Mol Gen Genet (1999) 261: 92–99

ORIGINAL PAPER

A. H. Hofmann · A. C. Codón · C. Ivascu · V. E. A. Russo
C. Knight · D. Cove · D. G. Schaefer
M. Chakhparonian · J.-P. Zrÿd

A specific member of the *Cab* multigene family can be efficiently targeted and disrupted in the moss *Physcomitrella patens*

Received: 2 February 1998 / Accepted: 15 October 1998

Abstract The analysis of phenotypic change resulting from gene disruption following homologous recombination provides a powerful technique for the study of gene function. This technique has so far been difficult to apply to plants because the frequency of gene disruption following transformation with constructs containing DNA homologous to genomic sequences is low (0.01 to 0.1%). It has recently been shown that high rates of gene disruption (up to 90%) can be achieved in the moss *Physcomitrella patens* using genomic sequences of unknown function. We have used this system to examine the specificity of gene disruption in *Physcomitrella* using a member of the *Cab* multigene family. We have employed the previously characterised *Cab* gene *ZLAB1* and have isolated segments of 13 other closely related members of the *Cab* gene family. In the 199-bp stretch sequenced, the 13 new members of the *Cab* family show an average of 8.5% divergence from the DNA sequence of *ZLAB1*. We observed 304 silent substitutions and 16 substitutions that lead to a change in the amino acid sequence of the protein. We cloned 1029 bp of the coding region of *ZLAB1* (including 177 of the 199 bp with high homology to the 13 new *Cab* genes) into a vector containing a selectable hygromycin resistance marker, and used this construct to transform *P. patens*. In three of nine stable transformants tested, the construct had inserted in, and disrupted, the *ZLAB1* gene. There was no discernible phenotype associated with the disruption. We have therefore shown that gene disruption is reproducible in *P. patens* and that the requirement for sequence homology appears to be stringent, therefore allowing the role of individual members of a gene family to be analysed in land plants for the first time.

Key words Gene disruption · Gene targeting · Homologous recombination · Plant transformation · *Physcomitrella patens*

Communicated by J. Schell

A. H. Hofmann (✉) · A. C. Codón · C. Ivascu · V. E. A. Russo
Max-Planck-Institut für Molekulare Genetik, Ihnestrasse 73
D-14195 Berlin, Germany
e-mail: Hofmann_a@mpimg-berlin-dahlem.mpg.de
Tel.: +49-30-8329766, Fax: +49-30-84131774

C. Knight · D. Cove
Department of Biology, The University of Leeds
Leeds LS2 9JT, UK

D. G. Schaefer · M. Chakhparonian · J.-P. Zrÿd
Laboratoire de Phytogénétique Cellulaire
University of Lausanne, CH-1015 Lausanne-Dorigny
Switzerland

Introduction

The integration of foreign DNA into eukaryotic genomes occurs mainly at random loci by illegitimate recombination, even when the introduced DNA has homology to endogenous sequences (Roth and Wilson 1988; Bollag et al. 1989; Capecchi 1989a; Puchta et al. 1994; Puchta and Hohn 1996). Exceptions to this generalization include *Saccharomyces cerevisiae* (Rothstein 1991), *Schizosaccharomyces pombe* (Grimm and Kohli 1988) and some filamentous fungi (Timberlake and Marshall 1989). Direct integration by homologous recombination (gene targeting) is rare compared with non-homologous integration, but would provide a powerful technique for the molecular genetic study of higher organisms and the determination of gene function.

Although gene targeting is now a well established tool for the specific inactivation or modification of genes in yeast and mouse embryonic stem cells (Struhl 1983; Capecchi 1989b; Rossant and Joyner 1989; te Riele et al. 1992), the development of similar techniques for plant systems is still at a very early stage (Puchta and Hohn 1996). Some groups have reported genomic integration by homologous recombination in plant cells using DNA introduced by direct gene transfer or *Agrobacterium*-mediated DNA transfer (Paszkowski et al. 1988; Offringa et al. 1990, 1993; Halfter et al. 1992). The highest rates of gene disruption so far published for flowering

plants are between 0.01 and 0.1% (Lee et al. 1990; Miao and Lam 1995; Kempin et al. 1997), and these are too low to allow routine gene disruption. However, it has recently been shown that integrative transformants resulting from homologous recombination can be obtained following polyethylene glycol (PEG)-mediated direct gene transfer into protoplasts of the moss *Physcomitrella patens*, at an efficiency comparable to that found for *S. cerevisiae* (Schaefer and Zrÿd 1997).

We have now investigated the specificity of homologous recombination by disrupting one of the members of a multigene family. We have studied the efficiency of gene disruption using the *ZLAB*1 gene (Long et al. 1989), which codes for a major chlorophyll-a/b-binding protein (CAB) in the light-harvesting complex of photosystem II. This gene belongs to a gene family containing at least 15 genes (Long et al. 1989; Reski et al. 1994).

Here we report that the *ZLAB*1 gene is specifically disrupted in *P. patens* in three out of nine (30%) integrative transformants analysed. Thus, gene disruption in *P. patens* is efficient, even when directed against a specific member of a multigene family, and shows a stringent requirement for sequence homology.

Materials and methods

Construction of the targeting plasmids

Physcomitrella patens genomic DNA was isolated according to Reski et al. (1994). A 1029-bp fragment of the *ZLAB1 Cab* gene (Long et al. 1989) was amplified by PCR using primer X (5′-CTC GGC GCC CTT GGC TGG GCA GA-3′), which is homologous to a region 32 bp downstream of the translation start (positions 1301–1324 of the sequence published by Long et al. 1989), and primer Y (5′-CGG GCG CCT GCA CGA AGA ACC CG-3′), which is complementary to a region 1042 bp downstream of the translation start (positions 2319–2342 of the published *ZLAB1* genomic sequence). Both primers contain a *Nar*I site (underlined), and mutated bases are indicated in bold face. The coding region of the *ZLAB1* gene starts at position 1277 and ends at position 2444 (Long et al. 1989). The amplified coding region therefore lacks the first 32 bp (coding for the first 11 N-terminal amino acids) and the last 104 bp (coding for the last 35 C-terminal amino acids).

PCR was performed in a thermocycler (Biometra) using 100 ng of genomic DNA in a total volume of 100 µl. The reaction mixture contained 20 mM TRIS-HCl pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 10 mM $MgSO_4$, 0.1% Triton X-100, 0.1 mg/ml BSA, 1 µM of each primer, 200 µM of each dNTP, 1 unit of Taq polymerase (Perkin Elmer), and 1 unit of Taq Extend PCR additive (Stratagene). DNA was denatured at 94° C for 2 min, and amplification was carried out in a two-step PCR reaction comprising 25 cycles of 45 s denaturation at 94° C and 2 min annealing/elongation at 65° C, followed by a three-step PCR reaction with 10 cycles of 45 s denaturation at 94° C, 1 min annealing at 65° C, and 2 min elongation at 75° C. The final elongation step was allowed to proceed for 5 min, and the reaction mixture was then cooled and stored at 4° C. A 10-µl aliquot of the reaction mixture was analysed on a 1.2% agarose gel in TAE buffer (Sambrook et al. 1989). The resulting 1042-bp fragment was cloned directly using the TA cloning kit (Invitrogen) according to the protocol provided. The resulting plasmid was digested with the restriction enzyme *Nar*I (Boehringer-Mannheim). After *Nar*I digestion, the 1035-bp fragment was isolated using Jetsorb (Genomed) according to the protocol provided, and subcloned into the plant transformation vector pGL-2 (Bilang et al. 1991), which had been digested with *Nar*I, resulting in the plasmid pGL*cab*-a (see Fig. 2A).

Plant tissue culture, transformation and selection of transgenic plants

The Gransden wild-type strain of *P. patens* was used in this study (Ashton and Cove 1977). Plant tissue culture conditions, protoplast isolation, PEG-mediated direct gene transfer to protoplasts, and selection of integrative transformants was performed as described by Schaefer et al. (1994) and Schaefer and Zrÿd (1997).

Plant genomic DNA extraction and Southern analysis

Plant genomic DNA was extracted with cetyl-trimethyl-ammonium-bromide (CTAB) (Schaefer and Zrÿd 1997). Aliquots (3 µg) of genomic DNA were digested to completion with *Eco*RI or *Xba*I according to the enzyme supplier's instructions and blotted onto nylon membranes (Biorad Zetaprobe) as previously described (Schaefer and Zrÿd 1997). The 5′ end-specific probe was isolated from the original plasmid containing the sequenced *ZLAB1* locus (Long et al. 1989; a gift from N. Nelson) as an *Xba*I-*Bst*XI fragment (bp 165 to 1053), purified by gel electrophoresis and isolated according to standard procedures (Sambrook et al. 1989). Radiolabelling of the probes and Southern hybridisation was performed as previously described (Schaefer and Zrÿd 1997).

PCR analysis

DNA was extracted as described above. PCR was performed in a thermocycler (Bio-med 60) using a total volume of 50 µl. The reaction mixture contained 350 µM of each dNTP (Pharmacia), 300 nM of each primer, 50 mM TRIS-HCl pH 9.2, 16 mM $(NH_4)_2SO_4$, 1.75 mM $MgCl_2$, and 2.5 units of Expand long-template enzyme mix containing *Taq* and *Pwo* polymerase (Boehringer Mannheim). The DNA was diluted 20-fold and 1 µl of the dilution (approximately 100 ng) was used for PCR. The DNA was denatured at 94° C for 2 min. Amplification was carried out for 30 cycles of 10 s denaturation at 94° C and 30 s annealing at 58° C followed by 2 min elongation at 68° C. The final elongation step was extended to 10 min, and the reaction mixture was then cooled and stored at 4° C. Aliquots (8 µl) of the reaction mixture were analysed on a 1.2% agarose gel in TAE (Sambrook et al. 1989) using various combinations of the following primers (see Fig. 4): g1, 5′-AGC ACT TAA AGG ACC GAT GTA ATC TCT C-3′ (positions 359–386 in the *ZLAB1* sequence published by Long et al. 1989); p1 (on pGL2), 5′-GCA TCA GAG CAG ATT GTA CTG AGA GTG-3′; g2, 5′-GTC CTA CTC TCA AGT GAT TTC GAC GGA-3′ (positions 2451–2477; Long et al. 1989); and p2 (on pGL2), 5′-CCA GCT GGC GTA ATA GCG AAG AGG C-3′.

Isolation of novel members of the *Cab* multigene family

DNA was extracted as described above and amplified with the primers cab-2147 (5′-ATG GGA GCC GTG GAG GGG TAC CGT G-3′; positions 2147–2172 in the *ZLAB1* sequence published by Long et al. 1989) and cab-2394rev (5′-ACC GGG TCC GCC AAG TGG TCG TTC-3′; positions 2370–2394). PCR was performed in a thermocycler (Bio-med 60) using a total volume of 50 µl. The reaction mixture contained 350 µM of each dNTP (Pharmacia), 300 nM of each primer, 50 mM TRIS-HCl pH 9.2, 16 mM $(NH_4)_2SO_4$, 1.75 mM $MgCl_2$, 2.5 units of Expand long-template enzyme mix containing *Taq* and *Pwo* polymerase (Boehringer Mannheim). Approximately 400 ng of template DNA was used for each reaction. The DNA was denatured at 92° C for 2 min. Amplification was performed for 30 cycles of 10 s denaturation at

92° C, 30 s annealing at 52° C, and 2 min elongation at 68° C, each successive elongation step being extended by 5 s. The final elongation step was for 10 min and the reaction mixture was then cooled and stored at 4° C. Aliquots (10 µl) of the reaction mixture were analysed on a 1% agarose gel as described above.

Sequence analysis

PCR fragments obtained using the primer pairs g1 + p2 and p1 + g2 were cloned directly using the TA cloning kit (Invitrogen) according to the protocol provided by the manufacturer. The resulting bacterial transformants were analysed by PCR with the above primer pairs, and those containing the correct inserts were expanded and plasmid DNA minipreps were prepared as described by Sambrook et al. (1989).

Sequence analysis was performed by cycle sequencing using 25 ng of DNA per 2 µl of d/ddNTP reaction mixture (Promega Silver sequencing kit), 1.25 µl of special sequencing buffer (a 50:1 mixture of the 5× sequencing buffer provided with the Promega Silver sequencing kit with Triton X-100), 2.5 pmol of fluorescently labelled primer, 2 µl of sterile water, and 0.25 µl of *Taq* polymerase (Promega Silver sequencing kit). The plasmid DNA was denatured at 93° C for 2 min. Amplification was for 36 cycles of 20 s denaturation at 94° C, 20 s annealing at 55° C, and 40 s elongation at 72° C. The final elongation step was for 10 min and the reaction mixture was then cooled and stored at 4° C after addition of 3 µl of sequencing stop buffer (Pharmacia) had been added. Sequences were read by ALF (Automated laser fluorescence, Pharmacia). Fluorescently labelled universal primer and reverse primer (Pharmacia) were used for sequencing.

Results

Isolation of new members of the *Cab* multigene family

Using PCR, the regions corresponding to the bases 2172 to 2370 of the *ZLAB1* gene were isolated from other members of the *Cab* multigene family, cloned and sequenced (see Materials and methods). Analysis of these clones revealed that there are at least 13 more genes within the *Cab* family that show a high degree of sequence identity to the *ZLAB1* gene (86–94% at the DNA level, see Fig. 1). The sequenced 3′ portions of the multigene family vary at the DNA level but base substitutions are silent in 95% of the cases; the few exceptions are indicated in bold face in Fig. 1. At the protein level, the products of the new *Cab* genes are 97–100% identical to that of *ZLAB1*.

Gene disruption strategy

We constructed the targeting plasmid pGL*cab*-a by cloning a portion of the *ZLAB1* coding region lacking

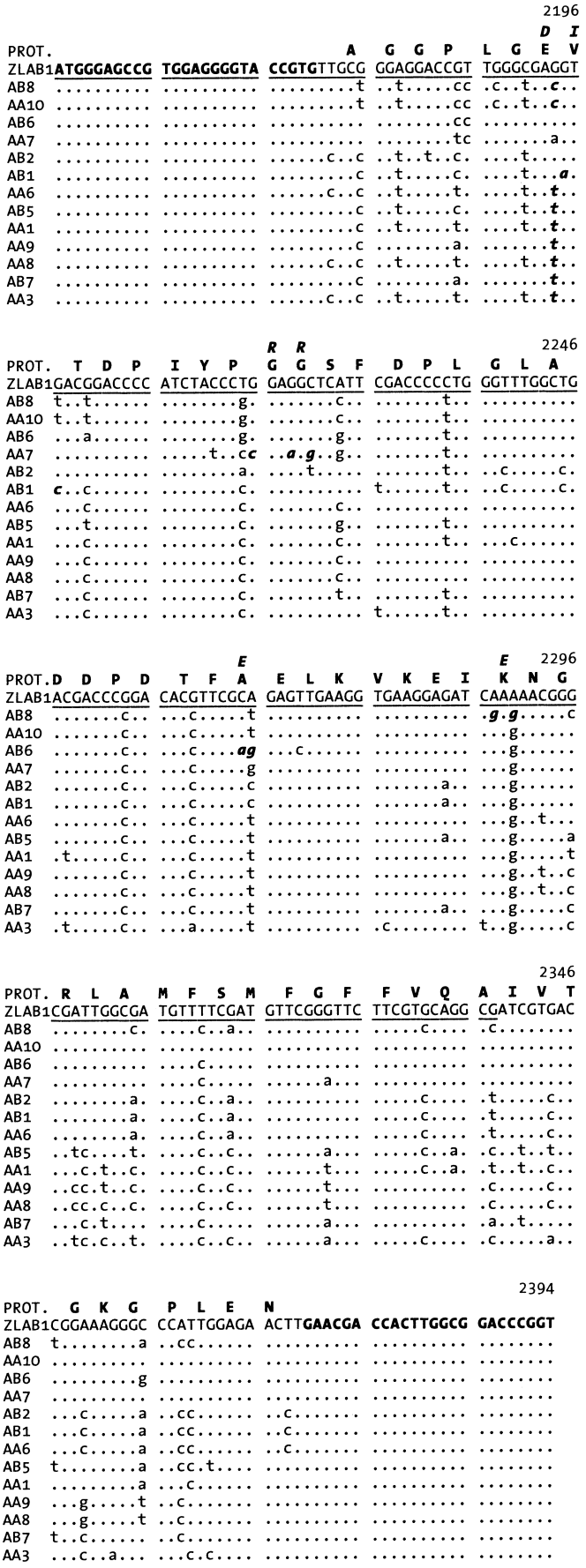


Fig. 1 Partial sequences of 13 new members of the *Cab* multigene family aligned with the published *ZLAB1* sequence (Long et al. 1989). These gene members are highly conserved at the DNA and protein level. All base substitutions are silent except for those marked in *bold face*. The sequences are numbered as for the *ZLAB1* gene (Long et al. 1989). The *underlined* sequence corresponds to the part of the *ZLAB1* contained in the plasmid pGL*cab*-a. The amino acids in *bold italics*, above the protein sequence, are the amino acids specified by the few non-silent substitutions

the bases coding for the 11 N-terminal and the last 35 C-terminal amino acids (see Materials and methods and Fig. 2A). Gene disruption should occur following homologous reciprocal exchange between the portion of the *ZLAB1* coding region in the targeting plasmid pGL*cab*-a and the corresponding region in the chromosomal *ZLAB1* locus, since neither of the two gene copies resulting from integration are full length – one is truncated at the 3′ end (missing the sequence coding for the last 35 amino acids) and the other at the 5′ end of the gene (lacking the whole promoter region and the sequence coding for the first 11 amino acids, see Fig. 2B).

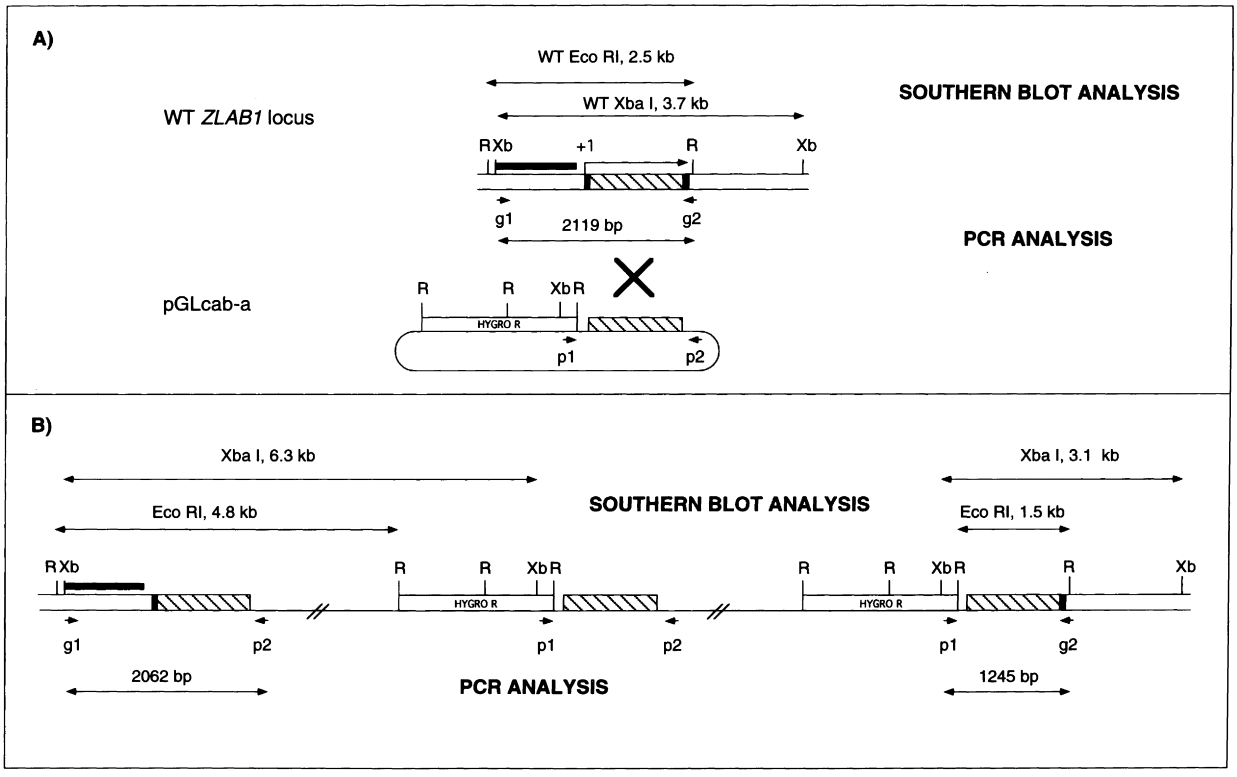
Following transformation of *P. patens* with the targeting plasmid pGL*cab*-a, stable transformants resulting from random integration of pGL*cab*-a should be distinguishable from those resulting from homologous recombination by Southern and PCR analysis (Fig. 2B). Sequencing of the junctions amplified by PCR allows the fidelity of homologous recombination in *P. patens* to be confirmed at the molecular level.

Isolation of integrative transformants

In *P. patens*, following transformation with vectors containing a gene coding for antibiotic resistance, the majority of resistant transformants are unstable (non-integrative; Schaefer et al. 1994). To distinguish stable (integrative) from unstable transformants, antibiotic-resistant plants are grown for 2 weeks on medium lacking antibiotic and then transferred back to selective medium. Only stable transformants retain the antibiotic-resistant phenotype following this procedure (Schaefer et al. 1994). The growth rate and development of stable integrative transformants on hygromycin-containing medium is comparable to that of wild-type *P. patens* grown on non-selective medium, whereas unstable transformants display much slower growth rates and formed gametophores at a very early stage.

We selected putative integrative transformants based on their unrestricted growth on selective medium, and confirmed the mitotic stability of their hygromycin resistance phenotype. No other phenotypic differences

Fig. 2A,B Structures of the wild-type genomic *ZLAB1* locus and the pGL*cab*-a plasmid (**A**), and the predicted structure of the *ZLAB1* locus after the insertion of two copies of pGL*cab*-a by a single cross-over homologous recombination event (**B**). Restriction sites: Xb, *Xba*I, R, *Eco*RI. The translation start site (ATG) of the *ZLAB1* gene (+1) corresponds to position 1277 of the genomic *ZLAB1* locus (Long et al. 1989). The *filled boxes* represent *ZLAB1* coding regions present only in the endogenous *Cab* gene, and the *hatched box* the truncated *ZLAB1* copy present in the targeting vector pGL*cab*-a. The *open box* indicates the hygromycin-resistance cassette. The fragment sizes obtained following digestion of the *ZLAB1* locus are indicated *above* the map in **A** and the predicted fragment sizes following insertion of pGL*cab*-a into the *ZLAB1* locus by a single cross-over are shown *above* the map in **B**. The 5′ probe is indicated by the *black bar* extending to the right from the *Xba*I site upstream of the *ZLAB1* locus (positions −1112 to −225 bp from the translation start site of *ZLAB1*). The 3′ probe (not depicted) extends from position −125 to +560 of *ZLAB1*. The locations of the primers g1, g2, p1 and p2 are shown, and the predicted sizes of PCR-amplified sequences are shown *below* the maps in **A** and **B**



were observed between the stable transgenic strains obtained with pGL*cab*-a and the wild type.

Plasmids containing *P. patens* DNA transform with a higher efficiency than plasmids without such sequences (Schaefer and Zrÿd, 1997). In the experiments reported here, we obtained no integrative transformants from 500,000 regenerating colonies transformed with the control plasmid pGL2 (which contains no *P. patens* DNA). Transformation with pGL2 usually yields stable transformants at a mean relative transformation frequency (RTF) of 1 in 10^5 regenerants (Schaefer and Zrÿd 1997). The RTF observed following transformation with pGL-108 (which contains 3.6 kb of *P. patens* DNA), was similar to that reported previously (146 integrative transformants among 684,000 regenerants, corresponding to 23.5 ± 14.4 transformants per 10^5 regenerants; see Schaefer and Zrÿd 1997). In all, 38 integrative transformants were obtained from 714,000 regenerants following transformation with pGL*cab*-a, corresponding to an RTF of 5.8 ± 4.2 transformants per 10^5 regenerants. This RTF is about three times higher than that routinely obtained with pGL2, and provides evidence that targeted integration of pGL*cab*-a is indeed occurring.

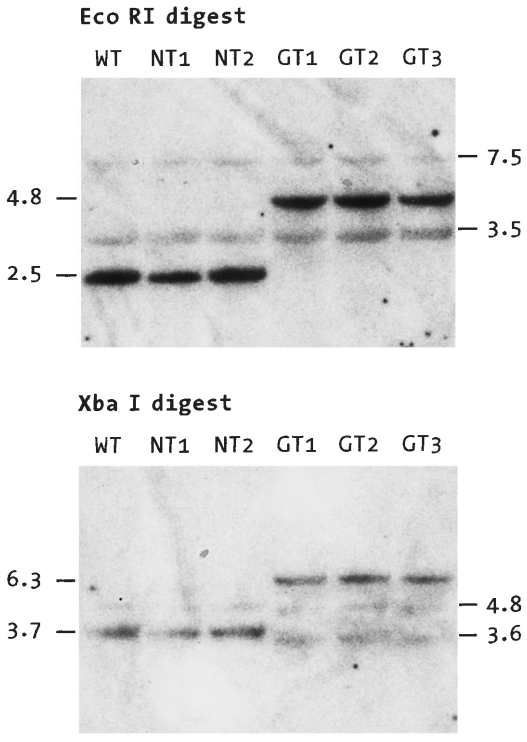


Fig. 3 Southern analysis of *Eco*RI- and *Xba*I-digested DNA from wild type (WT) and pGL*cab*-a transformants using the 5′ probe (see Fig. 2). NT1 and NT2 are non-targeted transformants; GT1-GT3 are targeted transformants. Replacement of the wild-type band by the predicted junction fragment is clearly observed in the three targeted transformants

Southern analysis

We isolated genomic DNA from nine randomly chosen integrative transformants and tested, by Southern analysis, whether the targeted *ZLAB1* gene had been disrupted. To avoid cross hybridisation with plasmid DNA and/or with other members of the *Cab* multigene family, a fragment covering 888 bp of upstream untranslated sequences of the *ZLAB1* gene (extending from the *Xba*I site at position −1112 to the *Bst*XI site at position −225 with respect to the translation start site – corresponding to bp 165 and bp 1053 of the sequence published by Long et al. (1989; Fig. 2) was used as a specific 5′ probe for the *ZLAB1* gene. Three of the nine transformants were found to be disrupted at the *ZLAB1* locus. The Southern analysis of three targeted (GT1 to GT3) and two non-targeted transformants (NT1 and NT2) is shown in Fig. 3. The other six non-targeted transformants gave results similar to those observed for NT1 and NT2. The predicted structure of a targeted *ZLAB1* locus is shown in Fig. 2 together with the sizes of bands expected to hybridise with the 5′ probe following digestion with *Eco*RI or *Xba*I.

As expected for DNA of the non-transformed wild type, the 5′ probe detected a major signal at 2.5 kb after *Eco*RI digestion, and a signal at 3.7 kb after *Xba*I digestion. In the non-targeted transformed strains NT1 and NT2, the 5′ probe also hybridised to a 2.5-kb band following *Eco*RI digestion and to a 3.7-kb band following digestion with *Xba*I (Fig. 3). Although the 5′ probe was assumed to be specific for *ZLAB1*, two additional weak bands were detected following digestion of wild-type DNA with *Eco*RI (at 3.5 and 7.5 kb) or *Xba*I (at 3.6 and 4.8 kb). Similar-sized bands were also detected in all stable transformants analysed. These bands may result from weak cross-hybridisation between the probe and other members of the *Cab* multigene family.

The major bands at 2.5 kb in *Eco*RI- and 3.7 kb in *Xba*I-digested DNA of the non-transformed wild type were not observed in the three targeted transformants GT1 to GT3 (Fig. 3). Instead, as would be predicted following disruption of the *ZLAB1* locus (see Fig. 2), bands could be seen at 4.8 kb in DNA digested with *Eco*RI and at 6.3 kb in DNA digested with *Xba*I (Fig. 3). Using a probe covering part of the 5′ untranslated sequence, part of the coding region coding for the N-terminal region of the *ZLAB1*product, and part of the intron (positions −124 to +560 with respect to the translation start site, corresponding to bp 1153–1837 of the sequence published by Long et al. (1989), the expected 3′ junctions (a 1.5-kb *Eco*RI fragment and a 3.1-kb *Xba*I band, (see Fig. 2) were also detected in all three targeted transformants, but not in the wild-type or the two non-targeted transformants (data not shown). pGL2 DNA was used as a probe to demonstrate that there were several direct repeats of pGL*cab*-a integrated into the genome of both targeted and non-targeted strains (data not shown). The strain GT1 contained five integrated plasmid copies, while the strains GT2 and

GT3 each contained about 20 integrated plasmid copies. These values are within the range reported previously (Schaefer and Zrd 1997).

These results provide molecular evidence for the successful disruption of the *ZLAB1* locus in approximately 30% of the plants analysed.

PCR amplification and sequence analysis of novel DNA junctions in the targeted strains

PCR amplification and sequence analysis of novel DNA junctions was carried out to determine the specificity of gene disruption. Only homologous recombination can juxtapose the PCR primer pairs g1 + p2 or g2 + p1 (see Fig. 2). One primer of each pair hybridises with the targeting vector pGL*cab*-a (p1 or p2) and the other with the 5′ non-coding region (g1) or the 3′ untranslated region (g2) of the endogenous *ZLAB1* gene. The g1 and g2 sequences are not present in the targeting vector. Since the CAB proteins in *P. patens* are encoded by a multigene family, we designed primer g1 to anneal at position −918 (with respect to the translation start site) of the 5′ untranslated region of *ZLAB1* (Fig. 2A), so that only specific PCR products corresponding to the *ZLAB1* gene would be amplified.

We performed PCR analysis on the nine transgenic plants analysed by Southern hybridisation, using primer pair g1 + g2, to amplify the wild-type *ZLAB1* gene, g1 + p2 to amplify the 5′ junction region between the integrated plasmid pGL*cab*-a and the genomic *ZLAB1* gene, and g2 + p1 to amplify the 3′ junction region (see Fig. 2). PCR with the primer pair g1 + g2 resulted in the amplification of the expected original 2119-bp fragment from the non-transformed wild type (Fig. 4). A similar result was obtained for those integrative transformants identified by Southern analysis as not having been targeted to the *ZLAB1* locus (NT1 and NT2 in Fig. 3; NT3-NT6 data not shown), indicating that the *ZLAB1* gene had not been disrupted. Amplification of this fragment was not observed for the pGL*cab*-a transformants GT1, GT2 or GT3 (Fig. 4).

PCR with primer pairs g1 + p2 for the 5′ junction region and g2 + p1 for the 3′ junction region (see Fig. 2) showed that gene disruption had occurred in the transgenic strains GT1-GT3. PCR analysis of genomic DNA from these transformants using either primers g1 + p2 or g2 + p1 resulted, respectively, in the amplification of the expected 2062-bp or 1245-bp products (Fig. 4). Wild-type moss DNA yielded no PCR products with these primer pairs, nor did DNA from transgenic strains that were not targeted at the *ZLAB1* locus (Fig. 4; NT3-NT6, data not shown).

The PCR-amplified fragments obtained from DNA of the targeted pGL*cab*-a transformants using the primer pairs g1 + p2 and p1 + g2 were cloned and subsequently sequenced for at least 300 bases in each direction. Sequences observed in the 5′ and 3′ regions of these fragments were identical to those expected to arise after targeted insertion of the pGL*cab*-a construct into the *P. patens ZLAB1* locus, and the predicted DNA junctions were identified clearly by sequencing (Fig. 5). The 3′ junction region (positions 2150–2450; Long et al. 1989) contains the sequence of *ZLAB1* shown in Fig. 1 (positions 2146–2394; Long et al.1989) that includes the region cloned and sequenced for the other 13 members of the *Cab* multigene family described above.

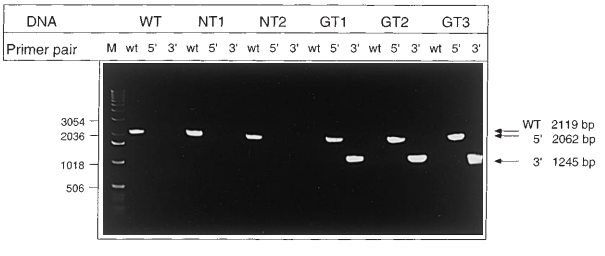


Fig. 4 PCR analysis of pGL*cab*-a transformants. The sizes of the PCR products are in agreement with the predicted sizes indicated on the right and with the predicted outcome of targeting depicted in detail in Fig. 2. DNA templates were from wild type (WT), the non-targeted transformants NT1 and NT2, and the targeted transformants GT1, GT2 and GT3. M, molecular weight marker. The primer pairs used were g1 + g2 (wt), g1 + p2 (5′), and p1 + g2 (3′). See Fig. 2 for locations

Discussion

The purpose of this study was to achieve the specific disruption of a member of the highly conserved *Cab* multigene family in *P. patens* with an insertion vector carrying a fragment of the targeted gene, *ZLAB1*, and to confirm the specificity of homologous recombination at the sequence level. For this purpose we have cloned and sequenced parts of 13 additional *Cab* genes. For these 13 genes, the region cloned, which lies within the coding sequence, shows 86–94% identity, at the DNA level, to *ZLAB1*. Such homology within *Cab* coding sequences has been well documented in the *Cab* multigene families of higher plants (Mcgrath et al. 1992).

In order to inactivate the chromosomal *ZLAB1* locus, we designed a construct containing a truncated *ZLAB1* fragment. Homologous recombination between this construct and the *ZLAB1* locus generates one copy of the *ZLAB1* gene which lacks the coding sequence for the 35 C-terminal amino acids, and a second copy of the *ZLAB1* gene copy lacking the promoter region and the coding sequence for the 10 N-terminal amino acids (see Fig. 2).

Three of the nine integrative transformants analysed in these experiments (GT1 to GT3) were disrupted at the *ZLAB1* locus. Sequence data from PCR-amplified DNA showed that, in targeted transformants, genomic DNA from outside the region of homology was contiguous with plasmid DNA, and confirmed successful gene disruption. In all targeted transgenics (GT1-GT3), the se-

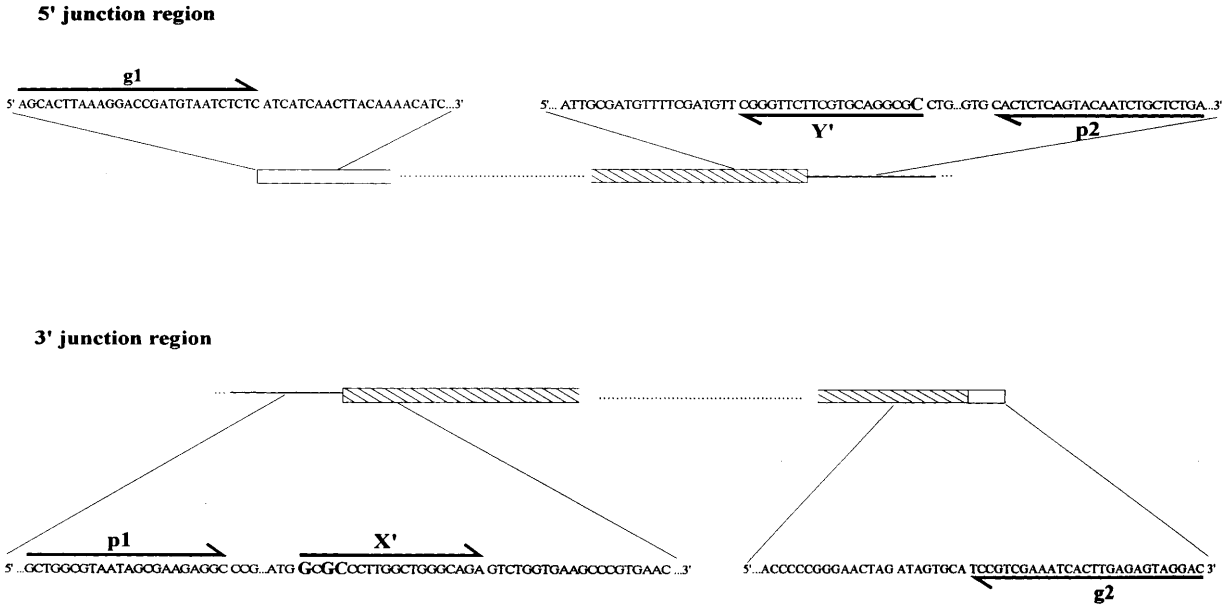


Fig. 5 Sequence analysis of the disrupted endogenous *ZLAB1* locus. The sequences of the 5′ and 3′ junction regions obtained after cloning of the corresponding fragments amplified from targeted transformants by PCR using g1 + p2 and p1 + g2 were identical to those expected following homologous recombination of the targeting pGL*cab*-a plasmid with the endogenous *ZLAB1* gene. Portions of the novel junctions arising from integration of the targeting vector pGL*cab*-a are shown in detail. Mutations were introduced into the oligonucleotide sequence in order to create a *Nar*I site for cloning (see Materials and methods), and these are indicated in *bold face*. X′ and Y′ denote the parts of primers X and Y, respectively, left after *Nar*I digestion. The locations of primers g1, g2, p1, and p2 are indicated. *Open boxes*, *P. patens* genomic DNA; *hatched boxes*, the truncated *Cab* sequence used in the targeting vector; *thin line*, pGL2 sequences

quence of the 3′ junction region (positions 2150 to 2450; Long et al. 1989) was identical to the corresponding sequence of *ZLAB1* and different from the other 13 Cab gene copies sequenced. (see Fig. 1). Thus, the specificity of homologous recombination could be established at the sequence level. We have not yet established whether the integrative transformants in which the *ZLAB1* locus was not targeted contain a plasmid insert in another member of the *Cab* multigene family. The weak bands detected by Southern analysis with the 5′ probe may correspond to other *Cab* genes which are closely related to *ZLAB1,* even in the promoter region. These bands were conserved in both targeted and non-targeted transformants.

The frequency of targeted events among integrative transformants following transformation with a vector which shares 1000 bp of sequence homology with the genome is therefore about 30%. This targeting efficiency is lower than, but comparable to, that reported for insertion vectors sharing more than 2 kb of sequence homology with the moss genome (Schaefer and Zrÿd 1997). However, a targeting efficiency of 30% using 1 kb of genomic sequence is considerably higher than that previously observed in higher plants (0.13%) (Kempin et al. 1997). The requirement for sequence homology for homologous recombination appears therefore to be stringent and comparable to that reported for mouse embryonic stem cells (te Riele et al. 1992).

We have detected no obvious phenotypic change in targeted transformants. This is not surprising in view of the complexity of the *Cab* multigene family. There are no data available on the level of expression of *ZLAB1* and other members of the family, and so their relative roles are as yet unknown. A more detailed phenotypic analysis of targeted transformants may however reveal more subtle changes in phenotype that we have not yet identified, and shed light on gene function within the *Cab* multigene family.

Acknowledgements We would like to thank Uta Marchfelder and Kerrin Huß for their contribution to this project. We also thank Francois Franceschi for critical reading of the manuscript. This work was funded by the Swiss OFES (Office fédéral pour l'éducation et la science) and by the European Union's BIOTECH programme as part of the EUROMOSS project, a component of the Project of Technological Priority 1993–1997. The experiments reported here comply with the current laws of Germany, United Kingdom and Switzerland, the countries in which the experiments were carried out.

References

Ashton NW, Cove DJ (1977) The isolation and preliminary characterisation of auxotrophic and analogue resistant mutants of the moss, *Physcomitrella patens*. Mol Gen Genet 154:87–95

Bilang R, Iida S, Peterhans A, Paszkowski J (1991) The 3′ terminal region of the hygromycin-B resistance gene is important for its activity in *Escherichia coli* and *Nicotiana tabacum*. Gene 100:247–250

Bollag RJ, Waldmann AS, Liskay RM (1989) Homologous recombination in mammalian cells. Annu Rev Genet 23:199–225

Capecchi MR (1989a) Altering the genome by homologous recombination. Science 244:1288–1292

Capecchi MR (1989b) The new mouse genetics: altering the genome by gene targeting. Trends Genet 5:70–76
Grimm C, Kohli J (1988) Observations on integrative transformation in *Schizosaccharomyces pombe*. Mol Gen Genet 215: 87–93
Halfter U, Morris PC, Willmitzer L (1992) Gene targeting in *Arabidopsis thaliana*. Mol Gen Genet 231:186–193
Kempin SA, Liljegren SJ, Block LM, Rounsley SD, Yanofsky MF (1997) Targeted disruption in *Arabidopsis*. Nature 389:802–803
Lee KY, Lund P, Lowe K, Dunsmuir P (1990) Homologous recombination in plant cells after *Agrobacterium*-mediated transformation. Plant Cell 2:415–425
Long Z, Wang S-Y, Nelson N (1989) Cloning and nucleotide sequence analysis of genes coding for the major chlorophyll-binding protein of the moss *Physcomitrella patens* and the halotolerant algae *Dunaliella salina*. Gene 76:299–312
Mcgrath JM, Terzaghi WB, Sridhar P, Cashmore AR, Pichersky E (1992) Sequence of the 4th and 5th photosystem II Type I chlorophyll a/b-binding protein genes of *Arabidopsis thaliana* and evidence for the presence of a full complement of the extended CAB gene family. Plant Mol Biol 19:725–733
Miao Z, Lam E (1995) Targeted disruption of the *TGA3* locus in *Arabidopsis thaliana*. Plant J 7:359–365
Offringa R, de Groot JA, Haagsman HJ, Does MP, van den Elzen PJM, Hooykaas PJJ (1990) Extrachromosomal recombination and gene targeting in plant cells after *Agrobacterium*-mediated transformation. EMBO J 9:3077–3084
Offringa R, Franke-van Dijk MEI, de Groot MJA, van den Elzen PJM, Hooykaas PJJ (1993) Non-reciprocal homologous recombination between *Agrobacterium* transferred DNA and a plant chromosomal locus. Proc Natl Acad Sci USA 90:7346–7350
Paszkowski J, Baur M, Bogucki A, Potrykus I (1988) Gene targeting in plants. EMBO J 7:4021–4026
Puchta H, Hohn B (1996) From centimorgans to base pairs: homologous recombination in plants. Trends Plant Sci 1:340–348
Puchta H, Dujon B, Hohn B (1996) Two different but related mechanisms are used in plants for the repair of genomic double strand breaks by homologous recombination. Proc Natl Acad Sci USA 93:5055–5060
Reski R, Faust M, Wang XH (1994) Genome analysis of the moss *Physcomitrella patens* (Hedw.) B.S.G. Mol Gen Genet 244:352–359
Riele H te, Maandag ER, Berns A (1992) Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. Proc Natl Acad Sci USA 89:5128–5132
Rossant J, Joyner A (1989) Toward a molecular-genetic analysis of mammalian development. Trends Genet 5:277–283
Roth D, Wilson J (1988) Illegitimate recombination in mammalian cells. In: Kucherlapati R, Smith GR (eds) Genetic recombination. American Society for Microbiology, Washington DC, pp 621–653
Rothstein R (1991) Targeting, disruption, replacement and allele rescue: integrative DNA transformation in yeast. Methods Enzymol. 194:281–301
Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual (2nd edn) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
Schaefer DG, Zrÿd J-P (1997) Efficient gene targeting in the moss *Physcomitrella patens*. Plant J 11:1195–1206
Schaefer DG, Bisztray G, Zrÿd JP (1994) Genetic transformation of the moss *Physcomitrella patens*. Biotech Agric For 29:349–364
Struhl K (1983). The new yeast genetics. Nature 305:391–397
Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW (1983) The double-strand-break repair model for recombination. Cell 33:25–35
Timberlake WE, Marshall MA (1989) Genetic engineering in filamentous fungi. Science 244:1313–1317