Evaluation of synthetic promoters in Physcomitrella patens

Peramuna, Anantha; Bae, Hansol; Rasmussen, Erling Koch; Dueholm, Bjørn; Waibel, Thomas; Critchley, Joanna H.; Brzezek, Kerstin; Roberts, Michael; Simonsen, Henrik Toft

Published in:
BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS

Link to article, DOI:
10.1016/j.bbrc.2018.04.092

Publication date:
2018

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Evaluation of synthetic promoters in *Physcomitrella patens*

Anantha Peramuna\(^1\), Hansol Bae\(^1,2\), Erling Koch Rasmussen\(^1\), Bjørn Dueholm\(^1\), Thomas Waibel\(^3\), Joanna H. Critchley\(^3\), Kerstin Brzezek\(^3\), Michael Roberts\(^3\), Henrik Toft Simonsen\(^1,2\)*

\(^1\)Department of Biotechnology and Biomedicine, Technical University of Denmark, Søltofts Plads 223, 2800 Kgs. Lyngby, Denmark

\(^2\)Mosspiration Biotech IVS, Hørsholm, Denmark

\(^3\)Synpromics, Edinburgh, Scotland

* hets@dtu.dk

**Highlights**

- *Cis*-elements can be randomly assembled to construct short plant promoters
- Synthetic promoters have higher mRNA expression level than the endogenous *PpAct7*
- Synthetic promoters have similar protein expression levels to *AtUBQ10* promoter
Abstract

Securing a molecular toolbox including diverse promoters is essential for genome engineering. However, native promoters have limitations such as the available number or the length of the promoter. In this work, three short synthetic promoters were characterized by using the yellow fluorescent protein Venus. All of the tested promoters were active and showed mRNA activity higher than housekeeping gene PpAct7, and similar protein expression level to AtUBQ10 promoter. This study shows that few cis-elements are enough to establish a strong promoter for continuous expression of genes in plants. Along with this study, enhance the number of available promoters to be used in P. patens. It also demonstrate the potential to construct multiple non-native promoters on demand, which would aid to resolve the bottleneck issue of multiple pathway expression in P. patens and other plants.

Keywords: Physcomitrella patens; Venus; Ubiquitin promoter; Actin promoter; Synthetic Biology; synthetic promoters
Introduction

Genes in eukaryotes consist of three major regions, the promoter, the coding strand and the terminator. The promoter controls the gene expression by regulating the binding of transcription factors to recruit RNA polymerase (Latchman 1997). Thus, the synthesis of mRNA is directly correlated with the promoter activity that may lead to the production of protein. The promoter is situated thousands of base pairs (bp) upstream from the transcription start site (TSS), to about 30 bp downstream from the TSS (Porto et al. 2014). Promoters come in different type of switches, some are constitutive, some react to specific stimuli, and some are inducible and maintain a strict on/off style switch.

Plants are known to have larger promoter sequences than those found in fungi and prokaryotes. Plant promoters typically range from 500 bp to over 2,000 bp (Liu and Stewart 2016). As the possibilities in synthetic biology advances, it is often necessary to introduce multiple genes and promoters to achieve the desired traits. However, endogenous plant promoters are often of limited use in plant synthetic biology as multiple copies of the same promoter can trigger homology-dependent gene silencing (Halpin 2005). Therefore, characterizing multiple promoters from heterologous species has become important for fine-tuning of multiple genes.

Physcomitrella patens is a plant model system that has been used extensively to study plant evolution, physiology, and development (Vesty et al. 2016). The full genome is sequenced (Rensing et al. 2007) and development growth media and transformation methods are well described (Bach et al. 2014). Its ability to perform efficient homologues recombination, that now can be explained with RecQ helicase function (Wiedemann et al. 2018), is unique among plants enabling in vivo assembly of multiple DNA fragments followed by targeted genome integration by homologues recombination (King et al. 2016). Collectively, such distinct features make P. patens attractive as an industrial
production platform for small natural products, which requires integration of numerous genes (Zhan et al. 2014; Ikram et al. 2015; Pan et al. 2015; Sabovljević et al. 2016; Khairul Ikram et al. 2017).

Monocot housekeeping gene promoters and the 35S promoter from Cauliflower mosaic virus (CaMV) have shown a high-level of gene expression in *P. patens* (Horstmann et al. 2004; Saidi et al. 2005).

However, the limited number of available promoters are a bottleneck in introducing complex pathways into *P. patens* and typically long plant promoters are difficult to handle in multiple gene integrations. Therefore, developing strong and short synthetic promoters has emerged as a major interest to provide a solution.

Synthetic promoters are relatively short (300-500 bp) and can be generated in countless number with similar strength or function, and could improve the genome stability (Roberts 2011). *Cis*-regulatory elements (CREs) of native promoters are non-conserved among genes that are similarly expressed. Thus, the synthetic promoters can be reasonably constructed to give high gene expression with a smaller size (Liu and Stewart 2016). The strength of synthetic promoters depends on the selection, spacing of CREs and the copy number. Using bioinformatic algorithms, novel CREs could be discovered, by comparing the upstream sequence of differentially regulated genes. These CREs could be put together rationally to design new synthetic promoters (Roberts 2011). Furthermore, the strength of the synthetic promoter could be enhanced by proper spacing and increasing the CRE copy number (Liu and Stewart 2016).

Here, we have tested three synthetic promoters in *P. patens*, developed using an automated high-throughput screening method. The promoters were built using computational analysis of large transcriptomic functional data set to identify *cis*-elements, which form the building blocks of synthetic promoter libraries. All three synthetic promoters had higher mRNA expression than the housekeeping gene *PpAct7* and showed similar protein expression pattern to *AtUBQ10* promoter.
Materials and Methods

Promoter construction

The synthetic promoter library was constructed at Synpromics using random assembly techniques of cis-elements through expression data analysis of Zea mays. Genes showing strong expression strength above Ubiquitin1 transcripts were labeled constitutive. Using transcription factor binding site database TRANSFAC, cis-elements of the constitutively expressed genes were identified from the 1,500bp upstream and 500bp downstream of the transcription start site. Subsequently, Synpromics Ltd Syn-score algorithm was applied to the identified regions to rank the cis-elements (Roberts et al. 2017). Further, a synthetic promoter library was constructed by, random assembly of the chosen cis-elements (300bp-800 bp), attached upstream to CaMV 35S minimal promoter (position -46 to +89). Later, functional promoters were identified by the expression analysis of the Luciferase gene.

Growth media

P. patens (Gransden ecotype, International Moss Stock Center #40001) was grown on solid and liquid PhyB media (Bach et al. 2014) under sterile conditions, with continuous 20–50 W/m² light intensity at 23°C.

DNA preparation and transformation protocol

DNA fragments for transformation were prepared in blocks as below. First block, a 2.7 kb region with 108 5’ neutral locus, G418 selection marker with CaMV 35S promoter/ CaMV poly(A) signal was amplified from the pRH004 plasmid. Second block, the synthetic promoter sequences developed by Synpromics, and the Arabidopsis Ubiquitin10 (AtUBQ10) promoter was amplified with 20–22 nt overhangs homologous to block one and three. Block 3, with the Venus fluorescent protein, OCS terminator and the 108 locus homologous recombination flanking region was amplified from pRH004
plasmid (Figure 1A). Purified 1.5 pmol of each DNA block (Figure 1B) was transfected into the isolated moss protoplasts during transformation process and selected for positive colonies according to previously published methods (King et al. 2016; Khairul Ikram et al. 2017).

Arabidopsis Ubiquitin10 promoter

Arabidopsis UBQ10 promoter with a length of 634 bp of (Grefen et al. 2010) was cloned from Arabidopsis genomic DNA using primer set; UBQ10 F 5’-GTCGACGAGTCAGTAATAAACGG-3’ and UBQ10 R 5’-CTGTTAATCAGAAAAACTCAGATTAATC-3’. For moss transformation, 22-nt overhangs that are identical to the next fragments (block one and three) were attached to both ends by second PCR using overhang primers.

Detection of Venus fluorescence

Venus fluorescence was detected on protonemal cells grown for seven days in PhyB liquid media. P. patens protonema cells were visualized and photographed using a confocal laser-scanning microscope. Z-stacks were performed on each line using the 488nm laser line and YFP emission filter. Z-stacks were put together using the Zeiss software built-in maximum projection function. Fluorescent level of each promoter lines was calculated from digital Images using the software ImageJ (https://imagej.nih.gov/ij/). A previously published method on fluorescent cell analysis was used to calculate the corrected total cell fluorescence (CTCF) levels (Burgess et al. 2010; McCloy et al. 2014).

For each cell, measurements were taken for the cell area, integrated density and mean grey value. Final corrected total cell fluorescence (CTCF) was calculated using the following formula.

\[ \text{CTCF} = \text{Integrated Density} - (\text{Area of selected cell} \times \text{Mean fluorescence of background reading}) \]
RNA extraction and qPCR

Total RNA was extracted from the appropriate lines (7 days after blending), using Spectrum™ Plant Total RNA Kit (Sigma, STRN250). To synthesize cDNA, 1 μg of extracted total RNA was reverse transcribed by iScript cDNA synthesis kit (Bio-rad, 1708891), followed by PCR amplification of the following transcripts, *PpAct7* and *Venus*. PCR reactions were carried out using (Qiagen kit name) and (Bio-rad machine name), by denaturation at 95°C 5 min, 40 cycles with 95°C for 10 sec and 60°C for 10 sec, and melting curve analysis to check the specificity. Relative *Venus* gene expression from each promoter line was analyzed by \( \text{Exp}_{\text{Venus}} = 2^{\Delta \text{Ct}[\text{promoter}]} \), \( \Delta \text{Ct}[\text{promoter}] = \text{Ct}[\text{Actin}] - \text{Ct}[\text{Venus}] \).
**Results and discussion**

*Assembly of cis-elements to construct constitutive plant promoters*

To generate synthetic promoters, functional cis-elements should be collected since cis-elements will form the building blocks of synthetic promoters. We used automated high-throughput screening method. In this method, computational analysis of large transcriptomic functional data sets of *Zea mays* was used to identify cis-elements from constitutively expressed genes. We ranked collected cis-elements by applying Syn-score algorithm and randomly assembled selected cis-elements to generate promoter library (Roberts et al, 2017). This technique has an advantage since elements are selected based on the requirements for the synthetic promoters (e.g., inducible, constitutive and tissue- or developmental stage-specific), which is a more focused approach than using completely random elements. The promoter candidates consist of randomly assembled cis-elements of varying lengths up to 30bp. Therefore, the promoter length and the position of cis-elements vary (Figure 1C). It has been shown that the position of cis-elements relative to each other markedly influences promoter strength (Rushton *et al.* 2002), which was shown in the transcription data of the synthetic promoters we tested (Figure 2).

The synthetic promoters can be used for gene stacking in *P. patens*, as they do not resemble each other on the sequence level and contain small cis-elements rather than large promoter fragments, which greatly reduces the risk of homology-induced gene silencing. As it is an automated system, it is reproducible, robust and faster than a manual approach. Thus, multiple functional promoters in *P. patens* could be constructed on demand.

*Promoter activity analysis of synthetic promoters*

To analyze the activity of synthetic promoters, we used fluorescence protein Venus. The coding sequence of *Venus* was placed downstream of each promoter including well-known constitutive
AtUBQ10 promoter, and all constructs were stably transfected to *P. patens* 108 neutral locus (Bach *et al.* 2014). We performed qPCR to compare the promoter activity. The mRNA expression level of *Venus* was calculated in relative to the endogenous housekeeping gene *Actin7* (*PpAct7*). Actin is an essential component of the plant cytoskeleton and is known to be a ubiquitous protein that is constitutively expressed in eukaryotes (Meagher *et al.* 1999). It is also shown that *PpAct7* is preferable to be used as a housekeeping gene (Le Bail *et al.* 2013) and perform dual functions as a control and a housekeeping gene in the data analysis of this experiment. Expression of human VEGF protein using the 5′ promoter region of the *PpAct7* depicted an eight-fold increase in the production of the VEGF protein compared to the constitutive CaMV 35S promoter (Weise *et al.* 2006). This suggests *PpAct7* is highly expressed compared to the 35S promoter that was quantified previously (Horstmann *et al.* 2004). Thus, the mRNA expression of *PpAct7* can be compared to the *Venus* transcript levels driven by the synthetic promoters. All three synthetic promoters showed higher expression than *PpAct7* (Figure 2). Synthetic promoter I2-10 yielded the highest mRNA expression level at 304.51 ± 4.21 fold relative to *PpAct7* followed by 45.91 ± 0.88, 124.93 ± 3.4, 258.38 ± 0.09 for I2-48, I2-79, and AtUBQ10, respectively. The I2-10 promoter was 2.43X and 6.63X expressed than I2-79 and I2-48. This high level of expression can be attributed to the presence of G-box in I2-10 (Figure 1A). The G-box is a regulatory element in plant promoters, playing an essential role in plant promoter responsiveness to light, stress, and hormones (Menkens *et al.* 1995). Some of the G-box motifs have been shown to aid high-level constitutive protein expression in some plant species (Ishige *et al.* 1999).

Protein expression analysis of synthetic promoters in *P. patens* protonema cells

After mRNA expression analysis, we wanted to identify if the protein expression would show similar pattern, because protein level does not always correspond to that of mRNA. The Venus protein
expression level was calculated for each synthetic promotor line. This was performed in 14 days old
*P. patens* protonema cells (Figure 3). Micrographs of \( I2-10 \), \( I2-48 \), \( I2-79 \), and \( AtUBQ10 \) promoter lines were processed to measure the fluorescence intensity of Venus via ImageJ.

All three synthetic promoters showed a similar level of Venus fluorescent protein expression (Figure 3). Compared to the medium-strength of the \( AtUBQ10 \) promoter (Grefen et al. 2010), the synthetic promoters, \( I2-10 \), \( I2-48 \) and \( I2-79 \) displayed 1.6X, 1.5X and 1.5X decrease in protein expression, respectively. All synthetic promoters share the identical sequence -123 bp from ATG, thus we excluded the possibility of ribosome entry for the discrepancy in expression of mRNA and protein level, and assume structure-related factor made the difference. These synthetic promoter candidates were derived from a screen conducted in Zea Mays, where Synpromics has shown that they mediate two- four-fold higher protein expression (firefly luciferase) compared to the ubiquitin-1 promoter. As one would expect we have seen that the promoters show maximum activity in the organism in which they were screened. The fact that promoters developed for another plant species show such high activity in \( P. patens \), bodes well for the further development of promoters for use in this chassis organism.

The short length of 634 bp makes the \( AtUBQ10 \) promoter an ideal control. However, the \( AtUBQ10 \) promoter consisted of the first 5’ intron expanding 304 bps, while synthetic promoters contained none. Deletion of the first 5’ intron in \( AtUBQ10 \) has shown to result in 3-fold lower protein activity than the first intron intact (Norris et al. 1993). Thus, adding an either synthetic or natural first 5’ intron section to the synthetic promoters would likely lead even higher expression of the following coding sequences.

In conclusion, we have tested the activity of three synthetic promoters in \( P. patens \). All three promoters showed high expression of mRNA compared to the \( PpAct7 \) and similar protein activity to
the medium-strength *AtUBQ10* promoter. Previously, published works have revealed that the addition of the first 5’ intron increase the stability of the mRNA and yield several folds of higher protein activity. Thus, adding a 5’ intron to the end of synthetic promoters would likely increase the strength of the promoters. Further, range of promoters mediating a range of different expression levels is essential in building genetic circuitry in synthetic biology applications, such that synthetic promoters can control the correct stoichiometry of different component proteins of the circuit at the transcriptional level.
Reference


**Figure text**

**Figure 1:** A) vector map of linearized pRH004 vector. B) Three PCR fragments (blocks) were constructed to transform the promoters to the *P. patens* 108 neutral loci. Block 1 was amplified from the pRH004 vector with the 108 neutral locus and resistance to Geneticin driven by CaMV 35S promoter. Block 2 depicts, PCR amplified *AtUBQ10, I2-10, I2-48* and *I2-79* promoter sequences. Followed by the final block 3, containing the Venus fluorescent protein with the OCS terminator sequence and the 108 loci. C) Schematic of randomly assembled cis regulatory elements to construct the synthetic promoters.

**Figure 2:** Overview of promoter strength based on *Venus* expression, using *PpAct7* promoter as background measurement.

**Figure 3:** Confocal images of Venus expression, driven by synthetic promoters and *AtUBQ10* promoter. All lines are grown and imaged in identical conditions. B) Representative, promoter strengths of *I2-10* (n=54), *I2-48* (n=41), *I2-79* (n=48) and *AtUBQ10* (n= 32) lines. Promoter strength was calculated by measurement of the fluorescence intensity for each cell by the use of confocal micrographs and ImageJ software. Scale bar =0.01 mm