APPLIED GENOMICS IN *PHYSCOMITRELLA*

HAUKE HOLTORF, WOLFGANG FRANK, RALF RESKI

Abstract. The moss *Physcomitrella* patens, which has become a versatile tool for plant reverse genetics, has been chosen for a large scale forward genetics mutagenesis approach to address gene function at a genome wide scale. Based on the high frequency of homologous recombination found in *Physcomitrella* targeted knockout approaches of single genes have lead to new insights into the mode of action of plant genes. Besides a targeted knockout approach of around 80 target genes which are of commercial interest we have developed a transposon-mediated mutagenesis platform to mutagenize cDNA libraries derived from different developmental stages of the *Physcomitrella* life cycle. These mutagenized cDNAs are transformed into *Physcomitrella* protoplasts to generate about 80.000 transgenic *Physcomitrella* lines. The mutagenized cDNAs are supposed to integrate into the nuclear DNA at their corresponding loci thereby disrupting the underlying gene. The basis for our experimental approach was the generation of a comprehensive EST database was build from 110.000 single cDNA sequences. To realize the generation of a saturated *Physcomitrella* mutant collection we have also adapted the production of moss material and the transformation protocols to comply with the requirements of a high-throughput mutagenesis approach. The derived transgenic plants were subjected to PCR-based molecular analysis to confirm stable integration of the selectable marker cassette. Furthermore, all transgenic lines were investigated in terms of their nuclear DNA content to detect doubling of the chcromosomal DNA that may have been occurred during the transformation procedure. Each mutant plant was also subjected to a detailed analysis with respect to its physiological, morphological and metabolic traits. The frequency of altered phenotypes in the *Physcomitrella* mutant collection was considerably higher than the reported rate in Arabidopsis. Within this chapter we outline the complete procedure of the generation of the *Physcomitrella* mutant collection and their initial characterization.

1. INTRODUCTION

Recently *Physcomitrella patens* has emerged as a model system to implement plant functional genomics. *Physcomitrella* is a non-vascular, multicellular land plant and a representative of the bryophytes (Reski 1998a, Reski 1998b, Reski 1999, Schaefer 2001, Schaefer 2002). Using *Physcomitrella* as a tool for functional genomics allows one to evade some of the apparent obstacles when working with higher plant model systems (Holtorf et al. 2002a). In particular, *Physcomitrella* stands out against other systems with respect to a dominant haploid gametophytic phase and a high rate of homologous recombination.

The dominant haploid state is advantageous when performing forward genetic screens because the loss of a certain gene function is not counterbalanced by a second allele. Molecular analysis of T-DNA or transposon-mutagenised Arabidopsis knockout mutants arising from reverse or forward genetics screens have been demonstrated to result in specific gene disruptions. Only a small number of them, however, exhibit an informative phenotype that would allow the clear definition of a gene function (Bouché

A. J. Wood, M. J. Oliver, D. J. Cove (eds.), New Frontiers in Bryology

and Bouchez 2001). A reason for this observation is most likely the result of a high degree of genetic redundancy in Arabidopsis (The Arabidopsis Genome Initiative 2000). Precise gene disruption in *Physcomitrella* that occur in a haploid background may overcome these difficulties because single gene knockout mutants are expected to show a more comprehensive phenotype. Apart from haploidy another advantage of *Physcomitrella* is even more striking. It is the only land plant studied so far that has the ability to integrate transforming DNA at a high frequency by way of gene targeting (Schaefer and Zryd, 1997), with the homologous recombination rate being several orders of magnitude higher than in seed plants (Reski 1998b, 1999).

Experiments performed by several independent groups have convincingly demonstrated that a precise disruption of a genomic locus in *Physcomitrella* correlates with a mutant phenotype that clearly reveals the biological function of the targeted gene (Girke et al. 1998, Strepp et al. 1998, Girod et al. 1999, Imaizumi et al. 2002, Koprivova et al. 2002, Zank et al. 2002).

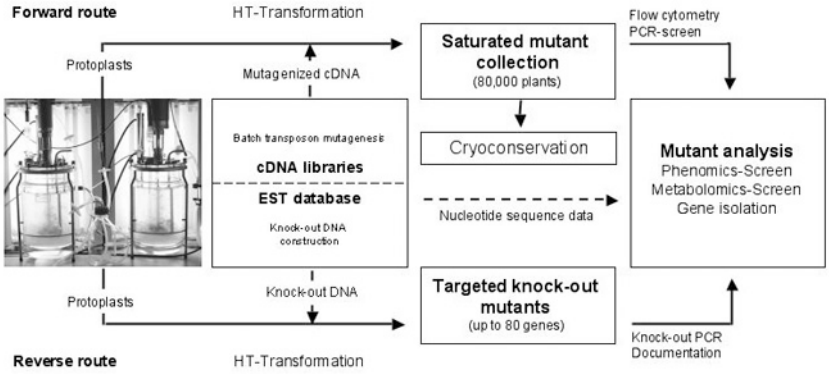

Figure 1

Schematic representation of the high-throughput platform for production of tagged Physcomitrella patens mutants. *This scheme outlines the processes underlying the production of tagged Physcomitrella mutants. Two strategies of mutant production are illustrated: a forward and a reverse route. The forward route uses anonymous cDNA clones to produce a saturated mutant collection of 80,000 plants. The cDNAs are batch-mutagenised using the transposon shuttle mutagenesis system (see Fig.2) to disrupt the DNA coding sequence within a selection cassette. Mutagenised plasmid DNA is transferred into Physcomitrella protoplasts by way of a high-throughput transformation platform. Mutants generated by the forward route are subjected to phenotypic and metabolic profiling. The reverse route is based on known moss candidate genes, the functions of which are investigated by direct gene knockout. Disruption constructs are transfected into protoplasts and resistant plants screened by PCR for targeted transgene insertion. All mutants are subjected to further molecular analyses.*

Single gene disruptions have become a routine approach in *Physcomitrella* but new strategies have also been developed to make use of the advantages of the system for large-scale reverse genetic approaches that have the potential to complement functional gene analyses in other plant species (Reski 1998b, Schaefer 2001). In close collaboration with our partner BASF Plant Science GmbH we have started to exploit the high gene-targeting rate of *Physcomitrella* to understand plant gene function on a genomic scale. The aim is to create genetic diversity by producing a saturated mutant collection in which every expressed *Physcomitrella* gene is knocked out. It is anticipated that a high proportion of mutants will have gene lesions resulting from homologous knockout integrations. By using anonymous cDNA clones derived from all stages of moss development the entire complement of expressed moss genes was used to generate disruption constructs (Fig.1). The group of Mitsuyasu Hasebe, Okazaki, Japan used a different approach to disrupt *Physcomitrella* genes using genomic DNA instead of cDNA for the large-scale production of knockout plants (Nishiyama et al. 2000). These researchers employed a shuttle mutagenesis protocol to perform the tagged mutagenesis of genomic library material. This objective, however, cannot rule out that transgenes will be targeted to genomic non-coding regions.

Following the classical forward genetic approach mutants, that have passed selection and quality control steps, are subjected to standardized phenotypic and metabolic screening protocols. Interesting phenotypes have been described and many more are expected (Egener et al. 2000). In parallel to producing a saturated mutant collection we have also conducted approximately 80 targeted knockouts of candidate genes that are of commercial interest. These genes were selected from the EST database on the basis of homology searches against public sequence databases.

This subchapter describes the essential parts of the collaboration between the Chair of Plant Biotechnology at Freiburg University and BASF Plant Science GmbH. During the course of this project it became necessary to scale up and standardize different protocols that constitute the technological platform. The essential steps involved will be discussed in the subsequent sections.

2. MASS PRODUCTION OF STANDARDIZED MOSS MATERIAL

When intending to set up a large scale moss functional genomics project one has to realize that scaling up the process of transformation implies the development of a high-throughput platform for mass production of plant material. The well-established protocols for the transformation of *Physcomitrella* protoplasts are based on polyethylene glycol (PEG)-mediated DNA transfer (Schaefer et al.1991, Reutter et al. 1998). Early protocols for the preparation and regeneration of protoplasts from *Physcomitrella* (Stumm et al. 1975) were improved by using the enzyme mixture Driselase for efficient cell wall digestion of *Physcomitrella* protonema tissue (Grimsley et al. 1977). As a source for protoplasts this protocol uses *Physcomitrella* plants grown on solidified medium in petri dishes. The medium is supplemented

with ammonium tartrate that allows cultivation of moss protonema predominantly in the chloronema stage (Ashton and Cove 1977, Jenkins and Cove 1983). Generally, chloronema cells from *Physcomitrella* are believed to be the best tissue type for the preparation of protoplasts whereas tissues from advanced developmental stages are not sufficiently susceptible to cell wall degrading enzymes (Lal 1984). The described protocols allow the preparation of sufficient amounts of protoplasts for a few transformation at a time but are not well-suited for large scale production as the cultivation of protonema on solidified medium is very labour intensive and a constant output of transformation-competent protoplasts is not possible.

To support the need for large quantities of standardized *Physcomitrella* material bioreactor cultures are preferable. Rother et al. (1994) demonstrated that *Physcomitrella* can be cultivated in liquid mineral medium without supplying ammonium tartrate. However, under such conditions *Physcomitrella* plantlets do not preferentially grow as protonemal tissue but also develop leafy gametophores. Preparing protoplasts from a mixture of tissues at different developmental stages would be unfavourable for achieving the desired high transformation rates. Consequently, an improved bioreactor protocol was required. Other early bioreactor cultures were also unsuitable as they were grown as batch or short continuous cultures and moreover produced cell material that was not suitable for protoplast isolation (Boyd et al. 1988, Cove et al. 1997, Reutter and Reski 1996). In an attempt to generate constant amounts of suitable moss material and to reduce manual labour, semicontinuous bioreactor cultures of *Physcomitrella* were developed. Protonema tissue was pre-cultured in liquid Knop medium according to Reski and Abel (1985) and equal amounts of tissue were used to inoculate stirred tank glass bioreactors to start semicontinuous liquid cultures under varying conditions of ammonium tartrate supplementation and pH-control (Hohe and Reski 2002; Hohe et al. 2001). Manipulation of ammonium tartrate supplementation and/or pH was required in order to approximate the desired protoplast generation and transformation rates. Supplementing the growth medium with 2.5 mM ammonium tartrate led to a six fold increase in protoplast yield over normal semicontinuous *Physcomitrella* bioreactor cultures. This increase was assumed to be the result of an inhibition of caulonema development in the presence of ammonium tartrate. In contrast, rigid control of pH at 4.5 enabled higher protoplast yields despite a lack of inhibition of caulonemal development.

It was assumed that pH-conditions in the medium may influence moss cell wall composition and that a constant pH-value of 4.5 may be advantageous for subsequent enzymatic digestion of cell walls. This notion was substantiated by analysing protoplast yields from protonema grown at a constant pH value of 7.5 that led to a severe drop in numbers of isolated protoplasts. Preculture conditions which use a controlled pH value of 4.5 instead of ammonium tartrate also positively influence gene targeting (GT) efficiencies (Hohe et al. 2002). Increasing the GT efficiency through modifying culture conditions is another prerequisite for producing a saturated mutant collection that is characterized by a high rate of homologous knockout integrations.

As a result of these studies a standard semicontinuous bioreactor culture at pH 4.5, without the addition of ammonium tartrate, was used for large scale production of plant material for protoplast isolation and transformation. Semicontinuous growth conditions ensure uniformity of the starting material and avoid batch-specific variance that may disturb subsequent analysis of the moss transformants. An average harvest of around 930 ml per day from a 5 l bioreactor allows preparation of protoplasts sufficient for 105-114 transformations.

3. PRODUCTION OF A PHYSOMITRELLA EST COLLECTION

Large collections of expressed sequence tag (EST) sequences are the basis for an array of different functional genomics projects covering a wide range of different plant species (Holtorf et al. 2002a).

Such EST databases have complemented genome sequencing projects performed in both Arabidopsis and rice (The Arabidopsis Genome Initiative 2000, Yu et al. 2002, Goff et al. 2002). The combined data sets, EST and genomic sequences, allow for comparative genomic and functional genomic studies. From an evolutionary point of view bryophytes are estimated to have diverged from ferns and seed plants around 450 million years ago (Theissen et al. 2001). Comparisons between the gene content of lower and higher plants will define a set of essential genes that have been conserved during evolution. Such genes will represent the minimal complement of genes present in any plant species. Among different plant lineages divergent genes can be identified that show differences with respect to the number of gene family members, the degree of sequence conservation, and expression patterns. Early studies have already demonstrated that *Physcomitrella* is a valuable source for novel genes (Reski et al. 1998, Machuka et al. 1999).

An essential part of the functional genomics project was the generation of novel EST sequence information from three cDNA libraries specific for protonemal, gametophoric, and sporophytic material. These libraries represent almost the complete transcriptome of *Physcomitrella*. in the preparation of these cDNA libraries plant material was used from all stages of the moss life cycle and, in addition, had undergone treatments with different plant hormones. The set of cDNA libraries served two functions. First, the resulting cDNA libraries were used to perform mass sequencing of individual cDNA clones to generate a comprehensive EST database. Second, the libraries were the source of cDNAs used in the production of a saturated mutant population. Transposon mutagenesis of cDNAs, in combination with high-throughput transformation and regeneration protocols, allows targeting of every expressed *Physcomitrella* gene by way of homologous recombination. The EST collection, which mirrors the content of the underlying cDNA libraries, contains sequences of over 110,000 cDNA clones (Rensing et al. 2002a, 2002b).

Prior to mass sequencing of individual clones the protonema, gametophore, and sporophyte-specific cDNAs libraries were normalized and

subtracted to drastically reduce redundancy (Rensing et al. (2002a). cDNA fragments were cloned directionally into either pUC or pBluescript plasmids and sequenced from the 3' end with primers specific to the respective vector sequences. In cases where the sequencing of a clone did not yield a valid sequence an additional sequence reaction was performed using a vector-specific 5' end primer. This was only necessary for about 3% of the sequences. Approximately 80% of the EST sequences were derived from the normalized and subtracted libraries, the remainder derive from a non-normalized protonemal library (Reski et al. 1994). In total, over 110,000 cDNA clones were sequenced representing more than five times the estimated number of *Physcomitrella* genes. This was done to ensure complete coverage of all expressed genes within the EST database. Bioinformatic tools were used for subsequent clustering of the EST sequences (Rensing et al. 2002). The removal of redundant sequences generated 33,581 sequence clusters that comprise 21,219 singletons (representing single sequences) and 12,362 contigs (built from at least two sequences). A detailed cluster analysis that generated a rate of 0.76 genes per cluster, estimated the total number of genes in *Physcomitrella* to be 25,500 +/-2,500 (expressed protein coding genes). This calculation matches the number of Arabidopsis genes, which is estimated to be 26,000, and underscores the notion that the moss EST database covers almost the entire transcriptome of *Physcomitrella*. While the *Physcomitrella* and Arabidopsis genomes contain a similar number of genes, the degree of similarity among all genes is suprisingly low. Rensing et al. (2002) estimated only a maximum of 50% of the *Physcomitrella* expressed protein genes have an Arabidopsis homologue. In addition, in a comparison of the *Physcomitrella* EST database with public sequence databases using different homology search programs, 11,600 *Physcomitrella* genes failed to match any known protein sequence. Therefore *Physcomitrella* is percieved to be a rich source of unknown plant genes, a high percentage of which could be expected to exhibit novel functions. Future investigations will certainly focus on this large set of unique moss genes.

Thorough bioinformatic analyses of the EST database generated important insights into the gene content of the *Physcomitrella* genome and its homology to other plant species. In addition the EST mass sequencing approach revealed, at an early stage, that the underlying cDNA libraries are of sufficiently high quality to be used as a gene source for transposon shuttle mutagenesis (Fig.2). Cluster analysis revealed that cDNA normalization and subtraction procedures were performed successfully. This is an important criterion when the aim is to produce a *Physcomitrella* mutant population that truly saturates the genome. The low redundancy of the cDNA clones within each library helps reduce the number of independent transformations needed to disrupt every expressed moss gene.

During the course of the project, the EST database was also used for selection of candidate genes for which targeted knockout mutagenesis is desirable (see subchapter 2.4.). It also serves to validate molecular mutant analyses (see subchapter 2.8.) by providing essential sequence information regarding the identity of isolated gene sequences.

This project was the first large scale attempt to produce a comprehensive and near complete EST database for *Physcomitrella*. Other *Physcomitrella* EST sequencing projects have been launched subsequently and have generated additional ESTs. A publicly available collection of 30,000 *Physcomitrella* ESTs was generated by a collaboration between researchers of Washington University and Leeds University (http://www.moss.leeds.ac.uk). Additionally, the group of Mitsuyasu Hasebe will reslease EST sequences into the public databases. In total, around 60,000 EST sequences are currently publically available. All EST collections will be valuable sources for plant functional genomics projects and will satisfy the growing interest in the model species *Physcomitrella.*

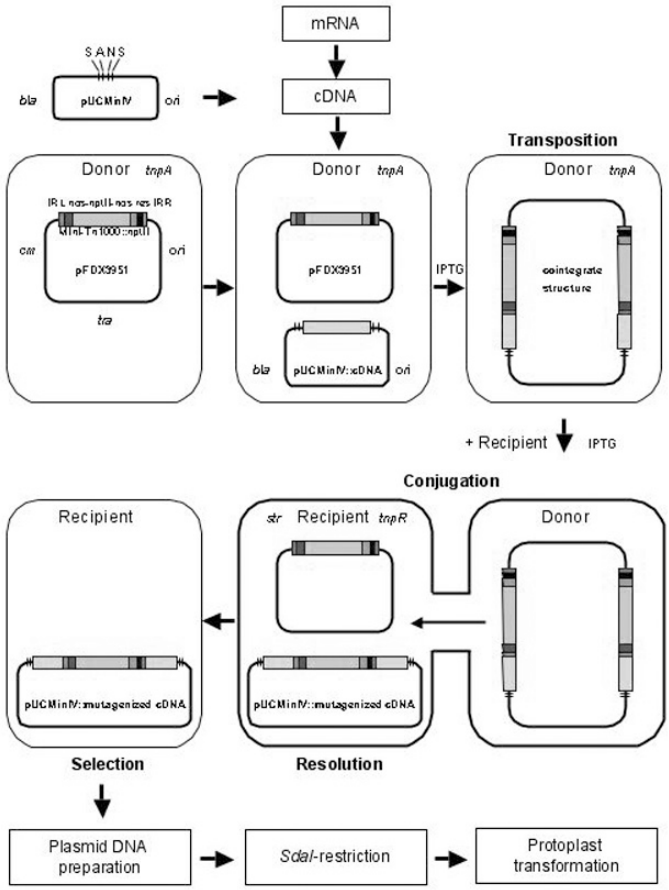

Figure 2 ***Transposon shuttle mutagenesis of Physcomitrella patens cDNA libraries.*** *Scheme outlining the process of transposon mutagenesis of Physcomitrella cDNA libraries. Low-redundancy moss cDNA populations are ligated into the minimal vector pUCMinIV. The ligation reactions are transformed into the E.coli donor strain R2217, which harbours the conjugative plasmid pFDX3951. This plasmid carries the transposon mini-Tn1000::nptII. The nptII gene is under control of nos promoter and*

termination signals. The donor strain also harbours the plasmid pFDX3957, which contains the bacterial transposase gene tnpA. TnpA is inducible through an IPTG-responsive promoter. IPTG induction activates expression of the bacterial transposase gene leading to transposition of mini-Tn1000::nptII. A cointegrate structure is formed between the plasmids pUCMinIV and pFDX3951. During conjugation of the donor with the recipient R1037 the cointegrate structure is transferred. Upon transfer the cointegrate is resolved by activity of the IPTG-induced resolvase gene in the recipient cell. After conjugation and resolution resultant cells are selected. Plasmid DNA is prepared from these cells and are digested with the rare-cutting restriction enzyme SdaI yielding DNA for transformation of Physcomitrella protoplasts. Plasmids are not to scale. tnpR: bacterial resolvase gene, tnpA: bacterial transposase gene, str: streptomycin resistance gene, cm: chloramphenicol resistance gene , bla: ampicillin resistance gene. Sites of restriction enzyme S (SdaI); A (AscI), N (NotI).

4. *IN VIVO* TRANSPOSON MUTAGENESIS OF MOSS cDNA LIBRARIES

The ongoing moss mutant production is aimed at taking advantage of the high rates homologous recombination in *Physcomitrella* for targeted insertion of gene disruption constructs into the plant nuclear genome. The disruption constructs used are cDNA molecules that are mutagenised by the insertion of a *nptII* selection marker gene. In contrast to the approach taken by Nishiyama et al. (2000) that used genomic fragments, we used chimeric cDNAs for targeting genomic coding sequences. The use of cDNAs instead of genomic fragments as insertional transgenes should preferentially direct insertion events to expressed *Physcomitrella* genes and circumvent the possibility of a high frequency of insertions into non-coding regions. This strategy enabled us to develop a collection of 80,000 *Physcomitrella* plants with a high percentage of targeted mutations that saturate the genome.

To allow rapid production of chimeric transgenes for every expressed *Physcomitrella* gene an efficient transposon-based shuttle mutagenesis system was developed for reliable introduction of the selectable marker gene cassette into anonymous clones from cDNA libraries (Rak et al. 2001). The system developed utilizes a bacterial transposon which has the advantage, in contrast to eukaryotic transposons, of ensuring that the transgene stays inactive inside the plant. The resulting gene disruption libraries were subsequently used to transform *Physcomitrella* protoplasts.

The process of transposon shuttle mutagenesis of *Physcomitrella* cDNA libraries is outlined in Figure 2. cDNA prepared from RNA extracted from protonemal material, grown in liquid culture for various time periods, was normalized to establish a *Physcomitrella* cDNA library representing genes expressed during vegetative growth before the onset of differentiation (Ko 1990). To reduce redundancy among cDNAs generated from gametophores and sporophytes, the reverse transcribed mRNAs were subjected to a subtraction procedure (Egener et al. 2002). The resulting low-redundancy moss cDNA populations were ligated into the minimal vector

pUCMinIV, a 1.7 kb derivative of plasmid pUC19 from which most non-essential DNA sequences were deleted. The smaller vector reduced the possibility of potential transposon insertion targets within the vector sequence. The ligation reactions were transformed into the *E.coli* donor strain R2217, which harbours the novel conjugative plasmid pFDX3951. This plasmid carries the transposon mini-Tn1000::*nptII*, and retained the chloramphenicol resistance gene (*cm*) and the *tra* function from the original conjugative plasmid R388 (Avila and Cruz 1988). The mini-Tn1000 transposon, a minimalized derivative of the bacterial transposon Tn1000, is flanked by the transposon Tn1000 border repeat sequences required for transposition (Sherratt 1989). The *nptII* selectable marker gene in the mini-Tn1000 transposon is under the control of the *nos* (nopaline synthase) promoter and termination signals (Egener et al. 2002). In *Physcomitrella* the bacterial *nos* promoter allows strong *nptII* gene expression at levels comparable to the standard CaMV 35S promoter (Holtorf et al. 2002b).

The donor strain R2217 also harbours the plasmid pFDX3957, which contains the bacterial transposase gene *tnpA* and the kanamycin resistance gene *kan*. TnpA is inducible through an IPTG-responsive promoter. IPTG induction activates expression of the bacterial transposase gene leading to transposition of mini-Tn1000::nptII. Upon transposition a cointegrate structure is formed between the plasmids pUCMinIV and pFDX3951. Conjugation of the donor with the recipient bacterial strain R1037 leads to transfer of the cointegrate structure into the recipient cell. Upon transfer the cointegrate is resolved by activity of the IPTG-induced resolvase gene *tnpR* of the recipient cell. After conjugation and resolution cells are selected on an appropriate antibiotic-containing medium and plasmid DNA prepared from those cells harbouring the pUCMinIV::mutagenised cDNA. Resulting plasmid DNA batches are digested with the rare-cutting restriction enzyme *SdaI* yielding linear DNA that is subsequently used for transformation of *Physcomitrella* protoplasts.

This transposon shuttle mutagenesis method has the advantage that anonymous cDNA populations can be rapidly manipulated *en masse* without the need for sequence information. From a single moss cDNA library a multitude of chimeric cDNA molecules, which harbour a selection cassette that disrupts the coding sequence, can be generated.

Shuttle mutagenesis leads to randomization of transposon-integration within the cDNA sequences, as there are no preferential sites for transposon insertion (Miller 1992). Non-biased transposon integration into moss cDNAs is illustrated in Figure 3. In this case 72 independent transposon insertion sites were mapped after mutagenesis by DNA sequence analysis of random individual clones from a single representative moss cDNA clone S_PP015059353 (808 bp). The experiment demonstrates that there is no apparent bias not only for the insertion of the transposon into a certain region of the cDNA but also for orientation of the transposon relative to the cDNA. In addition, the transposition of the mini-Tn1000 transposon into the cDNA occurred only once per plasmid molecule. Very few insertions into the pUCMinIV vector backbone were discovered.

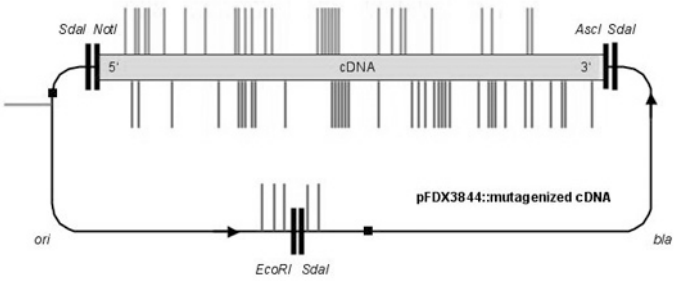

Figure 3
Distribution of transposon insertion sites in a representative *Physcomitrella patens* cDNA. Scheme depicting the orientation of 72 independent transposon insertion sites in a representative moss cDNA clone (ID: S_PP015059353; 808 bp). Prior to mutagenesis the cDNA was cloned into the minimal vector pUCMinIV. This defined plasmid was subjected to shuttle mutagenesis (see Fig. 2), and the transposon insertion sites for 72 resulting clones were mapped by DNA sequencing to assess the distribution of insertions. 41 insertions were found to be in "forward" orientation (*nptII* resistance marker on transposon and *bla* marker on vector in same orientation). "Forward" orientations are indicated by verticle lines below the cDNA, 31 "reverse" insertions by verticle lines below the cDNA. Most of the insertions (66 / 72, corresponding to 92%) occurred throughout the cDNA, without apparent strong preference for a certain insertion site or orientation. The *ori* and *bla* gene and restriction sites for *EcoRI*, *SdaI*, *AscI* and *NotI* are shown.

4.1 Fast forward production of single KO constructs

The usefulness of targeted gene knockouts for the elucidation of *Physcomitrella* gene function has recently been demonstrated (Strepp et al. 1998, Girke et al. 1998, Girod et al. 1999, Imaizumi et al. 2002, Koprivova et al. 2002, Zank et al., 2002). In these cases known genes were investigated for their function. For our program, based on the developed high-throughput transformation platform, the initial aim was to produce knockout plants for 80 candidate genes. Suitable target genes were identified on the basis of sequence homology to know3n genes by our industrial collaborator BASF Plant Science. These genes were chosen for their involvement in a variety of biochemical, developmental and signalling pathways. For each gene the corresponding EST clone was retrieved from the clone repository and disruption cassettes constructed manually. The resultant disruption constructs harboured the *nos-nptII-nos* selection cassette in an appropriate position in the central part of the cDNA. The routine transformation protocols (see subchapter 2.5.) allowed for the generation of close to 5,000 gene mutants in total, 4,898 transformants of which were analysed for knockout insertions by a high-throughput PCR-protocol using primers that flanked the predicted genomic insertion site (Fig.4). An undisrupted wild type gene yields a short, abundant PCR product whereas a much longer PCR fragment can be anticipated in the knockout plants. This longer fragment is often amplified inefficiently and thus a loss of the wild-type PCR band was taken to indicate a disrupted gene. Plants identified using this method of gene disruption were confirmation of the insertion event by Southern blot analysis. Preliminary results

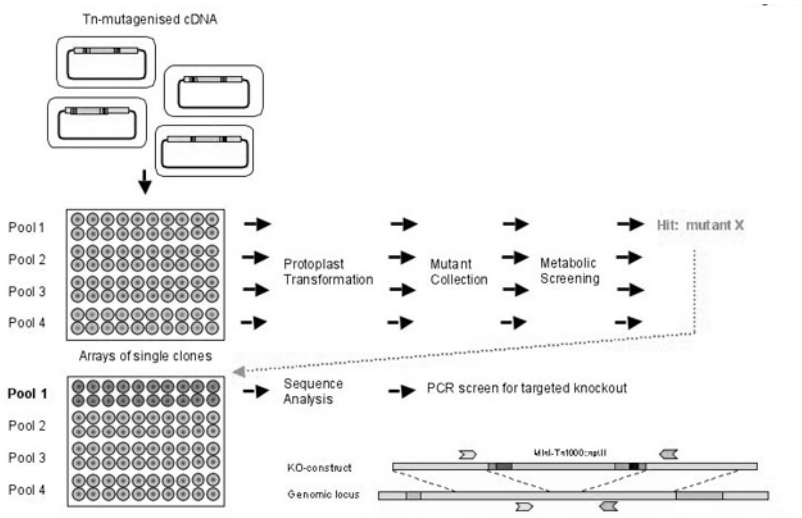

Figure 4
Production and analysis of mutants produced with batches of defined Tn-mutagenised cDNAs. Scheme outlining the production of moss mutants from batches of defined *Physcomitrella* cDNA clones. Pools containing 20 randomly selected cDNA clones are used for preparation of Tn-mutagenised plasmid DNA for protoplast transformation. After phenotypic screening mutants which show deviating metabolic profiles are selected for molecular analysis. Plasmid DNAs from clones of the corresponding pool are sequenced. Appropriate oligonucleotide primers are selected to flank the mini-Tn1000 transposon. Mutant DNA is screened for knockout insertions.

involving 61 candidate genes indicate that the average gene targeting (GT) efficiency was 27% . The highest GT efficiency observed for a single construct was 88%, however, a GT efficiency of less 5% was obtained for almost one fifth (13 of 61) of the target genes (T. Egener and J.M. Lucht, personal communication). The low GT frequency in 20% of the tested knockout constructs may be explained by the possibility that gene disruption affects viability in these cases or an unfavourable exon-intron structure may be formed within the genomic locus. As the genomic locus for the majority of the 61 candidate genes is not known, we cannot rule out nor prove either of these possibilities. We have observed, however, that gene targeting efficiency increases when using long continuous stretches of genomic sequence for the construction of the knockout transgenes (J. Granado personal communication, H. Holtorf unpublished results). Interestingly, there was no strong correlation between GT efficiencies and length of the cDNA stretches used in the knockout constructs.

The study demonstrates that routine, high-efficiency gene targeting using cDNA constructs is possible in *Physcomitrella*. A large number of candidate gene knockouts can be achieved in a relatively short period of time and mutants can be efficiently screened by PCR assays based on EST sequence fragments.

4.2. High throughput transformation protocol

The large-scale gene-function correlation study we have established is aimed at producing a collection of transgenics with insertion mutations that saturates the population of expressed *Physcomitrella* genes. To achieve this a robust transformation protocol that ensures reproducible output efficiencies and a high rate of knockout mutants must be developed. In this case, mutagenised cDNA for transformation was produced by transposon mutagenesis (Fig. 2) and the derived clones provided the transgene DNA. Plasmid DNA was prepared and transformed in batches. The mutagenised cDNA was integrated into the genome, preferentially by homologous recombination. The use of cDNA batches derived from the three different tissues mentioned earlier makes it reasonable to expect that all expressed moss genes will be disrupted with a targeting construct.

The polyethylene glycol (PEG)-based transformation protocol for direct DNA transfer to protoplasts (Strepp et al. 1998) had to be adapted in order to efficiently produce the required 80,000 mutants. Standard transformations utilize $3x10^5$ cells and 20 µg of linearized plasmid DNA. Transgenic plants are regenerated and selected on medium supplemented with various metabolites to facilitate survival of metabolic mutants (see subchapter 2.7.). The new protocol generated stable transformants in six weeks, using an initial selection on 25 µg/ml G418 for 2 weeks, a non-selective release step of 2 weeks, and an additional 2 weeks of G418 selection. Two rounds of selection on the antibiotic G418 are necessary to eliminate unstable transformants (Ashton et al. 2000).

Large-scale production of protoplast material was achieved by using a bioreactor culture protocol (see subchapter 2.1.) and improved culture conditions for optimal regeneration of protoplasts (Schween et al. 2002a). The transformation protocol was further optimized with respect to DNA conformation and pre-culture conditions which improved rates for plant output and gene targeting efficiency (Hohe et al. 2002). We compared transformation efficiencies for both circular and linear plasmids and demonstrated a difference in the yield of stable transformants. Transformation using circular DNA molecules resulted in only 0.2 % surviving plants after the second round of G418-selection, contrasting with 16 % for linear DNA. The low efficiency of stable transformation with circular DNA maybe a reflection of the finding that plasmids can be kept as extra-chromosomal replicating elements which rarely integrate into the genome (Ashton et al. 2000).

Disruption constructs from three different genes were used to generate optimal conditions for the pre-culture of plant material for transformation. Pre-culture conditions in the bioreactors were varied by supplementation with either ammonium tartrate or by controlling the pH (set point 4.5) of the culture solution. Both protocols resulted high protoplast yields but the transformation efficiencies were eight times higher in protoplasts derived from the pH controlled cultures. A pH-controlled

bioreactor pre-culture enabled the production of 48 stable transgenic moss plants per transformation. GT efficiencies mutants were measured by PCR-based screening procedures.

With these procedures it is possible to perform 18 independent transformations per day with a capacity of 9 plants per transformation, i.e. 160 plants per day. The output is highly reproducible and does not vary significantly for any of the mutagenised cDNA libraries that were used for transformation (A. Hohe personal communication). This high throughput transformation protocol will ensure the production a mutant collection of 80,000 plants, saturating the expressed genome of *Physcomitrella*.

Initially the overall number, identity, and nucleotide sequence of the different Tn-mutagenised cDNAs that made up a single batch of constructs used for transformation was not determined. This strategy hindered later characterization of the resulting transgenic plants. To correct this each batch of DNA was confined to 20 arbitrarily selected cDNA clones from a single cDNA library. The 20 clones constitute a single DNA pool. Each DNA pool and the individual clones contained therein are numbered and stored in a repository (Fig. 4). cDNA clone identification by sequence analysis is performed only when a mutant with an interesting phenotype emerges (see subchapter 2.8.) and the sequence of the 20 transgenes can be used to design oligonucleotide primers for PCR screening. Such cataloguing and tracking of clones enables a more rapid analysis of the gene lesion that causes the mutant phenotype.

4.3 Mutant quality control

To produce a mutant collection of 80,000 plants is not only challenging with respect to the number of plants that have to be produced but also with regards to mutant quality. Even routine protocols that underlie the production of standardized material cannot guarantee the absence of genetic changes that may arise during cell culture.

Two important quality traits of the generated plants are of special interest. First, all plants that have been selected on G418-containing medium must be tested to confirm the presence of the *nptII* selectable marker gene. Only these mutants can be considered as stable transformants. Second, the nuclear DNA content of all mutants must be determined in order to detect those plants that may arise from protoplast fusion events which double the chromosomal DNA content. Stable integration of the *nptII* selectable marker gene can be determined by a one-step genomic DNA extraction and a subsequent PCR assay (Schween et al. 2002c). Some of the mutants exhibit slow growth during the regeneration process which necessitated an adaptation of the DNA isolation protocol to accommodate small amounts of plant material (1 to 5mg). Furthermore, the genomic DNA preparation had to be fast and simple enough to to analyse 800 to1,000 plants per week (Schween et al. 2002c). The PCR protocol was optimized by addition of the linear multivalent polyamine spermidine and polyvinylprrolidone (PVP) to negate the inhibitory effects of polyphenols. The described PCR screen is fast and

accurate and can be used to analyse 300-400 *Physcomitrella* plants per day. In an analysis of 15,000 putative *Physcomitrella* transformants, 13.7% did not yield an *nptII*-specific PCR product. Such plants were subjected to a third round of selection on G418-containing medium. The majority of these plants survived the additional antibiotic treatment and were considered to be stable mutants. Multiple rounds of alternating selection and release treatments will only allow stably transformed plants to survive (Ashton et al. 2000). As a result 98% of plants were determined to be stable transgenics.

The PEG-transformation protocol occasionally yields polyploid plants, which may arise from protoplast fusion during the transformation process. To ensure a constant output of haploid mutants it is necessary to determine the nuclear DNA content of each transformant. Each mutant plant is required to be subjected to flow cytometric analysis (FCM). For routine mutant analysis by FCM the haploid genome of *Physcomitrella* was determined to be 511 Mbp in size (Schween et al. 2002c). In principle FCM analysis can be achieved with either protoplasts or isolated nuclei (Galbraith 1989; Ulrich and Ulrich 1991). For high throughput purposes a protocol described by Ulrich and Ulrich (1991) was adapted for *Physcomitrella*. This procedure required only 30 mg of fresh mutant tissue, from which intact nucleisuspensions were prepared by maceration of the material with a razor blade in a glass petri dish containing 2.5 mL of a DAPI-containing buffer. FCM measurements revealed that, using the standardized transformation and regeneration process, only 5% of 16,000 mutants tested were polyploid (Egener et al. 2002).

4.4 Screening for mutant phenotypes

Following the classical principle of forward genetics, all mutants produced from a Tn-mutagenised cDNA library were subjected to three different phenotypic screens: physiological, morphological, and metabolic. Screens for nutritional mutants of *Physcomitrella* have been described (Ashton and Cove 1997) and thus served as a model for an initial search for physiological alterations induced in moss mutants. To accomplish this mutant plants were regenerated and selected on both minimal and supplemented media. The supplemented media was Knop medium supplemented with MS microelements (Murashige and Skoog 1962), 4 mg/l myo-inositol, 2.8 mg/l choline chloride, 1 mg/l nicotinic acid, 0.5 mg/l thiamine-HCl, 0.25 mg/l pyridoxine, 0.01 mg/l biotin, 0.25 mg/l p-aminobenzoic acid, 1.9 mg/l Ca-D-pantothenate, 0.015 mg/l riboflavine, 6.76 mg/l adenine, 3.84 mg/l Na-palmitinic acid, 250 mg/l peptone, 920 mg/l ammonium tartrate and 50 g/l glucose.

Only 7.2% (1,163 of 16,203) of the regenerated plants experienced growth retardation when cultured on minimal medium in contrast to normal growth rates on the supplemented medium. Wild-type plants grew well on both media. The slow-growth phenotype on minimal medium suggests a metabolic defect that is rescued by supplying metabolites to the medium. A retarded growth during the culture on both minimal and supplemented meda,

observed in 3.1% (508 of 16,203) of the transformants, suggests either the absence of a required metabolite or an unknown defect that is reflected in a low growth phenotype.

To screen for developmental phenotypes, mutants were scored, relative to the wild-type, for visible morphological aberrations and growth behaviour. Digital images and microscopic observations for each mutant were recorded and archived. Mutants were scored directly following regeneration and during selection on supplemented Knop medium for 11 weeks followed by growth on minimal Knop medium for 8 weeks. Under these conditions, the filamentous protonema underwent a developmental switch resulting in the formation of a three-faced apical meristem bud. This bud differentiated into the gametophores that consist of a stem and associated leafy structures. Eventually the gametophores carried sex organs. Deviations from this normal developmental pathway were observed in 16.2% (2,631 of 16,203) of the gene-disruption library transformants. These mutants exhibited one or more of the following features: changed structure and colour of the moss plant, unusual coverage of the plant by gametophores, disturbed shape and uniformity of leaves, and altered arrangement of cells within the leaves (Egener et al. 2002). Control plants were continuously scored to ensure that the suspect phenotypes originated from the insertion of transgenes into the nuclear genome and not from some unknown factor experienced in culture. Morphological deviations did not occur spontaneously within 350 wild-type plants and in less than 1% of the 400 plants derived from protoplasts that had undergone a mock-transformation without DNA and regenerated in the absence of the antibiotic G418. In addition to scoring growth requirements and morphology the ploidy level was determined for each mutant (see subchapter 2.6.). There did not appear to be any strict correlation between any of these three parameters but various combinations of characteristics were observed (Egener et al. 2002).

Egener et al. (2002) describe an impressive collection of physiological and morphological mutant phenotypes uncovered by the described screening protocol. The frequency of altered phenotypes in the mutant collection was considerably higher than the fewer than 2% reported for Arabidopsis confirmed knockout-mutants (Bouche and Bouchez 2001), which might be attributable to haploidy and a lower degree of gene redundancy in *Physcomitrella* (Rensing et al. 2002). The reported phenotypes leads to the conclusion that a wide spectrum of morphological and developmental alterations do occur in *Physcomitrella* plants transformed with the Tn-mutagenised gene-disruption library.

Visible phenotypes that are revealed during physiological and morphological screens are scored with relative ease, however the loss of a functional gene may result in subtler and less visible alterations in the biochemical composition of a mutant. Certain metabolites can be lost; others in turn may increase in amount or be modified to give rise to new compounds. To identify metabolic mutants in the collection a third screen will focus on the comparative analysis of the composition of metabolites within the *Physcomitrella* mutant collection. This screen will take advantage of new technologies, developed by Metanomics, Berlin (www.metanomics.de), that

have emerged in the field of metabolomics. Such high-throughput approaches aim to identify gene function on the basis of a non-biased, simultaneous and rapid analysis of the metabolome (Trethewey et al. 1999; Trethewey 2001). The metabolic status of each mutant will be profiled with respect to a wide range of chemical constituents. Changed levels of certain compounds can reflect the presence or absence of gene products involved in synthesis or break-down of the compounds. Such analyses are predicted to aid in the identification of the immediate biochemical function of plant genes (Roessner et al. 2002).

Convincing examples of a link between a *Physcomitrella* gene knockout and a clear biochemical phenotype have been demonstrated (Girke et al. 1998, Zank et al. 2002, Sperling et al. 2002). These early studies demonstrated that *Physcomitrella* is a source for novel elongase and desaturase genes in the plant fatty acid biosynthetic pathway from significant changes of fatty acid profiles in knockout mutants. The alteration in fatty acid profiles allowed the clear assignment of a function to the genes under investigation.

Novel insights into another plant biochemical pathway were gained by "knocking out" the *Physcomitrella apr* gene. The phenotype of mutant the revived interest in an old proposed route for sulfate assimilation in plants (Koprivova et al. 2002).

The expectation of the metabolic analysis of the *Physcomitrella* mutant collection is that, as for the three cases mentioned above, a link between various biochemical phenotypes and the corresponding gene lesions will be established.

4.5. Molecular analysis of mutants

The overall purpose of the high-throughput, large-scale *Physcomitrella* genome project is to identify a large number of tagged *Physcomitrella* mutants that exhibit a commercially interesting phenotype. The high rate of homologous recombination, in combination with the *nptII* marker gene-tag, will enable rapid access to the disrupted gene sequence. The tagged loci are recoverble using a set of standard techniques that were adapted for molecular analysis of *Physcomitrella*. Efficiently screens for knockout insertions utilizing high-throughput PCR methods were developed (Schween et al. 2002b). In addition, by combining a newly developed DNA extraction protocol (Schlink and Reski 2002) with a non-radioactive detection method (Egener et al. 2002) the ease, speed and accuracy of gene knockout analysis by Southern blotting of genomic DNA has been improved.

Egener et al. 2002 assessed the overall quality of mutants, with respect to the number of integrated *nptII*-marker cassettes and integration loci. Molecular analyses were initially performed using Southern blots of isolated genomic DNA from randomly chosen transformants probed with suitable target sequences. This was achieved by digestion of the genomic DNA with *Pvu*II, to obtain an estimate for the number of gene-disruption constructs integrated in the transformed moss plants, since the *nptII* gene has a single *Pvu*II recognition site within its coding sequence. Restriction of this sequence

results in the formation of two hybridisation bands, which are detected by a *nptII*-derived radioactive probe. The derived hybridisation bands differ in length from mutant to mutant depending on the nature of the adjacent sequences. These sequences can be derived either from genomic sequences or from transgenic cDNA fragments. Such a Southern blot analysis of 154 transformants revealed that the average number of *nptII* genes per mutant was approximately 10 (K. Schlink, personal communication).

To evaluate the average number of insertion sites per mutant genomic DNA of 16 transformants was digested with several restriction enzymes that cut genomic *Physcomitrella* DNA frequently but do not cut within the *nptII* cassette, and only rarely within the cDNA sequences carried by the transforming DNA (Egener et al. 2002). Analysing the *nptII*-probed restriction patterns discriminates between independent integration events at different genomic positions and co-integrations at the same locus. The majority of fragments, detected after hybridisation, were more than 20 kb in size; only a few distinct *nptII*-hybridising bands were observed. This result suggested that most of the Tn-mutagenised transgenic cDNA molecules are co-integrated in close vicinity to each other at up to 3 independent genomic loci. This is in concurrence with observations that *Physcomitrella* integrates multiple transgene copies at one genomic locus, i.e. integration loci originating from a single homologous or illegitimate recombination event can carry multiple concatameric copies of transgenes (Schaefer and Zryd 1997, Nishiyama et al. 2000). Indeed, PCR amplification of several transgene sequences between the *nptII*-cassettes and subsequent sequence analysis of the products confirmed tandem integrations of different gene disruption constructs in close vicinity (Tanja Egener, Hauke Holtorf, unpublished results).

For gene-function assignment, however, the exact structure of the targeted locus is not of primary interest. Only a knowledge of the disrupted gene sequence at the border of the integration locus is relevant and needs to be determined. The challenging task is to employ suitable and rapid methods for identification of sequences that flank the insertion site. Various PCR-based methods can be used for this purpose (Ochman et al. 1993). In this project the borders were determined using a combination of PCR-based and plasmid rescue methods. For example, inverse PCR (iPCR, Hartl and Ochman 1994, Triglia et al. 2000), a combination of circularization of genomic fragments after restriction digestion and inversely oriented amplification, is one procedure used in the determination of genomic sequences that flank an inserted marker gene. Inverse PCR has been successfully used with moss genomic DNA and efficiently amplified sequences flanking the *nptII* gene. However, fragments generated by this method were small, making it difficult to unequivocally assign fragments to the gene insertion border.

A more suitable method is Thermal Asymmetric InterLaced PCR (TAIL-PCR) as described by Liu et al. 1995. TAIL-PCR is based on the combination of specific nested primers of (for example) a marker gene, with a set of nested, degenerate primers which can anneal in the genomic regions bordering the integrated transgene. TAIL-PCR is a very sensitive method that can be applied to large and complex genomes (Liu et al. 1995). This

procedure has been adapted to isolate and sequence flanking genomic sequences from *Physcomitrella* DNA.

The presence of concatameric structures at the insertion locus, however, necessitates thorough inspection of sequence data produced from amplified fragments. This problem arises from the fact that sequences from inside the concatamer, that neighbour a *nptII*-cassette, are likely to be amplified. This is also true for Arabidopsis, where TAIL-PCR has been reported to be a versatile tool for isolation and mapping of T-DNA insert junctions and other border sequences (Liu et al. 1995, He and Gan 2001, Okamoto and Hirochika 2001). TAIL-PCR will be successful as long as *nptII* marker gene sequences lie in the vicinity of the disrupted genomic locus. TAIL-PCR specific products can be used for direct sequencing (Mazars and Theillet 1996) and can be immediately used to test the integrity of genomic loci by Southern blotting.

Plasmid rescue is also a powerful tool for cloning plant sequences that flank a selection marker gene (Mathur et al. 1998, Nakazawa et al. 2001). A suitable plasmid vector has been developed for the *Physcomitrella project* that enables the trapping of sequences that flank the mini::Tn1000 transposon. The trap vector contains an incomplete N-terminal fragment of the *nptII* marker gene under the control of a bacterial promoter. To clone border sequences, genomic DNA of a mutant plant is digested with appropriate restriction enzymes and ligated into the trap vector. Upon selection with kanamycin, only a restored marker gene will confer antibiotic resistance to the bacteria. The transgenic nucleotide sequence flanking the *nptII* gene can then be determined by standard sequence analysis. Using this method several border sequences have been cloned (Tanja Egener, personal communication).

Plasmid rescue and TAIL-PCR protocols have successfully been used to assign border sequences to insertion sites in tagged mutants of Arabidopsis (McElver et al. 2001, Budziszewski et al. 2001). More than 25 genomic loci from *Physcomitrella* knockout mutants generated with known and anonymous disruption constructs have been isolated to date. Border sequence analysis was performed by comparing sequences of PCR products or fragments obtained by plasmid rescue with the public databases and our proprietary EST-database using BLAST (Altschul et al. 1997). Currently TAIL-PCR and plasmid rescue are the methods of choice to allow rapid isolation of tagged genomic loci in *Physcomitrella*.

Analysis of mutants from transformations, generated from defined cDNA pools (see subchapter 2.5.), is performed by the use the fast and comprehensive PCR-based screen to identify genes with a putative knockout insertion described here. Prior to the application of the protocols for isolation of tagged genes, each clone within the arrayed pool of clones is subjected to sequence analysis to identify the gene sequence and location of the transposon. It is the ability to identify the precise location of the transposon that is the advantage to using defined Tn-mutagenised cDNA batches: TO generate each mutant a defined pool of knockout constructs are used for transformation, and for each construct the position of the transposon is known, and only a maximum of 20 transgenes potentially integrate into the genome. The knowledge of the number and sequence of inserted transgenes allows for a much faster molecular analysis. Much like the candidate gene

knockout approach (see subchapter 2.4.), primers are chosen such that they flank the transposon sequence of the disruption construct and genomic mutant DNA is subjected to PCR screening to identify the loss of an amplification product that is specific for the wild-type gene (Fig.4). All 20 transgenes can be rapidly screened for a knockout integration in the corresponding gene. The knowledge of the construct used to generate a disrupted wild-type gene can subsequently be used in the design of a Southern analysis. A band shift in the mutant relative to the wild-type is indicative of a knockout insertion at the gene in question. For gene isolation, genomic knockout mutant DNA can be subsequently analysed using TAIL-PCR and plasmid methods.

5. OUTLOOK

The moss *Physcomitrella patens* has emerged as an attractive model system for plant biology and functional genomic analysis. It shares many biological features with higher plants but has the unique ability for efficient homologous recombination within its nuclear DNA. In comparison to other plant mode1s this advantage allows for precise genetic manipulation that can routinely produce targeted knockouts to study plant gene function. We have demonstrated that *Physcomitrella* is amenable to the large-scale functional genomics approache that allows the innovative creation, and efficient characterization, of plant genetic diversity to determine novel gene functions. The knowledge generated by such a model has the potential for application in crop plant improvement, especially in light of the high degree of genetic and physiological conservation between mosses and higher plants.

Plant Biotechnology, Albert-Ludwigs-Universität Freiburg,
Schänzlestrasse 1,
79106 Freiburg, Germany (RR; reski@uni-freiburg.de)

6. REFERENCES

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.,* 25, 3389-3402

Ashton, N.W., & Cove, D.J. (1977) The isolation and preliminary characterisation of auxotrophic and analogue resistant mutants of the moss *Physcomitrella patens*. *Mol. and Gen. Genet.,* 154, 87-95

Ashton, N.W., Champagne, C.E.M., Weiler, T., & Verkoczy, L.K. (2000) The bryophyte *Physcomitrella patens* replicates extrachromosomal transgenic elements. *New Phytol.,* 146, 391-402

Avila, P, & de la Cruz, F. (1988) Physical and genetic map of the IncW plasmid R388. *Plasmid,* 20, 155-157

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

Boyd, P.J., Hall, J., & Cove, D.J., (1988) An airlift fermenter for the culture of the moss *Physcomitrella* patens. In *Methods in bryology*, Proc. Bryol. Meth. Workshop (Mainz, J. & M. Glime eds). Nichinan: *Journal of the Hattori Botanical Laboratory*, 41-45

Budziszewski, G.J., Lewis, S.P., Wegrich-Glover, L., Reineke, J., Jones, G., Schlater-Ziemnik, L., Lonowski, J., Nyfeler, B., Aux, G., Zhou, Q., McElver, J., Patton, D.A., Martienssen, R.,

Grossniklaus, U., Ma, H., Law, M., & Levin, J.Z. (2001) Arabidopsis Genes Essential for Seedling Viability: Isolation of Insertional Mutants and Molecular Cloning. *Genetics* 159, 1765-1778

Cove, D.J., Knight, C.D., & Lamparter, T.L. (1997) Mosses as model systems. *Trends Plant Sci.,* 2, 99-105

Egener, T., Granado, J., Guitton, M.C., Hohe, A., Holtorf, H., Lucht, J.M., Rensing, S., Schlink, K., Schulte, J., Schween, G., Zimmermann, S., Duwenig, E., Rak, B., & Reski, R. (2002) High frequency of phenotypic deviations in *Physcomitrella patens* plants transformed with a gene-disruption library. *BioMed. Cen. Plant Biol*., 2, 6

Galbraith, D.W. (1989) Analysis of higher plants by flow cytometry and cell sorting. *Int. Rev. Cytol*. 116, 165-228

Girke, T., Schmidt, H., Zaehringer, U., Reski, R., & Heinz, E. (1998) Identification of a novel delta 6-acyl group desaturase by targeted gene disruption in *Physcomitrella patens*. *Plant J.* 15, 39-48

Girod, P.A., Fu, H., Zryd, J-P., & Viestra, R.D.(1999). Multiubiquitin chain binding subunit MCBI (RPN10) of the 26S proteasome is essential for developmental progression in *Physcomitrella patens*. *Plant Cell* 11, 1457-1472

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley. D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.L., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A., & Briggs, S. (2002) A draft sequence of the rice genome (L. ssp. *japonica*). *Science* 296, 92-100

Grimsley, N.H., Ashton, N.W., & Cove, D.J. (1977). The production of somatic hybrids by protoplast fusion in the moss *Physcomitrella patens*. *Mol. Gen. Genet*. 154:97-100

Hartl, D.L., & Ochman, H. (1994) Inverse polymerase chain reaction. *Methods Mol. Biol.*, 31, 187-96

He, Y., & Gan, S. (2001) Identical promoter elements are involved in regulation of the OPR1 gene by senescence and jasmonic acid in Arabidopsis. *Plant Mol. Biol*., 47, 595-605

Hohe, A., & Reski, R. (2002) Optimisation of a bioreactor culture of the moss *Physcomitrella patens* for mass production of protoplasts. *Plant Sci.,* 163, 69-74

Hohe, A., Schween, G., & Reski, R. (2001) Establishment of a semicontinuous bioreactor culture of *Physcomitrella patens* for mass production of protoplasts. *Acta Horticulturae.* 560, 425-428

Hohe, A., Egener, T., Lucht, J.M., Holtorf, H., Reinhard, C., Schween, G,, & Reski, R. (2002) DNA conformation and plant pre-culture influence both, transformation efficiency and gene targeting efficiency in *Physcomitrella*. *Submitted*.

Holtorf, H., Guitton, M.C., & Reski, R. (2002a) Plant Functional Genomics. *Naturwissenschaften* 89, 235-249

Holtorf, H., Hohe, A., Wang, H-L., Jugold, M., Rausch, T., Duwenig, E., & Reski, R. (2002b) Promoter subfragments of the sugar beet V-type H^+-ATPase subunit c isoform drive the expression of transgenes in the moss *Physcomitrella*. *Plant Cell Rep.,* In press.

Imaizumi, T., Kadota, A., Hasebe, M., & Wada, M. (2002) Cryptochrome light signals control development to suppress auxin sensitivity in the moss *Physcomitrella patens*. *Plant Cell* 14, 373-386

Jenkins, G.I., & Cove, D.J. (1983) Light requirements for regeneration of protoplasts of the moss *Physcomitrella patens*. *Planta* 157, 39-45

Ko, M.S. (1990) An 'equalized cDNA library' by the reassociation of short double- stranded cDNAs. Nucl. Acids Res., 18, 5705-5711

Koprivova, A., Meyer, A.J., Schween, G., Herschbach, C., Reski, R., & Kopriva, S. (2002) Functional knockout of the adenosine 5'-phosphosulfate reductase gene in *Physcomitrella patens* revives an old route of sulfate assimilation. *J Biol. Chem*. 277, 32195-32201

Lal M (1984) The culture of bryophytes including apogamy, apospory, parthenogenesis and protoplasts. In: *The Experimental Biology of Bryophytes*, (Dyer A.F. &. Duckett J.G eds.). London, Academic Press: 97-115

Liu, Y.G., Mitsukawa, N., Oosumi, T., & Whittier, R.F. (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* 8, 457-463

Machuka, J., Bashiardes, S., Ruben, E., Spooner, K., Cuming, A., Knight, C., & Cove, D. (1999). Sequence analysis of expressed sequence tags from an ABA-treated cDNA library identifies stress response genes in the moss *Physcomitrella patens*. *Plant Cell Phys.*, 40,378-387

Mathur, J., Szabados, L., Schaefer, S., Grunenberg, B., Lossow, A., Jonas-Straube, E., Schell, J., Koncz, C., & Koncz-Kalman, Z. (1998) Gene identification with sequenced T-DNA tags generated by transformation of Arabidopsis cell suspension. *Plant J.* 13, 707-16

Mazars, G.R., & Theillet, C. (1996) Direct sequencing by thermal asymmetric PCR. *Methods Mol. Biol.* 65, 35-40

McElver, J., Tzafrir, I., Aux, G., Rogers, R., Ashby, C., Smith, K., Thomas, C., Schetter, A., Zhou, Q., Cushman, M.A., Tossberg, J., Nickle, T., Levin, J.Z., Law, M., Meinke, D., & Patton, D. (2001) Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana. Genetics* 159, 1751-1763

Miller J.H. (1992) A short course in bacterial genetics. *Cold Spring Harbor, N. Y.*: Cold Spring Harbor Laboratory Press.

Murashige, T., & Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue culture. *Physiol. Plant*. 15:473–497

Nakazawa, M, Yabe, N., Ichikawa, T., Yamamoto, Y.Y., Yoshizumi, T., Hasunuma, K., & Matsui, M. (2001) DFL1, an auxin-responsive GH3 gene homologue, negatively regulates shoot cell elongation and lateral root formation, and positively regulates the light response of hypocotyl length. *Plant J.*, 25, 213-221

Nishiyama, T., Hiwatashi, Y., Sakakibara, I., Kato, M., & Hasebe, M. (2000) Tagged mutagensesis and gene-trap in the moss, *Physcomitrella patens* by shuttle mutagenesis. *DNA Res*. 7, 9-17

Ochman, H., Ayala, F.J., & Hartl, D.L. (1993) Use of polymerase chain reaction to amplify segments outside boundaries of known sequences. *Methods Enzymol.* 218, 309-321

Okamoto, H., & Hirochika, H. (2000) Efficient insertion mutagenesis of Arabidopsis by tissue culture-induced activation of the tobacco retrotransposon Tto1. *Plant J.*, 23, 291-304

Rak, B., Reski, R., Zimmermann, S., Guitton, M.C., Duwenig, E., & Freund, A. (2001) Method for the mutagenesis of nucleotide sequences from plants, algae and fungi. *WO 01/38509 A1*

Rensing, S.A., Rombauts, S., Hohe, A., Lang, D., Duwenig, E., Rouze, P., Van de Peer, Y., Reski, R. 2002a) The transcriptome of the moss *Physcomitrella patens*: comparative analysis reveals a rich source of new genes. Available online: http://www.plant-biotech.net/Rensing_et_al_transcriptome2002.pdf.

Rensing, S.A., Rombauts,S., Van de Peer, Y., & Reski, R. (2002b): Moss transcriptome and beyond. *Trends Plant Sci*.., in press.

Reski, R., Faust, M., Wang, X.H., Wehe, M., & Abel, W.O. (1994). Genome analysis of the moss *Physcomitrella patens* (Hedw.) BSG. *Mol. Gen. Genet.* 244, 352-359.

Reski, R. (1998a). Development, genetics and molecular biology of mosses. *Bot. Acta.* 111, 1-15

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

Reski, R. (1999). Molecular genetics of *Physcomitrella*. *Planta*, 208, 301-309

Reski, R. & Abel, W.O. (1985) Induction of budding on chloronemata and caulonemata of the moss, *Physcomitrella patens*, using isopentenyladenin. *Planta*, 165, 354-358

Reski, R. Reynolds, S., Wehe, M., Kleber-Janke, T., & Kruse, S. (1998) Moss (*Physcomitrella patens*) expressed sequence tags include several sequences which are novel for plants. *Bot. Acta* 111, 145-151

Reutter, K., Atzorn, R., Hadeler, B., Schmülling, T., & Reski, R. (1998). Expression of the bacterial ipt gene in *Physcomitrella* rescues mutations in *budding* and *plastid division*. *Planta*, 206, 196-203

Reutter, K., & Reski, R. (1996) Production of a heterologous protein in bioreactor cultures of fully differentiated moss plants. *Plant Tiss. Cult. Biotech.* 2, 142-147

Roessner, U., Willmitzer, L., & Fernie, A.R. (2002) Metabolic profiling and biochemical phenotyping of plant systems. *Plant Cell Rep*., In press

Rother, S., Hadeler, B., Orsini, J.M., Abel, W.O., & Reski, R. (1994). Fate of a mutant machrochloroplast in somatic hybrids. *J. Plant Physio*l., 143:72-77

Schaefer, D.G. (2001) Gene targeting in *Physcomitrella patens*. *Curr. Opin. Plant Biol*., 4, 138-141

Schaefer, D.G. (2002). A new moss genetics: targeted mutagenesis in *Physcomitrella patens*. *Ann. Rev. Plant Biol*., 53, 477-501

Schaefer, D.G., & Zryd, J.P. (1997) Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J*., 11, 1195-1206

Schaefer, D.G., Zryd, J.P., Knight, C.D., & Cove, D.J. (1991) Stable transformation of the moss *Physcomitrella patens*. *Mol. Gen. Genet*., 226, 418-424.

Schlink, K., & Reski, R. (2002): Preparing high-quality DNA from moss (*Physcomitrella patens)*. *Plant Mol. Biol. Rep.*, 20, In press

Schween, G., Hohe, A., Koprivova, A., & Reski, R. (2002a): Effects of nutrients, cell density and culture techniques on protoplast regeneration and early protonema development in *Physcomitrella patens*. *J. Plant Physiol*, In press.

Schween, G., Gorr, G., & Reski, R. (2002b) Unique tissue-specific G2/M-arrest in *Physcomitrella. Submitted*

Schween, G., Fleig, S., & Reski, R. (2002c). High-throughput-PCR screen of 15,000 transgenic *Physcomitrella* plants. *Plant Mol. Biol. Rep.* 20, 43-47

Sherratt, D.J. (1989) Tn3 and related transposable elements: site-specific recombination and transposition. In: *Mobile DNA* (Berg, D.E., & Howe, M.M. eds) Washington, D.C: American Society for Microbiology:163-184

Sperling, P., Egener, T., Lucht, J.M., Reski, R., Cirpus, P., & Heinz, E. (2002): Identification of a Δ5-fatty acid desaturase from *Physcomitrella patens*. *Advanced Researches on Plant Lipids*, In press.

Strepp, R., Scholz, S., Kruse, S., Speth, V., & Reski, R. (1998). 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*. 95, 4368-4373

Stumm, I., Meyer, Y., & Abel, W.O. (1975) Regeneration of the moss *Physcomitrella patens* (Hedw.) from isolated protoplasts. *Plants Sci. Lett*. 5, 113-118

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796-815

Theissen, G., Münster, T., & Henschel, K. (1999) Why don't mosses flower? *New Phytol.,* 150, 1-8

Trethewey, R.N., Krotzky, A.J., & Willmitzer, L. (1999) Metabolic profiling: a rosetta stone for genomics? *Curr. Opin. Plant Bio.l* 2,83-85

Trethewey, R.N. (2001) Gene discovery via metabolic profiling. *Curr Opin.Biotech*., 12, 135-138

Triglia, T. (2000) Inverse PCR (IPCR) for obtaining promoter sequence. *Methods Mol. Biol*. 130, 79-83

Ulrich, I., & Ulrich, W. (1991) High-resolution flow cytometry of nuclear DNA in higher plants. *Protoplasma,* 165, 212-215

Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., & Zhang, X.et al., (2002) A draft of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296, 79-92

Zank, T.K., Zahringer, U., Beckmann, C., Pohnert, G., Boland, W., Holtorf, H., Reski, R., Lerchl, J., & Heinz, E. (2002) Cloning and functional characterisation of an enzyme involved in the elongation of Δ6-polyunsaturated fatty acids from the moss *Physcomitrella patens*. *Plant J.* 31, 255-268