Theor Appl Genet (2004) 109: 815–826
DOI 10.1007/s00122-004-1692-y

ORIGINAL PAPER

A. S. Afolabi · B. Worland · J. W. Snape · P. Vain

A large-scale study of rice plants transformed with different T-DNAs provides new insights into locus composition and T-DNA linkage configurations

Received: 30 October 2003 / Accepted: 29 March 2004 / Published online: 15 May 2004

Abstract Transgenic locus composition and T-DNA linkage configuration were assessed in a population of rice plants transformed using the dual-binary vector system pGreen (T-DNA containing the *bar* and *gus*-genes)/pSoup (T-DNA containing the *aphIV* and *gfp* genes). Transgene structure, expression and inheritance were analysed in 62 independently transformed plant lines and in around 4,000 progeny plants. The plant lines exhibited a wide variety of transgenic locus number and composition. The most frequent form of integration was where both T-DNAs integrated at the same locus (56% of loci). When single-type T-DNA integration occurred (44% of loci), pGreen T-DNA was preferentially integrated. In around half of the plant lines (52%), the T-DNAs integrated at two independent loci or more. In these plants, both mixed and single-type T-DNA integration often occurred concurrently at different loci during the transformation process. Non-intact T-DNAs were present in 70–78% of the plant lines causing 14–21% of the loci to contain only the mid to right border part of a T-DNA. In 53–66% of the loci, T-DNA integrated with vector backbone sequences. Comparison of transgene presence and expression in progeny plants showed that segregation of the transgene phenotype was not a reliable indicator of either transgene inheritance or T-DNA linkage, as only 60–80% of the transgenic loci were detected by the expression study. Co-expression (28% of lines) and backbone transfer (53–66% of loci) were generally a greater limitation to the production of marker-free T_1 plants expressing the gene of interest than co-transformation (71% of lines) and unlinked integration (44% of loci).

Communicated by C. Möllers

A. S. Afolabi · B. Worland · J. W. Snape · P. Vain (✉)
John Innes Centre, Norwich Research Park,
Norwich, NR4 7UH, UK
e-mail: philippe.vain@bbsrc.ac.uk
Tel.: +44-1603-450612
Fax: +44-1603-450023

Introduction

Transgenic technologies play a key role in plant molecular genetic studies and crop improvement strategies, potentially offering world-wide benefits. However, important limitations still lie in the uncontrolled factors affecting the integration and behaviour of transgenes. Strategies to control transgene integration (Paszkowski et al. 1988; Albert et al. 1995; Terada et al. 2002) and to limit unwanted or multiple integrated DNA sequences (Hanson et al. 1999; Srivastava et al. 1999; Lu et al. 2001; Cotsaftis et al. 2002) are central to the development of advanced plant transformation technologies. Among these, the production of marker-free transgenic plants is particularly important. Selectable marker genes required for the transformation process are generally unwanted, especially in sequential transformation experiments or in transgenic crops.

To date, several approaches have been developed to remove or eliminate selectable marker genes from transgenic plants (for reviews see Yoder and Goldsbrough 1994; Ebinuma et al. 2001; Hare and Chua 2002): (1) co-integration of the gene of interest and the selectable marker gene at unlinked loci (De Framond et al. 1986), (2) excision of the selectable marker gene by site-specific recombination (Dale and Ow 1990), (3) transposon-mediated repositioning of the gene of interest (Goldsbourgh et al. 1993), and (d) excision of the selectable marker gene by homologous recombination (Zubko et al. 2000). Strategy (1) was the first to be implemented in plants (de Framond et al. 1986) and has been used in a wide range of species (Mc Knight et al. 1987; De Neve et al. 1997) including crops (De Block and Debrouwer 1991; Komari et al. 1996; Xing et al. 2000). It generally relies on *Agrobacterium tumefaciens*-mediated co-transformation of different transgenes carried by different T-DNAs. The different T-DNAs can either been present in one vector/one strain (Komari et al. 1996; Xing et al. 2000), two vectors/one strain (De Framond et al. 1986; Daley et al. 1998) or two vectors/two strains (McKnight et al. 1987; De Block and Debrouwer 1991; De Neve et al. 1997). This

multiple T-DNA approach differs from the other three in that unlinked integration of the different transgenes is achieved directly during the plant cell transformation process and does not require further transgenic locus alteration. The simplicity of this system is a great advantage but it also bears some limitations. The characteristics of the original integration sites may include unwanted DNA sequences (T-DNA and vector sequences including borders) or rearrangements (plant genomic DNA, repeats or truncated T-DNA) which are retained (Cotsaftis et al. 2002). The efficiency of this strategy also relies on the understanding of the mechanisms underlying T-DNA integration and T-DNA linkage in plants. Although the integration of T-DNAs in all possible configurations has been well documented (Jorgensen et al. 1987; Koncz et al. 1994; De Neve et al. 1997), to date there is only partial information available on transgenic locus composition and T-DNA linkage configuration in populations of transformed plants. Most previous studies on T-DNA linkage, with few exceptions (Bhattacharyya et al. 1994; Cluster et al. 1996; Lu et al. 2001; Matthew et al. 2001), were based on the molecular analysis of primary transformants (Southern blot analysis of T_0 plants) combined with the segregation analysis of the transgene phenotype. Previously, we showed that such an approach is prone to underestimate transgenic locus number and therefore to present a distorted view of T-DNA linkage groups in plants (Vain et al. 2003). Recently, it has been shown that 30–60% of the plants transformed with different T-DNAs can contain several types of T-DNA inserts (Lu et al. 2001; Miller et al. 2002) with some inserts containing either linked or unlinked T-DNAs or a mixture of both. The further development of the multiple T-DNA strategy relies, therefore, on a better understanding of locus constitution and T-DNA linkage in populations of transgenic plants.

In cereals, the development of marker-free transgenic plants has been hampered because efficient *Agrobacterium*-based transformation technologies have only recently been developed (Hiei et al. 1994). The multiple T-DNA approach was also the first to be implemented in cereals (Komari et al. 1996) and is now used to produce marker-free transgenic rice (Komari et al. 1996; Lu et al. 2001), barley (Matthew et al. 2001) and maize (Miller et al. 2002) plants. Rice plants containing only a rice ragged stunt virus resistance transgene have been produced using the one vector/one strain version of this strategy (Lu et al. 2001). The multiple T-DNA approach is likely to be one of the preferred type of "clean-gene" technology to produce transgenic cereal crops free of selectable marker genes in the near future.

In this study, we have characterised the integration profiles of different T-DNAs in populations of transgenic rice plants. We used the dual-binary vector system pGreen/pSoup (Hellens et al. 2000) to carry the two different T-DNAs in a single *Agrobacterium* strain (two vectors/one strain approach). pGreen is a small (3 kb) Ti binary vector able to replicate in *Escherichia coli* but unable to replicate in *Agrobacterium* without the presence of another binary plasmid, pSoup, in the same strain. pSoup provides replication functions in *trans* for pGreen (Hellens et al. 2000). Information on the pGreen/pSoup system can be found on the internet site (http://www.pgreen.ac.uk). In the past, the pGreen/pSoup system has mostly been used for the transformation of dicotyledonous (Hellens et al. 2000) or monocotyledonous species (Vain et al. 2003) using a single T-DNA present in either the pGreen or pSoup vector. In this study, we have transformed rice with both pGreen and pSoup vectors each containing a different T-DNA. Co-transformation and co-expression were quantified during single or dual T-DNA selection. Transgenic locus composition and T-DNA linkage configuration were determined by assessing the presence and expression of the four transgenes in large populations of T_0plants and their progeny. The contribution of the pGreen/pSoup dual-binary vector system to the understanding of T-DNA integration and to the development of "clean-gene" technology (Mc Cormac et al. 1999) in rice is discussed.

Materials and methods

PGreen/pSoup-based vectors

The binary vectors and *Agrobacterium* strain used for rice transformation are shown in Fig. 1. pRT18 (pGreen-based) contained the *gus* gene (plus intron) and the *bar* gene both driven by the maize 5′ ubiquitin region and the nopaline synthase terminator (Vain et al. 2003). pRT47 (pSoup-based) contained the *aphIV* gene (plus intron in 5′ UTR) and the *gfp* genes both driven by the CaMV35S promoter and the nopaline synthase or a soybean polyA terminator (Vain et al. 2003). Neither pRT18 nor pRT47 are super binary vectors (Hiei et al. 1994) as no additional virulence genes are present in the vector backbone.

Plasmids were transformed into *E. coli* strain DH5α using the PEG-transformation technique and into *Agro-*

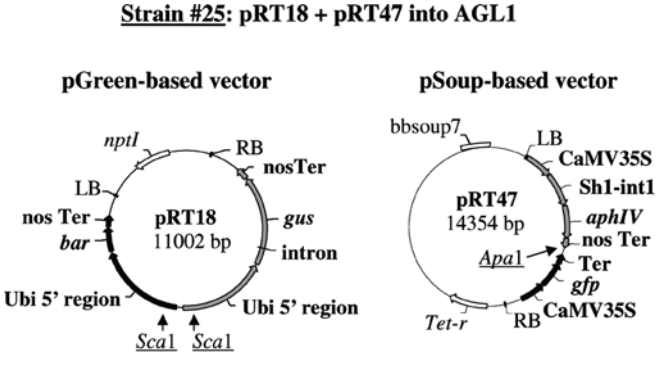

Fig. 1 pGreen/pSoup-based vectors used for rice transformation. Gene abbreviations are detailed in the experimental procedures. Strain No. 25 contains pRT18 and pRT47 binary vectors. The T-DNA of pRT18 (pGreen-based) contains the *bar* and *gus* genes. The T-DNA of pRT47 (pSoup-based) contains the *aphIV* and *gfp* genes. Probes for backbone sequences (*npt1* and BBsoup7) approximately 1 kb from the LB of each T-DNA are represented. In *Agrobacterium tumefaciens* pSoup-based vectors provide replication in *trans* for pGreen-based vectors (Hellens et al. 2000)

bacterium strain AGL1 using a freeze-thaw technique. As detailed in Fig. 1, strain no. 25 contains pRT18 (pGreen-based) and pRT47 (pSoup-based).

Rice transformation procedures

Embryogenic calli derived from mature seeds of rice (*Oryza sativa* L.) variety Nipponbare were used for transformation as previously described (Vain et al. 2003). Briefly, the embryos were aseptically removed from sterilised seeds and plated onto NBm medium for 3 weeks in the dark at 25°C. Loose embryogenic translucent globules (U), around 1 mm in size, were separated by rolling the callus grown from the original embryo onto the gelling agent. Globules were cultured for an additional 10 days on fresh NBm medium (~100 globules per plate) to produce embryogenic nodular units (ENUs; Bec et al. 1998), used as targets for transformation. Culture plates containing ENUs were flooded with bacterial suspension at an OD=1 (600 nm) for 5 min. ENUs were blotted and co-cultivated on NBm medium supplemented by 200 μM acetosyringone for 2 days. After co-culture, ENUs were put onto selection medium (NBm medium containing 5 mg/l phosphinotrycin (PPT) and/or 50 mg/l hygromycin for 2 weeks, then subcultured for an additional 3 weeks. Timentin (150 mg/l) was added to all selection media and L-glutamine was removed from culture media with PPT. Plants were regenerated by successive transfer onto pre-regeneration (PRm), regeneration (RNm) and germination selection media (Vain et al. 2003). Only one plant was regenerated from each original ENU (explant) to guarantee that each plant represented an independent transformation event. Transformed plants were transferred to a controlled environment room for growth to maturity. All transgenic plants produced were analysed to ensure the study of series of random, independent transformation events, with the widest spectrum of expression for the non-selected *gus* and *gfp* genes. Transgenic plants were cultured, sampled and analysed using a standard operating procedure (SOP, James et al. 2004).

Analysis of GUS and GFP activity

Fluorometric analysis for β-glucuronidase activity was carried out on leaf tissue from rice plants at the five-leaf stage, following the protocol of Jefferson (1987) and SOPs (James et al. 2004). Fluorescence was measured using a Titertek Fluoroskan II after 0, 30 and 60 min incubation. Each assay was performed in triplicate. Protein content was determined using a Bio-Rad protein assay kit. Data were expressed as pmol of 4-methylumbelliferone (MU) min^{-1} mg^{-1} of extracted protein. The background activity (33±4 pmol MU min^{-1} mg^{-1}protein) was subtracted from all fluorometric GUS measurements as previously described (Vain et al. 2002).

Visual detection of GFP fluorescence was performed using a MZ6 Leica dissecting microscope with a fluorescent module (Leica no. 10 446093) and appropriate wavelength filters (425/60 nm excitation filter, 470 nm dichromatic beam splitter and a G6457 emission barrier filter) over a high voltage mercury lamp.

Detection of transgenic plants by the polymerase chain reaction

DNA was isolated and polymerase chain reaction (PCR) reactions were carried out as previously described (Vain et al. 2002). Five primers sets were used on the DNA samples: (1) one to amplify the 1,200 bp single copy rice RFLP probe C213 (forward: 5′-AAAGGACCGGAATGACCACAA-3′; reverse: 5′-GAATGAACCACGCCCAAGAGT-3′) in order to ensure that each DNA sample was suitable for PCR amplification, (2) another to amplify a 421 bp fragment of the *bar* gene (forward: 5′-GGTCTGCACCATCGTCAACC-3′; reverse: 5′-GTCATGCCAGTTCCCGTGCT-3′), (3) another to amplify a 1,013 bp UBI:: *gus* fragment (forward: 5′-GGGCGGTCGTTCATTC-3′; reverse: 5′-TTCGGCGTGGTGTAGAGC-3′), (4) another to amplify a 727 bp fragment of the *aphIV* gene (forward: 5′-ACTCACCGCGACGTCTGTCG-3′; reverse: 5′-GCGCGTCTGCTGCTCCATA-3′) and (5) another to amplify a 527 bp fragment of the *gfp* gene (forward: 5′-GGAGAGGGTGAAGGTGATGCAA-3′; reverse: 5′-GGGCAGATTGTGTGGACAGGTA-3′).

Southern analysis

Genomic DNA extraction and Southern analyses were performed on primary transformed rice plants (T_0) as previously described (Vain et al. 2003). Membranes were hybridised with the following probes: 549 bp of the *bar* gene, 701 bp of the *gus* gene, 981 bp of the *aphIV* gene, 737 bp of the *gfp* gene, 421 bp of the pGreen backbone (888 bp from the left border) or 464 bp of the pSoup backbone (1,216 bp from the left border). The filters were analysed by using a Typhoon 8600 phospho-imager and the Typhoon Scanner Control version 1.0/ImageQuant version 5.1 software (Molecular Dynamics).

Dot blot analysis

The Qiagen DNeasy 96 high-throughput DNA isolation kit was used for genomic DNA extraction from rice plants. The dot blotting procedure was performed using 190 μl of the extracted DNA (approximately 1 μg) and the BIORAD Bio-Dot Microfiltration apparatus following the manufacturer's protocols and recommendations. Membranes were hybridised sequentially with probes for the *bar*, *gus, aphIV* and *gfp* genes, for the pGreen and pSoup backbones (as

described above) and for the R2272 rice RFLP genomic probe.

Transgene inheritance and segregation of transgene phenotypes

T_1 seeds were obtained by self-pollination of primary transformed rice (T_0) plants. Segregation analyses were conducted by germinating seeds on MSR6 medium (Vain et al. 2002) without selection. Transgene expression in the T_1 embryos or seedlings was assessed qualitatively by histochemical GUS staining (Jefferson et al. 1987), or by observation of GFP fluorescence in at least 64 random T_1 seedlings from each independently transformed T_0 plant. All T_1 plant progeny not expressing either the *gus* or *gfp* transgenes were analysed by PCR or dot blot for the presence of the *bar*, *gus, aphIV* and *gfp* genes. When PCR or dot blot reactions were negative for the transgene, an additional PCR analysis was conducted for the presence of the C213 or R2272 RFLP rice probes to confirm that the DNA extraction was suitable for PCR amplification or dot blot analysis.

Statistical analyses

Statistical analyses, following the requirements of each test, were performed using Minitab 13.1 or Genstat 5 software. Data sets were compared using ANOVA or the paired *t*-test. Observed segregation ratios were compared to Mendelian models using χ^2 analysis.

Results and discussion

Production of a population of transgenic rice plants using single or dual T-DNA selection

Rice plants were transformed using the *Agrobacterium* strain no. 25 containing the pGreen-based pRT18 vector (T-DNA containing the *bar*and *gus* genes) and the pSoup-based pRT47 vector (T-DNA containing the *aphIV* and *gfp*genes) (Fig. 1). Three selection regimes were applied during the transformation process with strain no. 25: (1) dual selection of pGreen and pSoup T-DNAs using phosphinothricin (PPT) and hygromycin, (2) selection of pSoup T-DNA only, using hygromycin alone, (3c) selection of pGreen T-DNA only, using PPT alone. The dual T-DNA selection regime was used to produce a large and random population of plants co-transformed and co-expressing all the transgenes present in the pGreen and the pSoup T-DNAs. This strategy was designed to by-pass the limitations of post-transformation screening for co-transformed and co-expressing lines in experiments designed to produce marker-free transgenic plants. Later on, the single T-DNA selection regimes were used to directly mimic experiments designed to produce marker-free transgenic plants. More than 400 independently transformed transgenic rice plant lines were produced (Table 1). Transgene copy numbers and the presence of backbone sequences were determined in 62 random T_0 plants, each representing an independently transformed plant line (50 lines with dual T-DNA selection, nine lines with pSoup T-DNA selection only, three lines with pGreen T-DNA selection only). Transgenic locus composition and T-DNA linkage configuration were determined in these 62 lines by assessing the presence and expression of the four transgenes in around 4,000 T_1 progeny plants. In fine T_0 and T_1 molecular and expression data were combined to provide an accurate picture of transgene structure and behaviour in populations of plant lines.

Dual T-DNA selection: transgene structure and expression in T_0 rice plants

Dual selection of pGreen and pSoup T-DNAs was undertaken using PPT (selecting for the *bar* gene in pRT18) and hygromycin (selecting for the *aphIV* gene in pRT47). As expected, after dual selection, 100% of the plant lines (208/208) were co-transformed for both T-DNAs (dot blot/PCR data not shown). Around 70% of the lines (146/208) co-expressed all four transgenes. The remaining lines were either not expressing *gfp* (24%), or *gus* (4%) or both unselected genes (2%). Fifty independent lines, co-transformed and co-expressing all transgenes present in the pGreen T-DNA (*bar* and *gus*) and the pSoup

Table 1 Summary of results of transformation experiments using different T-DNA selection regimes

	Selection regime		
pGreen-based T-DNA selected	pRT18	pRT18	–
pSoup-based T-DNA selected	–	pRT47	pRT47
No. of independent plant lines produced	10	208	185
No. of co-tranformed plant lines	80% (8/10)	100% (208/208)	71% (131/185)
No. of plant lines co-expressing the transgenes	88% (7/8)	70% (146/208)	24% (32/131)
No. of plant linef studied in T_1	3	50	9
Unlinked lines	100% (3/3)	48% (24/50)	44 % (4/9)
Unlinked loci	63% (5/8)	44% (43/98)	33% (7/21)
"G-S locus" (linked pGreen and pSoup T-DNAs)	38% (3/8)	56% (55/98)	67% (14/21)
"G locus" (pGreen T-DNA alone)	50% (4/8)	31% (30/98)	14% (3/21)
"S locus" (pSoup T-DNA alone)	12% (1/8)	13% (13/98)	19% (4/21)

T-DNA (*aphIV* and *gfp*), were randomly chosen. Transgene copy number and the presence of vector backbone was determined in these 50 lines by Southern blot hybridisation using a restriction enzyme that cuts in the middle of each T-DNA (*Sca*I for pRT18, *Apa*I for pRT47, Fig. 1) and using probes for either the right inner T-DNA border (*gus* in pRT18, *gfp* in pRT47) or the left inner T-DNA border (*bar* in pRT18, *aphIV* in pRT47) or backbone sequences (*npt1* in pRT18 and *BBsoup7* in pRT47). Transgene copy number was estimated from the number of integrated fragments (i.e., the number of bands in the hybridisation pattern, Fig. 2). When band intensity was taken into account, in addition to band number, this did not significantly change copy number determination (ANOVA, n=600, P=0.79).

The distribution profile of the copy number for each of the four transgenes was similar to those reported previously when pRT18 and pRT47 were used individually (not in combination) to transform rice (Vain et al. 2003). The four transgenes integrated at different frequencies into the plant genome. On average, transgenic rice plant lines contained 3.9 copies of *gus*, 2.7 copies of *bar*, 1.3 copies of *aphIV* and 2.5 copies of *gfp* genes. The *gus*gene integrated significantly more frequently than the *gfp* and *bar* genes; themselves integrating more often than the *aphIV* gene (ANOVA, n=200, P=0.001). Selectable marker genes *bar* and *aphIV* (present at the left inner border of the pRT18 and pRT47 T-DNAs, respectively) were significantly less present in the rice genome than the corresponding non-selected reporter genes *gus* and *gfp* (present at the right inner border of the pRT18 and pRT47 T-DNAs, respectively) (ANOVA,n=200, P<0.001). Depending upon the transgene studied, 30% (*bar* gene), 10% (*gus* gene), 28% (*gfp* gene) or 72% (*aphIV* gene) of the independently transformed plant lines contained a single transgene copy. As previously reported (Bhattacharyya et al. 1994; Ishida et al. 1996; Dong et al. 2001; Sallaud et al. 2003; Vain et al. 2003), the number of copies of transgenes originally present in the same T-DNA were often different in individually transformed plant lines (paired *t*-test,

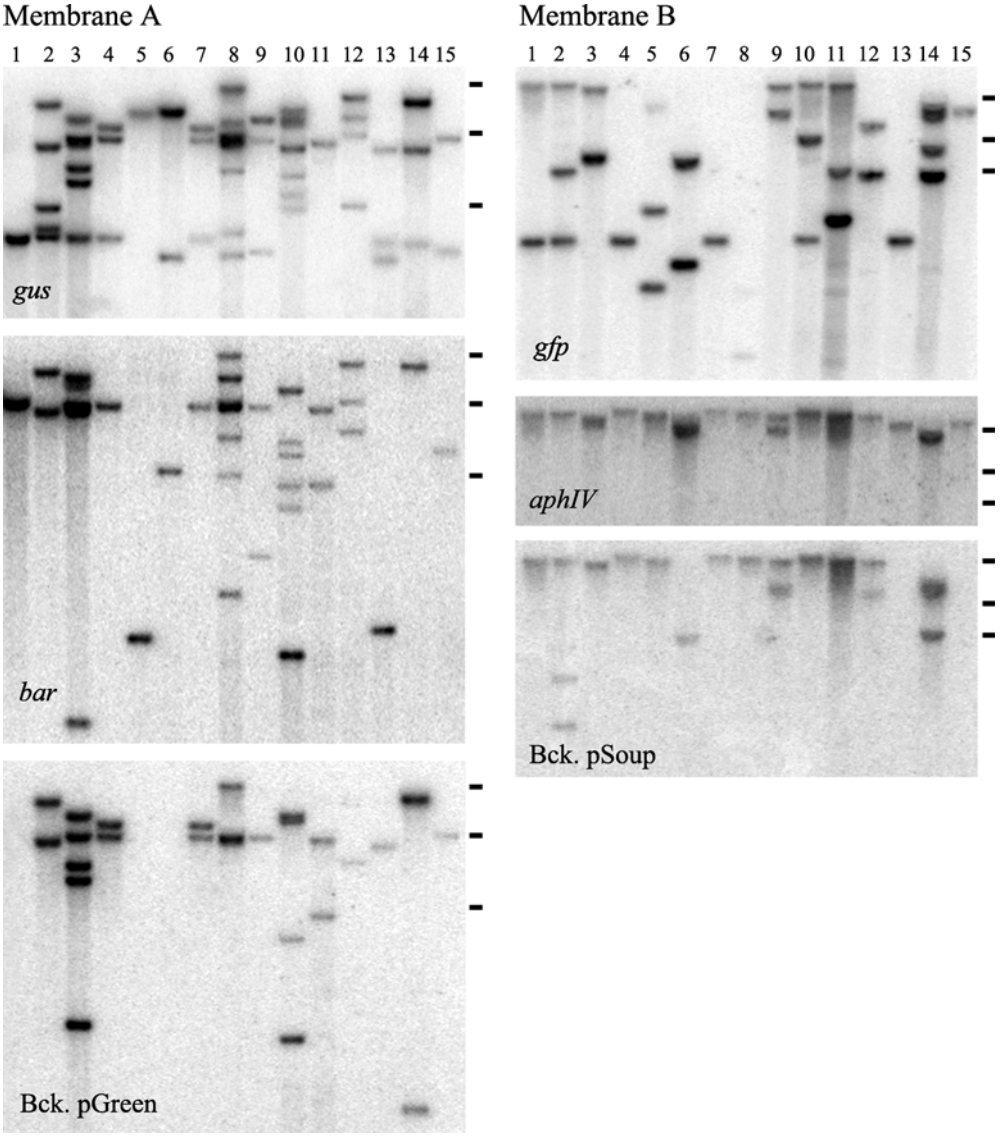

Fig. 2 Southern blot analysis of 15 T_0 rice plants independently transformed with *Agrobactrium* strain no. 25 (pRT18+pRT47 in AGL1). Membrane A (*left column*): Plant genomic DNA was digested using restriction enzyme *Sca*I (cutting in the middle of pRT18 T-DNA, see Fig. 1) and was probed with sequences from either the *gus* gene (*right border*) or the *bar* gene (*left border*) or the pGreen backbone. Membrane B (*right column*): Plant genomic DNA was digested using restriction enzyme *Apa*I (cutting in the middle of pRT47 T-DNA, see Fig. 1) and was probed with sequences from either *gfp*gene (*right border*) or the *aphIV* gene (*left border*) or pSoup backbone. Ladder (*right of pictures*) from top to bottom: 23.1, 9.4 and 6.6 kb

$P<0.05$, copy number *gus* versus *bar* and copy number *aphIV* versus *gfp*). Around 78% (39/50) of the lines contained different copy numbers for *gus* and *bar* genes (both present in pRT18 T-DNA) and 70% (35/50) of the lines contained different copy numbers for *aphIV* and *gfp* genes (both present in pRT47 T-DNA). These disparities could result from rearranged or truncated T-DNAs (Jorgensen et al. 1987; De Neve et al. 1997; Sallaud et al. 2003). Deletion or inverted repeat at the left T-DNA border could be responsible for the deficit of left-border selectable marker genes observed in this study. In the past T-DNA have been reported to be linked in all possible configurations (Jorgensen et al. 1987; Koncz et al. 1994) with a preference for at least one right border involved in the linkage (De Neve et al. 1997). Frequent T-DNA alteration can potentially generate a wide range of transgenic loci in the plant genome. Some loci could contain as little as only part of a given T-DNA, while other loci could contain one or more intact or rearranged copies of the T-DNAs. In the subset of 24 lines where all transgenes were integrated at a unique locus in the genome (as detailed in the next section), several copies of the same transgene were generally present. The plants contained on average 3.7 copies of the *gus*, 2.3 copies of the *bar*, 1.3 copies of the *aphIV* and 2.2 copies of the *gfp* genes. This profile of copy number distribution was nearly identical to that observed in 26 multilocus lines (containing on average 3.8 copies of the *gus*, 2.7 copies of the *bar*, 1.3 copies of the *aphIV* and 2.8 copies of the *gfp* genes ; ANOVA, $n=200$, $P=0.249$). This suggests that multilocus lines should contain, on average, less transgene copies per individual locus. The lower ratio of genes per locus observed in the multilocus lines could be due, in part, to the presence of loci containing only one type of T-DNA or only T-DNA fragment(s). Overall, all these data suggest a high level of heterogeneity between loci, even when present in a single plant genome.

As previously reported (Vain et al. 2003), backbone transfer was frequent in plant lines transformed with either pGreen-based, pRT18 (35/50 lines=70%) or with pSoup-based, pRT47 (40/50 lines=80%) vectors. In 8% (4/50) of the transgenic plant lines, no backbone sequences from either pRT18 or pRT47 were detected. Further transgene inheritance studies allowed the precise determination of transgenic locus number and composition.

Dual T-DNA selection: transgene inheritance and segregation of transgene phenotype

The presence and expression of transgenes were assessed in around 64 T_1 plants for each of the 50 co-transformed and co-expressing lines studied. A total of 3,109 T_1plants were scored for the expression of the non-selected *gus* (in pRT18) and *gfp* (in pRT47) genes (i.e., more than 6,200 phenotyping analyses). T_1 plants not expressing the transgenes were genotyped by dot blot or PCR analysis for the presence of the *bar*, *gus*, *aphIV* and *gfp* genes as well as for the genomic rice probe C213 (i.e., more than 3,340 genotyping analyses). Dot blot analysis for the four transgenes allowed clear identification of which T-DNA was present in each T_1 plant (Fig. 3). In fine each T_1 plant was scored as containing, or not, pRT18 T-DNA or pRT47 T-DNA, as well as expressing, or not, the unselected transgenes present in these T-DNAs. The set of 64 T_1 plants scored for each independent line allowed the calculation of an observed ratio for T-DNA inheritance and for the segregation of transgene expression.

Transgene inheritance was studied using the genotyping data of the 3,109 T_1 plants. The observed ratios for each line were compared statistically to those of 30 theoretical models representing all possible linkage configurations of pRT18 (pGreen-based) and pRT47 (pSoup-based) T-DNAs in up to four Mendelian loci. In the theoretical models, a locus containing at least one copy of pRT18 T-DNA and no pRT47 T-DNA was named a “G locus”, a locus containing at least one copy pRT47 T-DNA and no pRT18 T-DNA was named a “S locus” and a locus containing at least one copy of both pRT18 and pRT47 T-DNAs was named a “G-S locus”. For each independent line, a three step analysis was undertaken to compare the observed ratios to the ratios from the 30 theoretical models (1) elimination of models predicting the absence of a given type of progeny when some where observed among the T_1 plants (2) ranking all possible models according to their probability (using χ^2 analysis). In total, more than 6,000 individual tests were conducted to identify all probable models for each of the 50 independently transformed plant lines. The final step of the analysis (3) was to compare each probable model to the T_0 molecular data (i.e., *bar*, *gus*, *aphIV* and *gfp* gene copy numbers) and to the number of active loci determined by the segregation of transgene phenotype (as detailed in next section). For 54% of the

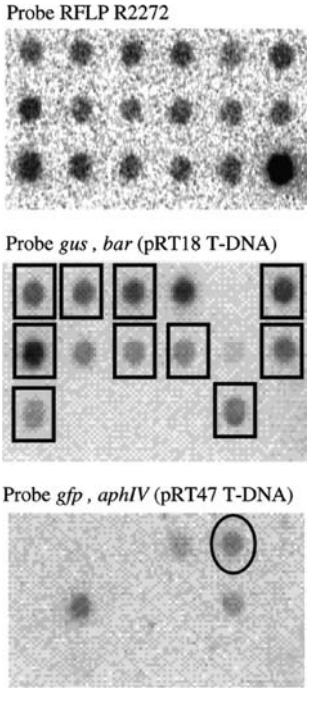

Fig. 3 Dot blot analysis of T_1 rice plants produced by self pollination of a T_0 plant transformed with *Agrobactrium* strain no. 25 harbouring pRT18 (containing the *bar* and *gus* genes) and pRT47 (containing the *aphIV* and *gfp* genes) in AGL1. The membrane was probed with the R2272 rice genomic RFLP probe, then probes for the *bar, gus, aphIV* and *gfp* genes. The *dot* in right bottom corner contains wild-type rice DNA. The *rectangles* indicates T_1 plants containing pRT18 and not pRT47 T-DNAs, *Oval* indicates T_1 plants containing pRT47 and not pRT18 T-DNAs

lines (27/50) only one model was possible and fitted all other data. For 22% of the lines (11/50) the most probable model (generally with half the χ^2 value of the next probable model) fitted all other data and was used. For the remaining 24% (12/50) the most probable model did not fit all other data, therefore one of the alternative probable models, with the best fit to transgene copy and active locus numbers for all four transgenes was used (Fig. 4). Only 14 of the 30 theoretical models corresponded to the 50 transformation events obtained in this part of the study. It is likely that additional transformation events would have fitted other models (see in last section of manuscript where T-DNA integration corresponded to other theoretical models: M10, M12 and M16). This suggests that many (if not all) T-DNA linkage configurations can be obtained during the transformation process.

The analysis was then conducted using the phenotyping data of the T_1 plants. Segregation of transgene expression using the same 3,109 T_1 plants showed striking differences to the transgene inheritance patterns previously based upon the presence of the transgene in the same progenies. Such differences were mostly due to T_1 plants containing but not expressing some transgenes. Segregation of transgene expression results matched those of the transgene inheritance study (above) in only 22% of the lines. The remaining 78% of the lines were either attributed to incorrect integration models (34% of the lines) or no model could fit the observed segregation ratios of transgene expression (44% of the lines). Around 40% of the multilocus lines were underestimated for locus number when using expression (instead of transgene presence) to conduct inheritance studies. Around 20–40% of the loci containing pRT18 and pRT47, respectively, were undetected by the expression study in the multilocus lines. This confirms that segregation of transgene phenotype should not be used as an indicator of transgene inheritance (Vain et al. 2002, 2003) and therefore is inadequate for loci number determination, identification of T_1plants free of selectable marker genes or for assessing T-DNA linkage. In the past, and with very few exceptions (Lu et al. 2001; Matthew et al. 2001), most studies on plant transformation with different T-DNAs relied mainly on progeny phenotyping to characterise T-DNA linkage groups in populations of transgenic plants. Our study shows that such an approach is prone to give a distorted view of T-DNA linkage configuration as it generally underestimates the number of transgenic loci in T_0 plants as well as overestimating T-DNA non-linkage in progenies (due to T_1 plants containing but not expressing the transgenes).

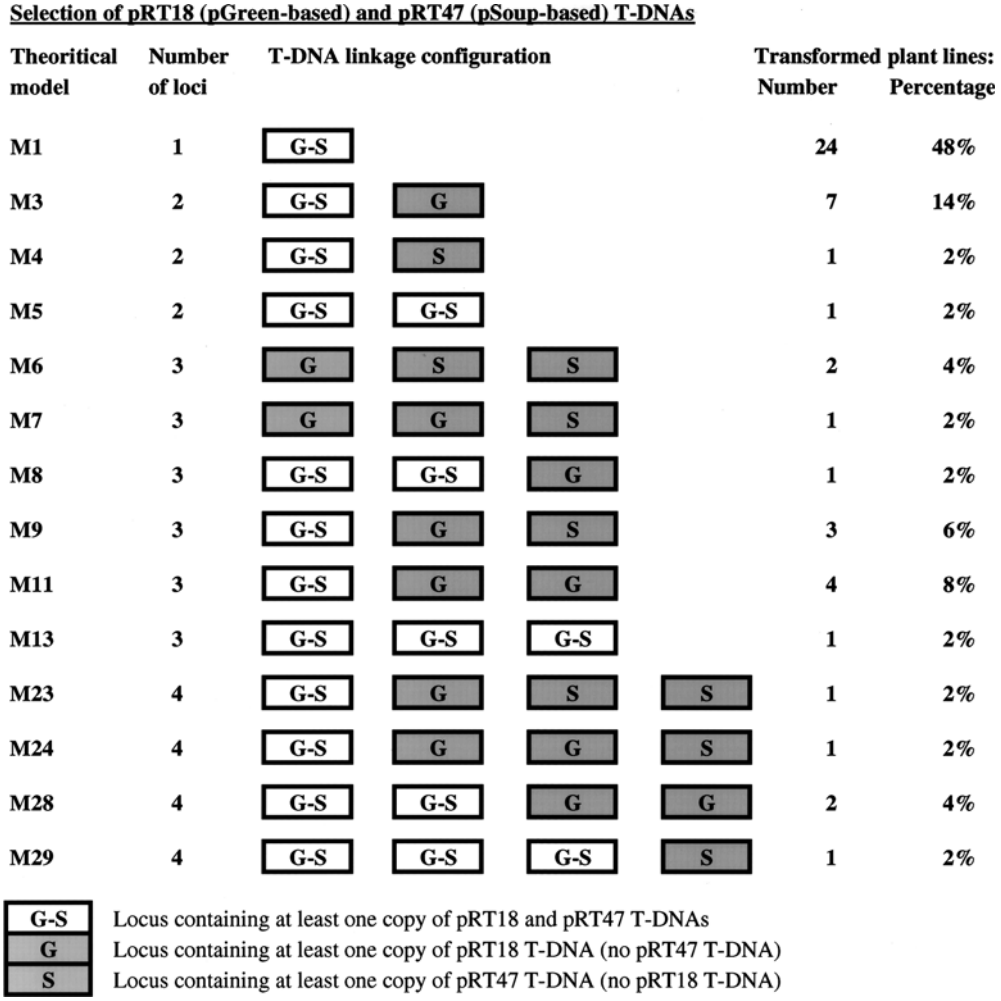

Fig. 4 T-DNA linkage configurations in a population of 50 independent rice plant lines co-transformed and co-expressing the transgenes (produced using dual T-DNA selection). Locus constitution was determined by assessing transgene presence and inheritance in 50 T_0 and 3,109 T_1 plants. Only 14 of 30 theoretical models, representing all T-DNA linkage configurations of pRT18 and pRT47 up to four independent Mendelian loci, corresponded to the independent plant lines produced

Across two generations, the transgene copy number had no significant effect on the level of transgene silencing of the unselected reporter *gus* and *gfp* genes (percentage of T_1plants silenced, ANOVA, n=56, P=0.36) or on the percentage of lines silenced (χ^2: P>0.05). Interestingly, multilocus lines were significantly more affected (80% of the lines) by gene silencing than single locus transgenic lines (25% of the lines) (χ^2: P<0.05). Non-expressing progenies from multilocus lines could result either from gene silencing in T_1 plants (especially in homozygous progenies, James et al. 2002) or from the segregation of some loci already inactive in primary T_0 plants (due to non-intact expression units, transcriptional interference from adjacent endogenous plant promoters and/or transcriptional gene silencing, Bhattacharyya et al. 1994). The latter hypothesis is supported by the predominant effect of locus number, rather than copy number, on gene silencing and by the fact that additional transgene copies are likely to be inactive (or slightly contributing at best) to transgene expression level in expressing primary T_0 rice plants (Vain et al. 2003).

Dual T-DNA selection: locus composition and T-DNA linkage configuration

Around 48% (24/50) of the lines contained only one locus, 18% (9/50) contained two loci, 24% (12/50) contained three loci and 10% (5/50) contained four loci (Fig. 5). Such frequent T-DNA integration at independent loci has previously been observed in dicotyledonous (Cluster et al. 1996) and monocotyledonous (Vain et al. 2003) species transformed with a single T-DNA. It could bring limitations to strategies exploiting insertional T-DNA mutagenesis in plants. T-DNA linkage configuration in the population of 50 independent lines was the following: 52% (26/50) of the lines contained only "G-S locus", 28% (14/50) contained a mixture of "G-S locus" and "G locus", 4% (2/50) contained a mixture of "G-S locus" and "S locus", 10% (5/50) contained a mixture of "G-S locus", "G locus" and "S locus" and 6% (3/50) contained a mixture of "G locus" and "S locus" (Fig. 4). Around half of the lines (48%) contained one or more loci harbouring a single-type T-DNA. Significantly more lines contained the pGreen T-DNA alone (22/50 lines with "G locus") than lines with a pSoup T-DNA alone (10/50 lines with "S locus", χ^2: P<0.05). As independent lines often contained a mixture of different types of loci, the population of loci was also studied as a whole. A total of 98 loci were present in the population of 50 plant lines: 56% (55/98) of the loci were "G-S", 31% (30/98) of the loci were "G" and 13% (13/98) of the loci were "S" (Fig. 5) with an average value of 2 loci per line. There was significantly more "G locus" than "S locus" (χ^2: P<0.05). Interestingly, the data analyses by line or by locus lead to near identical results. These characteristics bring new insight to the integration of different T-DNAs in co-transformed and co-expressing plants. Firstly, they demonstrate that in a population of plants co-transformed with pGreen and pSoup-based T-DNAs and co-expressing the transgenes present in these T-DNAs, nearly half of the lines and loci contain single-type T-DNA with twice the chance for this single-type of T-DNA to be pGreen-based than pSoup-based. Secondly, these results cannot be explained without hypothesising different behaviours of the two types of T-DNAs. A model considering a preference ratio of 3 pGreen:2 pSoup T-DNAs could generate the line and loci profiles observed in this study. This model could translate to any difference in delivery/stability/integration between these T-DNAs. The pGreen and pSoup T-DNAs can integrate linked or unlinked, in one or more loci, in one or more copies in the plant genome. This can be modelled by several rounds of combination between the pGreen and pSoup T-DNAs. For example, integration at a given locus, of any two T-DNAs among the three pGreen and two pSoup T-DNAs, would generate 48% (12/25) "G-S locus", 36% (9/25) "G locus" and 16% (4/25) "S locus". The lines containing a "G-S locus" could undertake, or not, a second round of integration, while the lines containing a "G locus" or a "S locus" would require further integration until the criteria of co-transformation is met. Integration of additional copies at the same or at a different locus in these lines (containing so far only one type of T-DNA) would increase the occurrence of "G-S" lines and loci while decreasing these of "G" and "S", bringing it near the 56% "G-S locus":31% "G locus":13% "S locus" observed in this study (Fig. 5). Simple ligation mechanisms occurring before or during the T-DNA integration process could provide a rational for this model. However, we expected the modelling of T-DNA linkage to be somewhat more complex as the pGreen and pSoup T-DNAs could also integrate alone or by more than two copies at a given locus. The preferences hypothesised in the 3 pGreen:2 pSoup model could originate from the vectors themselves, from different transgene content or from T-DNA size in the context of each vector.

Comparisons of locus and transgene copy numbers showed that T-DNA alteration leads 14% (pRT18 T-DNA) to 21% (pRT47 T-DNA) of the loci to contain only part of the T-DNA (i.e., more locus than copy number for a given transgene), which generally consisted of the mid to right border section of the T-DNA containing the reporter gene (*gus* gene for pRT18 and *gfp* gene for pRT47). Comparison of locus and backbone sequence copy number showed also that 34% (pRT47) and 47% (pRT18) of the loci did not contain backbone sequences. Southern analysis of each T_1 plant for the *bar*, *gus*, *aphIV* and *gfp* genes (in addition to assessing transgene presence, or absence, as in this study) would have allowed the precise determination of T-DNA copy number and alteration level, locus by locus. Unfortunately, we were unable to undertake Southern analysis of the thousands of T_1 plants that would be required for such a study.

Overall, large-scale genotyping and phenotyping of primary transformants and their progeny allowed us to build a new and accurate picture of transgenic locus composition and T-DNA linkage configuration in populations of rice plants transformed with two different T-

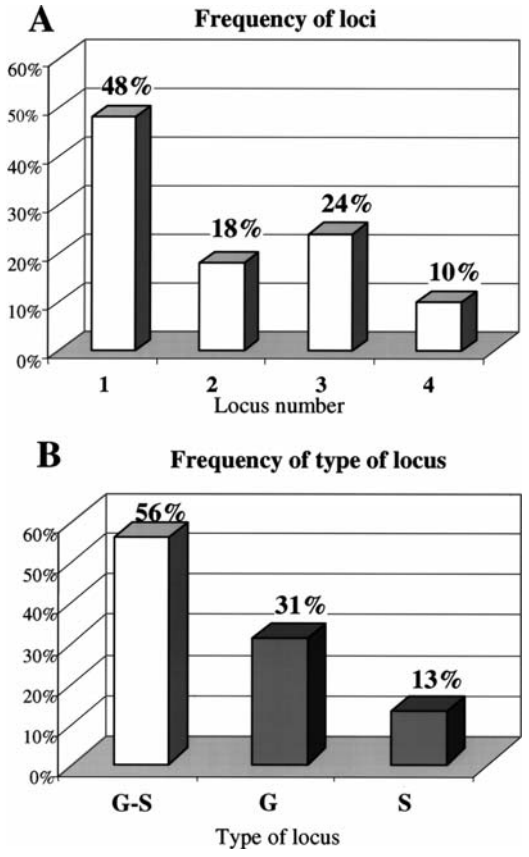

Fig. 5 Number (**a**) and type (**b**) of loci in a population of 50 independent rice plant lines co-transformed and co-expressing the transgenes (produced using dual T-DNA selection). "*G-S locus*": locus containing at least one copy of pRT18 and pRT47 T-DNAs, "*G locus*": locus containing at least one copy of pRT18 T-DNA (no pRT47 T-DNA), "*S locus*": locus containing at least one copy of pRT47 T-DNA (no pRT18 T-DNA). Based on the locus constitution of the 50 rice plant lines independently transformed described in Fig. 4 (50 T_0 and 3,109 T_1 plants studied)

DNAs. Our study showed that plant lines exhibited a wide variety of transgenic loci, transgene and T-DNA content. The most frequent form of integration was where both T-DNAs integrated at the same locus (56% of cases). This suggests that T-DNA repeats—after single or dual T-DNA transformation—originate often from either ligation or co-integration of separate T-DNAs (De Neve et al. 1997). When single-type T-DNA integration occurred (44% of loci), pGreen T-DNA was preferentially integrated. In around half of the plant lines (52%) the T-DNAs integrated at two independent loci or more. In these cases, mixed and single-type T-DNA integrations often occurred concurrently at different loci during the transformation process (in 42% of the plant lines). Non-intact T-DNA were present in 70–78% of the loci leading 14–21% of the loci to contain only the mid to right border part of a T-DNA. This situation generally limited straightforward molecular analysis of primary transformants as bands in hybridisation patterns could represent both an entire, or only part of a T-DNA, and could represent different loci in the plant genome. In 53–66% of the loci, T-DNA integrated with vector backbone sequences.

In the context of the production of maker-free transgenic plants, the co-transformed and co-expressing plants studied so far represented only a subset of the population of transformation events initially produced. It was therefore important to study how often such lines would occur in systems where only single T-DNA selection was used.

Single T-DNA selection: locus constitution and T-DNA linkage configuration

Single T-DNA selection was undertaken using PPT only (selecting for the *bar* gene present on pGreen-based pRT18) or using hygromycin only (selecting for the *aphIV* gene present in pSoup-based pRT47). Only a limited number of lines were produced by PPT selection, and the results could not be interpreted statistically. However, a large-scale experiment was undertaken using pSoup T-DNA selection alone. Transgene presence and expression were analysed in T_0 plants as described in previous sections. Around 70% (131/185) of the lines were co-transformed with the pGreen and pSoup T-DNAs but only 24% (32/131) of the co-transformed lines co-expressed all four transgenes (Table 1). The remaining co-transformed lines were either non-expressing for *gfp* (11%), or *gus* (35%) or both genes (30%). Overall, around 29% (53/185) of the lines did not express the *gfp* gene (present in the selected pRT47 T-DNA) which is similar to the levels obtained for the *gfp* gene during dual T-DNA selection (54/208=26%). However, 65% (85/131) of the lines did not express the *gus* gene, which is significantly higher than when pRT18 T-DNA was selected for (12/208=6% during dual selection, χ^2: $P<0.05$). Therefore in around two-thirds of the co-transformed lines, the transgenes present in the unselected T-DNA were non-expressing. This could be due to either non-intact expression units, transcriptional interference from adjacent endogenous plant promoters or silencing at the transcriptional or post-transcriptional levels (Bhattacharyya et al. 1994; Dong et al. 2001) of the transgenes present in the unselected T-DNA. Comparable levels of co-transformation (70–90%: Lu et al. 2001; Matthew et al. 2001) and co-expression (30–50%: De Block and Debrouwer 1991; Komari et al. 1996; Daley et al. 1998) were previously observed for transgenes carried by different binary vectors. This shows that in plants transformed with different T-DNAs, co-transformation is less of a limiting factor for the development of "clean-gene" technology than the constraint of co-expression of the non-selected gene(s). In total, nine lines obtained using pSoup T-DNA selection only, and three lines obtained using pGreen T-DNA selection only, co-transformed and co-expressing all transgenes present in pRT18 and pRT47 T-DNAs, were randomly chosen. Using the same protocol as before, transgene molecular and expression analyses of T_0 and T_1 plants for each of the 12 lines studied were conducted (i.e., 727 T_1 plants studied representing 1,454 and 1,005

phenotyping and genotyping analyses respectively). The T-DNA linkage configuration for each line is represented in Fig. 6. The overall integration profile is comparable to that obtained when co-transformed and co-expressing lines were directly obtained by dual selection of both T-DNAs (Fig. 4). Around 58% (7/12) of the lines contained a "G locus" or a "S locus" while overall 41% (12/29) of the loci contained single-type T-DNA(s). Progeny plants (T_1) containing a single type of T-DNA could therefore be obtained by selecting either for pGreen-based pRT18 T-DNA or for pSoup-based pRT47 T-DNA. This demonstrates that the pGreen/pSoup dual binary vector system is well suited for the production of marker-free transgenic plants and for the development of efficient "clean-gene" technology in rice. The overall efficiency for the production of a transgenic locus containing a single type of "active" T-DNA is estimated to be around 9% (co-transformation 139/195=71% × co-expression 39/139=28% × 55/127=43% unlinked loci; Table 1). The percentage of lines containing at least one "active" single T-DNA type locus is estimated to be around 10% (co-transformation 139/195=71% × co-expression 39/139=28% × 31/62=50% unlinked lines; Table 1). Backbone transfer (in 53%–66% of the loci for pRT18 and pRT47, respectively) should further reduce these overall efficiencies to around 5%. It is somewhat difficult to compare the results obtained in this study with previously published data as the assessment of "clean-gene" technology performances is directly linked to the accurate assessment of transgenic locus number and T-DNA linkage configurations. In the past, genotyping of large populations of plants transformed with different T-DNAs and their progeny have rarely been conducted (Matthew et al. 2001; Lu et al. 2001). However, unlinked T-DNA integration in 16% (15/95) to 50% (18/39) of independently transformed plant lines has been reported in barley (Matthew et al. 2001) and rice (Lu et al. 2001) respectively, using the one vector/one strain approach. It is therefore possible that the two vectors/one strain approach that we used could limit "clean-gene" efficiency when compared to the one vector/one strain approach.

Strategies to counter-select backbone transfer would constitute one of the first optimisation steps of "clean-gene" technologies based on plant transformation with multiple T-DNAs. Alternatively, new binary vectors could be designed. Each binary vector could be regarded as not one but two "transferable" DNA fragments separated by borders. One fragment (probably the T-DNA itself) could harbour only the gene(s) of interest (without any selectable marker gene) while the backbone could contain both the plant and the bacterial selectable marker genes as well as additional genes (*vir* etc.). The recovery of transgenic

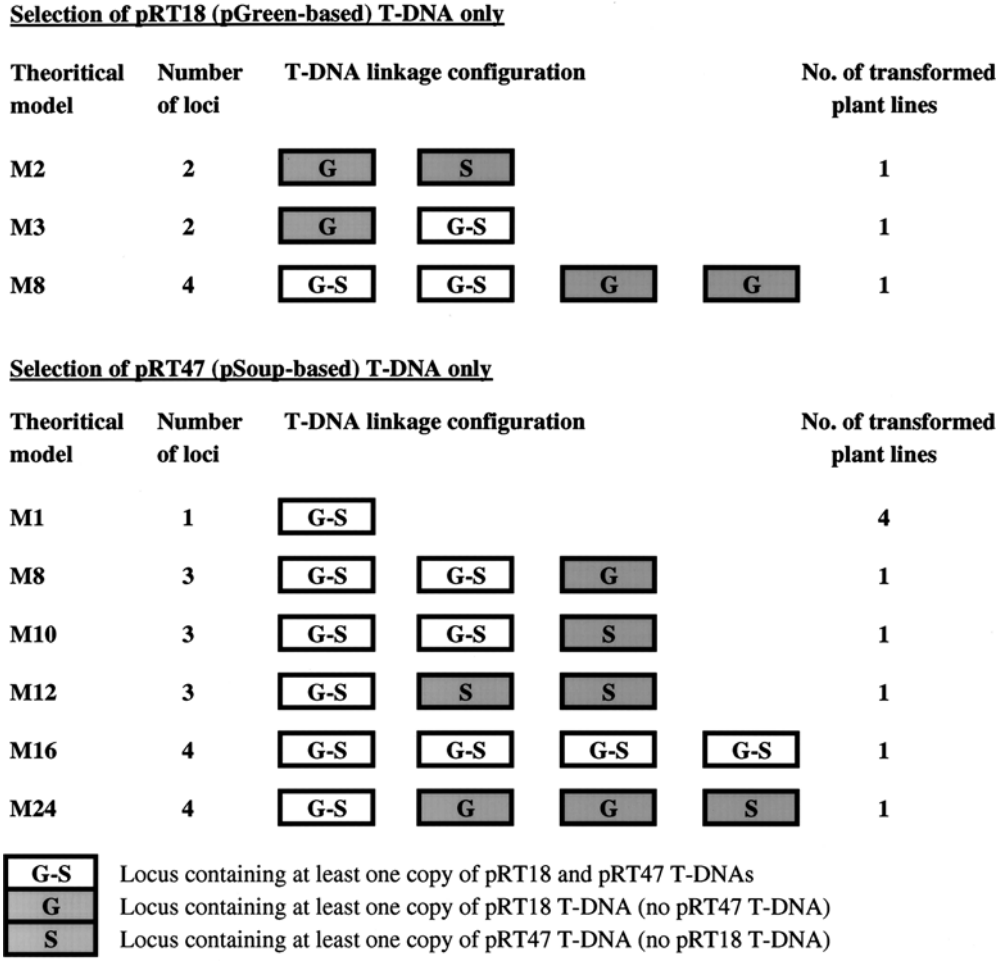

Fig. 6 T-DNA linkage configuration in a population of 12 independent rice plant lines transformed with single T-DNA selection, "clean-gene" technology. Locus constitution was determined by assessing transgene presence and inheritance in 12 T_0 and 727 T_1 plants. Only eight of 30 theoretical models representing all T-DNA linkage configurations of pRT18 and pRT47 up to four independent Mendelian loci corresponded to the independent plant lines produced

plants would be assured by co-transformation of the T-DNA and backbone sequences in a linked or unlinked fashion. Progeny plants free of selectable marker genes would be obtained by segregation of loci containing only the T-DNA. In this situation the dual binary vector system pGreen/pSoup could become a four component super binary vector system. Overall, this work provides new insights into transgenic locus composition and T-DNA linkage configuration in populations of plants transformed with different T-DNAs. Further understanding of transgene behaviour in terms of integration, expression and stability, is central to the future involvement of transformation technologies for the study of gene function and of the mechanisms underlying gene expression.

Acknowledgements We gratefully acknowledge The Rockefeller Foundation for its support. This document is an output from projects (Plant Sciences Research Programme R8031) funded by the UK Department for International Development (DFID) and administered by the Centre for Arid Zone Studies (CAZS) for the benefit of developing countries. The views expressed are not necessarily those of the DFID.

References

Albert H, Dale EC, Lee E, Ow DW (1995) Site-specific integration of T-DNA into wild-type and mutant lox sites in the plant genome. Plant J 7:649–659

Bec S, Chen L, Ferriere NM, Legave T, Fauquet C, Guideroni E (1998) Comparative histology of microprojectile-mediated gene transfer to embryonic calli in japonica rice (*Oryza sativa* L.): influence of the structural organization of target tissues on genotype transformation ability. Plant Sci 138:177–190

Bhattacharyya M, Stemer BA, Dixon RA (1994) Reduced variation in transgene expression from a binary vector with selectable markers at the right and left T-DNA borders. Plant J 6:957–968

Cluster PD, O'Dell M, Metzlaff M, Flavell RB (1996) Details of T-DNA structural organisation from a transgenic petunia population exhibiting co-suppression. Plant Mol Biol 32:1197–1203

Cotsaftis O, Sallaud C, Breitler JC, Meynard D, Greco R, Pereira A, Guiderdoni E (2002) Transposon-mediated generation of T-DNA and marker free rice plants expressing a Bt endotoxin gene. Mol Breed 10:165–180

Dale EC, Ow DE (1990) Intra and intermolecular site-specific recombination in plant cells mediated by bacteriophage P1 recombinase. Gene 91:79–85

Daley M, Knauf VC, Summerfelt KR, Turner JC (1998) Co-transformation with one *Agrobacterium tumefaciens* strain containing two binary plasmids as a method for producing marker-free transgenic plants. Plant Cell Rep 19:489–496

De Block M, Debrouwer D (1991) Two T-DNAs co–transformed into *Brassica napus* by a double *Agrobacterium* infections are mainly integrated at the same locus. Theor Appl Genet 82:257–263

De Framond AJ, Back EW, Chilton WS, Kayes L, Chilton M-D (1986) Two unlinked T-DNAs can transform the same tobacco plant cell and segregate in the F_1 generation. Mol Gen Genet 202:125–131

De Neve M, De Buck S, Jacobs A (1997) T-DNA integration patterns in co-transformed plant cells suggests that T-DNA repeats originate from co-integration of separate T-DNAs. Plant J 11:15–29

Dong J, Kharb P, Teng W, Hall TC (2001) Characterization of rice transformed via an *Agrobacterium*-mediated inflorescence approach. Mol Breed 7:187–194

Ebinuma H, Sugita K, Matsunaga E, Endo S, Yamada K, Komamine A (2001) Systems for the removal of a selection marker and their combination with a positive marker. Plant Cell Rep 20:383–392

Goldsbrough AP, Lastrella CN, Yoder J (1993) Transposition mediated repositioning and subsequent elimination of marker genes from transgenic tomato. BioTechnology 11:1286–1292

Hanson B, Engler D, Moy Y, Newman B (1999) A simple method to enrich an *Agrobacterium*-transformed population of plants containing only T-DNA sequences. Plant J 19:727–734

Hare P, Chua N-H (2002) Excision of selectable marker genes from transgenic plants. Nat Biotechnol 20:575–580

Hellens RP, Edwards A, Leyland NR, Bean S, Mullineaux PM (2000) pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. Plant Mol Biol 42:819–832

Hiei Y, Ohta S, Komari T, Kumashiro T (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries T-DNA. Plant J 6:271–282

Ishida Y, Saito H, Ohta S, Hiei Y, Komari T, Kumashiro T (1996) High efficiency transformation of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*. Nat Biotechnol 14:745–750

James VA, Avart C, Worland B, Snape JW, Vain P (2002) The relationship between homozygous and hemizygous transgene expression levels over generations in populations of transgenic rice plants. Theor Appl Genet 104:553–561

James VA, Worland B, Snape JW, Vain P (2004) Development of a standard operating procedure (SOP) for the precise quantification of transgene expression levels in rice. Physiol Plant 120:650–656

Jefferson RA, Kavanagh TA, Bevan MW (1987) β-glucuronidase as a sensitive and versatile fusion marker in higher plants. EMBO J 6:3901–3907

Jorgensen R, Snyder C, Jones JDG (1987) T-DNA is organized predominantly in inverted repeat structures in plants transformed with *Agrobacterium tumefaciens* C58 derivates. Mol Gen Genet 207:471–477

Komari T, Hiei Y, Saito Y, Murai N, Kumashiro T (1996) Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by *Agrobacterium tumefaciens*and segregation of transformants free from selection markers. Plant J 10:165–174

Koncz C, Nemeth K, Redei GP, Schell J (1994) Homology recognition during T-DNA integration into the plant genome. In: Paszkowski J (ed) Homologous recombination and gene silencing in plants. Kluwer, Dordrecht, pp 167–189

Lu HJ, Zhou X-R, Gong Z-X, Upadhyaya NM (2001) Generation of selectable marker-free transgenic rice using double right-border (DRB) binary vectors. Aust J Plant Physiol 28:241–248

Matthew PR, Wang M-B, Waterhouse PM, Thornton S, Fieg SJ, Gubler F, Jacobsen JV (2001) Marker gene elimination from transgenic barley, using co-transformation with adjacent "twin T-DNAs" on a standard *Agrobacterium* transformatiom vector. Mol Breed 7:195–202

McCormac AC, Elliot MC, Chen DF (1999) pBeCKS2000: a novel plasmid series for the facile creation of complex binary vectors, which incorporates "clean-gene" facilities. Mol Gen Genet 261:226–235

McKnight TD, Lillies MT, Simpson RB (1987) Segregation of genes transferred to one plant cell from two *Agrobacterium* strains. Plant Mol Biol 8:439–445

Miller M, Tagliani L, Wang N, Berka B, Bidney D (2002) High efficiency transgene segregation in co-transformed maize plants using an *Agrobacterium tumefaciens* 2 T-DNA binary system. Transgenic Res 11:381–396

Paszkowski J, Baur M, Bogucki A, Potrykus I (1988) Gene targeting in plants. EMBO J 7:4021–4026

Sallaud C, Meynard D, Van Boxtel J, Gay C, Bes M, Brizard JP, Larmande P, Ortega D, Raynal M, Portefaix M, Ouwerkerk PBF, Rueb S, Delseny M, Guiderdoni E (2003) Highly efficient production and characterization of T-DNA plants for rice (*Oriza sativa* L.) functional genomics. Theor Appl Genet 106:1396–1408

Srivastava V, Anderson OD, Ow DV (1999) Single-copy transgenic wheat generated through the resolution of complex integration patterns. Proc Natl Acad Sci USA 96:11117–11121

Terada R, Urawa H, Inagaki Y, Tsugane K, Lida S (2002) Efficient gene targeting by homologous recombination in rice. Nat Biotechnol 20:1030–1034

Vain P, James VA, Worland B, Snape JW (2002) Transgene behaviour across two generations in a large random population of transgenic rice plants produced by particle bombardment. Theor Appl Genet 105:878–889

Vain P, Afolabi A, Worland B, Snape JW (2003) Transgene behaviour in populations of rice plants transformed using a new dual binary vector system: pGreen/pSoup. Theor Appl Genet 107:210–217

Xing A, Zhang Z, Sato S, Staswick P, Clement T (2000) The use of two T-DNA binary system to derive marker-free transgenic soybeans. In Vitro Cell Dev Plant 36:456–463

Yoder JI, Goldsbrough AP (1994) Transformation systems for generating marker-free transgenic plants. BioTechnology 12:263–267

Zubko E, Scutt C, Meyer P (2000) Intrachromosomal recombination between *attP*regions as a tool to remove selectable marker genes from tobacco transgenes. Nat Biotechnol 18:442–445