Theor Appl Genet (2003) 107:210–217
DOI 10.1007/s00122-003-1255-7

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

Transgene behaviour in populations of rice plants transformed using a new dual binary vector system: pGreen/pSoup

Received: 10 September 2002 / Accepted: 12 December 2002 / Published online: 3 April 2003

Abstract Transgene integration, expression level and stability have been studied, across two generations, in a population of rice plants transformed using a new dual binary vector system: pGreen/pSoup. pGreen is a small Ti binary vector unable to replicate in *Agrobacterium* without the presence of another binary plasmid, pSoup, in the same strain. We engineered both pGreen and pSoup to contain each a different T-DNA. Transformation experiments were conducted using a pGreen vector containing the *bar* and *gusA* expression units (no transgene in pSoup) or with a pSoup vector containing an *aphIV* and *gfp* expression units (no transgene in pGreen). High plant transformation frequencies (up to 40%) were obtained using herbicide resistance (*bar*) or antibiotic resistance (*aphIV*) genes. Around 80% of the independently transformed plants expressed unselected reporter genes (*gusA* or *gfp*) present in the vectors. Backbone sequences transfer was frequent (45% of lines) and occurred often in multicopy lines. Around 15–20% of the rice plant lines contained a single T-DNA integration without backbone. Integration of additional transgene copies did not improve expression levels in either T_0 plants or T_1 progenies. Nearly all multicopy lines contained transgenes integrated at several loci in the plant genome, showing that T-DNAs from either pGreen or pSoup frequently integrated at unlinked loci. Precise determination of loci number required the analysis of transgene presence in progeny. Segregation of transgene phenotype was generally misleading and tended to underestimate the real number of transgenic loci. The contribution of this new dual-binary vector system to the development of high-throughput rice transformation systems and to the production of marker-free transgenic rice plants is discussed.

Communicated by C. Möllers

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

Keywords *Oryza sativa* · Transformation · *Agrobacterium* · pGreen/pSoup · Transgene expression · Transgene stability

Introduction

Since the mid-1990s, there has been great progress in *Agrobaterium*-based technologies for the transformation of recalcitrant crop species. Currently all cereal crops can be transformed via *Agrobacterium* (Hiei et al. 1994; Ishida et al. 1996; Cheng et al. 1997; Tingay et al. 1997). In rice, *Agrobacterium*-mediated transformation is currently efficient for a wide range of japonica and indica elite cultivars (Hiei et al. 1994; Aldemita and Hodges 1996; Dong et al. 1996). Even if some limitations remain associated with this technology (such as genotype recalcitrance, unwanted vector backbone transfer, T-DNA repeats and minor rearrangements at the border junctions), it offers key advantages over alternative DNA delivery methods based upon direct DNA transfer. *Agrobacterium*-based transformation technology generally produces a higher frequency of single-copy transgenic plants, which are key to most molecular genetic studies and crop improvement strategies. It also offers the possibility, when transgenes are present in separate T-DNAs, of obtaining unlinked co-transformed transgenes which allow the production of marker-free transgenic plants (Komari et al. 1996).

Advances in plant transformation using *Agrobacterium*-mediated techniques are linked, in part, to the improvement of Ti vector design (see Hellens and Mullineaux 2000 for review). In the past, important improvements in Ti vectors used for cereal crop transformation have been made (Hei et al. 1994; Komari et al. 1996; Wang et al. 1998). Nevertheless, there is still a demand for user-friendly vector systems offering small plasmid size, transformation selection flexibility, extensive multiple cloning sites, a flexibility to adapt to specific requirements (such as transfer of large pieces of DNA and transposon-based transgene reposition), allow-

ing multiple T-DNA transfer and leading to high transformation frequencies. In this study, we tested a new dual-binary vector system pGreen/pSoup (Hellens et al. 2000) for rice transformation. 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 species using a single T-DNA present on the pGreen vector only (Hellens et al. 2000). In this study, we engineered both pGreen and pSoup vectors to contain each a different T-DNA. We assessed the relationship between transgene copy number, loci number, expression level and stability, across two generations, in population of rice plants transformed with either pGreen- or pSoup-based vectors. The contribution of this new dual-binary system to the development of high-throughput rice transformation systems and to the production of marker-free transgenic rice plants is discussed.

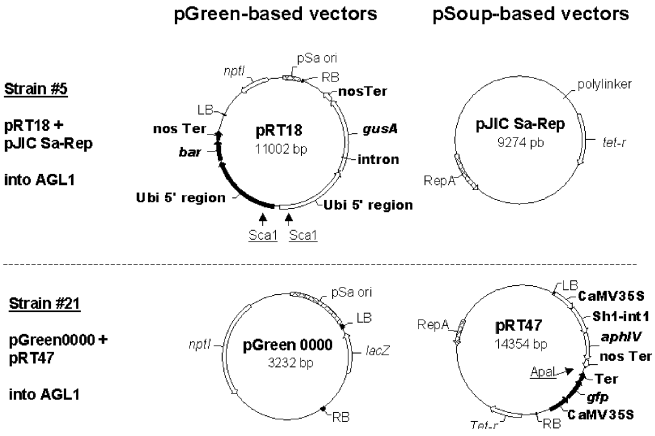

Fig. 1 pGreen/pSoup-based vectors used for rice transformation. RepA in pSoup-based vectors provided replication functions in *trans* for pGreen-based vectors (containing pSAori origin of replication) in *Agrobacterium*. pSoup-based vectors contained their own origin of replication and replicase (not represented here) (Hellens et al. 2000). Gene abbreviations are detailed in the Materials and methods. In strain no. 5, transgenes (*bar, gusA*) are present in the pGreen-based vector pRT18, in strain no. 21, transgenes (*aphIV, gfp*) are present in the pSoup-based vector pRT47

Materials and methods

pGreen/pSoup-based vectors

Binary vectors and *Agrobacterium* strains used for rice transformation are presented in Fig. 1. pGreen0000 and pSA-Rep are basic pGreen and pSoup vectors without transgenes (kindly provided by P. Mullineaux, John Innes Centre, UK). pRT18 harboured an intron-containing *gusA* gene and the *bar* gene both driven by the maize 5′ ubiquitin region and with a nopaline synthase terminator (kindly provided by D. Lonsdale, John Innes Centre, UK; pRT18 is also referenced as PAL156). pRT47 was constructed by transferring the T-DNA fragment from pGreen0000 into pSA-Rep and by inserting two expression units into the T-DNA. The first expression unit contained the CaMV35Spromoter::*gfp*::Ter from pGVT1 (kindly provided by V. Thole, John Innes Centre, UK). The second expression unit contained the CaMV35S promoter::Sh1-intron1::*aphIV*::nosTer from p35SshHPT (Vain et al. 1996). Neither pRT18 nor pRT47 are super binary vectors (Hiei et al. 1994, 1997) as no extra virulence genes are present in the vector backbone.

Plasmids were transformed into *E. coli* strain DH5α using the PEG-transformation technique and into *Agrobacterium* strains AGL1 or LBA4404 using a freeze-thaw technique (Hellens et al. 2000). As detailed in Fig. 1, *Agrobacterium* strain no. 5 harboured the two binary vectors, pRT18 (pGreen-based) and pSa-Rep (pSoup-based). *Agrobacterium* strain no. 21 harboured the two binary vectors pGreen0000 (pGreen-based) and pRT47 (pSoup-based).

Rice transformation procedures

Mature seeds of rice (*Oryza sativa* L.) var. Nipponbare were used for callus production using modified protocols from Sivamani et al. (1996), Wang et al. (1997) and Bec et al. (1998). Dehusked seeds were sterilised with half-strength commercial bleach for 15 min and rinsed three times with sterile distilled water. The embryos were aseptically removed under a dissecting microscope and plated onto NBm medium (macro-element N6, micro-elements B5, Fe-EDTA, 30 g l^{-1} sucrose, 2 mg l^{-1} 2,4-dichloroxyacetic acid (2,4-D), 300 mg l^{-1} casein hydrolysate, 500 mg l^{-1} L-glutamine, 500 mg l^{-1} L-proline, 2.5 g l^{-1} Phytagel, pH 5.8; filter-sterilised vitamins B5 added after autoclavage) for 3 weeks in the dark at 25 °C. Loose embryogenic transluscent 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 onto fresh NBm medium (approx. 100 globules per plate) to produce embryogenic nodular units (ENU, Bec et al. 1998), used as targets for transformation.

Agrobacterium strains were grown for 2 days at 28 °C on solid MG/L medium (Garfinkel and Nester 1980) supplemented by 200 μM acetosyringone, 50 mg l^{-1} kanamycin (selection for pGreen-based vectors) and 10 mg l^{-1} tetracyclin (selection for pSoup-based vectors). Bacteria cells were scooped up from the plate, re-suspended in 20 ml of SU4 liquid medium (macro-element N6, micro-elements MS, Fe-EDTA, 10 g l^{-1} sucrose, 10 g l^{-1} mannitol, pH 5.5, without antibiotics) and shaken for 1 h at 28 °C. Culture plates containing ENUs were flooded with bacterial suspension OD = 1 (600 nm) for 5 min. Liquid was removed, and each ENU was picked and blotted onto sterile filter paper before being placed onto co-cultivation medium (NBm medium supplemented by 200 μM acetosyringone) for 2 days in the dark at 25 °C. After co-culture, ENUs were put onto selection medium (NBm medium containing 150 mg l^{-1} timentin plus either 5 mg l^{-1} phosphinotrycin (PPT, selection pRT18) or 50 mg l^{-1} hygromycin (selection pRT47) in the dark at 28 °C. L-glutamine was removed from all culture media when PPT was included. After a 2-week-long culture, each callus (grown from an individual ENU) was split into two to five pieces. Each piece of callus could contain several independently transformed sectors but at this stage (2 weeks after transformation) the transformed sectors were too small to be divided. Pieces of callus were cultured for 3 additional weeks onto fresh NBm-based selection medium. Depending upon the selection regime, the resistant calli grown from individual ENU, after 2 + 3 weeks of selection, were either all grouped together or kept separated according to the separation undertaken at 2 weeks.

Transformed plants were regenerated from resistant calli using culture media all supplemented with 50 mg l^{-1} Timentin and containing either 5 mg l^{-1} PPT (selection pRT18) or 50 mg l^{-1} hygromycin (selection pRT47). The resistant calli were transferred to PRm pre-regeneration medium (NBm medium without 2,4-D but with 2 mg l^{-1} benzylaminopurine (BAP), 1 mg l^{-1} α naphthale-

neacetic acid (NAA), 5 mg l^{-1} abscicic acid (ABA) for 9 days in the dark at 28 °C. Calli showing clear differential growth were then transferred to regeneration medium RNm (NBm medium without 2,4-D but with 3 mg l^{-1} BAP, 0.5 mg l^{-1} NAA) for 2–3 weeks in the light at 28 °C. Unless otherwise stated, only one plant was regenerated from each orignal ENU to guarantee that each plant represented an independent transformation event. Plants were developed on MSR6 medium (Vain et al. 1998) for 2–3 weeks at 28 °C in the light. Transformed plants were transferred to a controlled environment room for growth to maturity. All transgenic plants produced were used in further experiments to ensure the study of randomised independent transformation events with the widest spectrum of expression for the non-selected *gusA* (in pRT18) or *gfp* (in pRT47) genes.

Analysis of β-glucuronidase (GUS) and green fluorescent protein (GFP) activity

Fluorometric analysis for GUS activity was carried out on leaf tissue from rice plants at the five-leaf stage, following the method of Jefferson (1987). Fluorescence was measured using a Titertek Fluoroskan II after a 0-, 30- and 60-min incubation. Each assay was performed in triplicate. Protein content was determined using a Bio-Rad (Herculus Calif.) protein assay kit. Data were expressed as picomoles of 4-methylumbelliferone (MU) per minute per milligram 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. 1999).

Visual detection of GFP fluorescence was performed using a MZ6 Leica dissecting microscope with a fluorescent module (Leica model 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 (PCR)

DNA was isolated and PCR reactions were carried out as previously described (Vain et al. 1998). Five primer sets were used on the DNA samples: (1) one to amplify the 1,200-bp single-copy rice genomic 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::*gusA* 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′); (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. 1999). Membranes were hybridised with the following probes: 549 bp of the *bar* gene, 701 bp of the *gusA* gene, 981 bp of the *aphIV* gene, 737 bp of the *gfp* gene, 421 bp of the pGreen backbone or 464 bp of the pSoup backbone. The filters were analysed by using a Typhoon 8600 phospho-imager and the Typhoon Scanner Control v1.0/ImageQuant v5.1 software (Molecular Dynamics).

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. 1998) without selection. Transgene expression in the T_1 embryos or seedlings was assessed qualitatively by histochemical GUS staining (Jefferson et al. 1987) or observation of GFP fluorescence in at least 40 random T_1 seedlings from each independently transformed T_0 plant. All T_1 plant progeny not expressing either the *gusA* or *gfp* transgene were analysed by PCR for the presence of the corresponding gene. When PCR reactions were negative for the transgene an additional PCR analysis was conducted for the presence of the C213 genomic rice probe to confirm that the DNA extraction was suitable for PCR amplification.

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 paired *t*-test. Observed segregation ratios were compared to Mendelian 3:1, 15:1, 63:1 and 255:1 ratios using chi-square analysis with Yate's correction.

Results and discussion

High-frequency rice transformation using either pGreen or pSoup vectors

pGreen and pSoup vectors were engineered to contain each a different T-DNA. Different selection and reporter gene combinations were introduced into pGreen and pSoup vectors (Fig. 1) to assess: (1) the possibility of using either pGreen or pSoup vectors for rice transformation; (2) the feasibility of developing rice transformation systems efficient enough to allow high-throughput rice transformation; (3) the integration and expression of transgenes delivered by either pGreen or pSoup vectors.

An *Agrobacterium*-mediated transformation system was developed and optimised using the model rice genotype Nipponbare and ENUs (Bec et al. 1998) derived from mature seeds. ENUs were used because they offered highly responsive target material previously used in particle bombardment experiments (Sivamani et al. 1996). The selection on herbicide (PPT) or antibiotic (hygromycin) of ENUs co-cultured with *Agrobacterium* strain no. 5 or no. 21 (Fig. 1) led to the production of numerous independently transformed callus clones. Depending upon the selection regime, each original ENU produced, on average, 1.3 (PPT selection) to 1.7 (hygromycin selection) transformed callus clones from which plants could be regenerated. Such production of multiple independent transformation events per original explant is crucial for the development of high-throughput transformation systems. However, it requires precise monitoring and optimisation of selection regimes in order to avoid clone duplication (when two plants from the same callus clone are regenerated) or clone mixing (when two different plants are regenerated from what is believed to be a single callus clone). Upadhyaya et al. (2000) reported that around 40% of rice callus clones generated via

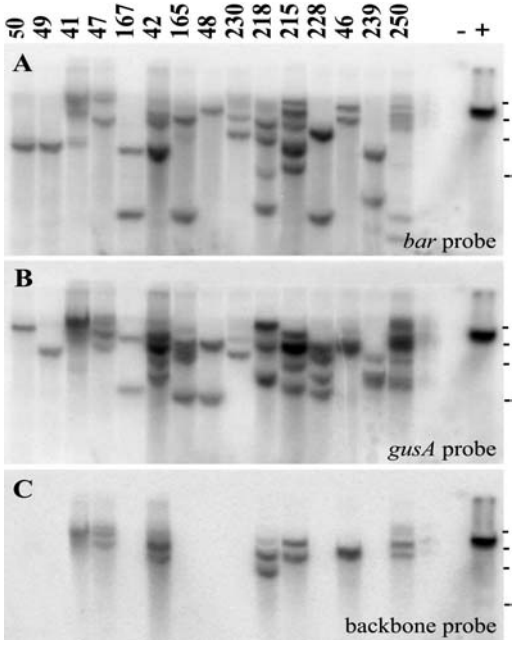

Fig. 2A–C Southern blot analysis of T_0 rice plants independently transformed with *Agrobactrium* strain no. 5 (pRT18 + pSa-Rep in AGL1). Plant genomic DNA was digested using restriction enzyme *Sca*I, which cuts in the middle of the pRT18 T-DNA (see Fig. 1). The membrane was probed with sequences from either: **A** the *bar* gene (left border), **B** the *gusA* gene (right border), or **C** the pGreen backbone. – Wild-type genomic DNA, + RT18. Ladder (*right* of picture) from *top* to *bottom*: 23.1 kb, 9.4 kb, 6.6 kb and 4.3 kb

Agrobacterium transformation produced plants originating from different transformation events. In this study, unless otherwise stated, only one plant was regenerated per original explant (ENU) for further molecular and generational studies to avoid such a problem.

Independently transformed rice plants were obtained at a high frequency using either the pGreen (15% transformation efficiency withPPT selection - pRT18)- or pSoup (24% transformation efficiency with hygromycin selection - pRT47)-based vectors (Table 1). The introduction of pRT18 and pSa-Rep vectors into *Agrobacterium* strain LBA4404 instead of AGL1 led to similar transformation efficiencies (24 transformed plant lines/159 ENUs = 15% using LBA4404). Strategies relying on hygromycin selection were, in general, easier to implement and more efficient than those relying on PPT due to a better differential growth between transformed and non-transformed tissues. When an optimised selection regime was used and more than one independent callus clone was kept per original explant (ENU), the transformation efficiency reached 38% (Table 1). This exceeds transformation efficiencies generally reported for moncotyledonous species using other binary vectors (up to 25% for model rice or maize genotypes; Hiei et al. 1994; Aldemita and Hodges 1996; Ishida et al. 1996). In this system, further improvements of strategies to individualize independently transformed callus clones should lead to a further increase of transformation efficiency. To date, the pGreen/pSoup transformation system routinely allows the production of more than 100 independent plant lines per man-experiment. This can lead to the production of thousands of plant lines per month suitable for complying with the high demands of some post-genomics applications such as these of high-throughput assessment of gene function.

Transgene structure

Transgene integration patterns and copy number were determined by genomic DNA blot analysis on primary transgenic plants (T_0) expressing the selectable marker genes and the reporter genes (*bar* and *gusA* for pRT18; *aphIV*and *gfp* for pRT47). Rice plants containing pRT18 (pGreen-based) were analysed using a restriction enzyme (*Sca*I) that cut in the middle of the T-DNA and using probes for either the *gusA* gene (right T-DNA border), the *bar* gene (left T-DNA border) or the p-Green backbone (Fig. 2). Rice plants containing pRT47 (pSoup-based) were analysed using a restriction enzyme (*Apa*I) that cut in the middle of the T-DNA and using probes for either the *gfp* gene (right T-DNA border), the *aphIV* gene (left T-DNA border) or the p-Soup backbone (Fig. 3).

Transgene copy number was estimated from the number of integrated fragments providing that all bands in the pattern exhibited comparable hybridisation intensity. Bands exhibiting stronger hydridisation signals were scored as more than one copy. Such stronger bands could result from multiple/rearranged T-DNA integration at a given locus (such as inverted repeats). The distribution profile of copy number for each of the four transgenes has

Table 1 Transformation frequency and transgene expression in rice plants

Selection	Number of independent experiments	Number of explants (ENU)	Number of callus clones selected per ENU	Number of clones producing transformed plants[a] (A)	Transformation frequency (%)	Number of clones producing plants expressing marker gene (B)	Percentage expressers (B/A)
PPT (strain no. 5)	3	568	1	86	15	75	87 GUS+
Hygromycin (strain no. 21)	3 1	553 90	1 1–4	144 44	26 38	106 32	89 GFP+ 73 GFP+

[a] Fully developed greenhouse plants containing transgene(s). Only one plant was regenerated per independent callus clone. *ENU* embryogenic nodular unit

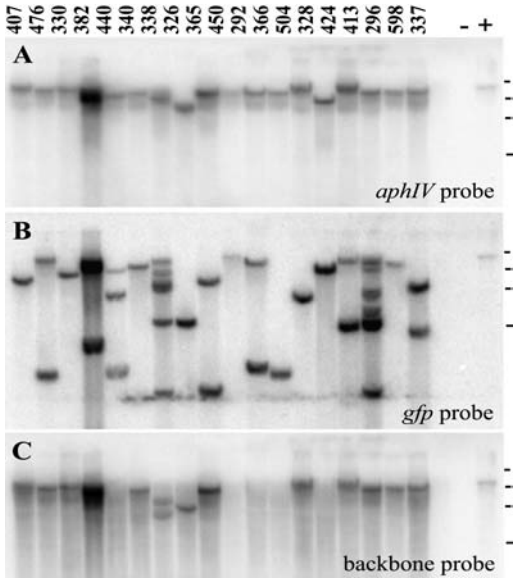

Fig. 3A–C Southern blot analysis of T_0 rice plants independently transformed with *Agrobactrium* strain no. 21 (pGreen0000 + pRT47 in AGL1). Plant genomic DNA was digested using restriction enzyme *Apa*I, which cuts in the middle of the pRT47 T-DNA (see Fig. 1). The membrane was probed with sequences from either: **A** the *aphIV* gene (left border), **B** the *gfp* gene (right border), or **C** the pSoup backbone. – Wild-type genomic DNA, + RT47. Ladder (*right* of picture) from *top* to *bottom*: 23.1 kb, 9.4 kb, 6.6 kb and 4.3 kb

been represented in Table 2. At least 27% of the independently transformed plant lines contained a single transgene copy. This is comparable to other binary vectors used for rice (20% single-copy, Hiei et al. 1994) or wheat (35% single-copy, Cheng et al. 1997) transformation. Surprisingly, the number of copies of transgenes originally present in the same T-DNA were often different in individually transformed plant lines (paired *t*-test, $P < 0.05$, copy number *gusA* versus *bar* and copy number *aphIV* versus *gfp*). Around 68% (51/75) of the lines contained different copy numbers for the *gusA* and *bar* genes (both present in pRT18 T-DNA), and 43% (9/21) of the lines contained different copy numbers for the *aphIV* and *gfp* genes (both present in pRT47 T-DNA). Such disparities have also been reported for other vectors used to transform rice (pJD4, Dong et al. 2001) or maize (pTOK233, Ishida et al. 1996). These disparities could result from partial T-DNA integration, head-to-head tandem repeats or limitations associated with copy-number determination based upon hybridisation pattern only (James et al. 2002, Vain et al. 2002). In this study, pSoup-based vector, pRT47, generally produced more single-copy integration events than the pGreen-based vector pRT18 (chi-square, $P < 0.05$, using one versus more than one gene copies, as shown in Table 2). The fact that both vectors carried different expression units and that different size populations were studied (75 lines for pRT18 and 21 lines for pRT47) could limit the comparison. The presence of the *aphIV* expression unit few nucleotides from the left border could also have favoured single T-DNA integration with intact left border (Fig. 3A). Around 19% (14/75) of the transgenic lines transformed with pRT18 (pGreen-based) contained a single-copy of both the *gusA* and *bar* genes, while 57% (12/21) of the transgenic lines transformed with pRT47 (pSoup-based) contained a single-copy of both the *aphIV* and *gfp* genes. Overall, pGreen- or pSoup-based vectors produced populations of transgenic plants exhibiting comparable profiles of trangene copy number to those obtained using other binary vectors for rice (pTOK233, Hiei et al. 1994), wheat (pMON18365, Cheng et al. 1997) or maize (pTOK233, Ishida et al. 1996) transformation.

Transgenic locus number can be determined by a combination of molecular, histological and segregation analyses. In this study, we used a transgene inheritance study at the structural level to determine the number of loci in independently transformed plant lines (see next section).

The presence of backbone sequences in transgenic plant lines was assessed using backbone probes near the left border of pRT18 (pGreen-based) or pRT47 (pSoup-based). The absence of a backbone sequence up to 1 kb from the left border should ensure that no unwanted antibiotic resistance gene (located 1 kb or more from the left border) has been introduced into transgenic plant lines. Backbone transfer was frequent in plant lines transformed with either pGreen-based, pRT18 (28/75

Table 2 Distribution profiles of transgene copy number in populations[a] of transgenic rice plants produced using pGreen or pSoup vectors

Transgenes in T-DNA	pGreen (pRT18) as in Fig. 2		pSoup (pRT47) as in Fig. 3	
	bar	*gusA*	*aphIV*	*gfp*
One copy	33%	27%	100%	57%
Two copies	39%	21%	0%	29%
Three copies or more	28%	52%	0%	14%
Average copy no.	2.3	2.6	1	1.8
Same copy no.[b]	32%		57%	
One copy T-DNA	19%		57%	
Backbone presence	37%		72%	
One copy T-DNA, no backbone	17%		19%	

[a] Study was conducted on 75 and 21 independently transformed plant lines for pRT18 and pRT47, respectively

[b] Percentage of independent lines containing the same number of copies of *bar* and *gusA* (pRT18) or *aphIV* and *gfp* (pRT47)

lines = 37%) or with pSoup-based, pRT47 (15/21 = 72%) vectors. These results are consistent with work using other binary vectors in transgenic dicotyledoneous (Martineau et al. 1994; De Buck et al. 2000) and monocotyledonous (Yin and Wang 2000; Upadhyaya et al. 2000) plant species. The transfer of vector backbone could be the result of read-through at the left border resulting in the transfer of the T-DNA plus vector backbone sequences. Read-through could proceed along the entire vector sequence up to the right or the left border, resulting in the transfer of one or two T-DNAs separated by the entire vector backbone (De Buck et al. 2000). The latter finding could explain the more frequent backbone transfer observed in multicopy lines (52%) than in single-copy lines (36%) observed in this study. Backbone transfer seemed also to occur at a lower frequency with pRT18 than with pRT47. Both vectors originally carried the same right and left border as well as the same inner right border sequences. They only differ by T-DNA sequence (i.e. different expression units), left inner border sequences (resulting from the cloning of the *aphIV* gene directly along the left border) and outer border sequences (i.e. different surrounding backbone sequences). This suggests that the sequences surrounding the border repeats or the size of the T-DNA (by itself or relative to the backbone size) could influence vector backbone transfer. De Buck et al. (2000) reported that inner border sequences strongly influence the frequency of vector backbone transfer in *Nicotiana* and *Arabidospsis*. McCormac et al. (2001) proposed that co-transformation of different T-DNAs in *Nicotiana* could be influenced by their relative sizes. In this study,17% (13/75) of the transgenic rice lines transformed with pRT18 (pGreen-based) and 19% (4/21) of the transgenic lines transformed with pRT47 (pSoup-based) contained a single T-DNA insert without backbone sequence (Table 2).

Transgene expression

Transgene expression in T_0 rice plants was assessed in 86 and 188 independently transformed lines containing pRT18 (pGreen-based), and pRT47 (pSoup-based) respectively (Table 1). Plants containing pRT18 were analysed for the expression of the unselected *gusA* gene by histochemical staining. Plants containing pRT47 were analysed for expression of the unselected *gfp* gene by scoring GFP fluorescence. For both pRT18 and pRT47, 70–90% of the independent T_0 plant lines expressed the transgenes (Table 1). This shows that large numbers of transgenic rice plants expressing non-selectable transgenes can be produced using either pGreen or pSoup binary vectors. A high percentage of expressing transformation events has also been observed in populations of transgenic rice (80–90% Hiei et al. 1994; Yin and Wang 2000; Dong et al. 2001), maize (60% Ishida et al. 1996) and barley (89% Tingay et al. 1997) plants created using other binary vectors. This similarity of transgene expression profiles is probably related to similar patterns of transgene integration in plants after *Agrobacterium*-mediated transformation using different vectors. In this study, non-expressing T_0 rice plant lines were not studied further. They could have contained either non-intact expression units or transgenes silenced at the transcriptional or post-transcriptional levels (Hiei et al. 1997; Dong et al. 2001).

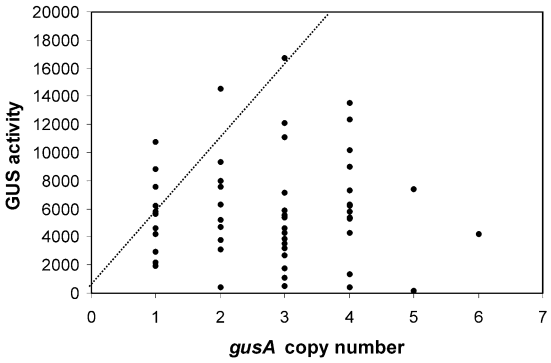

Fig. 4 Relationship between *gusA* copy number and expression level (in pmol MU min^{-1} mg^{-1} protein) in T_0 rice plants transformed with pRT18 and expressing the *gusA* gene. Each data point represents one independent plant line. The *dotted line* represents expected expression levels based upon zero and one copy average expression values

Transgene expression levels in T_0 rice plants transformed with pRT18 and exhibiting GUS activity was quantified using the fluorometric GUS assay. The relationship between *gusA* gene expression levels and copy number was analysed in 53 independently transformed plant lines (Fig. 4). There was no relationship between transgene copy number and expression level. Transgenic plants containing two or more copies showed no significant increase in transgene expression levels (ANOVA $P = 0.78$). This suggests that most transgenes present in multicopy expressing lines were either non-functional (non-intact expression unit) or silenced (at the transcriptional level), or they contributed to a reduction of overall transgene expression. These transgene copies could either be located at the same locus or at different loci than expressing gene copies.

Transgene inheritance and segregation of transgene phenotype

T_1 seeds were obtained by self-pollination of primary rice (T_0) plants transformed by pRT18 and expressing the *gusA* gene. More than 85% of the transformed plant lines exhibited reasonable fertility (more than 40 seeds). Transgene inheritance patterns were determined according to the presence of the transgenes in T_1 plants, while segregation of transgene phenotype was determined by assessing transgene expression in the same plants. Observed ratios were compared to Mendelian segregation models using chi-square analysis. Transgene inheritance studies were conducted for 42 out of the 75 independently

Table 3 Transgene inheritance and expression in T_1 rice plants[a] transformed with pRT18

Number of *gusA* copies (as determined in Fig. 2)	Number of plant lines studied[b]	Number of T_1 plant analysed[c]	Transgene inheritance[d]			Segregation of transgene phenotype[e]		
			Non-Mendelian (less 3:1)	1 locus (3:1)	2 loci (15:1) or more	Less 3:1 (suggest non-Mendelian)	3:1 (suggest 1 locus)	15:1 (suggest 2 loci)
1	14	531	0	14	0	1	13	0
2	10	373	0	0	10	2	4	4
3	9	334	0	1	8	3	3	3
4	5	236	0	0	5	0	2	3
5	4	250	0	0	4	0	0	4

[a] Observed segregation ratios were compared to Mendelian 3:1 or 15:1 ratios using Chi-square test ($P < 0.05$). For some lines, comparison with 63:1 (3 loci) and 255:1 (4 loci) ratios were performed using 80 and 130 T_1 plants, respectively
[b] Independent plant lines expressing the unselected *gusA* gene in T_0 plants
[c] At least 40 random T_1 seedlings were analysed from each independent plant line
[d] Based upon the presence of the *gusA* gene in T_1 plants
[e] Based upon the expression of the *gusA* gene in T_1 plants

transformed plant lines containing different copy numbers of the unselected g*usA* gene in T_0 plants (14 lines containing one copy, ten lines containing two copies, nine lines containing three copies, five lines containing four copies and four lines containing five copies; Table 3). All single-copy lines (14/14) segregated with a 3:1 ratio as expected for one Mendelian locus. Most multicopy lines (27/28 = 96%) exhibited segregation ratios not different from 15:1, 63:1 or 255:1, as expected from two, three or four Mendelian loci. This implies that many transgenes in multicopy lines were inserted at unlinked loci in the plant genome (at least two loci). Some transgenic lines, such as AF150, produced only transformed T_1 plants among 130 random progenies tested, suggesting a transgenic makeup of at least three unlinked loci. Only one line, containing three fragments in hybridisation patterns, segregated 3:1 as one Mendelian locus. Overall, multicopy lines were nearly all multilocus lines with some loci containing a single transgene copy and some loci containing multiple transgene copies. Similar results were obtained with eight independent plant lines transformed by pSoup-based pRT47 (data not shown). These results are difficult to compare to those obtained with other binary vectors as the number of transgenic loci in cereal plants produced via *Agrobacterium* has rarely been (if ever) determined in large populations of independent transformation events after structural analysis of progenies. Nevertheless, multilocus inserts are known to occur in cereal plants transformed by other binary vectors and have been described through transgene inheritance at the structural level or FISH analysis (Dong et al. 2001) in some plant lines. If multiple unlinked T-DNA integration was also occurring at high frequency using other binary vectors, this could have important implications for further exploitation of *Agrobacterium*-based transformation technologies.

Segregation of transgene phenotype using the same 1724 T_1 plants (Table 3) showed striking differences to transgene inheritance 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 the transgenes. Multicopy/multilocus lines were significantly more affected by this phenomenon (14/28) than single-copy transgene lines (1/14) (chi-square $P < 0.05$, Table 3). In fact, most of the non-expressing progenies from multicopy lines contained transgenes. This implies that segregation of transgene phenotype should not be trusted as a reliable indicator of transgene inheritance (James et al. 2002; Vain et al. 2002). In this study, it tended to underestimate the real number of transgenic loci, especially for multicopy lines. Around 20% (5/28) of the multilocus lines showed skewed (less than 3:1) segregation patterns of transgene phenotype. Around 30% (9/28) of the multilocus lines showed 3:1 segregation of the transgene phenotype, incorrectly suggesting a single transgenic locus. For the remaining mutlilocus lines (around 50%) the exact number of loci was also often underestimated. For lines such as AF150 containing at least three Mendelian loci, segregation of the transgene phenotype incorrectly suggested two loci (segregation of transgene phenotype not different from 15:1 but different from 63:1). Non-expressing progenies from multiloci lines could result either from gene silencing in T_1 plants (especially in homozygous progenies, James et al. 2002) or from segregating loci already not expressing in primary T_0 plants. The latter hypothesis is also consistent with the fact that additional transgene copies did not improve transgene expression level in expressing primary T_0 rice plants (Fig. 4). This would suggest that multiple unlinked integration of T-DNA are frequent in transgenic rice plants transformed by either pGreen or pSoup vectors and that only few of the transgenic loci (one or two) are generally expressing the transgenes while the other loci show no expression. In subsequent generations, the non-expressing loci would lead to the production of many plant progeny containing but not expressing the transgenes.

The high occurrence of multilocus/unlinked integration of T-DNA from either pGreen or pSoup vectors has important implications for the further exploitation of this dual binary system. Frequent multilocus integration could bring limitations to strategies exploiting insertional T-

DNA mutagenesis. However, multilocus integration of T-DNA delivered using either pGreen or pSoup vectors could be very useful for the production of transgenic plants free of marker genes. Further co-transformation experiments with one vector (pGreen) carrying the gene of interest and the other vector (pSoup) carrying the selectable marker gene will allow us to assess the frequency of unlinked integrations of pGreen and pSoup T-DNAs. Such unlinked integration would lead to the development of 'clean gene technology' allowing the production of plant progenies containing only the gene of interest but free of any selectable marker gene.

Acknowledgements We gratefully acknowledge The Rockefeller Foundation for its support. We thank Drs E. Guiderdoni, C. Sallaud and J.C. Breitler (CIRAD, France) for advice and insightful discussions. We thank V.A. James (University of Florida, USA) for proofing the manuscript. 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 DFID.

References

Aldemita R, Hodges TK (1996) *Agrobacterium tumefaciens*-mediated transformation of japonica and indica rice varieties. Planta 199:612–617

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

Cheng M, Fry JE, Pang S, Zhou H, Hironica CM, Duncan DR, Conner TW, Wan Y (1997) Genetic transformation of wheat mediated by *Agrobacterium tumefaciens*. Plant Physiol 115:971–980

De Buck S, De Wilde C, Van Montagu M, Depicker A (2000) T-DNA vector backbone sequences are frequently integrated into the genome of transgenic plants obtained by *Agrobacterium*-mediated transformation. Mol Breed 6:459–468

Dong J, Teng W, Buchholz G, Hall TC (1996) *Agrobacterium*-mediated transformation of Javanica rice. Mol Breed 2:267–276

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

Garfinkel M, Nester EW (1980) *Agrobacterium tumefaciens* mutants affected in crown gall tumorigenesis and octopine catabolism. J Bacteriol 144:732–743

Hellens R, Mullineaux P (2000) A guide to *Agrobacterium* binary Ti vectors. Trends Plant Sci 5:446–451

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

Hiei Y, Komari T, Kubo T (1997) Transformation of rice mediated by *Agrobacterium tumefaciens*. Plant Mol Biol 35:205–218

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 104:533–561

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

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

Martineau B, Voelker TA, Sanders RA (1994) On defining T-DNA. Plant Cell 6:1032–1033

McCormac AC, Fowler MR, Chen D, Elliot MC (2001) Efficient co-transformation of *Nicotiana tabacum* by two independent T-DNAs, the effect of T-DNA size and implications for genetic separation. Transgen Res 10:143–155

Sivamani E, Shen P, Opalka N, Beachy RN, Fauquet CM (1996) Selection of large quantities of embryogenic calli from indica rice seeds for production of fertile plants using the biolistic method. Plant Cell Rep 15:322–327

Tingay S, McElroy D, Kalla R, Fieg S, Wang M, Thornton S, Brettell R (1997) *Agrobacterium tumefaciens*-mediated barley transformation. Plant J 11:1369–1376

Upadhyaya NM, Surin B, Ramm K, Gaudron J, Schunmann PHD, Taylor W, Waterhouse PM, Wang MB (2000) *Agrobacterium*-mediated transformation of Australian rice cultivars Jarrah and Amarroo using modified promoters and selectable markers. Aust J Plant Physiol 27:201–210

Vain P, Finer K, Engler D, Pratt R, Finer JJ (1996) Intron-mediated gene expression enhancement in maize (*Zea mays* L.) and bluegrass (*Poa pratensis* L.). Plant Cell Rep 15:489–494

Vain P, Worland B, Clarke MC, Richard G, Beavis M, Liu H, Kohli A, Leech M, Snape JW, Christou P, Atkinson H (1998) Expression of an engineered proteinase inhibitor (Oryzacystatin-I Δd86) for nematode resistance in transgenic rice plants. Theor Appl Genet 96:266–271

Vain P, Worland B, Kohli A, Snape JW, Christou P, Allen GC, Thompson WF (1999) Matrix attachment regions increase transgene expression levels and stability in transgenic rice plants and their progeny. Plant J 18:233–242

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

Wang MB, Upadhyaya NB, Brettell RIS, Waterhouse PM (1997) Intron-mediated improvement of a selectable marker gene for plant transformation using *Agrobacterium tumefaciens*. J Genet Breed 51:325–334

Wang MB, Li ZY, Matthews PR, Upadhyaya NM, Waterhouse PM (1998) Improved vectors for *Agrobacterium tumefaciens*-mediated transformation of monocot plants. Acta Hortic 461:401–407

Yin Z, Wang GL (2000) Evidence of multiple complex patterns of T-DNA integration into the rice genome. Theor Appl Genet 100:461–470