Gene targeting in *Physcomitrella patens*

Didier G Schaefer

Gene-targeting efficiency in the land plant *Physcomitrella patens* (Bryophyta) can only be compared with that observed in *Saccharomyces cerevisiae*. Sequencing programs and microbiological molecular genetic approaches are now being developed to unravel the precise function of plant genes. *Physcomitrella patens*, as the new 'green yeast', might well become a major tool for functional genomic studies of multicellular eukaryotes.

Addresses
Institut d'écologie, Laboratoire de Phytogénétique Cellulaire, Bâtiment de Biologie, Université de Lausanne, CH-1015 Lausanne, Switzerland; e-mail: didier.schaefer@ie-pc.unil.ch

Current Opinion in Plant Biology 2001, **4**:143–150

1369-5266/01/$ – see front matter

Abbreviations

DAP	2,6-diaminopurine
DSB	double-strand breaks
ES	embryonic stem
EST	expressed sequence tag
gDNA	genomic DNA
GT	gene-targeting
GUS	β-glucuronidase
HR	homologous recombination
IR	illegitimate recombination
NHEJ	non-homologous end joining
RAD-52	radiation-sensitive-52
RTF	relative transformation frequencies

Introduction

The extensive DNA sequence data sets that have been made available to the scientific community provide an outstanding tool for the investigation of the precise function of genes. The next decade will be devoted to functional genomics and several approaches are currently being applied to unravel gene function in plants [1]. Gene expression profiles in response to different experimental conditions can be established using DNA microarrays [2,3] or by the analysis of populations of proteins separated by two-dimensional gel electrophoresis (i.e. through proteomics) [4]. Transgenesis is also extensively used for functional genomics. Along with overexpression studies, different insertional mutagenesis strategies including gene tagging [5•,6•,7] and gene trapping [8•] are being applied to identify the gene responsible for a detectable phenotype or to characterise the function of a specific gene.

In order to investigate gene function precisely, however, one needs to be able to generate specific mutations in a given gene and this is only possible when an efficient gene-targeting (GT) procedure is used. In GT experiments, the transforming DNA carries a mutated allele of the gene of interest along with a selectable marker. The homologous fragment targets the transforming DNA to the corresponding chromosomal locus where it integrates by homologous recombination (HR), resulting in either gene knock-out by insertion or point mutations by gene conversion (Figure 1).

Although GT is the standard method for the study of functional genetics in microbiological systems such as bacteria, yeast and some filamentous fungi, it can be applied to higher eukaryotic model systems only rarely. This is because the frequency of integration at random locations by illegitimate recombination (IR) in the higher eukaryotes is several orders of magnitude higher than that of targeted integration by HR (Table 1). The high ratio of targeted to random integration events (i.e. 0.1–1%) in murine embryonic stem (ES) cells relative to that in other mammalian systems accounts for the extraordinary development of the mouse ES cell system for functional genomics in mammalian biology over the past 10 years [9,10]. In plants, despite the fact that the feasibility of GT was demonstrated more than 10 years ago [11], the ratio of targeted to random integration events observed remains too low to enable systematic GT approaches [12,13••,14••] (Table *2*).

The moss *Physcomitrella patens* (Funariales, Bryophyta) is being developed as a model system for plant biological studies. The dominance of the haploid gametophyte in the life cycle of this moss facilitates genetic analysis. Both apical and caulinary growth can be studied at the single-cell level during gametophytic development. The metabolism and development of *P. patens* are controlled by the same growth substances (i.e. auxin, cytokinin and abscissic acid) and environmental signals as those acting in other land plants. However genetic transformation studies have revealed that, unlike in angiosperms, GT is the predominant pathway of genetic transformation in *P. patens*. In this review, we discuss the implications of this finding for functional genomic studies in plant biology.

Efficient GT is correlated with the dominant pathway of DNA double strand break repair

Genetic and biochemical studies conducted in *S. cerevisiae* provide a general model for the high efficiency of GT in yeast [15,16••]. Integration of foreign DNA sequences in the genome by HR or IR seems to be tightly correlated with the dominant pathway used by the cells to repair DNA double-strand breaks (DSB). In *S. cerevisiae*, DSB repair by HR is the predominant mechanism and the genes involved in this repair pathway have been assigned to the RAD-52 (radiation-sensitive-52) epistasis group. In contrast, DSB repair by non-homologous end joining (NHEJ) in *S. cerevisiae* is a minor pathway involving genes belonging to the RAD-50/Ku-70 epistasis group [16••]. Genes homologous to those of the *S. cerevisiae* RAD-52 and RAD-50/Ku-70 epistasis group have been identified in plant [17,18•] and mammalian systems [19], and their studies

Figure 1

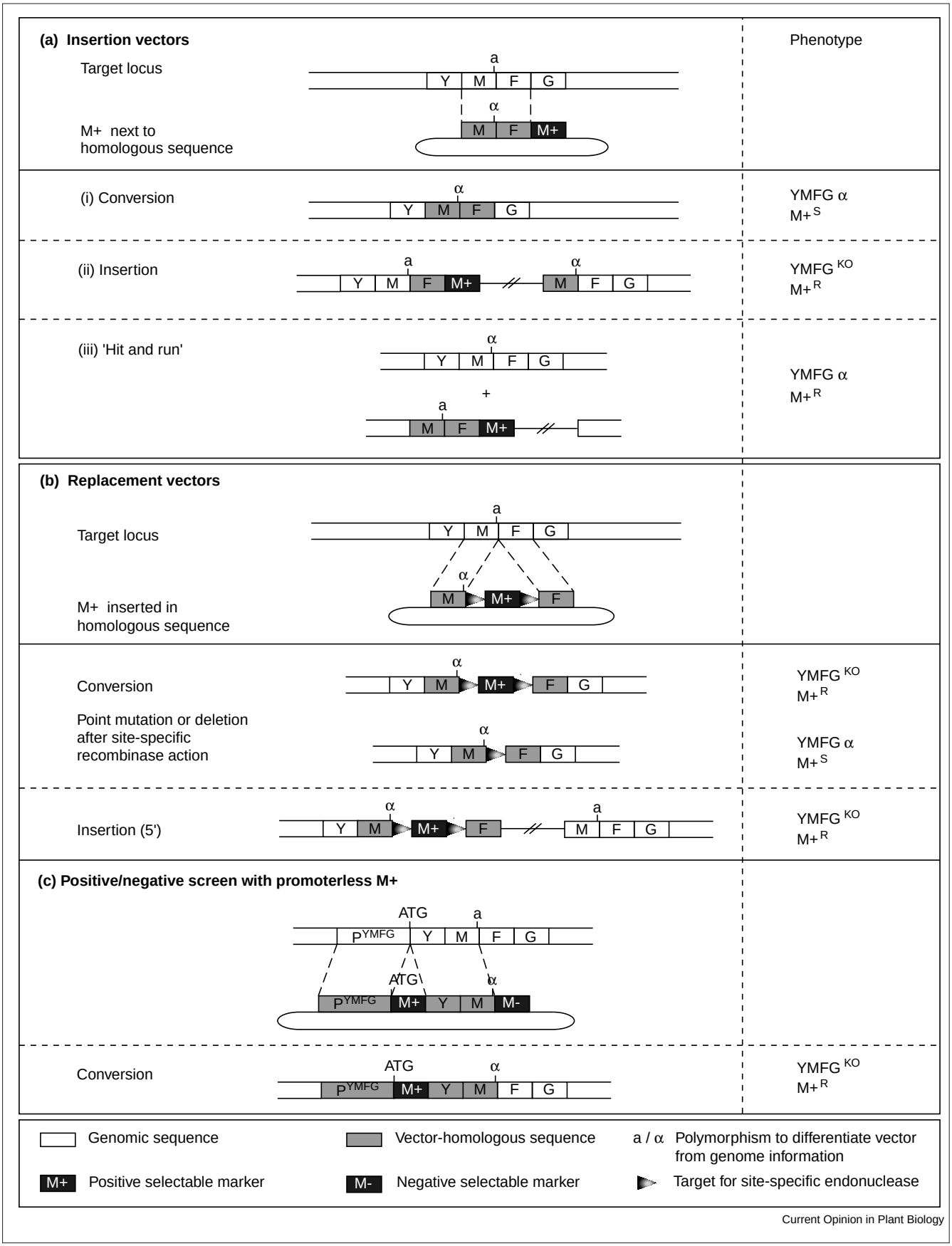
(a) Insertion vectors
Phenotype
Target locus
M+ next to homologous sequence
(i) Conversion
YMFG α
M+ S
(ii) Insertion
YMFG KO
M+ R
(iii) 'Hit and run'
YMFG α
M+ R
(b) Replacement vectors
Target locus
M+ inserted in homologous sequence
Conversion
YMFG KO
M+ R
Point mutation or deletion after site-specific recombinase action
YMFG α
M+ S
Insertion (5')
YMFG KO
M+ R
(c) Positive/negative screen with promoterless M+
Conversion
YMFG KO
M+ R
Genomic sequence
Vector-homologous sequence
a / α Polymorphism to differentiate vector from genome information
M+ Positive selectable marker
M- Negative selectable marker
Target for site-specific endonuclease
Current Opinion in Plant Biology

Figure 1 legend

Current gene targeting strategies use either insertion or replacement vectors. **(a)** An insertion vector carries, as a targeting fragment, an internal part of 'your most favourite gene' (YMFG) that lacks the amino- and carboxyl-terminal coding sequence cloned next to the positive selectable marker. The generation of a DSB in the homologous sequence generates an 'ends-in' vector, a structure that favours gene conversion events. Three types of homologous interactions have been observed with insertion vectors. (i) Conversion events are characterised by reciprocal or non-reciprocal transfer of genetic information from the vector to the target gene without integration of vector sequences. Conversion events can only be identified if mutations in the gene can be directly selected. This is the case for genes encoding acetolactate synthase, hypoxanthine phosphoribosyl transferase or adenine phosphoribosyl transferase: mutations in their coding sequences confer resistance to chlorsulfuron, hypoxanthine or adenine analogues, respectively. (ii) Insertion events are characterised by the integration of one or several copies of the transforming DNA in YMFG, generating an insertional disruption (i.e. a knock-out [$YMFG^{KO}$]) and conferring resistance to the positive selectable marker ($M+^{R}$). (iii) 'Hit and run' represents a third class of homologous interaction events observed in plant and animal cells. HR between transforming and genomic DNA results in reciprocal or non-reciprocal transfer of genetic information between both DNAs followed by illegitimate integration of the transforming DNA at ectopic locations. In this case, the mutated allele of YMFG and the positive marker will display independent segregation. **(b)** In a replacement vector, the positive selectable marker flanked by site-specific recombinase recognition sites (e.g. lox sites) is inserted in an internal mutated fragment of YMFG that lacks amino- and carboxyl-terminal coding sequences. DSB generated at the ends of homologous sequences generate an Omega or 'ends out' vector that will favour HR events. (i) Conversion events will disrupt YMFG or mutate it if the selectable marker is inserted in intron sequences. Subsequent transient expression of the site-specific recombinase (e.g. Cre recombinase) excises the positive selectable marker, resulting in a point mutation of YMFG. (ii) Insertion events mediated by HR within 5′ or 3′ homologous sequences can also occur, leading to insertional gene disruption. **(c)** To select against insertion events, positive/negative selection screens are applied. The most efficient counter selection is achieved when the positive selectable marker coding sequence is fused to the promoter of YMFG. With such a strategy, despite the fact that the population of cells to screen becomes very large, gene knock-outs have been isolated in mammalian cell lines that have a very low ratio of targeted to random integration events. $M+^{S}$, susceptibility to the positive selectable marker.

have further confirmed that inefficient GT in plants and mammals can be correlated with the fact that DSB are predominantly repaired by NHEJ in these organisms.

Yet, the genetic determinants that favour DSB repair by HR rather than by NHEJ are currently unknown. This is illustrated by the fact that a mutation that specifically alters the ratio of targeted to random integration events observed during integrative transformation has not yet been identified in yeast. Obviously, the identification of these genetic determinants will be an important advance in biology as it would not only allow GT approaches to be broadly applied in functional genomic studies, but also enable accurate gene therapy and the evaluation of engineered genes in their own chromosomal environment (i.e. it will facilitate protein design).

Table 1

Comparison of gene-targeting efficiencies in yeast, fungi and animals.

Species*	GT efficiency GT/GT+IR‡ (%)	GT efficiency per µg DNA	GT efficiency per living cell	Size of homology† (kb)	Reference
Yeasts					
S. cerevisiae	95	1–10	nd	0.1– 0.5	[41]
Schizosaccharomyces pombe	10–90	1–10	nd	1–2	[42]
Filamentous fungi					
Aspergillus nidulans	5–75	1–10	nd	4–6	[43]
Neurospora crassa	1–30	nd	nd	2–9	[44]
Protozoa					
Trypanosomatidae	95	100	10^{-4}	1–10	[45]
Dictyostelium discoideum	15	1	10^{-5}	5.3	[46]
Animals					
Caenorhabiditis elegans		No reports			
Zebra fish		No reports			
Drosophila melanogaster§	nd¶	nd	nd	1	[47••]
Mouse ES cell	0.1–1	nd	10^{-4}	10–15	[48]
Chicken B cell lines	12–100	0.2–1	nd	8–13	[49]
CHO cells	0.01	0.05	5×10^{-6}	4	[50]
Mammalian cells #	0.1–40	0.01	10^{-6}	2–8	[51•]

*Species with GT efficiencies above 1% are in bold. †Size of homologous sequence required to achieve GT at the mentioned efficiencies. ‡Ratio of GT events monitored in a population of integrative transformants (i.e. GT/GT+IR [%]). §Targeted integration occurred *in vivo* after illegitimate integration of the targeting sequence and its subsequent excision by FLP recombinase and linearisation by I-SceI endonuclease. This linear substrate generated *in vivo* subsequently integrates in the target locus by homologous recombination. #These high GT/IR ratios were obtained by a positive/negative selection screen in which the promoterless coding sequence of the selectable marker is fused to the promoter of the target gene (see Figure 1c). This strategy is a very efficient method of counter-selecting illegitimate integration events but does not change the low yield of GT events. ¶nd, not determined.

Table 2

Comparison of gene-targeting efficiencies monitored to date in plants and in *P. patens*.

	GT efficiency GT/GT+IR (%)	GT efficiency per µg DNA	GT efficiency per living cell	Size of homology* (kb)	Notes†	Reference(s)
Plants						
Nicotiana tabacum	0.01	0.02	10^{-6}	0.4	PEG (i)	[11]
	0.008	nd[#]	10^{-6}	1.9	At (i)	[52]
	0.02	nd	10^{-6}	3.6	At (i)	[53]
Arabidopsis thaliana	0.005	0.001	10^{-6}	1.0	PEG (i)	[54]
‡	0.08	nd	$5{\cdot}10^{-5}$	4 + 3	At (r)	[55]
‡	0.13	nd	10^{-4}	3 + 4	At (r)	[56]
Volvox carteri	0.02–0.1	nd	nd	6	Bi (i)	[57]
Chlamydomonas reinhardtii	0.1–1	nd	nd	8	PEG (i)	[58]
Lotus japonicus §	< 0.005	nd	nd	20	At (r)	[59]
Physcomitrella						
Artificial loci	75–100	0.3–30	10^{-3}–10^{-4}	1.8–3.6	(i)	[21,24]
Genomic loci	66–100	0.5–4	10^{-3}–10^{-4}	2.4–3.5	gDNA (i)	[25]
Cab *ZLAB1*	33	0.4	5×10^{-4}	1.0	gDNA (i)	[27•]
FtsZ	14	0.5	1.7×10^{-4}	0.3 + 0.6	cDNA (r)	[29]
Δ–6 desaturase	95	0.5	6×10^{-3}	0.9 + 1.1	gDNA (r)	[32]
Mcb1	4	0.04	9×10^{-4}	0.7 + 0.6	cDNA (r)	[34••]
Aprt	10	0.05	10^{-5}	0.2 + 0.5	cDNA (r)	[28]
	90	0.5–5	10^{-3}	0.6 + 0.7	gDNA (r)	(a)

*Size of homologous sequence required to achieve GT at the mentioned efficiencies. † Methodological annotations: At, *A. tumefaciens*-mediated transformation; Bi, biolistic transformation; cDNA or gDNA, nature of the targeting sequence; i, insertion vector; PEG, polyethylene glycol direct gene transfer to protoplasts; r, replacement vector. ‡ Targeting efficiencies in these reports are statistically questionable as only one and three targeted clones have been obtained. § Despite the use of a positive/negative screen and large extent of sequence homology, no targeted plants were identified among 20,000 transformants. [#]nd, not determined. (a) DG Schaefer, M Laloue, unpublished data.

Genetic transformation of *Physcomitrella*

Two classes of antibiotic-resistant colonies are observed following polyethylene-glycol-mediated direct transfer of bacterial plasmids carrying a plant antibiotic-resistance cassette into protoplasts [20]. The first class is characterised by resistant colonies displaying reduced growth rates on antibiotic-supplemented medium and a rapid loss of the resistance gene following a period of growth on non-selective medium. These antibiotic resistant colonies are obtained at relative transformation frequencies (RTF) as high as 10% (where RTF are the number of antibiotic resistant colonies divided by the number of surviving colonies after transformation), corresponding to several hundred clones per µg transforming DNA [21]. Subsequent molecular analyses of such strains have revealed that the transforming plasmid is concatenated to form high molecular weight episomal structures that are replicated in moss cells but display poor mitotic segregation [21,22,23•]. These 'replicative transformants' are obtained in every transformation experiment and provide the basis for the future development of shuttle vectors and moss artificial chromosomes [21].

The second class is characterised by antibiotic resistant colonies displaying mitotic stability and Mendelian segregation of antibiotic resistance genes. In the absence of sequence homology between the transforming DNA and the moss genome, these integrative transformants are recovered at RTF in the range of 10–5 (*circa* one clone per 10–50 µg DNA) [20]. Considering the efficient replicative transformation frequencies observed in these experiments, these data suggest that integration of foreign DNA sequences at random locations of the genome by illegitimate recombination is possible but not very efficient.

GT is as efficient in *Physcomitrella* as in *S. cerevisiae*

Experiments designed to target artificial loci previously generated in the moss genome by illegitimate recombination provide the first evidence to suggest that homologous recombination is the predominant pathway of integration of DNA sequences in the genome of *P. patens*. Two plasmids that are homologous except for the antibiotic resistance gene (i.e. carrying a kanamycin or a hygromycin resistance gene, respectively) have been used independently to generate kanamycin- or hygromycin-mono-resistant transgenic strains. These strains have subsequently been retransformed with the other plasmid to generate kanamycin and hygromycin resistant transgenic strains. These experiments revealed that RTF are one order of magnitude higher in mono-resistant transgenic strains than those in the wild-type strain and that this increase is correlated with co-segregation of both antibiotic selectable markers in 87% of the tested plants ([21,22,24]; Table 2). Subsequent molecular analyses have confirmed integration of the second plasmid into each target artificial

locus by homologous recombination (DG Schaefer, S Vlach, unpublished data).

To rule out the possibility that such high GT efficiency could be associated with unusual features connected with artificial loci, subsequent GT experiments have been performed on three single-copy unsequenced moss genomic sequences [25]. RTF are again one order of magnitude higher with plasmids carrying moss genomic sequences as compared to those observed with plasmids lacking sequence homology to the moss genome. Southern blot analyses have confirmed targeted integration by HR between the transforming DNA and the genomic homologous sequences at every tested locus and in 66–100% of the plants analysed (Table 2). Consequently, it has been concluded that HR is the predominant pathway used by *Physcomitrella* to integrate foreign DNA sequences in the genome [25,26]. Subsequently, the optimal and minimal extents of sequence homology required to target genes in *Physcomitrella* have been assessed, as has the specificity of GT in this moss. The specificity of GT in *Physcomitrella* was determined in experiments designed to target a specific member of the highly conserved chlorophyll a/b binding protein multigene family (i.e. the *cab ZLAB1* locus). Although sequence homology between 11 members of the *cab* multigene family is as high as 88-93% at the nucleotide level, successful targeting of the *ZLAB1* locus has been achieved in 30% of the transgenic plants analysed (Table 2). Sequence analysis of the hybrid junctions between chromosomal and plasmid sequences has confirmed the fidelity of HR-mediated targeted integration at the base-pair level [27•].

The minimal and optimal extent of sequence homology required for GT was determined in experiments designed to target the adenine phosphoribosyl transferase gene of *Physcomitrella* (Pp*apt*). These experiments took advantage of the fact that loss of function of the *apt* gene can be directly selected as it confers resistance to the adenine analogue 2,6-diaminopurine (DAP; DG Schaefer, M Laloue *et al*., unpublished data). Using cDNA-based replacement vectors, conversion of the *apt* gene mediated by HR events occurring within stretches of continuous homology as short as 53 and 191 basepairs (bp) have been observed, although at low frequencies ([28]; Table 2). Using genomic DNA (gDNA)-based gapped replacement vectors, stretches of homology ranging from 500–700 bp are sufficient to target the *apt* locus at RTF as high as 10^{-3} (up to one clone per μg DNA). These experiments have also revealed that targeted insertion versus conversion events occur at similar frequencies, demonstrating that point mutagenesis is directly accessible in *Physcomitrella*. These optimal and minimal sizes of homologous sequence required for GT correspond to those observed in yeast and have been independently confirmed in gene knock-out experiments on the *ftsZ*, Δ-6 desaturase and *mcb1* loci, which will be discussed below (Table 2). Thus, we can conclude from these data that GT is as efficient in *P. patens* as in *S. cerevisiae*, much easier in *P. patens* than in mouse ES cells and several orders of magnitude more efficient in *P. patens* than in most other eukaryotes. In fact, *P. patens* ranks among the most efficient multicellular eukaryotes for functional genomics studies using GT (Tables 1,2).

Table 3

Non-exhaustive list of moss ESTs that are homologous to genes involved in different biological processes in plants.

Biological process	ESTs	Species*
Plant development	MADS box genes	P
	Homeobox genes	P
Hormone-mediated responses	Auxin binding protein ABP1	C
	Auxin resistance protein AXR1	P
	ABA-responsive protein	P, C
Transduction pathways	Ser/threo phosphatase 2A, 2C	P, C
	G-proteins subunit	P
	GTPase rac1, rac4	P
	Protein kinase C, MAP, cAMP	P, C
Plant–pathogen interactions	CF-4/9 haplotype resistance gene	P
	CF-9 resistance gene cluster	P
	Lipoxygenase lox1, lox3	P, C
Transcription factors	Myb 3R-1	P
	Myc	P
Cellular ultrastructure	Actin	P, C
	Tubulin α, β, γ	P
	FtsZ1 and FtsZ2	P
DNA metabolism	DNA polymerase γ	P
	RNA polymerase I, II, III subunits	P, C
	Topoisomerase I	P
Cell wall metabolism	Pectin methylesterase	P, C
	Cellulose synthase	P, C

*The moss EST database (URL http:www.ncbi.nih.gov) contains mainly sequences from *Physcomitrella patens* (P) but also some sequences of another species from the Funariales family, *Ceratodon purpureus* (C). *C. purpureus* is currently used in studies of phototropic responses and of the mechanism underlying the establishment of cell polarity. ABA, abscisic acid; CF-4/9, *Cladosporium fulvum*-4/9; MAP, mitogen-activated protein.

Why is GT so efficient in *Physcomitrella*?

Obviously, achieving GT in eukaryotes is not a 'just so story' and two hypotheses have been proposed to account for the high efficiency of GT in *P. patens*. First, efficient GT may be correlated with the dominance of the gametophytic haplophase in the life cycle of Bryophytes [21,25]. This proposition is based on the observation that a ratio of targeted to random integration events above 1% seems to be restricted to primitive and/or haploid eukaryotes, such as protozoa, yeast, some green algae and filamentous fungi (Table 1). Yet such a ratio is not sustained in wild-type *S. cerevisiae*, which is naturally diploid. The proposition is also supported by the assumption that DSB in a haploid genome have to be repaired by HR in order to maintain the integrity of the genome. GT experiments performed on the gametophytic cells of vascular plants may validate this correlation.

The second hypothesis is based on the observation that *P. patens* chloronemal cells, from which protoplasts are isolated for transformation experiments, divide synchronously every 24 hours and are arrested for most of the day at the $G_2 \rightarrow M$ boundary [29,30]. Reski [30] proposed that efficient GT could be correlated with this cell cycle arrest but this assumption is not supported by controversial data from other biological systems. Experiments conducted in *S. cerevisiae* have not shown any correlation between GT efficiency and cell cycle stages, except that DSB repair by NHEJ induced by the site-specific HO endonuclease drops nearly 100-fold during G_1 [16••], which would argue against Reski's proposition. In contrast, in vertebrate cells, HR contributes to DSB repair predominantly during $S \rightarrow G_2$ interphase and NHEJ is the dominant pathway during G_1/early S phase [31]. The possibility of a correlation between GT competency and cell cycle stages in plants can be tested experimentally in different synchronised plant cell lines that have been arrested at different stages of the cell cycle. Finally, functional analysis of moss homologues to the genes of the RAD-50/Ku-70 and RAD-52 epistasis group would allow the investigation of homologous and illegitimate recombination processes in *P. patens*.

Gene disruption in *Physcomitrella* generates predicted and unpredicted phenotypes that may be complemented by genes from other plant species

Strepp and co-workers successfully disrupted the *ftsZ1* gene of *P. patens*, which encodes an ancestral tubulin that is thought to be involved in chloroplast division [29]. As predicted, the *ftsZ* knock-out phenotype is characterised by the presence of a single giant chloroplast per cell instead of the 50 chloroplasts that are usually found in wild-type cells. Girke and co-workers have generated Δ-6 desaturase knock-out strains whose phenotype is a severe alteration of the normal *P. patens* fatty-acid pattern [32]. Even more interestingly, Girod *et al.* [33,34••] have shown that disruption of the multi-ubiquitin chain binding gene *mcb1* (encoding a protein of the 26S proteasome) leads to developmental arrest in *P. patens*. This results in the generation of strains that have abnormal filamentous caulonemata and are unable to progress through the developmental transition from uni-dimensional apical growth to three-dimensional caulinary growth. This phenotype cannot be observed in unicellular yeast in which Δmcb1 strains display normal growth on rich medium but are more sensitive than the wildtype to amino-acid analogues. These findings emphasise the fact that loss-of-function mutations of genes involved in general biological processes may have more dramatic effects in multicellular eukaryotes than in yeast. Finally, *apt* knock-out strains of *P. patens* display, in addition to resistance to DAP, a developmental phenotype characterised by altered development of the leafy shoot (i.e. gad for gametophore altered development) (DG Schaefer, M Laloue *et al.*, unpublished data). *Arabidopsis apt1* mutants are male sterile but do not otherwise display such a strong developmental phenotype [35], probably because *apt* genes form a small multigene family in *Arabidopsis* and because *apt* null mutants have not yet been obtained. Nevertheless, complementation of the moss developmental phenotype with the *Arabidopsis Atapt1* and *Atapt2* genes has been successful, providing evidence that angiosperm genes can effectively complement mutations in *P. patens* (DG Schaefer, M Laloue *et al.*, unpublished data). Thus, the characterisation of loss-of-function mutations in *Physcomitrella* can provide additional information to that previously obtained in yeast and *Arabidopsis*.

Physcomitrella, a simple plant model system with modern molecular genetic features

It is now well recognised that bryophytes and angiosperms display the same basic biological processes [36,37,38•]. The initiation of a large-scale EST sequencing program in *P. patens* [39••] has so far provided almost 10,000 ESTs, representing the same number of directly accessible putative knockouts (Table 3). Angiosperms' genes and EST sequences available in the database can be used either as heterologous probes or to design degenerate PCR primers in order to identify and isolate the moss homologues and immediately initiate fine functional studies of these genes by GT in *P. patens*. Even more promising, the successful development of tagged mutagenesis and gene-trap approaches in *P. patens* by shuttle mutagenesis [40••] allows the use of a molecular genetic methodology in plants that, up to now, has only been applied in yeast. The data reviewed here clearly support the idea that *Physcomitrella* advantageously complements *Arabidopsis* and rice as a model system for plant biological studies.

Conclusions

Efficient GT in *P. patens* finally provides the missing genetic tool required for fine functional genomic studies in plants. Research with *P. patens* is entering an expansion phase. This is the appropriate time to consider the opportunity to initiate a genome-sequencing program for *P. patens*, which has a genome size only twice that of *Arabidopsis* and similar to that of rice, that would provide the scientific community with information indispensable for future research. Considering the importance of efficient GT for functional genomics and the developmental complexity of *P. patens* compared to the few other eukaryotes in which this technology is possible, further progress may place this plant at the forefront of biological studies in multicellular eukaryotes.

Acknowledgements

I would like to thank M Laloue, M Chakhaparonian, N Houba-Hérin, K von Schwartzenberg and S Vlach for sharing results prior to publication and J-P Zrÿd for fruitful discussions. Part of the work reported here was supported by the Swiss Office Fédéral pour l'Education et la Science (OFES), in the frame of the Euromoss project of the EU PTP 1993–1996.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Bouchez D, Höfte H: **Functional genomics in plants.** *Plant Physiol* 1998, **118**:725-732.

2. Kehoe DM, Villand P, Somerville S: **DNA microarrays for studies of higher plants and other photosynthetic organisms.** *Trends Plant Sci* 1999, **4**:38-41.

3. Richmond T, Somerville S: **Chasing the dream: plant EST microarrays.** *Curr Opin Plant Biol* 2000, **3**:108-116.

4. Thiellement H, Bahrman N, Damerval C, Plomion C, Rossignol M, Santoni V, de Vienne D, Zivy M: **Proteomics for genetic and physiological studies in plants.** *Electrophoresis* 1999, **20**:2013-2026.

5. • Weigel D, Ahn JH, Blazquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferrandiz C, Kardailsky I, Malancharuvil EJ, Neff MM *et al.*: **Activation tagging in *Arabidopsis*.** *Plant Physiol* 2000, **122**:1003-1013.

The authors of this study describe progress towards the isolation of gain-of-function mutants of *Arabidopsis*. A collection of plants transformed with T-DNA carrying multimerised transcriptional enhancers from the cauliflower mosaic virus 35S promoter was produced. Dominant mutated phenotypes resulting either from constitutive ectopic or enhanced expression of a gene induced by the 35S enhancers are observed in *circa* 0.1% of the transformants. The subsequent identification of the mutated gene is facilitated by the T-DNA tag.

6. • Jeon JS, Lee S, Jung KH, Jun SH, Jeong DH, Lee J, Kim C, Jang S, Yang K, Nam J *et al.*: **T-DNA insertional mutagenesis for functional genomics in rice.** *Plant J* 2000, **22**:561-570.

The first description of a large collection of T-DNA tagged insertional mutants in rice. The T-DNA used in this work also carries a promoterless β-glucuronidase (GUS) gene that functions as a gene trap, enabling detection of fusion products between GUS and the endogenous tagged gene.

7. Parinov S, Sundaresan V: **Functional genomics in *Arabidopsis*: large-scale insertional mutagenesis complements the genome sequencing project.** *Curr Opin Biotech* 2000, **11**:157-161.

8. • Springer PS: **Gene traps: tools for plant development and genomics.** *Plant Cell* 2000, **12**:1007-1020.

A detailed and comprehensive review of the different strategies used for gene trapping in plants. The advantages and disadvantages of different types of vectors (i.e. T-DNA or transposon) and reporter genes (i.e. GUS or green fluorescent protein [GFP]) are highlighted.

9. Capecchi MR: **Altering the genome by homologous recombination.** *Science* 1989, **244**:1288-1292.

10. Capecchi MR: **The new mouse genetics: altering the genome by gene targeting.** *Trends Genet* 1989, **5**:70-76.

11. Paszkowski J, Baur M, Bogucki A, Potrykus I: **Gene targeting in plants.** *EMBO J* 1988, **7**:4021-4026.

12. Puchta H, Hohn B: **From centiMorgans to base pairs: homologous recombination in plants.** *Trends Plant Sci* 1996, **1**:340-348.

13. •• Mengiste T, Paszkowski J: **Prospects for the precise engineering of plant genomes by homologous recombination.** *Biol Chem* 1999, **380**:749-758.

This review summarises the current status of gene targeting and homologous recombination in plants. Special emphasis is given to the biochemical models of DSB repair and to progress using genetic approaches, which involve the isolation of radiation-sensitive mutants, aimed at identifying genes involved in recombination processes in plants.

14. •• Vergunst AC, Hooykaas PJJ: **Recombination in the plant genome and its application in biotechnology.** *Crit Rev Plant Sci* 1999, **18**:1-31.

An extensive review on homologous recombination and gene targeting in plants that complements [13••]. This review also provides information on putative problems associated with *Agrobacterium*-mediated targeted transformation, extrachromosomal and intrachromosomal recombination in plants, site-specific mutations induced by the *Cre/lox* system from bacteriophage P1 and attempts to improve GT efficiency by overexpressing recombination enzymes such as RecA, a major recombination protein from *Escherichia coli*.

15. Struhl K: **The new yeast genetics.** *Nature* 1983, **305**:391-397.

16. •• Paques F, Haber JE: **Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*.** *Microbiol Mol Biol Rev* 1999, **63**:349-404.

An extremely detailed and comprehensive review on DSB repair in *S. cerevisiae*. A number of topics are discussed including the mechanism of DSB repair by HR; proteins involved in HR and NHEJ; the stimulation of mitotic recombination in relation to DNA topology, structure and replication; meiotic recombination; the role of mismatch repair in recombination; and DSB repair check points. This review is a must for everybody who wishes to have a comprehensive view of recombination in eukaryotes.

17. Britt AB: **Molecular genetics of DNA repair in higher plants.** *Trends Plant Sci* 1999, **4**:20-25.

18. • Gorbunova V, Levy AA: **How plants make ends meet: DNA double strand break repair.** *Trends Plant Sci* 1999, **4**:263-269.

Recent studies showing that the predominant pathway of DSB repair in plants involves NHEJ are reviewed. The authors also discuss the fact that error-prone DSB repair may be an important driving force in plant genome evolution.

19. Kanaar R, Hoeijmakers JHJ, van Gent DC**: Molecular mechanisms of DNA double-strand break repair.** *Trends Cell Biol* 1998, **8**:483-489.

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

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

22. Schaefer DG, Bisztray G, Zrÿd J-P: **Genetic transformation of the moss *Physcomitrella patens*.** In *Plant Protoplasts and Genetic Engineering V*. Edited by Bajaj YPS. Berlin: Springer Verlag; 1994: 349-364.

23. • Ashton NW, Champagne CEM, Weiler T, Verkoczy LK: **The bryophyte *Physcomitrella patens* replicates extrachromosomal transgenic elements.** *New Phytol* 2000, **146**:391-402.

The results reported here confirm those obtained in [21] suggesting that plasmids can be concatenated to form high molecular weight extrachromosomal structures that are replicated in moss cells. This finding provides the starting point to the development of moss artificial chromosomes and moss shuttle vectors. The discrepancies in transformation frequencies described in this report and in [21] are probably caused by differences in the transformation protocols.

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

25. Schaefer DG, Zryd J-P: **Efficient gene targeting in the moss *Physcomitrella patens*.** *Plant J* 1997, **11**:1195-1206.

26. Puchta H: **Towards targeted transformation in plants.** *Trends Plant Sci* 1998, **3**:77-78.

27. • Hofmann AH, Codon AT, Ivascu C, Russo VEA, Knight C, Cove DJ, Schaefer DG, Chakhaparonian M, Zrÿd J-P: **A specific member of the *cab* multigene family can be efficiently targeted and disrupted in the moss *Physcomitrella patens*.** *Mol Gen Genet* 1999, **261**:92-99.

The accuracy of GT in *Physcomitrella* is demonstrated on a specific member of a highly conserved multigene family, and HR events are documented at the sequence level.

28. Houba-Hérin N, Reynolds S, Schaefer D, von Schwartzenberg K, Laloue M: **Molecular characterisation of homologous recombination events in the *moss Physcomitrella patens*.** In *Plant Sciences 1997, 3e Colloque Général de la Société Française de Physiologie Végétale*. Edited by Petch JC, Latché A, Bouzayen M. Toulouse: SFPV-INRA; 1997:22-23.

29. Strepp R, Scholz S, Kruse S, Speth V, Reski R: **Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin.** *Proc Natl Acad Sci USA* 1998, **95**:4368-4373.

30. Reski R: ***Physcomitrella* and *Arabidopsis*: the David and Goliath of reverse genetics.** *Trends Plant Sci* 1998, **3**:209-210.

31. Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, Utsumi H, Yamaguchiiwai Y, Shinohara A, Takeda S: **Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the**

maintenance of chromosomal integrity in vertebrate cells. *EMBO J* 1998, **17**:5497-5508.

32. Girke T, Schmidt H, Zähringer U, Reski R, Heinz E: **Identification of a novel Δ6-acyl-group desaturase by targeted disruption in *Physcomitrella patens*.** *Plant J* 1998, **15**:39-48.

33. Girod P-A, Fu H, Schaefer DG, Zrÿd J-P, Vierstra RD: **Targeted disruption of *Mcb1* leads to inhibition of gametophore development in the moss *Physcomitrella patens*.** Abstract 365 of the 9th Conference on *Arabidopsis* Research, 1998 June 24–28, Madison, Wisconsin.

34. •• Girod PA, Fu HY, Zryd JP, Vierstra RD: **Multiubiquitin chain binding subunit MCB1 (RPN10) of the 26S proteasome is essential for developmental progression in *Physcomitrella patens*.** *Plant Cell* 1999, **11**:1457-1471.
This work demonstrates that disruption of a subunit of the 26S proteasome leads to developmental arrest in *P. patens*, an observation that was not feasible in yeast. Partial complementation of the developmental-arrest phenotype by plant hormones is reported, suggesting that the ubiquitin/26S proteasome proteolytic pathway is involved in plant developmental processes that are triggered by hormones.

35. Gaillard C, Moffat BA, Blacker M, Laloue M: **Male sterility associated with APRT deficiency is *Arabidopsis thaliana* results from a mutation in the gene *APT1*.** *Mol Gen Genet* 1998, **257**:348-353.

36. Cove DJ, Knight CD, Lamparter T: **Mosses as model systems.** *Trends Plant Sci* 1997, **2**:99-105.

37. Reski R: **Development, genetics and molecular biology of mosses.** *Bot Acta* 1998, **111**:1-15.

38. • Reski R: **Molecular genetics of *Physcomitrella*.** *Planta* 1999, **208**:301-309.
Several aspects of molecular genetics in *Physcomitrella patens* are discussed. Data on nuclear, plastidian and mitochondrial genomes and on the codon usage of this moss are included.

39. •• Quatrano R, Bashiardes S, Cove D, Cuming A, Knight C, Clifton S, Marra M, Hillier L, Pape D, Martin J *et al.*: **Leeds Washington University Moss EST Project.** URL http://www.ncbi.nlm.nih.gov
This database provides the scientific community with the first large set of EST sequences from the mosses *Physcomitrella* and *Ceratodon*.

40. •• Nishiyama T, Hiwatashi Y, Sakakibara K, Kato M, Hasebe M: **Tagged mutagenesis and gene-trap in the moss, *Physcomitrella patens* by shuttle mutagenesis.** DNA Research 2000, **7**:9-17.
Shuttle mutagenesis was applied for the first time in plants to generate a mutagenised moss genomic library that was designed to tag or trap genes by targeted integration. The minitransposons used for shuttle mutagenesis in *E. coli* carry either a selectable neo cassette (for tagged mutagenesis) or a *GUS* gene (for gene trapping). The mutagenised library is used to transform *P. patens*. Preliminary data indicate that single-locus-targeted insertions are predominant. Mutated phenotypes are recovered in 4% of the transformants, whereas 3% of the transformants exhibit cell-specific or constitutive GUS expression.

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

42. Grimm C, Kohli J: **Observations on integrative transformation in *Schizosaccharomyces pombe*.** *Mol Gen Genet* 1988, **215**:87-93.

43. Miller BL, Miller KY, Timberlake WE: **Direct and indirect gene replacements in *Aspergillus nidulans*.** *Mol Cell Biol* 1985, **5**:1714-1721.

44. Asch DK, Kinsey JA: **Relationship of vector insert size to homologous integration during transformation of *Neurospora crassa* with the cloned *am* (GDH) gene.** *Mol Gen Genet* 1990, **221**:37-43.

45. Papadopoulou B, Dumas C: **Parameters controlling the rate of gene targeting frequency in the protozoan parasite *Leishmania*.** *Nucleic Acids Res* 1997, **25**:4278-4286.

46. De Lozanne A, Spudich JA: **Disruption of the *Dictyostelium* myosin heavy chain gene by homologous recombination.** *Science* 1987, **236**:1086-1091.

47. •• Rong YS, Golic KG: **Gene targeting by homologous recombination in *Drosophila*.** *Science* 2000, **288**:2013-2018.
The first report of successful gene targeting in *Drosophila*. This success was obtained with a sophisticated experimental design that combines the use of site-specific recombinase and site-specific endonuclease. The targeting sequence is initially introduced at ectopic locations by standard transformation via P elements. It is subsequently excised *in vivo* as a circular molecule by the FLP recombinase (from the FLP/frt site-specific recombination system from *S. cerevisiae*) and linearised within targeting homologous sequences by the I-SceI site-specific endonuclease from *S. cerevisiae*. This linear substrate finally reintegrates at the targeted location by HR. This approach might be applied to other species recalcitrant to GT but the general validity of such a strategy is still a matter for discussion (see [60]).

48. Deng C, Capecchi MR: **Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus.** *Mol Cell Biol* 1992, **12**:3365-3371.

49. Buerstedde J-M, Takeda S: **Increased ratio of gene targeted to random integration after transfection of chicken B cell lines.** *Cell* 1991, **67**:179-188.

50. Adair GM, Scheerer JB, Brotherman A, McConville S, Wilson JH, Nairn RS: **Targeted recombination at the Chinese hamster APRT locus using insertion versus replacement vectors.** *Somat Cell Mol Genet* 1998, **24**:91-105.

51. • Sedivy JM, Dutriaux A: **Gene targeting and somatic cell genetics — a rebirth or a coming of age?** *Trends Genet* 1999, **15**:88-90.
The authors of this article discuss the potential of positive/negative screens using a promoterless positive selectable marker fused to promoter sequences of the target gene. Selection against illegitimate integration is efficient with such vectors. Yet, the generation of a specific knock-out with such a screen in *Arabidopsis*, for example, would require an experiment of the size of that performed to generate a large collection of T-DNA tagged mutants.

52. Lee KY, Lund P, Lowe K, Dunsmuir P: **Homologous recombination in plant cells after *Agrobacterium*-mediated transformation.** *Plant Cell* 1990, **2**:415-425.

53. Offringa R, de Groot JA, Haagsman HJ, Does MP, van den Elzen PJM, Hooykaas PJJ: **Extrachromosomsal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation.** *EMBO J* 1990, **9**:3077-3084.

54. Halfter U, Morris P-C, Willmitzer L: **Gene targeting in *Arabidopsis thaliana*.** *Mol Gen Genet* 1992, **231**:186-193.

55. Miao Z-H, Lam E: **Targeted disruption of the *TGA3* locus in *Arabidopsis thaliana*.** *Plant J* 1995, **7**:359-365.

56. Kempin SA, Liljegre SJ, Block LM, Roundsley SD, Yanofsky MF, Lam E: **Targeted disruption in *Arabidopsis*.** *Nature* 1997, **389**:802-803.

57. Hallmann A, Rappel A, Sumper M: **Gene replacement by homologous recombination in the multicellular green alga *Volvox carteri*.** *Proc Natl Acad Sci USA* 1997, **94**:7469-7474.

58. Sodeinde OA, Kindle KL: **Homologous recombination in the nuclear genome of *Chlamydomonas reinhardtii*.** *Proc Natl Acad Sci USA* 1993, **90**:9199-9203.

59. Thykjaer T, Finnemann J, Schauser L, Christensen L, Poulsen C, Stougaard J: **Gene targeting approaches using positive–negative selection and large flanking regions.** *Plant Mol Biol* 1997, **35**:523-530.

60. Engels WR: **Reversal of fortune for *Drosophila* geneticists?** *Science* 2000, **288**:1973-1975.