Maize Transposable Elements *Ac/Ds* as Insertion Mutagenesis Tools in *Candida albicans*

Kevin Mielich,*,[1] Ella Shtifman-Segal,†,[1] Julia C. Golz,* Guisheng Zeng,‡ Yue Wang,‡ Judith Berman,† and Reinhard Kunze*,[2]

*Institute of Biology, Dahlem Centre of Plant Sciences, Free University pf Berlin, 14195, Germany, †Department of Molecular Microbiology and Biotechnology, George Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv 69978, Israel, and ‡Institute of Molecular and Cell Biology, Agency for Science, Technology and Research, 138673, Singapore

ORCID IDs: 0000-0002-1211-2822 (Y.W.); 0000-0002-8577-0084 (J.B.); 0000-0002-3304-5550 (R.K.)

ABSTRACT In nonmodel systems, genetic research is often limited by the lack of techniques for the generation and identification of gene mutations. One approach to overcome this bottleneck is the application of transposons for gene tagging. We have established a two-element transposon tagging system, based on the transposable elements *Activator* (*Ac*)/*Dissociation* (*Ds*) from maize, for *in vivo* insertion mutagenesis in the fungal human pathogen *Candida albicans*. A nonautonomous *Ds* transposon carrying a selectable marker was constructed into the *ADE2* promoter on chromosome 3 and a codon usage-adapted *Ac* transposase gene was inserted into the neutral *NEUT5L* locus on chromosome 5. In *C. albicans* cells expressing the transposase, the *Ds* element efficiently excised and reintegrated elsewhere in the genome, which makes the *Ac/Ds* transposons promising tools for saturating insertion mutagenesis in clinical strains of *C. albicans*.

The autonomous transposable element *Activator* (*Ac*) from *Zea mays*, originally discovered as the first transposon by McClintock (1951), is one of three prototypic elements of the hAT superfamily of "cut-and-paste" transposons. Soon after its discovery, the potential of *Ac* and its nonautonomous derivatives, termed *Dissociation* (*Ds*), for gene tagging and insertion mutagenesis in maize and heterologous plants was recognized (Walbot 1992). *Ac/Ds* elements are convenient to use in heterologous plants because the single *Ac*-encoded transposase protein (TPase) is sufficient for transposition, the transposon ends with the essential TPase-binding sites are only ∼240 bp in length, and the nonautonomous *Ds* elements can carry > 15 kb of cargo DNA. *Ac/Ds* has been used in > 20 higher plants, among them model plants and agro-

nomically important species [reviewed in Kunze and Weil (2002)]. *Ac/Ds* elements tend to transpose into genetically linked sites [Dooner and Belachew 1989; Vollbrecht *et al.* 2010; reviewed in Lazarow *et al.* (2013)], but they do not have the target sequence preference exhibited by other transposons (Spradling *et al.* 2011). More recently, the focus of insertion mutagenesis approaches has shifted from single-gene tagging toward the generation of large insertion mutant populations. The use of transposons for this purpose is highly attractive for the study of plants because transposons enable the formation of a large number of insertion mutations in plant species where ballistic or *Agrobacterium tumefaciens*-based Transfer DNA (T-DNA) transformation is inefficient [reviewed in Lazarow *et al.* (2013)]. *Ac/Ds* transposons also function in vertebrate animals (Emelyanov *et al.* 2006; Quach *et al.* 2015; Vrljicak *et al.* 2016) as well as in the budding yeast *Saccharomyces cerevisiae* (Weil and Kunze 2000). Recently, *Ac/Ds* was employed for saturated transposon mutagenesis in *S. cerevisiae*, facilitating the identification of conditionally essential genes and identifying important functional protein domains (Michel *et al.* 2017). An important advantage is the facility of generating libraries of insertion mutations in mutant strain backgrounds. A similar approach, using the hAT element Hermes from the house fly (*Musca domestica*), achieved saturated transposon insertion mutagenesis in the fission yeast *Schizosaccharomyces pombe* (Guo *et al.* 2013).

Transposable elements in eukaryotic genomes are mostly inactive and transpose only rarely. This relative stability of transposon insertions

is due to epigenetic silencing, the low activity of the TPase, as well as post-translational negative autoregulation, such that increased TPase levels result in reduced transposition frequency [reviewed by Bire *et al.* (2013)]. This phenomenon, termed TPase "overexpression inhibition," presumably underlies the "inverse dose effect" of *Ac/Ds* first discovered by McClintock (1951) and later also reported in heterologous plants (Scofield *et al.* 1993, Heinlein *et al.* 1994). The naturally low activity of transposons limits the efficiency of transposon mutagenesis experiments; hyperactive TPase variants that catalyze more frequent transposition reactions have increased the efficiency of transposon mutagenesis with several prokaryotic and eukaryotic transposons (Goryshin and Reznikoff 1998; Lampe *et al.* 1999; Beall *et al.* 2002; Zayed *et al.* 2004; Mátés *et al.* 2009; Yusa *et al.* 2011; Lazarow *et al.* 2012). For example, *Ac*TPase4x, a hyperactive *Ac* TPase, catalyzes a higher *Ds* excision frequency compared to the wild-type *Ac*TPase [100-fold higher in budding yeast and sixfold higher in *Arabidopsis thaliana* (Lazarow *et al.* 2012)].

Here, we asked if the maize *Ac/Ds* transposon system could be applied to nonmodel micro-organisms more distantly related to the model budding and fission yeasts. The commensal ascomycete *Candida albicans* is the most common fungal pathogen of humans (Kim and Sudbery 2011) and differs in several genetic and morphological features from *S. cerevisiae*. In addition, *C. albicans* diverges from the universal genetic code at one codon: CTG encodes serine instead of leucine, which complicates the design of heterologous molecular genetic systems. Furthermore, *C. albicans* was long thought to be an obligate diploid with the ability to undergo a "parasexual cycle" between generations without conventional meiosis from diploid to tetraploid and back [reviewed by Noble and Johnson (2007)]. However, mating-competent haploid cells were discovered recently (Hickman *et al.* 2013). Haploids appear to arise via a concerted chromosome loss mechanism rather than meiosis. The initial isolates grew slower than diploids and were relatively unstable, reverting to "autodiploids" with high frequency (Hickman *et al.* 2013). Here, we describe a haploid strain selected for its faster growth rate and use it to test the design of an inducible two-component *Ac/Ds* transposon system for *C. albicans*, consisting of a codon-adapted hyperactive TPase gene and a nonautonomous *Ds* element carrying a selection marker. We demonstrate that the TPase mobilizes the *Ds* element and that the *Ds* excision footprint sequences exhibit subtle differences compared to those seen in *S. cerevisiae* and plants.

MATERIALS AND METHODS

Construction of *Ac/Ds* components
The codon-adapted *Ac*TPase4xCa open reading frame (Supplemental Material, Figure S1 in File S1) was constructed by fusing three *de novo*-synthesized DNA fragments (Thermo Fisher Scientific GENEART GmbH, Regensburg, Germany) and cloned into pJET1.2 via the CloneJET kit (Thermo Fisher Scientific). The *CaMAL2* promoter (Backen *et al.* 2000), amplified from pKB2019 with primers oMAL2-F and oMAL2-R, and the *CaADH1* terminator, amplified from pMG2120 (Gerami-Nejad *et al.* 2012) with primers oADHT-F and oADHT-R, were also ligated into pJET1.2. The *Ac*TPase4xCa expression cassette was assembled by fusing the promoter, *Ac*TPase4xCa-coding region, and terminator in pJET1.2. From this plasmid, the *Ac*TPase4xCa expression cassette was excised with *Eco*91I and ligated into *Eco*91I-linearized pDUP5 (Gerami-Nejad *et al.* 2013), yielding pKM300 (Figure 1A, top line).

To construct a nonautonomous *miniDs* element with a selectable marker inserted into the 5′-UTR of the *C. albicans* ADE2 gene, the *miniDs* from pBBDsXho (Laufs *et al.* 1990) was excised with *Sal*I and cloned into pUC19. The resulting plasmid was pRK86. The nourseothricin acetyl

N-transferase gene *NAT1*, under the control of the *TEF* promoter, was amplified from pDUP3 (Gerami-Nejad *et al.* 2013) with the primers oTEFmut-F and oTEFmut-R, and the PCR product was digested with *Sal*I and ligated into *Xho*I-linearized, dephosphorylated pRK86 to form pRK86-NAT1, carrying the transposable element *Ds-NAT1*. A 528-bp segment of *C. albicans* chromosome 3, centered around an artificial *Sal*I site 24-bp upstream of the *ADE2* start codon, was *de novo* synthesized (Thermo Fisher Scientific) and ligated into pJET1.2. The resulting plasmid, pRK401, was linearized with *Sal*I and dephosphorylated with FastAP (Thermo Fisher Scientific). *Ds-NAT1* was excised with *Sal*I from pRK86-NAT1 and ligated into *Sal*I-linearized pRK401, followed by the exchange of two nucleotides flanking the *Ds-NAT1* 5′-end using primers oKM34 and oKM35, and the Q5 site-directed mutagenesis kit (New England Biolabs, Beverly, MA). The resulting plasmid pRK403a was used for chromosomal integration via homologous recombination of *Ds-NAT1* into the 5′-UTR of the *C. albicans* ADE2 gene (Figure 1B, top line).

Generation of the *MTLa* haploid *Candida albicans* strain GZY896
To isolate a *MTLa* haploid strain from YJB12881 (Hickman *et al.* 2013), single colonies were subjected to repeated rounds of subcultivation in glucose minimal medium supplemented with 40 μg ml^{-1} histidine (GMM + His) medium (Zeng *et al.* 2014). Haploid colonies were identified by flow cytometry analysis and then stored at −80°. Strains were revived on GMM + His plates and several more subcultivation cycles were performed, yielding a haploid strain, GZY892, that appeared stable by flow cytometry. A *ura3*Δ derivative of the strain was constructed by excising the *ura3*Δ::*HIS4* cassette from plasmid pYGS1023 (Hickman *et al.* 2013) and transforming it into GZY892, according to the method of Zeng *et al.* (2014), where it inserted by homologous recombination at the *URA3* locus. The resulting strain was termed GZY896 (*MTLa ura3*Δ::*imm434 his4 gal1*Δ:: *ura3*Δ::*HIS4*).

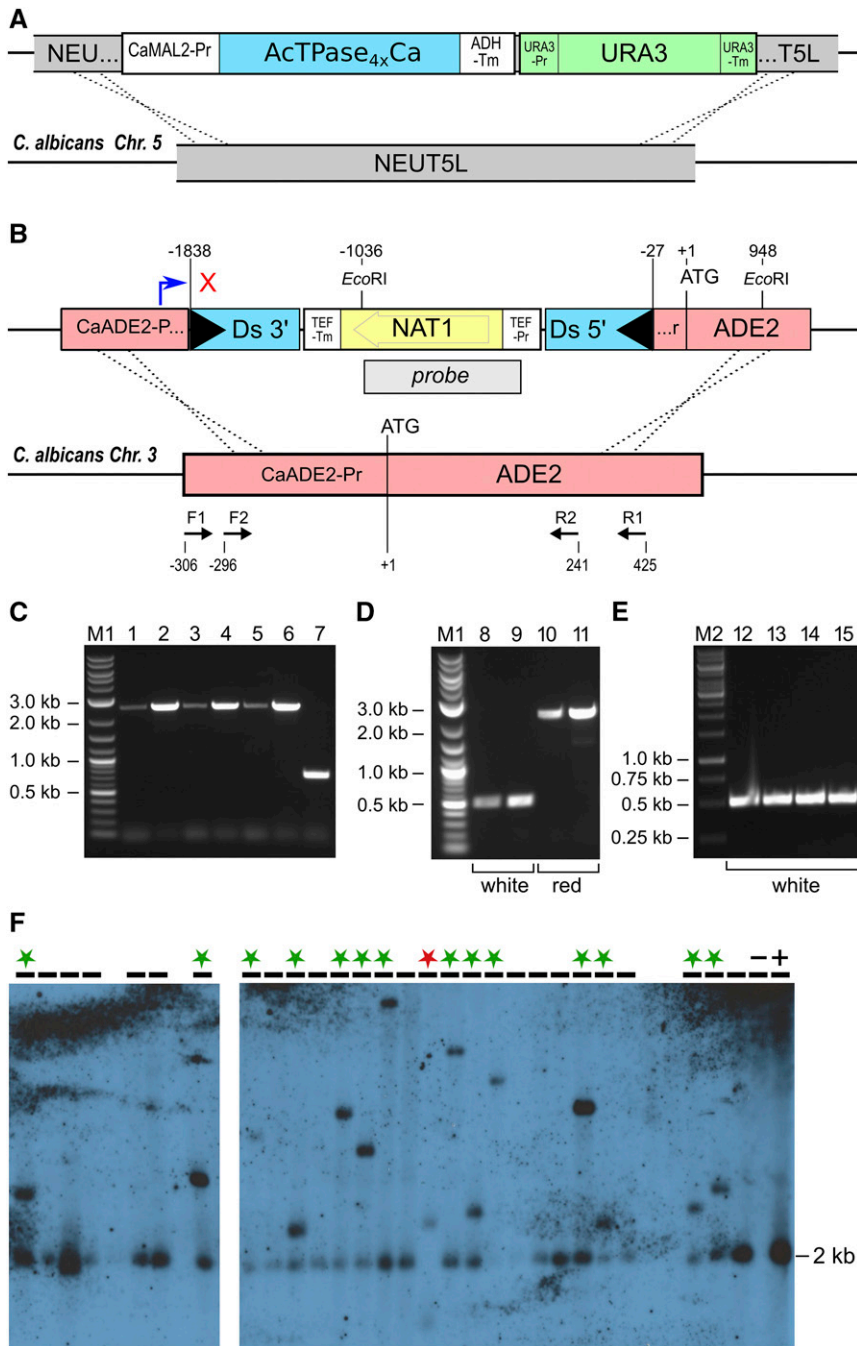### Generation of *Ac/Ds* transposable element-carrying *Candida albicans* strains KMY100 and KMY103G1
GZY896 transformation was performed according to the method of Zeng *et al.* (2014). First, linearized pKM300 was used to integrate the *Ac*TPase4xCa expression cassette via homologous recombination into the *NEUT5L* locus (Gerami-Nejad *et al.* 2013) (Figure 1A). Transgenic cells were selected on complete synthetic defined minimal (SDC) plates lacking uridine. In a second step, these transformants and, in parallel, GZY896 cells were transformed with linearized pRK403a to integrate the *Ds-NAT1* into the 5′-UTR of *ADE2* on chromosome 3 (Figure 1B). Transformants were selected by plating the cells on yeast extract peptone (YP) medium + adenine + glucose with 0.4 mg/ml nourseothricin. The resulting strains carrying the *ade2*::*Ds-NAT1* construct alone, or in combination with the *Ac*TPase4xCa expression construct, were termed KMY100 and KMY103G1, respectively.

Isolation of *Ds* transposants
KMY103G1 cells were grown in SDC + 3% maltose liquid medium at 30° for 24 hr, and then plated on SDC − adenine + glucose plates to select for *ADE2* revertants. To select for reintegration of *Ds-NAT1* elements, *ADE2* revertant colonies were replica-plated on YP + adenine + glucose plates with 0.4 mg/ml nourseothricin and grown for 2 days at 30°.

Flow cytometry
Flow cytometry was performed as described previously (Hickman *et al.* 2013) using a MACSQuant FLOW cytometer (Miltenyi Biotec GmbH, Bergisch Gladbach, Germany).

Figure 1 *Ac/Ds* (*Activator/Dissociation*) transposition in *C. albicans*. (A) The transposase expression construct consists of the *AcTPase4xCa ORF* (blue) inserted between the *CaMAL2* promoter and the *CaADH1* terminator, flanked by the *CaURA3* gene (green), and inserted into the *NEUT5L* locus for homologous recombination into that locus on *C. albicans* chromosome 5. (B) The *ade2*::*Ds-NAT1* cassette consists of a nonautonomous mini*Ds* transposon (blue) that carries a *NAT1* selectable marker gene (yellow). The transposon is inserted into the 5'-UTR of the *CaADE2* gene (red). This cassette was inserted by homologous recombination into the *ADE2* gene on chromosome 3. The blue bent arrow indicates the transcriptional start site of the *CaADE2* promoter and the red cross symbolizes transcriptional interruption by the *Ds-NAT1* transposon. TEF-Pr, *TEF* promoter; TEF-Tm, *TEF* terminator; Black triangles, *Ds* terminal inverted repeats; Probe, hybridization probe used for DNA gel blot analysis shown in (F); F1/2 and R1/2, *CaADE2* primers used for PCR analyses shown in (C–E). (C) PCR products obtained with primers F1 and R1 on genomic DNA from KMY100 colonies after four successive cell passages (lane 1, passage 0; lane 2, passage 1; lane 3 passage 2; lane 4, passage 3; and lanes 5–6, passage 4) and from the progenitor GZY896 strain (lane 7). (D) PCR products obtained with primers F2 and R2 on genomic DNA from white (lanes 8–9) and red (lanes 10–11) KMY103G1 colonies grown on inducing SDC (synthetic defined minimal) plates (+ adenine + maltose). (E) PCR products obtained with primers F2 and R2 on genomic DNA from white KMY103G1 colonies (lanes 12–15) grown on inducing SDC plates (+ adenine and + maltose). These PCR products were cloned and sequenced for *Ds* excision footprint analyses. M1, New England Biolabs Quick-Load 2-Log DNA Ladder and M2, Thermo Fisher Scientific GeneRuler 1 kb DNA Ladder. (F) DNA gel blot analysis of *Ds-NAT1* element distribution in *ADE2* revertant clones. Genomic DNA from KMY103G1 *ADE2* revertant colonies was *Eco*RI-digested, size-fractionated, blotted, and hybridized with a *NAT1*-probe. Samples in the indicated lanes were evaluable. Lanes marked with green asterisks: an additional band hybridized, indicating *Ds-NAT1* excision and reintegration. Unmarked lanes: only the 2-kb fragment of the *ade2*::*Ds-NAT1* allele hybridized. Lane marked with red asterisk: two novel bands indicate excision and reintegration of both *Ds-NAT1* elements. Lane marked "−": GZY896 DNA (negative control). Lane marked "+": KMY100 DNA (positive control).

Ds excision and reintegration site analysis

ADE2 revertant colonies were picked, resuspended in 10 μl H$_2$O containing 2.5 U Zymolyase (Zymo Research, Irvine, CA) and incubated for 15 min at 30°. Next, 2 μl of a 1:10 dilution of these suspensions was used as template for PCR amplification of *Ds* excision sites with *CaADE2* primers F2 and R2 (Table S1 in File S1) and Q5 DNA Polymerase (New England Biolabs). PCR products were purified on NucleoSpin Gel and PCR Clean-up columns (MACHEREY-NAGEL GmbH, Düren, Germany) and sequenced.

Reintegration sites of *Ds-NAT1* elements were mapped by modified FPNI-PCR (Fusion Primer and Nested Integrated PCR; Wang *et al.* 2011) on *ADE2*+-Nat^R cells to amplify genomic DNA flanking the *Ds*

3′-end. Target site duplications at three *Ds* reinsertion sites were identified by amplification of novel *Ds-NAT1* insertion sites with pairs of *Ds*-flanking primers (Table S1 in File S1) and sequencing of the PCR products.

DNA gel blot analysis

Genomic DNAs were isolated from 2 ml cultures of *ADE2* revertants grown in SDC + 2% glucose liquid medium at 30° for 24 hr with the Yeast DNA Preparation Kit (Jena Bioscience GmbH, Jena, Germany). The quality of the isolated DNAs was analyzed on a 1% agarose gel. Next, 200 ng DNA of each sample were digested with 10 U of *Eco*RI (Thermo Fisher Scientific) for 4 hr at 37° and size fractionated on a 0.9% agarose gel overnight at 2 V/cm. The gel was successively soaked for 20 min in 0.2 M HCl, 20 min in 0.5 M NaOH/1.5 M NaCl, and 30 min in 1 M Tris-Cl/1.5 M NaCl pH 7.4. Following overnight capillary transfer of the DNA to a Hybond NX membrane (GE Healthcare, Little Chalfont, UK) the DNA was UV cross-linked (120 J at 254 nm). The membrane was successively prehybridized, hybridized with a Digoxigenin (DIG)-labeled *NAT1* probe, washed, and detected with anti-DIG-AP, Fab fragments, and CDP-*Star* (Roche Applied Science, Penzberg, Germany) following the manufacturer's protocols. The DIG-labeled *NAT1*-probe was amplified with primers oKM222 and oKM223 using pRK403a as template.

Data availability

Plasmids and strains are available upon request.

RESULTS

An *Ac/Ds* two-component transposon system in *C. albicans*

C. albicans belongs to the CUG clade of pathogenic yeasts, in which the trinucleotide CTG specifies serine instead of the commonly encoded leucine (Santos and Tuite 1995). Furthermore, *C. albicans* does not have naturally occurring or engineered plasmids that are maintained autonomously. Therefore, the *Ac/Ds* transposon components previously constructed for *S. cerevisiae* (Lazarow *et al.* 2012) were not suitable for application in *C. albicans* and a new *Ac/Ds* two-component system for *C. albicans* was required. The system includes a nonautonomous *Ds* element carrying the nourseothricin N-acetyl transferase (*NAT1*) gene as a selectable marker gene, *Ds-NAT1*, and a codon-adapted ("candidized") hyperactive TPase-coding sequence (*Ac*TPase4xCa) under control of the inducible *CaMAL2* promoter. The original *Ac*TPase4x-coding region with an N-terminally-fused SV40 nuclear localization sequence (NLS) was candidized by replacing two CTG codons with TTG to ensure translation into leucine and by exchanging all codons that are rarely ($\leq$ 11%) used in *C. albicans* with more frequently used codons to ensure efficient translation (Figure S1 in File S1). The *Ac*TPase4xCa expression cassette was inserted into an intergenic region on chromosome 5 (*NEUT5L*; Gerami-Nejad *et al.* 2013) of the haploid *C. albicans* strain GZY896 (Figure 1A). The *Ds-NAT1* element was cloned into the 5′-UTR of *ADE2*, 26-bp upstream of the *ADE2* start codon on chromosome 3 (Figure 1B), and recombined into GZY896 with or without the *Ac*TPase4xCa gene, respectively. The resulting strains KMY100 (*ade2::Ds-NAT1*) and KMY103G1 (*Ac*TPase4xCa::*NEUT5L ade2::Ds-NAT1*) grow a little slower than *ADE2* cells and form red colonies on SDC medium.

*Ac*TPase4xCa expression induces *Ds-NAT1* excision

The stability of the *Ds-NAT1* insertion in the *ADE2* 5′-UTR was examined by propagating KMY100 cells for four passages on YP + adenine + glucose medium. After each passage, we plated three colonies and observed only red colonies, suggesting that the *ade2::Ds-NAT1* locus is stable. Accordingly, amplification of the *Ds-NAT1* from cells of each passage with flanking *ADE2*-primers yielded a single 2.6-kb band that was indistinguishable from that obtained from cells after the first passage (Figure 1C), indicating that the *Ds-NAT1* insertion remained stably inserted in the *ADE2* promoter.

Similarly, KMY103G1 cells grew as red colonies on SDC plates after propagation for 24 hr in noninducing SDC + adenine + glucose liquid medium. In contrast, when cells were precultured in glucose-free SDC + adenine + maltose medium to induce *Ac*TPase4xCa expression, white *ADE2* revertant colonies appeared with a frequency ranging from 0.0001 to 0.001%.

From 128 independent *ADE2* revertant KMY103G1 colonies, the *Ds-NAT1* excision sites were amplified. The PCR bands from all revertants were indistinguishable in size from the 510-bp band in the progenitor strain GZY896, suggesting that the *Ds-NAT1* had been excised from the *ADE2* promoter locus (Figure 1, D and E). Sequencing of the PCR bands revealed 28 different *Ds* excision footprints (Figure 2A). Two predominant footprints, which differed in a single nucleotide at the fusion site, were recovered in 35 and 18 clones, respectively. In addition to these two major footprints, 26 other, different footprints were observed less frequently, with half of them appearing only once. Despite the limited number of footprints analyzed, the bias in excision site repair products was obvious. All footprints exhibited short deletions or palindromic insertions (also termed P-nucleotide insertions; Huefner *et al.* 2011) at either side of the breakpoint, which closely resembled the *Ac/Ds* excision footprints in plants and in budding yeast.

Ds-NAT1 efficiently reinserts into the genome

Reintegration of the *Ds-NAT1* transposon was investigated by cultivating 10 individual red KMY103G1 colonies in SDC + adenine + maltose liquid medium and spreading aliquots on SDC − adenine + glucose plates, yielding 1722 *ADE2* revertant colonies. The white *ADE2* colonies were replica-plated onto nourseothricin-containing SDC − adenine + glucose plates and grown for 2 days at 30°. In total, 1528 colonies ($\sim$90%) grew on nourseothricin. This unexpectedly high frequency (in plants and *S. cerevisiae* reinsertion frequencies of $\sim$60% were observed) prompted us to scrutinize the ploidy of the cells. All five KM103G1 colonies that we tested by flow cytometry were diploid. A DNA gel blot analysis of genomic DNA from white *ADE2* revertant colonies revealed that, with only one exception, all clones still contained the *ade2::Ds-NAT1* allele (Figure 1F). Fourteen clones exhibit an additional, differently sized band, indicating reinsertion of the *Ds-NAT1* transposon elsewhere in the genome. One exceptional clone showed no *ade2::Ds-NAT1* band, with one smaller and one larger band. These data demonstrate that: (i) the KMY103G1 progenitor cells had undergone autodiploidization prior to transposon mobilization, (ii) the *Ds-NAT1* transposon had excised from only one of the two alleles with the exception of one clone, and (iii) in 15 of the 28 *ADE2* revertant clones, the transposon reintegrated in different positions in the *Candida* genome. In the one exceptional clone (marked with a red asterisk), the *Ds-NAT1* transposons apparently transposed from both alleles and reintegrated in two novel chromosomal sites.

For three *ADE2* revertant clones, the genomic reinsertion sites were identified by FPNI-PCR of the 3′-end of *Ds-NAT1* (Figure 2B and Table S2 in File S1). Two of the *Ds-NAT1* elements reintegrated in chromosome 3, 174- and 266-kb distal and proximal to the *ADE2* locus, respectively. The third element transposed to chromosome 5. Amplification and sequencing of both *Ds-NAT1* flanking sequences revealed the canonical 8-bp target site duplications generated upon integration

A *ade2::Ds-NAT1*

```
                     5'                3'
       ATCGTTGAGG[CAG...   Ds-NAT1   ...CTA]TGTCGACGGTGA        F

1      ATCGTTGAGCC                    AGTCGACGGTGA            35
2      ATCGTTGAGC                     AGTCGACGGTGA            18
3      ATCGTT                        CAGTCGACGGTGA            10
4      ATCGTTGA                     ACAGTCGACGGTGA             9
5      ATCGTTGAGCCT                    ACGGTGA                8
6      ATCGTTGAGCCT                    GTCGACGGTGA            6
7      ATCGTTGAGCC                     TCGACGGTGA             5
8      ATCGTT[GA]                    CAGTCGACGGTGA             5
9      ATCGTTGAGC                      CGACGGTGA             4
10     ATCGTT[G]                     ACAGTCGACGGTGA            3
11     ATCGTTGA                    GACAGTCGACGGTGA            3
12     ATCGTTGAGCCT                    AGTCGACGGTGA            3
13     ATCGTTGAGCC                   CAGTCGACGGTGA             2
14     ATCGTTGAGCC                     GTCGACGGTGA            2
15     ATCGTTGAGCC                     GaCGACGaTGA            2
16     ATCGTT                          GTCGACGGTGA            1
17     ATCGTTG[AG]                     TCGACGGTGA             1
18     ATCGTTGAGtCT                    ACGGTGA                1
19     ATCGTTGAGCC        g            GTCGACGGTGA            1
20     ATCGTTGAG          a          AGTCaACGGTGA            1
21     ATCGTTG                       CAGTCGACGGTGA            1
22     ATCGTTGAGCC                     TCcACGGTGA            1
23     AT[C]                          AGTCGACGGTGA            1
24     ATCGTT                        CAGTCcACGGTGA            1
25     ATCGTTGAGCC                     GTCaACGGTGA            1
26     ATCGTT[G]                       TCGACGGTGA            1
27     AT                            AGTCGACGGTGA            1
28     ATCGTTGAG           t         CAGTCGACGGTGA            1
```

B Target Site Duplications

```
"1"  TGTTGGTTGCTGATGTTGTTGTCATTGG[CAG            CTA]GTCATTGGTGGGACTCTAGTAATGTTTG
"2"  CCATCACCTAATTTGACTGACTGATATT[CAG  Ds-NAT1  CTA]CTGATATTATCATAGACTTCTTCCAAAT
"3"  TTCACTCTATTGGAAAAATCGTGTCAAG[CAG            CTA]GTGTCAAGTCCAATTGGCAGCACAAGCA
```

Figure 2 DNA sequences at the *Dissociation* (*Ds*)-*NAT1* excision and reinsertion sites. (A) *Ds* excision footprints of 128 independent KMY103G1 *ADE2* revertant colonies. The top row shows the ade2::Ds-NAT1 sequence on chromosome 3. Rows 1–28 show the recovered *Ds* excision footprints and their incidence (*F*) in independent *ADE2* revertants from *Ac*TPase-4xCa-expressing KMY103G1 cells grown in maltose-containing medium. Arrows above the sequences indicate inverted repeats centered around the complementary bases C and A (boldface red letters) of the nucleotides bordering the *Ds-NAT1* transposon that result from the resolution of intermediate hairpin structures. Lower case letters indicate nucleotides that are not explained with the hairpin model. (B) Target site duplications at *Ds-NAT1* reinsertion sites in transposants "1," "2," and "3"; see Table S2 in File S1.

of the *Ds-NAT1* transposable elements (Figure 2B). Thus, transposition occurred, both close to the original insertion site as well as on a different chromosome.

DISCUSSION

Approaches for genome-wide, untargeted insertion mutagenesis of eukaryotic genomes are frequently based on chemical or ballistic transformation with selectable DNA marker molecules [in fungi, sometimes boosted by restriction enzyme-mediated integration (REMI) reviewed in Jiang *et al.* (2013)], on *A. tumefaciens* infection and T-DNA insertion in plants, or by transposon tagging. In micro-organisms that can be efficiently transformed, large mutant collections can be generated by transformation with plasmid-based mutant libraries produced by

in vitro transposition of bacterial transposons (see for example Kumar *et al.* 2004). However, transformation is inefficient in the majority of eukaryotes. This limitation can be avoided by using *in vivo* transposon tagging that exploits endogenous or heterologous transposons. The advantage of using a heterologous transposon is that it will be unique in the genome and is easily detected. The diploid *C. albicans* WO-1 genome is composed of 2.1% retroelements and contains only two copies each of a Tc1/mariner-like DNA transposon and a (defective) Mutator-like element (Butler *et al.* 2009). However, endogenous transposition activity has not been reported for any of these elements, despite a report that *C. albicans* transposons are not modified by DNA methylation (Mishra *et al.* 2011). In this study, we explored the suitability of the maize *Ac/Ds* transposons for *in vivo* transposon insertion mutagenesis in *C. albicans*.

Like in all heterologous host organisms tested, the nonautonomous *Ds* element in *C. albicans* is genetically stable and its mobilization strictly depends on expression of the *Ac* TPase. The *Ds* excision footprints display sequence features similar to those that have been reported for *Ac/Ds* in other plants and in *S. cerevisiae* (Lazarow *et al.* 2012): deletions and palindromic P-nucleotide insertions at the repaired DNA joint are centered around the complement of the base adjacent to the transposon ends. This indicates that in haploid *C. albicans*, the DNA double-strand break (DSB) repair pathway forms a DNA hairpin followed by hairpin resolution and nonhomologous end joining repair of the empty donor sites, essentially as has been suggested for plants and *S. cerevisiae* (Coen *et al.* 1986; Lazarow *et al.* 2012). However, an interesting distinction is that the structural heterogeneity of the footprints in *C. albicans* is lower than in *Arabidopsis* and budding yeast. For example, the longest P-nucleotide insertions were three nucleotides and occurred only in a single footprint type three times among the 128 footprints (2.3%) (Figure 2A, line 11), whereas in *S. cerevisiae*, the maximal palindrome arm length was 20 nucleotides and 66 of 140 footprints (47%) contained palindromes with ≥3 nucleotide-long arms (Lazarow *et al.* 2012). Also in *Arabidopsis*, the frequency of palindromes at least three nucleotides in length was much higher than in *C. albicans* (20–30%) (Huefner *et al.* 2011; Lazarow *et al.* 2012). The shorter palindrome lengths in *C. albicans* indicate that the (unknown) hairpin-opening enzymes introduce nicks closer to the apex of the hairpin than in *S. cerevisiae* and plants. In *S. cerevisiae*, mutations in the Artemis-related Pso2 protein, which cleaves DNA hairpins (Tiefenbach and Junop 2012), reduced *Ds* excisions 10-fold and increased atypical *Ds* excision site repair products (Yu *et al.* 2004). It will be interesting to determine if the *C. albicans* and *S. cerevisiae* PSO2 orthologs have different activities, or if other enzymes are responsible for the distinct footprints and excision frequencies.

Ac/Ds excision footprints in both *S. cerevisiae* and *A. thaliana* are accompanied by putative microhomologies (MHs) of the repair intermediates after DNA hairpin opening (Yu *et al.* 2004; Huefner *et al.* 2011; Lazarow *et al.* 2012). The frequency of footprints with MHs ranged from 85 to almost 100% in wild-type *A. thaliana* (Huefner *et al.* 2011; Lazarow *et al.* 2012) and from 88 to 89% in *S. cerevisiae* (Yu *et al.* 2004; Lazarow *et al.* 2012). In contrast, only 51% of the footprints in *C. albicans* had putative MHs. This suggests that DNA hairpin opening and DSB repair proceeds somewhat differently in *C. albicans* than in the model yeast *S. cerevisiae* and plants. Shorter P-nucleotide insertions in *C. albicans* may indicate that the nuclease that binds to and opens DNA hairpins is more stringently locked to the hairpin apex in *C. albicans* than the functional homologs in plants and budding yeast. While MHs are not essential for the fusion of open ends, terminal MH in double-stranded DNA with short 3′ overhanging single strands stabilizes DSB repair intermediates and enhances the efficiency of repair (Gu *et al.* 2007). Since DNA ligase IV (encoded by *LIG4*) is critically involved in the repair of DSBs in all eukaryotes and also can ligate DNA ends lacking MHs (Gu *et al.* 2007), the lower frequency of MHs at *Ds* excision footprints suggests that *LIG4*-dependent DSB repair is likely to be more active in *C. albicans* than in the other organisms tested.

The high efficiency of the *Ac/Ds* transposon system demonstrated in this study makes it a potentially superior molecular tool for genome-wide *in vivo* insertion mutagenesis in *C. albicans* compared to alternative approaches. Several *in vitro* insertion mutant libraries were constructed in diploid *C. albicans* using bacterial *Tn7* or *Tn5* derivatives that were transposed into isolated genomic DNA fragments, cloned, amplified in *Escherichia coli*, and then transformed into *C. albicans*, where the transposon together with flanking genomic DNA integrated by homologous recombination into the chromosomes (Uhl *et al.* 2003; Oh *et al.* 2010; Bharucha *et al.* 2011). However, this approach was labor-intensive and yielded 18,000, 3633, and 6528 heterozygous transformants, respectively, in the three studies. Thus, the *in vitro* approach is much less practical for repeated application in multiple different strains, for example clinical isolates. Another advantage of the transposon system described here is that, in addition to generating total loss-of-function alleles, it should be useful to identify haploinsufficient genes in diploid cells and other conditionally essential genes, as was recently demonstrated for *Ac/Ds* insertions in *S. cerevisiae* (Michel *et al.* 2017). Furthermore, it can easily be modified for activation-tagging applications, as has been demonstrated in plants [reviewed in Lazarow *et al.* (2013)], as well as for the generation of overexpression and epitope tagging libraries. The beauty of the system is the ease in generating new libraries with thousands, to hundreds of thousands, of new mutants in a single strain background. When coupled with deep sequencing technologies, this provides the potential for facile, rapid generation of large libraries of random insertion mutants without the need for efficient transformation frequencies.

Author contributions: K.M. conducted sequence design, cloning, transformation of *C. albicans*, footprint and reinsertion site analyses, DNA gel blot, interpretation of data, and writing the manuscript. E.S.-S. performed transformation of *C. albicans*, flow cytometry, FPNI PCR, and reinsertion site analyses. J.C.G. performed DNA gel blotting. G.Z. and Y.W. generated the GZY896 strain. J.B. conceived and designed the work and interpreted data. R.K. conceived and designed the work, interpreted data, and wrote the manuscript. The authors declare no conflict of interest.

LITERATURE CITED

Backen, A. C., I. D. Broadbent, R. W. Fetherston, J. D. Rosamond, N. F. Schnell et al., 2000 Evaluation of the CaMAL2 promoter for regulated expression of genes in Candida albicans. Yeast 16: 1121–1129. https://doi.org/10.1002/1097-0061(20000915)16:12<1121::AID-YEA614>3.0.CO;2-U

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.

Bharucha, N., Y. Chabrier-Rosello, T. Xu, C. Johnson, S. Sobczynski et al., 2011 A large-scale complex haploinsufficiency-based genetic interaction screen in Candida albicans: analysis of the RAM network during morphogenesis. PLoS Genet. 7: e1002058. https://doi.org/10.1371/journal.pgen.1002058

Bire, S., S. Casteret, A. Arnaoty, B. Piegu, T. Lecomte et al., 2013 Transposase concentration controls transposition activity: myth or reality? Gene 530: 165–171. https://doi.org/10.1016/j.gene.2013.08.039

Butler, G., M. D. Rasmussen, M. F. Lin, M. A. Santos, S. Sakthikumar et al., 2009 Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature 459: 657–662. https://doi.org/10.1038/nature08064

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. https://doi.org/10.1016/0092-8674(86)90451-4

Dooner, H. K., and A. Belachew, 1989 Transposition pattern of the maize element *Ac* from the *bz-m2(Ac)* allele. Genetics 122: 447–457.

Emelyanov, A., Y. Gao, N. I. Naqvi, and S. Parinov, 2006 Trans-kingdom transposition of the maize Dissociation element. Genetics 174: 1095–1104. https://doi.org/10.1534/genetics.106.061184

Gerami-Nejad, M., A. Forche, M. McClellan, and J. Berman, 2012 Analysis of protein function in clinical C. albicans isolates. Yeast 29: 303–309. https://doi.org/10.1002/yea.2910

Gerami-Nejad, M., L. F. Zacchi, M. McClellan, K. Matter, and J. Berman, 2013 Shuttle vectors for facile gap repair cloning and integration into a neutral locus in Candida albicans. Microbiology 159: 565–579. https://doi.org/10.1099/mic.0.064097-0

Goryshin, I. Y., and W. S. Reznikoff, 1998 Tn5 in vitro transposition. J. Biol. Chem. 273: 7367–7374. https://doi.org/10.1074/jbc.273.13.7367

Gu, J., H. Lu, B. Tippin, N. Shimazaki, M. F. Goodman et al., 2007 XRCC4: DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. EMBO J. 26: 1010–1023. https://doi.org/10.1038/sj.emboj.7601559

Guo, Y., J. M. Park, B. Cui, E. Humes, S. Gangadharan et al., 2013 Integration profiling of gene function with dense maps of transposon integration. Genetics 195: 599–609. https://doi.org/10.1534/genetics.113.152744

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. https://doi.org/10.1111/j.1365-313X.1994.00705.x

Hickman, M. A., G. Zeng, A. Forche, M. P. Hirakawa, D. Abbey et al., 2013 The 'obligate diploid' Candida albicans forms mating-competent haploids. Nature 494: 55–59 (erratum: Cell 530: 242). https://doi.org/10.1038/nature11865

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. https://doi.org/10.1016/j.dnarep.2011.07.011

Jiang, D., W. Zhu, Y. Wang, C. Sun, K. Q. Zhang et al., 2013 Molecular tools for functional genomics in filamentous fungi: recent advances and new strategies. Biotechnol. Adv. 31: 1562–1574. https://doi.org/10.1016/j.biotechadv.2013.08.005

Kim, J., and P. Sudbery, 2011 Candida albicans, a major human fungal pathogen. J. Microbiol. 49: 171–177. https://doi.org/10.1007/s12275-011-1064-7

Kumar, A., M. Seringhaus, M. C. Biery, R. J. Sarnovsky, L. Umansky et al., 2004 Large-scale mutagenesis of the yeast genome using a Tn7-derived multipurpose transposon. Genome Res. 14: 1975–1986. https://doi.org/10.1101/gr.2875304

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. American Society for Microbiology (ASM) Press, Washington, DC. https://doi.org/10.1128/9781555817954.ch24

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. https://doi.org/10.1073/pnas.96.20.11428

Laufs, J., U. Wirtz, M. Kammann, V. Matzeit, S. Schaefer et al., 1990 Wheat dwarf virus Ac/Ds vectors: expression and excision of transposable elements introduced into various cereals by a viral replicon. Proc. Natl. Acad. Sci. USA 87: 7752–7756. https://doi.org/10.1073/pnas.87.19.7752

Lazarow, K., M. L. Du, R. Weimer, and R. Kunze, 2012 A hyperactive transposase of the maize transposable element Activator (Ac). Genetics 191: 747–756. https://doi.org/10.1534/genetics.112.139642

Lazarow, K., M. L. Doll, and R. Kunze, 2013 Molecular biology of maize Ac/Ds elements: an overview. Methods Mol. Biol. 1057: 59–82. https://doi.org/10.1007/978-1-62703-568-2_5

Mátés, 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. https://doi.org/10.1038/ng.343

McClintock, B., 1951 Mutable loci in maize. Carnegie Inst. Wash. Year Book 50: 174–181.

Michel, A. H., R. Hatakeyama, P. Kimmig, M. Arter, M. Peter et al., 2017 Functional mapping of yeast genomes by saturated transposition. Elife 6: e23570. https://doi.org/10.7554/eLife.23570

Mishra, P. K., M. Baum, and J. Carbon, 2011 DNA methylation regulates phenotype-dependent transcriptional activity in Candida albicans. Proc. Natl. Acad. Sci. USA 108: 11965–11970. https://doi.org/10.1073/pnas.1109631108

Noble, S. M., and A. D. Johnson, 2007 Genetics of Candida albicans, a diploid human fungal pathogen. Annu. Rev. Genet. 41: 193–211. https://doi.org/10.1146/annurev.genet.41.042007.170146

Oh, J., E. Fung, U. Schlecht, R. W. Davis, G. Giaever et al., 2010 Gene annotation and drug target discovery in Candida albicans with a tagged transposon mutant collection. PLoS Pathog. 6: e1001140. https://doi.org/10.1371/journal.ppat.1001140

Quach, H. N., S. Tao, P. Vrljicak, A. Joshi, H. Ruan et al., 2015 A multifunctional mutagenesis system for analysis of gene function in zebrafish. G3 (Bethesda) 5: 1283–1299. https://doi.org/10.1534/g3.114.015842

Santos, M. A., and M. F. Tuite, 1995 The CUG codon is decoded in vivo as serine and not leucine in Candida albicans. Nucleic Acids Res. 23: 1481–1486. https://doi.org/10.1093/nar/23.9.1481

Scofield, S. R., J. J. English, and J. D. G. Jones, 1993 High level expression of the Activator transposase gene inhibits the excision of Dissociation in tobacco cotyledons. Cell 75: 507–517. https://doi.org/10.1016/0092-8674(93)90385-4

Spradling, A. C., H. J. Bellen, and R. A. Hoskins, 2011 Drosophila P elements preferentially transpose to replication origins. Proc. Natl. Acad. Sci. USA 108: 15948–15953. https://doi.org/10.1073/pnas.1112960108

Tiefenbach, T., and M. Junop, 2012 Pso2 (SNM1) is a DNA structure-specific endonuclease. Nucleic Acids Res. 40: 2131–2139. https://doi.org/10.1093/nar/gkr1059

Uhl, M. A., M. Biery, N. Craig, and A. D. Johnson, 2003 Haploinsufficiency-based large-scale forward genetic analysis of filamentous growth in the diploid human fungal pathogen C.albicans. EMBO J. 22: 2668–2678. https://doi.org/10.1093/emboj/cdg256

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. https://doi.org/10.1105/tpc.109.073452

Vrljicak, P., S. Tao, G. K. Varshney, H. N. Quach, A. Joshi et al., 2016 Genome-wide analysis of transposon and retroviral insertions reveals preferential integrations in regions of DNA flexibility. G3 (Bethesda) 6: 805–817. https://doi.org/10.1534/g3.115.026849

Walbot, V., 1992 Strategies for mutagenesis and gene cloning using transposon tagging and T-DNA insertional mutagenesis. Annu. Rev. Plant Physiol. 43: 49–78. https://doi.org/10.1146/annurev.pp.43.060192.000405

Wang, Z., S. Ye, S. Li, B. Zheng, M. Bao et al., 2011 Fusion primer and nested integrated PCR (FPNI-PCR): a new high-efficiency strategy for rapid chromosome walking or flanking sequence cloning. BMC Biotechnol. 11: 109. https://doi.org/10.1186/1472-6750-11-109

Weil, C. F., and R. Kunze, 2000 Transposition of maize Ac/Ds transposable elements in the yeast Saccharomyces cerevisiae. Nat. Genet. 26: 187–190. https://doi.org/10.1038/82827

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. https://doi.org/10.1128/MCB.24.3.1351-1364.2004

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. https://doi.org/10.1073/pnas.1008322108

Zayed, H., Z. Izsvak, O. Walisko, and Z. Ivics, 2004 Development of hyperactive sleeping beauty transposon vectors by mutational analysis. Mol. Ther. 9: 292–304. https://doi.org/10.1016/j.ymthe.2003.11.024

Zeng, G., Y. M. Wang, F. Y. Chan, and Y. Wang, 2014 One-step targeted gene deletion in Candida albicans haploids. Nat. Protoc. 9: 464–473. https://doi.org/10.1038/nprot.2014.029

Communicating editor: B. Andrews