

## Virus-like particle formation and translational start site choice of the plant retrotransposon Tto1

Gudrun Böhmendorfer<sup>a,1</sup>, Kerstin Luxa<sup>a</sup>, Andrea Frosch<sup>b</sup>, Karin Garber<sup>b,2</sup>, Andrea Tramontano<sup>a</sup>, Srecko Jelenic<sup>b,3</sup>, Martina Weber<sup>b</sup>, Andreas Bachmair<sup>a,b,c,\*</sup>

<sup>a</sup> Max Planck Institute for Plant Breeding Research, Department of Plant Developmental Biology, Carl-von-Linné-Weg 10, D-50829 Cologne, Germany

<sup>b</sup> Institute of Botany, Univ. of Vienna, Rennweg 14, A-1030 Vienna, Austria

<sup>c</sup> Max F. Perutz Laboratories, University of Vienna, Department of Chromosome Biology, Dr. Bohr-Gasse 1, A-1030 Vienna, Austria

Received 2 November 2007; accepted 27 November 2007

Available online 8 January 2008

### Abstract

Ty1/copia group retrotransposon Tto1 from tobacco was put under control of an inducible promoter for expression in *Arabidopsis thaliana*. The system was used to analyze intermediates of the transposition process. The Tto1 RNA 5' region has a complex structure and contains several AUG codons. We therefore sought to experimentally define the translation initiation site. Constructs starting at various positions within the structural gag region were expressed *in planta* and functionally characterized. We found that gag proteins starting at the first ATG of the gag-pol ORF (ATG1), but also those starting at the third ATG of the gag-pol ORF (ATG3), can form virus-like particles (VLPs). However, gag protein expressed by the inducible Tto1 element had a size similar to gag starting at ATG1, and mutation of ATG1 in the inducible element abolished reverse transcription. This suggested that translation initiation at ATG1 is essential for the Tto1 life cycle. To support this conjecture, gag protein starting at ATG1, or gag protein shortened amino-terminally by nine amino acids (starting at the second ATG of the gag region, ATG2), was co-expressed with Tto1 carrying mutations at ATG1 and ATG2. *Trans*-complementation of the defective Tto element by gag starting at ATG1, but not by gag starting at ATG2, defines ATG1 as the functional translation initiation site.

© 2007 Elsevier Inc. All rights reserved.

**Keywords:** Inducible retrotransposon; Reverse transcription; Virus-like particles; Translation initiation

### Introduction

Tto1 is an LTR-containing retrotransposon from tobacco that belongs to the Ty1/copia group (Hirochika, 1993; Bachmair et al., 2004; for classification of retroelements, see Capy, 2005; Havecker et al., 2004; Malik and Eickbush, 2001; Peterson-Burch and Voytas, 2002; Wicker et al., 2007). Similar to copia, it has a single long open reading frame that encodes both the structural gag protein, and the enzyme components protease, integrase, and reverse transcriptase/RNase H. Its transposition cycle is expected to match the consensus model: complementarity to iMet-tRNA at the minus strand primer binding site suggests that this tRNA is used as a primer for first strand synthesis. Enzymatic steps are presumably carried out in virus-like particles by the element-encoded enzymes. Tto1 inserts preferentially into coding regions. Activity both in monocot and

**Abbreviations:** VLPs, virus-like particles; ORF, open reading frame; Dex, Dexamethasone.

\* Corresponding author. Max F. Perutz Laboratories, Department of Chromosome Biology, Dr. Bohr Gasse 1, A-1030 Vienna, Austria. Fax: +43 1 4277 9744.

**E-mail addresses:** [gudrun.boehmendorfer@gmi.oeaw.ac.at](mailto:gudrun.boehmendorfer@gmi.oeaw.ac.at) (G. Böhmendorfer), [luxa@mpiz-koeln.mpg.de](mailto:luxa@mpiz-koeln.mpg.de) (K. Luxa), [andrea.frosch@univie.ac.at](mailto:andrea.frosch@univie.ac.at) (A. Frosch), [garber@dialog-gentechnik.at](mailto:garber@dialog-gentechnik.at) (K. Garber), [tramonta@mpiz-koeln.mpg.de](mailto:tramonta@mpiz-koeln.mpg.de) (A. Tramontano), [sjelen@zg.biol.pmf.hr](mailto:sjelen@zg.biol.pmf.hr) (S. Jelenic), [martina.weber@univie.ac.at](mailto:martina.weber@univie.ac.at) (M. Weber), [andreas.bachmair@univie.ac.at](mailto:andreas.bachmair@univie.ac.at) (A. Bachmair).

<sup>1</sup> Current address: Gregor Mendel Institute of Plant Molecular Biology, Dr. Bohr-Gasse 3, A-1030 Vienna, Austria.

<sup>2</sup> Current address: Dialog Gentechnik, Dr. Bohr-Gasse 3, A-1030 Vienna, Austria.

<sup>3</sup> Current address: Department of Molecular Biology, Univ. of Zagreb, HR-10000, Rooseveltov trg. 6, Zagreb, Croatia.

dicot hosts indicates a broad host range (Hirochika et al., 1996; Okamoto and Hirochika, 2000), which would make it an ideal tool for insertional mutagenesis in plant species with large genomes. Transposition activity, however, is restricted to stressful conditions. Generally, pathogen infection or adverse growth conditions (abiotic stress) are frequent inducing conditions for plant elements (Brookfield, 2005; Feschotte et al., 2002; Grandbastien, 1998; Sabot and Schulman, 2006). Likewise, growth of plant cells in tissue culture is effective in induction of most plant retrotransposons including Tto1. However, while convenient for some investigations, a need for tissue culture makes the use for insertional mutagenesis difficult in plant species without efficient regeneration protocol. In addition, levels of intermediates in tissue culture cells may still be too low for biochemical studies.

We therefore initiated investigations to understand regulatory steps that restrict Tto1 activity to stress conditions. In previous work, we found that the Tto1 mRNA 5' leader sequence is complex and that choice of the transcript initiation site may act as a filter that down-regulates Tto1 activity under many non-stress conditions (Böhmdorfer et al., 2005). In this work, we investigated the translation initiation site of Tto1 by analysis of Tto1 gag protein and found that an inducible Tto1 element allows cDNA accumulation in wild type plants in the absence of tissue culture.

## Results

### *Tto1 has a complex 5' untranslated leader sequence*

We have previously shown that Tto1 has a complex leader sequence and that a tightly folded region close to the 5' end, starting at the mapped *in vivo* transcription initiation site at nt 200 of Tto1, imposes restrictions on the mode of translation. To generate a short stretch of unstructured RNA, the Tto1 mRNA 5' end was extended to nt 172. This change greatly enhanced biological activity, presumably by allowing better presentation of the cap structure to the ribosome (Böhmdorfer et al., 2005). Fig. 1A shows the leader sequence of Tto1 RNA.

The necessity for an accessible 5' end suggests ribosome assembly at this position. However, it is unclear whether Tto1 is translated by the standard ribosomal scanning mechanism because two short ORFs precede the long enzyme-encoding gag-pol ORF. Such short ORFs, if translated efficiently, would lead to dissociation of the ribosome and correspondingly poor translation initiation at ATGs further downstream. Fig. 1B lists the sequence environment of Tto1 ATGs. The first open reading frame, called short ORF A in Fig. 1A, consists of only 9 amino acids. Its start ATG has a moderately favorable sequence environment in comparison to *Arabidopsis* genes (Alexandrov et al., 2006; Rogozin et al., 2001). The start codon of the second ORF (called short ORF B in Fig. 1A) is in a poor sequence context. The first four in-frame ATGs of the long, gag-pol encoding ORF (called ATG1, ATG2, ATG3 and ATG4 in the following; cf. Fig. 1A) are in a moderately favorable sequence context. At least one of the two favored bases, A at -3 and G at +4, is present for several ATGs. These characteristics are

apparently shared with cellular genes that have ATG codons in the 5' leader sequence (Rogozin et al., 2001). Sequence context does therefore not offer a strong indication regarding the biologically relevant initiation codon. Moreover, known sequence context preferences of the ribosome generally refer to a scanning mode of mRNA translation. If translation employs any other means of recognition of the initiation site, such as ribosome shunting (Hohn et al., 2001) or an internal ribosome entry-based mechanism (Fernandez et al., 2005; Hellen, 2007), definition of the initiation codon by sequence context may be misleading. We therefore tried to obtain experimental evidence regarding the translation start site of Tto1.

### *Virus-like particle formation in Arabidopsis thaliana*

The best understood contribution of gag to the retrotransposon life cycle is the generation of VLPs as secluded areas dedicated to the formation of reverse transcripts. To observe such particles for Tto1, we made constructs to overexpress Tto1 gag in plants. gag of Tto1 is exclusively derived from a polyprotein by proteolytic processing (Tto1 does not encode a frameshift or an intron). The exact cleavage site is currently unknown. We therefore included a larger part of Tto1 in the constructs (Fig. 2A). This part contains the protease domain to facilitate auto-processing. Experiments described later in this work, and additional control experiments (data not shown), indicated that processing does indeed occur by the co-translated protease domain. Experiments in yeast (data not shown) indicated that Tto1 gag starting at ATG4 (ATG4-gag) does not readily accumulate, whereas Tto1 gag starting at ATG1, ATG2 or ATG3 could be expressed to the expected levels. The suggested instability of ATG4-gag, possibly caused by folding defects, made ATG4 an unlikely start codon of Tto1 gag. We therefore concentrated further efforts to define the initiation site on gag proteins starting at ATG1, ATG2 and ATG3. Constructs for overexpression of Tto1 gag, starting either at ATG1 or at ATG3, and terminating after the protease domain (nt 1993 of Tto1; accession D83003), were introduced into *A. thaliana*. We modified a procedure published by Zheng et al. (2000) for Tto1 gag particle enrichment from plant extracts. Briefly, anti-gag antibody was spread on gold grids and incubated with extract from Tto1 gag-expressing plants. To enhance contrast, samples were prepared for electron microscopy by lightly staining with uranyl acetate. Fig. 2 shows the images obtained. Extracts from *Arabidopsis* plants that did not express Tto1 served as controls. Particles were detected for both gag starting at ATG1 (ATG1-gag) and at ATG3 (ATG3-gag). Similar to published images of Tfl1 or copia (Bachmann et al., 2004; Teysset et al., 2003), the particles appeared uniformly electron-dense, not revealing internal structures. Interestingly, particles made of ATG3-gag had a size of approximately 20 nm and were more uniform in size (19 nm average,  $s=4$  nm,  $n=11$ ), whereas those isolated from ATG1-gag were approximately 30 nm in diameter (29 nm average,  $s=13$  nm,  $n=15$ ) and had a broader size distribution. Many particles were slightly oval, the larger diameter being approximately 3 nm larger than the smaller diameter. These experiments provided a first

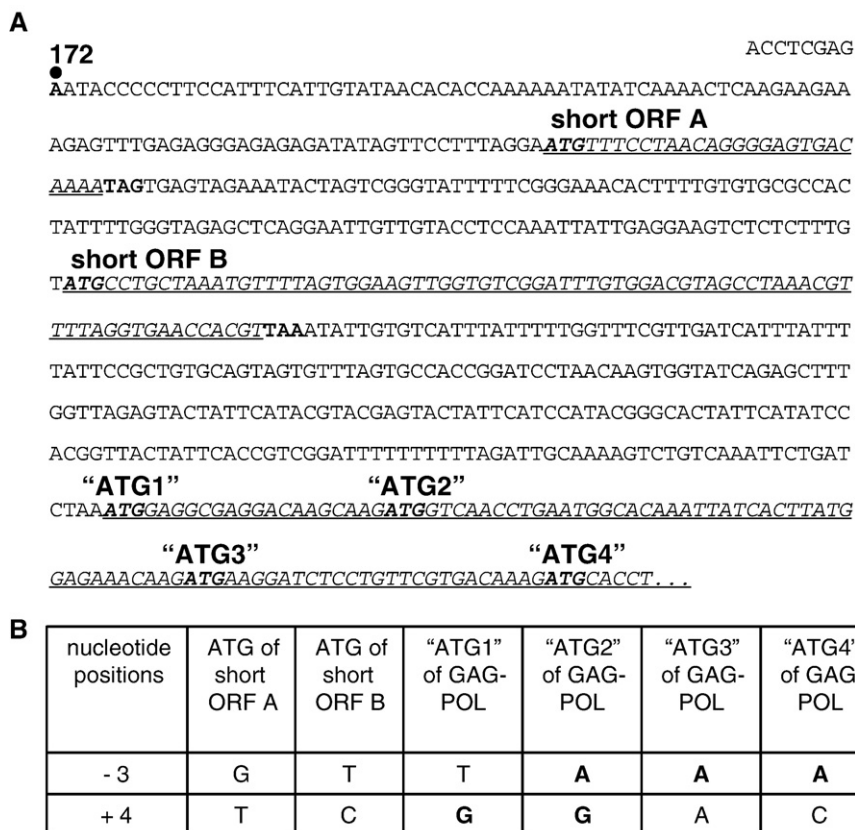


Fig. 1. Sequence context of Tto1 start codons. (A) Sequence of the transcribed region. Homology to Tto1 in biologically active transcripts of this work starts at nt 172 of Tto1 (bold face A, with dot above the letter). Start and stop codons are in bold face, putative open reading frames in italics and underlined. Additional features are described in the text. (B) Sequence context of potential start codons. Adenine (A) at position -3 and Guanine (G) at position +4 enhance translation initiation of a scanning ribosome.

definition of the VLPs of Tto1 made in a permissive host and indicated flexibility in assembly of Tto1 gag. They did, however, not allow a further functional assignment of the precise translational initiation site. We therefore sought to establish further functional assays with Tto1 in *Arabidopsis*.

#### Inducible expression of Tto1 in *Arabidopsis*

The expression of native Tto1 does not readily support biochemical analysis. Firstly, the natural Tto1 promoter allows expression only under stress conditions. Secondly, continued expression is expected to invoke RNA-based defense mechanisms at the transcriptional and posttranscriptional level, which down-regulate the element by RNA degradation and/or promoter silencing (Cheng et al., 2006; Ding et al., 2007; Miura et al., 2001; for review, see Matzke and Birchler, 2005). Recent insight into the biology of Tto1 indicated that Tto1 can be functionally expressed if the mRNA starts at nucleotide position 172 (Fig. 1A; Böhmendorfer et al., 2005). Based on this finding, we supplemented Tto1 with a Dexamethasone (Dex)-inducible promoter system (Aoyama and Chua, 1997). Fig. 3A shows schematically the construct used for inducible expression of Tto1. PCR amplification showed that cDNA is formed after induction with Dex because primers flanking the introns amplify the intron-less fragments, but not before exposure to

Dex (Fig. 3B). Fig. 3C shows a DNA gel blot, using undigested genomic DNA. Genomic DNA was randomly sheared to ca. 20–200 kb fragments and migrated at the separation limit of the gel system. In addition, transgenic *Arabidopsis* lines displayed a distinct band after Dex induction, which can be assigned to un-integrated Tto1 cDNA, as produced from RNA in VLPs. Quantitative assessment of the blot by phosphorimager indicated that induced plant cells contained up to six times more extrachromosomal copies of Tto1 than chromosomal copies inserted into the *Arabidopsis* genome (data not shown). This quantity is consistent with results obtained by the much simpler PCR assay of Fig. 3B.

#### Comparison of Tto1 gag protein initiation sites in *Arabidopsis*

The formation of cDNA by the inducible Tto1 construct implied that the gag protein of induced plants supported functional VLPs. Using Western blotting, it was possible to detect gag protein from induced Tto1, which was compared to Tto1 gag-prot expression constructs starting at different ATG codons (cf. Fig. 2A). Fig. 4A shows the single band detected after Tto1 induction, which was absent from un-induced extract. Fig. 4B shows a side-by-side comparison of plant extracts containing ATG1-gag induced Tto1 gag and ATG3-gag. The expression constructs were identical to those employed for VLP

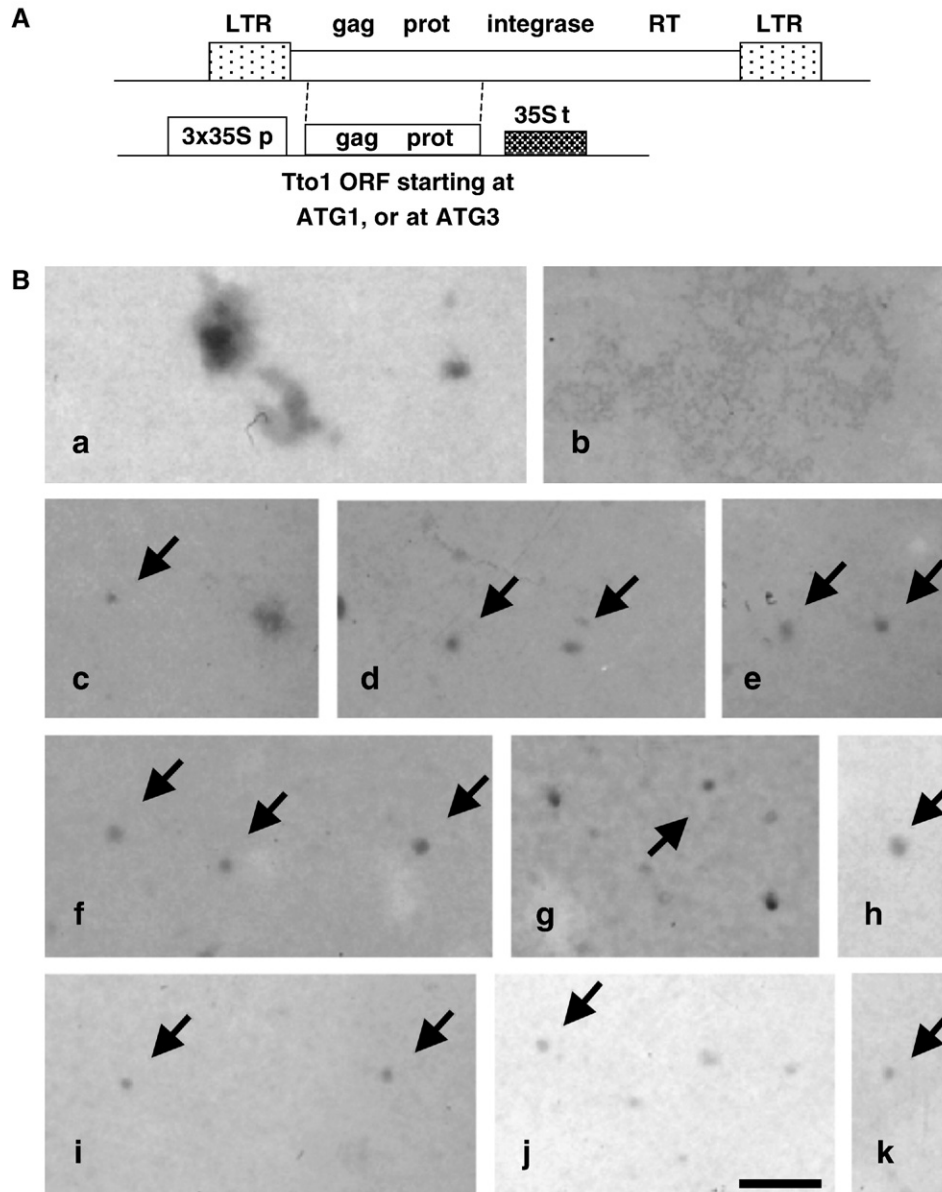


Fig. 2. Tto1 VLP formation in *A. thaliana*. (A) Scheme of constructs used for expression of gag protein, aligned with the native Tto1 element. A gag-prot fusion protein is cleaved auto-catalytically to release the structural gag moiety. gag, structural protein of Tto1; prot, protease; RT, reverse transcriptase; LTR, long terminal repeat. (B) Electron microscopic pictures of VLPs. (a, b) Control spreads from plant cells not expressing Tto1 gag protein show occasionally larger assemblies of electron-dense material (a). Alternatively, a slightly granular background can be seen (b). (c–h) Particles visualized from *Arabidopsis* cells expressing ATG1-gag. (i–k) Particles from *Arabidopsis* cells expressing ATG3-gag. The latter particles are smaller than the former ones (average 30 nm versus 20 nm). Arrows point to typical particles. Scale bar shown in panel j indicates 150 nm and is valid for all panels. Magnification was 48,000 $\times$ .

isolation. The experiment suggested that Tto1 gag from the inducible, active construct is similar in size to ATG1-gag. Up to this point, we did not employ constructs starting at ATG2 in plants because ATG1 and 2 are only 8 amino acids away from each other. Moreover, the VLP formation assay (Fig. 2) had indicated that gag assembly is robust to an extent that allows even the more truncated ATG3 construct to form such particles. The results of Figs. 4A and B, which used the inducible Tto1 construct, prompted us to also test ATG2 as a start codon. An ATG2-gag construct for expression of gag-prot with ATG2 as initiation codon was made and introduced into *A. thaliana*. Fig. 4C shows a side-by-side comparison of plant extract from

ATG2-gag expressing plants and extract from the inducible Tto1 plants. The result was suggestive of a small size difference between ATG2-gag and gag of inducible Tto1 but could not serve as a definitive proof. The general conclusion was therefore that Tto1 gag, as produced by an element capable of Tto1 cDNA production, most probably starts with ATG1, but the experiments did not rule out ATG2 as the start codon.

#### *In planta test of Tto1 variants*

To detect differences between ATG1 and ATG2 as translational start points, we introduced mutations into the inducible

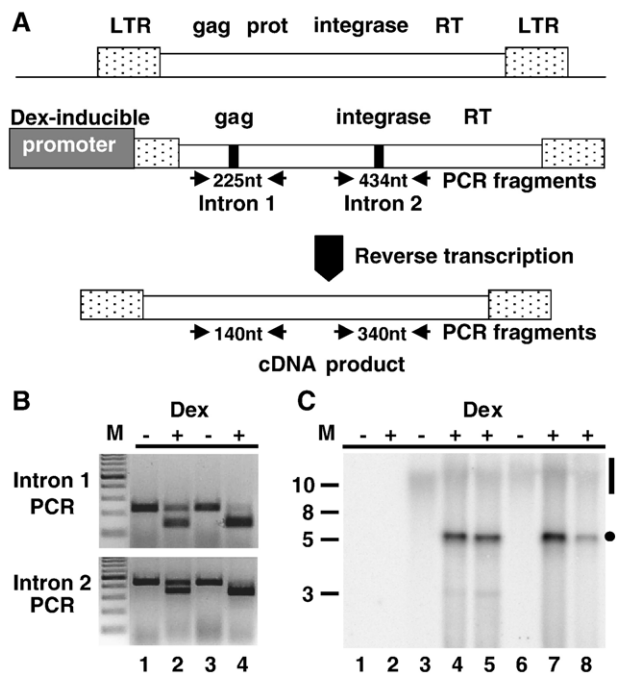


Fig. 3. Inducible expression of Tto1 in *A. thaliana*. (A) Schematic diagram of native Tto1 (top), an inducible Tto1 construct as present on inserted T-DNA copies (middle), and the cDNA formed upon induction of the T-DNA-encoded element (bottom), which is identical to native Tto1. Approximate primer positions and PCR fragment sizes are also indicated. (B) PCR-based detection of cDNA is possible by choice of primers that flank the introns of the inducible construct (fragment positions and sizes indicated in panel A). T-DNA encoded copies serve as an internal standard in semi-quantitative analysis. Lanes 1, 2 and 3, 4 represent two independent transgenic plant lines, either non-induced (lanes 1 and 3) or induced by addition of chemical inducer Dex (2, 4). M, molecular weight marker lane (100 bp steps, strong band at 500 bp). (C) DNA gel blot experiment with undigested DNA shows that a significant portion of the produced cDNA is present as apparently full-length, extrachromosomal copies (predicted size, 5.3 kbp). Lanes 1 and 2, DNA prepared from untransformed Col-0 plants. Lanes 3–5, DNA from line carrying the inducible Tto1 construct integrated into the genome. Lanes 6–8, same as lanes 3–5, but different transgenic line. Plants used for lanes 2, 4, 5, 7 and 8 were treated with Dex. Molecular weight marker sizes are indicated to the left (in kbp). Line to the right indicates position of sheared genomic DNA, dot indicates size of full-length Tto1 cDNA.

Tto1 construct that eliminate ATG1. A similar construct, eliminating ATG1 and ATG2, was also made for testing. Figs. 5A and B show that the mutations interfered with cDNA production, implying that productive VLPs are not formed in plants containing these constructs. RT-PCR controls confirmed that mRNA expression was adequate in the plants (lanes 6 to 10 of Figs. 5A and B). This result is consistent with ATG1 being the only biologically active translation initiation site.

#### Trans-complementation of gag variants

The inability of Tto1 constructs with mutations in ATG1 and ATG2, or in ATG1 alone, to produce cDNA could be due to production of gag protein with inadequate functionality. However, the sequence changes may also have deleterious consequences for RNA structure so that the changes decrease translation or interfere with later steps such as RNA packaging

or reverse transcription. We made an attempt to distinguish between decreased functionality of gag and compromised functionality of the RNA by providing gag proteins in *trans*. Although retroelements generally show a *cis* preference for association of RNA with the proteins encoded by this RNA, the potential of retroelement proteins for activity in *trans* is suggested by the existence of incomplete, yet active element families (Dewannieux et al., 2003; Sabot and Schulman, 2006; Tanskanen et al., 2007). Furthermore, experiments in human cells, in *Drosophila* and in yeast, suggested that *trans*-complementation can sometimes be observed upon overexpression of one component (Bolton et al., 2005; Busseau et al., 1998; Nikolaitchik et al., 2006; Wei et al., 2001). We tested this route by crossing plants expressing Tto1 ATG1-gag or ATG2-gag with those containing the Tto1 variant with changes in ATG1 and ATG2. Fig. 6A shows schematically the employed transgenes. The two settings differ only by the amino-terminal end of the gag construct, which was 8 amino acids shorter for ATG2-gag than for ATG1-gag. F1 plants were induced by Dex and used for DNA extraction. DNA from plants that screened positive for both transgenes was subsequently used for PCR with primers flanking the first intron (the pre-screening was necessary because not all Tto1 expressing constructs of the parental plants were in the homozygous state). To rule out contamination from previously amplified spliced DNA fragments, a combination of primers exclusive for this experiment was employed. Fig. 6B shows that a low level of Tto1 cDNA formation is observed in seedlings of an F1 population containing ATG1-gag, but not in seedlings containing ATG2-gag. This result implies weak complementation by, and thus functionality of, ATG1-gag, but not of ATG2-gag. The finding

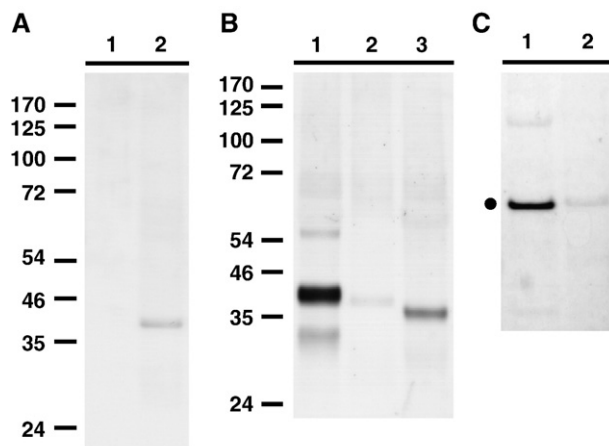


Fig. 4. Tto1 gag protein expression in *A. thaliana*. Constructs for expression of Tto1 gag starting at different positions were compared to an inducible Tto1 construct that produces functional gag. (A) Lane 1, inducible Tto1 in non-induced state does not express gag protein. Lane 2, induction results in appearance of a single gag band. (B) Extract from plants with induced Tto1 element (lane 2) is flanked by ATG1-gag (lane 1) and by ATG3-gag (lane 3), both expressed in plants. Constructs for expression of these proteins are as shown in Fig. 2A. Note that the protein present in plant cells with active reverse transcription activity is larger than ATG3-gag. (C) Side-by-side comparison of ATG2-gag (lane 1) and gag from induced Tto1 plants (lane 2). Dot indicates position of Tto1 gag.

that individual ATG1-gag plants show a variable extent of complementation may be attributable to the fact that two transgenes with related sequences can invoke silencing mechanisms. An additional control for the PCR reactions (lanes C of Fig. 6B) was the use of a mixture of DNA from plants containing only the inducible Tto1 construct with mutations in ATG1 and ATG2 and of DNA from plants expressing only ATG1-gag (lane 15) or only ATG2-gag (lane 7). This mixing experiment did not result in the smaller size intron-less cDNA-derived band. The negative result rules out template switching from the gag-prot construct to the full-length construct as a source of the spliced bands.

## Discussion

In this work, we analyzed Tto1, a member of the Ty1/copia group of LTR retrotransposons. A Dexamethasone-inducible construct allowed *in planta* production of Tto1 cDNA and was used to identify the biologically active translation initiation site of Tto1. The main conclusions of the experiments were: (i) Tto1 VLPs have a size of approximately 30 nm. (ii) Tto1 gag protein is flexible regarding VLP formation because amino-terminal deletion of 23 amino acids, resulting from translation initiation at ATG3 (cf. Fig. 1A), still allows VLP formation. These particles are, however, smaller, and not fully functional. (iii) The biologically active translation product of Tto1 starts at ATG1.

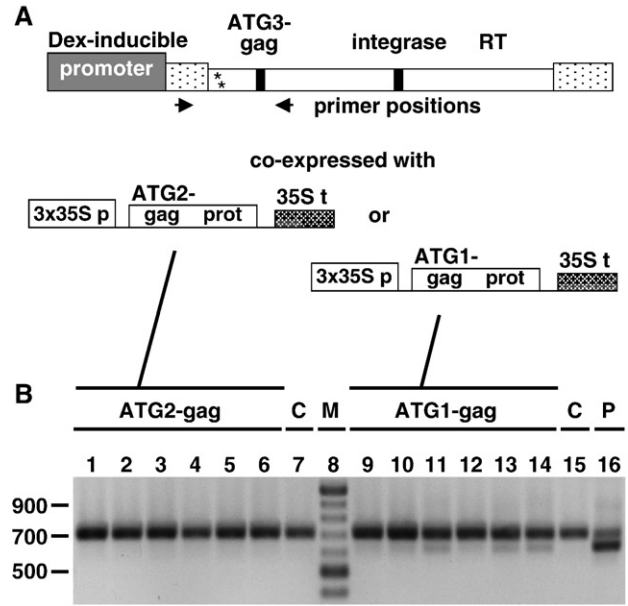


Fig. 6. Trans-complementation by ATG1-gag, but not by ATG2-gag. Plants expressing ATG2-gag or ATG1-gag, respectively, were crossed to plants harboring a Tto1 variant with mutated ATG1 and ATG2. F1 progeny were selected for DNA extraction and PCR analysis (primers flanking intron 1). (A) Constructs expressed *in planta* for the complementation assay. Asterisks in the top construct symbolize mutations at ATG1 and ATG2 as shown in Fig. 5A. (B) Results of PCR assays. Lanes 1–6, co-expression of ATG2-gag together with Tto1 mutated at ATG1 and ATG2; lanes 9–14, co-expression of ATG1-gag together with Tto1 mutated at ATG1 and ATG2. Lanes 7 and 15 (label C) were controls in which DNA from plants containing the ATG2-gag construct only (lane 7) or the ATG1-gag construct only (lane 15) was mixed with DNA from plants harboring exclusively Tto1 with mutated ATG1 and ATG2 and used for PCR reaction. Lane 16 (label P), positive control (DNA from induced plants carrying a non-mutated Tto1 construct was used for PCR). Lane 8, DNA size marker (100 bp steps). Marker sizes (in bp) are indicated to the left.

This conclusion is based on the size of the protein in SDS-PAGE and on functional complementation. Tto1 elements with mutated ATG codons 1 and 2 (which can only use the third in-frame ATG to produce a polyprotein) do not proceed to the stage of reverse transcription. Reverse transcription can be partially restored by co-expression of gag protein starting at ATG1, but not of gag protein starting at ATG2.

In contrast to retroviruses, which can produce high titers of replication intermediates, retrotransposon activity is generally much more restricted due to the fact that excessive activity may kill the host and thus the element itself. In order to obtain a detailed picture of the life cycle of Tto1, we developed an inducible expression system for the plant *A. thaliana*. It was shown previously that *Arabidopsis* allows transposition of wild type Tto1 in tissue culture cells or in mutants defective in DNA methylation (Okamoto and Hirochika, 2000; Takeda et al., 2001). The heterologous host *A. thaliana* thus offers the possibility to study all steps of the Tto1 life cycle. While the original host, tobacco, carries ca. 30 copies of Tto1 in its genome, a heterologous host should allow to precisely score functional changes because all element copies are introduced by transformation, and all detectable products are derived from those introduced copies. A low copy number of the transgene,

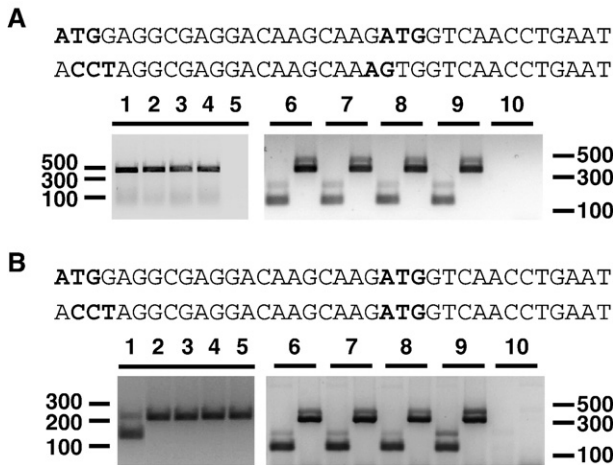


Fig. 5. Test of Tto1 variants. Constructs utilized are as shown in Fig. 3A, but with the sequence variations indicated at the top of each panel (lower line, mutated sequence, upper line, WT sequence for comparison). (A) Sequence changes at ATG1 and ATG2 abolish activity. Four different transgenic lines (lanes 1 to 4) were induced by Dex treatment and used for intron 2 PCR. Lane 5 shows reaction without DNA addition. Lanes 6 to 9, RNA was prepared for RT-PCR from plant material as used for lanes 1 to 4, to demonstrate production of mRNA. Each sample was used for intron 1 PCR (left lane) and for intron 2 PCR (right lane). Lane 10, control without cDNA. (B) Similar to (A), but sequence change only at ATG1, and primers flanking intron 1 were used for lanes 1–5. Lane 1, PCR bands obtained from induced positive control (no sequence change in inducible Tto1). Lanes 2–5, PCR with DNA from different transgenic lines with sequence change at ATG1 after Dex induction. Lanes 6–9, RNA from the induced lines used for lanes 2–5 was subjected to RT-PCR with primers flanking intron 1 (left lane) and intron 2 (right lane) to demonstrate transcription of the transgene upon induction. Lanes 10, non-transgenic plants were used for RNA preparation, reverse transcription and PCR.

and the chemically inducible promoter introduced in this work, should also side-step silencing mechanisms, which normally interfere with retrotransposon activity.

Previous investigations had shown that the Tto1 mRNA leader sequence is complex to an extent that predictions regarding translation initiation are uncertain (Böhmdorfer et al., 2005). In particular, two short open reading frames may make ribosome scanning inefficient (Fig. 1). Interestingly, the related Cauliflower Mosaic Virus uses a mode of translation that skips parts of its RNA (Hohn et al., 2001). A mechanism with similar characteristics operates in translation of mammalian cat-1 Arg/Lys transporter (Fernandez et al., 2005). In both cases, the choice of start site is not determined by the linear sequence and therefore difficult to predict from sequence features alone. Because a similarly unconventional mechanism may be used for Tto1 mRNA translation, we sought experimental evidence to determine the start ATG. Four in-frame ATGs were initially considered as potential translation initiation sites. However, translation initiation at ATG4 was considered rather unlikely because a construct starting at ATG4 did not accumulate upon overexpression in yeast, suggesting problems with folding or assembly (G. B., unpublished observations).

Tto1 gag proteins starting at the first three in-frame ATGs of the long ORF (Fig. 1) could be expressed in *Arabidopsis* and were detected in Western experiments using anti-gag antibodies (Fig. 4). The same plant material was used for visualization of VLPs (Fig. 2). Particles appeared uniformly electron-dense, which is similar to published images of Tfl and copia, but differs from images of Ty1, which also revealed internal structures (Bachmann et al., 2004; Brookman et al., 1995; Burns et al., 1992; Teysset et al., 2003). The size of 30 nm diameter for Tto1 VLPs corresponds well to the published size of particles of the related yeast element Ty1, which have a size distribution between 20 and 35 nm (Burns et al., 1992), or of fission yeast Tfl particles, which were also reported to be below 50 nm (Teyssset et al., 2003). Likewise, particles of barley retrotransposon BARE-1 were found to be ca. 35 nm in size (Jääskeläinen et al., 1999). In contrast, particles of yeast retrotransposon Ty3, which belongs to the gypsy group of retroelements, have a size range of 45–50 nm (Hansen et al., 1992; Kuznetsov et al., 2005). Interestingly, Tto1 VLPs made by an amino-terminally shortened gag protein were smaller than VLPs made by full-length Tto1 gag (ca. 20 nm versus 30 nm).

An inducible promoter allowed expression of Tto1 and detection of cDNA (Fig. 3) and of gag protein (Fig. 4). The size of this biologically active protein suggested translation initiation at ATG1. In support of this assumption, mutational changes of ATG1, or of both ATG1 and ATG2, abolished Tto1 activity (Fig. 5). We cannot rule out that the mutational changes necessary to destroy ATG1 and ATG2 also changed important *cis* elements of the RNA sequence. However, the mutationally changed element could be partially reactivated by co-expression of gag starting at ATG1. This complementation experiment indicated that the essential consequence of the mutations in ATG1 and ATG2 was an altered protein sequence, and not changes on the RNA sequence level (Fig. 6). In contrast, ATG2-gag, although differing from ATG1-gag by only 8 amino-

terminal residues, was not able to complement, suggesting that it was biologically inactive.

Generally, a major motivation for the experiments presented was to understand regulatory steps in retrotransposition beyond transcriptional induction and RNA-mediated silencing. The inducible system for Tto1 expression could circumvent these two regulatory steps and thus offered the opportunity to study previously uncharacterized steps of the Tto1 life cycle. We started with an assessment of the translation initiation site. Knowledge of the Tto1 translational start site, and the potential for *trans*-complementation, should make it now possible to analyze *cis*-acting sequence motifs of Tto1 RNA without the additional need to preserve the protein coding capacity. Furthermore, accumulation of cDNA upon Tto1 induction should allow to study steps of the Tto1 life cycle that occur after reverse transcription.

## Materials and methods

### Biological material and growth conditions

*A. thaliana* accession Col-0 was transformed by the floral dip method (Clough and Bent, 1998). Seeds were germinated on agar medium containing MS salts and, if appropriate, 15 mg/l of Hygromycin (Roche). Dexamethasone (Dex) induction was carried out by transfer of seedlings from plate to 24 well microtiter plates containing 1 ml of MS medium with 1% (w/v) sucrose, supplemented with up to 10  $\mu$ M Dex as indicated. Plants transferred to soil were kept in long day light conditions as described (Yin et al., 2007). Plants used for Fig. 3 were grown on soil and sprayed with aqueous solution of Dex (10  $\mu$ M) twice with a 3-day interval prior to harvesting. Crosses for Fig. 6 were made with the ATG1-gag or ATG2-gag containing plants as male, and as female parents, with identical results (in Fig. 6B, DNA for lanes 1–3 and 9–11 was from crosses with ATG-gag plants as male parent, and DNA for lanes 4–6 and 12–14 was from crosses with ATG-gag plants as female parents).

### DNA constructs

An intron was inserted into the gag domain of Tto1 in the following way: Vector pK19 was digested with *AseI*, treated with Klenow fragment of DNA polymerase I and religated. A *Bam*HI *Sal*I fragment (Tto1 nt 570–1100) was inserted to give plasmid p19TtoBS. The latter plasmid was digested with *Ecl*136II and *Sma*I and religated. The ensuing construct was cleaved with *Bst*XI and *Ase*I to replace the excised fragment by a PCR-generated *Bst*XI *Ase*I fragment that had been amplified with primers CGTGCCACATTTCTGGTGTGACACATGCAAG-GTCGTTATGGGACAAGCTCGAAGAGCTCTATGC and 756A (sequence CCGACAACCTGGTTCGACAATCCCTT), using Tto1 DNA as a template. The ensuing plasmid, p19TtoBSS, contains a unique *Sac*I site that does not change the Tto1 protein sequence. It was digested with *Ase*I and *Sac*I and used to insert an *Ase*I and *Sac*I digested PCR fragment, amplified with primers CGGAGCTCTATGCCTCTAAAACAGGTATATTCAAGACAA-TAATT and GGCAAATTAATTTTGTCAAGTAAATAATTTGTT-GTTACCTACAAGGAAAGTTTAT together with *Arabidopsis*

DNA. The fragment contains an 85 nt intron from RNA polymerase II largest subunit (nt 4188–4272 of accession X52954) between Tto1 nucleotides 998 and 999. The ensuing plasmid, p19TtoBlnS, was digested with *BsiWI* and *SalI*. The insert fragment was cloned between *BsiWI* and *SalI* sites of plasmid p2RT172 (a construct similar to p2RT200T, but starting at nt 172 of Tto1; Böhmendorfer et al., 2005) to give plasmid p2RT1723.I. A *HindIII* *SalI* fragment from p2RT1723.I was used to replace the analogous fragment of pBTtoP+S (Böhmendorfer et al., 2005) to give plasmid pBTto3.I. An *XhoI* *SmaI* fragment from the latter vector was inserted into vector pTAs, which had been digested with *SpeI*, treated with Klenow fragment of DNA polymerase, and digested with *XhoI*. Vector pTAs is identical to vector pTA7002 (Aoyama and Chua, 1997) except for a 24 bp deletion after the transcription initiation site. It was constructed by replacing an *AflIII* *XhoI* fragment by an *AflIII* *XhoI* digested PCR fragment generated with oligonucleotides *GTTTCTTAAGATTGAATCCTGTT* and *CCGCTCGAGTCCTCTCCAAATGAAA*, together with pTA7002 as a template. The Tto1 containing plasmid, which allows inducible expression of a Tto1 element that contains two introns, was called pTAsTto3.I. To mutate ATG codons in an inducible Tto1 construct, primers *CCGCCTAGGCGAGGACAAGCAAAGTGGTCAAC* and 756A were employed together with Tto1 DNA in a PCR reaction. The ensuing fragment was cloned into *SmaI* digested vector pSK+ to give plasmid pSKAvS. A second PCR reaction, using primers *GCCAGCCTAGGTTTAGATCAGAATTTGACAGACT* and *CCCTCGAGAAGCTTGGGAGAGGTAAGCAATGAT* and Tto1 DNA, generated a fragment that was digested with *XhoI* and *AvrII* and inserted between *XhoI* and *AvrII* sites of pSKAvS to give pSKXAvS. A *BsiWI* *BstXI* fragment from this vector was inserted between *BsiWI* *BstXI* sites of p19TtoBlnS. From there, a series of cloning steps similar to those employed to generate construct pTAsTto3.I led to plasmid pTAs172A3. This construct has the nucleotide changes at ATG1 and ATG2 shown in Fig. 5A and contains an *AvrII* site at the position of former ATG1. A Tto1 construct with changes only in ATG1 was obtained by using primers *CCCGGGCCTAGGCGAGGACAAGCAAGATGGTC* and 756A with Tto1 DNA to obtain a fragment that contains the nucleotide changes shown in Fig. 5B. A series of intermediate steps as exemplified above allowed assembly of construct pTAsTto1A2, which has a mutation only at ATG1. For high level expression of Tto1 gag with initiation site ATG1 in plants, a fragment of Tto1 starting with *CTCGAGAATTACTATTTACAATTACCATG* (the initiation codon ATG1 at nt 716 of Tto1, underlined, is preceded by a segment of the tobacco mosaic virus leader to enhance translation efficiency) and ending with *TGGCATA-GACGTTGACTAAGTAGTCTAGA* (the underlined TGA follows nt 1994 of Tto1 to create a stop codon) was assembled between *XhoI* and *XbaI* sites of pSK+ to give pSKΩM1 and thereafter inserted between *XhoI* and *XbaI* sites of plant binary vector pHi (Schlögelhofer and Bachmair, 2002) to give construct pHiΩM1. The construct contains gag and protease domains of Tto1 to release mature gag upon translation. To obtain a construct for expression of gag starting at ATG2, primers *CCCGGGCCATGGT-CAACCTGAATGGCAC* and 756A were used with Tto1 DNA as template to generate a fragment flanked by *NcoI* and *SalI* sites. The

fragment was used to replace an *NcoI* *SalI* fragment of pSKΩM1 to obtain pSKΩM2, which served as a source for insertion of an *XhoI* *XbaI* fragment into pHi, to give pHiΩM2. A PCR fragment generated with primers *CCGCTCGAGAATTACTATTTACAAT-TAAGATGAAGGATCTCCT* and 756A and with Tto1 DNA as a template is flanked by *XhoI* and *SalI* sites and was used to replace an *XhoI* *SalI* fragment of pSKΩM1 to give pSKΩM3(n). An *XhoI* *XbaI* fragment from the latter construct was used to generate pHiΩM3(n), which is similar to pHiΩM1, but with start of the Tto1 sequence at ATG3 (nt 782).

#### Protein analysis

Protein analysis by Western blotting was carried out essentially as described (Stary et al., 2003), except that plant tissue was ground to a powder in liquid nitrogen before further processing, and NuPAGE 4–12% Bis–Tris gels (Invitrogen) were employed. Primary antibody directed against gag was produced in rabbits (Feik, 2000).

#### VLP isolation and detection

In an adaptation of the method described by Zheng et al. (2000), 50 mesh gold grids were incubated overnight at 4°C with sterile-filtered anti-gag antiserum (dilution 1:200 in phosphate-buffered saline; Feik, 2000). On the next day, gold grids were washed three times with 500 µl sterile-filtered phosphate-buffered saline (PBS, Sigma). Amberlite XAD4 (Sigma) was stirred with 200 mM Na-phosphate buffer (pH 7) for 2 h at room temperature with two buffer changes. 150 mg plant material was ground in liquid nitrogen and combined with 150 µl 200 mM Na-phosphate buffer and 150 mg of wet Amberlite resin. The sample was mixed for 1 min and centrifuged for 5 min (4 °C, 18,000 ×g). The centrifugation step was repeated with the supernatant, which was thereafter applied to the processed gold grids (15 µl extract per grid) for overnight incubation at 4°C. Gold grids were subsequently washed three times with filtered distilled water (400 µl each) and stained with 2% (w/v) uranyl acetate in water (50 µl per grid) at room temperature for 15 min. After four washes in water as above, grids were dried and used for visualization on a Zeiss EM 109 electron microscope.

#### Plant DNA analysis

Plant DNA for gel blotting was prepared using the DNeasy Plant Maxi kit (Qiagen). Gel blot analysis was carried out as described (Böhmendorfer et al., 2005). For PCR reactions, DNA was prepared from fresh material either manually as described (Böhmendorfer et al., 2005) or with a cell disrupter for microtiter format (Retsch MM301) followed by DNA isolation with BioSprint 96 (Qiagen). PCR reaction with primers flanking intron 2 was as described (Böhmendorfer et al., 2005). Detection of intron 1 was similar, but with primers *AGCTCGAA-GAGTTGTATGCCTCT* (forward) and 756A (reverse). For the experiments of Fig. 6, oligonucleotide *GGTGAACCACGT-TAAATATTGTGTC* was used as a forward primer instead.



## RNA analysis

RT-PCR was carried out as described (Böhmendorfer et al., 2005).

## Acknowledgments

We wish to thank M. Lehnen, A. Krull and R. D. Hirtz for expert technical help and N. Feik for anti-gag antibody. This work was funded by the Austrian Science Foundation FWF (grant P15136), the German Research Foundation DFG (grant Ba1158/4-1) and by the Max Planck Society.

## References

- Alexandrov, N.N., Troukhan, M.E., Brover, V.V., Tatarinova, T., Flavell, R.B., Feldmann, K.A., 2006. Features of *Arabidopsis* genes and genome discovered using full-length cDNAs. *Plant Mol. Biol.* 60 (1), 69–85.
- Aoyama, T., Chua, N.-H., 1997. A glucocorticoid-mediated transcriptional induction system in transgenic plants. *Plant J.* 11 (3), 605–612.
- Bachmair, A., Garber, K., Takeda, S., Sugimoto, K., Kakutani, T., Hirochika, H., 2004. Biochemical analysis of long terminal repeat retrotransposons. In: Miller, W., Capy, P. (Eds.), *Methods in Molecular Biology. Mobile Genetic Elements: Protocols and Genomic Applications*, vol. 260. Humana Press, Totowa, NJ, pp. 73–82.
- Bachmann, A.S., Corpuz, G., Hareld, W.P., Wang, G., Collier, B.-A., 2004. A simple method for the rapid purification of copia virus-like particles from *Drosophila* Schneider 2 cells. *J. Virol. Methods* 115 (2), 159–165.
- Böhmendorfer, G., Hofacker, I., Garber, K., Jelenic, S., Nizhynska, V., Hirochika, H., Stadler, P., Bachmair, A., 2005. Unorthodox mRNA start site to extend the highly structured leader of retrotransposon Tto1 mRNA increases transposition rate. *RNA* 11 (8), 1181–1191.
- Bolton, E.C., Coombes, C., Eby, Y., Cardell, M., Boeke, J.D., 2005. Identification and characterization of critical *cis*-acting sequences within the yeast Ty1 retrotransposon. *RNA* 11 (3), 308–322.
- Brookfield, J.F.Y., 2005. The ecology of the genome—mobile DNA elements and their hosts. *Nat. Rev., Genet.* 6 (2), 128–136.
- Brookman, J., Stott, A.J., Cheeseman, P.J., Adamson, C.S., Holmes, D., Cole, J., Burns, N.R., 1995. Analysis of TyA protein regions necessary for formation of the Ty1 virus-like particle structure. *Virology* 212 (1), 69–76.
- Burns, N.R., Saibil, H.R., White, N.S., Pardon, J.F., Timmins, P.A., Richardson, S.M.H., Richards, B.M., Adams, S.E., Kingsman, S.M., Kingsman, A.J., 1992. Symmetry, flexibility and permeability in the structure of yeast retrotransposon virus-like particles. *EMBO J.* 11 (3), 1155–1164.
- Busseau, I., Malinsky, S., Balakireva, M., Chaboissier, M.-C., Teninges, D., Bucheton, A., 1998. A genetically marked I element in *Drosophila melanogaster* can be mobilized when ORF2 is provided in *trans*. *Genetics* 148 (1), 267–275.
- Capy, P., 2005. Classification and nomenclature of retrotransposable elements. *Cytogenet. Genome Res.* 110 (1–4), 457–461.
- Cheng, C., Daigen, M., Hirochika, H., 2006. Epigenetic regulation of rice retrotransposon *Tos17*. *Mol. Genet. Genomics* 276 (4), 378–390.
- Clough, S.J., Bent, A.F., 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16 (6), 735–743.
- Dewannieux, M., Esnault, C., Heidmann, T., 2003. LINE-mediated transposition of marked Alu sequences. *Nat. Genet.* 35 (1), 41–48.
- Ding, Y., Wang, X., Su, L., Zhai, J.X., Cao, S.Y., Zhang, D.F., Liu, C.Y., Bi, Y.P., Qian, Q., Cheng, Z.K., Chu, C.C., Cao, X.F., 2007. SDG714, a histone H3K9 methyltransferase, is involved in *Tos17* DNA methylation and transposition in rice. *Plant Cell* 19 (1), 9–22.
- Feik, N., 2000. Expression des Retrotransposons Tto1 in Hefe und *A. thaliana*. Diploma Thesis, Univ. of Vienna, Austria.
- Fernandez, J., Yaman, I., Huang, C., Liu, H., Lopez, A.B., Komar, A.A., Caprara, M.G., Merrick, W.C., Snider, M.D., Kaufman, R.J., Lamers, W.H., Hatzoglou, M., 2005. Ribosome stalling regulates IRES-mediated translation in eukaryotes, a parallel to prokaryotic attenuation. *Mol. Cell* 17 (3), 405–416.
- Feschotte, C., Jiang, N., Wessler, S.R., 2002. Plant transposable elements: where genetics meets genomics. *Nat. Rev., Genet.* 3 (5), 329–341.
- Grandbastien, M.-A., 1998. Activation of plant retrotransposons under stress. *Trends Plant Sci.* 3 (5), 181–187.
- Hansen, L.J., Chalker, D.L., Orlinsky, K.J., Sandmeyer, S.B., 1992. Ty3 GAG3 and POL3 genes encode the components of intracellular particles. *J. Virol.* 66 (3), 1414–1424.
- Havecker, E.R., Gao, X., Voytas, D.F., 2004. The Diversity of LTR Retrotransposons. <http://genomebiology.com/2004/5/6/225>.
- Hellen, C.U.T., 2007. Bypassing translation initiation. *Structure* 15 (1), 4–6.
- Hirochika, H., 1993. Activation of tobacco retrotransposons during tissue culture. *EMBO J.* 12 (6), 2521–2528.
- Hirochika, H., Otsuki, H., Yoshikawa, M., Otsuki, Y., Sugimoto, K., Takeda, S., 1996. Autonomous transposition of the tobacco retrotransposon Tto1 in rice. *Plant Cell* 8 (4), 725–734.
- Hohn, T., Corsten, S., Dominguez, D., Fütterer, J., Kirk, D., Hemmings-Mieszczak, M., Pooggin, M., Schärer-Hernandez, N., Ryabova, L., 2001. Shunting is a translation strategy used by plant pararetroviruses (Caulimoviridae). *Micron* 32 (1), 51–57.
- Jääskeläinen, M., Mykkanen, A.H., Arna, T., Vicent, C.M., Suoniemi, A., Kalendar, R., Savilahti, H., Schulman, A.H., 1999. Retrotransposon BARE-1: expression of encoded proteins and formation of virus-like particles in barley cells. *Plant J.* 20 (4), 413–422.
- Kuznetsov, Y.G., Zhang, M., Menees, T.M., McPherson, A., Sandmeyer, S., 2005. Investigation by atomic force microscopy of the structure of Ty3 retrotransposon particles. *J. Virol.* 79 (13), 8032–8045.
- Malik, H.S., Eickbush, T.H., 2001. Phylogenetic analysis of ribonuclease H domains suggests late, chimeric origin of LTR retrotransposable elements and retroviruses. *Genome Res.* 11 (7), 1187–1197.
- Matzke, M.A., Birchler, J.A., 2005. RNAi-mediated pathways in the nucleus. *Nat. Rev., Genet.* 6 (1), 24–35.
- Miura, A., Yonebayashi, S., Watanabe, K., Toyama, T., Shimada, H., Kakutani, T., 2001. Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. *Nature* 411 (6834), 212–214.
- Nikolaitchik, O., Rhodes, T.D., Ott, D., Hu, W.-S., 2006. Effects of mutations in the human immunodeficiency virus type 1 *gag* gene on RNA packaging and recombination. *J. Virol.* 80 (10), 4691–4697.
- Okamoto, H., Hirochika, H., 2000. Efficient insertion mutagenesis of *Arabidopsis* by tissue culture-induced activation of the tobacco retrotransposon Tto1. *Plant J.* 23 (2), 291–304.
- Peterson-Burch, B.D., Voytas, D.F., 2002. Genes of the Pseudoviridae (Ty1/copia retrotransposons). *Mol. Biol. Evol.* 19 (11), 1832–1845.
- Rogozin, I.B., Kochetov, A.V., Kondrashov, F.A., Koonin, E.V., Milanexi, L., 2001. Presence of ATG triplets in 5′ untranslated regions of eukaryotic cDNAs correlates with a ‘weak’ context of the start codon. *Bioinformatics* 17 (10), 890–900.
- Sabot, F., Schulman, A.H., 2006. Parasitism and the retrotransposon life cycle in plants: a hitchhiker’s guide to the genome. *Heredity* 97 (6), 381–388.
- Schlögelhofer, P., Bachmair, A., 2002. A test of fusion protein stability in the plant *Arabidopsis thaliana* reveals degradation signals from ACC synthase and from the plant N-end rule pathway. *Plant Cell Rep.* 21 (2), 174–179.
- Stary, S., Yin, X.-j., Potuschak, T., Schlögelhofer, P., Nizhynska, V., Bachmair, A., 2003. PRT1 of *Arabidopsis* is a ubiquitin protein ligase of the plant N-end rule pathway with specificity for aromatic amino-terminal residues. *Plant Physiol.* 133 (3), 1360–1366.
- Tanskanen, J.A., Sabot, F., Vicent, C., Schulman, A.H., 2007. Life without GAG: the *BARE-2* retrotransposon as a parasite’s parasite. *Gene* 390 (1–2), 166–174.
- Takeda, S., Sugimoto, K., Kakutani, T., Hirochika, H., 2001. Linear DNA intermediates of the Tto1 retrotransposon in gag particles accumulated in stressed tobacco and *Arabidopsis thaliana*. *Plant J.* 28 (3), 307–317.
- Teyssset, L., Dang, V.-D., Kim, M.K., Levin, H.L., 2003. A long terminal repeat-containing retrotransposon of *Schizosaccharomyces pombe* expresses a gag-like protein that assembles into virus-like particles which mediate reverse transcription. *J. Virol.* 77 (9), 5451–5463.

- Wei, W., Gilbert, N., Loon Ooi, S., Lawler, J.F., Ostertag, E.M., Kazazian Jr., H.H., Boeke, J.D., Moran, J.V., 2001. Human L1 retrotransposition: *cis* preference versus *trans* complementation. *Mol. Cell. Biol.* 21 (4), 1429–1439.
- Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P., Schulman, A.H., 2007. A unified classification system for eukaryotic transposable elements. *Nat. Rev., Genet.* 8 (12), 973–982.
- Yin, X.-J., Volk, S., Ljung, K., Mehlmer, N., Dolezal, K., Ditengou, F., Hanano, S., Davis, S.J., Schmelzer, E., Sandberg, G., Teige, M., Palme, K., Pickart, C., Bachmair, A., 2007. Ubiquitin Lysine 63 chain-forming ligases regulate apical dominance in *Arabidopsis*. *Plant Cell* 19 (6), 1898–1911.
- Zheng, H., Yu, L., Wei, C., Hu, D., Shen, Y., Chen, Z., Li, Y., 2000. Assembly of double-shelled, virus-like particles in transgenic rice plants expressing two major structural proteins of rice dwarf virus. *J. Virol.* 74 (20), 9808–9810.