Curr Genet (2005) 47: 111–120
DOI 10.1007/s00294-004-0555-z

RESEARCH ARTICLE

Wolfgang Jost · Sandra Link · Verena Horstmann
Eva L. Decker · Ralf Reski · Gilbert Gorr

Isolation and characterisation of three moss-derived beta-tubulin promoters suitable for recombinant expression

Received: 27 August 2004 / Revised: 12 November 2004 / Accepted: 16 November 2004 / Published online: 17 December 2004

Abstract The moss *Physcomitrella patens* is an excellent tool to study plant gene–function relationships due to its high rate of homologous recombination (HR). It has also been shown to be very useful in the production of recombinant proteins which are secreted into a simple medium. Thus, there is a need for suitable promoters functional in this well established model organism. We isolated genomic flanking regions of the beta-tubulin gene family from *Physcomitrella*, concentrating on those family members showing high transcript abundance integrated over gametophytic tissues. Using a novel, fast and reliable quantification assay based on the transient expression and secretion of a recombinant human protein, three genomic upstream regions were characterised in serial deletion constructs. Expression rates were up to three times higher than those obtained with the 35S cauliflower mosaic virus (35S) promoter, which served as a reference.

Keywords Recombinant expression · Tubulin promoter · Bryophyte · *Physcomitrella* · Gametophyte

Introduction

Because of their simple morphology with only a few cell types, the direct access to their cell lineage and their clear differentiation patterns, bryophytes like *Physcomitrella patens* and *Ceratodon purpureus* have been used as plant model organisms for physiological, developmental and mutational studies from as early as the beginning of past century (for a general review on mosses, see Cove et al. 1997; Reski 1998). More recently, moss was discovered as a tool for reverse genetics due to the highly efficient homologous recombination (HR) possible in its nucleus (Schaefer and Zryd 1997; Strepp et al. 1998). This opened so far unequalled potential for studying plant gene function (for a review, see Schaefer 2002). Furthermore, in the case of *P. patens*, expressed sequence tags (ESTs) covering more than 95% of the transcriptome have been analysed in detail (Rensing et al. 2002; Nishiyama et al. 2003), with over 100,000 ESTs available in public databases.

Nevertheless, there remains a lack of regulatory sequences capable of driving recombinant expression in lower plants and especially in mosses. Appropriate regulatory sequences would be useful not only for ectopic expression of proteins for physiological, developmental and/or genetic studies, but also for the production of complex biopharmaceuticals (Decker and Reski 2004) - including biopharmaceuticals with a humanised N-glycosylation (Koprivova et al. 2004). Promoters per se and notably promoters of variable strength are very useful in designing optimised production lines expressing one or even several recombinant genes in a tuned manner.

The 35S cauliflower mosaic virus (35S) promoter and derivatives thereof, the nopalin synthase promoter and the rice actin 1 5′ region are most commonly used for recombinant expression in higher plants. With the exception of the latter, these heterologous promoters have shown to be rather weak in mosses (Zeidler et al. 1999; Horstmann et al. 2004). As opposed to the diploid sporophyte of higher plants, the dominant life cycle phase of mosses is the haploid, photoautotrophic gametophyte. Moss endogenous expression promoting regions (EPRs) can be expected to result in better expression rates than the above-mentioned heterologous promoters. However, reintroducing endogenous sequences, e.g. genomic 5′ flanking regions of essential single-copy genes, into a haploid recipient showing outstanding rates of HR holds a high risk to set lethal, apparent or even minor/subliminal phenotypes. Even

Communicated by S. Hohmann

W. Jost · S. Link · G. Gorr (✉)
greenovation Biotech GmbH, Bötzinger Strasse 29b,
79111 Freiburg, Germany
E-mail: ggorr@greenovation.com
Tel.: +49-761-47099130
Fax: +49-761-47099190

W. Jost · V. Horstmann · E. L. Decker · R. Reski
Plant Biotechnology, Faculty of Biology, University of Freiburg,
Schänzlestrasse 1, 79104 Freiburg, Germany

though HR is not quantitative, this risk could be avoided or at least reduced to a minimum by using EPRs of a functionally redundant gene family.

Here, we describe the isolation and characterisation of moss endogenous EPRs derived from the beta-tubulin gene family of *P. patens*. It consists of six highly conserved members (*PpTub1–PpTub6*; Jost et al. 2004). Their deduced amino acid sequences with identities of 96.4–99.5% strongly suggest a functional redundancy within the family. In a first step, we performed a detailed analysis of public EST data and found that *PpTub1–PpTub4*, in contrast to *PpTub5* and *PpTub6*, are most frequently presented in public databases, implying a proper transcriptional activity of these four genes in the tissues under investigation. Since *PpTub4* has already been shown to be down-regulated in adult gametophytes (Jost et al. 2004), we focused on the endogenous transcript levels of *PpTub1–PpTub3* under the relevant culture conditions important for the production of recombinant proteins. Besides the determination of transcriptional initiation (transcription start points; TSPs) and a search for common upstream motifs in silico, we established a fast, reliable and non-invasive transient expression assay based on the secretion and direct quantification of recombinant human vascular endothelial growth factor ($rhVEGF_{121}$) in the culture supernatant of non-regenerating *Physcomitrella* protoplasts.

We directly compared different deletion constructs, including promoter and corresponding 5′ untranslated region (5′UTR) sequences by quantifying the levels of secreted product with a commercially available ELISA. Obtained values were evaluated relative to values generated by the 35S promoter in each transformation and reached 100% of the 35S promoter for short- and up to over 300% for full-length constructs. This resulted in several functional moss EPRs variable in strength but with a reduced risk of introducing unintended phenotypes upon HR in transgenic/reverse genetic approaches.

Materials and methods

Standard plant culture conditions

P. patens (Hedw.) B.S.G. cultures were grown axenically under sterile conditions in 500 ml Erlenmeyer flasks containing 200 ml of modified inorganic liquid Knop medium, as described by Reski and Abel (1985).

Bioreactor cultures were grown in stirred tank glass bioreactors (Applikon) with a working volume of 5 l (Hohe et al. 2002) in 10% Knop medium.

Preparation of genomic DNA, total RNA and first-strand cDNA

P. patens genomic DNA was isolated from 13-day-old protonema following the CTAB protocol, as described by Richter et al. (2002). *P. patens* total RNA was prepared by disintegration of either 100 mg tissue (fresh weight) or protoplasts pooled from five transformations transiently transformed with water instead of DNA, under liquid nitrogen in a mixer mill (MM 300, Retsch) and use of RNeasy plant mini kit (Qiagen) following the manual of the manufacturer. DNase-treated (GIBCO BRL) total RNA was used for reverse transcription with oligo-d(T) primers and Superscript II RNase H reverse transcriptase (GIBCO BRL), following the manual of the manufacturer.

Molecular cloning of *PpTub1–PpTub3* genomic regions via inverse PCR

In general, conventional molecular biology protocols were applied as described by Sambrook et al. (2000). All primers mentioned in this study were synthesised by MWG Biotech AG. The corresponding sequences can be given upon request.

Genomic DNA was digested with various restriction endonucleases (all MBI Fermentas), using one endonuclease per digest. Purified DNA (PCR purification columns; Qiagen) was religated (T4 ligase; MBI Fermentas). After ethanol precipitation and two washes, the religated DNA was resuspended in water and 0.01 vol. of this religated genomic DNA together with gene-specific and/or degenerate primers was used per inverse PCR (Advantage cDNA polymerase mix; BD Biosciences Clontech). For a detailed protocol, see Horstmann et al. (2004). PCR products were eluted from agarose gels and either cloned directly in TOPO TA vectors (pCR4-TOPO; Invitrogen) or used as templates for reconfirmation in nested PCRs. In total, for each *PpTub* gene, at least two independent inverse PCR clones were generated. Based on the obtained sequence data, several conventional proofreading PCR-clones for each genomic region, raised on undigested genomic DNA as a template and covering the corresponding genomic regions completely, were generated in order to confirm the inverse PCR sequence data.

Transcript detection

For PCRs on bioreactor and protoplast cDNA, the identical upstream/forward primer, located within exon 3, was used for amplification of all three *PpTub* genes. Downstream/reverse primers were gene-specific and located within the corresponding 3′UTRs. Length, G/C content and calculated annealing temperature were in the same range for all primer pairs, as were the expected lengths of the PCR products. Gene-specificity for each primer pair was shown in PCRs using cloned DNA of *PpTub1–PpTub6* as template (data not shown). The absence of contaminating genomic DNA was verified by control PCRs for each cDNA with primers located on either side of an intron (data not shown). Amplificates of

two to three different cycle-numbers for each PCR were gel-analysed to identify those cycle-numbers for quantification in which amplification did not quite reach the stationary phase. Two independent reverse transcriptions for each RNA preparation were analysed. All PCRs were repeated at least once on each independent cDNA, always starting from mastermixes and using Taq recombinant polymerase (MBI Fermentas). PCR samples were loaded quantitatively on ethidium bromide-stained agarose gels. After electrophoresis, pictures were taken under ultraviolet light (BioDocAnalyze; Biometra).

For real-time PCRs (Lightcycler; Roche), upstream/forward gene-specific primers located within the corresponding 5′UTRs and an identical downstream/reverse primer located within exon 1, together with the LightCycler-FastStart DNA master SYBR green I kit (Roche) were used. After 10 min of denaturation and 35 cycles of amplification (15 s at 95°C, 5 s at 60°C, 20 s at 70°C), a melting curve analysis showed only one product per positive PCR. Furthermore, gel-analysis of all PCRs after their run revealed a correct length of all amplificates (data not shown).

Cloning of reporter gene constructs

The vector pRT101VEGFP21 (Gorr 1999) was used for reporter gene expression under the control of the 35S promoter (the cDNA coding for $rhVEGF_{121}$ was kindly provided by Dr. Weich, GBF, Germany). A modified version of pRT101VEGFP21, with removed 35S promoter, was kindly provided by Dr. A. Weise and Dr. M. Rodriguez-Franco (greenovation, Germany). This vector, called pRT101new, still harboured the coding region for $rhVEGF_{121}$ and 35S terminator sequences and was used for the cloning of putative EPRs of *PpTub1–PpTub3* via *Xho*I or *Hin*cII, precisely in front of the start codon of the reporter gene (for a more detailed description, see Weise et al. 2004).

The genomic fragments of constructs 1-0, 1-1, 1-3, 1-4, 2-0 and 3-0 were amplified with Pfu native polymerase (MBI Fermentas). In the cases introduced by the primers, the PCR products were either cut with *Xho*I and ligated into pRT101new via *Xho*I/*Hin*cII or not cut at all and ligated into *Hin*cII-opened pRT101new. All generated clones were verified by sequencing. Reporter gene constructs 1-2 (*Xho*I/*Eco*RI), 2-1 (*Bgl*II), 2-2 (*Sal*I), 2-3 (*Eco*RI/*Sal*I), 2-4 (*Xho*I/*Sal*I), 3-1 (*Sal*I), 3-2 (*Eco*147I/*Hin*cII) and 3-3 (*Xho*I/*Sal*I) were generated by internal deletions of the corresponding longer clones. The remaining vectors were gel-eluted and religated. In the cases when single-strand overhangs did not fit, ligation was performed after filling-in the recessed 3′ termini with a Klenow fragment (MBI Fermentas).

PpTub1 luciferase constructs were cloned by overlapping proofreading PCRs (Pfu native polymerase; MBI Fermentas). In a first set of PCRs, the *PpTub1* genomic upstream regions, the *PpTub1* genomic downstream regions and the luciferase gene were amplified on genomic DNA or cloned luciferase DNA (pSP-luc+; Promega) with primers for genomic upstream and downstream amplifications that overlapped with the luciferase gene. Three corresponding gel-eluted PCR products together (upstream region, luciferase gene, downstream region), were used as a template in a second, overlapping PCR. Three such amplificates varying in upstream (promoter, 5′ UTR) and downstream (3′ UTR, terminator) lengths were cloned in *Eco*RV linearised pZero-2 (Invitrogen), resulting in three different constructs (luc1, luc2, luc3). All generated clones were verified by sequencing. For the creation of the *Renilla* luciferase control plasmid (pRluc) and the firefly luciferase plasmid (p35S/luc), see Horstmann et al. (2004).

Transient transformations and product quantification

Moss tissue for protoplast isolation was cultivated for at least 7 days in Knop medium with reduced (10%) $Ca(NO_3)_2$ content. Protoplasts were generated essentially as described by Rother et al. (1994). Subsequently, protoplasts were sedimented by centrifugation at room temperature for 10 min at 55 *g* and resuspended in 3 M medium (15 mM $CaCl_2{\cdot}2H_2O$, 0.48 M mannitol, 0.1% MES, pH 5.6, 540 mOsm, sterile-filtered; Schaefer et al. 1991) to a concentration of $1.2{\times}10^6$ protoplasts ml^{-1}. Then, 250 µl of this protoplast suspension were mixed with 50 µl of DNA solution (equimolar amounts for the different constructs or water only were used with a minimum of 20 µg for the smallest construct) and gently mixed with 250 µl of polyethylene glycol (PEG) solution (40% PEG 4000, 0.4 M mannitol, 0.1 M $Ca(NO_3)_2$, pH 6.0). After incubation for 6 min at room temperature, the suspension was diluted progressively by adding 1–4 ml of 3 M medium. After centrifugation for 10 min at 55 *g* at 20°C and resuspending in 400 µl of 3 M medium, the transformed protoplasts were incubated in 48-well plates (Cellstar; Greiner Bio-One) in dim light (4.6 $\mu mols^{-1}\ m^{-2}$) at 25°C. Under these conditions, the protoplasts are viable for several weeks (Armin Baur, personal communication) but do not regenerate. Samples were taken after 24 h and 48 h, by carefully replacing half of the medium (200 µl) by fresh medium. Samples were stored at −20°C. Transiently expressed $rhVEGF_{121}$ from the 48-h samples was quantified by ELISA (R&D Systems), according to the instructions of the manufacturer. The ELISA reader was a Multiskan RC (Labsystems).

For the luciferase assay, protoplasts after transformation were incubated in regeneration medium (Knop medium, 3% mannitol, 5% glucose, pH 5.8) to achieve standard conditions for the luciferase assay (Horstmann et al. 2004). Under these conditions, the protoplasts start to build up cell walls and subsequently undergo cell division. Cells were harvested 48 h after transformation and were always measured in triplicates. The relative

luciferase activity was normalized between the firefly luciferase and the control *Renilla* luciferase activity, both driven by the 35S promoter. For a detailed protocol, refer to the dual-luciferase reporter assay system described by Promega and Horstmann et al. (2004).

EST search

In order to identify all *Physcomitrella* ESTs with significant homology to beta-tubulins (GenBank accession numbers can be given upon request), the GenBank EST database (http://www.ncbi.nlm.nih.gov) was searched by NCBI/blastn with the *AtTub1* gene as query. Settings were as given by the browser with database "EST others", limit by entrez query "Physcomitrella" and an expected E-value range of 0–0.01 (for a public BLAST service against clustered EST data from the moss *P. patens*, see http://www.cosmoss.org). The resulting ESTs were grouped for each particular gene by multiple sequence alignments, resulting in 5′ and 3′ contigs for each locus. From each contig, 5′UTR and 3′UTR sequences were used for additional analogous BLAST searches in order to identify ESTs outside the coding regions. The total numbers of ESTs for each library were determined by searching the NCBI nucleotide browser. Search terms were "taxon identifier" (txid3217[Organism:exp]) and "EST" for the total amount of ESTs. Additional search terms were used for the different libraries, e.g. "AND normalized" or "AND full length NOT normalized" or "NOT full length".

The nucleotide sequence data reported here (*PpTub1*, *PpTub2*, *PpTub3* genomic upstream/downstream regions) are available in the DDBJ/EMBL/GenBank databases under accession numbers AY724467, AY724468, AY724469, AY724470, AY724471 and AY724472. For *PpTub1*, *PpTub2* and *PpTub3* genomic coding regions, see also AY382287, AY382288 and AY382289.

Software

Sci Ed Central, Clone Manager Suite was used for primer design, general sequence analysis and pairwise sequence alignments. Lasergene, DNASTAR SeqMan Ver. 5 was used for analysing and editing the sequence raw data.

Results

Analysis of public EST data

The beta-tubulin gene family of *P. patens* consists of six highly conserved members (*PpTub1–PpTub6*; Jost et al. 2004). Their deduced amino acid sequences are nearly identical (96.4–99.5%). In contrast to beta-tubulin families of other species, this conservation includes the very C-termini, which are the major sites of modifications (MacRae 1997) and interaction (Littauer et al. 1986). It is thought that the C-termini may not be required for microtubule assembly per se, but may be necessary for conferring properties required for isotype-specific functions (for a review on plant tubulins, see Breviaro and Nick 2000). Thus, the overall sequence conservation and especially the C-terminal conservation strongly suggest a functional redundancy within the entire family (Jost et al. 2004).

A detailed analysis of all public available EST data revealed that *PpTub1–PpTub4* are roughly five to ten times more frequently represented in public EST databases than *PpTub5* or *PpTub6* (see Table 1), indicating a higher transcriptional activity of these four genes in the tissues under investigation (gametophyte, mainly protonema).

Among these four genes, *PpTub4* is known to be down-regulated in adult gametophytes (Jost et al. 2004). We focused our interest on *PpTub1–PpTub3*, which so far have not exhibited differential regulation in gametophytic tissues (Jost et al. 2004). However, the possibility of a still uncovered down-regulation of *PpTub1–PpTub3* - in particular under conditions of interest - had to be taken into account since they might affect the efficiency of corresponding EPRs in ectopic recombinant expression.

Transcriptional activity of endogenous *PpTub* promoters

We performed a two-step reverse transcriptase PCR to analyse endogenous transcript levels of *PpTub1–PpTub3* in tissues generated from a *Physcomitrella* bioreactor culture. In comparison with their corresponding primer fidelities tested on genomic DNA, *PpTub1–PpTub3* reached comparable levels of expression in all repetitions performed (Fig. 1, top).

To exemplarily verify the two-step reverse transcriptase PCR expression studies more precisely, we followed the transcript abundance of *PpTub2* and *PpTub3* in total RNA isolated from a flask-grown culture (exhibiting identical tissue compositions, as compared with bioreactors) in real-time PCR. As shown in Fig. 1, bottom, in these experiments both genes exhibit identical transcript quantities. Taken together, no major decrease in transcript abundance of any of the three genes could be detected regardless of different culture conditions; and, thus, all three genes were confirmed as candidates for further analysis of their corresponding EPRs.

Besides the expression of heterologous proteins in bioreactor or flask culture, endogenous moss EPRs could be used in a transient production system using non-regenerating *Physcomitrella* protoplasts. Furthermore, in such a transient system, putative EPRs could easily be quantified. Plant beta-tubulin genes encode for a major microtubule protein involved in a variety of intracellular functions, e.g. morphogenesis, cell growth and division. Prior to the cell wall formation, the cortical

Table 1 Number of public *Physcomitrella* ESTs[a] with homology to beta-tubulins

library	NFL[b]	FL[c]	PP[d]	total(5')
No. of ESTs	65263	17808	19264	102335
orientation	5' + 3'	5' + 3'	5' + 3'	
PpTub1	20 + 19	7 + 9	0 + 1	56(27)
PpTub2	12 + 7	6 + 7	2 + 2	36(20)
PpTub3	9 + 8	5 + 4	0 + 1	27(14)
PpTub4	20 + 20	3 + 5	0 + 0	48(23)
PpTub5	1 + 0	1 + 2	0 + 0	4(2)
PpTub6	2 + 2	0 + 0	0 + 2	6(2)
total	120	49	8	177

[a] As of May, 2004; NCBI/dbEST; Accession nos can be given upon request.

[b] Two normalised full length cDNA libraries (Tadasu Shin-i, Japan); mainly protonema with some rhizoid-like protonema or malformed buds; see Nishiyama et al. 2003 for details.

[c] Full length cDNA library (Tadasu Shin-i, Japan); protonema and young gametophores; see Nishiyama et al. 2003 for details.

[d] All other libraries (mainly protonema).

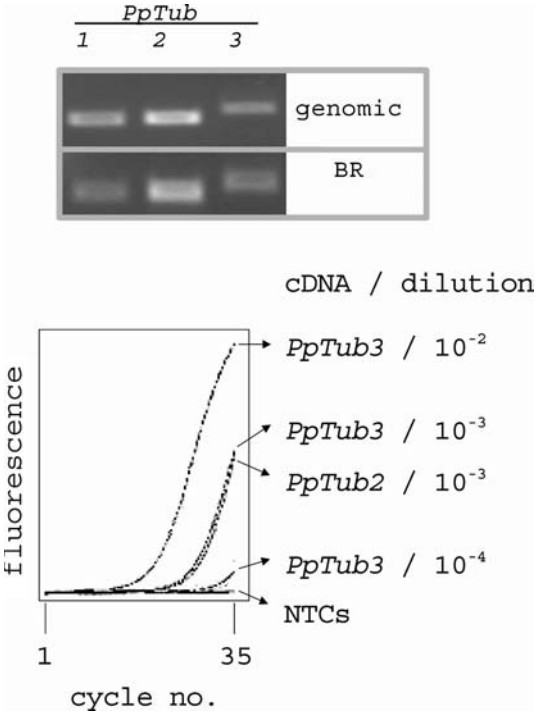


Fig. 1 *Top* Representative result of two-step reverse transcriptase PCR experiment on RNA preparations isolated from bioreactor tissue (*BR*) with reference to the corresponding primer fidelity in genomic DNA. *Bottom* Real-time PCR comparing *PpTub2* and *PpTub3* expression in flask cultures. *NTCs* No template controls

microtubule system is known to be rearranged in higher plants and is believed to control the direction in which new cellulose microfibrils are deposited. Following cell wall synthesis, cells start to divide under regenerating conditions, which again is accompanied by plant-specific microtubule rearrangements (preprophase band, phragmoplast; see Breviaro and Nick 2000). Thus, we were interested in whether the endogenous tubulin genes are transcribed at all in non-regenerating protoplasts. Total RNA was isolated from independent protoplast charges on different days, each charge pooled from five parallel mock transformations (without DNA). Both pools of protoplasts were treated alike in reporter gene expression experiments. Several PCR repetitions on either of the prepared cDNAs showed that all three *PpTub* genes analysed are expressed in non-regenerating protoplasts (data not shown).

Molecular cloning of *Physcomitrella* beta-tubulin genomic regions

A genomic walk into the flanking regions of *PpTub1–PpTub3* was used to isolate potential EPRs. The published sequence data of *PpTub1–PpTub3* (Jost et al. 2004) and additional EST sequences were used to design primers for use in inverse PCR experiments on religated genomic digests. Obtained sequence data were confirmed by independent PCR-generated clones covering the

corresponding genomic regions. In total, 1–2 kbp upstream and downstream of the coding regions of the three genes were confirmed. A schematic overview of all three genomic regions is given in Fig. 2.

In silico prediction of TSPs

The great majority of all *PpTub* ESTs belongs to full-length cDNA libraries that were prepared in a way that allows the determination of transcriptional initiation (TSPs), e.g. by using biotinylation of cap structures (Nishiyama et al. 2003; Seki et al. 1998; Sugahara et al. 2001). This helped to define the TSPs in silico. A multiple alignment of 5′ ESTs revealed that *PpTub1–PpTub3* have a precise transcriptional initiation; and 21 out of 27 5′ ESTs for *PpTub1*, 16 out of 20 5′ ESTs for *PpTub2*, 9 out of 14 5′ ESTs for *PpTub3* start at the same and most-upstream position (Fig. 2, marked by +1). Only in the case of *PpTub2* does a single EST start even 68 bp further upstream of the defined TSP. An analogous multiple alignment of all 3′ ESTs suggested more than one cleavage/poly(A) site as typical for many plant genes (for a review, see Rothnie 1996). In addition, the consensus sequences were not as clearly defined as in mammalian genes, where the sequence AAUAAA is nearly ubiquitous (for a review, see Wahle and Ruegsegger 1999).

Promoter structure and conservation

A detailed comparison among the *PpTub1–PpTub3* upstream sequences and comparison with 5′ regions of other known plant beta-tubulin genes revealed no overall conservation. However, some interesting shorter matches of conservation could be detected: (a) The TSPs of *PpTub1–PpTub3* fall within the consensus sequence T/C C A(+1) G/T T G T G C (cf. Fig. 2), (b) a weakly conserved 8 bp TATA box consensus sequence can be found in all three genes at 22–24 bp upstream of the TSPs (which is within the typical distance for plant TATA promoters) and, at a comparable position, this stretch can even be found in genomic upstream sequences of beta-tubulin 1 from *Pisum sativum* (*PsTub1*; Fig. 2), (c) *PpTub2* harbours a 40 bp stretch with 80% A content (from −328 to −367), (d) *PpTub1* harbours a nearly 900 bp very A/T-rich upstream region (80% A/T ; from position −1,306 to −417), opening the possibility for the location of scaffold/matrix attachment regions (Liebich et al. 2002) upstream of this gene.

Functional characterisation and quantification of *PpTub* regulatory sequences

Functional quantification of genomic upstream sequences of *PpTub1–PpTub3* was performed in a transient expression system using non-regenerating *Physcomitrella* protoplasts. For each upstream region, four to five constructs of decreasing length and including the 5′UTRs were brought precisely in front of the start codon of a recombinant gene coding for rhVEGF$_{121}$, which was secreted into the medium via its own signal peptide. The amount of rhVEGF$_{121}$ in the supernatant of the moss culture was quantified by ELISA. In each transformation, the mean value obtained for the 35S promoter was set as 100% and used as a reference for all other values generated within the same transformation. Each construct was transformed a minimum of six times in two to three different transformation experiments on different days and using separate protoplast charges. A schematic overview of the deletion constructs along with a summary of the results is given in Table 2. Highest expression was achieved with the longest construct of *PpTub3* (3-0) with more than three-fold expression, as compared with the 35S promoter. Relatively short fragments (250–270 bp prior to the TSPs) in all three cases (constructs 1-3, 2-3, 3-1) still gave fair expression

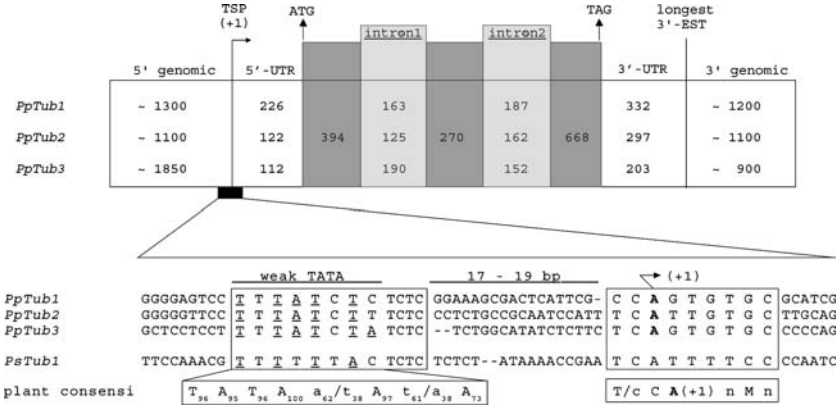
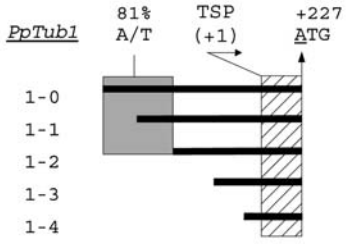
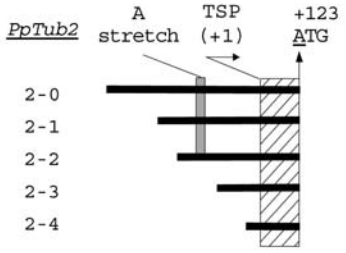
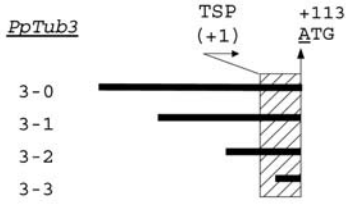


Fig. 2 *Top* Schematic overview of genomic regions of *PpTub1–PpTub3* (numbering given in basepairs). *Bottom* Alignment of regions around predicted transcriptional initiation (+1 indicated in *bold*) and putative TATA boxes. Note that, in the case of *PsTub1* (*Pisum sativum* beta-tubulin1; GenBank accession number X54844), the TSP has not been determined. *Underlined characters* within putative TATA boxes indicate identity with the consensus sequence. Consensus sequences were generated from 171 unrelated TATA plant promoters (see plantProm database:http://mendel.cs.rhul.ac.uk/mendel.php?topic=plantprom)

Table 2

VEGF assay: *PpTub* expression promoting regions in relation to 35S promoter

Name Constructs	Graphical Scheme Constructs	Promoter (5' of TSP)	n (trafos)	Relative Activity (%)	+/-
35S / VEGF		35S promoter	34 (7)	100	19
Δ35S / VEGF		no promoter	9 (3)	1	2
PpTub1					
1-0		-1307	6 (2)	212	23
1-1		- 985	6 (2)	161	37
1-2		- 416	6 (2)	165	19
1-3		- 248	6 (2)	111	41
1-4		- 83	6 (2)	39	12
PpTub2					
2-0		-1075	6 (3)	113	31
2-1		- 676	6 (3)	125	22
2-2		- 425	7 (3)	131	14
2-3		- 245	6 (2)	92	28
2-4		- 67	6 (2)	18	7
PpTub3					
3-0		-1274	6 (2)	332	52
3-1		- 767	6 (2)	287	69
3-2		- 272	6 (2)	112	49
3-3		+ 53	6 (2)	37	15

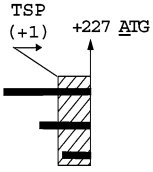
rates similar to the 35S promoter. Construct 1-2 (−417 bp), construct 2-2 (−424 bp) and construct 3-1 (−765 bp) can be defined as the shortest fragments for each EPR still resulting in high/maximum expression rates. In the case of *PpTub1*, when constructs 1-0 and 1-2 are compared, introduction of the entire A/T-rich upstream region (over 80% A/T for nearly 900 bp) results in a further enhancement of reporter gene expression. Interestingly, there is no effect on expression when only the partial A/T-rich region is inserted, as in construct 1-1.

Recently, Horstmann et al. (2004) successfully used the dual-luciferase intracellular reporter assay, which is well established in mammalian expression systems, for promoter quantification in protoplasts of *Physcomitrella* under regenerating conditions. We exemplarily tested the reliability of our novel rhVEGF$_{121}$ quantification system by measuring three *PpTub1* constructs in this assay. Values are given in Table 3 and correspond nicely with the rhVEGF$_{121}$ data, even though in these experiments regenerating protoplasts were used and not the nopalin synthase terminator but endogenous *PpTub1* 3′ sequences were used as transcriptional terminators.

Discussion

In contrast to a growing demand, there is a lack of well defined nucleus-derived regulatory sequences of bryophytes. This may be due to the lack of public bryophytic genome sequences, especially sequences with the potential for promoting the expression of recombinant proteins. We isolated and characterised endogenous genomic upstream regions, including 5′UTRs, capable of driving reporter gene expression in the well established model organism *P. patens*. In order to reduce the likelihood of unintended phenotypic variations by reinserting such endogenous sequences into a haploid gametophytic tissue highly capable of HR, we focused on EPRs of the redundant beta-tubulin gene family of

Table 3 Luciferase assay: *PpTub1* expression promoting regions in relation to 35S promoter

Name Constructs	Graphical Scheme Constructs	Promoter (5' of TSP)	Terminator [a]	n (trafos)	Relative Activity (%)	+/-
35S / luc		35S promoter	35S term.	20 (6)	100	12
PpTub1						
luc1		-360	633	17 (4)	144	51
luc2		- 98	368	20 (6)	30	34
luc3		+ 10	214	20 (5)	0	0

[a] numbering in bp 3' of stop codon

Physcomitrella, where functional redundancy can be expected.

The most promising candidates were *PpTub1–PpTub3*, due to their high representation in public cDNA libraries and the results of expression studies. Taken together, transcription of *PpTub1–PpTub3* was shown not to be differentially regulated, when different gametophytic tissue cultures (normal flask cultures, chloronema enriched flask cultures, adult gametophyte) are compared (Jost et al. 2004). Additionally, we show here that transcription of these three genes is comparable in bioreactor cultures and - concerning *PpTub2* and *PpTub3* - at identical levels in normal flask cultures (Fig. 1). Moreover, expression of all three genes is detectable in non-regenerating protoplasts, providing the possibility to quantify their upstream sequences in transient reporter gene expression systems.

A comparison of the isolated genomic upstream regions and determination of transcriptional initiation in silico revealed clearly defined TSPs and surrounding consensus sequences, an important aspect for the proper characterisation of promoters (Fig. 2). In our expression constructs, endogenous 5′ leader sequences (5′UTRs) are included. Plant 5′ leader sequences have been shown to affect gene expression at the level of translation, mRNA stability or even transcription (Hulzing et al. 2002; Hua et al. 2001; Bolle et al. 1994, 1996; Bate et al. 1996; Chabouté et al. 2002). In this context, it is remarkable that the *PpTub3* construct 3-3, harbouring only 60 bp of the 5′UTR, still exhibits around one-third of the 35S promoter activity (Table 2). In contrast, a nearly complete version of the *PpTub1* 5′UTR (from +10 to +126) does not show any detectable activity (Table 3), even though five out of 27 5′ ESTs start at position +83, which is included in this construct. This preliminary finding points to a putative role of the truncated *PpTub3* 5′UTR in the enhancement of gene expression.

Scaffold/matrix attachment regions, generally A/T-rich and several hundred basepairs in length (Liebich et al. 2002), have been shown to be able to enhance the transcription of adjacent genes in plants, primarily when integrated into the genome (Schöffl et al. 1993; van der Geest and Hall 1994). However, in some cases and to a lesser extent, this effect was also described in transient expression experiments (Allen et al. 1993, 1996). Although A/T richness per se is not a sufficient criterion to define scaffold/matrix attachment regions, it is striking that the addition of the *PpTub1* upstream sequence, which is nearly 1 kbp in length and shows an A/T content of over 80%, enhances transient expression significantly only if inserted completely (cf. Table 1, constructs 1-0, 1-1, 1-2).

In this work, we describe a novel and highly reproducible expression system based on the secretion of $rhVEGF_{121}$ from transiently transformed non-regenerating *Physcomitrella* protoplasts. Notably non-regenerating protoplasts, in which no cell wall synthesis and cell division takes place, provide an ideal tool for quantitative studies, since they are derived from one cell type only, the chloronema cell, mostly sharing the same cell cycle phase (G2; Schween et al. 2003). In contrast to regenerating protoplasts, which are undergoing cell wall synthesis followed by mitosis, non-regenerating protoplasts hold this homogenous status. The $rhVEGF_{121}$ assay under non-regenerating conditions interestingly gave similar results as compared to the dual-luciferase reporter assay, which was performed under regenerating conditions. Accordingly, the $rhVEGF_{121}$ assay resulted in low standard deviations. Furthermore, quantification of a secreted and therefore processed protein, as in our novel assay, reflects the situation of the target application more accurately than quantifying intracellular product levels.

In summary, we provide a fast and reliable tool for the quantitative analysis of putative EPRs, which in principle could be used not only in mosses; and we provide several moss EPRs of variable strength, with up to three-fold the strength of the 35S promoter. A next step in analysing these beta-tubulin EPRs in order to verify their suitability in expressing recombinant genes is to generate stably transformed plants. Preliminary investigations indicate that corresponding EPRs efficiently drive expression in stable lines: e.g. quantification of a secreted product in culture supernatants of several lines gave, respectively, up to 10 μg and 13 μg rhVEGF

per g dry weight for the *PpTub1* construct 1-2 and the *PpTub2* construct 2-2.

The endogenous transcript levels of *PpTub1–PpTub3* turned out to be very similar. However, our transient quantification data for the full-length constructs 1-0, 2-0 and 3-0 showed clear differences in strength. This could most easily be explained by the fact that transient quantifications were done in highly homogeneous non-regenerating protoplasts (not in complex tissues). Nevertheless, it cannot be excluded that additional regulatory sequences outside our constructs are used in the endogenous transcriptional regulation of these three genes.

Based on the data presented, further studies on the regulation of beta-tubulin genes in *Physcomitrella* become feasible. The targeted introduction of reporter genes into the genome by basepair-exact gene replacement via HR can be used to define the entire expression pattern of a particular beta-tubulin gene. Furthermore, stably transformed lines with a modified expression pattern for a given family member by removal, replacement or modification of its regulatory regions can help to resolve the developmental relevance of a given regulation pattern.

Acknowledgements We are grateful to Hermann J. Thierse who enabled the real-time PCR experiments. This work was partially supported by grant no. 0312624 from the German Federal Ministry of Science and Education (BMBF). All EPRs presented in this study are freely available for non-commercial use.

References

Allen GC, Hall GE Jr, Childs LC, Weissinger AK, Spiker S, Thompson WF (1993) Scaffold attachment regions increase reporter gene expression in stably transformed plant cells. Plant Cell 5:603–613

Allen GC, Hall G Jr, Michalowski S, Newman W, Spiker S, Weissinger AK, Thompson WF (1996) High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco. Plant Cell 8:899–913

Bate N, Spurr C, Foster GD, Twell D (1996) Maturation-specific translational enhancement mediated by the 5′UTR of a late pollen transcript. Plant J 10:613–623

Bolle C, Herrmann RG, Oelmuller R (1996) Different sequences for 5′ untranslated leaders of nuclear genes for plastid proteins affect the expression of the beta-glucuronidase gene. Plant Mol Biol 32:861–868

Bolle C, Sopory S, Lubberstedt T, Herrmann RG, Oelmuller R (1994) Segments encoding 5′ untranslated leaders of genes for thyklakoid proteins contain cis-elements essential for transcription. Plant J 6:513–523

Breviario D, Nick P (2000) Plant tubulins: a melting pot for basic questions and promising applications. Transgenic Res 9:383–393

Chabouté ME, Clèment B, Philipps G (2002) S phase and meristem-specific expression of the tobacco *RNR1b* gene is mediated by an E2F element located in the 5′ leader sequence. J Biol Chem 277:17845–17851

Cove JD, Knight CD, Lamparter T (1997) Mosses as model systems. Trends Plant Sci 2:99–105

Decker EL, Reski R (2004) The moss bioreactor. Curr Opin Plant Biol 7:166–170

Geest AH van der, Hall TC (1994) The beta-phaseolin 5′ matrix attachment region acts as an enhancer facilitator. Plant Mol Biol 33:553–557

Gorr G(1999) Biotechnologische Nutzung von *Physcomitrella patens* (Hedw.) B.S.G. PhD thesis, University of Hamburg, Hamburg

Hohe A, Decker EL, Gorr G, Schween G, Reski R (2002) Tight control of growth and cell differentiation in photoautotrophically growing moss (*Physcomitrella patens*) bioreactor cultures. Plant Cell Rep 20:1135–1140

Horstmann V, Hüther CM, Jost W, Reski R, Decker EL (2004) Quantitative promoter analysis in *Physcomitrella patens*: a set of plant vectors activating gene expression within three orders of magnitude. Biotechnology 4:13

Hua XJ, Cotte B van de, Van Montagu M, Verbruggen N (2001) The 5′ untranslated region of the *At-P5R* gene is involved in both transcriptional and post-transcriptional regulation. Plant J 26:157–169

Hulzing RJM, Groot PFM de, Croes AF, Quaedvlieg W, Twell D, Wullems GJ, Herpen MMA van (2002) The 5′ untranslated region of the *ntp303* gene strongly enhances translation during pollen tube growth, but not during pollen maturation. Plant Physiol 129:342–353

Jost W, Baur A, Nick P, Reski R, Gorr G (2004) A large plant beta-tubulin family with minimal C-terminal variation but differences in expression. Gene 340:151–160

Koprivova A, Stemmer C, Altmann F, Hoffmann A, Kopriva S, Gorr G, Reski R, Decker EL (2004) Targeted knockouts of *Physcomitrella* lacking plantspecific immunogenic *N*-glycans. Plant Biotechnol J 2:517–523

Liebich I, Bode J, Reuter I, Wingender E (2002) Evaluation of sequence motifs found in scaffold/matrix-attached regions (S/MARs). Nucleic Acids Res 30:3433–3442

Littauer UZ, Giveon D, Thierauf M, Ginzburg I, Ponstingl H (1986) Common and distinct tubulin binding sites for microtubule-associated proteins. Proc Natl Acad Sci USA 83:7162–7166

MacRae TH (1997) Tubulin post-translational modifications. Enzymes and their mechanisms of action. Eur J Biochem 244:265–278

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

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

Reski R, Abel WO (1985) Induction of budding on chloronemata and caulonemata of the moss, *Physcomitrella patens*, using isopentenyladenine. Planta 165:354–358

Reski R, Cove DJ (2004) *Physcomitrella patens*. Curr Biol 14:261–262

Reski R (1998) Development, genetics and molecular biology of mosses. Bot Acta 111:1–15

Richter U, Kiessling J, Hedtke B, Decker EL, Reski R, Börner T, Weihe A (2002) Two RpoT genes of *Physcomitrella patens* encode phage-type RNA polymerases with dual targeting to mitochondria and plastids. Gene 290:95–105

Rother S, Hadeler B, Orsini JM, Abel WO, Reski R (1994) Fate of a mutant macrochloroplast in somatic hybrids. J Plant Physiol 143:72–77

Rothnie HM (1996) Plant mRNA 3′ end formation. Plant Mol Biol 32:43–61

Sambrook J, MacCallum P, Russell D (2000) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Schaefer DG, Zryd JP, Knight CD, Cove DJ (1991) Stable transformation of the moss *Physcomitrella patens*. Mol Gen Genet 226:418–424

Schaefer DG, Zryd JP (1997) Efficient gene targeting in the moss *Physcomitrella patens*. Plant J 11:1195–1206

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

Schoeffl F, Schroeder G, Kliem M, Rieping M (1993) An SAR sequence containing 395 bp DNA fragment mediates enhanced, gene-dosage-correlated expression of a chimaeric heat shock gene in transgenic tobacco plants. Transgenic Res 2:93–100

Schween G, Gorr G, Hohe A, Reski R (2003) Unique tissue-specific cell cycle in *Physcomitrella*. Plant Biol 5:50–58

Seki M, Carninci P, Nishiyama Y, Hayashizaki Y, Shinozaki K (1998) High-efficiency cloning of Arabidopsis full-length cDNA by biotinylated CAP trapper. Plant J 15:707–720

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

Sugahara Y, Carninci P, Itoh M, Shibata K, Konno H, Endo T, Muramatsu M, Hayashizaki Y (2001) Comaparative evaluation of 5′ end-sequence quality of clone CAP trapper and other full-length-cDNA libraries. Gene 263:93–102

Wahle E, Rueegsegger U (1999) 3′ End processing of pre-mRNA in eukaryotes. FEMS Microbiol Rev 23:277–295

Weise A, Rodriguez-Franco M, Timm B, Hermann M, Link S, Jost W, Gorr G (2004) Isolation of four members of a plant actin gene family with remarkable gene structures and the use of their 5′ regions for high transgene expression. (in preparation)

Zeidler M, Hartmann E, Hughes J (1999) Transgene expression in the moss Ceratodon purpureus. J Plant Physiol 154:641–650