A new approach for the identification and cloning of genes: the pBACwich system using Cre/*lox* site-specific recombination

Sangdun Choi*, Dilara Begum[1], Heather Koshinsky[2], David W. Ow[2] and Rod A. Wing[1]

Division of Biology 147-75, California Institute of Technology, Pasadena, CA 91125, USA, [1]Departments of Agronomy and Biological Sciences, Clemson University, Clemson, SC 29634, USA and [2]Plant Gene Expression Center, USDA/ARS-UC Berkeley, 800 Buchanan Street, Albany, CA 94710, USA

Received February 17, 1999; Revised October 8, 1999; Accepted November 10, 1999

ABSTRACT

With current plant transformation methods (*Agrobacterium*, biolistics and protoplast fusion), insertion of DNA into the genome occurs randomly and in many instances at multiple sites. Associated position effects, copy number differences and multigene interactions can make gene expression experiments difficult to interpret and plant phenotypes less predictable. An alternative approach to random integration of large DNA fragments into plants is to utilize one of several site-specific recombination (SSR) systems, such as Cre/*lox*. Cre has been shown in numerous instances to mediate *lox* site-specific recombination in animal and plant cells. By incorporating the Cre/*lox* SSR system into a bacterial artificial chromosome (BAC) vector, a more precise evaluation of large DNA inserts for genetic complementation should be possible. Site-specific insertion of DNA into predefined sites in the genome may eliminate unwanted 'position effects' caused by the random integration of exogenously introduced DNA. In an effort to make the Cre/*lox* system an effective tool for site-directed integration of large DNAs, we constructed and tested a new vector potentially capable of integrating large DNA inserts into plant and fungal genomes. In this study, we present the construction of a new BAC vector, pBACwich, for the system and the use of this vector to demonstrate SSR of large DNA inserts (up to 230 kb) into plant and fungal genomes.

INTRODUCTION

The isolation of agriculturally important genes is an important goal in plant molecular biology. Since most agriculturally important genes are known only by phenotype, techniques have been developed to isolate such genes. Currently, map-based cloning (or positional cloning), insertional mutagenesis and subtraction cloning are three of the best-developed strategies (1). Since the late 1980s, map-based cloning has been successfully used to isolate human genetic disease genes (2–6) and plant genes (7–9). Map-based gene cloning is based on prior knowledge of the location of a target gene on a genetic map. For successful map-based cloning of a gene, one must complete the following steps: (i) isolate molecular markers genetically linked to the target gene; (ii) generate a physical map of the region surrounding the target gene with the most closely linked molecular markers; (iii) isolate a set of overlapping DNA fragments that encompass the target gene, i.e. chromosome walking/landing; (iv) identify the target gene using a variety of techniques (e.g. sequencing and genetic complementation).

Once a contig of overlapping clones has been assembled, it then becomes necessary to determine which of the overlapping clones contains the target gene. This step, 'the gene identification step', is daunting and tedious. It normally requires subcloning of a large DNA fragment containing a target gene and introduction of each small fragment into a plant. By developing a vector that can accommodate large DNA fragments and is suitable for plant transformation, we can significantly streamline the process of gene identification. The ability to introduce relatively large fragments of DNA into plant cells also increases the chance that an entire gene is contained within the introduced DNA.

With current plant transformation methods (*Agrobacterium*, biolistics and protoplast fusion), insertion of DNA into the genome occurs randomly and in many instances at multiple sites. Associated position effects, copy number differences and multigene interactions can make gene expression experiments difficult to interpret and plant phenotypes less predictable. *Agrobacterium*-mediated transformation is the most commonly used method and relies on the efficient transfer of DNA from bacteria to plants. Hamilton *et al.* (10) constructed a binary plant transformation vector that is capable of transferring large DNA fragments from *Agrobacterium* into plants. However, the primary drawback to *Agrobacterium* transformation is that integration is random throughout the genome and multiple transformants have to be generated to compensate for variation in position effects. In addition, the stability of large plant DNA sequences in *Agrobacterium* has not been fully tested and may pose other potential problems.

An alternative approach to random integration of large DNA fragments into plants is to utilize one of several site-specific recombination systems such as Cre/*lox*. Site-specific

*To whom correspondence should be addressed. Tel: +1 626 395 4504; Fax: +1 626 796 7066; Email: schoi@its.caltech.edu

recombination (SSR) is the enzyme-mediated cleavage and ligation of two defined deoxynucleotide sequences. A number of conservative site-specific recombination systems have been described in both prokaryotic and eukaryotic organisms (11). These systems generally use one or more proteins and act on unique asymmetric DNA sequences (12). Cre, the product of the *cre* gene, is a 38.5 kDa recombinase that can reciprocally exchange the DNA at 34 bp *lox* sites (13). Cre has been shown in numerous instances to mediate *lox* site-specific recombination in animal and plant cells (14–18). The products of the recombination event depend on the relative orientation of these asymmetric sequences (12). If well-characterized recombination systems function in higher eukaryotic cells, it should be possible to use the reciprocal exchange of strategically placed recombination sites for a variety of applications. By incorporating the Cre/*lox* SSR system into a BAC vector, a more precise evaluation of large DNA inserts for genetic complementation should be possible. Site-specific insertion of DNA into predefined sites in the genome may eliminate unwanted 'position effects' caused by the random integration of exogenously introduced DNA.

The bacteriophage P1 Cre/*lox* site-specific recombination system has been used to integrate DNA specifically at *lox* sites previously placed in the tobacco genome (14). When the Cre recombinase is supplied by transient *cre* expression, the *lox* site yields integration events. Integration of the insertion plasmid results in promoter displacement and the presumed termination of *cre* transcription.

In an effort to make the Cre/*lox* system an effective tool for site-directed integration of large DNAs, we constructed and tested a new vector potentially capable of integrating large DNA inserts into plant and fungal genomes. In this study, we present the construction of a new BAC vector for the system and the use of this vector to demonstrate SSR of large DNA inserts into plant and fungal genomes.

MATERIALS AND METHODS

Vector constructs

The pBeloBAC11 vector (19) was first modified to have a *35S* promoter plus the *NPTII* (neomycin phosphotransferase coding region) gene (pBACKAN; Dr S. Reddy, Texas A&M University, unpublished). pBACKAN was digested with *Not*I and purified with a Qiaex II Gel Extraction Kit (Qiagen, USA). Plasmid p24 (containing a *lox* site, a promoterless *hpt* gene and a *nos 3′* terminator in pUC19; Drs H. Koshinsky and D. Ow, USDA/ARS Plant Gene Expression Center) was digested with *Not*I and *Hin*dIII. After gel electrophoresis, a 1.9 kb DNA fragment containing a *lox* site, a promoterless *hpt* (hygromycin phosphotransferase coding region) gene and a *nos 3′* terminator from plasmid p24 was isolated, gel purified and ligated on the 3′-end at the *Not*I site of pBACKAN with T4 DNA ligase (US Biochemical) (1× ligation buffer comprising 66 mM Tris–HCl, pH 7.6, 6.6 mM $MgCl_2$, 10 mM DTT, 66 mM ATP). At the 5′-end, the fragment was inserted by site-specific recombination *in vitro* at the *lox* site already present on pBACKAN using Cre recombinase (Dr D. Ow). The Cre reaction was carried out in the same ligation solution (after overnight ligation) with the addition of 33 mM NaCl and 0.2 µg of Cre enzyme in a 20 µl reaction volume for 2 h. After construction, all of the cloning junctions and the *hpt* gene were verified by DNA sequencing. The new vector is called pBACwich.

Generation of a promoter–*lox*–Cre integration site

Tobacco. In tobacco, a line (1999.5) harboring a *35S–lox–cre* construct was used (14), kindly provided by Dr D. Ow.

Magnaporthe grisea. Plasmid pGGB1001 (Drs E. Fang and R. Dean, Clemson University) was digested with *Nco*I and treated with mung bean nuclease (Promega) to remove the protruding 5′-ends. A 5.8 kb fragment containing BAR (Basta resistance), Pgpd (the glyceraldehyde 3-phosphate dehydrogenase promoter from *Aspergillus nidulans*) and TtrpC (the *A.nidulans trpC* terminator) was gel purified. A 1.1 kb fragment containing *lox76*–cre was amplified from plasmid pED98 (Dr D. Ow) by PCR. PCR reactions were performed in a total volume of 25 µl containing 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl, 0.1 mg/ml gelatin, 200 mM dNTPs, 1 mM primers (lox, 5′-P-GAG GAT CCA TAA CTT CGT ATA GC-3′; nos, 5′-P-CTT GCC TTT CCC CGC ATG AA-3′) and 1 U *Taq* DNA polymerase. The following PCR program was used: 94°C for 1 min for 1 cycle, then 94°C for 1 min, 55°C for 1 min 10 s, 72°C for 2 min for 30 cycles. After the last cycle, the polymerization step at 72°C was extended for an additional 7 min. Both the 5.8 and 1.1 kb fragments were ligated to create plasmid pSD. pSD was transformed into *M.grisea* by electroporation (Mr H. Zhu of Dr R. Dean's laboratory). Several transgenic *M.grisea* harboring a copy of pSD and expressing the *cre* gene were confirmed by Southern analysis of pSD and northern analysis of *cre* mRNA.

Tobacco transformation

Tobacco seeds (1999.5, harboring a single *35S–loxP–npt* construct) were sterilized in 100% ethanol for 1 min and in bleach solution (1:2 dilution of household bleach in sterile ddH_2O with one drop of Tween 20 per 10 ml of diluted solution). The seeds were washed three to four times with sterile ddH_2O and then germinated on Murashige Minimal Organics Medium (Gibco BRL) containing kanamycin at 100 µg/ml to select against segregants.

Biolistic transformation conditions for tobacco were selected as 1 µm diameter gold microparticles and a helium pressure of 900 p.s.i., based on preliminary experiments. Gold microparticles were prepared and coated with plasmid pBACwich (2 µg/bombardment) essentially as described by Sanford *et al*. (20). After biolistic bombardment (Bio-Rad, USA) with pBACwich50, -150 and -230, leaf explants (≈1 cm^2) were placed on shoot-inducing Murashige-Skoog medium containing hygromycin at 20 µg/ml. Plants from leaf explants which gave rise to shoots were considered as having hygromycin-resistant cells.

Molecular analyses

PCR analysis incorporated two primer pairs (s and c; w and h), with one primer from the *35S* sequence (primer s, 5′-GTT CAT TTC ATT TGG AGA GG-3′), one from the pBACwich *hpt* gene (primer h, 5′-GGT GTC GTC CAT CAC AGT TTG CCA G-3′), one from the pBACwich backbone (primer w, 5′-GAT GGC CTC CAC GCA CGT TGT G-3′) and the other from the *cre* gene (primer c, 5′-CTA ATC GCC ATC TTC CAG CAG

G-3′). The reaction conditions were 94°C for 1 min, 55°C for 1 min, 72°C for 1 min 10 s, 35 cycles, and 72°C for 10 min.

For Southern analysis, tobacco DNA was isolated from young leaves (21), digested with various restriction enzymes (~8 µg for each sample), separated on 0.9% agarose gels and probed with various probes, including internal fragments from the cotton DNA inserts. For CHEF (contour-clamped homogeneous electric field) electrophoresis analysis, tobacco nuclei were isolated from young leaves and embedded in agarose plugs. After enzyme digestion, DNA fragments were separated with a CHEF Mapper (Bio-Rad).

DOP (partially degenerate oligonucleotide primer) PCR

To isolate the vector–insert junction sequences from BAC clones, DOP (22,23) was used in combination with vector primers. The DOP primer (5′-CCG ACT CGA GNN NNN NAT GTG G-3′) has six bases specified at the 3′-end, six degenerate bases in the middle and 10 arbitrary bases at the 5′-end. An aliquot of 1 µl (5 ng/µl) of the BAC DNA was added to 7 µl of DOP1 mix [2 µl of 5× Sequenase buffer (100 mM potassium phosphate, pH 7.4, 5 mM DTT, 0.5 mM EDTA; US Biochemical), 1 µl of 2.5 mM dNTP, 1 µl of 10 mM DOP primer and 3 µl of water]. The following thermocycler program was used: 96°C for 3 min, 30°C pause (at this step, 2 µl of 1 U/µl Sequenase was added), a 30–37°C ramp over 1 min, 37°C for 3 min and 72°C for 10 min.

An aliquot of 40 µl of DOP2 mix [5 µl of 10× PCR buffer (500 mM KCl, 100 mM Tris–HCl, pH 8.3, 0.01% gelatin; Perkin-Elmer), 2 µl of 2.5 mM dNTP, 1 µl of 10 µM DOP primer, 1 µl of 10 µM BAC vector forward (5′-TGG GTA ACG CCA GGG TTT TC-3′) or backward (5′-CGG CTC GTA TGT TGT GTG GAA-3′) primer, 1 µl of 1 U/µl Ampli*Taq* (Perkin-Elmer) and 30 µl of water] was added to each tube. PCR was executed for 35 cycles at 94°C for 30 s, 58°C for 30 s and 72°C for 1 min. In the secondary PCR for further enrichment of junction sequences, 0.5 µl of primary PCR product was used in a 50 µl reaction under the same conditions except that the nested forward (5′-CCT GCA GGC ATG CAA GCT T-3′) or nested backward (5′-CAC TAT AGA ATA CTC AAG C-3′) vector primer was used.

RESULTS

pBACwich vector construction

We constructed a new BAC vector with a promoterless hygromycin phosphotransferase gene that was designed to be activated and selected for in a plant cell upon site-specific integration. As outlined in Figure 1, plasmid p24, which contains a *lox* site, a promoterless *hpt* gene and a *nos 3′* terminator, was digested with *Not*I and *Hin*dIII to obtain a 1.9 kb DNA fragment containing a *lox* site, a promoterless *hpt* gene and a *nos 3′* terminator. The 1.9 kb fragment was ligated on the 3′-end at the *Not*I site of pBACKAN with T4 DNA ligase. The 5′-end of the fragment was inserted by SSR *in vitro* at the *lox* site already present on pBACKAN using Cre recombinase as diagrammed in Figure 1. The new vector was named pBACwich (derived from the constructors, Wing and Choi) (Fig. 2). After construction, all of the cloning junctions were verified by DNA sequencing.

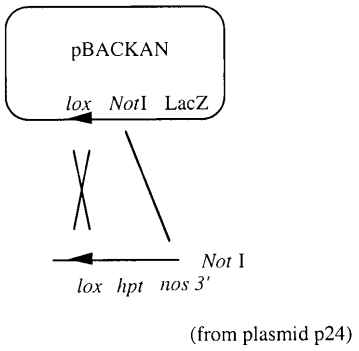

Figure 1. Construction of pBACwich. pBACKAN was modified by inserting a DNA fragment that contains a *lox* site, a promoterless *hpt* (hygromycin phosphotransferase coding region) gene and the *nos 3′* polyadenylation sequence (*lox–hpt–nos 3′*). The DNA fragment from plasmid p24 (Dr D. Ow, Plant Gene Expression Center) was ligated on the 3′-end at the *Not*I restriction site. At the 5′-end, the fragment was inserted by site-specific recombination *in vitro* at the *lox* site already present on the BAC vector using Cre recombinase. After construction, all of the cloning junctions were verified by DNA sequencing.

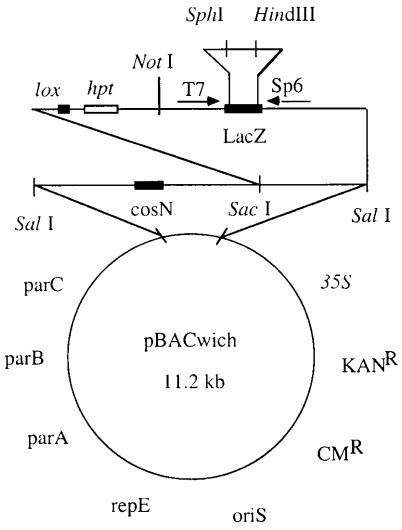

Figure 2. Diagram of the pBACwich vector. The plasmid is based on pBeloBAC11. CM^R, chloramphenicol resistance; *35S*, 35S dual enhancer promoter; KAN^R, kanamycin resistance; *hpt*, hygromycin resistance. *35S* and KAN^R can be used for random transformation in plants and *hpt* can be used for site-specific integration in plants or animals.

pBACwich library construction

We used pBACwich to construct a 32 729 clone cotton BAC library with an average insert size of 110 kb as described in Choi and Wing (21). pBACwich is designed to integrate into a *lox* site, located in a plant or fungal genome, by Cre/*lox* site-specific recombination.

To verify the transformation potential of the pBACwich system within a convenient time period and to apply the system to fungal genome research, we chose the fungus *Magnaporthe grisea*, which is known as the causal agent of blast disease of rice (24). A plasmid (pSD) containing the *A.nidulans gpdA*

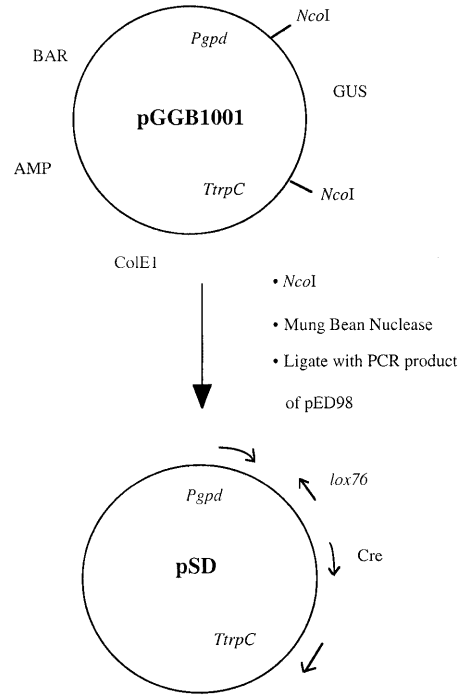

Figure 3. Construction of pSD. Details are provided in Materials and Methods.

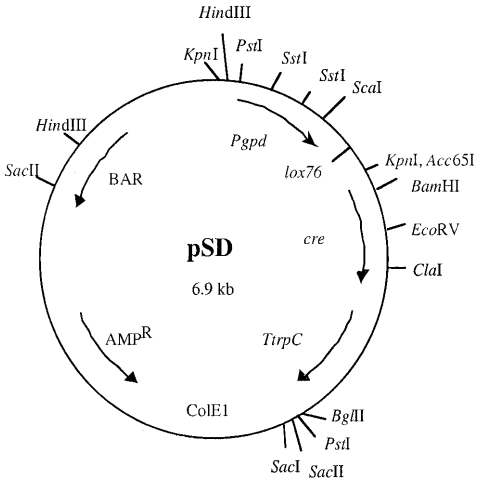

Figure 4. The structure of plasmid pSD to create a *lox* site in *M.grisea*. Pgpd, *A.nidulans gpdA* (glyceraldehyde-3-phosphate dehydrogenase) promoter; *lox76*, 5′-ATAACTTCGTATAGCATACATTATACGcccggta-3′; *cre*, Cre recombinase; *hpt*, hygromycin phosphotransferase; TtrpC, *A.nidulans trpC* terminator; AMP^R, ampicillin resistance; BAR, Basta resistance.

(constitutively regulated) promoter fused to a *lox76–hpt* cassette using the *bar* gene as a selectable marker was constructed as outlined in Figures 3 and 4. Plasmid pGGB1001 was digested with *Nco*I and treated with mung bean nuclease (Promega) to remove the protruding 5′-ends. The 5.8 kb fragment containing BAR, Pgpd and TtrpC was ligated into the 1.1 kb fragment (containing *lox76–cre*) PCR amplified from plasmid pED98. The new vector was named pSD.

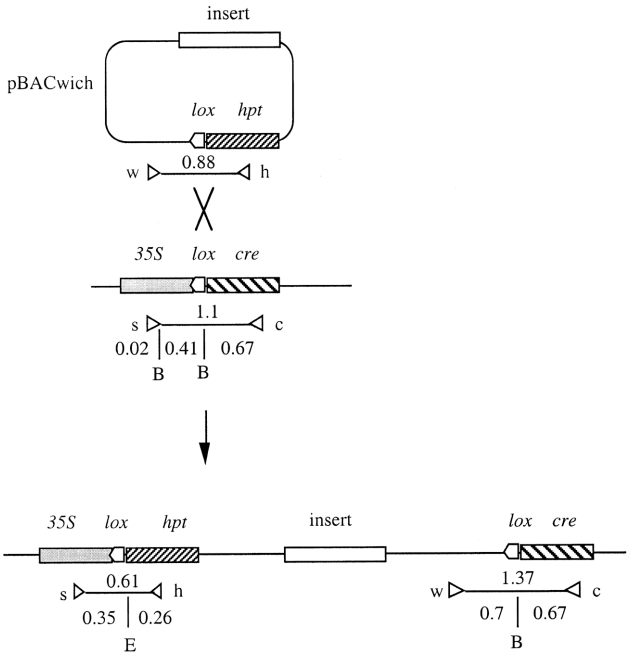

Figure 5. Experimental design of Cre-mediated site-specific recombination. PCR primers (s, h, w and c) are shown as triangles, with the numbers indicating the sizes in kb of the expected PCR products and the DNA fragments after digestion with *Eco*RI (E) and *Bam*HI (B).

Six appropriate recipient *M.grisea* strains were created by transformation with pSD. These recipient strains were transformed with pBACwich while selecting for hygromycin resistance to test the efficiency of Cre/*lox* recombination in *M.grisea*. Five *Magnaporthe* transformants resistant to hygromycin (Hpt^R) were propagated and the DNAs were isolated and analyzed by PCR and Southern hybridization to verify that pBACwich integrated into the *Magnaporthe* genome by site-specific recombination. pBACwich (two out of five transformants) was successfully integrated into the *lox* site of *M.grisea* (H.Zhu, personal communication). These results demonstrate that pBACwich can be integrated into a fungal genome by SSR.

To determine if this system could be used for integration of large insert DNA, we tested the system in tobacco. We selected three cotton pBACwich clones containing DNA inserts of increasing size, 50 (pBACwich50), 150 (pBACwich150) and 230 kb (pBACwich230), to transform into a *lox* site within the tobacco genome using biolistic transformation. We transformed tobacco plants previously transformed with a *35S–lox–cre* construct with these test constructs by biolistic particle bombardment. Transformants were selected for growth on hygromycin, the plants regenerated and the cotton DNA inserts analyzed to determine if the entire cotton DNA fragment was transferred into tobacco.

Figure 5 shows a diagram of how pBACwich is predicted to integrate into the plant genome. A *35S–lox–cre* construct, where *35S* is a cauliflower mosaic virus promoter, was placed in the tobacco genome (14). The pBACwich plasmid bears a promoterless *lox–hpt* gene, where *hpt* encodes hygromycin phosphotransferase, followed by a *nos 3′* polyadenylation site. Integration of the plasmid via *lox* site recombination would

produce a *35S–lox–hpt–lox–cre* linkage, resulting in a hygromycin-resistant (Hpt^R) phenotype. Termination of *cre* transcription should result in an insertion event that is stably maintained.

Particle bombardment of tobacco line 1999.5 with the selected pBACwich clones resulted in tens of Hpt^R calli per transformation. Table 1 shows the number of hygromycin-resistant calli obtained per bombardment. The reason why we recovered fewer resistant calli for the 50 kb BAC may be that increased degradation by vortexing of pBACwich50 occurred during the bombardment step (data not shown).

Table 1. Tobacco transformants obtained from biolistic bombardment

	pBACwich50	pBACwich150	pBACwich230
Bombardment 1	10–20	40–50	<10
Bombardment 2	30–40	40–50	10–20
Bombardment 3	20–30	60–70	10–20

We verified the presence of the cotton pBACwich DNA (150 kb) in the tobacco genome using PCR and Southern blot hybridization. PCR analysis incorporated two primer pairs (s and c; w and h; Fig. 5), with one primer from the *35S* promoter region (primer s) and another from the pBACwich hygromycin resistance region (primer h). The other primers were from the pBACwich backbone region (primer w) and the *cre* coding region (primer c). Figure 6 shows PCR and restriction digestion analysis of the PCR products. A specific band (1.1 kb) was amplified using primers s (5′-GTT CAT TTC ATT TGG AGA GG-3′) and c (5′-CTA ATC GCC ATC TTC CAG CAG G-3′) from the recipient plant (1999.5, harboring a single *35S–loxP–npt* construct), but not from plants transformed with pBACwich150. Two bands, obtained with the primer combinations s and h (5′-GGT GTC GTC CAT CAC AGT TTG CCA G-3′) (0.61 kb) and w (5′-GAT GGC CTC CAC GCA CGT TGT G-3′) and c (1.37 kb), were amplified from pBACwich150-transformed tobacco but not from the recipient tobacco. The size distribution of the PCR products indicates that there was a *lox* site-specific recombination event between pBACwich and the recipient tobacco plant. In addition, these PCR products have a single characteristic *Eco*RI or *Bam*HI site, respectively (Fig. 5). *Bam*HI cleaved the 1.1 kb s+c product from the recipient into three fragments of 0.02, 0.41 and 0.67 kb. The 0.61 kb s+h product from the pBACwich150-transformed plant was cleaved by *Eco*RI into two fragments of 0.35 and 0.26 kb. The 1.37 kb w+c product from the pBACwich150-transformed tobacco was also cleaved by *Bam*HI into two fragments of 0.7 and 0.67 kb (Fig. 6). The same analysis was performed with 14 individual pBACwich150-transformed plants showing Hyg^R (12 from one explant, one from another explant and one from the third explant). In all cases the DNAs produced the exact predicted PCR pattern (data not shown).

Plants containing the pBACwich DNA in a random position are sensitive to hygromycin due to the lack of *hpt* transcription. The recipient plant (1999.5) has the construct *35S–loxP–cre*, where the coding region of the Cre recombinase (*cre*) is preceded by a *loxP* site (wild-type *lox* site) followed by the *35S* promoter. Expression of the *cre* gene upon integration of

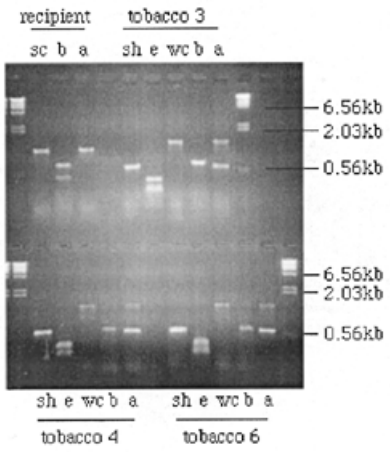

Figure 6. PCR analysis and restriction analysis of the PCR products. A specific band (1.1 kb) was amplified using primers s and c (sc) from the recipient plant (1999.5, harboring a single *35S–loxP–npt* construct) but not from the pBACwich150-transformed plants (tobacco 3, 4 and 6) (see Fig. 5). All primers (s, c, h and w) were used in lane a for each DNA to see if any fragments would be amplified by any combinations. Two bands, obtained with primers s and h (sh, 0.61 kb) and w and c (wc, 1.37 kb), were amplified from pBACwich150-transformed tobacco but not from the recipient tobacco. *Bam*HI cleaved the 1.1 kb sc product from the recipient into three fragments of 0.02, 0.41 and 0.67 kb (b in recipient). The 0.61 kb sh product from the pBACwich150-transformed plant was cleaved by *Eco*RI into two fragments of 0.35 and 0.26 kb (e in tobacco 3, 4 and 6). The 1.37 kb wc product from the pBACwich150-transformed tobacco was also cleaved by *Bam*HI into two fragments of 0.7 and 0.67 kb (b in tobacco 3, 4 and 6).

pBACwich into the *lox* site will promote recombination between the *lox* sites in pBACwich and in plants. The *35S* promoter is disengaged from the *cre* coding region and fused to the *hpt* (hygromycin phosphotransferase) coding sequence. The lack of a functional promoter in front of the *cre* gene will terminate *de novo* synthesis of Cre enzyme and will eventually curtail further recombination of the *lox* sites. All PCR products have a predicted insert size and should be present if the BAC clone has integrated at the *lox* site within the tobacco genome (Fig. 6).

To determine if the entire 150 kb fragment integrated intact into the SSR site of tobacco line 1999.5, Southern blot analysis was performed on tobacco DNA digested with various restriction enzymes and hybridized with various probes (*hpt* and *cre*), including the internal cotton insert fragments (Figs 7 and 8). Figure 7 shows Southern blot analyses of tobacco genomic DNA from five R_0 tobacco plants transformed with pBACwich150. The recipient tobacco is referred to as 1999.5 and 3, 12, 4, A and B refer to tobacco plants transformed with pBACwich150. As shown in Figure 7, the restriction fragments that hybridized with the *cre* probe were shifted in all transformants, indicating DNA rearrangement in the *cre* gene area. In 1999.5 the *Hin*dIII digest shows a 6.1 kb band when hybridized with either *35S* or *cre* probes (Dr H. Koshinsky, personal communication). In the pBACwich150 transformants the *Hin*dIII digest gave a band of a different size when probed with *cre*. The *hpt* probe hybridized to the transformants but not to the recipient, indicating the presence of *hpt* in the transformants.

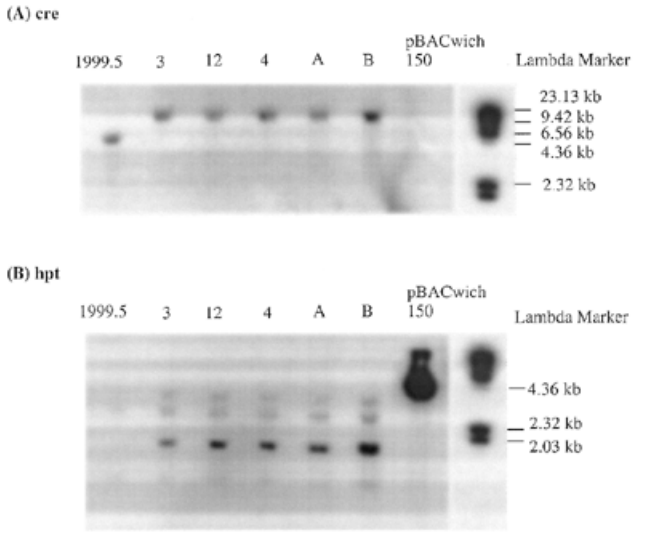

Figure 7. Southern blot analysis of tobacco genomic DNA from R_0 tobacco plants transformed with pBACwich150. All samples were digested with *Hin*dIII. (**A**) The restriction fragments hybridized with *cre* probe were shifted in all transformants, indicating DNA rearrangement (see Fig. 5). (**B**) The *hpt* probe hybridized to transformants that included the *hpt* gene but not to the recipient. The recipient tobacco is referred to as 1999.5, and 3, 12, 4, A and B refer to transformed tobaccos with pBACwich150.

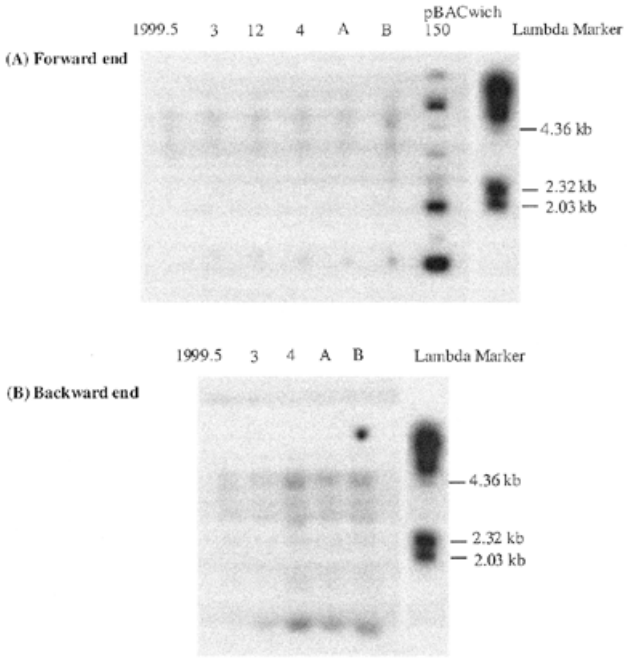

Figure 8. Analysis of tobacco genomic DNA from R_0 tobacco plants transformed with the cotton pBACwich150 insert. (**A**) All samples were digested with *Hin*dIII and hybridized with the left forward DOP PCR product of pBACwich150. (**B**) All samples were digested with *Hin*dIII and *Eco*RV and hybridized with the right backward DOP PCR product of pBACwich150. 1999.5, recipient; 3, 12, 4, A and B, transformants.

Figure 8 shows that the transformants hybridized with the forward and backward DOP PCR products of pBACwich150, indicating the presence of the ends of the 150 kb cotton DNA fragment in the tobacco transformants.

No significant differences in the numbers of integration events recovered were found among the 50, 150 and 230 kb pBACwich transformants. Between two and 14 independent integration events from each type of transformation were assayed for the correct border sequences on either side by PCR. In every case the *35S–lox–cre* junction was no longer present and a new *35S–lox–hpt* junction was found (Fig. 6). Southern blotting also confirmed DNA rearrangements, as shown in Figure 7. Copy number was assessed from Southern analysis of the integrant plants (one band on *hpt* probing; Fig. 7B). In all cases, only a single copy of the pBACwich insert was found integrated at the genomic *lox* target. The stability of the large DNA inserts is being assessed by analyzing tobacco genomic DNA from hygromycin-resistant plants over several generations of selfing and crossing with the wild-type (Wi38).

DISCUSSION

The pBACwich system

We have provided preliminary evidence that the Cre/*lox* system can be used to integrate large intact foreign DNA (up to 230 kb, Table 1) into a specific *lox* site in tobacco. Taken together, these preliminary results with the pBACwich system using Cre/*lox* recombination may offer a method for precise, single copy insertion of large DNA fragments into genomic targets.

Site-specific integration events are now rapidly being achieved in plant cells (14,16,17). The introduction of site-specific recombination systems into cells of higher eukaryotes presents many exciting new possibilities in molecular biology. The advantage of site-specific recombination versus alternative integration methods is that plant DNA will always integrate into a specific genomic location. Thus the phenotype of all integrated DNA is assessed from a single genomic location instead of several locations, as with *Agrobacterium*-mediated transformation. Site-specific integration will enable insertion of large DNA fragments at predetermined locations in a genome and may thereby allow avoidance of the so-called 'position effects' by selection of well-placed insertion sites.

By introducing all candidate BAC clones into the same genomic locus, site-specific integration of large genomic clones can be used to streamline the positional cloning and transfer of agronomically important genes into crop plants. The ability of pBACwich to clone large fragments (up to 350 kb) will reduce the number of subclones and transgenic plants required to determine which clones contain the target gene. By streamlining the gene identification step, target genes could be identified more rapidly, thus leading to progress in the isolation of genes based solely on map position and a more efficient system with which to identify agriculturally important genes isolated by map-based cloning strategies. It will also make possible the cloning and manipulation of large enough DNA fragments to

study the composition and function of elements such as heterochromatin, centromeres and telomeric regions. Recently Choi developed an improved method of construction of BAC libraries and began constructing a series of BAC libraries with much larger insert size from human (202 kb) and a variety of organisms, including *Arabidopsis* (190 kb) and maize (164 kb) (25,26; http://www.tree.caltech.edu). Currently Choi has constructed the Caltech *Arabidopsis* pBACwich library and is working on developing this technique in *Arabidopsis*.

A set (eight) of *35S–lox–cre* transgenic *Arabidopsis* lines, where *35S* is the cauliflower mosaic virus (CaMV) 35S RNA promoter and *lox* is placed in the leader sequence of the *cre* gene, has already been produced (27). The map locations of six *35S–lox–cre* insertions have been determined by RFLP mapping and YAC library hybridization. Site-specific integration into these lines will enable insertion of large DNA fragments at predetermined locations in the *Arabidopsis* genome and avoid 'position effects' by the selection of specific insertion sites. The pBACwich vector also has a CaMV 35S promoter and the *NPTII* gene and can be used for random transformation of any plants (successful with the 150 kb insert in cotton; unpublished work).

The use of site-specific recombination in conjunction with transposable elements (the maize *Ac–Ds* element) opens the way to many additional applications, especially for gene complementation and gene identification. If the maize *dissociation* element (*Ds*) is present within pBACwich, it can be induced to transpose in the presence of the maize *Activator* (*Ac*), which provides a transposase. Since the maize transposons *Ac* and *Ds* tend to transpose to genetically linked sites close to their original position in several plant systems, including maize (28), tobacco (29), tomato (30,31) and *Arabidopsis* (32), once *Ds* is activated by a stable allele of *Ac* (31) there is a high probability that it will transpose into the large DNA insert that has been shown to contain a target gene. In some instances the *Ds* element will transpose into the target gene, thereby creating a mutation in the target gene. By combining site-specific integration with transposon mutagenesis, not only can it be proved that the target gene has been cloned by genetic complementation, but the target gene can also be identified by transposon tagging. In addition, transposon mutagenesis will also create new mutant alleles which may help pinpoint important regulatory or coding regions not yet discovered within the target locus.

ACKNOWLEDGEMENTS

We thank Dr Melvin Simon for critical reading of the manuscript and helpful discussions, Dr Sam Reddy for the gift of plasmid pBACKAN and Drs Eric Fang and Ralph Dean for the gift of plasmid pGGB1001.

REFERENCES

1. van Haaren,M.J. and Ow,D.W. (1993) *Plant Mol. Biol.*, **23**, 525–533.
2. Riordan,J.R., Rommens,J.M., Kerem,B., Alon,N., Rozmahel,R., Grzelczak,Z., Zielenski,J., Lok,S., Plavsic,N. and Chou,J.L. (1989) *Science*, **245**, 1066–1073.
3. Rommens,J.M., Iannuzzi,M.C., Kerem,B., Drumm,M.L., Melmer,G., Dean,M., Rozmahel,R., Cole,J.L., Kennedy,D., Hidaka,N. *et al*. (1989) *Science*, **245**, 1059–1065.
4. Collins,F.S. (1990) *Harvey Lect.*, **86**, 149–164.
5. van Bokhoven,H., van den Hurk,J.A., Bogerd,L., Philippe,C., Gilgenkrantz,S., de Jong,P., Ropers,H.H. and Cremers,F.P. (1994) *Hum. Mol. Genet.*, **3**, 1041–1046.
6. Chandrasekharappa,S.C., Guru,S.C., Manickam,P., Olufemi,S.E., Collins,F.S., Emmert-Buck,M.R., Debelenko,L.V., Zhuang,Z., Lubensky,I.A., Liotta,L.A., Crabtree,J.S., Wang,Y., Roe,B.A., Weisemann,J., Boguski,M.S., Agarwal,S.K., Kester,M.B., Kim,Y.S., Heppner,C., Dong,Q., Spiegel,A.M., Burns,A.L. and Marx,S.J. (1997) *Science*, **276**, 404–407.
7. Arondel,V., Lemieux,B., Hwang,I., Gibson,S., Goodman,H.M. and Somerville,C.R. (1992) *Science*, **258**, 1353–1355.
8. Martin,G.B., Brommonschenkel,S.H., Chunwongse,J., Frary,A., Ganal,M.W., Spivey,R., Wu,T., Earle,E.D. and Tanksley,S.D. (1993) *Science*, **262**, 1432–1436.
9. Cai,D., Kleine,M., Kifle,S., Harloff,H.J., Sandal,N.N., Marcker,K.A., Klein-Lankhorst,R.M., Salentijn,E.M.J., Lange,W., Stiekema,W.J., Wyss,U., Grundler,F.M.W. and Jung,C. (1997) *Science*, **275**, 832–834.
10. Hamilton,C.M., Frary,A., Lewis,C. and Tanksley,S.D. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 9975–9979.
11. Craig,N.L. (1988) *Annu. Rev. Genet.*, **22**, 77–105.
12. Ow,D.W. and Medberry,S.L. (1995) *Crit. Rev. Plant Sci.*, **14**, 239–261.
13. Hoess,R., Abremski,K., Irwin,S., Kendall,M. and Mack,A. (1990) *J. Mol. Biol.*, **216**, 873–882.
14. Albert,H., Dale,E.C., Lee,E. and Ow,D.W. (1995) *Plant J.*, **7**, 649–659.
15. Qin,M., Lee,E., Zankel,T. and Ow,D.W. (1995) *Nucleic Acids Res.*, **23**, 1923–1927.
16. Medberry,S.L., Dale,E., Qin,M. and Ow,D.W. (1995) *Nucleic Acids Res.*, **23**, 485–490.
17. Stuurman,J., de Vroomen,M.J., Nijkamp,H.J. and van Haaren,M.J. (1996) *Plant Mol. Biol.*, **32**, 901–913.
18. Akagi,K., Sandig,V., Vooijs,M., Van der Valk,M., Giovannini,M., Strauss,M. and Berns,A. (1997) *Nucleic Acids Res.*, **25**, 1766–1773.
19. Kim,U.-J., Birren,B.W., Slepak,T., Mancino,V., Boysen,C., Kang,H.L., Simon,M.I. and Shizuya,H. (1996) *Genomics*, **34**, 213–218.
20. Sanford,J.C., Smith,F.D. and Russell,J.A. (1993) *Methods Enzymol.*, **217**, 483–509.
21. Choi,S. and Wing,R.A. (1999) In Gelvin,S. and Schilperoort,R. (eds), *Plant Molecular Biology Manual*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 1–32.
22. Telenius,H., Carter,N.P., Bebb,C.E., Nordenskjold,M., Ponder,B.A. and Tunnacliffe,A. (1992) *Genomics*, **13**, 718–725.
23. Wu,C., Zhu,S., Simpson,S. and de Jong,P.J. (1996) *Nucleic Acids Res.*, **24**, 2614–2615.
24. Ou,S.H. (1985) In *Rice Disease*, 2nd Edn. Commonwealth Mycological Institute, Kew, UK.
25. Choi,S., Sablowski,R., Meyerowitz,E. and Simon,M. (1999) In *Plant and Animal Genome VII*, January 17–21, San Diego, CA, abstract P113.
26. Choi,S., Shizuya,H. and Simon,M. (1999) In *The Seventh DOE Contractor and Grantee Genome Program Workshop*, January 12–16, Oakland, CA, abstract P68.
27. Chou,I., Kobayashi,J., Lee,E., Koshinsky,H., Medberry,S. and Ow,D. (1999) In *Plant and Animal Genome VII*, January 17–21, San Diego, CA, abstract P133.
28. Ralston,E., English,J. and Dooner,H.K. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 9451–9455.
29. Jones,J.D., Carland,F., Lim,E., Ralston,E. and Dooner,H.K. (1990) *Plant Cell*, **2**, 701–707.
30. Osborne,B.I., Corr,C.A., Prince,J.P., Hehl,R., Tanksley,S.D., McCormick,S. and Baker,B. (1991) *Genetics*, **129**, 833–844.
31. Healy,J., Corr,C., DeYoung,J. and Baker,B. (1993) *Genetics*, **134**, 571–578.
32. Long,D., Goodrich,J., Wilson,K., Sundberg,E., Martin,M., Puangsomlee,P. and Coupland,G. (1997) *Plant J.*, **11**, 145–148.