Available online at www.sciencedirect.com
SCIENCE DIRECT
Plant Science 171 (2006) 308–322
PLANT SCIENCE
www.elsevier.com/locate/plantsci

Transgene structures suggest that multiple mechanisms are involved in T-DNA integration in plants

Qian-Hao Zhu [a,b], Kerrie Ramm [a,b], Andrew L. Eamens [a,1], Elizabeth S. Dennis [a,b], Narayana M. Upadhyaya [a,b,*]

[a] *CSIRO Plant Industry, Canberra, ACT 2601, Australia*
[b] *New South Wales Agricultural Genomics Centre, Wagga Wagga, NSW 2678, Australia*

Received 19 September 2005; received in revised form 20 January 2006; accepted 28 March 2006
Available online 24 April 2006

Abstract

To gain further understanding of the mechanisms involved in *Agrobacterium*-mediated genetic transformation and T-DNA integration, we analysed 156 T-DNA/rice, 69 T-DNA/T-DNA and 11 T-DNA/vector backbone (VB) junctions, which included 171 left borders (LB) and 134 right borders (RB). Conserved cleavage was observed in 6% of the LB and 43% of the RB. Terminal microhomology (1–10 bp) was identified in 58% of T-DNA/rice, 43% of T-DNA/T-DNA and 82% of T-DNA/VB junctions, and this occurred particularly at the LB junctions. Approximately 32% of both T-DNA/rice and T-DNA/T-DNA junctions harboured 1–344 bp of filler DNA that was derived mainly from the T-DNA region adjacent to the breakpoint and/or from the rice genome flanking the T-DNA integration site. Structure of the filler DNA was more complicated at the T-DNA/T-DNA junction than at the T-DNA/rice junction, indicating the presence of T-DNA recombination or rearrangement prior to or during T-DNA integration. When two T-DNAs were integrated in the inverted repeat configuration, significant truncation was always observed in one of the two T-DNAs whereas with direct repeat configuration, a large truncation was less frequent. Most integration events analysed in this study could be addressed by previously proposed models; however, the characteristics of the T-DNA repeats and the complicated filler DNA between two T-DNA copies imply that multiple mechanisms are involved in the formation of T-DNA repeats as well as in T-DNA integration in plants.

Keywords: Filler DNA; Rice; T-DNA integration; T-DNA junction; Terminal microhomology

1. Introduction

Agrobacterium tumefaciens is well known for its capacity for inter-kingdom DNA transfer [1]. It harbours a tumour-inducing (Ti) plasmid that encodes most of the major functions required for transferring an oncogenic segment of DNA, the transferred DNA (T-DNA), into the host cell [2,3]. The T-DNA itself does not include genes required for this transfer process and is delimited by two 25-bp imperfect terminal repeats, termed T-DNA left and right borders (LB and RB). The wild-type T-DNA sequence between the borders has been deleted and specifically modified to provide a range of vectors for introducing genes of interest into dicotyledonous and monocotyledonous plants, and as a genetic tool for functional genomics in plants. In view of the importance of *Agrobacterium*-mediated plant genetic engineering, a great deal of effort has been devoted to unravelling the mechanism of T-DNA integration.

T-DNA integration initiates with the induction of *Vir* genes in *Agrobacterium*. A single-stranded (ss) T-strand is released from the Ti plasmid by the border-specific action of the VirD1/VirD2 protein complex, and the VirD2 moiety covalently attaches to the 5′ end of the T-strand. The T-strand is then delivered across the bacterial envelope and into the plant cell using a type IV secretion system (T4SS). Recently, two classes of plant proteins (BTI and AtRAB8) that may participate in the

Abbreviations: ds, double-stranded; DSB, double-strand break; DSBR, double-strand break repair; LB, T-DNA left border; NHEJ, non-homologous end-joining; RB, T-DNA right border; ss, single-stranded; SSA, single-strand annealing; SDSA, synthesis-dependent strand annealing; SSGR, single-strand gap repair; T-DNA, transferred DNA; T-strand, transferred DNA strand; VB, vector backbone; Vir, virulence

* Corresponding author. Tel.: +61 2 6246 5491; fax: +61 2 6246 5000.
E-mail address: Narayana.upadhyaya@csiro.au (N.M. Upadhyaya).
[1] Present address: Department of Biology, University of York, Heslington, York YO10 5DD, UK.

0168-9452/$ – see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.plantsci.2006.03.019

initial steps of the T-DNA transfer process by interacting with the components of T4SS have been identified [4]. Once in the plant cell, the T-strand is thought to be imported into the nucleus as a DNA–protein complex containing VirD2, VirE2 [5] and their host cell interactors, such as AtKAP1 and VIP1 [6,7]. The T-strand DNA–protein complex (T-complex) is then dissociated by the action of VirF on targeted intra-nuclear proteolysis of AtVIP1 and VirE2 [8]. Finally, the T-DNA is integrated into the plant genome via single-stranded or double-stranded intermediate [3,9,10].

Based on the characteristics of the junctions and structure of the integrated T-DNA, it has been shown that T-DNA integration in plants is a result of illegitimate recombination [11,12]. Two models have been proposed for T-DNA integration: the single-strand gap repair (SSGR) model and the double-strand break repair (DSBR) model [12].

The SSGR model suggests that T-DNA integration initiates with the production of a nick in the bottom strand of the target plant DNA, which then leads to a gap. Both ends of the ss T-strand anneal to the ss upper strand of the target plant DNA within any microhomologous regions and non-annealed overhangs of the T-strand are then removed. Following integration of the T-strand into the target plant DNA, a nick in the upper strand of the target plant DNA is produced and extended to a gap. The integration is completed by gap repair using the inserted T-strand as a template [12]. The SSGR model was further revised to suggest that the integration is initiated by microhomology between the 3′ end, or the region adjacent to the 3′ end, of the T-strand and the upper strand of the target plant DNA. This is followed by annealing of the VirD2-attached 5′ end of the T-strand to the 3′ end of the target plant DNA by microhomology [13]. This model relies on microhomology and the ligase function of VirD2 within the plant cells. SSGR and the revised SSGR models explain the integration events with terminal microhomology at the LB/plant junction and with or without terminal microhomology at the RB/plant junction.

The DSBR model initially suggested that T-DNA integration results from annealing of the 5′ and 3′ overhangs of a ds T-DNA intermediate to the ss ends of a double-strand break (DSB) in the plant DNA in areas of microhomology [12]. This model explains the integration events in which terminal microhomology are observed at both the LB/plant and the RB/plant junctions. Compelling evidence supports DSBR playing an important role in T-DNA integration in plants. For example, low-dose X-ray irradiation, a cause of DSB, enhances transformation rates in plants [14]. T-DNA is found to be captured at DSB in tobacco and transgene integration is increased when DSBs are intentionally induced [15]. ss T-strand is thought to be the main intermediate for T-DNA integration, but recent reports have shown that ss T-strand molecule can be converted to ds T-DNA before its integration into a DSB [9,10]. Accordingly, the DSBR model was extended based on synthesis-dependent strand annealing (SDSA) to account for the T-DNA integration events in which both ends of the T-DNA are partly truncated and terminal microhomology exists between the T-DNA borders and the pre-insertion site [15]. The filler DNA found at the T-DNA/plant junction has been explained by abortive gap repair through a SDSA mechanism [16]. Like the effects of DSBR via non-homologous end-joining (NHEJ) in plant cells [17], the complex filler DNA structure at the T-DNA/plant junction was postulated to result from multiple template switching by the free 3′ end of T-DNA or genomic DSBs [16].

These models together explain the characteristics of single-copy T-DNA integration, but they do not address the formation of the complex T-DNA repeats that are frequently observed in transgenic plants. Based on the finding of precise fusion between two RBs in some T-DNA repeats and the assumption that ss T-strands cannot recombine in head-to-head configuration [18,19], a model via a ds T-DNA intermediate has been proposed to address the formation of T-DNA repeats in which two or more T-DNA copies, with or without filler DNA between them, are arranged in direct or inverted orientation [18,19]. According to this model, T-DNA recombination occurs prior to T-DNA integration. This model is supported by the result that T-DNAs derived from different *Agrobacterium* strains can co-integrate at a single locus [20], but a two-phase integration mechanism in which a primary T-DNA integration event stimulates further T-DNA integration at the same locus, or interaction of two simultaneously integrated T-DNAs cannot be ruled out.

To further understand the mechanisms involved in T-DNA integration, we sequenced and analysed 236 T-DNA/rice, T-DNA/T-DNA and T-DNA/VB junctions isolated from independent transgenic rice plants. We found that the characteristics of all T-DNA/rice junctions were similar to previously reported results in rice and other plants [21–23], and these integration events could be explained by SSGR and/or DSBR models [12,13,15,16,24]. However, new features were found at the junction of two T-DNA copies, in which one of the two T-DNA copies truncated significantly, especially in inverted T-DNA repeats, and a more complicated filler DNA structure was frequently observed at the junction between two linked T-DNA copies. These results suggest that multiple mechanisms may be involved in the formation of the complex filler DNA structures and T-DNA repeats.

2. Materials and methods

2.1. Plant materials

Independent transgenic rice (*Oryza sativa* L. cv. Nipponbare) plants, which were transformed with binary vectors pSK100/200, pEU334NA/NB (GenBank Acc. No. AY488510 or AY488511) or pNU393B2 (GenBank Acc. No. DQ225749) (Fig. 1) [23,25,26], were selected based on Southern blot results (data not shown). Detailed transformation and tissue culture procedures have been previously described [27]. Regenerants were grown under controlled glasshouse conditions at 25 ± 3 °C and 16 h of light.

2.2. Genomic DNA extraction and T-DNA flanking sequence isolation

Genomic DNA was extracted from leaf samples of transformants using the PureGene total nucleic acid isolation