Generation of "backbone" free, low transgene copy plants by launching T-DNA from the *Agrobacterium* chromosome

Heiko Oltmanns[1], Bronwyn Frame[2], Lan-Ying Lee[1], Susan Johnson[1,3], Bo Li[1], Kan Wang[2], and Stanton B. Gelvin[1]

[1]Department of Biological Sciences, Purdue University, West Lafayette, IN 47907-1392

[2]Department of Agronomy and Plant Transformation Facility, Iowa State University, Ames, IA 50010-1010

[3]Present address: Monsanto Company, 800 North Lindbergh Blvd., Saint Louis, MO 63141

Correspondence should be addressed to Stanton B. Gelvin:

gelvin@bilbo.bio.purdue.edu

Telephone: 765-494-4939

FAX: 765-464-1693

Abstract

In both applied and basic research, *Agrobacterium*-mediated transformation is commonly used to introduce genes into plants. We investigated the effect of three *Agrobacterium tumefaciens* strains and five T-DNA origins of replication on transformation frequency, transgene copy number, and the frequency of integration of non-T-DNA portions of the T-DNA-containing vector ("backbone") into the genome of *Arabidopsis thaliana* and *Zea mays*. Launching T-DNA from the *picA* locus of the *Agrobacterium* chromosome increases the frequency of single transgene integration events and almost eliminates the presence of vector backbone sequences in transgenic plants. Along with novel *Agrobacterium* strains we have developed, our findings are useful for improving the quality of T-DNA integration events.

Since the generation of transgenic plants ~25 years ago, *Agrobacterium tumefaciens* has been widely used for introducing genes into plants for purposes of basic research as well as for generation of commercially used transgenic crops. For plant transformation, the gene of interest is placed between the left and right border repeats of *Agrobacterium* T-(transferred) DNA[1]. The T-DNA region harbouring the transgene is stably integrated into the plant genome by using an appropriate plant transformation protocol. T-DNA originates from the *Agrobacterium* Ti-(tumor inducing) plasmid. Because Ti-plasmids are large and difficult to manipulate, smaller T-DNA binary vectors are currently predominately used for generation of transgenic plants[2,3].

Although *Agrobacterium* has been used for plant transformation for more than two decades, problems using this bacterium remain. *Agrobacterium*-mediated transformation generally results in lower transgene copy numbers than do other transformation methods such as particle bombardment or polyethylene glycol-mediated transformation[4,5]. However, transformation frequently results in unwanted high copy number T-DNA integration events[5,6,7]. Multiple integration events, often coupled with inverted repeat T-DNA integration patterns, may affect the stability of transgene expression by silencing mechanisms[8]. An additional problem with *Agrobacterium*-mediated transformation is the propensity for DNA sequences outside the T-DNA region to integrate into the plant genome[5,9,10]. Integration of such vector backbone sequences can occur with high frequency. For example, Kononov et al.[9] detected backbone sequences in 75 % of tested transgenic

tobacco plants, and very often the entire vector backbone is introduced into the plant genome[11]. T-DNA vector backbones usually harbour bacterial antibiotic resistance genes that can create governmental regulatory concerns.

Here we show that launching T-DNA from the *A. tumefaciens* chromosome reduces integrated transgene copy number and almost eliminates the presence of T-DNA backbone sequences. We describe several plasmids and bacterial strains to facilitate use of this methodology.

Results

A. tumefaciens strains and T-DNA constructions

Our investigation utilized various combinations of the commonly used *A. tumefaciens* strains EHA101, GV3101, and LBA4404 with five different T-DNA binary systems. These *Agrobacterium* strains are non-oncogenic ("disarmed") and have been used for transformation of a large variety of plants. EHA101[12] harbours a derivative of the agropine/L,L-succinamopine-type Ti-plasmid pTiBo542, GV3101[13] a derivative of the nopaline-type Ti-plasmid pTiC58, and LBA4404[14] a derivative of the octopine-type Ti-plasmid pTiAch5. The tested T-DNA vectors contain an identical T-DNA region plus an *aadA* gene for bacterial selection for spectinomycin resistance. However, they contain different origins of replication (*ori*): the pVS ori, the pSa ori, the RK2 ori, and the pRiA4b ori. We furthermore analyzed the effect of launching T-DNA from the *Agrobacterium* C58 chromosome at the *picA* locus in strains EHA101 and GV3101. Disruption of this locus does not affect transformation[15] and we have generated vectors specifically designed to integrate genes into this locus[16].

We analyzed 14 different *A. tumefaciens* strain by replication origin combinations. The T-DNA region, derived from the binary vector pTF101.1[17], harbours a *bar* gene as a plant selectable marker under the control of a CaMV double 35S promoter (Fig. 1). Utilizing identical T-DNA regions with the same plant selectable marker and the identical non-T-DNA sequence proximal to the T-DNA left border in all constructions enabled us directly to compare results obtained for transformation frequencies, integrated transgene copy numbers, and backbone integration for all strain by construct combinations.

Effect of binary vector replication origin on binary vector copy number in *Agrobacterium*

We determined the copy number in *A. tumefaciens* of the four binary vectors used in our study. We placed a *bar* gene into the EHA105 chromosome and separately introduced each of the four T-DNA binary vectors into this strain. The resulting strains were incubated with or without acetosyringone to induce *vir gene* expression, and total bacterial DNA was extracted and subjected to DNA blot analysis using the *bar* gene as a probe. T-DNA binary vector copy numbers were determined by comparison of the signal intensity of the chromosomal band (one per cell) to the T-DNA binary vector band. Figure S1 shows that

plasmids containing the pSa origin are maintained at ~4 copies per cell. The copy numbers of plasmids containing the RK2 and pVS origins are 7-10 per cell, and plasmids containing the pRi origin replicate to 15-20 copies per cell. No significant differences were seen when the strains were incubated under inducing or non-inducing conditions.

Effect of *A. tumefaciens* strain and T-DNA *ori* on transformation frequency

We determined the effect of 14 *Agrobacterium* strain-by-construct combinations on transformation frequency of *Arabidopsis thaliana* and *Zea mays*. *Arabidopsis* was transformed using a floral dip protocol[18]. At least five transformation experiments were conducted for each vector-by-strain combination, and transformation frequencies were determined by analyzing 1500-4500 seeds per experiment (Fig. 2A). Transformation frequency was highly dependent upon the *A. tumefaciens* strain utilized. GV3101 resulted in the highest transformation frequencies (0.97-2.11%), whereas EHA101 and LBA4404 effected medium (0.09-0.58%) and low (0.01-0.12%) transformation frequencies, respectively. T-DNA replication origin had little effect on transformation frequency with one exception: Launching T-DNA from the *Agrobacterium* chromosome of EHA101 or GV3101 resulted in transformation frequencies lower than those of the other four T-DNA binary systems of the respective strain. Because the *picA* sequence of the Ach5 chromosome of LBA4404 does not share sufficient homology with the *picA* sequence of the C58-derived recombination vector we used[16], we were unable to integrate the T-DNA region into the *picA* locus of LBA4404.

Using an embryo inoculation protocol[19], we conducted five independent maize transformation experiments from which relative transformation frequencies for each of 12 strain-by-replication origin combinations (except chromosomal) were established (Fig. 2B). Four experiments were conducted to establish relative transformation frequencies for the two chromosomal replication origin combinations (in EHA101 and GV3101, Fig. 2B). As with *Arabidopsis*, launching T-DNA from the chromosomes of EHA101 and GV3101 resulted in low transformation frequencies (0.9%), whereas transformation frequencies were considerably higher (5-15%) when T-DNA was placed on a plasmid binary vector.

Launching T-DNA from the *Agrobacterium* chromosome results in a high percentage of plants containing a single integrated T-DNA copy

We investigated the number of copies of integrated T-DNA in transgenic events by DNA dot blot hybridization. T1 generation (heterozygous for T-DNA) *Arabidopsis* leaf samples were analysed using a *bar* gene-specific fragment (Fig. 1). Fig 3A shows, strikingly, that launching T-DNA from the *Agrobacterium picA* chromosomal locus resulted in 77-78% of the events containing a single transgene copy. The percentage of single transgene copy events resulting from use of "conventional" T-DNA binary vectors was much lower. Correspondingly, the average transgene copy number for events generated using "conventional" T-DNA binary vectors was higher (3.3-4.9 copies/genome) than was the integrated transgene copy number of events generated using strains with T-DNA launched from the bacterial chromosome (1.3-1.6 copies/genome; Supplemental Table 1).

Maize T-DNA copy number determinations were made using heterozygous T0 generation plants (one plant per event). Because the *bar* gene resulted in background hybridization signals in maize (H. Oltmanns, unpublished), we used the CaMV double 35S promoter fragment as the probe (Fig. 1). Fig. 3B shows that for events generated using "conventional" T-DNA binary vectors, the percentage of events carrying a single transgene copy ranged from 16-48%. Average T-DNA copy numbers resulting from the use of binary vectors ranged from 2.1-4.2 copies per cell (Supplemental Table 1). As with *Arabidopsis,* use of *Agrobacterium* strains containing chromosomal integration of T-DNA resulted in a higher percentage (58-64%) of single transgene copy number events; the average T-DNA copy number in these transgenic events was 1.7 copies per cell (Supplemental Table 1).

Although antibiotics were used to eradicate *Agrobacterium* after infection, bacterial cells might still contaminate selected transgenic *Arabidopsis* and maize plants. To eliminate the possibility of contaminating *Agrobacterium* DNA falsely increasing the apparent T-DNA copy number, we hybridized several membranes with the *Agrobacterium* chromosomal *picA* gene. We did not detect a hybridization signal using this probe (data not shown).

Launching T-DNA from the *Agrobacterium* chromosome mitigates integration of T-DNA backbone sequences in transgenic plants

Integration of T-DNA backbone sequences into the genome of transgenic plants can present regulatory problems, especially when bacterial antibiotic resistance genes are transferred. We investigated whether the *Agrobacterium* strain or T-DNA replication origin affects the frequency of backbone integration events. We used the spectinomycin resistance (*aadA*) gene immediately outside the T-DNA left border (Fig. 1) as a hybridization probe to detect backbone sequences within the genome of transgenic plants.

Figures 4A and 4B show the percentage of transgenic plants with backbone integration events for the 14 strain-by-origin combinations in *Arabidopsis* and maize, respectively. For *Arabidopsis*, one plant from 14 EHA101 events (7%) and 12 GV3101 events (8%), respectively, contained backbone sequences when they were generated by *Agrobacterium* strains in which the T-DNA was launched from the chromosome. For maize, no plants from 13 EHA101 events (0%) and one plant from 40 GV3101 events (3%), respectively, contained the vector backbone. On the other hand, the use of "conventional" T-DNA binary vectors resulted in a relatively high percentage of plants containing backbone sequences (47-67% for *Arabidopsis*; 19-55% for maize).

Vectors and *A. tumefaciens* strains to facilitate integration of T-DNA into the *Agrobacterium* chromosome

We generated two systems to facilitate integration of T-DNA into the *Agrobacterium* chromosome (Fig. 5). We first introduced a T-DNA region into the *picA* locus of the C58 chromosome. This T-DNA contains a CaMV 35S-*bar* gene as a plant selection marker, and a small region of pBluescript to provide homology for recombination with a variety of pBluescript-derived plasmids, such as the pSAT series of expression vectors[20]. Because plasmids harbouring the ColE1 *ori* cannot replicate in *A. tumefaciens*, ampicillin/carbenicillin-resistant bacteria can only be selected when the introduced plasmid co-integrates into the homologous T-DNA region (Fig. 5A). This bacterial strain contains an *aadA* gene directly outside the T-DNA left border to detect transfer of non-T-DNA sequences. In addition, the strain can be eliminated from co-cultivation reactions by using β-lactam antibiotics containing clavulanate, such as Timentin (L.-Y. Lee, unpublished).

Although the strain described in Fig. 5A is easy to use, it has the potential regulatory disadvantage of introducing a β-lactamase antibiotic resistance gene into the plant when T-DNA integrates. To eliminate such potential regulatory problems for plants destined for field release, we generated a second system to launch T-DNA from the *Agrobacterium* chromosome. A region homologous to the *Agrobacterium pgl*/*picA* locus was cloned into a plasmid containing a T-DNA. Gene expression cassettes from, e.g., pSAT vectors can be cloned into the rare-cutting multiple cloning site of this vector, and the entire vector can be introduced into *A. tumefaciens*. Spectinomycin resistance conferred by this plasmid results from homologous recombination with the *picA*/*pgl* region of the *A. tumefaciens* C58 chromosome (Fig. 5B). Thus, integration of T-DNA into the *A. tumefaciens* chromosome occurs without introduction of an antibiotic resistance gene into T-DNA.

Discussion

We studied the effect of three commonly used *A. tumefaciens* strains and five T-DNA replication origins on transformation frequency and the "quality" of T-DNA integration events in *Arabidopsis* and maize. Launching T-DNA from the *Agrobacterium* chromosome results in fewer integrated transgene copies and almost eliminates the presence of T-DNA binary backbone sequences in recovered transgenic events. However, these two advantageous aspects of plant transformation are accompanied by decreased transformation frequency. In *Arabidopsis*, this decrease is slight (2- to 4-fold) but it can be greater (approximately 10-fold) in maize. Whether scientists are willing to compensate decreased transformation frequency with a higher "quality" transformation event will depend upon the ease, time, and cost in generation of multiple transgenic events for different plant species.

Stable and predictable transgene expression is a major objective for both basic and applied research. Multiple integrated T-DNA copies, especially when combined with complex T-DNA integration patterns, can trigger transgene silencing[7,8,21]. The routine generation of single-copy transgenic events is therefore a major goal for agricultural biotechnology. Launching T-DNA from the *Agrobacterium* chromosome may provide one approach for achieving this goal.

Several studies have analyzed T-DNA locus and/or copy numbers in transgenic *Arabidopsis*. Feldmann[22] concluded that the average number of independently segregating, active transgene loci in his initial library of T-DNA tagged plants is 1.4. This value is similar to that of other T-DNA tagged collections in *Arabidopsis*[23,24] and rice[25,26]. However, the number of active loci in these plants is generally less than the number of integrated T-DNA molecules. T-DNA insertions frequently occur as partial or complete multimers in direct or inverted repeat orientation[7,22]. Bechtold et al.[27] showed that 70% of tested *Arabidopsis* transformants generated by a vacuum infiltration protocol carried direct or indirect tandem repeat copies of T-DNA. In this study the average T-DNA copy number in *Arabidopsis* ranged from 1.0-4.9, and in maize from 1.3-3.9 per diploid genome (Fig. 1, supplemental data). Our results agree with those of Galbiati et al.[28] who investigated 38,000 transgenic *Arabidopsis* plants generated by a floral dip method. Interestingly, transformation using *A. tumefaciens* LBA4404 resulted in an average transgene copy number lower than that resulting from transformation using the other tested strains. Grevelding et al.[29] investigated whether the transformation method affected transgene copy number in *Arabidopsis*. Most transgenic

plants produced by a leaf-disc inoculation method contained multiple T-DNA insertions, whereas root transformation resulted mostly in single T-DNA insertions. Therefore, the *Agrobacterium* strain, transformation method, and plant target tissue may influence the number of integrated T-DNA molecules.

Although T-DNA integration into the plant genome was experimentally shown almost 30 years ago, little is known about how many T-DNA strands are produced in *Agrobacterium* and transferred to the plant cell. It is likely that considerably more T-strands are transferred than are integrated[30,31]. T-DNA copy number of the chromosomal integration construction in the bacterial cell is '1' (except during replication before cell partition). Low integrated transgene copy numbers in plants may result from a limited number of T-strands transferred to the plant cell. We might therefore have expected to see a correlation between bacterial and plant T-DNA copy number using the different T-DNA replication origins because they replicate to different extents in the bacteria. However, we did not find such a correlation.

Integration of binary vector backbone sequences in transgenic plants is a common phenomenon[9.10,11]. Launching T-DNA from the *Agrobacterium* chromosome almost eliminates the presence of integrated T-DNA backbone sequences. In contrast, 47-67% of the *Arabidopsis* plants generated by *Agrobacterium* strains harbouring a T-DNA binary vector contained integrated vector backbone sequences. Although elimination of these sequences from transgenic plants is a major goal for agricultural biotechnology, only one previous report described a methodology to effect this result. By incorporating a lethal *barnase* gene into the non-T-DNA region of the binary vector, Hanson et al.[32] reduced the number of plants harbouring backbone sequences. However, up to 18% of the transgenic plants still carried backbone sequences.

Transfer of binary vector backbone sequences can occur when the T-DNA left border repeat is not recognized by the VirD2 endonuclease during processing of the T-DNA strand. It can also occur as a result of VirD2 linkage to the 5' end of the vector DNA directly outside the T-DNA left border, followed by transfer of the backbone in a manner analogous to that of T-DNA transfer[9,11,33]. If the T-DNA strand were derived from a binary plasmid and during T-DNA processing the left border repeat is skipped, T-DNA processing will either end at a sequence in the backbone that resembles a T-DNA border or, due to the circular nature of binary vectors, when the right border repeat is reached. However, if T-DNA is integrated into the bacterial chromosome, read-through at the T-DNA left border repeat could result in very long T-DNAs, theoretically as long as the *Agrobacterium* chromosome itself if no adequate

termination site were present. Although transfer of long T-DNA molecules is possible[34,35], it is less frequent than transfer of small T-DNAs. The observation that large T-DNAs only integrate into the plant genome very rarely offers a possible explanation for why chromosomal integration of T-DNA results in transgenic plants lacking backbone sequences. If the T-DNA left border were skipped during T-DNA strand processing (or if DNA transfer initiates from sequences directly to the left of the left T-DNA border), the resulting T-DNA would be too long for efficient transfer to the plant or integration into the plant genome. Although there might be concern that sequences from the bacterial chromosome next to the T-DNA right border could be integrated into the plant genome, probing of the DNA membranes with an *Agrobacterium picA* fragment, located immediately to the right of the right T-DNA border, failed to detect its presence (data not shown).

It is tempting to speculate whether integration of vector backbone sequences into plants is a consequence of simplifying *Agrobacterium*-mediated plant transformation by using small T-DNA binary vectors. Transfer of non-T-DNA portions of a large Ti-plasmid to plants is possible but rare: On average only one out of 80 transgenic tobacco calli contained a *nptII* gene positioned outside the T-DNA left border[36]. In contrast, Kononov et al.[9] detected vector "backbone" sequences in ~75% of transgenic tobacco plants generated using an *Agrobacterium* strain carrying a small T-DNA binary vector. These results suggest that backbone integration occurs more frequently when a small T-DNA binary vector is used. Although T-DNA binary vectors are ubiquitously used because of their ease of handling, we present here two vector systems to simplify launching T-DNA from the *Agrobacterium* chromosome.

Methods

Agrobacterium tumefaciens growth conditions

Agrobacterium tumefaciens strains were grown on solidified or liquid AB sucrose or yeast extract peptone medium[37] supplemented with appropriate antibiotics (rifampicin, 10 μg/ml; spectinomycin, 100 μg/ml; kanamycin, 25 μg/ml; gentamicin, 25 μg/ml).

T-DNA constructions

The T-DNA region and the bacterial *aadA* (spectinomycin resistance) gene (Fig.1) used in all T-DNA binary constructions derives from pTF101.1[17]. pTF101.1 contains a pVS1 origin of replication. To generate the various binary vectors, we replaced the pVS1 replication origin (*ori*) with those from other plasmids. For introducing the RK2 *ori*, we removed the pVS1 origin from pTF101.1 using *Sca*I and *Not*I and replaced it with a *Not*I/*Nru*I fragment from pBIN19[38], generating pTF::Bin19. For introducing the pSa ori, we removed the pVS1 ori from pTF101.1 using *Sca*I and *Nsi*I and replaced it with a *Pst*I/*Sac*II fragment from pUCD2[39]. All overhanging ends were made blunt using T4 DNA polymerase (New England Biolabs, Ipswich, MA, USA) to enable ligation. The resulting plasmid was designated pTF::UCD2. The pRiA4b origin isolated from *A. rhizogenes* A4[40] and cloned as a *Bam*HI-*Hin*dIII fragment into pBluescriptII KS^+, generating pBluescript::Ri. To confirm that the cloned pRi replication origin effects replication in *Agrobacterium*, pBluescript::Ri was transformed into *A. tumefaciens* by electroporation and plasmid DNA was isolated from carbenicillin-resistant colonies. pBluescript::pRi was digested with *Cla*I and the overhanging ends were made blunt using T4 DNA polymerase. The product was subsequently digested with *Not*I and cloned into pTF101.1 prior digested with *Sca*I and *Not*I to remove the pVS1 replication origin. The resulting plasmid was designated pTF::Ri.

The four T-DNA vectors pTF101.1, pTF::Bin19, pTF::UCD2 and pTF::Ri were separately transformed into *A. tumefaciens* EHA101[12], GV3101[13], and LBA4404[14] by electroporation.

Construction of *Agrobacterium* strains containing T-DNA integrated into the *pgl*/*picA* locus of the C58 chromosome

A 4.2 kb *Sca*I-*Nsi*I fragment containing the T-DNA region plus the *aadA* gene of pTF101.1 was inserted into the blunted *Spe*I and *Pst*I sites of the integration vector pE1931, generating pE2759. pE2759 was separately introduced into *A. tumefaciens* EHA101 and GV3101, generating *A. tumefaciens* At1586 and At1588, respectively. The eviction plasmid pPH1JI or pVK102 was introduced into *A. tumefaciens* At1586 and At1588, respectively, and colonies were selected on gentamicin and carbenicillin. Tetracycline-sensitive colonies (which had lost pE2759) were selected, generating *A. tumefaciens* At1589 and At1591, respectively, and recombination of the T-DNA region into the *pgl*/*picA* locus of the *Agrobacterium* chromosome was confirmed by DNA blot hybridization[16].

Construction of an *Agrobacterium* strain to facilitate integration of expression cassettes into the T-DNA region on the C58 chromosome

*A 1.549 kbp Pvu*II-*Sca*I fragment from pBluescript was cloned into the *Sma*I site of pTF101.1, generating pE3265. A 5.8 kbp *Sca*I-*Nsi*I fragment from pE3265 containing T-DNA, part of pBluescript, and the *aadA* gene was cloned into blunted *Spe*I and *Pst*I sites of pE1770, generating pE3349. pE3349 was introduced into *A. tumefaciens* EHA105, generating *A. tumefaciens* At1687. *E. coli* strains containing pVK102 and pRK230, respectively, were used to conjugate with *A. tumefaciens* At1687. The recombinant *Agrobacterium* strain containing the T-DNA borders, *aadA* gene, and a portion of pBluescript sequence was named *A. tumefaciens* At1702.

Construction of an integration binary vector to facilitate launching T-DNA from the A. tumefaciens C58 chromosome

A blunted *Eco*RI fragment containing the *pgl*/*picA* locus was cloned into the blunted *Nde*I site of pE3055. Removal of the *Nde*I fragment from pE3055 resulted in the loss of the pVS *ori* to create pE3361, a binary vector containing RCS (rare cloning sites), a plant selection marker (*bar* gene), a bacterial selection marker (*aadA* gene), and a ColE1 *ori*. This plasmid cannot replicate in *Agrobacterium*.

Determination of binary vector copy number in *Agrobacterium*

The four T-DNA vectors pTF101.1, pTF::Bin19, pTF::UCD2 and pTF::Ri were separately transformed into *A. tumefaciens* EHA101, grown under *vir* gene inducing or non-inducing conditions[41] and total DNA isolated by phenol:chloroform extraction. Equal amounts of DNA were digested with *Bgl*II and *Eco*RI and separated by electrophroesis through 0.9%

agarose gels. DNA was transferred to a Hybond N+ nylon membrane (Amersham Pharmacia Biotech, Piscataway, NJ, USA) and hybridized with a *bar* gene PCR fragment using conditions described below. Hybridization signal intensity was determined by scanning autoradiograms and using Labworks 4.6 Image Acquisition and Analysis Software (UVP, Upland, CA).

Agrobacterium-mediated transformation of _Arabidopsis thaliana_ and _Zea mays_

Arabidopsis thaliana (ecotype Wassilewskija, Ws-2) was transformed by a floral dip protocol[18]. Phosphinothricin-resistant plants were selected on Gamborg's B5 medium (Caisson Laboratories, Rexburg, ID) supplemented with 10 μg/ml phosphinothricin and 100 μg/ml timentin. Transformation frequencies were calculated as follows: [Number of phosphinothricin resistant plants/number of seeds tested] x 100. 10 mg of seeds correspond to ~500 seeds. Maize transformation was as previously described[19].

Genomic DNA extraction and DNA dot blot hybridization

Genomic DNA was extracted from 3-5 leaves of 3-4 week old *Arabidopsis thaliana* plants or approximately 100 mg maize leaf tissue according to Murray and Thompson[42]. Genomic DNA was quantified using a Gemini XPS microplate spectrofluorometer (Molecular Devices, Sunnyvale, CA, USA, excitation: 488 nm emission: 525 nm) using a Quant-iT™ PicoGreen® dsDNA Assay Kit (Invitrogen Inc.; Carlsbad, CA, USA).

DNA dot blots were prepared as follows: *Arabidopsis* (75 ng) or maize (1.5 μg) genomic DNA samples were denatured by adding NaOH and EDTA to final concentrations of 0.4 M and 10 mM, respectively, followed by 10 min incubation in boiling water. A Hybond N+ nylon membrane (Amersham Pharmacia Biotech, Piscataway, NJ, USA) was pre-wetted with water and placed between the layers of the dot blot apparatus. Samples were applied to the wells of the dot blot apparatus, incubated for 30 min, then drawn onto the nylon membrane using a gentle vacuum. DNA was cross-linked to the membrane by using a CL-1000 UV crosslinker (UVP, Upland, CA). The membranes were incubated in 2xSSC for 10 min and dried.

Probes for dot blot hybridizations were generated using random prime Ready-To-Go DNA Labelling Beads and ^{32}P-dCTP (both Amersham Pharmacia Biosciences). Unincorporated radioactive nucleotides were removed by Sephadex G-100 gel filtration. Membranes were pre-hybridized in 7% (w/v) SDS, 0.5 M sodium phosphate (pH 7.2) and 10

mM EDTA (pH 8.0) at 65°C for two hours. Hybridization was conducted overnight at 65°C. After hybridization, membranes were washed 2 times with 2xSSC, 0.1% (w/v) SDS, 10 mM EDTA (pH 8.0), then 2 times with 1xSSC, 0.1% (w/v) SDS, 10 mM EDTA and finally 2 times with 0.1xSSC, 0.1% (w/v) SDS, 10 mM EDTA at 65°C. Membranes were exposed at -80°C for autoradiography. Integrated dot density was determined using Labworks 4.6 Image Acquisition and Analysis Software (UVP, Upland, CA). For re-probing, membranes were stripped using boiling 0.1% (w/v) SDS twice. Calculations for T-DNA copy number standards were based on an *Arabidopsis thaliana* genome size of 125 Mbp[43] and a *Zea mays* genome size of 2500 Mbp[44]. The size of pTF101.1 is 9189 bp[17].

DNA dot blot membranes were hybridized sequentially with various probes. For determining T-DNA copy numbers of *Arabidopsis* plants, dot blot membranes were hybridized with a *bar* gene PCR fragment (Fig.1). In order to determine T-DNA copy numbers of maize plants, membranes were probed with a 759 bp *Pst*I/*Xho*I fragment from pTF101.1 harboring the double *CaMV 35S* promoter (Fig. 1). To detect vector "backbone" sequences in both *Arabidopsis* and maize DNA, a 656 bp PCR fragment derived from the non-T-DNA region next to the T-DNA left border of pTF101.1 (Fig. 1) was used as a probe. The fragment was amplified using pTF101.1 as a template and 5'-TCA CCG TAA CCA GC AAA TCA-3' and 5'-CTC GGC ACA AAA TCA CCA CT-3' as primers. A 3.1 kb *Eco*R1 fragment containing the *pgl*/*picA* locus[15] was used as a probe to check for the presence of contaminating *Agrobacterium* DNA in plant genomic DNA. All DNA fragments were gel purified prior to labelling using a QIAEX II Gel Extraction Kit (QIAGEN, Hilden, Germany). In order to normalize amounts of DNA in each dot, membranes were hybridized with genomic *Arabidopsis* or maize DNA.

Acknowledgments

This work was supported by a NSF Plant Genome grant DBI-0110023. We thank Marcy Main, Jennifer McMurray, and Tina Paque for their assistance in maize transformation experiments.

References

1. Gelvin, S.B. *Agrobacterium*-mediated plant transformation: The biology behind the "gene-jockeying" tool. *Microbiol. Mol. Biol. Rev*. **67**, 16-37 (2003).
2. de Frammond, A.J., Barton, K.A. & Chilton, M.-D. Mini-Ti-: A new vector strategy for plant genetic engineering. *Bio/technology* **5**, 262-269 (1983).
3. Lee, L.-Y. & Gelvin, S.B. T-DNA binary vectors and systems. *Plant Physiol.* **146**, 325-332 (2008).
4. Kohli, A. et al. Transgene integration, organization and interaction in plants. *Plant Mol. Biol*. **52**, 247–258 (2003).
5. Shou, H., Frame, B.R., Whitham, S.A. & Wang, K. Assessment of transgenic maize events produced by particle bombardment or *Agrobacterium*-mediated transformation. *Mol. Breeding* **13**, 201-208 (2004).
6. Deroles S.C. & Gardner, R.C. Analysis of the T-DNA structure in a large number of transgenic petunias generated by *Agrobacterium*-mediated transformation. *Plant Mol. Biol*. **11**, 365–377 (1988).
7. Jorgensen, R., Snyder, C. & Jones, J.D.G. T-DNA is organized predominantly in inverted repeat structures in plants transformed with *Agrobacterium tumefaciens* C58 derivatives. *Mol. Gen. Genet*. **207**, 471-477 (1987).
8. Jorgensen, R.A. et al. Chalcone synthase cosuppression phenotypes in petunia flowers: Comparison of sense vs. antisense constructs and single-copy vs. complex T-DNA sequences. *Plant Mol. Biol*. **31**, 957-973 (1996).
9. Kononov, M.E., Bassuner, B. & Gelvin, S.B. Integration of T-DNA binary vector "backbone" sequences into the tobacco genome: evidence for multiple complex patterns of integration. *Plant J*. **11**, 945-957 (1997).
10. Wenck, A., Czako, M., Kanevski, I. & Marton, L. Frequent collinear long transfer of DNA inclusive of the whole binary vector during *Agrobacterium*-mediated transformation. *Plant Mol. Biol.* **34**, 913 922 (1997).
11. De Buck, S., de Wilde, C., van Montagu, M. & Depicker, A. T-DNA vector backbone sequences are frequently integrated into the genome of transgenic plants obtained by *Agrobacterium*-mediated transformation. *Mol. Breed*. **6,** 459-468 (2000).
12. Hood, E.E., Fraley, R.T., & Chilton, M.-D. Virulence of *Agrobacterium tumefaciens* strain A281 on legumes. *Plant Physiol*. **83**, 529-534 (1987).
13. Koncz, C. & Schell, J. The promoter of TL-DNA gene 5 controls the tissue-specific

expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* **204**, 383–396 (1986).

14. Ooms, G., Hooykaas, P.J.J., Moolenaar, G. & Schilperoort, R.A. Crown gall plant tumors of abnormal morphology, induced by *Agrobacterium tumefaciens* carrying mutated octopine Ti plasmids; analysis of T-DNA functions. *Gene* **14**, 33-50 (1981).
15. Rong, L. et al. *picA*, a novel plant-inducible locus on the *Agrobacterium tumefaciens* chromosome. *J. Bacteriol.* **172**, 5828–5836 (1990).
16. Lee, L.-Y., Humara, J.M. & Gelvin, S.B. Novel constructions to enable the integration of genes into the *Agrobacterium tumefaciens* C58 chromosome. *Mol. Plant-Microbe Interact.* **14**, 577-579 (2001).
17. Paz, J.M. et al. Assessment of conditions affecting *Agrobacterium*-mediated soybean transformation using the cotyledonary node explant. *Euphyt.* **136**, 167-179 (2004).
18. Clough, S.J. & Bent, A.F. Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).
19. Frame, B.R. et al. *Agrobacterium tumefaciens*-mediated transformation of maize embryos using a standard binary vector system. Plant Physiol. **129**, 13-22 (2002).
20. Chung, S.-M., Frankman, E.L. & Tzfira, T. A versatile vector system for multiple gene expression in plants. *Trends Plant Sci.* **10**, 357-361 (2005).
21. Stam, M. et al. 1997. Post-transcriptional silencing of chalcone synthase in *Petunia* by inverted transgene repeats. *Plant J.* **12**, 63-82 (1997).
22. Feldmann, K.A. T-DNA insertion mutagenesis in *Arabidopsis*: Mutational spectrum. *Plant J.* **1**, 71-82 (1991).
23. Alonso, J.M. et al. Genome-Wide Insertional Mutagenesis of *Arabidopsis thaliana*. *Science* **301**, 653-657 (2003).
24. McElver, J. et al. Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. *Genetics* **159**, 1751-1763 (2001).
25. Barakat, A. et al. The distribution of T-DNA in the genomes of transgenic *Arabidopsis* and rice. *FEBS Lett.* **471**, 161-164 (2000).
26. Jeon, J.-S. et al. T-DNA insertional mutagenesis for functional genomics in rice. *Plant J.* **22**, 561-570 (2000).
27. Bechtold, N., Ellis, J. & Pelletier, G. *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C. R. Acad. Sci. Paris Life Sci.* **316**, 1194-1199 (1993).
28. Galbiati, M. et al. Large-scale T-DNA mutagenesis in *Arabidopsis* for functional genomic

analysis. Funct. Integr. *Genomics* **1**, 25–34 (2000).

29. Grevelding C. et al. Single-copy T-DNA insertions in *Arabidopsis* are the predominant form of integration in root-derived transgenics, whereas multiple insertions are found in leaf discs. *Plant Mol. Biol.* **23,** 847–860 (1993).

30. Atmakuri, K. et al. *Agrobacterium* ParA/Mind-like VirC1 spatially coordinates early conjugative DNA transfer reactions. *EMBO J.* **26**, 2540-2551 (2007).

31. Virts, E.L. & Gelvin, S.B. Analysis of transfer of tumor-inducing plasmids from *Agrobacterium tumefaciens* to *Petunia* protoplasts. *J. Bacteriol.* **162**, 1030-1038 (1985).

32. Hanson, B. et al. A simple method to enrich an *Agrobacterium*-transformed population for plants containing only T-DNA sequences. *Plant J.* **19,** 727-734 (1999).

33. Durrenberger, F., Crameri, A., Hohn, B. & Koukolikova-Nicola, Z. Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. *Proc. Natl. Acad. Sci. USA* **86**, 9154-9158 (1989).

34. Hamilton, C. M., Frary, A., Lewis, C. & Tanksley, S.D. Stable transfer of intact high molecular weight DNA into plant chromosomes. *Proc. Natl. Acad. Sci. USA* **93,** 9975-9979 (1996).

35. Miranda, A. et al. *Agrobacterium tumefaciens* transfers extremely long T-DNAs by a unidirectional mechanism. *J. Bacteriol.* **174**, 2288-2297 (1992).

36. Ramanathan, V. & Veluthambi, K. Transfer of non-T-DNA portions of the *Agrobacterium tumefaciens* Ti plasmid pTiA6 from the left terminus of TL-DNA. *Plant Mol. Biol.* **28**, 1149-1154 (1995).

37. Lichtenstein, C. & Draper. J. DNA cloning: A practical approach. In *Genetic Engineering of Plants*, (ed. Glover, D.M.) 67-119 (IRL Press, Washington, D.C. Vol. 2, 1986).

38. Bevan, M. Binary *Agrobacterium* vectors for plant transformation. *Nucl. Acids Res.* **12**, 8711-8721 (1984).

39. Close, T.J., Zaitlin, C. & Kado, C.I. Design and development of amplifiable broad-host-range cloning vectors: Analysis of the *vir* region of *Agrobacterium tumefaciens* plasmid pTiC58. *Plasmid* **12**, 111-118 (1984).

40. Jouanin, L., Tourneur, J., Tourneur, C. & Casse-Delbart, F. Restriction maps and homologies of the three plasmids of *Agrobacterium rhizogenes* strain A4. *Plasmid* **16**, 124-134 (1986).

41. Gelvin, S.B. *Agrobacterium* virulence gene induction. In *Methods in Molecular Biology:*

Agrobacterium protocols. (ed. Wang, K.). 77-84 (Humana Press, Totowa, NJ. Vol. 44, 2006.

42. Murray, M.G. & Thompson, W.F. Rapid isolation of high-molecular-weight plant DNA. *Nucl. Acids Res.* **8**, 4321-4325 (1980).

43. The *Arabidopsis* Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815 (2000).

44. Arumuganathan, K. & Earle, E.D. Estimation of nuclear DNA content of plants by flow cytometry. *Plant Mol. Biol. Rep.* **9,** 229-233 (1991).

Figure legends

Figure 1. Schematic map of the T-DNA and neighbouring regions of the base vector pTF101.1[17] used in constructing the various vectors. Black bars represent fragments used for DNA blot hybridizations experiments. LB and RB, T-DNA left and right borders, respectively; *aadA*, gene encoding spectinomycin resistance; E, translational leader from Tobacco Etch Virus; Tvsp, VSP terminator sequence; *bar*, gene conferring resistance to Basta®/Bialophos®/phosphinothricin; 2x P35S, CaMV double 35S promoter sequence.

Figure 2. Transformation frequencies of *Arabidopsis thaliana* **(A)** and maize **(B)** with different *Agrobacterium* strain-by-origin of replication combinations. *Arabidopsis thaliana* was transformed by a floral dip protocol, and maize by embryo inoculation. Error bars represent the standard error among different transformation experiments. At least four independent transformation experiments were conducted for both *Arabidopsis thaliana* and maize. The total number of seeds screened **(A)** or embryos infected **(B)** to calculate transformation frequencies is indicated above each error bar. For maize, 5 independent experiments were conducted to establish relative transformation frequencies for each of 12 strain-by-replication origin combinations (except chromosomal). Embryos from each of the 22 ears dissected were shared 12 ways at infection. Subsequently, transformation frequencies for the two chromosomal replication origin combinations (in EHA101 and GV3101) were established by inoculating them beside the pTF101.1 replication origin vector in EHA101 and GV3101. In four independent experiments, embryos from 21 dissected ears were shared four ways at infection. Throughout the remainder of this study, additional maize transformation experiments using these 14 strain-by-replication origin combinations were conducted as needed to recover adequate numbers of transgenic events for molecular analysis.

Figure 3. The percentage of events containing a single copy of integrated T-DNA following transformation by the 14 T-DNA replication origin-by-*Agrobacterium* strain combinations for *Arabidopsis thaliana* (**A**) and maize (**B**). Copy number reconstructions were done by DNA dot blot experiments using T-DNA specific probes (Fig. 1). Analysis was performed on heterozygous T1 generation *Arabidopsis thaliana* plants and heterozygous regenerated T0 generation maize plants. Numbers over the bars represent the number of independent events analyzed.

Figure 4. The percentage of independent events containing vector "backbone" sequences following transformation by the 14 T-DNA replication origin-by-*Agrobacterium* strain combinations for *Arabidopsis thaliana* (**A**) and maize (**B**). Blots were hybridized with a "backbone"-specific probe, the *aadA* gene (Fig. 1). Numbers over the bars represent the number of independent events analyzed.

Figure 5. Systems for introducing T-DNA harbouring transgene expression cassettes into the *pgl*/*picA* region of the *A. tumefaciens* chromosome. **(A)** T-DNA containing a CaMV 35S promoter-*bar* gene plant selection marker and a fragment of pBluescript, and an *aadA* gene to the left of the T-DNA left border, has been inserted into the *A. tumefaciens* C58 chromosome. When a pUC-derived plasmid containing an expression cassette is introduced into this strain, carbenicillin-resistant colonies can only be obtained if homologous recombination occurs resulting in a co-integrate between this plasmid and the pBluescript region of the T-DNA; **(B)** In this system, a transgene expression cassette is first cloned into a T-DNA region located on a plasmid harbouring a ColE1 *ori*. The plasmid also contains the *pgl*/*picA* region of the *A. tumefaciens* C58 chromosome. When this plasmid is introduced into *A. tumefaciens* C58-derived strains, carbenicillin-resistant colonies can only be obtained if homolgous recombination occurs between the *pgl*/*picA* regions of the plasmid and the chromosome. LB, T-DNA left border; RB, T-DNA right border; ori, origin of replication; amp^r, β-lactamase gene conferring ampicillin/carbenicillin resistance upon the bacterium; aadA, spectinomycin resistance gene; 2X35S, double CaMV 35S promoter; bar, gene conferring resistance to Basta®/Bialophos®/phosphinothricin; P_{ocs}, octopine synthase promoter; T_{ocs}, octopine synthase terminator; RCS, multiple rare cutting sites flanking various pSAT vectors. Broken crossed lines indicate homologous recombination events.

Supplemental Figure 1. T-DNA binary vector copy numbers caused by various replication origins in *A. tumefaciens*

A. tumefaciens cells harboring a fragment of the *bar* gene in the chromosome and a T-DNA binary vector were grown under *vir* gene inducing and non-inducing conditions. DNA blot analysis was done to determine copy numbers of various T-DNA binary vectors. Membranes were probed with a *bar* gene fragment. The upper band represents the chromosomal integrated *bar* gene fragment (copy number per cell: 1), the lower band the T-DNA binary vector.

Figure 1

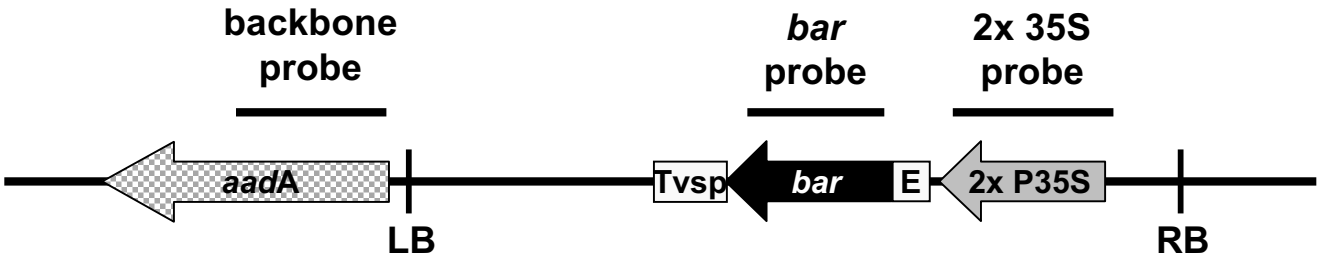

backbone probe
bar probe
2x 35S probe
aadA
LB
Tvsp
bar
E
2x P35S
RB

Figure 2

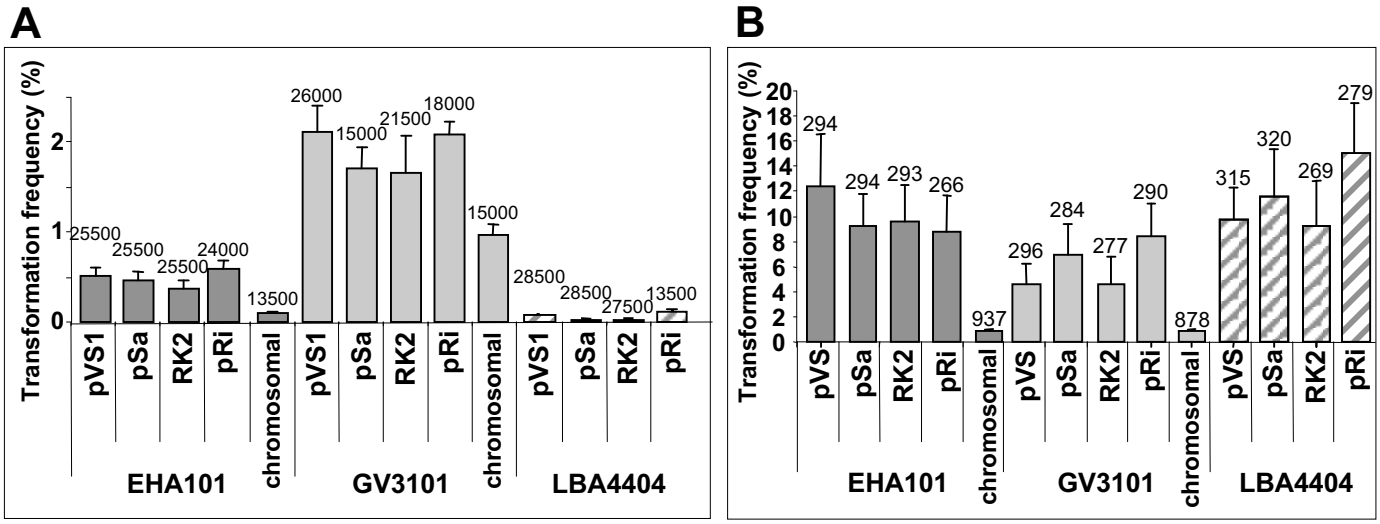

A
B
Transformation frequency (%)
25500
25500
25500
24000
13500
26000
15000
21500
18000
15000
28500
28500
27500
13500
294
294
293
266
937
296
284
277
290
878
315
320
269
279
pVS1
pSa
RK2
pRi
chromosomal
pVS
EHA101
GV3101
LBA4404

Figure 3

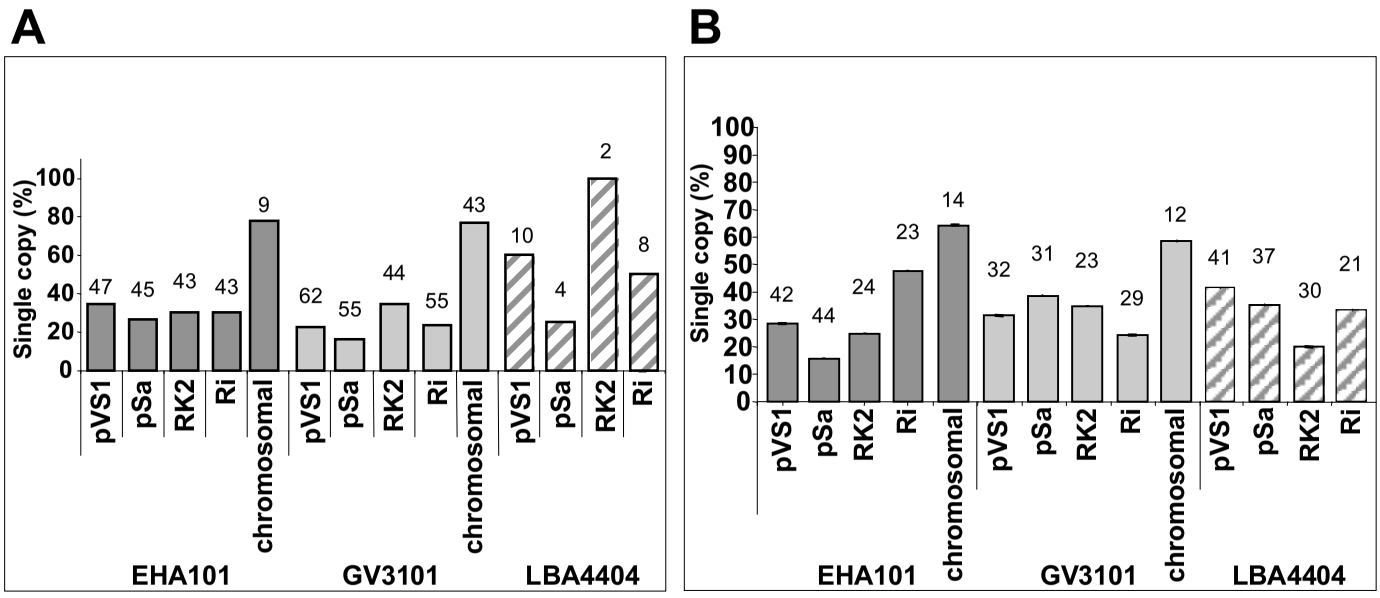

A
B
Single copy (%)
100
90
80
70
60
50
40
30
20
10
0
47
45
43
43
9
62
55
44
55
43
10
4
2
8
42
44
24
23
14
32
31
23
29
12
41
37
30
21
pVS1
pSa
RK2
Ri
chromosomal
EHA101
GV3101
LBA4404

Figure 4

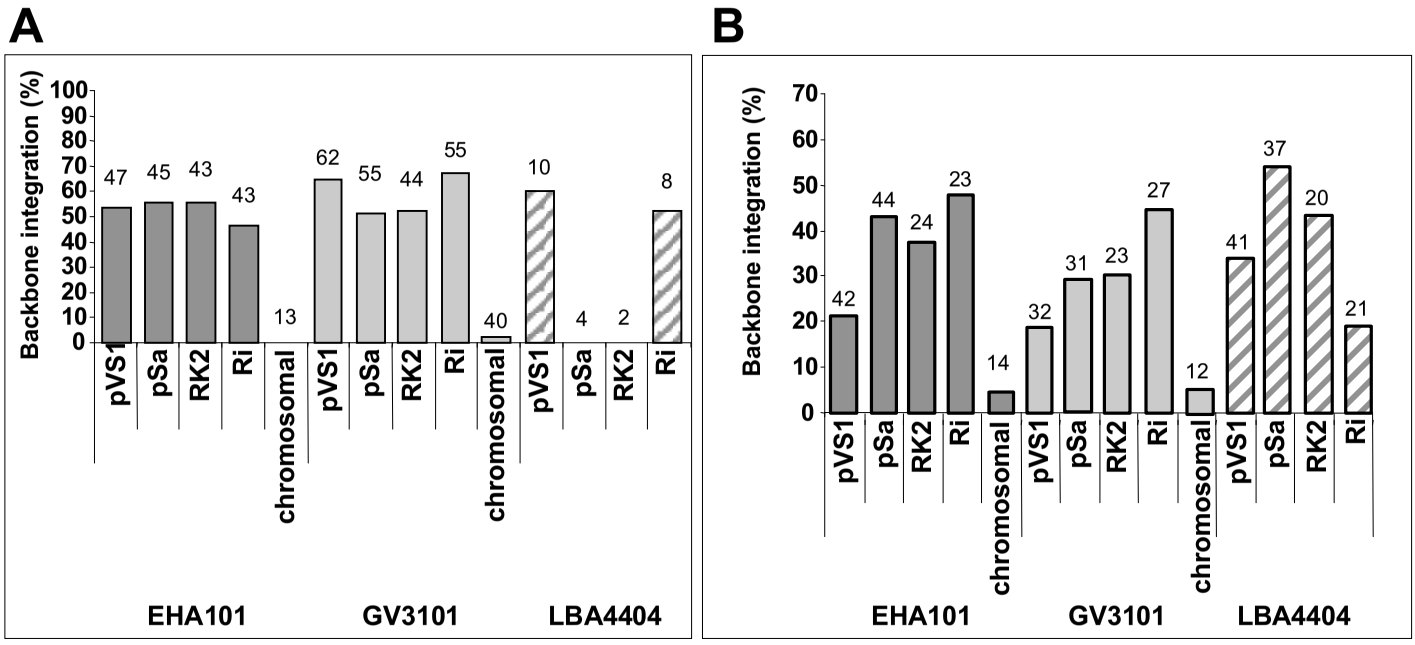
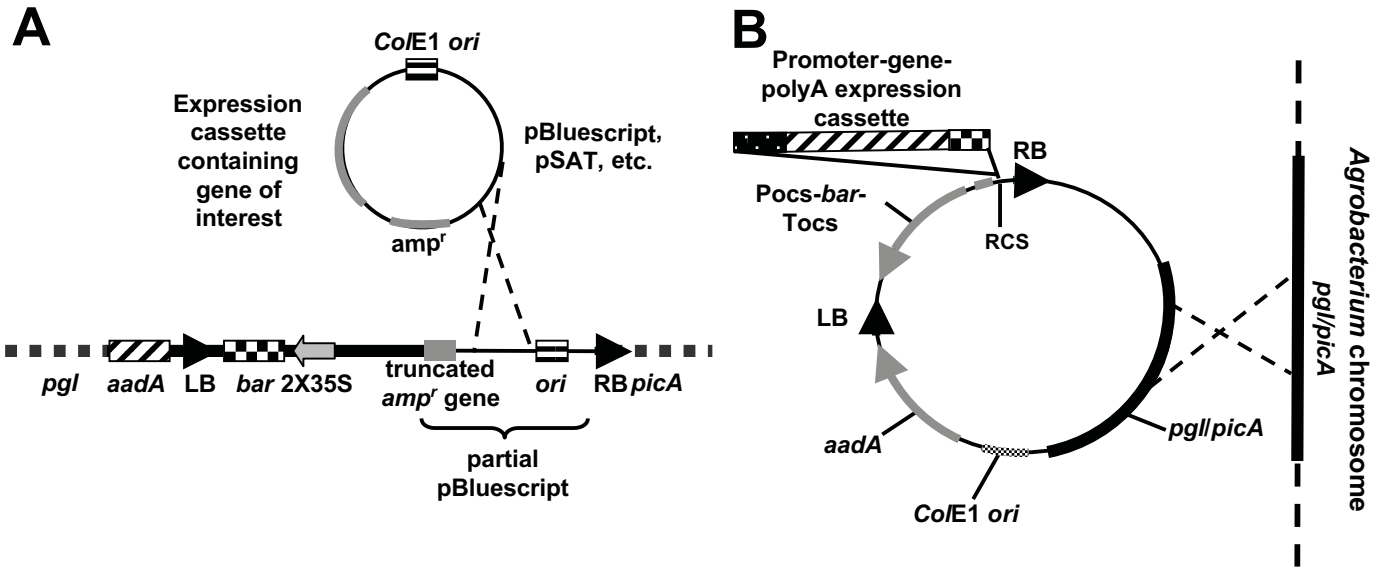

A
B
Backbone integration (%)
pVS1
pSa
RK2
Ri
chromosomal
EHA101
GV3101
LBA4404
47
45
43
43
13
62
55
44
55
40
10
4
2
8
42
44
24
23
14
32
31
23
27
12
41
37
20
21

Figure 5


A
ColE1 ori
Expression cassette containing gene of interest
pBluescript, pSAT, etc.
amp^r
pgl
aadA
LB
bar 2X35S
truncated amp^r gene
ori
RB
picA
partial pBluescript
B
Promoter-gene-polyA expression cassette
RB
Pocs-bar-Tocs
RCS
LB
aadA
ColE1 ori
pgl/picA
Agrobacterium chromosome
pgl/picA