

Research article

Open Access

## Quantitative promoter analysis in *Physcomitrella patens*: a set of plant vectors activating gene expression within three orders of magnitude

Verena Horstmann<sup>1</sup>, Claudia M Huether<sup>1</sup>, Wolfgang Jost<sup>2</sup>, Ralf Reski<sup>1</sup> and Eva L Decker\*<sup>1</sup>

Address: <sup>1</sup>University of Freiburg, Plant Biotechnology, Schaezlestr. 1, 79104 Freiburg, Germany and <sup>2</sup>greenovation Biotech GmbH, Boetzing Str. 29b, 79111 Freiburg, Germany

Email: Verena Horstmann - verena.horstmann@biologie.uni-freiburg.de; Claudia M Huether - claudia.huether@biologie.uni-freiburg.de; Wolfgang Jost - wjost@greenovation.com; Ralf Reski - ralf.reski@biologie.uni-freiburg.de; Eva L Decker\* - eva.decker@biologie.uni-freiburg.de

\* Corresponding author

Published: 07 July 2004

Received: 18 February 2004

BMC Biotechnology 2004, 4:13 doi:10.1186/1472-6750-4-13

Accepted: 07 July 2004

This article is available from: <http://www.biomedcentral.com/1472-6750/4/13>

© 2004 Horstmann et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

### Abstract

**Background:** In addition to studies of plant gene function and developmental analyses, plant biotechnological use is largely dependent upon transgenic technologies. The moss *Physcomitrella patens* has become an exciting model system for studying plant molecular processes due to an exceptionally high rate of nuclear gene targeting by homologous recombination compared with other plants. However, its use in transgenic approaches requires expression vectors that incorporate sufficiently strong promoters. To satisfy this requirement, a set of plant expression vectors was constructed and equipped with either heterologous or endogenous promoters.

**Results:** Promoter activity was quantified using the dual-luciferase reporter assay system. The eight different heterologous promoter constructs tested exhibited expression levels spanning three orders of magnitude. Of these, the complete rice actin I gene promoter showed the highest activity in *Physcomitrella*, followed by a truncated version of this promoter and three different versions of the cauliflower mosaic virus 35S promoter. In contrast, the *Agrobacterium tumefaciens* nopaline synthase promoter induced transcription rather weakly. Constructs including promoters commonly used in mammalian expression systems also proved to be functional in *Physcomitrella*. In addition, the 5' -regions of two *Physcomitrella* glycosyltransferases (i.e.  $\alpha$ 1,3-fucosyltransferase and  $\beta$ 1,2-xylosyltransferase) were identified and functionally characterised in comparison to the heterologous promoters. Furthermore, motifs responsible for enhancement of translation efficiency – such as the TMV omega element and a modified sequence directly prior the start codon – were tested in this model.

**Conclusion:** We developed a vector set that enables gene expression studies, both in lower and higher land plants, thus providing valuable tools applicable in both basic and applied molecular research.

## Background

The moss *Physcomitrella patens* (Hedw.) B.S.G. is characterised by its simple development and well-defined differentiation steps, including a dominant haploid gametophyte. *Physcomitrella* has proven to be an appropriate model system for studying gene function in molecular and cellular development as well as for analysis of differentiation processes in plants. The suitability of this model system is attributable to it being capable of undergoing highly efficient homologous recombination events. This characteristic allows for gene function analysis by targeted knockout [1-3]. Furthermore, mosses are acknowledged as comparable to higher plants in terms of gene content, expression, and regulation [4], although they are estimated to have diverged from other groups of land plants approximately 450 million years ago [5]. To enable functional genomic approaches, EST databases covering more than 95% of the *Physcomitrella* transcriptome have been created [6,7] and two different approaches of targeted mutagenesis via transposon tagging have been performed [8,9]. In previous studies, the influence of phytohormones in *Physcomitrella* was examined by physiological analyses [10,11] and in a recent survey the auxin distribution in *Physcomitrella* was characterised with the help of transgenic plants [12].

In addition to the extensive basic research employing *Physcomitrella*, this system represents a promising model for applied approaches. For instance, a saturated mutant collection has been created to uncover targets for crop plant improvement [8,13]. Additionally, *Physcomitrella* has been used as a bioreactor for the production of complex biopharmaceuticals [14] <http://www.greenovation.de>. *Physcomitrella* is suitable for low cost and high volume production of recombinant proteins. This system has the added capability to extensively posttranslationally process proteins, including the formation of disulfide bridges and complex glycosylation [15].

For basic as well as applied research, potent and flexible expression systems are necessary. Transgenic technologies are dependent upon genetic tools among which sufficiently strong promoters for the construction of expression vectors are very important. As hardly any promoter sequences of moss genes are currently available and only a few heterologous promoters have been reported to function reliably in *Physcomitrella*, moss researchers must use the latter for expression of interesting transgenes [e.g. [16]]. Specifically, systems involving various gene cassettes require different promoter sequences to minimise molecular interaction by recombination processes between the different constructs. There is also preliminary evidence that two transgenes driven by the same promoter may be silenced by a mechanism that remains to be determined [17].

In other moss studies based upon transient or stable transformations, the ability of certain promoters to drive the expression of widely used plant-selectable markers and reporter genes have been demonstrated. Among these are the neomycin phosphotransferase gene *nptII* [e.g. [1,8]], the hygromycin phosphotransferase gene *hpt* [18], as well as the  $\beta$ -glucuronidase (GUS; *uidA*) gene [12,19], the green fluorescent protein (*gfp*) [e.g. [20]] and the firefly luciferase (*luc*) gene [16,21].

Among the heterologous promoters analysed so far are the cauliflower mosaic virus (CaMV)35S promoter, the *Agrobacterium tumefaciens* nopaline synthase (*nos*) promoter, the rice actin1 gene (*Act1*) promoter, a soybean heat-shock promoter and an artificial TOP10 promoter [summarised in [16]].

Although these studies address the individual strengths of the various promoters, there is a fundamental lack of a quantitative comparison between them. To identify sufficiently strong promoters that can drive expression of transgenes in *Physcomitrella*, Holtorf et al. [16], for example studied promoter activity in relation to the number of independent transformants surviving the selection process.

In this study we compared suitability of different heterologous and endogenous promoters to express foreign genes in *Physcomitrella patens*. Additionally, the effect of two translational enhancement motifs (the TMV omega element and a modified sequence directly prior the start codon) were analysed.

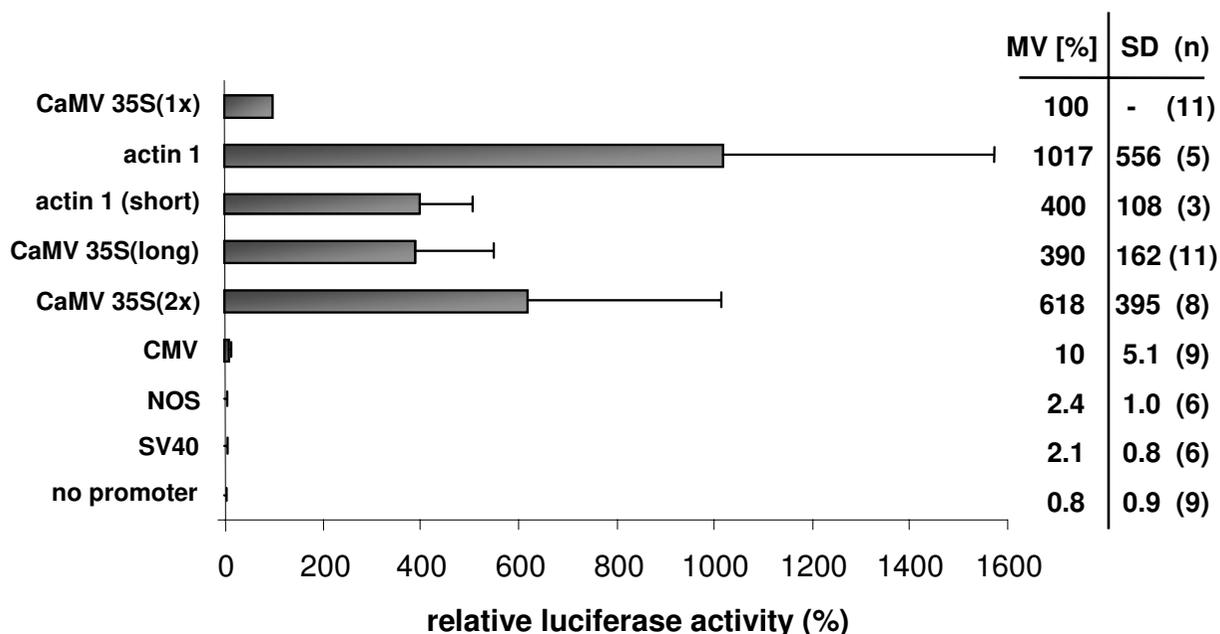
## Results and Discussion

### Quantitative analysis of heterologous promoter activities in transiently transfected *Physcomitrella* protoplasts

We quantified the ability of heterologous and endogenous promoters in *Physcomitrella* to drive foreign gene expression in plants. For this purpose, a set of diverse reporter constructs was created. The final basic constructs used in this study are shown in Fig. 1A.

Firefly (*Photinus pyralis*) luciferase was used as a highly sensitive reporter enzyme to quantify gene expression in a transient expression assay. For normalisation of transfection efficiency, *Renilla reniformis* luciferase was used. The dual-luciferase assay, which is commonly used in mammalian systems [e.g. [22]], was improved for use in a plant system by constructing a *Renilla* control plasmid in which the reporter gene is driven by the CaMV 35S promoter (Fig. 1B). To our knowledge, this dual-luciferase assay has been adapted for use in a plant system for the first time in our current study.





**Figure 2**  
**Comparison of different heterologous promoters in *Physcomitrella patens*.** Firefly luciferase light emission was normalised by *Renilla* luciferase induced by the control plasmid. The strengths of various promoters are shown in relation to the activity of the CaMV 35S promoter, which was set as 100%. Each column represents the mean value (MV) of at least three independent experiments (n: number of replication) where all samples were measured in triplicates. SD: standard deviation.

In addition to the 1230 bp complete rice *Act1* promoter, activity of a truncated version was analysed. This truncated form is missing 374 bp of the 5'-end resulting in a loss of about 60% in promoter activity, compared with the complete *Act1* promoter (Fig. 2). Similar results have been obtained previously by McElroy et al. [23] in transiently transformed rice protoplasts where the complete promoter possessed twice the strength of the cut version.

The CaMV 35S(2x) promoter also supported strong reporter expression, approximately six times higher than that of the CaMV 35S(1x) promoter (Fig. 2). In transformed tobacco plants, Kay et al. [25] observed an approximately tenfold higher transcriptional activity of a CaMV 35S variant constructed by tandem duplication of 250 base pairs of the transcription-activating sequences upstream of the TATA element. The CaMV 35S(long) promoter was about four times as effective as the single 35S promoter and therefore comparable to the cut rice *Act1* promoter (Fig. 2).

The *nos* promoter of the *Agrobacterium tumefaciens* nopaline synthase gene exhibited an activity of about

2.4% compared with activity of the CaMV 35S(1x) promoter. The *nos* promoter has been commonly used to drive reporter genes or selectable marker genes in *Physcomitrella patens* [e.g. [8,17]], resulting in sufficient expression levels. The relationship observed between the strength of the *nos* promoter and the 35S promoter in *Physcomitrella* is similar to observations made in petunia plants [26] where the *nos* promoter was 30 times less active than the 35S promoter.

The quantification of the different promoter activities shows that the ratios between these activities are comparable between *Physcomitrella* and higher plants. This confirms the observation that, despite the evolutionary distance between mosses and seed plants, the basic organisation and the regulatory mechanisms of *Physcomitrella* are similar to those of higher land plants.

The human cytomegalo virus (CMV) early promoter exhibited about 10 percent of the activity of the single 35S promoter. The simian virus (SV) 40 promoter resulted in an expression rate of about 2 percent compared with the 35S(1x) promoter. This implies that in *Physcomitrella*,

promoters that have commonly been used for transfection of human or animal cells, are stronger than the *nos* promoter or at least comparable to it and therefore offer an appropriate alternative for foreign gene expression in *Physcomitrella*.

#### **TMV omega element does not mediate translation enhancement in *Physcomitrella***

The TMV (tobacco mosaic virus) omega element has been demonstrated to be a translational enhancer in both eukaryotes and prokaryotes, increasing translation both *in vivo* and *in vitro* [27]. However, various experiments illustrated that the effects of the omega element are highly species-specific. Enhancement was strongest in dicotyledonous plant cells and moderate in cultured mammalian cells as well as in monocotyledonous plant cells [27]. An effect of the omega element on yeast translation was demonstrated in the presence of the tobacco heat shock protein HSP101, which binds to the omega element mRNA [28]. Results describing the effects of the TMV omega element in lower land plants have not been reported thus far.

We analysed the translation enhancing effect of the TMV omega element in *Physcomitrella* by introducing the element directly downstream of the CaMV 35S(long) promoter and the CMV promoter. Neither the quite strong long version of the 35S promoter nor the relatively weak CMV promoter exhibited a significantly augmented reporter signal with the introduction of this element (Fig. 3).

Previous studies have shown that activity of the CaMV 35S promoter was enhanced two- to threefold with the addition of the translational enhancer TMV omega element in stably transformed *Arabidopsis thaliana* plants [29] as well as in transiently transformed tobacco protoplasts [30].

Prior studies indicate that the enhancing effect of this omega element is related to reduction of secondary structures at the 5'-end of the RNA [30]. Gallie [31] pointed out that the reason underlying the omega's element's translational enhancement capability is its functional similarity to a 5'-cap and a poly(A) tail. This characteristic would allow the omega element to recruit the eukaryotic initiation factor eIF4E, thus enhancing translation from mRNA. So, if *Physcomitrella* already has an effective translation mechanism, it might not need the structures provided by the omega element. Gallie et al. [32] assumed that specific functional differences between ribosomes could explain the variation of the omega element's enhancing potential in various organisms. According to this theory, functional features of *Physcomitrella* ribosomes, in contrast to other eukaryotic ribosomes, could be responsible for the lacking translational enhancer effect. It is also possible that a

translation enhancing effect of the omega element in *Physcomitrella* would require an additional mRNA-binding protein as it was shown for yeast [28].

#### **Modifying the start codon context slightly enhances expression in *Physcomitrella***

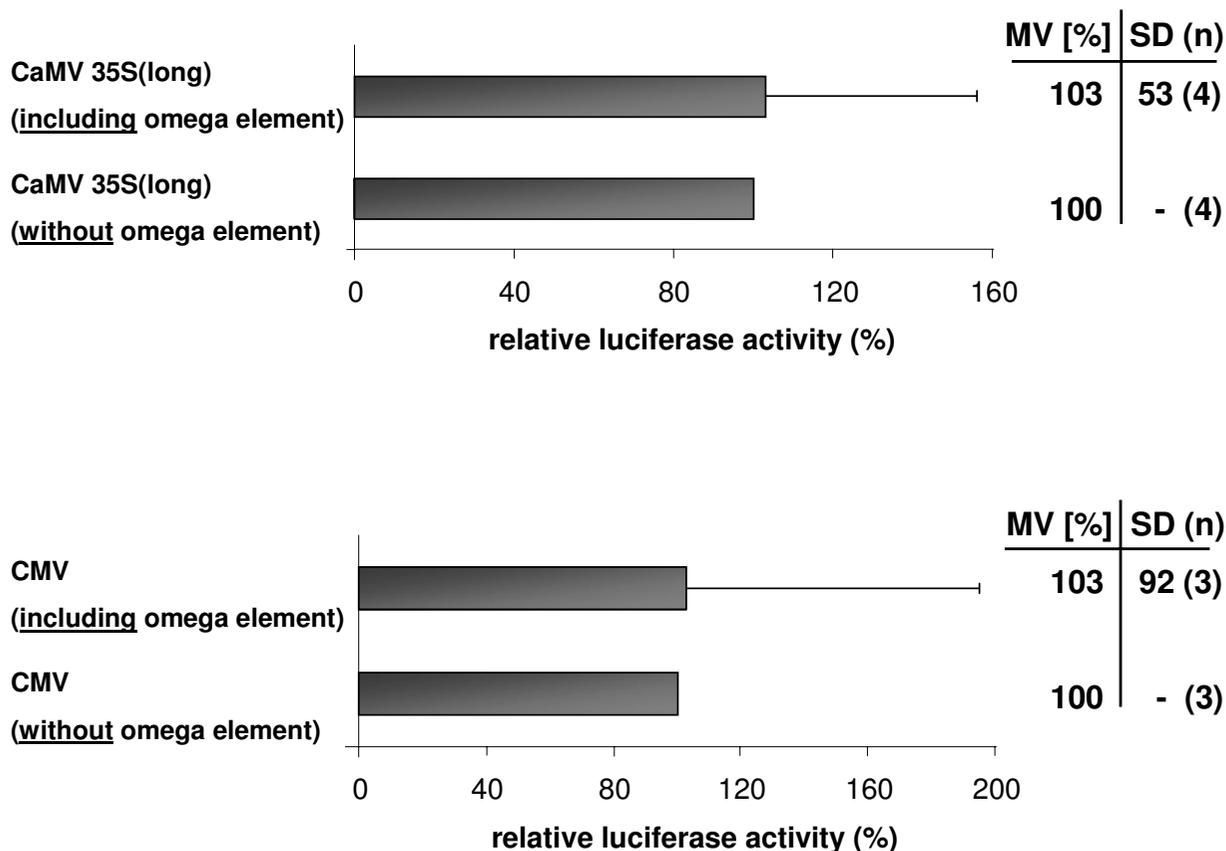
The sequences flanking the initiation codon in eukaryotic mRNAs have been described to be conserved to a certain extent. By analysing 5'-non-coding sequences of various vertebrate mRNAs, Kozak [33] described a consensus sequence, which was composed of the most frequently observed nucleotides in positions -6 to +1 (the A of the AUG codon providing position +1). For vertebrates the consensus sequence was revealed to be GCC(A/G)CCAUGG [33] and for plants AACCAAUGGC [34]. Joshi et al. [35] performed an extensive study identifying the consensus context for various categories of plants such as monocots, dicots, lower and higher plants. They suggest therefore that the derived sequence CAA(A/C)AAAUGGCG is the general plant consensus sequence, whereas in lower plants the consensus sequence is C(A/C)A(A/C)AAAUGGC(C/G).

The bases at the -3 (purine) and +4 (G) positions are considered to be the most critical ones for initiation codon selection in plants and mammals [34,36,37]. Mutations at these positions modulated initiation of translation over a 20-fold range [36].

In the context of our promoter studies we changed the ATG-flanking sequence by introducing five additional nucleotides in front of the initiation codon. Consequently, the original sequence AGCCATATGG was altered to AACCAATGG. The change from A to C at position -1 has been shown to be of little consequence for transient transgene expression in tobacco plants [38].

*Physcomitrella* protoplasts transfected with the single 35S promoter and the 35S(long) promoter, including this modified sequence around the ATG, showed increased expression levels by about 52% and about 33%, respectively (Fig. 4). The relatively weak effect of the modified sequence in our analyses may be explained by the original sequence of the plasmid, which had a G in the most critical positions -3 and +4. Thus, these constructs do not differ very much from the consensus sequence described by Kozak [33].

Based upon wheat germ in-vitro translation experiments, Luetcke et al. [34] suggested a less important role for the AUG context in plants than in other organisms. Structural and functional findings of that study indicate that the factors that select AUG initiation codons in plants and animals differ significantly.



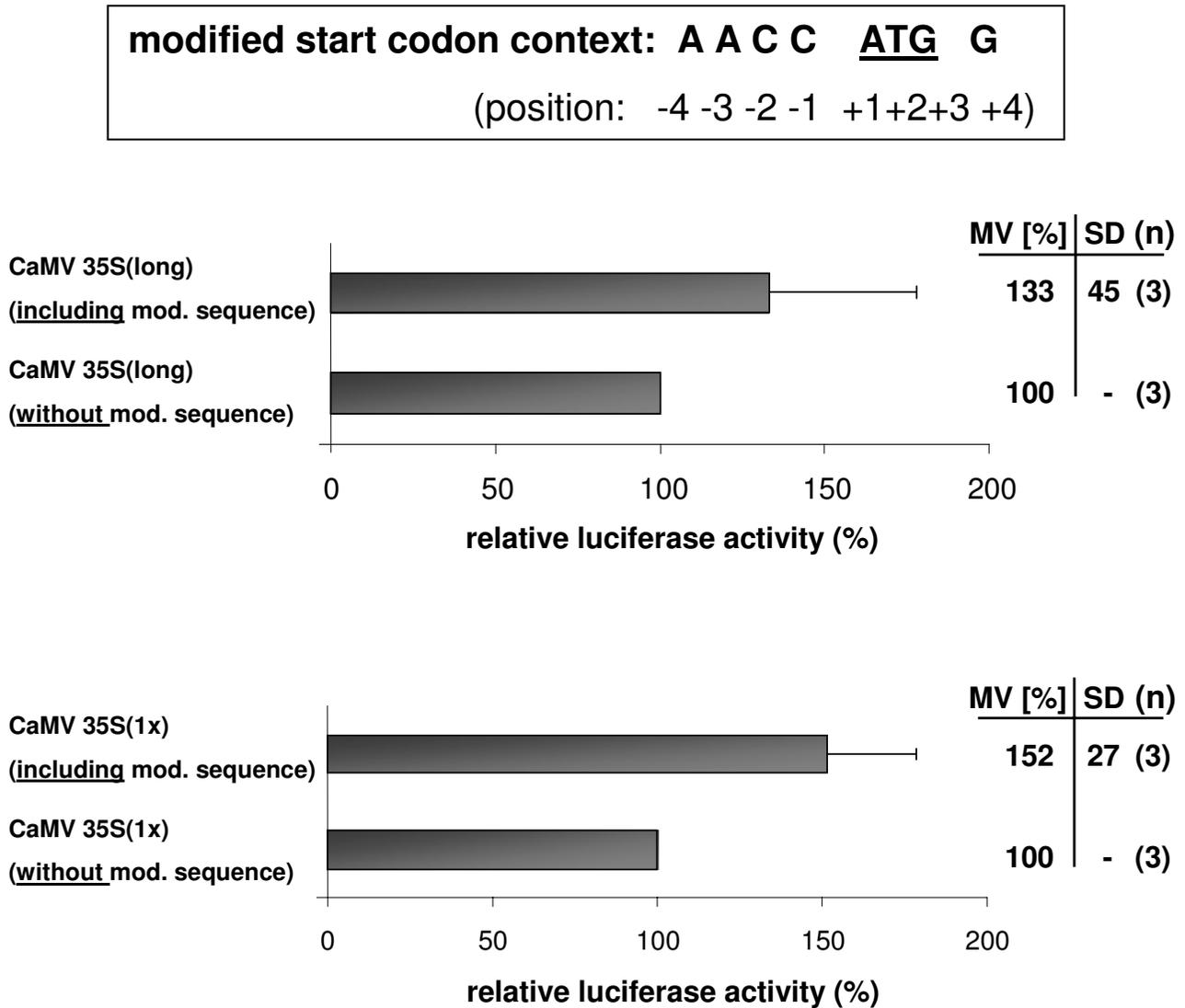
**Figure 3**  
**The TMV omega element did not mediate translation enhancement in *Physcomitrella patens*.** The expression strengths of the promoters CaMV 35S(long) and CMV with an additional omega element are shown in relation to the expression of these promoters without the enhancer. The expression rate driven without enhancer was set as 100%. Mean values (MV), standard deviations (SD) and numbers of replication (n) are indicated.

Other factors such as 5'-untranslated leader structures, upstream AUG codons, and sensitivity of plant ribosomes to secondary structure [35,39], play an important role in translation efficiency as well. The fact that these various promoter studies are based upon diverse plasmids resulting in different structural conditions should be taken into account when comparing their results.

**Characterisation of endogenous promoters in *Physcomitrella patens***

In addition to the heterologous promoters, we characterised the activity of two native *Physcomitrella* promoters,

namely  $\alpha$ 1,3-fucosyltransferase and  $\beta$ 1,2-xylosyltransferase (*fuc-t*, *xyl-t*). Four consecutive inverse PCRs were carried out to determine 2087 bp of the 5'-flanking region of the *fuc-t* sequence (EMBL accession no. AJ618932, see Additional file 1), whereas 1430 bp of the 5'-flanking-sequence of the *xyl-t* gene were identified by a single inverse PCR using genomic DNA cut with *Nco*I (EMBL accession no. AJ618933, see Additional file 2).

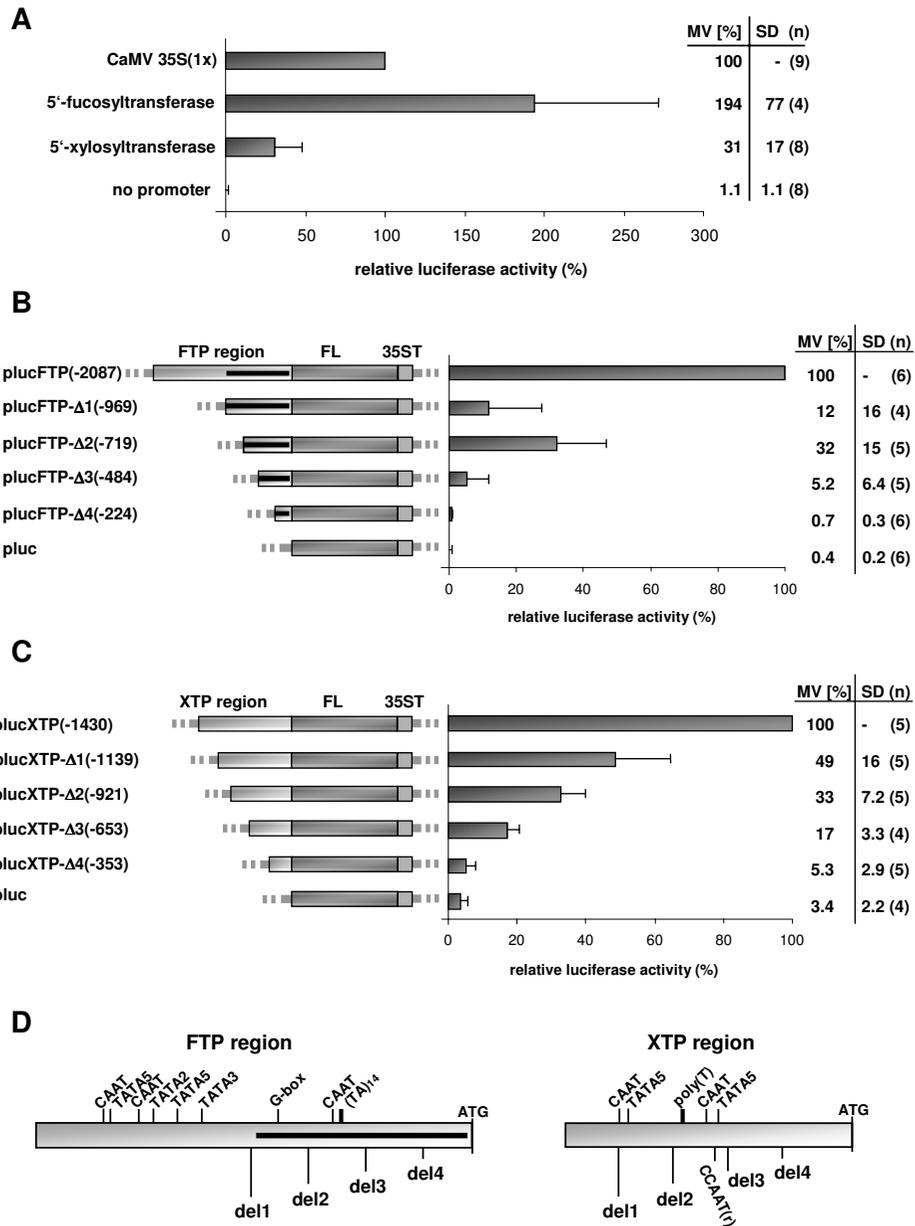


**Figure 4**

**A modified sequence directly upstream of the start codon slightly increases transcription efficiency.** Transcriptional activation by the promoters CaMV 35S(long) and CaMV 35S(1x) is shown with and without modified ATG context sequence (GCCATGG to a ACCATGG). The expression rates driven without this sequence were defined as 100%. Mean values (MV), standard deviations (SD) and numbers of replication (n) are indicated.

To ensure valid comparison with the heterologous promoter activity studies (Fig. 2), the CaMV 35S activity was again used as the standard. The two promoters identified from *Physcomitrella patens*, the 5'-sequences of the *fuc-t* and *xyl-t*, exhibited expression levels of 194% and 31%, respectively (Fig. 5A).

Four deletion constructs were examined for the functional analysis of the *fuc-t* promoter region (Fig. 5B). The first deletion construct plucFTP-Δ1 included 969 bp of the 5'-region, of which 930 bp were identified as an intron within the 5'-untranslated region. This construct showed an activity of 12% compared with the entire 5'-sequence. This type of 5'-intron has reportedly evolved to contain



**Figure 5**

**Characterisation of endogenous promoters in *Physcomitrella patens*.** (A) Quantification of fucosyltransferase promoter (FTP) and xylosyltransferase promoter (XTP) activity. Promoters again are compared with the activity of the CaMV 35S promoter (cp. Fig. 2). Deletion constructs of FTP (B) and XTP (C) were functionally analysed. Maps of the various deletion constructs are shown next to their corresponding luciferase expression levels. The results of each deletion construct are given as percentage of the FTP and XTP activity, respectively. The begin of each deletion construct relative to the ATG start codon is given in brackets. The promoter regions of the genes include 5'-untranslated regions and in the case of the fucosyltransferase also the first intron of the gene, shown as black bar. All values are the average of four to six independent experiments, whereas all samples were measured in triplicates. The activities measured with the deletion fragments were expressed as a percentage of the activity obtained with the complete FTP and XTP, respectively. Bars represent standard deviations (SD). FL: firefly luciferase; 35ST: CaMV 35S terminator; n: number of replication. (D) Illustration of possible transcription influencing elements discussed in the text (identified by: <http://www.dna.affrc.go.jp/htdocs/PLACE/signalscan.html>. r:reverse strand).

transcriptional regulators such as enhancer elements, to further extend their involvement in the control of gene expression [e.g. [40]]. Buchman and Berg [41], for example, observed that the expression of some genes is highly dependent on the presence of a 5'-intron. Given this connection, intron-mediated enhancement is not restricted to the original promoter or coding region [42].

To further characterise the role of the *fuc-t* 5'-intron, three 5'-deletion constructs (plucFTP- $\Delta$ 2, plucFTP- $\Delta$ 3 and plucFTP- $\Delta$ 4) were created, each about 250 bp shorter than the previous construct. Compared to the first deletion construct, plucFTP- $\Delta$ 1, plucFTP- $\Delta$ 2 caused an almost threefold increase in luciferase activity (Fig. 5B; from 12 to 32%). Deletion construct plucFTP- $\Delta$ 3 showed a residual activity of around 5% compared to plucFTP, a significant decrease of luciferase activity in relation to plucFTP- $\Delta$ 2 (Fig. 5B). The shortest deletion construct, plucFTP- $\Delta$ 4, did not induce sufficient gene expression (below 1%; Fig. 5B). Previous studies analysing the rice *Act1* promoter, which is also included in our study [43], demonstrated that introns significantly contribute to gene expression levels in rice and maize [42]. However, McElroy et al. [23] did not detect promoter activity in the intron itself. The intron-mediated stimulation of gene expression was associated with an in-vivo requirement for efficient intron splicing [43]. PlucFTP- $\Delta$ 1, where the potential promoter sequence mainly consists of the 5'-intron, contained 17 bp and 22 bp, respectively, of untranslated sequence information at the intron-flanking sides. These sequences may ensure effective intron splicing, in contrast to the intron sequence analysed in the study of McElroy et al. [23]. This result corresponds to earlier observations [e.g. [44]] suggesting that the degree of enhancement depends upon the exon sequences flanking the intron; additionally that the sequences surrounding the introns do influence efficiency of intron splicing.

Nevertheless, the increase in luciferase activity induced by plucFTP- $\Delta$ 2, compared with plucFTP- $\Delta$ 1, suggests that the 250 bp missing in plucFTP- $\Delta$ 2 comprise repressor-binding elements. Therefore the intron sequence by itself must influence the level of gene expression. Accordingly, a G-box motif was identified at position -852, relative to the translation start site, which was described to act as transcriptional repressor <http://www.dna.affrc.go.jp/htdocs/PLACE/signalscan.html> (an assortment of possible transcription influencing elements is illustrated in Fig. 5D). The significant decrease in reporter expression observed with plucFTP- $\Delta$ 2 and plucFTP- $\Delta$ 3 strongly suggests that the sequence from position -719 to position -484 is essential for expression efficiency of the *fuc-t* first intron. Moreover, we identified a region of 14 TA repeats in the middle of the intron (position -583), which might serve as an alternative binding site for the basal transcription initia-

tion complex. Additionally, potential regulatory elements were recognised in the sequence between position -719 and position -484, for instance two CAAT boxes at position -1457 and -1552. Regarding the complete 5'-sequence of the fucosyltransferase, several presumptive TATA boxes were identified at position -1741, -1603, -1513 and -1409. They are all located in the 5'-region of the complete fucosyltransferase promoter sequence and missing in the deletion constructs. Outside of the intron sequence, several CAAT boxes were identified, for example at position -1762, -1673 and -631.

Since the deletion constructs plucFTP- $\Delta$ 2, plucFTP- $\Delta$ 3 and plucFTP- $\Delta$ 4 are probably not able to remove the leader intron, one could assume that the relative activity of the luciferase reporter could result from altered translational efficiency. However, we think that the differences in luciferase activity observed with the deletion constructs are due to altered transcript abundance rather than an altered 5'-UTR sequence. Several features of leader sequences located within the 5'-untranslated region (UTR), like stable hairpin structures [45,46], and the presence of AUG triplets [47], have been shown to influence mRNA translational efficiency. However, the average length of eukaryotic 5'-UTRs varies from 90 to 170 bases [47]. The 5'-UTRs of plucFTP- $\Delta$ 2, plucFTP- $\Delta$ 3 and plucFTP- $\Delta$ 4 encompass about 250 to 750 bases, whereas the last 250 bp are identical in all deletion constructs. Therefore it seems unlikely that features located further upstream strongly influence the translation efficiency.

To determine the minimal length of the putative *xyl-t* promoter sequence required for maximal luciferase activity, four 5'-deletion constructs were created (Fig. 5C). Partial deletion of the 5'-sequence resulted in a concomitant reduction of reporter gene expression. PlucXTP- $\Delta$ 1 exhibited half of the activity of plucXTP, suggesting that important regulatory elements have to be located between position -1430 and -1139. Plasmid plucXTP- $\Delta$ 1 showed 49%, plucXTP- $\Delta$ 2 33% and plucXTP- $\Delta$ 3 17% activity compared with the complete putative *xyl-t* promoter sequence. PlucXTP- $\Delta$ 4 only showed a very weak reporter gene expression equivalent to about 5% of plucXTP activity. Due to a steady decrease in expression efficiency displayed by the 5'-deletion constructs, increased activity with the introduction of additional 5'-sequence information is to be expected. In the xylosyltransferase upstream region, two possible TATA boxes were identified at position -1107 and -704. At position -873 there is a 30-base-pair poly(T) region that can serve as a functional genetic element [48]. Three putative CCAAT boxes were recognised, one of them located at position -735 in the reverse strand (r). Several CAAT boxes are present, for example at position -1141 and -816.

Both of the *Physcomitrella* 5'-upstream sequences in question are AT-rich (AT content is 61% in case of the fucosyltransferase and 57% for the xylosyltransferase 5'-untranslated region), while the AT content of translated sequences is below 50% [6].

Thus, we identified two endogenous promoter fragments that can be used to drive the expression of heterologous genes in *Physcomitrella*. The entire 5'-sequence of the *fuc-t* gene, which encompasses 2087 bp, induces reasonable expression levels of the luciferase reporter gene at 194% compared to the CaMV 35S promoter. Even the second deletion construct plucFTP- $\Delta$ 2, containing 716 bp, still showed 32% of the activity of plucFTP and therefore about three quarters of the activity of the CaMV 35S promoter. 1430 bp of the 5'-sequence of the *xyl-t* confers an expression efficiency of 31% in relation to the CaMV 35S promoter. Therefore these sequences are much more efficient to drive gene expression than the *nos* promoter (2.4%), for example.

## Conclusions

The various heterologous and endogenous promoters, whose activity in *Physcomitrella* was quantified in these studies provide valuable tools for basic plant research and biotechnology applications. Availability of a set of differentially active promoters in a particular system is important both for the construction of expression vectors used in analysis of gene function as well as for large scale production of recombinant proteins. The results presented here provide evidence that the general transcriptional and translational mechanisms of the moss *Physcomitrella* resembles those in other eukaryotic organisms, despite the evolutionary distance between them. Beyond the specific applications described in the *Physcomitrella* system, our set of expression vectors will provide a valuable resource for genetic engineering of lower and of higher plants.

## Methods

### Plasmid construction

All constructs were based on the vector pUC18 supplemented by several restriction sites (*EcoRV*, *NaeI*, *NotI*, *NheI*, *BglII*, *XhoI*, *NcoI* and *SpeI*). For this purpose, the vector pUC18 was cut inside the multiple cloning site (*Sall* and *HindIII*) and a double-stranded oligonucleotide containing the additional restriction sites was inserted.

The sequence of the firefly luciferase and the CaMV 35S terminator from plasmid pGN35S-luc<sup>+</sup> (kindly provided by Prof. Gunther Neuhaus, Freiburg University) was introduced into the modified pUC18 vector via *NcoI* (providing the ATG luciferase start codon) and *PstI*.

A synthetic polyadenylation (polyA) signal for background reduction was inserted into the *EcoRI* site upstream of the additionally introduced restriction sites. The signal was PCR-amplified from the plasmid pGL3-basic (Promega, Mannheim, Germany) (primers: polyA-a and polyA-b; all oligonucleotides used as primers are shown in Tab. 1, see Additional file 3) and subcloned into the vector pCR<sup>®</sup>4-TOPO<sup>®</sup> (Invitrogen, Karlsruhe, Germany) out of which it was cut with *EcoRI* for further cloning. This yielded the firefly luciferase construct pluc into which the transcriptional and translational enhancer elements were introduced afterwards (Fig. 1A).

The TMV (tobacco mosaic virus) omega translation enhancer [49] was cut out of pGN35S-luc<sup>+</sup> and inserted into the basic vector pluc using the restriction sites *XhoI* and *NcoI* to yield pluc- $\Omega$  (Fig. 1A). With the aim of changing the sequence precisely in front of the translation initiation codon ATG, in accordance to Kozak [33] and Luetcke et al. [34], the five nucleotides CTCAA were inserted between the restriction sites *XhoI* and *NcoI* (*NcoI* including the translation initiation codon) of pluc to give pluc-enh (Fig. 1A).

All of the promoter sequences were PCR-amplified and the amplification products subcloned into the vector pCR<sup>®</sup>4-TOPO<sup>®</sup> (Invitrogen, Karlsruhe, Germany). Fragments of the heterologous promoters, except CaMV 35S(long), were excised with *Sall* and *XhoI* and subsequently inserted between the *Sall* and *XhoI* sites of the three basic vectors pluc, pluc- $\Omega$  and pluc-enh, respectively. The CaMV 35S(long) promoter was inserted into the basic vectors using the restriction sites *EcoRI* and *XhoI*. The 5'-sequences of the *Physcomitrella* glycosyltransferases *fuc-t* and *xyl-t* were introduced into pluc using the restriction sites *Sall* and *NcoI*. To avoid background activity resulting from the vector backbone, all firefly luciferase promoter plasmids were linearized using restriction enzyme *EcoRI*, which cuts 33 bp upstream the start of the corresponding promoter sequence, at the 5'-end of the multiple cloning site.

The sequence of the CaMV 35S(1x) promoter was amplified by PCR from the vector pRT101 [50] using the primers 35S(1x)-a and 35S(1x)-b. For the amplification of CaMV 35S(long) promoter sequence, the vector mAV4 [51] was used as a PCR-template with the primers 35S(long)-a and 35S(long)-b. PGN35S-luc<sup>+</sup> employed for amplification of the CaMV 35S(2x) promoter sequence using the primers 35S(2x)-a and 35S(2x)-b. For the construction of plasmids carrying the rice *Act1* promoter [43], the *Act1* 5'-region was PCR-amplified from the vector pDM302 [52], kindly provided by Prof. Ray Wu, Cornell University Ithaca, New York], using the primers Act1-a and Act1-b. PEGFP-N1 (BD Biosciences Clontech, Heidel-

berg, Germany) served as template for the amplification of the human cytomegalo virus (CMV) promoter with the primers CMV-a and CMV-b. The *nos* promoter sequence was PCR-amplified with the primers nos-a and nos-b using plasmid pBSN3 [8] as a template. Amplification of the simian virus (SV) 40 promoter sequence from the vector pSG5 (Stratagene, Amsterdam, The Netherlands) was carried out with the primers SV40-a and SV40-b.

The 5'-sequences of the *Physcomitrella* glycosyltransferases identified by inverse PCR were amplified using the primers FT-P-a and FT-P-b in the case of the fucosyltransferase promoter and XT-P-a and XT-P-b for the xylosyltransferase promoter, respectively. The promoter region of deletion construct plucFTP-Δ1 was amplified with the primers FTP-Δ1 and FT-P-b. PlucFTP-Δ1 served as a template for the amplification of the promoter sequences of plucFTP-Δ2, plucFTP-Δ3 and plucFTP-Δ4 with the forward primers FTP-Δ2, FTP-Δ3 or FTP-Δ4, respectively, and the reverse primer FT-P-b. The promoter regions of the xylosyltransferase deletion constructs were amplified with the forward primers XTP-Δ1, XTP-Δ2, XTP-Δ3 or XTP-Δ4, respectively, and XT-P-b as reverse primer, using plucXTP as template. The resulting amplification products were inserted into pluc using restriction sites *Sall* and *XhoI*.

Taken together, the XT fusion contains the sequence upstream of position -1 relative to the translation start codon. For cloning reasons the G at position -1 was replaced by the nucleotides CC followed by the luciferase coding sequence. The FT fusion contains the 5'-untranslated sequence upstream of position -2. The nucleotides A and T at positions -2 and -1 were replaced by CC followed by the luciferase coding sequence.

For the creation of the *Renilla* luciferase control plasmid (pRLuc) the firefly luciferase sequence was exchanged with the *Renilla* luciferase sequence in the vector pluc-35S(long), using *XhoI* and *XbaI* (Fig. 1B). The sequence of the luciferase control reporter gene was amplified from the plasmid pRL-CMV (Promega), using the primers RLuc-a and RLuc-b.

#### Identification of endogenous promoter sequences via inverse PCR

The 5'-regions of the *Physcomitrella patens* fucosyl- and xylosyltransferase genes were identified by inverse PCR (I-PCR). Genomic DNA (3–5 μg) was digested with 30 units of various restriction endonucleases in a total volume of 30 μl at 37°C for two hours. One endonuclease (e.g. *Bam*HI, *Bsp*HI, *Nco*I, *Nde*I, *Sph*I) was used per approach. The digested DNA was purified and eluted in a volume of 50 μl of elution buffer (EB; Qiagen, Hilden, Germany). Prior to any further treatment, 10 μl of the eluate were analysed on an agarose gel (0.5%).

Purified DNA was re-ligated in a total volume of 300 μl for two hours at room temperature and for an additional two days at 4°C. Before the addition of the enzyme ligation mixture, the DNA was incubated at 50°C for five minutes and then quenched on ice, in order to melt sticky end base pairing. After ethanol precipitation with 0.3 M Na-acetate (pH 4.8) and two washes with 70% ethanol, the re-ligated DNA was resuspended in 200 μl elution buffer (theoretically: ~10 ng/μl).

I-PCR was done with 0.25 μl Advantage™ cDNA Polymerase Mix (implying proofreading activity) and buffer (including 3.5 mM Mg (OAc)<sub>2</sub>, both BD Biosciences Clontech), 0.2 mM of each primer and 0.2 mM dNTPs and 1 to 3 μl of the re-ligated genomic DNA in a total volume of 25 μl. Cycling conditions were: an initial step of 2 minutes at 96°C, then 20 seconds at 94°C, 10 seconds initially at 61°C and 10 minutes at 68°C, with 35 to 40 repetitions, followed by a terminal step of 20 minutes at 68°C and cooling to 4°C at the end of the program. The primers were designed based on genomic fucosyl- and xylosyltransferase sequences, respectively [15]. PCR products were eluted from agarose gels (elution was done in a volume of 30 μl) and either cloned directly in pCR®4-TOPO® vector or used as templates for reconfirmation via nested PCRs. In the latter case, the gel-eluted, nested PCR products were cloned in pCR®4-TOPO®. Sequences derived from the I-PCR were confirmed by amplification of undigested genomic DNA with primers located at the 5'-end of the fucosyl- or xylosyltransferase on the one hand and at the 5'-end of the newly identified 5'-region on the other hand. The resulting products were also cloned in pCR®4-TOPO® and served as templates for the cloning of the luciferase-promoter constructs.

#### Plant material, protoplast isolation and transfection

The moss *Physcomitrella patens* (Hedw.) B.S.G. was propagated as an axenic suspension in modified Knop medium (1.84 mM KH<sub>2</sub>PO<sub>4</sub>, 3.36 mM MgSO<sub>4</sub> × 7H<sub>2</sub>O, 1.02 mM KCl, 4.24 mM Ca(NO<sub>3</sub>) × 4H<sub>2</sub>O, 0.045 mM FeSO<sub>4</sub> × 7H<sub>2</sub>O, pH 5.8). Uniform plant material grown as semi-continuous photoautotrophic bioreactor culture [53] was utilised. After isolating the protoplasts, about 3 × 10<sup>5</sup> cells were transfected with 35 μg of the various linearised firefly luciferase promoter plasmids and 0.5 μg of circular *Renilla* luciferase plasmid as a control for transfection efficiency. After polyethylene-glycol (PEG)-mediated transfection, the protoplasts were incubated in regeneration medium (Knop medium, 3% mannitol, 5% glucose, pH 5.8) for 48 h [1].

#### Luciferase assay

Cells were harvested by a 10 min-centrifugation at 160 × g in a microcentrifuge at room temperature. After discarding the supernatant, the cells were frozen in liquid nitro-

gen and stored at  $-80^{\circ}\text{C}$ . The cells were lysed by incubation in 80  $\mu\text{l}$  passive lysis buffer (PLB; Promega) and additionally homogenised in a 2 ml tube using a micropestle at maximum speed (2000 rpm) for 45 seconds on ice. Measurement of luciferase activity in the lysates (20  $\mu\text{l}$  each) was performed using the dual-luciferase reporter assay system, according to the manufacturer's recommendation (Promega). Light emission was measured immediately at room temperature using a luminometer (Berthold Lumat LB 9507, Bad Wildbach, Germany). The initial 10-second-integral of light emission was recorded. The relative luciferase activity was calculated as the ratio between the firefly luciferase and the control *Renilla* luciferase activity. This transient expression assay was repeated independently 3 to 11 times for each construct with all samples measured in triplicates.

### Authors' contributions

VH essentially performed the described experiments and prepared the manuscript. CMH participated in the analysis of the FTP deletion constructs. WJ optimised the inverse PCR assay. RR leads the chair plant biotechnology at Freiburg University. ELD was responsible for the design and development of these studies; consequently she is the corresponding author. All authors read and approved the final manuscript.

### Additional material

#### Additional File 1

**Genomic sequence of the *Physcomitrella patens*  $\alpha 1,3$ -fucosyltransferase locus** The genomic sequence of the *Physcomitrella patens*  $\alpha 1,3$ -fucosyltransferase is presented with its flanking 5'- and 3'-regions. The translation start and stop codons as well as the introns are indicated. The sequence corresponds to the sequence with the EMBL accession number AJ618932.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1472-6750-4-13-S1.PDF>]

#### Additional File 2

**Genomic sequence of the *Physcomitrella patens*  $\beta 1,2$ -xylosyltransferase locus** The genomic sequence of the *Physcomitrella patens*  $\beta 1,2$ -xylosyltransferase including the upstream region is presented. The translation start and stop codons as well as the introns are indicated. The sequence corresponds to the sequence with the EMBL accession number AJ618933.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1472-6750-4-13-S2.PDF>]

#### Additional File 3

**Table 1: Oligonucleotides used as primers** The sequences of the primers used in this study are listed with the corresponding PCR-templates.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1472-6750-4-13-S3.PDF>]

### Acknowledgements

This work was supported by a grant from German Federal Ministry of Science and Education (BMBF) No. 0312624.

### References

1. 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 U S A* 1998, **95**:4368-4373.
2. Schaefer DG: **Gene targeting in *Physcomitrella patens*.** *Curr Opin Plant Biol* 2001, **4**:143-150.
3. Koprivova A, Altmann F, Luethje R, Gorr G, Kopriva S, Reski R, Decker E: **Targeted knockout plants lacking the "immunogenic residues" of N-glycans: opening of the perspectives for molecular farming.** 2003.
4. Reski R: **Molecular genetics of *Physcomitrella*.** *Planta* 1999, **208**:301-309.
5. Theissen Gunter, Munster Thomas, Henschel Katrin: **Why don't mosses flower?** *New Phytol* 2001, **150**:1-5.
6. Rensing SA, Rombauts S, Van de Peer Y, Reski R: **Moss transcriptome and beyond.** *Trends Plant Sci* 2002, **7**:535-538.
7. Nishiyama T, Fujita T, Shin IT, Seki M, Nishide H, Uchiyama I, Kamiya A, Carninci P, Hayashizaki Y, Shinozaki K, Kohara Y, Hasebe M: **Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: implication for land plant evolution.** *Proc Natl Acad Sci U S A* 2003, **100**:8007-8012.
8. Egener T, Granada J, Guitton MC, Hohe A, Holtorf H, Lucht JM, Rensing SA, Schlink K, Schulte J, Schween G, Zimmermann S, Duwenig E, Rak B, Reski R: **High frequency of phenotypic deviations in *Physcomitrella patens* plants transformed with a gene-disruption library.** *BMC Plant Biol* 2002, **2**:6.
9. Nishiyama T, Hiwatashi Y, Sakakibara I, Kato M, Hasebe M: **Tagged mutagenesis and gene-trap in the moss, *Physcomitrella patens* by shuttle mutagenesis.** *DNA Res* 2000, **7**:9-17.
10. Cove DJ: **Regulation of development in the moss *Physcomitrella patens*.** In *Development: The molecular genetic approach* Edited by: Russo V E A, Brody S, Cove D and Ottolenghi S. Berlin, Springer-Verlag; 1992:179-193.
11. Schulz P, Reski R, Maldiney R, Laloue M, v. Schwartzberg K: **Kinetics of cytokinin production and bud formation in *Physcomitrella*: Analysis of wild type, a developmental mutant and two its ipt transgenics.** *Journal of Plant Physiology* 2000, **156**:768-774.
12. Bierfreund NM, Reski R, Decker EL: **Use of an inducible reporter gene system for the analysis of auxin distribution in the moss *Physcomitrella patens*.** *Plant Cell Rep* 2003, **21**:1143-1152.
13. Holtorf H, Guitton MC, Reski R: **Plant functional genomics.** *Naturwissenschaften* 2002, **89**:235-249.
14. Decker EL, Reski R: **The moss bioreactor.** *Curr Opin Plant Biol* 2004, **7**:(online available).
15. Koprivova A, Altmann F, Gorr G, Kopriva S, Reski R, Decker EL: **N-glycosylation in the moss *Physcomitrella patens* is organized similarly to higher plants.** *Plant Biol* 2003, **5**:582-591.
16. Holtorf H, Hohe A, Wang HL, Jugold M, Rausch T, Duwenig E, Reski R: **Promoter subfragments of the sugar beet V-type H<sup>+</sup>-ATPase subunit c isoform drive the expression of transgenes in the moss *Physcomitrella patens*.** *Plant Cell Reports* 2002, **21**:341.
17. Schaefer DG: **A new moss genetics: targeted mutagenesis in *Physcomitrella patens*.** *Annu Rev Plant Biol* 2002, **53**:477-501.
18. Imaizumi T, Kadota A, Hasebe M, Wada M: **Cryptochrome light signals control development to suppress auxin sensitivity in the moss *Physcomitrella patens*.** *Plant Cell* 2002, **14**:373-386.
19. Knight CD, Sehgal A, Atwal K, Wallace JC, Cove DJ, Coates D, Quatrano RS, Bahadur S, Stockley PG, Cuming AC: **Molecular responses to abscisic acid and stress are conserved between moss and cereals.** *Plant Cell* 1995, **7**:499-506.
20. Bezanilla M, Pan A, Quatrano RS: **RNA interference in the moss *Physcomitrella patens*.** *Plant Physiol* 2003, **133**:470-474.
21. Zeidler M, Hartmann E, Hughes J: **Transgene expression in the moss *Ceratodon purpureus*.** *J Plant Physiol* 1999, **154**:641-650.
22. Decker EL, Skerka C, Zipfel PF: **The early growth response protein (EGR-1) regulates interleukin-2 transcription by syner-**

- gistic interaction with the nuclear factor of activated T cells. *J Biol Chem* 1998, **273**:26923-26930.
23. McElroy D, Blowers AD, Jeness B, Wu R: **Construction of expression vectors based on the rice actin I (Act1) 5' region for use in monocot transformation.** *Mol Gen Genet* 1991, **231**:150-160.
  24. Zeidler M, Gatz C, Hartmann E, Hughes J: **Tetracycline-regulated reporter gene expression in the moss *Physcomitrella patens*.** *Plant Mol Biol* 1996, **30**:199-205.
  25. Kay R, Chan A, Daly M, McPherson J: **Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes.** *Science* 1987, **236**:1299-1302.
  26. Sanders PR, Winter JA, Barnason AR, Rogers SG, Fraley RT: **Comparison of cauliflower mosaic virus 35S and nopaline synthase promoters in transgenic plants.** *Nucleic Acids Res* 1987, **15**:1543-1558.
  27. Gallie DR, Walbot V: **Identification of the motifs within the tobacco mosaic virus 5'-leader responsible for enhancing translation.** *Nucleic Acids Res* 1992, **20**:4631-4638.
  28. Wells DR, Tanguay RL, Le H, Gallie DR: **HSP101 functions as a specific translational regulatory protein whose activity is regulated by nutrient status.** *Genes Dev* 1998, **12**:3236-3251.
  29. Holtorf S, Apel K, Bohlmann H: **Comparison of different constitutive and inducible promoters for the overexpression of transgenes in *Arabidopsis thaliana*.** *Plant Mol Biol* 1995, **29**:637-646.
  30. Dowson Day MJ, Ashurst JL, Mathias SF, Watts JW, Wilson TM, Dixon RA: **Plant viral leaders influence expression of a reporter gene in tobacco.** *Plant Mol Biol* 1993, **23**:97-109.
  31. Gallie DR: **The 5'-leader of tobacco mosaic virus promotes translation through enhanced recruitment of eIF4F.** *Nucleic Acids Res* 2002, **30**:3401-3411.
  32. Gallie DR, Walbot V, Hershey JW: **The ribosomal fraction mediates the translational enhancement associated with the 5'-leader of tobacco mosaic virus.** *Nucleic Acids Res* 1988, **16**:8675-8694.
  33. Kozak M: **An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs.** *Nucleic Acids Res* 1987, **15**:8125-8148.
  34. Luetcke HA, Chow KC, Mickel FS, Moss KA, Kern HF, Scheele GA: **Selection of AUG initiation codons differs in plants and animals.** *Embo J* 1987, **6**:43-48.
  35. Joshi CP, Zhou H, Huang X, Chiang VL: **Context sequences of translation initiation codon in plants.** *Plant Mol Biol* 1997, **35**:993-1001.
  36. Kozak M: **Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes.** *Cell* 1986, **44**:283-292.
  37. Taylor JL, Jones JDG, Sandler S, Mueller GM, Bedbrook JR, Dunsmuir P: **Optimizing the expression of chimeric genes in plant cells.** *Mol Gen Genet* 1987, **210**:572-577.
  38. Guerineau F, Lucy A, Mullineaux P: **Effect of two consensus sequences preceding the translation initiator codon on gene expression in plant protoplasts.** *Plant Mol Biol* 1992, **18**:815-818.
  39. Gallie DR: **Translational control of cellular and viral mRNAs.** *Plant Mol Biol* 1996, **32**:145-158.
  40. Bornstein P, McKay J, Morishima JK, Devarayalu S, Gelinis RE: **Regulatory elements in the first intron contribute to transcriptional control of the human alpha 1(I) collagen gene.** *Proc Natl Acad Sci U S A* 1987, **84**:8869-8873.
  41. Buchman AR, Berg P: **Comparison of intron-dependent and intron-independent gene expression.** *Mol Cell Biol* 1988, **8**:4395-4405.
  42. Callis J, Fromm M, Walbot V: **Introns increase gene expression in cultured maize cells.** *Genes Dev* 1987, **1**:1183-1200.
  43. McElroy D, Zhang W, Cao J, Wu R: **Isolation of an efficient actin promoter for use in rice transformation.** *Plant Cell* 1990, **2**:163-171.
  44. Matsumoto K, Wassarman KM, Wolffe AP: **Nuclear history of a pre-mRNA determines the translational activity of cytoplasmic mRNA.** *Embo J* 1998, **17**:2107-2121.
  45. Niepel M, Ling J, Gallie DR: **Secondary structure in the 5'-leader or 3'-untranslated region reduces protein yield but does not affect the functional interaction between the 5'-cap and the poly(A) tail.** *FEBS Letters* 1999, **462**:79-84.
  46. Kozak M: **Pushing the limits of the scanning mechanism for initiation of translation.** *Gene* 2002, **299**:1-34.
  47. Rogozin IB, Kochtov AV, Kondrashov FA, V. Koonin E., Milanezi L: **Presence of ATG triplets in 5' untranslated regions of eukaryotic cDNAs correlates with a 'weak' context of the start codon.** *Bioinformatics* 2001, **17**:890-900.
  48. Struhl K: **Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast.** *Proc Natl Acad Sci USA* 1985, **82**:8419-8423.
  49. Gallie DR, Sleat DE, Watts JW, Turner PC, Wilson TM: **A comparison of eukaryotic viral 5'-leader sequences as enhancers of mRNA expression in vivo.** *Nucleic Acids Res* 1987, **15**:8693-8711.
  50. Toepfer R, Schell J, Steinbiss HH: **Versatile cloning vectors for transient gene expression and direct gene transfer in plant cells.** *Nucleic Acids Res* 1988, **16**:8725.
  51. Kircher S, Wellmer F, Nick P, Rugner A, Schafer E, Harter K: **Nuclear import of the parsley bZIP transcription factor CPRF2 is regulated by phytochrome photoreceptors.** *J Cell Biol* 1999, **144**:201-211.
  52. Cao J, Duan X, McElroy D, Wu R: **Regeneration of herbicide resistant transgenic rice plants following microprojectile-mediated transformation of suspension culture cells.** *Plant Cell Reports* 1992, **11**:586-591.
  53. Hohe A, Decker EL, Gorr G, Schween G, Reski R: **Tight control of growth and cell differentiation in photoautotrophically growing moss (*Physcomitrella patens*) bioreactor cultures.** *Plant Cell Reports* 2002, **20**:1135-1140.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

