Chapter #IV

PRINCIPLES OF TARGETED MUTAGENESIS IN THE MOSS *PHYSCOMITRELLA PATENS*.

Didier Schaefer, Jean-Pierre Zrÿd
Laboratory of Plant Cell Genetics,Department of Plant Molecular Biology, University of Lausanne, CH 1015 Lausanne, Switzerland

Abstract: Highly efficient gene targeting is a unique feature of the moss *Physcomitrella patens.* It is now possible, in a multicellular eukaryote, to use the gene replacement technology which has been so successful in yeast. Careful design of transforming vector is nevertheless a key to successful generation of specific mutations.

Key words: Gene targeting, vector design, homologous recombination, allele replacement, protein tagging, *Physcomitrella patens.*

1. INTRODUCTION

Functional genomics is a major field of modern biology that aims to understand the relationship between biological systems and the presence and activity of genes; model systems such as yeast have provided the biological context to explore this relationship. The recent decipheration of whole genomes from an increasing number of organisms provides a complete description of the genetic information available to fulfil their biological functions. Two major strategies are conducted in parallel to unravel the function of genes and genetic networks: (1) high throughput computational-assisted global analyses of gene expression (transcriptomics, proteomics, phenomics, reviewed in (Holtorf et al., 2002) and (2) the generation and fine characterisation of mutants (Bouche and Bouchez, 2001). The former provides an extremely valuable description of whole genome activity in

response to determined experimental conditions whereas the latter enables a detailed physiological and genetic characterisation of the function of specific genes. Combining the data obtained with both approaches will provide an accurate understanding of the function of genes and genetic networks.

Over the last decades, transgene-mediated insertional mutagenesis has replaced traditional chemical- or radiation- induced mutagenesis since subsequent identification of the mutated gene is facilitated by the presence of the inserted tag (Jeon et al., 2000; Parinov and Sundaresan, 2000). The scientific knowledge gathered from the characterisation of tagged mutations identified in the large collections of insertional mutants available in the model plant *Arabidopsis thaliana* is meaningful in that view (Bouche and Bouchez, 2001). However insertional mutagenesis by stochastic transgenesis has several methodological limitations (for a discussion see Schaefer, 2002) which can be circumvented by the capacity to target transforming sequences to specific genomic locations, i.e. to achieve efficient gene targeting (GT). The principle of GT is based on the ability to address a transforming DNA carrying a homologous targeting sequence to its corresponding chromosomal locus where it integrates by homologous recombination (HR). Gene conversion events induced this way enable the generation of virtually any type of mutations ranging from the generation of null or weak alleles, to the fusion of molecular markers that allow the spatio-temporal expression of a gene to be followed in vivo, or to the addition of a molecular tags to the proteins to facilitate subsequent biochemical purification and characterisation.

GT is used routinely for functional studies in micro-organisms such as bacteria, yeast and filamentous fungi but has so far found limited application in multicellular eukaryotes since illegitimate integration of transgenes at random locations occurs orders of magnitude more frequently than targeted integration at homologous loci. The current model proposes that the ratio of targeted to random integration upon transgenesis correlates with the major pathway of double strand break (DSB) repair used by the cells: most DSB are repaired by a non-homologous end joining pathway in plants and animal cells whereas yeast repairs DSB essentially by a homologous recombination pathway. Yet the factors that determine this ratio are currently poorly understood and have been discussed extensively in several reviews (Mengiste and Paszkowski, 1999; Paques and Haber, 1999; Puchta, 2002; Hohe and Reski, 2003); this topic will not be discussed here. Mouse embryonic stem (ES) cells provide the only model system in animal biology proficient for GT approaches and this accounts for its exponential development over the last fifteen years (Müller, 1999). The moss *Physcomitrella patens* is the only plant in which functional genomic studies by targeted transgenesis is feasible today (Schaefer and Zrÿd, 1997;

Schaefer, 2001). We will review here the current knowledge of the features of GT in *Physcomitrella patens* and provides a description of the experimental strategies that can be used to achieve various types of mutagenesis illustrated with example from recently obtained data.

2. CONSIDERATIONS ON VECTOR STRUCTURE AND DESIGN

Strategies of targeted mutagenesis have been initially defined in budding yeast (Rothstein, 1991) and mouse ES cells (Müller, 1999) The structure of the construct plays a critical role for the transformation outcome, and the addition of a site specific recombination system such as Cre/lox (Sauer, 1993) within the vector enables subsequent modifications of the transgenic locus (for a recent review of Cre/lox applications in plants, see Ow, 2002). Strategies combining GT with site specific recombination provide the ultimate genetic tool for targeted mutagenesis (Müller, 1999) and form the conceptual framework for those described here for functional studies in *Physcomitrella patens*.

Two types of vectors are currently used for transformation: insertion and replacement vectors. Insertion vectors carry a sole homologous targeting sequence cloned next to the selectable marker (figure 1A). They insert in the target locus by a single homologous recombination event, generating an insertional mutation flanked by 2 copies of the targeting sequence. Cre-mediated removal of sequences flanked by LoxP sites will eliminate selectable markers and vector repeats but leave a duplicated target sequence in the genome. Insertion vectors clearly do not provide the optimal structure for the generation of subtle mutations since gene conversion is impossible. They are also not ideal for the generation of null alleles since the remaining duplicated target sequences present in the genome may enable the expression of truncated or mutated forms of the protein which can display transdominant properties (for a discussion of the problems observed in mouse ES cells, see Müller, 1999). The prevalence of the GT pathway upon *Physcomitrella* transformation was demonstrated with insertion vectors (Schaefer and Zrÿd, 1997).

In replacement vectors, the selectable marker is flanked by two homologous targeting sequences which do not have to be contiguous on the target locus but have to be in the same 5’ to 3’ orientation (figure 1B). Replacement vectors can integrate by two homologous recombination events

and thus generate true gene conversions. Cre-mediated removal of DNA sequences flanked by LoxP sites enables the elimination of selectable markers and vector repeats, leaving a single LoxP site between the two targeting sequences, i.e. a conversion of the target sequence with the in-vitro modified one. Replacement vectors can also insert in the target locus by a single homologous recombination event taking place within one of the targeting sequence; the considerations mentioned above for insertion vectors are also valid in this case. This type of vectors has been used to assess the main parameters of GT in *Physcomitrella* and their structure provides the basis for the experimental strategies described below.

3. FEATURES OF GENE TARGETING IN *PHYSCOMITRELLA*

3.1 The history of gene targeting

The first successful transformation of *Physcomitrella patens* was achieved by polyethylene-glycol (PEG) mediated direct DNA transfer into protoplasts (Schaefer et al., 1991). In the absence of sequence homology between the transforming DNA and the moss genome, relative transformation frequencies (RTF) are in average 1 in 100 000 regenerating colonies for supercoiled plasmids and ca. 5-10 times higher with linearised DNA (Schaefer, 1994). Molecular and genetic analyses demonstrate that transformation is achieved by the integration of several (2 to 30 in average) tandem repeats of the transforming plasmid usually at single location of the genome. These low RTF, compared to those observed in tobacco for example (0.1-1%), are not associated with a sub-optimal transformation protocol since transient expression of a GFP reporter reproducibly yields 5 – 25% labelled cells within 48 hours. PEG-mediated direct DNA transfer into protoplasts remains the method of choice to transform *Physcomitrella*.

The first evidence for the prevalence of the GT pathway in *Physcomitrella* were obtained when several moss strains transformed with a plasmid conferring hygromycin resistance were retransformed with a second plasmid conferring kanamycin resistance (Schaefer, 1994). In this situation sequence homology was provided by the plasmid backbone common to both vectors (insertion type). We observed that RTF were in average one order of magnitude higher in transformed strains than in wild type, and that the antibiotic resistance markers were genetically linked in most double resistant

strains obtained by sequential transformation. Integration of tandem repeats of the second plasmid by HR with the sequences integrated in the chromosome was demonstrated at the molecular level subsequently (Schaefer and Reynolds, unpublished data). At the same time, co-segregation of resistance markers involved in sequential transformation experiments was also reported independently (Kammerer and Cove, 1996). Definitive evidence for the dominance of the GT pathway in *Physcomitrella* were provided by experiments designed to target 3 independent unique sequences of the moss genome (Schaefer and Zrÿd, 1997). This seminal work demonstrates that in up to 90% of the case, insertion vectors sharing 2.3 to 3.5 kb of sequence homology with the moss genome integrates in the target locus by HR. RTF observed in these experiments raised up to 1 in 1000 regenerating colonies and targeted transformation was associated with the insertion of direct tandem repeats of the vector. This work opened the way to the direct generation of knock-outs in *Physcomitrella* (for recent reviews see Puchta, 1998; Schaefer and Zrÿd, 2001; Puchta, 2002; Schaefer, 2002; Hohe and Reski, 2003).

3.2 The parameters influencing gene targeting

Yet the development of GT technology requires further methodological refinements to achieve gene conversions required for fine mutagenesis. The following parameters have to be assessed: (1) the relation between the extent of sequence homology on the transforming DNA and GT efficiency, (2) the patterns of targeted integration following transformation with replacement vectors and (3) the suitability of a site specific recombination system such as Cre/lox to remove plasmid repeats and recycle selectable markers. In collaboration with the group of Michel Laloue (INRA Versailles), we addressed these questions in a pilot study on the adenine phosphoribosyl transferase (aprt) locus of *Physcomitrella* (manuscript in preparation). This gene encodes for an enzyme of the purine salvage pathway that recycles adenine into AMP and provides a suitable target for GT studies since its loss of function confers resistance to the adenine analogue 2,6-diaminopurine (DAP). It is unique in the moss genome and is encoded at the genomic level by 7 exons with an average size of 100 bp.

The first two parameters were assessed with 3 different replacement vectors. In the first one, a positive selectable marker was inserted in the middle of the ca. 900 bp aprt cDNA. This vector was thus displaying discontinuous sequence homology with the target locus. The two other vectors were carrying 1.2 or 2.9 kb of aprt genomic sequence in the middle

of which part of the aprt coding sequence was deleted (including a specific restriction site used for subsequent molecular analyses). The relation between GT efficiencies and the extent of sequence homology could be established as follows (Figure 2a). (1) GT is possible with cDNA based replacement vectors but not very efficient. Nevertheless HR events occurring within stretches of 51 and 187 bp of continuous homology have been observed. These values probably represent the lowest amount of sequence homology required for GT in *Physcomitrella*. (2) RTF (around 1 in 1000 regenerating colonies) and targeting efficiencies (95 %) were high and similar with both genomic DNA based vectors indicating that 2 stretches of 600 bp or more are sufficient to saturate GT in *Physcomitrella*. (3) Linearisation of replacement vectors at the edges of the targeting sequences are required for optimal targeting frequencies. These values are comparable to those reported in recent independent gene disruption studies (Hohe and Reski, 2003). Such a targeting efficiency can only be compared with that observed in budding yeast and signifies that *Physcomitrella* ranks number one for the efficiency of GT among multicellular eukaryotes (Schaefer, 2001).

A detailed molecular analysis of these strains revealed a strikingly predictable pattern of DNA integration. Insertion in (1 HR) and conversion of (2 HR) the aprt locus were observed at similar frequencies with the 3 replacement vectors tested. The pattern of insertion events was characterised by the presence of tandem direct repeats (2-30) of the whole vector as previously observed with insertion vectors. Replacement of the resident locus with a single copy of the replacement cassette accounted for 10-20% of the conversion events, and we have been able to generate specific deletions of 150 and 1100 bp in the aprt locus this way. The presence of tandem direct repeats of the whole plasmid and/or of the replacement cassette only was observed in the other cases (Figure 2b). The 1:1 ratio of insertion versus conversion events observed with replacement vectors was confirmed in a recent study conducted at the University of Leeds (Y. Kamisugi, pers. com.) whereas the integration pattern is consistent with that described in recent independent gene disruption studies. This pattern differs from that observed in budding yeast where single copy replacement is the major class of events, but resemble that previously described for GT events observed in plants (Puchta, 2002) and animal CHO cells (Adair et al., 1998).

3.3 Getting clean gene replacement

The frequent insertion of replacement vectors requires the sorting of conversion events in the transformed population. This could in principle be

easily achieved using a positive-negative screening procedure that counter select insertion events, but the lack of a suitable negative selectable marker for *Physcomitrella* does not allow such approach. Therefore PCR analysis must be performed with primers specific for the recombined junction to identify gene conversion events. The presence of the selectable marker and of multiple copies of the transforming DNA is another major obstacle to the generation of fine mutations. The highly predictable integration pattern enables this situation to be solved using the site specific recombination Cre/lox system. In this system, the Cre recombinase is able to excise or invert DNA sequences flanked by two 34 bp asymmetrical LoxP sites positioned in direct or inverted orientation, respectively (Sauer, 1993). The Cre/lox system is extensively used in transgenesis to recycle selectable markers, to eliminate plasmid repeats or to generate conditional mutations using Cre expression cassettes driven by inducible promoters. Yet the possible presence of cryptic LoxP sites in genomes may lead to additional genetic rearrangements and the use of the Cre recombinase in transient expression assay is preferable.

Among the replacement vectors tested on the aprt locus, one of them was carrying a neo selectable marker flanked by 2 LoxP sites in direct orientation cloned in the middle of 2.9 kb of aprt genomic sequences. Transgenic strains obtained with this construct were used to assess the efficiency of Cre-mediated site specific recombination in *Physcomitrella* (Chakhparonian, 2001; Schaefer et al., manuscript in preparation). Transient expression of a constitutive Cre expression cassette in protoplasts of seven strains carrying 3 to 15 direct repeats of the vector was performed. Protoplasts were regenerated at low density on non-selective medium for 2 weeks at what time fragments of single protoplast derived colonies were replica-plated on selective and non-selective medium. Loss of antibiotic resistance was scored after an additional week of growth and the replicates growing on non selective medium were used for subsequent amplification and molecular analyses. In this time frame, PCR and Southern blot data can be obtained within 4 and 6 weeks after transient expression, respectively. A typical example of replica plate and PCR analyses is presented in figure 2c.

For all the strains tested, phenotypic and molecular analyses provided evidence for Cre-mediated elimination of the selectable marker in 5 to 25% of single protoplast derived colonies. Resolution of multiple copies in simple insertion or conversion events was observed in ca. 25% antibiotic sensitive colonies whereas one or 2 copies of the plasmid without the selectable cassette could be identified in the others, suggesting that the Cre recombinase excises marker repeats sequentially. Our data also indicated that the proportion of simple integration patterns recovered after transient

Cre expression was higher in strains carrying an initial lower plasmid copy number and the choice of a transformants with few integrated repeats is recommended to optimise the probability of getting clean conversions. Nevertheless these data demonstrate that Cre-mediated recycling of selectable marker and elimination of plasmid repeats is very efficient in *Physcomitrella* and that it enables the generation of clean conversions in the moss genome in a very short time scale. It also permits the generation of multiple mutations in the same strain by sequential transformation since the elimination of plasmid repeats and selectable markers remove putative target sites for the second round of transformation and allow the same selectable marker to be used several times. Therefore LoxP sites must be integrated in the design of transformation vectors for *Physcomitrella* and we have recently developed in collaboration with Fabien Nogué (INRA, Versailles) vectors carrying multiple cloning sites and selectable cassette flanked by LoxP sites that are freely available upon request. With these tools in hands, the most sophisticated strategies of functional genomic by targeted mutagenesis can be applied to *Physcomitrella*.

4. STRATEGIES OF TARGETED MUTAGENESIS

4.1 Sequence requirement for a complete functional study of a specific gene

Major progresses have been accomplished over the last years in the amount of genetic information available for *Physcomitrella*. Two publicly funded ESTs sequencing programs have provided the scientific community with a remarkable public database covering more than 100000 ESTs representing ca. 16000 genes (Nishiyama et al., 2003). In parallel, another database covering 100000 ESTs was also generated in a program associating the University of Freiburg with the private company BASF Plant Sciences (Rensing et al., 2002). A detailed analysis of the function of a gene requires a complete knowledge and access to its sequence including the flanking genomic sequences. Since the whole genome of *Physcomitrella* is not yet sequenced, PCR based strategies have to be conducted starting from the available EST information to isolate the corresponding genomic sequence. Starting from a partial EST, 3’ and 5’ RACE with nested primers on cDNA libraries, mRNA or by RT-PCR enables the isolation of the corresponding full length cDNA. Since the exon intron structure of genes is frequently

conserved between *Physcomitrella* and *Arabidopsis* or rice, the putative genomic structure of the moss gene can be deciphered from sequence comparison with the structure of the angiosperm's homologue. Amplification of the genomic sequence corresponding to the full length cDNA can easily be achieved by PCR on moss genomic DNA but care should be taken on primer design to avoid that they match with putative exon junctions. Finally the isolation of flanking genomic sequences can be achieved by tailed PCR or RAGE PCR (Cormack and Somssich, 1997) using nested primers based on the previously obtained genomic information. With nowadays PCR technologies, this strategy is faster than the classical screen of cDNA and genomic libraries and enables the isolation of a complete genomic sequence within 6 to 8 weeks. And the isolation of a complete genomic locus is necessary for a thorough functional genomic study of your favourite gene.

4.2 The generation of mutations

A comprehensive study of gene function must entail the generation and characterisation of different types of mutants. Null alleles will provide information on complete loss of function, weak alleles will enable the identification of functional domains or critical amino-acids within the protein, whereas the expression pattern of the gene can be studied in vivo in strains carrying in its natural chromosomal environment translational fusion with suitable reporters. We describe below the different GT strategies to adopt to generate these different types of mutations (Figure 3).

The best way to generate null phenotype is upon deletion of the entire coding sequence. This guarantees the absence of truncated or mutant forms of the protein that may acquire unexpected properties such as transdominant interactions. It is the strategy that has been adopted in the *Saccharomyces* Genome Deletion Project that aims to generate null alleles for every single gene identified in the yeast genome (Giaever et al., 2002). To delete the entire coding sequence of a gene in *Physcomitrella*, the replacement vector should carry a LoxP selectable cassette inserted between the 5' and the 3' flanking genomic sequences. Cre-mediated removal of selectable marker and repeats will leave a single LoxP site between the 5' and 3' flanking sequence.

The generation of weak alleles is not a just so story; it is difficult to make reliable predictions on the best strategy to adopt. Modifications of the regulatory elements of a gene may alter its expression, and we observed this situation in the experiments conducted on the aprt locus. In this work, we

have obtained null and weak aprt alleles, and characterisation of the weak phenotypes revealed in each case that it was associated with the presence of an entire aprt coding sequence flanked by a modified 3' UTS. We hypothesised that the stability of the mRNA could be impaired and thus lead to a reduced translation level which accounts for the weak phenotype.

We describe below 2 strategies to generate weak alleles: the first one will lead to the formation of truncated forms of the protein and is suitable for the functional dissection of the role of different domains within a protein; the second one is probably of more general application and enables the generation of point mutations (figure 3). To achieve C-terminal protein truncation, the replacement vector should carry (from the 5' to the 3' end) (a) the N-terminal gene sequence to which an in frame Stop codon has been added at the 3' end, (b) the LoxP selectable cassette and (c) the 3' UTS. The mutagenised locus will carry, after Cre recombination, the truncated gene driven by its natural regulatory elements. To achieve N-terminal truncations, the selectable marker should be flanked in 5' by the 5' flanking genomic sequences and in 3' by the C-terminal coding sequence to which an in frame ATG has been added at the 5'end. Protein truncation can also be obtained with insertion vectors; we have recently observed a weak phenotype that was associated with C-terminal truncation of a structural cytoskeleton gene (Pegoraro and Schaefer, unpublished results). Protein truncation thus provides a convenient way to generate weak alleles.

The generation of point mutations offers the most sophisticated approach for unravelling gene function as it permits the modification of any specific amino acid within a protein. The strategy proposed here is similar to that currently used in mouse ES cells. The replacement vector carries the LoxP selectable marker cloned in an intron and flanked by the adjacent genomic sequences in which the desired mutation has been generated in vitro. The addition of a restriction fragment polymorphic trait (i.e. elimination or creation of a restriction site) to the mutated codon is recommended since it facilitates subsequent identification of the mutation. Conversion of the wild type sequence with the in vitro mutated one and elimination of selectable marker and plasmid repeats by the Cre recombinase will leave a single LoxP site in the intron. In yeast, the transfer of genetic information during HR between a transforming DNA and the chromosomal sequence is reciprocal and occurs statistically at equal frequencies. This means that ca 50 % of GT events are associated with the transfer of the mutation to the resident chromosomal information. These features are probably similar in *Physcomitrella*: this implies that the mutated strains have to be identified in the transformed population. PCR amplification of the mutagenised sequence followed by restriction analysis or DNA sequencing easily allows the identification of mutated strains. In the experiments conducted on the aprt

locus, successful transfer of genetic information (in this case deletion) from the transforming DNA to the genome was reproducibly observed which provides experimental support for the feasibility of such strategy in *Physcomitrella*.

Finally, efficient GT in *Physcomitrella* allows the direct creation of translational fusion of the gene under study with reporter genes (GUS, GFP; etc..) or protein tags (6 x His, TAP-tag, HA, c-Myc, etc…); this facilitates both the study of the expression profile of the protein, and its purification. The choice between N-terminal or C-terminal translational fusion cannot be predicted and this question must be addressed experimentally for each situation. The presence of additional protein sequence to the gene may lead to unexpected side effects and a careful phenotypic analysis of the strains must be performed to validate the approach. The structure of replacement vectors designed to generate N-ter or C-ter translational fusion of moss genes with tags is depicted in figure 3. This type of approach has been successfully achieved in several studies conducted in the laboratory of M. Hasebe in Okasaki (Sakakibara et al., 2003, and personal communication).

5. CONCLUSION

The remarkable GT efficiency observed in *Physcomitrella patens* provides a unique opportunity to apply the most sophisticated genetic tool, targeted mutagenesis, to investigate the function of plant genes. The data reviewed here clearly demonstrates that it is possible to day to combine GT with site specific recombination in order to modify accurately any sequences in the moss genome. We have described several GT strategies that can be applied in this moss for the generation of different types of mutations. This description is not exhaustive and the limit of such approaches probably resides within the limit of the investigator’s imagination. It is a fact that *Physcomitrella* ranks number one for the efficiency of GT among eukaryotes model systems and this moss will be used as such for the detailed study of gene function.

Figure legends

Figure 1

Basic features of insertion and replacement vectors.

A) An insertion vector carries a single targeting sequence (TS boxes) cloned next to a selectable cassette (R box) flanked by 2 LoxP sites (stars). Targeted insertion is mediated by a single HR reaction (hatched bar). Cre-mediated deletions are indicated by the hatched lines. The structure of open reading frames and truncated protein after targeted insertion are illustrated with arrows, with the circle representing the ATG codon and the arrowhead the Stop codon. Note that the ATG or STOP codon must be absent from the targeting sequence since their occurrence will lead to the formation of a complete ORF with modified 5' or 3' untranslated sequences at one junction.

B) A replacement vector carries the selection cassette in between 2 targeting sequence. Conversion events are mediated by 2 HR reactions and the subsequent removal of plasmid repeats leaves a single LoxP site in between targeting sequences.

Figure 2

Parameter of GT in *Physcomitrella*

A) Transformation frequencies (RTF) and targeting efficiencies observed with different replacement vectors on the aprt locus of *Physcomitrella*. RTF observed with supercoiled and linear plasmids correspond to the black and dotted bars, respectively.

B) Southern blot analysis of DNA isolated from representative strains obtained following conversion (C) or insertion (In 3' and In 5') of the replacement vector in the aprt locus. DNA was digested with EcoRV and hybridised with a complete aprt sequence. The 2 WT bands (3.7 and 4.0 kb) are converted into a single band (7.5 kb) upon gene conversion, demonstrating that the deletion of 200 bp generated in the transforming DNA and containing the internal EcoRV site was effectively transferred into the genome. Only one of the WT band is shifted to high molecular weight upon targeted insertion demonstrating that targeted insertion was mediated by a single HR event occurring within one of the 2 WT bands. The HMW signal observed here corresponds to plasmid repeats which do not contain Eco RV restriction sites.

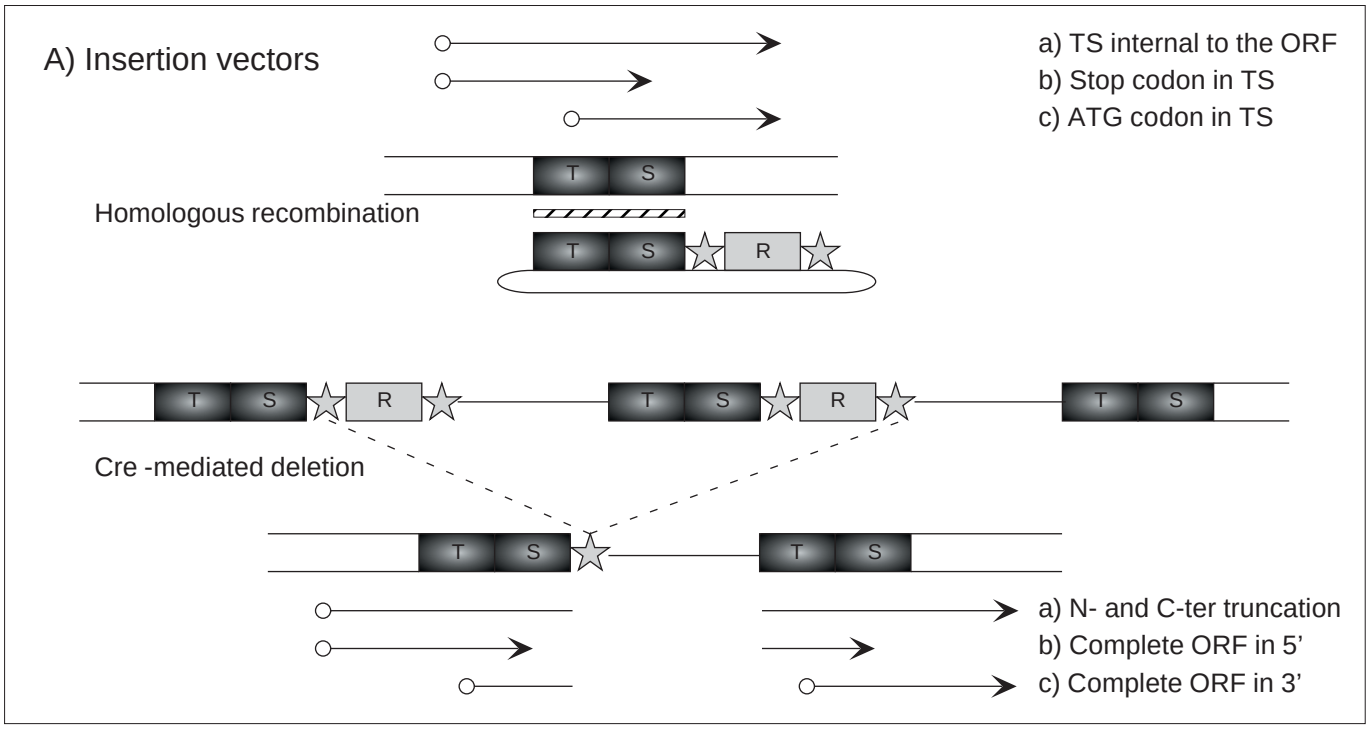
C) Typical data set of the analysis of single protoplast derived colonies after transient expression of the Cre recombinase. The strains that have lost antibiotic resistance (S) upon Cre expression are circled on the selective

plate. PCR analyses reveal that the resistant strains (R) carry the 2.3 kb band characteristic of the AP-Lox neo Lox–RT fragment whereas sensitive carry a new fragment of 0.7 kb which corresponds to the AP-Lox-RT fragment.

Figure 3

Schematic structure of disruption vectors designed to generate different types of mutation in the moss genome. The wild-type locus of Your Favourite Gene (YFG) is illustrated on top with 5' and 3' flanking sequences represented with thin grey bars, exons with the Y, F and G boxes and introns with thin lines. Point mutation in the F box is marked with an ∗ on top of the box, the other symbols are identical to those used in figure 1. Strategies for the generation of null allele by ORF deletion, of week allele by protein truncation or point mutation and of tagged proteins are shown.

Figure 1


A) Insertion vectors
a) TS internal to the ORF
b) Stop codon in TS
c) ATG codon in TS
T
S
Homologous recombination
T
S
R
T
S
R
T
S
R
T
S
Cre -mediated deletion
T
S
T
S
a) N- and C-ter truncation
b) Complete ORF in 5'
c) Complete ORF in 3'

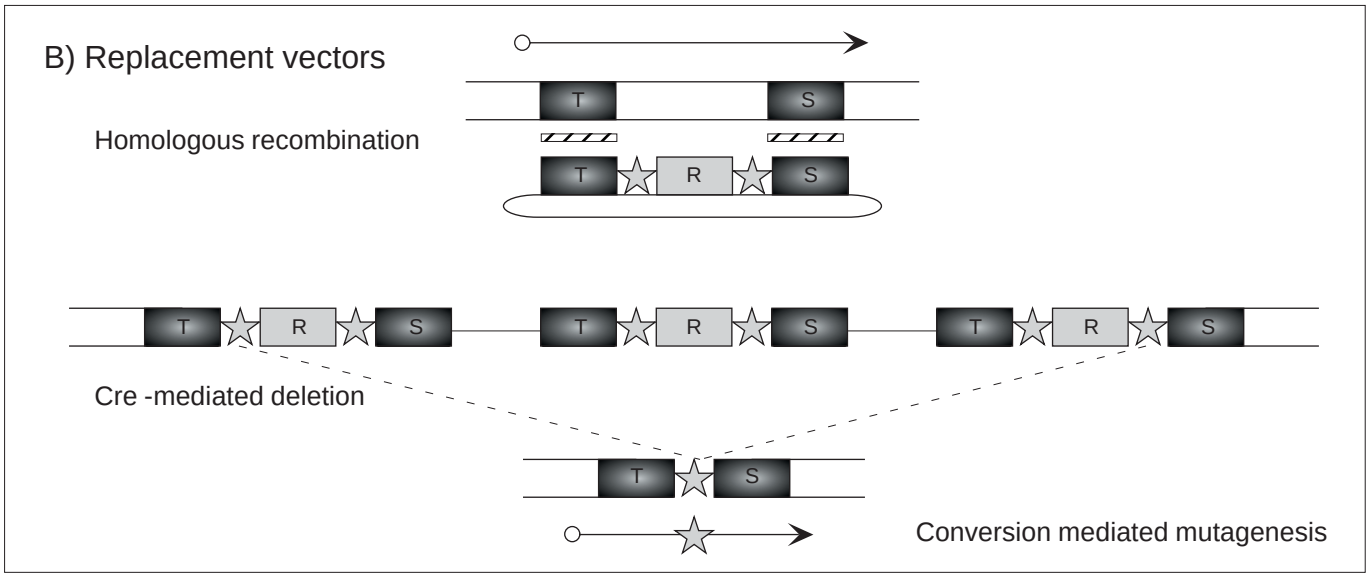
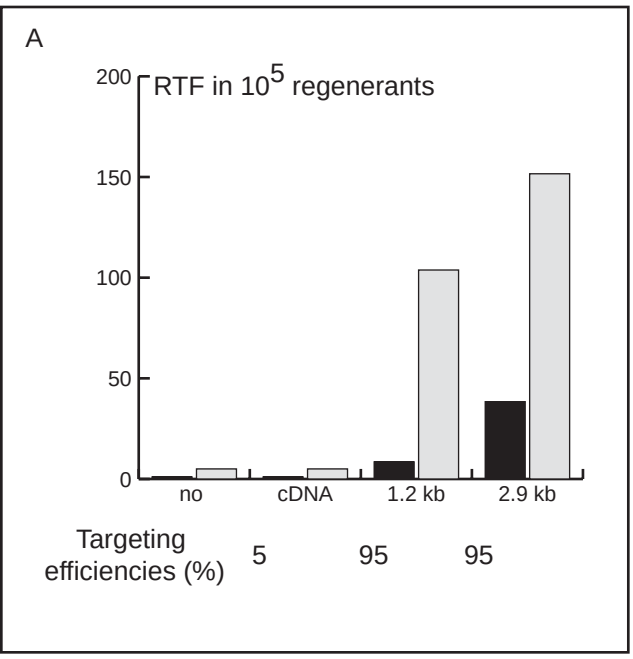
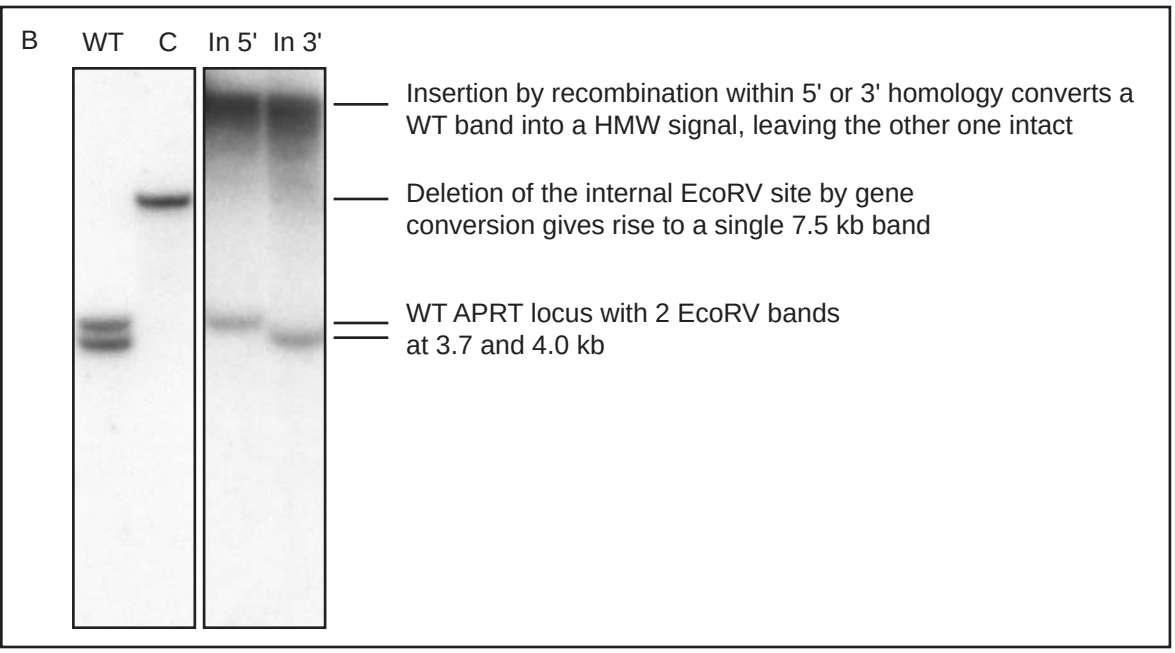

B) Replacement vectors
T
S
Homologous recombination
T
R
S
T
R
S
T
R
S
T
R
S
Cre -mediated deletion
T
S
Conversion mediated mutagenesis

Figure 2


A
RTF in 10^5 regenerants
200
150
100
50
0
no
cDNA
1.2 kb
2.9 kb
Targeting efficiencies (%)
5
95
95


B
WT
C
In 5'
In 3'
Insertion by recombination within 5' or 3' homology converts a WT band into a HMW signal, leaving the other one intact
Deletion of the internal EcoRV site by gene conversion gives rise to a single 7.5 kb band
WT APRT locus with 2 EcoRV bands at 3.7 and 4.0 kb

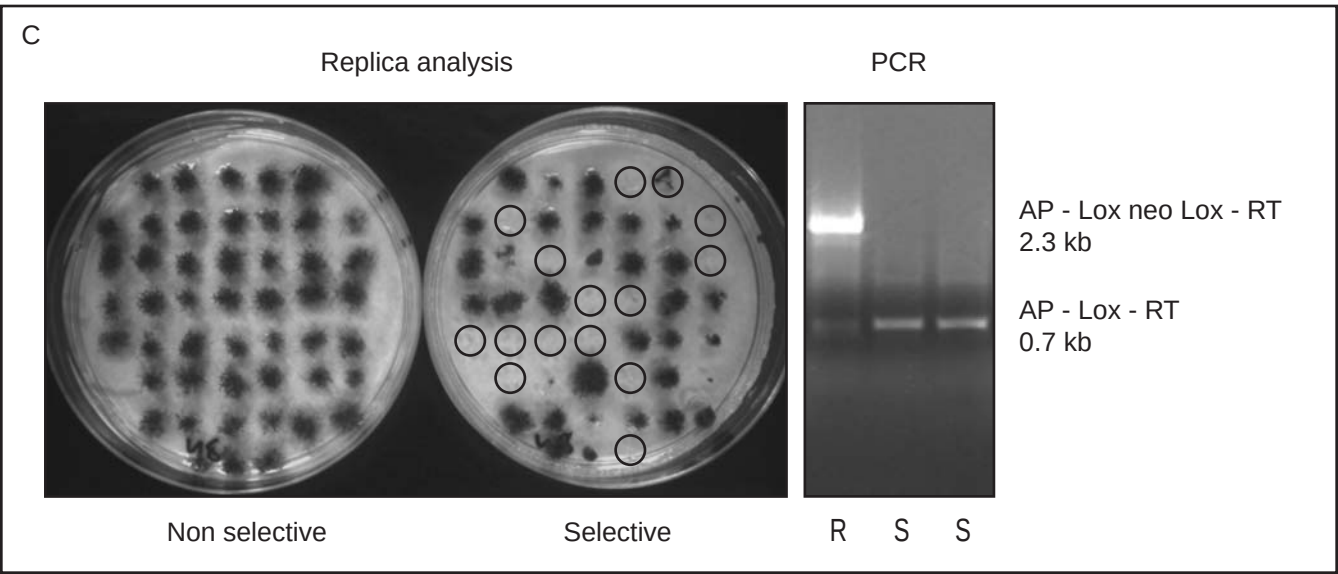
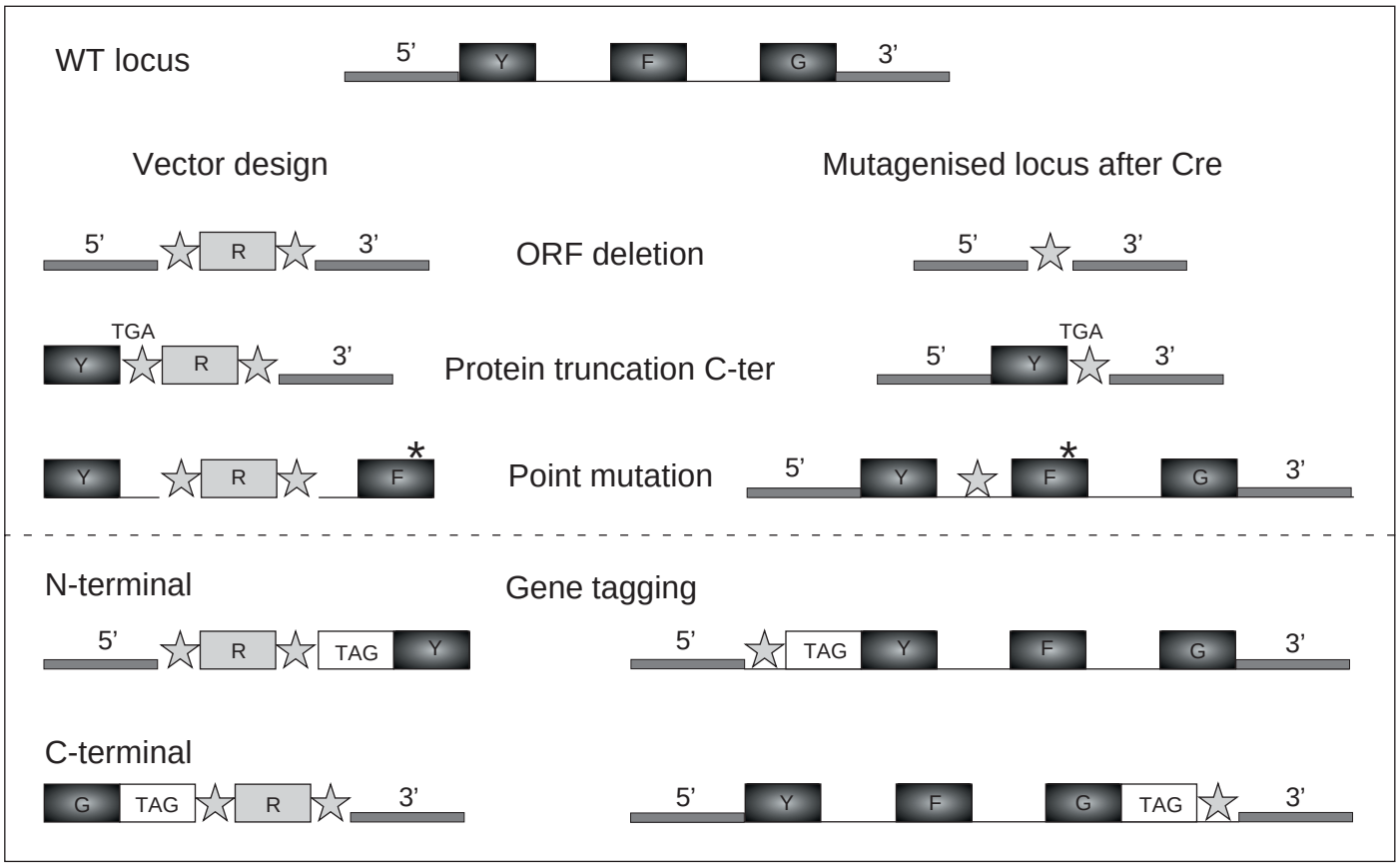

C
Replica analysis
PCR
AP - Lox neo Lox - RT
2.3 kb
AP - Lox - RT
0.7 kb
Non selective
Selective
R
S
S

Figure 3


WT locus
5’
Y
F
G
3’
Vector design
Mutagenised locus after Cre
5’
R
3’
ORF deletion
5’
3’
TGA
Y
R
3’
Protein truncation C-ter
TGA
5’
Y
3’
Y
R
F
*
Point mutation
5’
Y
F
*
G
3’
N-terminal
Gene tagging
5’
R
TAG
Y
5’
TAG
Y
F
G
3’
C-terminal
G
TAG
R
3’
5’
Y
F
G
TAG
3’

Bibliography

Adair GM, Scheerer JB, Brotherman A, McConville S, Wilson JH, Nairn RS (1998) Targeted recombination at the Chinese hamster APRT locus using insertion versus replacement vectors. Somatic Cell & Molecular Genetics 24: 91-105

Bouche N, Bouchez D (2001) *Arabidopsis* gene knockout: phenotypes wanted. Curr Opin Plant Biol 4: 111-117

Chakhparonian M (2001) Développement d'outils de la mutagenèse ciblée par recombinaison homologue chez *Physcomitrella patens*. PhD thesis. Université de Lausanne,Switzerland "http://www.unil.ch/lpc/docs/theses/Thesemicha.pdf"

Cormack RS, Somssich IE (1997) Rapid amplification of genomic ends (RAGE) as a simple method to clone flanking genomic DNA. Gene 194: 273-276

Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K, Deutschbauer A, Entian KD, Flaherty P, Foury F, Garfinkel DJ, Gerstein M, Gotte D, Guldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kotter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang CY, Ward TR, Wilhelmy J, Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M, Philippsen P, Davis RW, Johnston M (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. Nature 418: 387-391

Hohe A, Reski R (2003) A tool for understanding homologous recombination in plants. Plant Cell Rep 21: 1135-1142

Holtorf H, Guitton MC, Reski R (2002) Plant functional genomics. Naturwissenschaften 89: 235-249

Jeon JS, Lee S, Jung KH, Jun SH, Jeong DH, Lee J, Kim C, Jang S, Yang K, Nam J, An K, Han MJ, Sung RJ, Choi HS, Yu JH, Choi JH, Cho SY, Cha SS, Kim SI, An G (2000) T-DNA insertional mutagenesis for functional genomics in rice. Plant Journal 22: 561-570

Kammerer W, Cove Dj (1996) Genetic analysis of the result of re-transformation of transgenic lines of the moss, *Physcomitrella patens*. Mol. Gen. Genet. 250: 380 - 382

Mengiste T, Paszkowski J (1999) Prospects for the precise engineering of plant genomes by homologous recombination [Review]. Biological Chemistry 380: 749-758

Müller U (1999) Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis [Review]. Mechanisms of Development 82: 3-21

Nishiyama T, Fujita T, Shin IT, Seki M, Nishide H, Uchiyama I, Kamiya A, Carninci P, Hayashizaki Y, Shinozaki K, Kohara Y, Hasebe M (2003) Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: implication for land plant evolution. Proc Natl Acad Sci U S A 100: 8007-8012

Ow DW (2002) Recombinase-directed plant transformation for the post genomic era. Plant Molecular Biology 48: 183 - 200

Paques F, Haber JE (1999) Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae* [Review]. Microbiology & Molecular Biology Reviews 63: 349-404

Parinov S, Sundaresan V (2000) Functional genomics in *Arabidopsis*: large-scale insertional mutagenesis complements the genome sequencing project [Review]. Current Opinion in Biotechnology 11: 157-161

Puchta H (1998) Towards targeted transformation in plants. Trends in Plant Science 3: 77-78

Puchta H (2002) Gene replacement by homologous recombination in plants. Plant Mol Biol 48: 173-182

Rensing SA, Rombauts S, Van de Peer Y, Reski R (2002) Moss transcriptome and beyond. Trends Plant Sci 7: 535-538

Rothstein R (1991) Targeting, disruption, replacement and allele rescue: integrative DNA transformation in yeast. Methods Enzymol. 194: 281 - 301

Sakakibara K, Nishiyama T, Sumikawa N, Kofuji R, Murata T, Hasebe M (2003) Involvement of auxin and a homeodomain-leucine zipper I gene in rhizoid development of the moss *Physcomitrella patens*. Development 130: 4835-4846

Sauer B (1993) Manipulation of the transgene by site-specific recombination: use of cre recombinase. Methods Enzymol. 225: 890-900

Schaefer DG (1994) Molecular genetic approaches to the biology of the moss *Physcomitrella patens*. PhD Thesis. University of Lausanne, Switzerland "http://www.unil.ch/lpc/docs/DSThesis.htm"

Schaefer DG (2001) Gene targeting in *Physcomitrella patens*. Current Opinion in Plant Biology 4: 143-150

Schaefer DG (2002) A new moss genetics: targeted mutagenesis in *Physcomitrella patens*. Annu Rev Plant Biol 53: 477-501

Schaefer DG, Zrÿd J-P (1997) Efficient gene targeting in the moss *Physcomitrella patens*. The Plant Journal 11: 1195-1206

Schaefer DG, Zrÿd J-P (2001) The moss *Physcomitrella patens* now and then. Plant Physiology 127: 1430-1438

Schaefer DG, Zrÿd J-P, Knight CD, Cove DJ (1991) Stable transformation of the moss *Physcomitrella patens*. Mol. Gen. Genet. 226: 418-424