 are essential for *Ac*TPase catalytic function and supports the notion that these residues form the DDE motif of the *Ac*TPase. Mutation of any of the catalytic DDE residues should completely abolish transposase activity as has been shown for retroviral integrases, bacterial and *Tc1/mariner* superfamily transposases, and the *hAT* element *Hermes* transposase (reviewed in Haren *et al.* 1999; Zhou *et al.* 2004).

Quadruple mutant AcTPase$_{4x}$ is hyperactive in yeast

In the course of functional analyses of the *Ac*TPase, we have generated various mutant *Ac*TPase proteins with single amino acid substitutions and screened their excision activities in the yeast transposition assay (Figure 2). Of these, the four *Ac*TPase derivatives E249A, E336A, D459A, and D545A induced three- to fivefold increases in *Ds* excision frequency when compared to the native *Ac*TPase (Figure 2B). To investigate if the combination of these mutations leads to a further enhancement, we first constructed the two double mutants *Ac*TPase$_{E249A/E336A}$ and *Ac*TPase$_{D459A/D545A}$. With 11- and 13-fold increases in *Ds* excision frequency, both double mutants exhibit a more-than-additive rise in activity (Figure 2B). Subsequently, we generated the quadruple mutant *Ac*TPase$_{E249A/E336A/D459A/D545A}$ (*Ac*TPase$_{4x}$). The expression of this protein resulted in a 10-fold increase in *Ds* excisions relative to the two double mutants and an almost 100-fold increase compared to wild-type *Ac*TPase (Figure 2B), indicating a strong synergistic effect of the mutations. The transposase amounts in the yeast cells expressing *Ac*TPase$_{4x}$ and *Ac*TPase$_{E249A}$ are slightly reduced (Figure 2C). Although *Ac*TPase is subject to overexpression inhibition (Heinlein *et al.* 1994; Weil and Kunze 2000), this slightly reduced transposase level is unlikely to account for the significant increase in *Ds* mobilization activity of *Ac*TPase$_{4x}$.

To investigate if the enhanced *Ac*TPase$_{4x}$ activity is accompanied by altered processing of the *Ds* excision site, we analyzed the empty donor sequences in 93 and 100 Ade$^+$ revertants generated by *Ac*TPase and *Ac*TPase$_{4x}$, respectively (Figure 3). In 43 *Ac*TPase- and 92 *Ac*TPase$_{4x}$-derived Ade$^+$ revertants, the *Ds* excision site sequences are consistent with the DNA hairpin model for *Ds* footprint formation (Kunze and Weil 2002). Two footprints formed by *Ac*TPase$_{4x}$ (Figure 3, rows 15 and 17) and one formed by *Ac*TPase (Figure 3, row 30) conflict with this model. We cannot distinguish if these sequences result from aberrant DNA repair after *Ds* excision or from PCR errors. In three other Ade$^+$ revertants, two generated by *Ac*TPase$_{4x}$ and one by the unmutated *Ac*TPase, 11–42 bases from the transposon 5′ end are left behind at the excision site (Figure 3, rows 32–34), whereas the 3′ end has been processed according to the hairpin model. In these revertants, no reintegration of *Ds* was detected, suggesting that they result from abortive excision events. In the remaining 4 *Ac*TPase$_{4x}$- and 49 *Ac*TPase-generated Ade$^+$ revertants, *Ds* was only partially lost (Figure 3, row 35). In all these "type 35" revertants, the same 42 bases from the *Ds* 5′ end and 19 bases from the *Ds* 3′end, including both terminal inverted repeats, are left behind. These revertants could result from homologous recombination at the 12-bp direct repeat TACCGACCGTTA present in the *Ds* 5′ and 3′ subterminal regions. Consistent with this assumption, no reintegration of *Ds* was observed in any of these revertants. The wild-type *Ac*TPase apparently gave rise to 12-fold more type 35 revertants than the *Ac*TPase$_{4x}$. Although the overall number of the 93 *Ac*TPase-generated Ade$^+$ revertants were isolated from seven independent experiments with, on average, 16 revertants per culture, we cannot exclude that the high number of type 35 revertants partially results from clonal propagation of early events after induction of transposase expression in the yeast culture.

Not counting the type 35 revertants, in *Ac*TPase$_{4x}$- and *Ac*TPase-expressing cells, 25 and 19 different *Ds* excision footprint sequences were isolated. Both transposases generate the same two predominant footprint types accounting for 40% (n = 96) of the footprints formed by *Ac*TPase$_{4x}$ and 41% (n = 44) of the footprints formed by *Ac*TPase (Figure 3, rows 1 and 2), suggesting that *Ac*TPase$_{4x}$ does

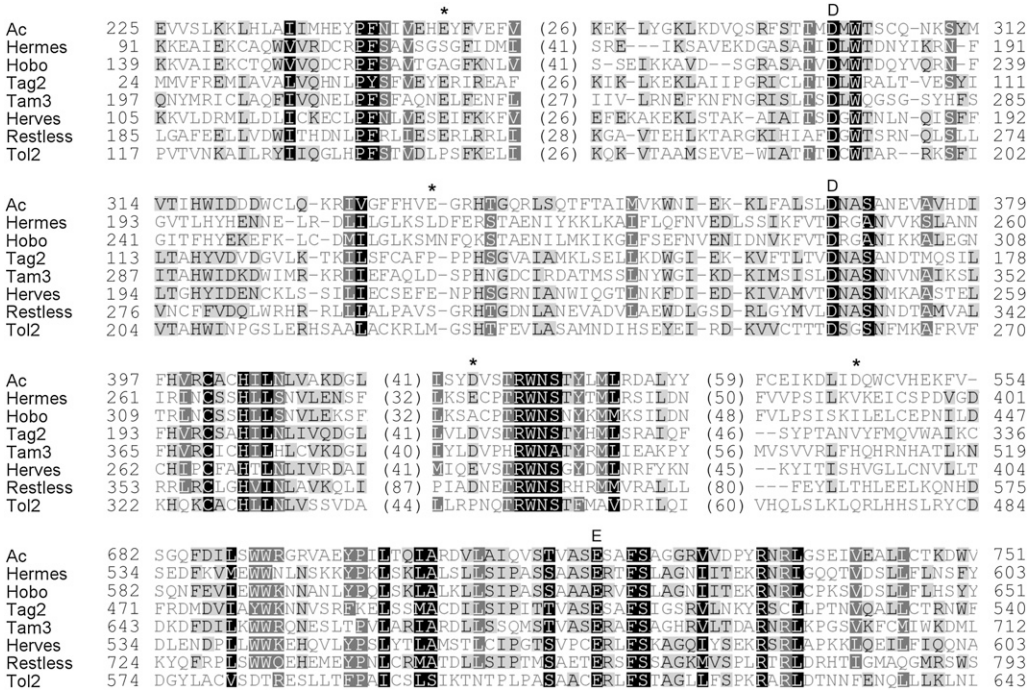

Figure 1 Alignment of eight *hAT* transposases. The DDE triad is marked with letters above the alignment. Asterisks indicate the four amino acids E249, E336, D459, and D545 mutated in the hyperactive *Ac*TPase$_{4x}$.

not differ from *Ac*TPase in the biochemistry of *Ds* excision. Moreover, at most *Ds* excision sites, generated by either version of the transposase, the chromosome breaks seem to have been resolved by microhomology-dependent end-joining (Yu *et al.* 2004) (Figure 3).

Ds reintegration rate affected by AcTPase$_{4x}$ is unchanged in yeast

Although the two DNA single-strand cleavage steps at the transposon donor site and the strand-transfer step to the target DNA in nonreplicative transposition are catalyzed by the same active center of the transposase, they can be uncoupled in transposase mutants (Bolland and Kleckner 1996). We therefore compared the reintegration frequency of *Ds* elements in Ade$^+$ revertants obtained with *Ac*TPase or *Ac*TPase$_{4x}$. DNA gel-blot analysis revealed that *Ac*TPase$_{4x}$ catalyzes *Ds* reintegration as efficiently as the wild-type protein: ignoring the type 35 revertants, 57% (n = 44) of the excision events conducted by the native *Ac*TPase and 57% (n = 88) of the events induced by *Ac*TPase$_{4x}$ were followed by reintegration of *Ds* into a new genomic position. In all cases, a single reintegrated *Ds* copy was observed (Figure 4A). Consequently, not only *Ds* excisions but also the total number of *Ds* reintegrations per yeast cell effected by the *Ac*TPase$_{4x}$ protein is ~100-fold higher than with the *Ac*TPase.

AcTPase$_{4x}$ exhibits an altered Ds insertion sequence preference

Upon integration at new genomic positions, *hAT* transposable elements generate 8-bp target site duplications (TSD). By sequencing the flanking genomic DNA on both sides of 14 *Ds* reinsertions formed by *Ac*TPase$_{4x}$ and 7 reinsertions formed by wild-type *Ac*TPase, we confirmed that in all cases the expected TSD was created (Table S1), indicating that the architecture of the transpososome is not or is only marginally altered with the *Ac*TPase$_{4x}$ protein.

Recently, in maize a preference for *Ds* insertions into DNA target sites with an elevated guanine–cytosine (GC) content and highly nonrandom protein–DNA twist values was detected (Vollbrecht *et al.* 2010). The average GC content of the 18 and 35, 8-bp target sites chosen by *Ac*TPase and *Ac*TPase$_{4x}$ is 42.4 and 38.9%, respectively (Table S1). The GC content in a 400-bp window centered on the 18 *Ds* insertion sites chosen by *Ac*TPase is 40.2% and that for the 35 *Ds* reinsertions promoted by *Ac*TPase$_{4x}$ is 38.9%. Thus, the GC content of 8-bp *Ds* insertion sites and the chromosomal vicinity targeted by wild-type *Ac*TPase is slightly higher than that of the *Ac*TPase$_{4x}$ landing regions, which closely match the average yeast genomic GC content of 39%. We then analyzed the 8-bp *Ds* insertion site sequences ±20 bp on either side for protein DNA-twist prediction (Olson *et al.* 1998). The insertion sites chosen by the *Ac*TPase display a pattern of alternating protein–DNA twist values with a twofold symmetry about the center of the target site that is highly similar to the pattern observed in maize. Interestingly, the target sites chosen by the *Ac*TPase$_{4x}$ do not exhibit such a pattern (Figure 5).

Preference for linked transposition and insertion into genes is conserved for AcTPase$_{4x}$

Mapping of the *Ds* reinsertion sites on the yeast chromosomes indicates for both transposases a weak preference for linked transposition (Figure 4B). Seven of the 34 mapped *Ac*TPase$_{4X}$-driven *Ds* transpositions (21%) reinserted on chromosome XV, 4 of them <72 kb distant from the *Ds* launch pad in the *ade2* locus (Figure 4B; Table S1 and Table

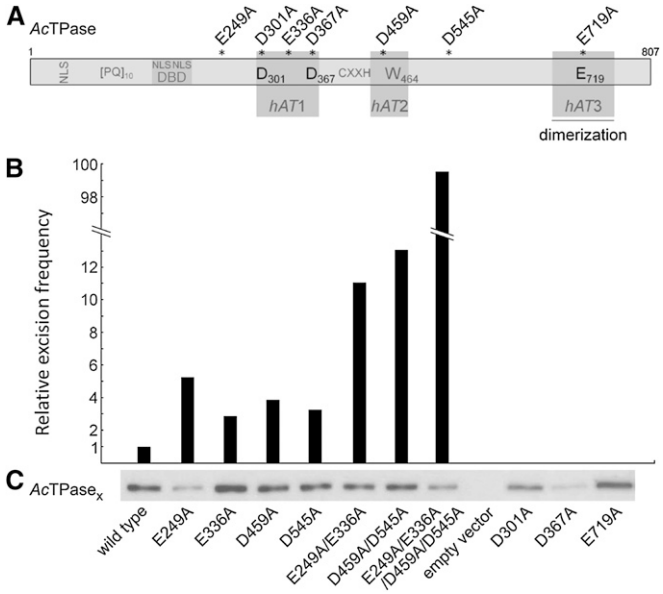

Figure 2 *Ds* excision activities of *Ac*TPase mutants. (A) Schematic of the *Ac* transposase. The three amino acid sequence regions *hAT*1 (aa 293–376), *hAT*2 (aa 442–490) and *hAT*3 (aa 678–759) are conserved in *hAT* transposases. The *hAT*3 region is a dimerization domain. Residues D301, D367, and E719 presumably form the catalytic triad, and tryptophan W464 is supposedly involved in hairpin formation. NLS, nuclear localization signals. DBD, DNA-binding domain. $(PQ)_{10}$, dipeptide repeat essential for *Ac*TPase function. (B) Relative *Ds* excision frequencies triggered by *Ac*TPase derivatives in yeast. The values are the median of nine independent experiments for the wild-type *Ac*TPase and the quadruple *Ac*TPase mutant and of five independent experiments for the other *Ac*TPase derivatives. For the relative values, *Ds* excision frequency induced by wild-type *Ac*TPase was set to 1. The absolute median excision frequency promoted by wild-type *Ac*TPase was 2.9×10^{-5} Ade$^+$ cells/total living cells and for the quadruple mutant 288×10^{-5} Ade$^+$ cells/total cells. (C) Protein gelblot analysis of wild-type *Ac*TPase and mutant derivative expression in yeast.

S2). When *Ds* was mobilized by the wild-type *Ac*TPase, 3 of 17 reintegrations (18%) occurred on chromosome XV, one of them at a distance of only 20 kb (Figure 4B; Table S1 and Table S2). Accordingly, both transposases deployed twice as many *Ds* insertions into the chromosomal vicinity than expected for random distribution in the genome (Table S2). Even though the yeast genome with 70% open reading frames has a much higher fraction of coding sequences than plants, it is conspicuous that 33 (94%) of the 35 analyzed *Ac*TPase$_{4X}$-promoted and 17 (94%) of the 18 *Ac*TPase-promoted *Ds* insertions occurred in annotated genes or predicted open reading frames. Each transposase produced one insertion into moderately repeated sequences, yeast retrotransposon TYA, and the long terminal repeat of TY3, respectively.

Hyperactive AcTPase$_{4X}$ induces more frequent Ds excision in planta

To investigate whether *Ac*TPase$_{4X}$ is also hyperactive in plants, we generated two independent *A. thaliana* starter lines that carry a 3-kb *Ds* element flanked by short, GC-rich sequences from the original *wx-m7* maize locus including the TSD. The *Ds* starter lines were transformed with expression cassettes for either *Ac*TPase or *Ac*TPase$_{4x}$ under control of the *Agrobacterium tumefaciens* TR-octopine DNA 2′ promoter. One to two rosette leaves from 382 individual *Arabidopsis* plants containing both the *Ds* element and the *Ac*TPase or *Ac*TPase$_{4x}$ expression construct were analyzed for somatic excision of *Ds* (Table 1 and Figure 6).

From 10 (6%) of the 170 *Ac*TPase-expressing plants, *Ds* footprints could be amplified and sequenced (Table 1). These samples were all derived from progeny of parental plant no. 20. In progeny from parental lines no. 30 and 31, no empty donor sites could be amplified although they expressed *Ac*TPase. In contrast, from 70 (33%) of the 212 descendants of the three independent *Ac*TPase$_{4x}$-expressing parental plants, empty donor sites were amplified (Table 1). From 15 of these plants, two to five different *Ds* excision products were recovered, indicating that multiple independent somatic *Ds* excisions had occurred in the leaves. Ten empty donor-site amplificates from *Ac*TPase$_{4x}$-expressing plants could not be sequenced across the *Ds* excision site because several sequences were superimposed distal to it. However, all 10 sequence reads, starting from primer P2 in the 3′-flanking DNA (Figure 6A), terminate at the former *Ds* position with a cytosine that is the complement of the guanine immediately flanking the *Ds* 3′ end. This cytosine is generated during hairpin resolution and thus proves that transposase-promoted *Ds* excision had occurred at these sites. Taken together, we detected a more than sixfold higher somatic *Ds* excision frequency in *Arabidopsis* plants expressing *Ac*TPase$_{4x}$ compared to plants expressing the unmutated protein, indicating that the *Ac*TPase$_{4x}$ is also hyperactive in plants.

Both transposases give rise to the same predominant "gc" footprint (Table 1 and Figure 6C). It was detected in 60% of the *Ac*TPase-generated *Ds* excision sites and in 52% of the *Ac*TPase$_{4x}$-generated *Ds* excision sites (49 of 95 sequence reads including those with superimposed sequences behind the *Ds*). These frequencies closely match the results from a recent deep-sequencing survey of *Ac/Ds* excision from a donor locus flanked by the identical 8-bp TSD in *Arabidopsis*, where the same predominant gc footprint was found to account for ~55% of all reads (Huefner *et al.* 2011).

Unique *Ds* excision footprints were detected in four *Ac*TPase- and in 21 *Ac*TPase$_{4x}$-expressing plants, including 13 plants with multiple excision events (Figure 6C). The generation of the gc footprint and the majority of unique footprints can be explained by the hairpin model in conjunction with the microhomology-dependent nonhomologous end-joining (NHEJ) model for DNA break repair (Figure 6C) (Yu *et al.* 2004). The residual nine footprints (plant nos. 1031, 1055, 730, 753, 754, 769, 871, 910, and 942) consist of segments derived by processing of the flanking DNA hairpin and additional nonpalindromic nucleotides that originate neither from the *Ds* nor from the flanking DNA. *Ds* excision sites from plant nos. 753 and 754 (Figure 6C) contain insertions

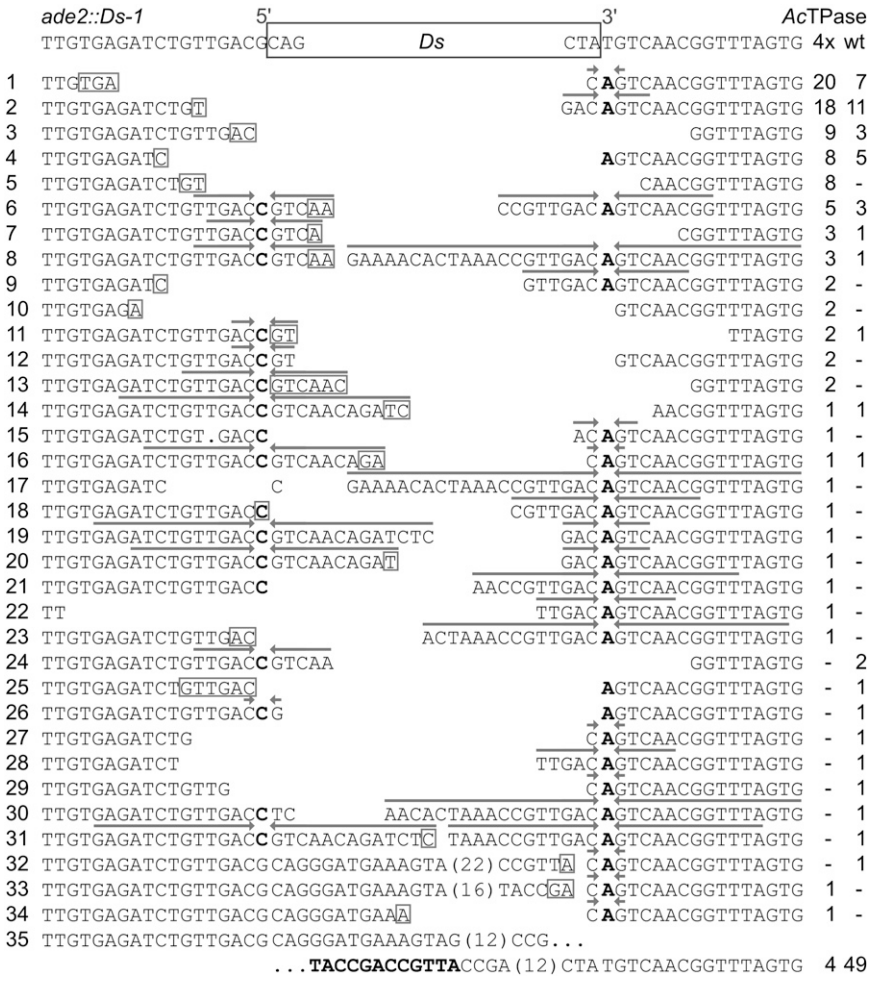

Figure 3 *Ds* excision footprints formed by *Ac*TPase and *Ac*TPase$_{4x}$ in yeast. The top row shows the sequence at the *Ds* insertion site in the yeast *ADE2* gene. Rows 1–35 show the recovered *Ds* excision footprints and at the right their incidence in independent Ade$^+$ revertants from wild-type *Ac*TPase (wt) or hyperactive *Ac*TPase$_{4x}$ (4x) expressing cells. Putative microhomologies at flanking DNA fusion sites are indicated as boxed nucleotides. Arrows above sequences highlight inverted repeats centered around the complementary bases C and A of the nucleotides bordering the *Ds* element that result from resolution of intermediate hairpin structures formed at the *Ds*-flanking host DNA during excision. Boldface letters in row 35 indicate the remaining copy of the 12-bp direct repeat in the *Ds* ends.

of 57 and 71 bases, respectively, which partially match *Arabidopsis* genomic sequences. In 7 *Ac*TPase$_{4x}$ plants (7.4% of the 95 sequence reads), we detected excision sites where the *Ds* element and one copy of the duplicated target site were precisely eliminated, resulting in restoration of the original target-site sequence ("restoring excision"; Figure 6C).

Discussion

The efficiency of transposon insertional mutagenesis approaches is frequently hampered by low transposase activity, which may have evolved to protect the host genome against unfavorably high mutation frequencies. Consistent with this idea, various hyperactive transposases were identified by mutagenesis in prokaryotic and animal transposons (Weinreich *et al.* 1994; Lampe *et al.* 1999; Beall *et al.* 2002; Baus *et al.* 2005; Mates *et al.* 2009; Yusa *et al.* 2011). The maize *Activator* transposase mutants reported here are the first examples of a hyperactive plant transposase. The alignment of eight *hAT* element transposases, most of which were shown to be active, illustrates that none of the amino acids substituted in the hyperactive *Ac*TPase$_{4x}$ are in highly conserved positions (Figure 1). The hyperactive mutations act synergistically as their combination into a quadruple mutant *Ac*TPase$_{4x}$ results in a 100-fold higher *Ds* excision activity in yeast. Synergistic enhancement of hyperactivity following the combination of two to nine single amino acid substitutions has also been observed for Tn*5*, *piggyBac*, *Himar1*, and *Sleeping Beauty* transposases (Weinreich *et al.* 1994; Lampe *et al.* 1999; Baus *et al.* 2005; Yusa *et al.* 2011).

Except for Tn*5*, little is known about the biochemical mechanisms of transposase hyperactivity. In the case of SB100X, it is hypothesized that the mutations alter the folding properties of the transposase (Mates *et al.* 2009). As the four amino acid substitutions in *Ac*TPase$_{4x}$ are located in a region that is assumed to contain a multimerization interface (Essers *et al.* 2000) and with increasing concentration the *Ac*TPase activity is limited by progressive aggregation into nonfunctional complexes (Heinlein *et al.* 1994), we tested whether *Ac*TPase$_{4x}$ has altered aggregation properties. However, upon expression in *Escherichia coli*, yeast, and petunia protoplasts, we did not detect any difference in the wild-type *Ac*TPase (our unpublished results). The hyperactive mutation D459A is only two amino acids apart from a tryptophan residue that is highly conserved among *hAT* transposases (Figure 1; Zhou *et al.* 2004). In *Hermes* and *piggyBac* transposases, this tryptophan supposedly acts

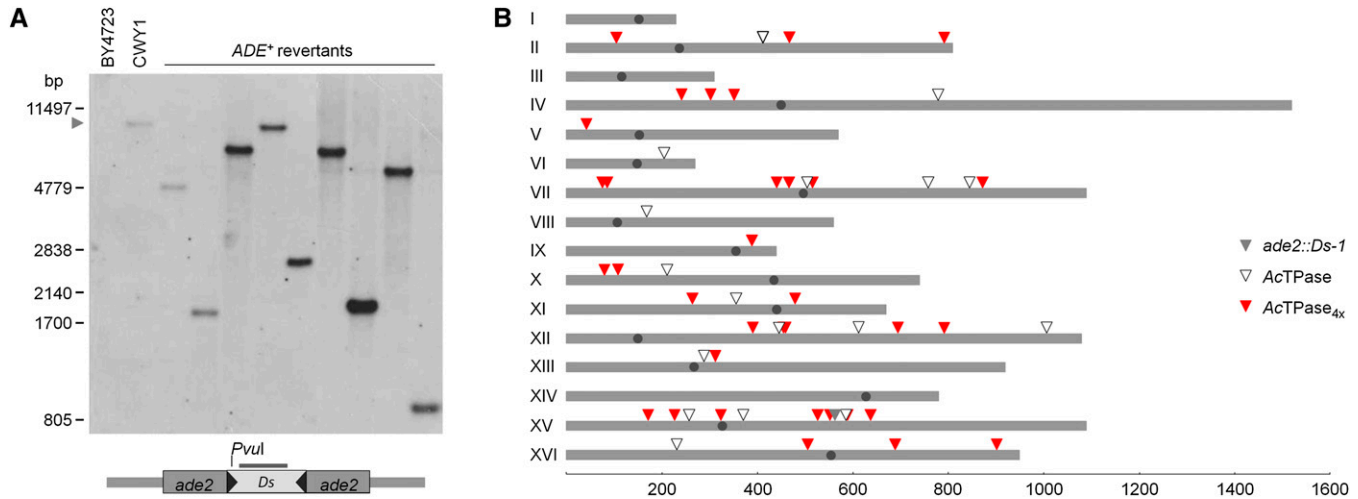

Figure 4 Reintegration of *Ds* elements in yeast. (A) Examples for reintegrated *Ds* elements from *Ac*TPase$_{4x}$-promoted transposition. DNA gel blot with *Pvu*I-digested genomic DNA from yeast strains BY4723, CWY1, and nine individual Ade$^+$ revertants, hybridized with a *Ds*-specific probe. The gray arrowhead marks the *Ds* donor site in the *ade2* locus of yeast strain CWY1. (B) Distribution of *Ds* insertion sites in the yeast genome. The 16 yeast chromosomes are drawn to scale as gray bars with the centromere for each chromosome indicated by a dot. The gray triangle marks the initial position of the *Ds* element in the *ade2* locus on yeast chromosome XV. Red triangles indicate reintegration of the *Ds* element promoted by the hyperactive *Ac*TPase$_{4X}$, and open triangles by the *Ac*TPase. Scale bar is in kilobase pairs.

in DNA hairpin formation and base flipping (Zhou *et al.* 2004; Mitra *et al.* 2008).

*Ac*TPase$_{4x}$ is also hyperactive in *Arabidopsis*, although apparently not as pronounced as in yeast. However, the comparative quantification of *Ac*TPase activity in sporophytic plant tissue by amplification of *Ds* excision sites is inherently imprecise and underestimates the excision frequency, as excisions in single cells and very small tissue

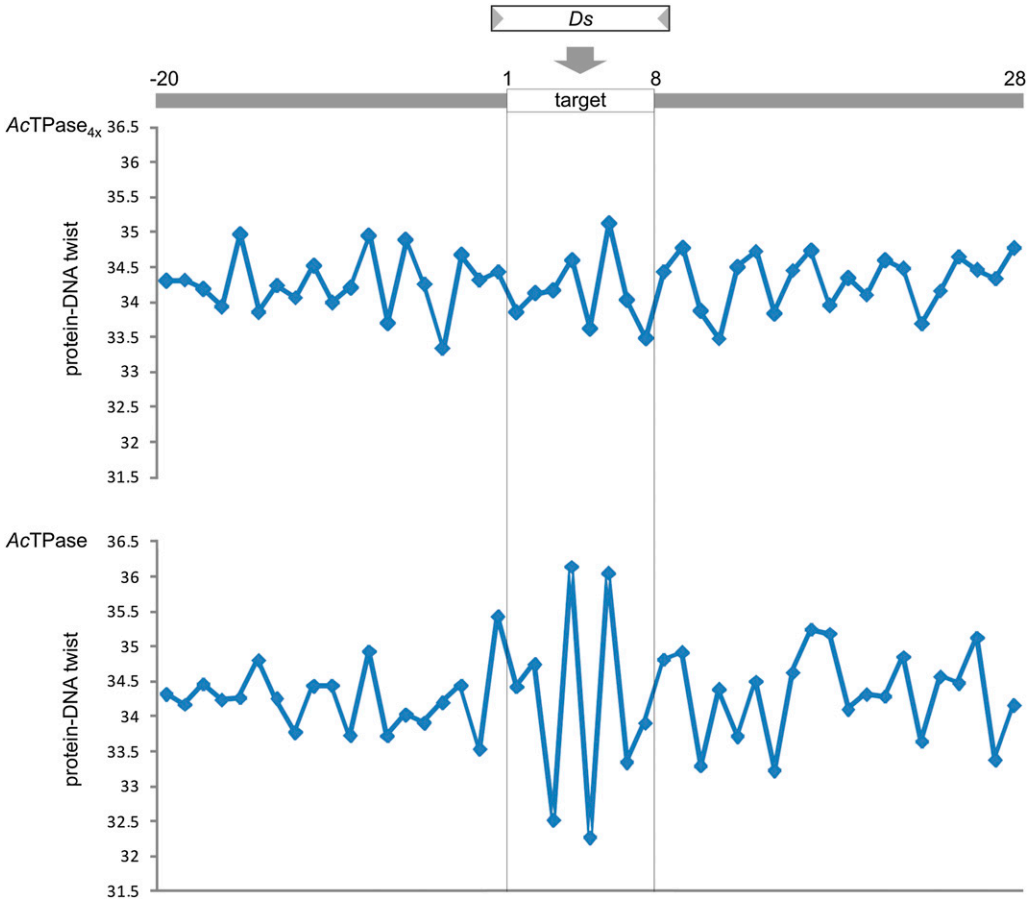

Figure 5 Predicted twist-angle torsion at *Ds* insertion sites in yeast. Average values for protein-DNA twist (*y* axis) are plotted for each nucleotide position against the sequence of the 8-bp *Ds* insertion sites ±20 bp adjoining sequence (*x*-axis) from 35 *Ac*TPase$_{4x}$-generated *Ds* insertions (top) and 18 *Ac*TPase-generated *Ds* insertions (bottom). The borders of the 8-bp *Ds* target sites are indicated by the vertical lines.

Table 1 *Ds* excision in *Arabidopsis*

Transposase	Line[b]	Plants[c]	Excisions[d]	*Ds* excision site[a]			
				A	B	C	D
*Ac*TPase	30	54	—	—	—	—	—
	31	50	—	—	—	—	—
	20	66	10	6	4	—	—
Σ[e]		**170**	**10**	**6**	**4**	—	—
*Ac*TPase$_{4x}$	29	68	21	7	5	6	3
	42	72	21	13	—	5	3
	41	72	28	17	3	4	4
Σ		**212**	**70**	**37**	**8**	**15**	**10**

[a] Classification of individual plants according to the *Ds* excision sites isolated. A: number of plants with "gc" footprint. B: number of plants showing unique footprints including "restoring" excision products. C: number of plants with multiple footprints. D: sequence was unreadable distal to the *Ds* excision site.
[b] The three independent plant lines analyzed for each transposase are, with the exception of 29 (Ds5), all derived from *Ds* starter line Ds24.
[c] Number of plants analyzed for excision of *Ds*.
[d] Number of individual plants with empty *Ds* donor sites.
[e] Σ = total number of *Ac*TPase and *Ac*TPase$_{4x}$ plants and excisions, respectively.

sectors may remain undetected and multiple independent excisions generating the same (predominant) footprint are counted as one event. On the other hand, it is also possible that hyperactivity is unequally expressed in different host systems. Of 18 *piggyBac* transposase mutants showing hyperactivity in yeast, only five exhibited hyperactivity in mammalian cells (Yusa *et al.* 2011). Similarly, hyperactivity of Sleeping Beauty transposase SB10 in HeLa cells was stronger than in mice (Baus *et al.* 2005).

DNA hairpin formation, resolution, repair, and end joining at the empty donor site, first suggested by Coen *et al.* (1986) for the snapdragon *Tam3* element, is the universal excision site processing pathway for *hAT* transposons (Coen *et al.* 1986; Kunze and Weil 2002; Yu *et al.* 2004). This mechanism results in footprints with short deletions and palindromes at the repaired DNA joint centered around the complement of the base adjacent to the transposon ends. The vast majority (90%) of footprints, including the predominant ones, generated by *Ac*TPase$_{4x}$ and *Ac*TPase in yeast and in *Arabidopsis*, can be explained by this model. Fifteen of the 22 distinct footprint types that had been reported by Yu *et al.* (2004) using the wild-type *Ac*TPase were also recovered from wild-type *Ac*TPase- and *Ac*TPase$_{4x}$-expressing yeast cells in this study, suggesting that the basic excision mechanism remains unaffected by the four amino acid substitutions in *Ac*TPase$_{4x}$. This conclusion is corroborated by the *Ds* excision footprint spectrum *in planta*. The frequency of the predominant gc footprint generated by the quadruple mutant *Ac*TPase$_{4x}$, the low frequency of 4%

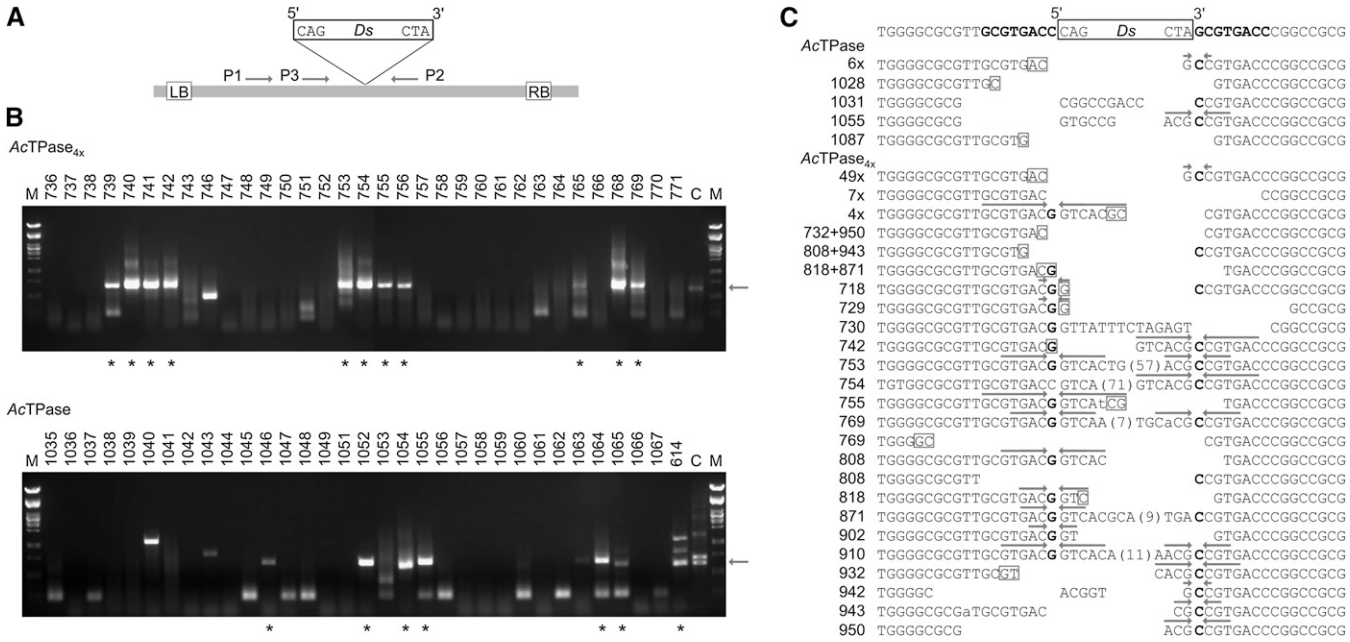

Figure 6 Analysis of *Ds* excision sites in *Arabidopsis*. (A) Schematic of the T-DNA in pCAM10ATG. LB, left T-DNA borders. RB, right T-DNA borders. *Ds*, 3-kb *Ds* element. P1–P3, primers used for amplification of empty *Ds* donor sites. (B) Screen for somatic *Ds* excision events. For each transposase, a representative agarose gel for screening and recovery of *Ds* excision sites from individual *Arabidopsis* plants is shown. Arrows at the right indicate the expected size for empty *Ds* donor-site amplificates of ~687 bp. Asterisks indicate PCR products from individual plants that approximate this size and were subjected to sequence analysis. (C) *Ds* excision footprints formed by *Ac*TPase and *Ac*TPase$_{4x}$ in *Arabidopsis*. The top line shows the sequence at the *Ds* donor site. Letters in boldface flanking the *Ds* indicate the 8-bp target-site duplication. The lines below show transposon footprints in individual plants after *Ds* excision by *Ac*TPase or *Ac*TPase$_{4x}$. The predominant footprint gc was detected in 6 *Ac*TPase ("6×") and in 49 *Ac*TPase$_{4x}$ individuals ("49×"). A "restoring" excision product was recovered from seven *Ac*TPase$_{4x}$ plants ("7×"). Four individual *Ac*TPase$_{4x}$ plants share the same footprint ("4×"). From plant nos. 769, 818, 871, 943, and 950 two and from plant no. 808 three distinct footprints were isolated, respectively. Bases with weak signal strength in the sequence reads are indicated in lowercase. Putative microhomologies at flanking DNA fusion sites are indicated as boxed nucleotides. Arrows above sequences highlight inverted repeats centered around the complementary bases G and C of the nucleotides bordering the *Ds* element that result from resolution of intermediate hairpin structures formed at the *Ds*-flanking host DNA during excision.

restoring *Ds* excisions, and the occurrence of excision products with short sequence insertions, suggesting that transposon excision-induced double-strand break repair is not strictly dependent on the canonical NHEJ pathway factors Ku70 and DNA ligase IV (Lig4), are concordant with the results from a recent deep sequencing survey of somatic *Ds* excision catalyzed by the wild-type *Ac* transposase in *Arabidopsis* (Huefner *et al.* 2011).

Among the Ade^+ revertants from *Ac*TPase$_{4x}$ and, more frequently, from *Ac*TPase-expressing cells, we detected the same type of apparently incomplete *Ds* excision products (Figure 3, row 35) that had also been found by Yu *et al.* (2004). This product could emerge either from a homologous recombination event between two perfect 12-bp direct repeats in both transposon ends or from gene conversion and microhomology-dependent NHEJ after excision of *Ds* during replication from one daughter chromatid. As this product was never obtained in the absence of transposase, we speculate that its formation is promoted by the transposase in a similar way as in plants, where the presence of an active *Ac* element can greatly enhance intramolecular homologous recombination between direct repeat sequences in maize and *Arabidopsis* (Athma and Peterson 1991; Xiao *et al.* 2000). In the transpososome, the *Ac*TPase synapses the transposon ends, which might stimulate recombination between directly repeated sequences in the flanking host DNA and possibly also within the transposon ends.

Consistent with other *Ac/Ds* transposition studies in plants and yeast, we do not recognize a target sequence preference for the *Ds* insertion site selection with the *Ac*TPase or *Ac*TPase$_{4x}$. The analysis of 1741 *Ds* insertion sites in maize revealed, however, that the elements have a preference for insertion into sequences with a high GC content and that the insertion sites exhibit a pattern of alternating nucleotide pairs with more and less than average DNA deformability, respectively (Vollbrecht *et al.* 2010). This suggests that *Ac/Ds* insertion is guided by structural features of the DNA rather than by the primary DNA sequence. Interestingly, for the wild-type *Ac*TPase, we observe in yeast, similar to that in maize, a slightly elevated GC content at the insertion sites and a periodicity in the protein-DNA twist, whereas the hyperactive *Ac*TPase$_{4x}$ apparently has largely lost the preference for target sites with these features.

In plants, 50–80% of excised *Ac/Ds* elements reinsert into the genome (reviewed in Kunze *et al.* 1997; Kunze and Weil 2002). The hyperactive *Ac*TPase$_{4x}$ induces the same reintegration frequency as the wild-type protein, suggesting that excision and reintegration activities are effected by different amino acids. Similarly, in a hyperactive *piggyBac* transposase, the excision reaction was enhanced, whereas the integration reaction remained unchanged (Yusa *et al.* 2011).

In plants, the propensity of *Ac* and *Ds* elements to transpose to genetically linked sites and to insert into coding regions and the lack of insertion-site specificity have proven to be advantageous. In this study, we find that these properties are largely conserved in yeast. In comparison, the transposase of the insect transposon *Hermes* directs only 40 to 45% of insertions into ORFs of the *Saccharomyces* genome (Gangadharan *et al.* 2010). Although it has not been shown yet whether the *Ac*TPase$_{4x}$ excision and reintegration properties are quantitatively similar in plants, the protein is a promising candidate to improve transposon mutagenesis and gene-tagging efficiencies.

Acknowledgments

We thank Anne Herrmann for excellent technical support. This work was supported by Deutsche Forschungsgemeinschaft grant KU-715/9.

Literature Cited

Athma, P., and T. Peterson, 1991 Ac induces homologous recombination at the maize P locus. Genetics 128: 163–173.

Baker, B., G. Coupland, N. V. Fedoroff, P. Starlinger, and J. Schell, 1987 Phenotypic assay for excision of the maize controlling element Ac in tobacco. EMBO J. 6: 1547–1554.

Baus, J., L. Liu, A. D. Heggestad, S. Sanz, and B. S. Fletcher, 2005 Hyperactive transposase mutants of the Sleeping Beauty transposon. Mol. Ther. 12: 1148–1156.

Beall, E. L., M. B. Mahoney, and D. C. Rio, 2002 Identification and analysis of a hyperactive mutant form of Drosophila *P*-element transposase. Genetics 162: 217–227.

Belzile, F., M. W. Lassner, Y. Tong, R. Khush, and J. I. Yoder, 1989 Sexual transmission of transposed Activator elements in transgenic tomatoes. Genetics 123: 181–189.

Bolland, S., and N. Kleckner, 1996 The three chemical steps of Tn*10*/IS*10* transposition involve repeated utilization of a single active site. Cell 84: 223–233.

Boon Ng, G. H., and Z. Gong, 2011 Maize Ac/Ds transposon system leads to highly efficient germline transmission of transgenes in medaka (Oryzias latipes). Biochimie 93: 1858–1864.

Capy, P., T. Langin, D. Higuet, P. Maurer, and C. Bazin, 1997 Do the integrases of LTR-retrotransposons and class II element transposases have a common ancestor? Genetica 100: 63–72.

Clough, S. J., and A. F. Bent, 1998 Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. 16: 735–743.

Coen, E. S., R. Carpenter, and C. Martin, 1986 Transposable elements generate novel spatial patterns of gene expression in *Antirrhinum majus*. Cell 47: 285–296.

Dean, C., C. Sjodin, T. Page, J. D. G. Jones, and C. Lister, 1992 Behavior of the maize transposable element Ac in Arabidopsis thaliana. Plant J. 2: 69–81.

Emelyanov, A., Y. Gao, N. I. Naqvi, and S. Parinov, 2006 Trans-kingdom transposition of the maize Dissociation element. Genetics 174: 1095–1104.

Enoki, H., T. Izawa, M. Kawahara, M. Komatsu, S. Koh *et al.*, 1999 Ac as a tool for the functional genomics of rice. Plant J. 19: 605–613.

Essers, L., R. H. Adolphs, and R. Kunze, 2000 A highly conserved domain of the maize *Activator* transposase is involved in dimerization. Plant Cell 12: 211–224.

Froschauer, A., D. Sprott, F. Gerwien, Y. Henker, F. Rudolph *et al.*, 2012 Effective generation of transgenic reporter and gene trap lines of the medaka (Oryzias latipes) using the Ac/Ds transposon system. Transgenic Res. 21: 149–162.

Gangadharan, S., L. Mularoni, J. Fain-Thornton, S. J. Wheelan, and N. L. Craig, 2010 DNA transposon Hermes inserts into DNA in

nucleosome-free regions in vivo. Proc. Natl. Acad. Sci. USA 107: 21966–21972.
Goryshin, I. Y., and W. S. Reznikoff, 1998 Tn5 in vitro transposition. J. Biol. Chem. 273: 7367–7374.
Haren, L., B. Ton-Hoang, and M. Chandler, 1999 Integrating DNA: transposases and retroviral integrases. Annu. Rev. Microbiol. 53: 245–281.
Heinlein, M., T. Brattig, and R. Kunze, 1994 In vivo aggregation of maize Activator (Ac) transposase in nuclei of maize endosperm and Petunia protoplasts. Plant J. 5: 705–714.
Huefner, N. D., Y. Mizuno, C. F. Weil, I. Korf, and A. B. Britt, 2011 Breadth by depth: expanding our understanding of the repair of transposon-induced DNA double strand breaks via deep-sequencing. DNA Repair (Amst.) 10: 1023–1033.
Jones, J. D. G., F. Carland, L. Harper, E. Lim, and H. Dooner, 1990 Genetic Properties of the Maize Transposon *Activator* (*Ac*) in Tobacco, pp. 59–64 in *Plant Gene Transfer - UCLA Symposia on Molecular and Cellular Biology, New Series,* Vol. 129, edited by C. J. Lamb and R. N. Beachy. Wiley-Liss Inc., New York.
Kasaras, A., and R. Kunze, 2010 Expression, localisation and phylogeny of a novel family of plant-specific membrane proteins. Plant Biol (Stuttg) 12(Suppl. 1): 140–152.
Kempken, F., and F. Windhofer, 2001 The hAT family: a versatile transposon group common to plants, fungi, animals, and man. Chromosoma 110: 1–9.
Keravala, A., D. Liu, E. R. Lechman, D. Wolfe, J. A. Nash *et al.*, 2006 Hyperactive Himar1 transposase mediates transposition in cell culture and enhances gene expression in vivo. Hum. Gene Ther. 17: 1006–1018.
Koprek, T., D. McElroy, J. Louwerse, R. Williams-Carrier, and P. G. Lemaux, 2000 An efficient method for dispersing Ds elements in the barley genome as a tool for determining gene function. Plant J. 24: 253–263.
Kulkosky, J., K. S. Jones, R. A. Katz, J. P. G. Mack, and A. M. Skalka, 1992 Residues critical for retrovirus integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Mol. Cell. Biol. 12: 2331–2338.
Kunze, R., and C. F. Weil, 2002 The hAT and CACTA superfamilies of plant transposons, pp. 565–610 in *Mobile DNA II*, edited by N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz. ASM Press, Washington, D.C.
Kunze, R., P. Starlinger, and D. Schwartz, 1988 DNA methylation of the maize transposable element *Ac* interferes with its transcription. Mol. Gen. Genet. 214: 325–327.
Kunze, R., U. Behrens, U. Courage-Franzkowiak, S. Feldmar, S. Kühn *et al.*, 1993 Dominant transposition-deficient mutants of maize *Activator* (*Ac*) transposase. Proc. Natl. Acad. Sci. USA 90: 7094–7098.
Kunze, R., H. Saedler, and W.-E. Lönnig, 1997 Plant transposable elements, pp. 331–470 in *Advances in Botanical Research*, Vol. 27, edited by J. A. Callow. Academic Press, San Diego.
Lampe, D. J., B. J. Akerley, E. J. Rubin, J. J. Mekalanos, and H. M. Robertson, 1999 Hyperactive transposase mutants of the Himar1 mariner transposon. Proc. Natl. Acad. Sci. USA 96: 11428–11433.
Lazarow, K., and S. Lütticke, 2009 An Ac/Ds-mediated gene trap system for functional genomics in barley. BMC Genomics 10: 55.
Mates, L., M. K. Chuah, E. Belay, B. Jerchow, N. Manoj *et al.*, 2009 Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. Nat. Genet. 41: 753–761.
McClintock, B., 1951 Chromosome organization and genic expression. Cold Spring Harb. Symp. Quant. Biol. 16: 13–47.
Mitra, R., J. Fain-Thornton, and N. L. Craig, 2008 piggyBac can bypass DNA synthesis during cut and paste transposition. EMBO J. 27: 1097–1109.
Olson, W. K., A. A. Gorin, X. J. Lu, L. M. Hock, and V. B. Zhurkin, 1998 DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. Proc. Natl. Acad. Sci. USA 95: 11163–11168.
Ros, F., and R. Kunze, 2001 Regulation of Activator/Dissociation transposition by replication and DNA methylation. Genetics 157: 1723–1733.
Schmidt, R., and L. Willmitzer, 1989 The maize autonomous element Activator (Ac) shows a minimal germinal excision frequency of 0.2%–0.5% in transgenic Arabidopsis thaliana plants. Mol. Gen. Genet. 220: 17–24.
Uren, A. G., H. Mikkers, J. Kool, L. van der Weyden, A. H. Lund *et al.*, 2009 A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. Nat. Protoc. 4: 789–798.
Vollbrecht, E., J. Duvick, J. P. Schares, K. R. Ahern, P. Deewatthanawong *et al.*, 2010 Genome-wide distribution of transposed Dissociation elements in maize. Plant Cell 22: 1667–1685.
Wang, L., and R. Kunze, 1998 Transposase binding site methylation in the epigenetically inactivated Ac derivative Ds-cy. Plant J. 13: 577–582.
Wang, L., M. Heinlein, and R. Kunze, 1996 Methylation pattern of Activator (Ac) transposase binding sites in maize endosperm. Plant Cell 8: 747–758.
Weil, C. F., and R. Kunze, 2000 Transposition of maize Ac/Ds transposable elements in the yeast Saccharomyces cerevisiae. Nat. Genet. 26: 187–190.
Weinreich, M. D., A. Gasch, and W. S. Reznikoff, 1994 Evidence that the *cis* preference of the Tn*5* transposase is caused by nonproductive multimerization. Genes Dev. 8: 2363–2374.
Xiao, Y. L., X. Li, and T. Peterson, 2000 Ac insertion site affects the frequency of transposon-induced homologous recombination at the maize p1 locus. Genetics 156: 2007–2017.
Yu, J., K. Marshall, M. Yamaguchi, J. E. Haber, and C. F. Weil, 2004 Microhomology-dependent end joining and repair of transposon-induced DNA hairpins by host factors in Saccharomyces cerevisiae. Mol. Cell. Biol. 24: 1351–1364.
Yuan, Y. W., and S. R. Wessler, 2011 The catalytic domain of all eukaryotic cut-and-paste transposase superfamilies. Proc. Natl. Acad. Sci. USA 108: 7884–7889.
Yusa, K., L. Zhou, M. A. Li, A. Bradley, and N. L. Craig, 2011 A hyperactive piggyBac transposase for mammalian applications. Proc. Natl. Acad. Sci. USA 108: 1531–1536.
Zhou, L., R. Mitra, P. W. Atkinson, A. B. Hickman, F. Dyda *et al.*, 2004 Transposition of hAT elements links transposable elements and V(D)J recombination. Nature 432: 995–1001.

Communicating editor: S. E. Bickel

GENETICS

Supporting Information
http://www.genetics.org/content/suppl/2012/05/04/genetics.112.139642.DC1

A Hyperactive Transposase of the Maize Transposable Element *Activator* (*Ac*)

Katina Lazarow, My-Linh Du, Ruth Weimer, and Reinhard Kunze

DOI: 10.1534/genetics.112.139642

Table S1 ***Ds*** **insertion sites in yeast**

Ade$^+$rev	Chr	Position	*Ds* target site[a]	Accession	E-value[d]	Gene
***Ac*TPase$_{4x}$**						
1	XIII	311423	TTGGCCAA	NM_001182514.1	3e -19	S288c Putative protein with similarity to human PEX5Rp
2	XII	458235	CAAAGAGT	FN554374.1	8e -144	Partial 5S rRNA gene
3	XI	264064	CCCCAGAT	X71133.1	1e -75	HAP4, AAT1, GFA1, LAP4 and MBR1 genes
4	XV	323245	**GCACTTTC**	Z74744.1	0.0	Chromosome XV reading frame ORF YOL002c
5	XII	693673	**GTTTCATC**	U09242.1	0.0	DNA replication protein Cdc46p (CDC46) gene
6	VII	440638	CTTATTTT	Z72551.1	3e -32	Chromosome VII reading frame ORF YGL029w
7	II	792133	**AAAATGTT**	AB200246.1	0.0	BIO6 gene for biotin biosynthesis enzyme
8	XI	478727	**GACTCTTT**	NM_001179810.1	7e -112	S288c Component of GARP (Golgi associated retrograde protein) complex
9	VII	465531	**GAGAATAG**	S58126.1	7e -51	LEU1-ATE1 loci: LEU1...YGL029
10	IV	302350	AAAAAAGG	Z74134.1	1e -105	Chromosome IV reading frame ORF YDL086w
11	XV	637133[e]	GCCCTAAT	NM_001183580.1	7e -170	S288c Pns1p (PNS1)
12	XVI	688009	TAGAAAAT	NM_001184168.1	2e -85	S288c Putative membrane protein
13	XV	552278[e]	**AGTTTTCT**	M23369.1	0.0	Profilin (PFY) gene
14	XII	390865	**AAAATACG**	Z73293.1	0.0	Chromosome XII reading frame ORF YLR121c
15	IV	239896	**GTGGCCAA**	X56956.1	4e -80	CDC48 gene for cell cycle protein CDC48p
16	XV	525094[e]	ATCCGAAT	NM_001183528.1	0.0	S288c Inp53p (INP53)
17	IX	388640	GTACAGAA	BK006942.2	8e-34	S288c chromosome IX complete sequence
18	XV	170355	CCCCTTGC	M83139.1	1e -101	RAS1 and RAS2 attenuator (IRA2) gene, 5' end
19	XII	457081	**GCCAACCG**	BK006945.1	0.0	Isolate OC21 18S ribosomal RNA gene
20	II	466515	GTCCACGT	X66247.1	5e -146	YSA1, SSN6, RAD16, and LYS2 genes
21	XV	226759	TTCATATG	Z74797.1	5e -146	Chromosome XV reading frame ORF YOL055c
22	V	41348	CATTTGGG	NM_001178875.1	0.0	S288c Prb1p (PRB1), mRNA, (Protease B)
23	X	106834	**ATTTATTT**	U15300.1	7e -22	Hal5p (HAL5) mRNA
24	XVI	501953	CTGCCGGC	U30613.1	0.0	Suppressor kinase of snf3, Sks1p (SKS1) gene
25	XV	587811[e]	**GGCATGGT**	NM_001183559.1	0.0	S288c Sfl1p (SFL1)
26	II	10413	TTTCTAAG	NM_001178347.1	199	S288c Putative protein of unknown function
27	VII	871291	GAGCTCAA	U27358.1	6e -130	Hgh1p (HGH1) gene
28	VII	73860	TTAATTAC	Z72748.1	0.0	Chromosome VII reading frame ORF YGL226w
29	XII	790299	AATTCATT	Z49198.1	2e -72	CHS5 (CAL3) gene
30	IV	352297	TGAATGTC	NM_001180116.1	1e -06	S288c Putative protein of unknown function
31	VII	84707	GGTAAAAT	Z72741.1	3e -81	Chromosome VII reading frame ORF YGL219c
32	X	78731	**GCCTCGTA**	NM_001181620.1	4e -59	S288c Protein kinase
33	VII	514478	TAATTACT	Z72797.1	7e -72	Chromosome VII reading frame ORF YGR012w
34	XVI	901730	**GGCAGAAT**	J04184.1	1e -63	Dolichol phosphate mannose synthase (DPM1) gene
35	-[b]	-	**TATGAATG**	NM_001184395.1	1e -121	S288c Retrotransposon TYA Gag gene
***Ac*TPase**						
1	VI	204983	**CACACGCC**	U09479.1	4e -147	Histidinolphosphatase (HIS2) gene
2	XV	256401	ATTTCCAC	Z74779.1	0.0	Chromosome XV reading frame ORF
3	X	210344	GAGACGAG	NM_001181543.1	3e -26	S288c GATA zinc finger protein and Dal80p homolog
4	XII	1078175	GGTACAGA	NM_001182321.1	8e -50	S288c Calcineurin A
5	XVI	230894	**ACTCTAAG**	Z73525.1	4e -42	CXVI reading frame ORF YPL169c
6	XII	611863	CATGGTTT	M24939.1	0.0	DNA topoisomerase (TOP3) gene
7	VIII	167614	**TGTATATA**	X15484.1	1e -14	Dipeptidyl aminopeptidase B (DPAP B)
8	VII	842662	GTTTGAAA	X14629.1	5e -24	MSM1 gene for mitochondrial methionyl-tRNA synthase
9	XII	445181	**AGGATATT**	S79456.1	7e -93	tRNA-Gln, ACS2=acetyl-coenzyme A synthetase 2
10	XIII	287458	AAAGAAAC	M33270.1	1e -94	High affinity hexose transporter-2 (HXT2) gene
11	II	410391	TCTAGCGC	J03724.1	0.0	Mitochondrial C-1-Tetrahydrofolate synthase gene (MIS1)
12	IV	778863	**GCTTTACG**	NM_001180467.1	1e -41	S288c SPS plasma membrane sensor system (Ssy1p-Ptr3p-Ssy5p)
13	VII	503923	AATATGGG	NM_001181133.1	0.0	S288c Peroxisomal integral membrane protein

14	XV	369772	**CGACCCAT**	NM_001183438.1	9e -92	S288c Protein of unknown function
15	XI	354898	TACATATC	X71621.1	0.0	Genes ELM1 and PRI2
16	VII	757585	**AATGTGTC**	X65470.1	7e -40	PAS2 gene
17	XV	586815[e]	GCCACAGT	Z75047.1	9e -87	Chromosome XV reading frame ORF YOR139c
18	-[c]	-	TGAACGTT	M18354.1	2e -35	Sigma repetitive element (TY3)

[a] *Ds* insertion sites where 5' and 3' flanking sequences were isolated are bold typed.

[b] In case of insertion into TYA retrotransposon the exact chromosomal localization is not possible (XVI, XV,XII, XVII, V, IV or I).

[c] In case of insertion into yeast sigma repetitive element (TY3) the exact chromosomal localization is not possible (XIV, XI, IX, VII or I).

[d] Sequences were BLASTed against GenBank (All GenBank+EMBL+DDBJ+PDB sequences but no EST, STS, GSS,environmental samples or phase 0, 1 or 2 HTGS sequences)

[e] Short range transposition; *Ds* reintegration occurred on the same chromosome in a maximal distance of 72 kb from the donor site.

Table S2 Distribution of *Ds* insertions on yeast chromosomes

Yeast chromosome			*Ds* insertions					
		% of	*Ac*TPase$_{4x}$		*Ac*TPase		Total	
No.	Size in kb	genome	Observed	Expected[b]	Observed	Expected[b]	Observed	Expected[b]
I	230	1.9	-	0.6	-	0.3	-	1
II	813	6.7	3	2.3	1	1.1	4	3
III	315	2.6	-	0.9	-	0.4	-	1
IV	1522	12.6	3	4.3	1	2.1	4	6
V	575	4.8	1	1.6	-	0.8	1	2
VI	270	2.2	-	0.8	1	0.4	1	1
VII	1091	9.0	6	3.1	3	1.5	9	5
VIII	562	4.7	-	1.6	1	0.8	1	3
IX	440	3.7	1	1.2	-	0.6	1	2
X	745	6.2	2	2.1	1	1.0	3	3
XI	666	5.5	2	1.9	1	0.9	3	3
XII	1078	8.9	5	3.0	3	1.5	8	5
XIII	924	7.7	1	2.6	1	1.3	2	4
XIV	784	6.5	-	2.2	-	1.1	-	3
XV[a]	1091	9.1	7	3.1	3	1.5	10	5
XVI	948	7.9	3	1.3	1	1.3	3	4
Σ	12057	100	34	34	17	17	51	51

[a] Yeast chromosome with *Ds* launch pad in *ade2* locus.
[b] Number of expected insertions was calculated based on the size of each chromosome.

Table S3 Primer sequences (5' to 3')

Primers for site directed mutagenesis	
E249A	CAATATTGTAGAACATGCGTACTTTGTTGAG CTCAACAAAGTACGCATGTTCTACAATATTG
D301A	CTCGCTTCAGTACAACTATGGCTATGTGGACA ACTGAACATCTTTTAGTTTTCCATACAACTTTTCTTTTTC
E336A	GCTTTTTTCATGTTGCAGGGCGCCACACTGG CCAGTGTGGCGCCCTGCAACATGAAAAAAGC
D367A	TGCCTTGTCTTTGGCTAATGCTAGTGC AACAATTTTTTCTCAATGTTCCACTTAACCA
D459A	GGGATCTCATATGCTGTCTCAACTAGATGG CCATCTAGTTGAGACAGCATATGAGATCCC
D545A	GGATTTGATTGCCCAATGGTGTGTTC GAACACACCATTGGGCAATCAAATCC
E719A	CTGTTGCTTCTGCGTCTGCGTTCAGTGC GCACTGAACGCAGACGCAGAAGCAACAG
Splinkerette-PCR gene specific primers	
primary PCR *Ds* 5' end	GGATAAAACTAACAAAATCGGTTATACG
primary PCR *Ds* 3' end	CCGTATTTATCCCGTTCGTTTTCGTTA
secondary PCR *Ds* 5' end	CGATAACGGTCGGTACGGGATTTTC
secondary PCR *Ds* 3' end	CCCGTTTTCGTTTCCGTCCCGCAAG
Primers for amplification of empty *Ds* donor sites in *Arabidopsis*	
P1	GGAAGTTCATTTCATTTGGAG
P2	TCGCCCTCGCCCTCGCCGGAC
P3	GATCTCGGTGACGGGCAGGAC