MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2011) **11**, 890–902 doi: 10.1111/j.1755-0998.2011.03017.x

Efficient screening of transgenic plant lines for ecological research

KLAUS GASE, ARNE WEINHOLD, TOHIR BOZOROV, STEFAN SCHUCK and IAN T. BALDWIN
Department of Molecular Ecology, Max-Planck-Institute for Chemical Ecology, Hans-Knoell-Strasse 8, 07745 Jena, Germany

Abstract

Plants stably transformed to manipulate the expression of genes mediating ecological performance have profoundly altered research in plant ecology. *Agrobacterium*-mediated transformation remains the most effective method of creating plants harbouring a limited number of transgene integrations of low complexity. For ecological/physiological research, the following requirements must be met: (i) the regenerated plants should have the same ploidy level as the corresponding wild-type plant and (ii) contain a single transgene copy in a homozygous state; (iii) the T-DNA must be completely inserted without vector backbone sequence and all its elements functional; and (iv) the integration should not change the phenotype of the plant by interrupting chromosomal genes or by mutations occurring during the regeneration procedure. The screening process to obtain transformed plants that meet the above criteria is costly and time-consuming, and an optimized screening procedure is presented. We developed a flow chart that optimizes the screening process to efficiently select transformed plants for ecological research. It consists of segregational analyses, which select transgenic T_1 and T_2 generation plants with single T-DNA copies that are homozygous. Indispensable molecular genetic tests (flow cytometry, diagnostic PCRs and Southern blotting) are performed at the earliest and most effective times in the screening process. qPCR to quantify changes in transcript accumulation to confirm gene silencing or overexpression is the last step in the selection process. Because we routinely transform the wild tobacco, *Nicotiana attenuata*, with constructs that silence or ectopically overexpress ecologically relevant genes, the proposed protocol is supported by examples from this system.

Keywords: ecological genomics, field releases, plant transformation, ploidy, reverse genetics, screening

Received 14 December 2010; revision received 14 March 2011; accepted 17 March 2011

Introduction

Stably transformed plants have proven a powerful tool in ecological research to investigate the ecological relevance of particular genes. For this, the inserted sequence elements are designed either to silence intrinsic or to overexpress intrinsic or heterologous genes of interest. This reverse genetics approach allows for the creation of transgenic lines with either abnormally low or high levels of transcripts for a particular gene in an otherwise isogenic background. These isogenic lines provide a particularly efficient means of studying the fitness consequences of a given gene's expression (Steppuhn *et al.* 2004; Zavala *et al.* 2004; Kang *et al.* 2006; Schwachtje *et al.* 2008).

Two general strategies are used for plant transformation: *Agrobacterium*-based transformation and a group of unrelated techniques collectively referred to as 'direct DNA transfer' (Kohli *et al.* 2003). Direct DNA transfer methods such as particle bombardment (Christou 1992) often result in transgenic loci with a high transgene copy number (often more than 40) (Kohli *et al.* 2003; Latham *et al.* 2006). *Agrobacterium*-based transformation procedures produce lines with less complex transgenic loci, but still, the integration of multiple T-DNA copies into a limited number of loci is common (De Buck *et al.* 2009; Bhat & Srinivasan 2002). The transformation mediated by *Agrobacterium* involves the transfer of the T-DNA molecule to the eukaryotic host cell and its integration into the host genome. The machinery required for this process comprises proteins encoded by bacterial chromosomal genes and Ti-plasmid virulence genes as well as the host intracellular transport and DNA repair machinery (Tzfira & Citovsky 2006; Gelvin 2009; Lacroix & Citovsky 2009). The T-DNA region is defined by its left and right border sequences, two 25-bp inverted repeats, originally present on the Ti plasmid (Tzfira *et al.* 2004). To allow the creation of transgenic plant lines with desired insertions, the

Correspondence: Ian T. Baldwin, Fax: +49 3641 571102; E-mail: baldwin@ice.mpg.de

T-DNA borders were transferred to a binary vector, and the sequences that should be integrated into the plant genome are cloned between these border sequences (Hoekema *et al.* 1983).

T-DNA integration occurs randomly throughout the plant genome (Gelvin & Kim 2007; Kim *et al.* 2007) by non-homologous end-joining (Gheysen *et al.* 1991) and is accompanied by deletions and rearrangements of the T-DNA flanks and of the target DNA near the integration site (Latham *et al.* 2006; Muller *et al.* 2007; Gambino *et al.* 2009). The integration of DNA from binary plant transformation vectors is not always limited to the region between the T-DNA borders. 'Read-through' events can occur and result in the unwanted cotransfer of vector backbone sequences, such as bacterial resistance genes. After *Agrobacterium*-mediated T-DNA transfer, binary vector backbone sequences can be detected in up to 75% of the transgenic plants (Kononov *et al.* 1997; Wenck *et al.* 1997; McCormac *et al.* 2001; Kohli *et al.* 2003; Lange *et al.* 2006; Gambino *et al.* 2009). Regulatory agencies that govern the release of transgenic organisms do normally not permit the release of plants carrying such sequences. *Agrobacterium* infection and, in most cases, the regeneration process necessary to regenerate a transgenic plant via cell culture and callus formation from a single transformed plant cell can lead to somoclonal variations (Bhat & Srinivasan 2002), including genome-wide mutations (Latham *et al.* 2006) and polyploidization (Bubner *et al.* 2006).

Agrobacterium-based transformation has proven to be the best-suited approach to produce transgenic plant lines with single-copy T-DNA insertions (De Buck *et al.* 2009; Bhat & Srinivasan 2002; O'Malley & Ecker 2010, Meza *et al.* 2002; Sallaud *et al.* 2003; Olhoft *et al.* 2004; Yu *et al.* 2010). Transformed plant lines may contain two or more independent transgenic loci, but sufficient lines with single T-DNA insertions in a single locus can be expected (De Buck *et al.* 2004). This makes *Agrobacterium*-based transformation the preferred method to create transgenic plants for ecological research.

If the transgenic plants are to be used to answer questions about the organismic-level consequences of a particular gene's expression, then the transformants must fulfil a number of strict criteria: The transgenic line should have the same ploidy level as the plant that was transformed (Schwachtje & Baldwin 2008). The insert should comprise the complete T-DNA originating from the binary vector used for transformation without deletions, rearrangements or other mutations. In each line, only one transgenic locus with a single T-DNA copy should be present, and each line should be homozygous with respect to this locus. The integration site of the transgenic DNA should not disrupt other functional genes, which could confound the analysis of phenotypes associated with the gene of interest. To exclude mutations arising from T-DNA insertion or regeneration, at least two independent lines created with the same T-DNA should be evaluated, and both should exhibit the same phenotype, as the chances that the T-DNA inserted twice into the same functional gene are vanishingly small (Schwachtje *et al.* 2008).

For many plant scientists interested in the function of genes at a whole plant level, the utilization of genetically modified plants silencing or overexpressing a particular gene of interest is the most powerful means of answering functional questions. One important challenge for this approach is the lack of transformation and selection procedures for many plant species of ecological interest. This protocol paper will help ecological researchers to create and select transgenic lines that fulfil the requirements of their research questions. We have developed an *Agrobacterium*-mediated transformation system for the wild tobacco *Nicotiana attenuata* Torrey ex Watson, an ecological model plant (Baldwin 2001). Based on our experiences with this method in producing transgenic lines for use in ecological research, we describe a protocol that optimizes the efficiency of the transformation and selection system and can be applied to other plants of ecological interest.

The transformation procedure is laborious and comprises the construction of appropriate binary plant transformation vectors, *Agrobacterium*-based transformation and the selection and regeneration of the first transgenic generation (T_0 plants). Because the T_0 generation originates from regenerated calli, growth regulators used during regeneration may have a lasting effect on plant performance (Bhat & Srinivasan 2002). It also cannot be excluded that T_0 plants are chimeras regenerated from two or more transformed plant cells, or have increased ploidy levels (Bubner *et al.* 2006). Because of these properties, T_0 plants should not be used for experiments. The first non-chimeric transgenic plant generation that did not undergo the regeneration process (T_1) is produced after self-pollination of the T_0 flowers, and the resulting T_1 seeds should be germinated on medium containing the selective antibiotic. This allows the identification of individuals carrying T-DNA insertions and indicates possible silencing problems because of promoter methylation that may interfere with the transcription of the transgene (Stam *et al.* 1997).

The T_1 generation represents an important intermediate stage in the selection of transgenic lines that fulfil the criteria for ecological research. In most cases, this generation contains homozygous lines with a single T-DNA insertion locus. The ploidy level, the arrangement of the T-DNA insertion and often the level of transcription of the transgene will—if no further chromosomal rearrangements occur—remain unchanged in subsequent inbred generations.

Further inheritance studies with the T_2 generation, obtained by self-pollination of the T_1 generation and germination, are necessary to identify candidate transgenic T_1 lines that harbour single transgene insertions in a homozygous state. Each of the selected lines should be evaluated by a number of molecular genetic analyses to confirm that the above criteria are met. These analyses include flow cytometry to determine the ploidy level of the transgenic lines, appropriate diagnostic polymerase chain reactions (PCR) to ensure the insertion of the complete T-DNA into the plant chromosome and to exclude vector backbone integrations, Southern blotting to confirm single T-DNA insertions into the plant chromosome and qPCR to evaluate silencing efficiency or the level of transcript accumulation of the transgene.

The transformation procedure and the inheritance studies necessary to produce and identify the required transgenic lines consume an enormous amount of human and material resources. The time to grow one generation from germination to seed ripening can take, depending on the species, several months, in the case of *N. attenuata* 3–4 months. It is therefore important to limit the screening effort to what is absolutely necessary. In this study, we present a workflow protocol that enables researchers to produce and select transgenic lines in an optimized screening process. We analyse each step of the selection process using examples from the screening of *N. attenuata* lines and discuss possible variations.

Materials and methods

Construction of plasmids for the transformation of N. attenuata

The initial vector for the construction of the pSOL8/pSOL9 plasmid series was pSOL3LOX (10.6 kb) (Bubner *et al.* 2006). After replacing the *nptIII* (aminoglycoside phosphotransferase class III) gene with *nptII* from pAcGFP1-1 (http://www.clontech.com) and replacing the inverted repeat gene fragments with *N. attenuata DCL2* cDNA (HQ698849) fragments, thereby providing optimized cloning sites, cloning vector pSOL8DCL2 (10.2 kb; HQ698851) was created.

The pSOL8 series inverted repeat gene silencing plasmids (Fig. 1a) were created by replacing the *Xho*I-*Sac*I and *Pst*I-*Hin*dIII inverted repeat fragments of pSOL8DCL2 with inverted repeat PCR fragments (0.3–0.6 kb) of the following *N. attenuata* genes: pSOL8DC3 [10.8 kb; RNA-dependent RNA polymerase 1 gene (DQ988990) combined with WRKY6 gene (AY456272)], pSOL8PNRP [10.2 kb; gene similar to *Arabidopsis thaliana* putative nematode resistance protein mRNA (AY080778)], pSOL8AEP65 [10.2 kb; gene similar to *N. tabacum* Avr9/Cf-9 rapidly elicited protein 65 mRNA (AF211539)] and pSOL8AEP150 [10.1 kb; gene similar to *N. tabacum* Avr9/Cf-9 rapidly elicited protein 150 mRNA (AY775041)].

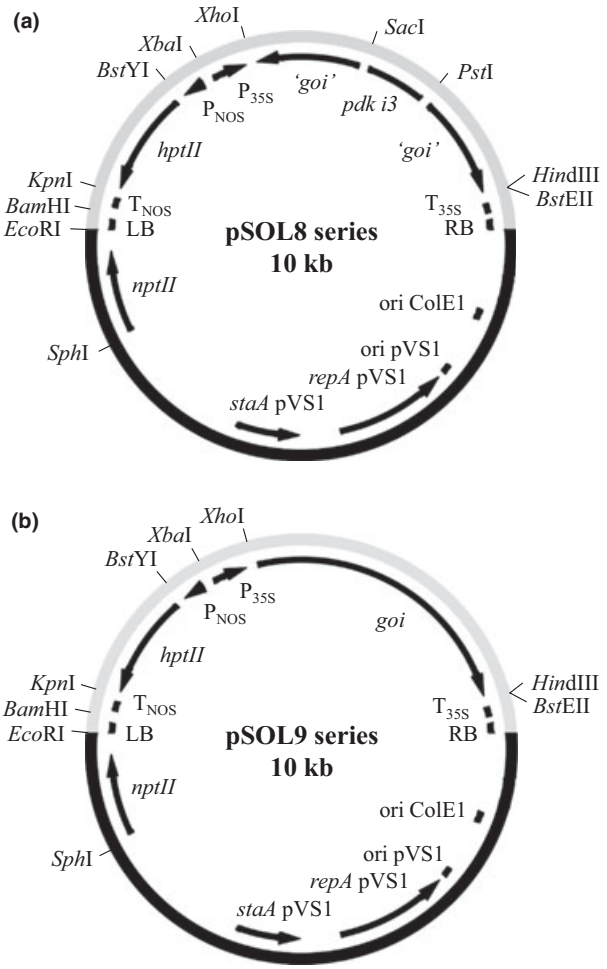

Fig. 1 The pSOL8 series (a) and pSOL9 series (b) binary plant transformation vectors. Abbreviations: LB/RB, left/right border of T-DNA; P_{NOS}/T_{NOS}, promoter/terminator of the nopaline synthase gene from the Ti plasmid of *Agrobacterium tumefaciens*; P_{35S}/T_{35S}, 35S promoter/terminator from cauliflower mosaic virus; *hptII*, hygromycin phosphotransferase gene from pCAMBIA-1301 (AF234297); *goi*, gene of interest; i, intron 3 of *Flaveria trinervia pdk* gene for pyruvate, orthophosphate dikinase; *nptII*, aminoglycoside phosphotransferase class II; ori, origin of replication.

The following pSOL9 series gene expression plasmids (Fig. 1b) were obtained by replacing the 1.4 kb *Xho*I-*Hin*dIII-fragment of pSOL8DCL2 with appropriate PCR fragments (0.2–0.4 kb) allowing the overexpression of the following antimicrobial protein genes: pSOL9CAP [9.2 kb; *Capsella bursa-pastoris* antimicrobial peptide mRNA (HQ698850)], pSOL9ESC [9.0 kb; synthetic gene similar to *Rana plancyi fukienensis* mRNA for esculentin-1P precursor protein (AJ968397)], pSOL9FAB [9.0 kb; synthetic gene similar to *Vicia faba* fabatin precursor

mRNA (EU920043)], pSOL9ICE [9.0 kb; *Mesembryanthemum crystallinum* antimicrobial peptide 1 precursor mRNA (AF069321)], pSOL9LEA [9.1 kb; synthetic gene similar to *Leonurus japonicus* antimicrobial protein mRNA; (AY971513)], pSOL9PNA [9.1 kb; synthetic gene similar to *Ipomoea nil* antifungal protein mRNA (U40076)], pSOL9SSP [9.0 kb; synthetic gene for *Aptenodytes patagonicus* Spheniscin-2 (P83430)] and pSOL9VRD [9.0 kb; synthetic gene similar to *Vigna nakashimae* defensin-like protein gene (AY856095)]. As reference examples, the sequences of pSOL8DC3 and pSOL9CAP have been submitted to GenBank (HQ698853 and HQ698852).

Plant transformation, regeneration and cultivation

Transformation of *N. attenuata* was performed as described in Kruegel *et al.* 2002. In brief, hypocotyls from 8- to 10-day-old seedlings were cut into 3-mm-long pieces with a scalpel that previously had been dipped into a culture of *A. tumefaciens* LBA4404 (Invitrogen, http://www.invitrogen.com) carrying the binary plant transformation vector. After 3 days of cocultivation with *Agrobacterium*, the transgenic tissue went through the following regeneration steps on specific phytagel-based media, containing the selective antibiotic hygromycin B (20 g/l) from Duchefa, http://www.duchefa.com (H0192) and the antibacterial antibiotic ticarcillin disodium/clavulanate potassium (125 mg/l) (Duchefa T0190): callus induction (14–21 days), shoot regeneration (14–21 days) and shoot maturation (14–21 days). Subsequently, plantlets were cultured for at least 21 days on rooting medium without both antibiotics. After root regeneration, plants were grown on soil, first in Magenta boxes (http://www.bio-world.com), and finally in 2-litre pots for flowering, self-pollination and seed production in the glasshouse.

Flow cytometry

Flow cytometry was performed with leaf material from *N. attenuata* on a flow cytometer CCA-II (Partec, http://www.partec.com) as described in Bubner *et al.* 2006.

Germination of N. attenuata *and screening for individuals with T-DNA insertions*

Germination of *N. attenuata* was performed as described in Kruegel *et al.* 2002 with the exception that 60 seeds were germinated per plate. If screening for individuals with T-DNA insertions should be performed, the selective antibiotic hygromycin B (Duchefa H0192) was added at a concentration of 35 mg/l to the germination medium. After 10 days, the ratio of seedlings surviving the antibiotic selection was determined.

Diagnostic PCRs for integrity of T-DNA insertions

Genomic DNA (gDNA) was isolated from leaves or seedlings of *N. attenuata* by a modified cetyltrimethylammonium bromide method (Bubner *et al.* 2004). PCR was performed with DreamTaq™ DNA Polymerase (Fermentas, http://www.fermentas.com) according to the instructions of the manufacturer with 1–100 ng of gDNA per sample. The following primer pairs were used: PROM FOR/INT REV and INT FOR/TER REV for the inverted repeat gene silencing constructs; PROM FOR/TER REV for the gene overexpression constructs; and DCL2GF1/DCL2GR1 or GGPP22-22/GGPP23-21 as positive controls amplifying a 334-bp or a 241-bp fragment of *DCL2* (GenBank GU479998) or *ggpps* (GenBank EF382626). Cycles were 5 min 95 °C (30s 95 °C, 30s 55–60 °C, 1 min 72 °C), repeated 30 times, 5 min 72 °C, hold 20 °C. Primer sequences are given in Table S1 (Supporting Information).

Southern blotting

Southern blotting was performed as described in Jassbi *et al.* (2008), with the exception that a 287-bp hygromycin phosphotransferase gene (*hptII*) probe obtained by PCR with primer pair HYG1-18/HYG2-18 (Table S1, Supporting Information) was used. Labelling was performed with the GE Healthcare (http://www.gehealthcare.com) Readyprime DNA labelling system and ProbeQuant g-50 microcolumns according to the instructions of the manufacturer; 7 μg of genomic DNA was digested with restriction enzymes from New England Biolabs (http://www.neb.com) and blotted onto a nylon membrane (GeneScreenPlus; PerkinElmer, http://www.perkinelmer.com) according to the manufacturer's protocol.

qPCR

Plant material was ground in liquid nitrogen with mortar and pestle. Total RNA was extracted with TRI reagent™ (SIGMA, http://www.sigmaaldrich.com) according to the manufacturer's instructions. RNA quality was checked on a 1% agarose gel, and concentration was measured spectrophotometrically at 260 nm.

For qPCR analysis, at least three replicated biological samples were used. One microgram of total RNA obtained from each sample was reverse transcribed using an oligo(dT) primer and Superscript II reverse transcriptase (Invitrogen) for a total volume of 20 μl according to the instructions of the manufacturer. cDNA samples were diluted 1:10 and used for SYBR®Green-based qPCR, carried out on a Stratagene MX3005P™ using qPCR™ Core Kits for SYBR®Green No ROX (Eurogentec, http://www.eurogentec.com) according to the instructions of

the manufacturer. Analysis of data was carried out according to the comparative Ct ($2^{-\Delta\Delta Ct}$) (Bubner & Baldwin 2004) method or by standard curves (Jassbi *et al.* 2008). The actin cDNA was amplified with primer pair Actin-F1/Actin-R1 (Table S1, Supporting Information) and used as an internal standard for normalizing cDNA concentration variations. For the determination of transcript abundances of the genes of interest, appropriate gene-specific primers were used.

Results and discussion

Here, we discuss the most efficient protocol to produce genetically modified plants utilizable in ecological research; the workflow that we describe is summarized in Fig. 2.

Construction of binary plant transformation vectors

Binary plant transformation vectors consist of two general regions—one representing the T-DNA, defined by the left and right border repeats, and one with the regions and genes necessary for replication in *Escherichia coli* and *Agrobacterium tumefaciens*. To avoid unwanted side effects and to achieve a high stability of the T-DNA inserts in the plant genome, the T-DNA should only contain the elements necessary for the intended function of the transgene. These elements are the two expression cassettes for

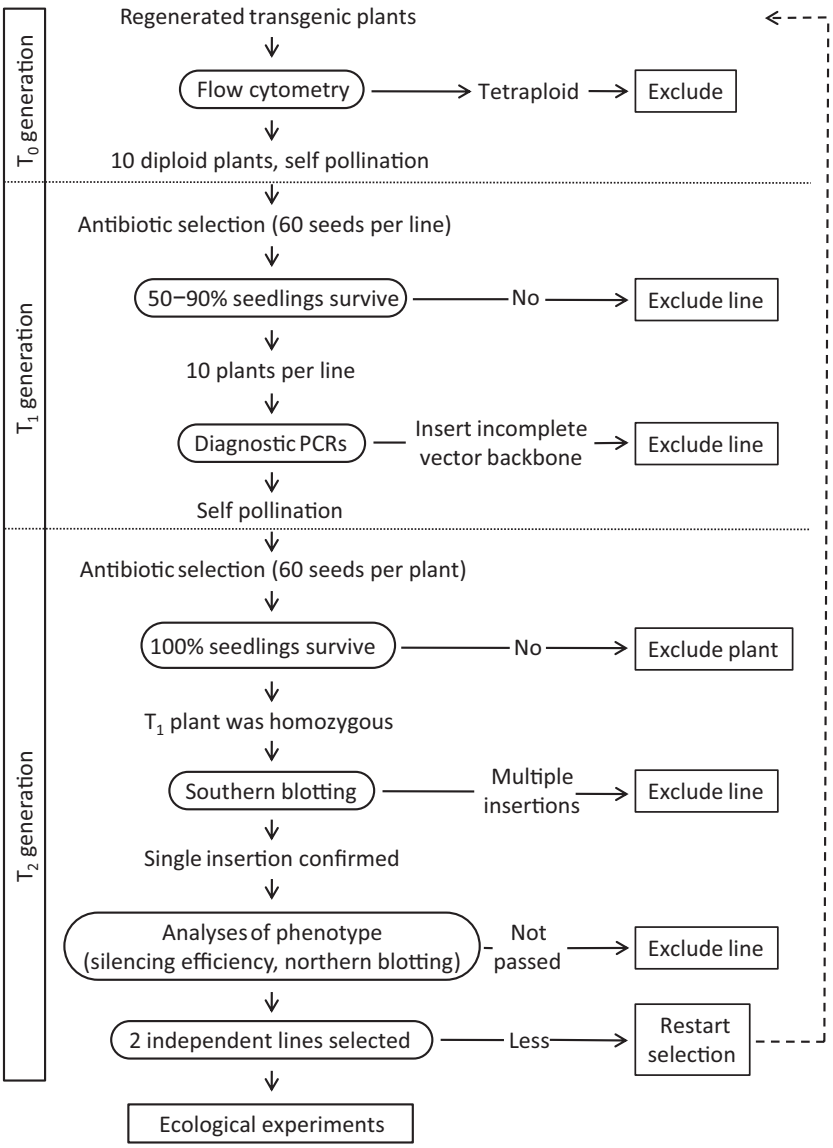

Fig. 2 Workflow for efficient screening of transgenic lines for ecological research. Each test is performed as early as possible during plant screening to obtain a fast and reliable selection.

the transgene and the plant selectable marker gene, both containing promoter, transgene sequence and terminator. In gene silencing vectors, the transgene region should consist of an inverted repeat of a partial sequence of the gene to be silenced, separated by a functional intron larger than 100 bp. The presence of this non-repeated sequence allows replication of the plasmid in bacteria despite the long inverted repeat (Warren & Green 1985). Splicing of the intron from the mRNA in the host plant greatly increases the probability that a dsRNA molecule is formed, which in turn efficiently initiates the silencing of the target gene. In our experience, target gene fragments with sizes ranging from 150 bp to about 1 kb can be used for efficient gene silencing. Routinely, our silencing constructs carry inverted repeat fragments of about 300 bp. If a member of a gene family or a certain allele should be silenced, the choice of the gene sequence requires special consideration. Nucleic acid homology of 23 nt is sufficient to direct post-transcriptional silencing of a gene (Thomas *et al.* 2001). To silence a single gene, sequence homology of more than 22 nt should be avoided. To silence a gene family, a sequence with homologies of more than 22 nt should be chosen. To enhance transgene stability, the use of different promoters and terminators, e.g. promoter/terminator of the nopaline synthase gene from the Ti plasmid of *A. tumefaciens* and 35S promoter/terminator from the cauliflower mosaic virus, on the same T-DNA is advisable.

The bacterial part of the binary transformation vector should contain an origin of replication functional in *E. coli*, e.g. from the ColE1 plasmid (construction of the binary vectors is performed in *E. coli*), an origin of replication and the genes that are necessary for plasmid replication in *A. tumefaciens*, e.g. from plasmid pVS1 and an antibiotic resistance marker both selectable in *E. coli* and *A. tumefaciens*. Because T-DNA border 'read-through' events are quite common during T-DNA integration, the use of a bacterial resistance gene that is already widely spread in nature, e.g. the *nptII* kanamycin resistance gene, is advisable.

Examples for binary plant transformation vectors constructed according to the principles described above are the pSOL8 gene silencing series (Fig. 1a) and the pSOL9 gene overexpression series (Fig. 1b), both extensively used to transform *N. attenuata*.

Transformation and regeneration

Agrobacterium-based transformation is the preferred method to create transgenic plants for ecological research. After transformation, the regeneration of differentiated plants from the transformed cells is necessary. Establishing transformation and regeneration procedures is probably the most challenging step in the utilization of transgenic approaches for non-model plants. Published transformation and regeneration procedures [as examples see *N. tabacum* (Horsch *et al.* 1985; Gallois & Marinho 1995), *Beta vulgaris* (Lindsey & Gallois 1990), *A. thaliana* (Valvekens *et al.* 1988; Clough & Bent 1998) and *Hordeum vulgare* (Tingay *et al.* 1997)] suggest that protocols specifically adapted for each species, and sometimes for each cultivar of each species (Valvekens *et al.* 1988) to be transformed, need to be worked out. With considerable effort, including as many as 10 people-years, we developed a transformation and regeneration procedure for *N. attenuata* (described in the 'Materials and methods' section). Because of different media and hormone requirements of this organism, established procedures from other closely related species like *N. tabacum* could not be applied.

Determination of ploidy level

One of the heritable somoclonal variations that may occur during tissue culture is autopolyploidy (Bubner *et al.* 2006). The extent of polyploidization can be substantial: In diploid tomato, 24.5–80% of transformants were found to be tetraploid [depending on cultivar and method; (Ellul *et al.* 2003)], and up to 92% of originally triploid bermuda grass *Cynodon dactylon* × *transvaalensis* cv. Tif Eagle transformants were found to be hexaploid (Goldman *et al.* 2004). To ensure the comparability and the relevance of the results obtained from experiments with transgenic plants and control plants, it is essential that only transgenic plants with the same ploidy level as the plant that has been transformed are selected. The ploidy level of the first transgenic generation (T_0) is preserved in all following generations produced by self-pollination. Ploidy-level determination of the lines of this generation should be the first step in selecting transgenic lines. This approach allows the early elimination of unwanted lines and saves resources.

The most efficient and conclusive method of determining ploidy levels is by flow cytometry. Because T_0 plants may be chimeric, material from leaves close to the flowers is determined to obtain results indicative of the ploidy level of the T_1 generation. Ideally, 10 independent transgenic lines with the correct ploidy level should be identified from this step in the screening process. According to our experience with *N. attenuata*, plants with doubled sets of chromosomes can often be visually distinguished from plants with the original set of chromosomes, as these plants commonly have 10% broader leaves and larger trichomes, stomates, flowers and seeds. They produce fewer seeds (20–40 vs. 100–200 seeds per capsule in diploid individuals) and have abnormal growth forms. Critical assessment of the phenotypes of the T_0 plants with respect to measured ploidy values has

in our experience greatly facilitated reliable elimination of tetraploid plants.

The ploidy analyses performed with transgenic *N. attenuata* lines from transformation experiments with more than 200 different constructs document that the frequency of occurrence of tetraploid plants substantially varies depending on the construct used for transformation. This effect occurred with all inbred lines (6th to the 30th inbred generation) used as starting material for the transformation. An example for the frequency of occurrence of tetraploid plants is shown in Fig. 3. Plants from the 30th inbred generation of a genotype collected from Utah (Glawe *et al.* 2003) were transformed with a series of binary vectors (pSOL9CAP, pSOL9ESC, pSOL9FAB, pSOL9ICE, pSOL9LEA, pSOL9PNA, pSOL9SSP and pSOL9VRD), all ectopically overexpressing heterologous genes coding for proteins with antimicrobial activities. The ploidy levels of plants from 11 to 25 independent T_0 lines per construct were measured using flow cytometry. Depending on the transgene, 0–57% of the transgenic plants were tetraploid. All together, 159 plants were analysed, 37% were tetraploid and 63% diploid. The increase in the portion of tetraploid plant lines occurring after transformation with particular constructs may require that a larger number of transgenic lines need to be tested to obtain the desired number of diploid lines.

Screening for homozygous lines with single-copy T-DNA

Transgenic plant lines used for ecological research should carry a single T-DNA copy in a single transgenic locus. We have found that a single copy is sufficient for the expected phenotypic effect (gene silencing or overexpression). To avoid segregation, these lines should be homozygous with respect to the T-DNA insertion. Multiple independent transgenic loci would dramatically increase the effort required to generate and identify non-segregating, homozygous lines. In the case of overexpression, multiple T-DNA copies or even strong promoters of a single T-DNA copy may lead to transgene mRNA concentrations above a critical level and elicitation of silencing of the overexpressed gene (Lechtenberg *et al.* 2003; Eamens *et al.* 2008; Hirai *et al.* 2010).

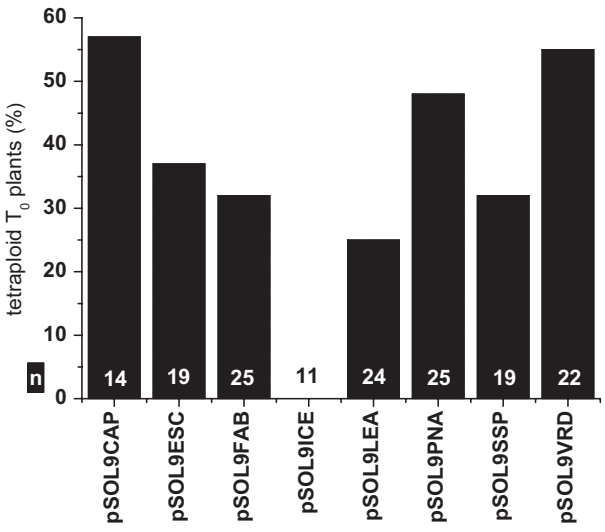

Fig. 3 The fraction of tetraploid *Nicotiana attenuata* plants after transformation of the 30th Utah inbred line with pSOL9 vectors. The numbers of tested plants are indicated.

Transgenic plants with two or more independent transgenic loci can generally be identified by inheritance studies (Vain & Thole 2009). Non-independent transgenic loci (T-DNA insertions at a distance smaller than 50 cM on the same chromosome) and complex insertions into one locus need to be identified by Southern analysis (protocol section 7) at later stages in the screening process. Any T-DNA insertion into T_0 plants will usually be hemizygous, because the probability of simultaneous insertions into the same locus of two homologous chromosomes is very low.

To produce homozygous lines with respect to a transgenic locus, self-pollination of T_0 plants is required. Screening for T-DNA insertions can be performed most efficiently by growing seedlings on medium containing the selective antibiotic. Only plants carrying the T-DNA and expressing the resistance marker gene will grow. The expected Mendelian inheritance ratio for a single transgenic locus in the T_1 offspring should be 1 (homozygous) to 2 (hemizygous) to 1 (wild type); thus, 75% of the offspring would carry the T-DNA. A second independent transgenic locus would lead to a ratio of 15 (any transgenic locus) to 1 (wild type), and 93.75% of the offspring would carry the T-DNA. Additional transgenic loci would further increase this portion.

If the ratio of transgenic vs. non-transgenic seedlings is lower than expected (<50%), silencing of the selectable marker gene may be the reason. In this case, and only if a simultaneous silencing of the transgene can be excluded, screening for the selectable marker gene could be performed by PCR genotyping of seedlings grown on non-selective medium. To assess the extent to which silencing of the selectable marker gene occurs during the screening process, the T_1, T_2 and T_3 inbred generations of the *N. attenuata* pSOL9 overexpression lines were analysed (Fig. 4). The initial number of T_1 lines was 10–12 depending on the construct. Because according to our observations (data not shown) silencing of the selectable marker gene is associated with silencing of the transgene and will, once initiated, reoccur in the following generations, all plant lines exhibiting signs of silencing were excluded from further screening. Silencing of the resistance gene

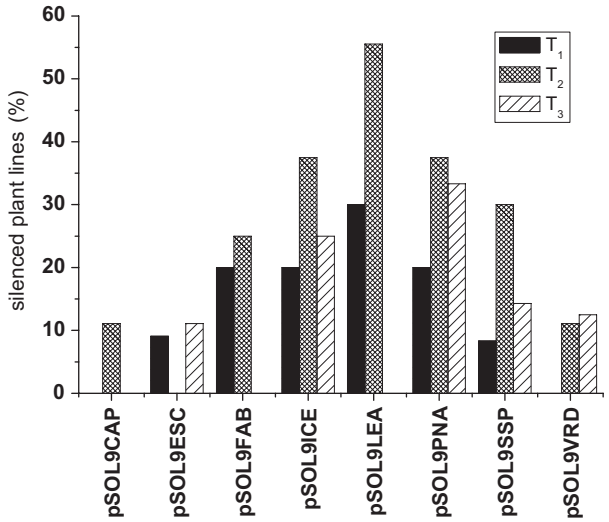

Fig. 4 Portion of lines in which silencing of the *hpt*II gene occurred during different stages of screening (T_1, T_2 or T_3 generation) as detected by hygromycin sensitivity. The initial number of T_1 lines was 10–12 depending on the construct. Plants in T_1 and T_2 stage were considered as silenced if more than 50% of seedlings were hygromycin sensitive. Homozygous T_3 plants were considered as silenced if segregation occured.

occurred unpredictably and irregularly. The mean portion of silenced lines was 13.4% for T_1, 26.0% for T_2 and 12.0% for T_3 generations. Gene silencing in sense overexpression lines is most likely caused by promoter methylation (Weber *et al.* 1990; Stam *et al.* 1997). This form of epigenetic regulation can occur in each offspring generation of a previously unsilenced line. To guarantee a high expression level of the transgene, the functionality of the resistance marker gene should be examined for each new generation by germination on medium containing the selective antibiotic.

T-DNA insertions may interrupt genes essential for embryo or gametophyte development, thus leading to embryo- (Errampalli *et al.* 1991) or gametophyte-lethal lines (Feldmann *et al.* 1997; Howden *et al.* 1998). This will result in exceptional segregation, characterized by progeny segregating for fewer T-DNA carrying seedlings than predicted by Mendelian principles. Exceptional segregation may occur in about 9% of the transgenic lines (Feldmann *et al.* 1997) and should be accounted for in the segregation analyses.

The inheritance-based screening of transgenic plant lines is always performed via self-pollination and germination of 60 seedlings per plant. In the first step, the T_1 generation seeds from 10 independent T_0 lines with the correct ploidy level are germinated. A T_0 plant is considered an appropriate candidate line for carrying a single T-DNA insertion if 50–90% (75% calculated) of the T_1 offspring are transgenic. More transgenic seedlings point to multiple T-DNA integrations, while fewer indicate gene silencing or exceptional segregation.

In the second screening step, T_2 seeds from 10 T_1 plants of each single T-DNA insertion candidate line are germinated. If all T_2 seedlings from a T_1 plant carry the T-DNA, this plant represents the desired homozygous genotype of the respective line and can be used for further screening (Fig. 2). Plants with extreme, unexpected phenotypes should not be used for seed production.

At this point, it should be mentioned that because of extremely long generation times or self-incompatibility in some ecologically interesting systems (e.g. trees), self-pollination cannot be applied in the screening process. Each system may present its specific challenges, which have to be overcome in a specific way, for instance, by deeper genetic analyses of the T_0 generation.

Confirmation of complete T-DNA integrations using PCR

Strand breaks followed by deletions, inversions and rearrangements at the T-DNA flanking regions are some of the events during *Agrobacterium*-mediated T-DNA integration that may result in non-functional transgenes (Latham *et al.* 2006; Muller *et al.* 2007; Gambino *et al.* 2009). Selection for non-functional integrations of the transgene occurs when a plant is transformed with a transgene the product of which interferes with plant regeneration. This process leads to a dramatically reduced transformation efficiency. When this occurs, the regenerated lines often carry large deletions of the T-DNA flank containing the transgene, in our experience with *N. attenuata*. In contrast, lines that carry no functional resistance marker are eliminated by selection for antibiotic resistance during tissue culture.

It is essential to demonstrate that each candidate line that should be used in further screening carries a complete, functional T-DNA. Antibiotic selection guarantees the intactness of the T-DNA flank harbouring the selectable marker gene, but the integrity of the T-DNA transgene flank still has to be demonstrated. This is efficiently carried out with a diagnostic PCR-based analysis of the T_1 genome from which selected fragments of the transgene are amplified. The presence of a PCR product at the expected size provides strong evidence of a complete T-DNA insertion. Lines that do not yield the expected PCR product carry an incorrect transgene, most likely a deletion of the region to be amplified, and should be excluded from further screening. The results of this analysis are only informative for plants with single T-DNA insertions, because in plants with multiple insertions, a positive result does not exclude the presence of an additional defective transgene. If appropriate, this test can already be performed on T_0 plants to exclude lines

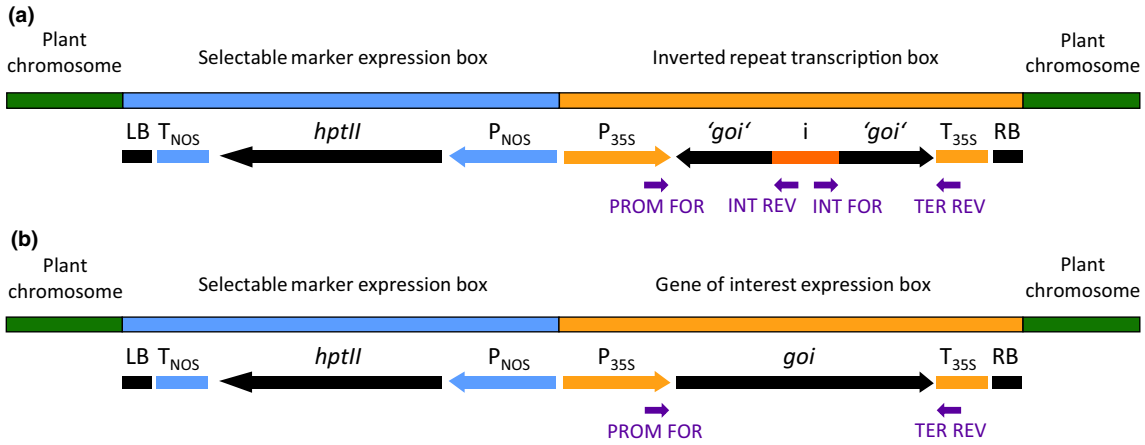

Fig. 5 The positions of annealing sites of the diagnostic PCR primers at the T-DNA integrated into the plant genome: (a) inverted repeat silencing vectors, (b) overexpression vectors. Abbreviations: see Fig. 1.

without functional T-DNA insertions or on plants from later generations to confirm the presence of a correctly inserted transgene.

To test whether the transgene expression cassettes of different *N. attenuata* lines transformed with pSOL8 or pSOL9 constructs were integrated completely, we designed four universal primers that bind to the cassettes' functional elements (Fig. 5): PROM FOR (35S promoter), INT REV (*pdk*-intron 5′-end) and INT FOR (*pdk*-intron 3′-end) TER REV (35S terminator). The *N. attenuata* silencing lines were analysed with primer pairs PROM FOR/INT REV (amplifying the transgene 5′ copy) and INT FOR/TER REV (amplifying the transgene 3′ copy adjacent to the right T-DNA border). These allow for the amplification of separate fragments for each transgene copy, thus avoiding low PCR efficiency because of amplification of an inverted repeat. For the analyses of the transgenic *N. attenuata* lines that were transformed with pSOL9 overexpression constructs, we used primer pair PROM FOR/TER REV, amplifying the DNA between promoter and terminator of the transgene expression box.

Examples for diagnostic PCR are shown in Fig. 6. The results of all diagnostic PCR performed with 24 transgenic *N. attenuata* lines transformed with the inverted repeat silencing constructs pSOL8DC3, pSOL8PNRP, pSOL8AEP65 and pSOL8AEP150 are summarized in Table S2 (Supporting Information). The rate of incomplete insertions varied between 0 and 60% depending on the transgene construct. Any transgenic line yielding a negative PCR result for the amplification of the 5′ transgene copy always yielded the same result for the 3′ transgene copy, whereas lines with the combination negative PCR result for the 3′ transgene copy and positive result for the 5′ transgene copy were found. These results demonstrate that deletions at the left border T-DNA flank

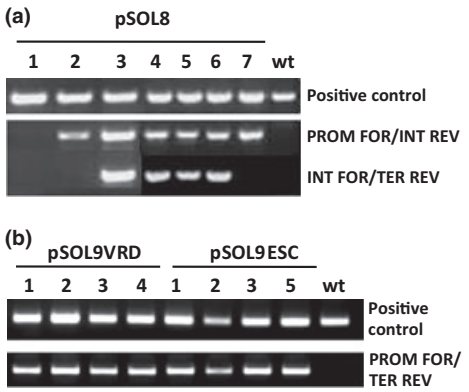

Fig. 6 Diagnostic PCRs with chromosomal DNA from transgenic *Nicotiana attenuata* plants transformed with (a) pSOL8 inverted repeat vectors (1–3: pSOL8DC3; 4: pSOL8PNRP; 5: pSOL8AEP65; 6: pSOL8AEP150: 7 pSOL8PNRP; wt: wild type) and (b) pSOL9 overexpression vectors (lines indicated). Primer pairs are shown.

are common in *N. attenuata*. The highest rates of incomplete insertions were found in plants transformed with pSOL8PNRP and pSOL8DC3 for the silencing of the PNRP gene and for the combined silencing of the RdR1 and WRKY3 genes. We interpret this result as being consistent with selection for incomplete insertions when the product of a transgene is detrimental to the plant.

Confirming the lack of vector backbone

The absence of binary vector backbone sequences in the genome is an important quality criterion established by many regulatory agencies for plants that are to be used in field releases. Demonstrating that a transgenic line does not contain vector sequences outside the T-DNA can be done by a PCR-based approach or—parallel to the

determination of T-DNA copy number described in the following chapter—by Southern analysis. Because vector backbone integration often occurs after a T-DNA border 'read-through', the lack of PCR products from primer pairs designed for the amplification of the vector backbone adjacent to the T-DNA borders demonstrates the absence of vector backbone. For Southern analysis, these primer pairs are used to amplify probes from the transformation vector. Genomic DNA of plants, which does not allow the detection of a DNA fragment with these probes, is considered to contain no vector backbone. If the regulatory agencies require that the absence of critical vector regions, such as the bacterial antibiotic resistance gene or plasmid origins of replication, is to be explicitly demonstrated, primers for these regions can be designed and used in the PCR analyses accordingly. All transgenic lines that contain unwanted vector sequences should be eliminated.

A recent study (Oltmanns *et al.* 2010) demonstrated that starting plant transformation from the *A. tumefaciens* chromosome could be a way to reduce the portion of transgenic plant lines with vector backbone integrations dramatically, but so far, this procedure to produce *Agrobacterium* plant transformation strains is laborious and not applicable to the widely used strain *A. tumefaciens* LBA4404.

Determination of T-DNA copy number by Southern analysis

Multiple transgenic loci and complex T-DNA insertions into one locus present in the genome of transgenic plant lines can, in most cases, be reliably detected by Southern analysis. To confirm single T-DNA insertions in lines selected so far, the probe should be identical to a part of the antibiotic resistance marker gene. This gene does not exist in the wild-type plant genome, but should be present in the genome of all transgenic plants as a consequence of antibiotic selection during regeneration. Moreover, this allows the same probe for the screening of different transgenic plant lines generated with T-DNA from different binary vectors to be used, as long as the selectable marker gene is the same.

For Southern analysis, chromosomal DNA from two or more homozygous T_2 generation individuals of each transgenic line that passed all previous screening tests is completely digested in separate reactions with at least two different restriction enzymes. These enzymes must not cut the T-DNA on both sides or inside the probe sequence, but in order to detect multiple insertions into a single transgenic locus and to reduce the expected size of the genomic fragments carrying this sequence, the T-DNA should be cut once. Under these conditions, all transgenic lines that yield in the Southern analysis multiple bands with any of the restriction enzymes harbour multiple transgenic loci or multiple T-DNA insertions at one locus. Incomplete T-DNA insertions or unwanted sequence rearrangements of the transgenic locus are indicated when fragments smaller than the minimal possible T-DNA size calculated from T-DNA borders and the restriction sites on it are detected. Single bands equal to or larger than the calculated minimal size obtained with all used restriction enzymes are indicative of a single T-DNA insertion. Transgenic lines yielding this pattern are chosen for further screening. An example for a Southern blot of each four independent lines from three different constructs is shown in Fig. 7. However, the existence of additional T-DNA fragments in the genome of the chosen lines, not detectable with the selected probe, cannot be definitively excluded.

Confirmation of transgene function

Before a transgenic line that has been demonstrated to harbour a correct single T-DNA insertion in homozygous stage can be used in ecological experiments, the function of the transgene should be confirmed on the level of RNA. For overexpression lines, the mRNA abundance of the transgene is quantified by qPCR. Choosing an amplicon comprising the stop codon will provide an additional control that the full-length gene is expressed. For gene silencing lines, the silencing efficiency is determined. The

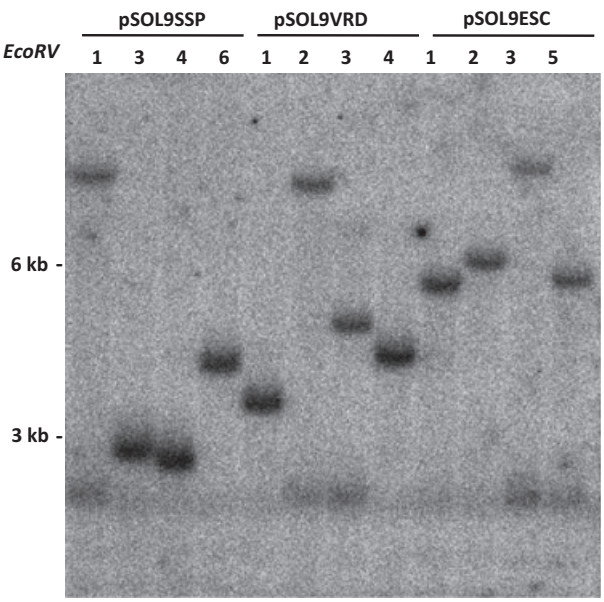

Fig. 7 Southern blotting with nucleic acid from each four independent transgenic *Nicotiana attenuata* T_2 lines transformed with vectors pSOL9SSP, pSOL9VRD or pSOL9ESC. The DNA for the Southern blot was digested with *Eco*RV. A fragment of the marker gene (*hpt*II) served as probe.

relative transcript abundance of the gene of interest both for the transgenic line and wild type plants is measured by qPCR using an amplicon originating from the target gene outside the inverted repeat. Silencing efficiency is then calculated as 100% minus the ratio (in percentage) of relative transcript abundance in the transgenic line and in wild type plants grown under the same conditions and treated in the same way. In our experience with *N. attenuata*, only lines that allow efficient silencing of the target gene (>20%) are valuable for ecological experiments.

qPCR is a very robust means of determining how efficiently an endogenous plant gene is silenced. Nevertheless, the choice of the intrinsic plant gene that shows constant transcript abundance in all the experiments performed and serves as standard gene for normalization is essential for the reliability of the results obtained with this method (Czechowski *et al.* 2005; Gutierrez *et al.* 2008).

If appropriate, transgene function should be confirmed on a phenotypic level by studying the phenotypes associated with silencing or overexpression of the targeted gene. Depending on the insertion site, the strength of expression of the transgene can vary considerably because of 'position effects' (Prols & Meyer 1992; Matzke & Matzke 1998; *Qin et al.* 2003). Different lines with the same T-DNA insertion may thus have modulated phenotypes with different strengths. This titration of phenotypes can be a powerful means of examining the fitness consequences of a gene. In certain cases, the functional transgene will lead to morphological changes, which may allow a pre-selection during the screening process.

Conclusions

Reverse genetics is a powerful tool in plant ecology. To take advantage of this tool, transgenic plants that fulfil the requirements for the structure and stable inheritance of the transgene need to be created and selected. The screening of transgenic plants is a costly and time-consuming procedure. We developed a flow chart protocol (Fig. 2) that allows for the efficient production and selection of transgenic plants for ecological research. We encourage groups working in the field of ecology to make use of the resources described in this study, and the authors will be happy to share plasmids, plant lines and experience with all interested groups.

Acknowledgements

We thank Susan Kutschbach, Wibke Kröber and Antje Wissgott for help with vector construction and plant transformation, Thomas Hahn for sequencing and the Max Planck Society for funding.

References

Baldwin IT (2001) An ecologically motivated analysis of plant-herbivore interactions in native tobacco. *Plant Physiology*, **127**, 1449–1458.

Bhat SR, Srinivasan S (2002) Molecular and genetic analyses of transgenic plants: considerations and approaches. *Plant Science*, **163**, 673–681.

Bubner B, Baldwin IT (2004) Use of real-time PCR for determining copy number and zygosity in transgenic plants. *Plant Cell Reports*, **23**, 263–271.

Bubner B, Gase K, Baldwin IT (2004) Two-fold differences are the detection limit for determining transgene copy numbers in plants by real-time PCR. *BMC Biotechnology*, **4**, 14.

Bubner B, Gase K, Berger B, Link D, Baldwin IT (2006) Occurrence of tetraploidy in *Nicotiana attenuata* plants after *Agrobacterium*-mediated transformation is genotype specific but independent of polysomaty of explant tissue. *Plant Cell Reports*, **25**, 668–675.

Christou P (1992) Genetic-transformation of crop plants using microprojectile bombardment. *Plant Journal*, **2**, 275–281.

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of Arabidopsis thaliana. *Plant Journal*, **16**, 735–743.

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiology*, **139**, 5–17.

De Buck S, Windels P, De Loose M, Depicker A (2004) Single-copy T-DNAs integrated at different positions in the *Arabidopsis* genome display uniform and comparable beta-glucuronidase accumulation levels. *Cellular and Molecular Life Sciences*, **61**, 2632–2645.

De Buck S, Podevin N, Nolf J, Jacobs A, Depicker A (2009) The T-DNA integration pattern in Arabidopsis transformants is highly determined by the transformed target cell. *Plant Journal*, **60**, 134–145.

Eamens A, Wang MB, Smith NA, Waterhouse PM (2008) RNA silencing in plants: yesterday, today, and tomorrow. *Plant Physiology*, **147**, 456–468.

Ellul P, Garcia-Sogo B, Pineda B *et al.* (2003) The ploidy level of transgenic plants in *Agrobacterium*-mediated transformation of tomato cotyledons (*Lycopersicon esculentum* L.Mill.) is genotype and procedure dependent. *Theoretical and Applied Genetics*, **106**, 231–238.

Errampalli D, Patton D, Castle L *et al.* (1991) Embryonic lethals and T-DNA insertional mutagenesis in *Arabidopsis*. *Plant Cell*, **3**, 149–157.

Feldmann KA, Coury DA, Christianson ML (1997) Exceptional segregation of a selectable marker (Kan(R)) in Arabidopsis identifies genes important for gametophytic growth and development. *Genetics*, **147**, 1411–1422.

Gallois P, Marinho P (1995) Leaf disk transformation using *Agrobacterium tumefaciens*: expression of heterologous genes in tobacco. *Methods in Molecular Biology*, **49**, 39–48.

Gambino G, Chitarra W, Maghuly F *et al.* (2009) Characterization of T-DNA insertions in transgenic grapevines obtained by *Agrobacterium*-mediated transformation. *Molecular Breeding*, **24**, 305–320.

Gelvin SB (2009) *Agrobacterium* in the Genomics Age. *Plant Physiology*, **150**, 1665–1676.

Gelvin SB, Kim SI (2007) Effect of chromatin upon Agrobacterium T-DNA integration and transgene expression. *Biochimica Et Biophysica Acta-Gene Structure and Expression*, **1769**, 410–421.

Gheysen G, Villarroel R, Vanmontagu M (1991) Illegitimate recombination in plants—a model for T-DNA integration. *Genes & Development*, **5**, 287–297.

Glawe GA, Zavala JA, Kessler A, Van Dam NM, Baldwin IT (2003) Ecological costs and benefits correlated with trypsin protease inhibitor production in *Nicotiana attenuata*. *Ecology*, **84**, 79–90.

Goldman JJ, Hanna WW, Fleming GH, Ozias-Akins P (2004) Ploidy variation among herbicide-resistant bermudagrass plants of cv. TifEagle transformed with the bar gene. *Plant Cell Reports*, **22**, 553–560.

Gutierrez L, Mauriat M, Pelloux J, Bellini C, van Wuytswinkel O (2008) Towards a systematic validation of references in real-time RT-PCR. *Plant Cell*, **20**, 1734–1735.

Hirai S, Takahashi K, Abiko T, Kodama H (2010) Loss of sense transgene-induced post-transcriptional gene silencing by sequential introduction of the same transgene sequences in tobacco. *FEBS Journal*, **277**, 1695–1703.

Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA (1983) A binary plant vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature*, **303**, 179–180.

Horsch RB, Fry JE, Hoffmann NL *et al.* (1985) A simple and general-method for transferring genes into plants. *Science*, **227**, 1229–1231.

Howden R, Park SK, Moore JM *et al.* (1998) Selection of T-DNA-tagged male and female gametophytic mutants by segregation distortion in *Arabidopsis*. *Genetics*, **149**, 621–631.

Jassbi AR, Gase K, Hettenhausen C, Schmidt A, Baldwin IT (2008) Silencing geranylgeranyl diphosphate synthase in *Nicotiana attenuata* dramatically impairs resistance to tobacco hornworm. *Plant Physiology*, **146**, 974–986.

Kang JH, Wang L, Giri A, Baldwin IT (2006) Silencing threonine deaminase and JAR4 in *Nicotiana attenuata* impairs jasmonic acid-isoleucine-mediated defenses against Manduca sexta. *Plant Cell*, **18**, 3303–3320.

Kim SI, Veena, Gelvin SB (2007) Genome-wide analysis of Agrobacterium T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *Plant Journal*, **51**, 779–791.

Kohli A, Twyman RM, Abranches R *et al.* (2003) Transgene integration, organization and interaction in plants. *Plant Molecular Biology*, **52**, 247–258.

Kononov ME, Bassuner B, Gelvin SB (1997) Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome: evidence for multiple complex patterns of integration. *Plant Journal*, **11**, 945–957.

Kruegel T, Lim M, Gase K, Halitschke R, Baldwin IT (2002) Agrobacterium-mediated transformation of *Nicotiana attenuata*, a model ecological expression system. *Chemoecology*, **12**, 177–183.

Lacroix B, Citovsky V (2009) Agrobacterium aiming for the host chromatin. *Communicative & Integrative Biology*, **2**, 42–45.

Lange M, Vincze E, Moller MG, Holm PB (2006) Molecular analysis of transgene and vector backbone integration into the barley genome following *Agrobacterium*-mediated transformation. *Plant Cell Reports*, **25**, 815–820.

Latham JR, Wilson AK, Steinbrecher RA (2006) The mutational consequences of plant transformation. *Journal of Biomedicine and Biotechnology*, **2006**, 25376.1–25376.7.

Lechtenberg B, Schubert D, Forsbach A, Gils M, Schmidt R (2003) Neither inverted repeat T-DNA configurations nor arrangements of tandemly repeated transgenes are sufficient to trigger transgene silencing. *Plant Journal*, **34**, 507–517.

Lindsey K, Gallois P (1990) Transformation of sugar-beet (*Beta vulgaris*) by *Agrobacterium tumefaciens*. *Journal of Experimental Botany*, **41**, 529–536.

Matzke AJM, Matzke MA (1998) Position effects and epigenetic silencing of plant transgenes. *Current Opinion in Plant Biology*, **1**, 142–148.

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

Meza TJ, Stangeland B, Mercy IS *et al.* (2002) Analyses of single-copy Arabidopsis T-DNA-transformed lines show that the presence of vector backbone sequences, short inverted repeats and DNA methylation is not sufficient or necessary for the induction of transgene silencing. *Nucleic Acids Research*, **30**, 4556–4566.

Muller AE, Atkinson RG, Sandoval RB, Jorgensen RA (2007) Microhomologies between T-DNA ends and target sites often occur in inverted orientation and may be responsible for the high frequency of T-DNA-associated inversions. *Plant Cell Reports*, **26**, 617–630.

Olhoft PM, Flagel LE, Somers DA (2004) T-DNA locus structure in a large population of soybean plants transformed using the Agrobacterium-mediated cotyledonary-node method. *Plant Biotechnology Journal*, **2**, 289–300.

Oltmanns H, Frame B, Lee LY *et al.* (2010) Generation of backbone-free, low transgene copy plants by launching T-DNA from the *Agrobacterium* chromosome. *Plant Physiology*, **152**, 1158–1166.

O'Malley RC, Ecker JR (2010) Linking genotype to phenotype using the Arabidopsis unimutant collection. *Plant Journal*, **61**, 928–940.

Prols F, Meyer P (1992) The methylation patterns of chromosomal integration regions influence gene activity of transferred DNA in *Petunia hybrida*. *Plant Journal*, **2**, 465–475.

Qin HX, Dong YZ, von Arnim AG (2003) Epigenetic interactions between Arabidopsis transgenes: characterization in light of transgene integration sites. *Plant Molecular Biology*, **52**, 217–231.

Sallaud C, Meynard D, van Boxtel J *et al.* (2003) Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. *Theoretical and Applied Genetics*, **106**, 1396–1408.

Schwachtje J, Baldwin IT (2008) Why does herbivore attack reconfigure primary metabolism? *Plant Physiology*, **146**, 845–851.

Schwachtje J, Kutschbach S, Baldwin IT (2008) Reverse genetics in ecological research. *PLoS ONE*, **3**, e1543.

Stam M, Mol JNM, Kooter JM (1997) The silence of genes in transgenic plants. *Annals of Botany*, **79**, 3–12.

Steppuhn A, Gase K, Krock B, Halitschke R, Baldwin IT (2004) Nicotine's defensive function in nature. *Plos Biology*, **2**, 1074–1080.

Thomas CL, Jones L, Baulcombe DC, Maule AJ (2001) Size constraints for targeting post-transcriptional gene silencing and for RNA-directed methylation in *Nicotiana benthamiana* using a potato virus X vector. *Plant Journal*, **25**, 417–425.

Tingay S, McElroy D, Kalla R *et al.* (1997) *Agrobacterium tumefaciens*-mediated barley transformation. *Plant Journal*, **11**, 1369–1376.

Tzfira T, Citovsky V (2006) Agrobacterium-mediated genetic transformation of plants: biology and biotechnology. *Current Opinion in Biotechnology*, **17**, 147–154.

Tzfira T, Li JX, Lacroix B, Citovsky V (2004) *Agrobacterium* T-DNA integration: molecules and models. *Trends in Genetics*, **20**, 375–383.

Vain P, Thole V (2009) Gene Insertion Patterns and Sites. In: *Methods in Molecular Biology, Transgenic Wheat, Barley and Oats* (eds Jones HD & Shewry PR), vol. 478, pp. 203–226. © Humana Press, a part of Springer Science + Business Media, LLC, New York.

Valvekens D, van Montagu M, Van Lijsebettens M (1988) *Agrobacterium tumefaciens*-mediated transformation of Arabidopsis thaliana root explants by using kanamycin selection. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 5536–5540.

Warren GJ, Green RL (1985) Comparison of physical and genetic properties of palindromic DNA-sequences. *Journal of Bacteriology*, **161**, 1103–1111.

Weber H, Ziechmann C, Graessmann A (1990) *In vitro* DNA methylation inhibits gene-expression in transgenic tobacco. *EMBO Journal*, **9**, 4409–4415.

Wenck A, Czako M, Kanevski I, Marton L (1997) Frequent collinear long transfer of DNA inclusive of the whole binary vector during Agrobacterium-mediated transformation. *Plant Molecular Biology*, **34**, 913–922.

Yu JG, Lee GH, Kim JS, Shim EJ, Park YD (2010) An insertional mutagenesis system for analyzing the Chinese cabbage genome using *Agrobacterium* T-DNA. *Molecules and Cells*, **29**, 267–275.

Zavala JA, Patankar AG, Gase K, Baldwin IT (2004) Constitutive and inducible trypsin proteinase inhibitor production incurs large fitness costs in *Nicotiana attenuata*. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 1607–1612.

Data accessibility

Tables S1 and S2 (Supporting Information): Dryad repository http://dx.doi.org/10.5061/dryad.8951.

DNA sequences: GenBank accessions GU479998; HQ698849-HQ698853.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 Sequences of the primers used for diagnostic PCRs, Southern blotting, and qPCR.

Table S2 Results of diagnostic PCRs with 24 different transgenic *N. attenuata* lines.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.