# Molecular genetic approaches to study the roles of SHORTLEAF and SCARECROW in moss gametophytic development.

A thesis submitted to

Indian Institute of Science Education and Research, Pune in partial fulfilment of the requirements for the MS degree


Indian Institute of Science Education and Research, Pune

Dr. Homi Bhabha road, Pashan, Pune, 411008, India

**Gargi Chaturvedi**

**20172010**

**Biology Department, IISER Pune**

**Under the supervision of**

**Prof. Anjan Banerjee**

**Indian Institute of Science Education and Research, Pune**

# CERTIFICATE

This is to certify that this dissertation entitled "**Molecular genetic approaches to study the roles of SHORTLEAF and SCARECROW in moss gametophytic development**" towards the partial fulfilment of the MS degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Gargi Chaturvedi at IISER Pune under the supervision of Prof. Anjan Banerjee, Professor, Indian Institute of Science Education and Research during the academic year 2019-2020.

30/6/2020

Signature of the Supervisor

प्रो. अंजन के. बैनर्जी/Prof. Anjan K. Banerjee
प्राध्यापक, जीवशास्त्र कार्यक्रम
Professor, Biology Program
भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान
Indian Institute of Science Education & Research
पुणे, भारत /Pune- 411 008, India

Gargi Chaturvedi
30-6-2020

Signature of the student

# DECLARATION

I hereby declare that the matter embodied in the report entitled "**Molecular genetic approaches to study the roles of SHORTLEAF and SCARECROW in moss gametophytic development**" are the results of the work carried out by me at the Biology Department, Indian Institute of Science Education and Research, Pune, under the supervision of Prof. Anjan Banerjee and the same has not been submitted elsewhere for any other degree.

30/6/2020

Signature of the Supervisor

(Prof. Anjan Banerjee)

Date:

प्रो. अंजन के. बॅनर्जी/Prof. Anjan K. Banerjee
प्राध्यापक, जीवशास्त्र कार्यक्रम
Professor, Biology Program
भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान
Indian Institute of Science Education & Research
पुणे, भारत /Pune- 411 008, India

Gargi Chaturvedi
30-6-2020

Signature of the student

(Gargi Chaturvedi)

Date:

# ABSTRACT

*Physcomitrella patens* or moss, of the phyla Bryophyta belong to an ancient lineage of plants, which colonised land around 400mya after diverging from the vascular plant lineage. Leading a haplo-diplontic life cycle, with a dominant haploid stage, this model organism has a very simple yet sturdy body plan. The dominant body plan in moss is the gametophyte, which bears structures similar to those borne by sporophytic vascular plants. Moss bears leaf like structures known as phyllids, (which lack a proper vasculature) and root like organs known as Rhizoids. A forward genetics approach allows for the creation of mutants which target random regions in the genome. One such event carried out in the lab lead to the generation of a *shlf* (shortleaf) mutant, which as the name suggests bore phyllids which were smaller in size. The gene region that was targeted in the forward genetics approach was characterised to a certain extent during my study. *SHLF* contains four repeats , with each repeat being highly similar, one of my targets was to understand the functionality of a single repeat. *SHLF* also has paralogs in the sister clade of Marchantiophyta. Another one of my objectives aimed at understanding the extent of conservation of the gene across the Bryophyta lineage. I also wanted to study the localisation of the protein generated by this gene by raising GFP tagged lines. Understanding the localisation might help us understand the potential function of the protein.

Reverse genetics is another methodological approach to create mutants by specifically targeting a known region of the genome. Another such event targeting *SCARECROW* lead to the generation of a mutant bearing needle like leaves. This gene, in vascular plants is usually involved int patterning the development of the root vasculature and leaves to a certain extent. My study on this mutant included characterising the promoter activity by generation of fluorescent protein (tdTomato) tagged lines, followed by understanding the cell division pattern in the mutant phyllid.

Based on the results, the localisation of the protein SHLF could be determined, the functionality of a single repeat of the gene was better understood and a partial conservation of the paralog of *SHLF* could be concluded.

# CONTENTS

**Fig no. Figure Title Page no**

| Table no. | Table Title | Page no. |
|---|---|---|
| 1. | List of primers involved in the amplification of *SHLF* from pTFH15.3_*SHLF* | 19 |
| 2. | A table summarising required constructs and their purpose | 23 |
| 3. | Primer list for the amplification of MpHP50 from the cDNA of Tak1 | 38 |
| 4. | Primer list for the amplification of *SCR2* and *SCR3* promoter | 43 |

# ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Prof. Anjan Banerjee for his continuous support, patience, motivation, enthusiasm and wonderful guidance. I feel privileged to have been mentored by him during this project work.

Besides my mentor I want to thank Dr. Amey Bhide, Dr. Boominathan Mohanasundaram, Shirsa Palit for their guidance and encouragement. I would also like to thank my lab members – Harpreet Singh Kalsi, Nilam Malankar, Nikita Patil, Apurva Kute, Anindita Karkhanis and Tanvi Gupta for their insightful comments and hard questions.

Lastly, I would like to thank my family and friends, who have constantly supported and encouraged me in this endeavor.

# TOPICS

**1.Generation of GFP fusion lines of SHLF to study its sub-cellular localization in the moss, *Physcomitrella patens.***

**2.Functional validation of the first repeat of moss *SHLF* in the mutant *shlf* background**

**3. Overexpression of a high expressing *SHLF* paralog from *Marchantia polymorpha* in *shlf* mutant background**

**4. Understanding the role of *SCARECROW* in *Physcomitrella patens* for leaf development.**

# 1. A BRIEF INTRODUCTION

The Bryophyte lineage of plants diverged from the vascular plant lineage nearly 400 mya. They are one of the early land colonisers which evolved structures akin to those present in the vascular plants seen today (Morris *et al.* 2018). This could be a result of convergent evolution. Structures like leaves and roots which are borne by vascular plants have non vascular counterparts in the leafy bryophytes including leaf like organs known as phyllids and root like organs known as rhizoids.

Mosses (Physcomitrella) and some of the liverworts (Haplomitrium) have a leaf-like body plan. These are leaf-like lateral organs are arranged spirally around the stock of the gametophore in case of mosses and in rows in case of liverworts (Ligrone et al 2012). Leaves are known to assist the plant body in functions like transpiration, gaseous exchange, temperature homeostasis, photosynthesis etc. In vascular plants, leaves are the lateral appendages

borne by the dominant sporophyte (diploid plant body) and can be defined by the following features- vascularization, determinate growth, adaxial-abaxial polarity and a definite arrangement- phyllotaxy. Due to the important roles played by the leaves their origin has been studied in great detail. Studies on extant euphyllophytes suggests that leaves have evolved independently across plant lineages and can be broadly categorised as megaphylls, microphylls and phyllids. The leafless ancestors of all vascular plants, from the late Silurian and the early Devonian period can be divided into two clades- the zosterophylls and trimerophytes. They evolved microphylls (lycophytes) and meagaphylls (euphyllophytes) respectively. (Tomescu,2009). Phylogenetic and morphological analysis has suggested that megaphylls have evolved, independently, a total of 9 times, whereas microphylls have a single common origin. According to Zimmerman's (1930) telome theory, megaphylls, have evolved from an algal like ancestor, which branched dichotomously. The term telome was assigned to each terminal region of the branch. On the other hand, the widely accepted theory about the evolution of microphylls is the enation theory (Bower,1935). This theory states that a leafless Rhynia-like ancestor developed outgrowths, or enations which were later enriched with a vascular system. Theories dealing with the origin of phyllids remain unanswered leaving us with many possibilities which can be explored.

Phyllids in Bryophytes arise as result of division of the gametophore apical cell. This division leads to the formation of a wedge-shaped cell known as the leaf founder. This divides in an oblique fashion giving rise to proximodistal and medio-lateral polarity. (Harrison et al 2009).

The tetrahedral apical cell divides on three of its faces in a spiral phyllotaxy. This spiral phyllotaxy follows the golden angle ratio (Harrison et al. 2009). Such a spiral arrangement of the phyllids allows each one to be adequately exposed to the sunlight. The phyllid lack any known vasculature. This could

be one of the reasons why the plant body is restricted to millimetres in height. The phyllids also bear a midrib which divides in a particular fashion governed by auxin gradients to allow for the development of the typical oblanceolate shape.

## 1. Generation of GFP fusion lines of SHLF to study its sub-cellular localization in the moss, *Physcomitrella patens.*

*Physcomitrella patens* has been used as a model system for non-seed plants, due to the ease in its culturing under laboratory conditions. The dominant haplontic nature of this plant system allows it to be manipulated as a tool for genetic studies due to the ease of homologous recombination (Rensing et al ,2020).

In biological systems it is essential to understand the functioning of the genes in order to determine how the mechanics of the system works. Forward genetics is an unbiased tool allowing for generation of mutants in order to understand their functional importance. This technique involves randomly mutagenizing the genome of the organism in order to look at all potential physical outcomes of this process. Once a random mutant with an altered phenotype is generated, the only job left to do is trace the mutant phenotype back to a possible region in the genome which is affected (Moresco et al., 2013). This approach has led to some important novel discoveries. A forward genetics mutagenesis event carried out in the lab on a wild type background of *Physcomitrella patens* (Gransden ecotype) lead to the generation of multiple mutants (Mohanasundaram et al, 2019). One such mutant displayed a phenotype of unusually short and was thus given the tern *shlf* (Short leaf). This mutant displayed a couple of traits including- two times shorter leaves, increased internodal distance and a decrease in the

overall number of plasmodesmata connections across the phyllid. (Mohanasundaram et al, 2019, Under revision).

*Agrobacterium* mediated Tnt1 mutagenesis of moss was the strategy adopted for the generation of mutant lines. Overexpressing the gene *SHLF* resulted in the reversal of the phenotype to WT. In order to characterize the function of any gene forming a protein it is essential to understand where within the subcellular space is the protein being localised to. The structure of the *SHLF* gene consists of 2169bp long stretch, with a 72 bp long 'N' terminal region.

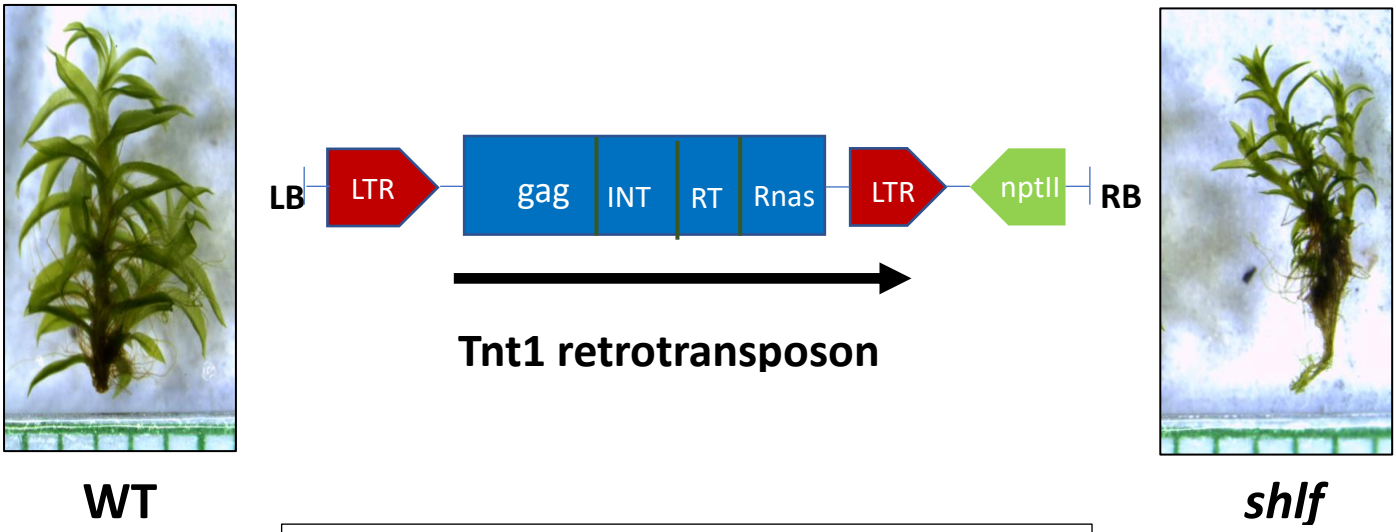



Figure1: Tnt1 retrotransposon mediated mutagenesis in Moss

This N- terminal region contains an eight amino acid long ER localisation signal as detected by the Signal P software. In order to understand and visualise subcellular localisation of the protein GFP fusion constructs were generated. A 'Two Fragment Gateway Cloning' strategy was used to prepare the required constructs in order to be expressed in the *shlf* background.

**RATIONALE: To study the subcellular localisation of the protein *SHLF* in order to understand its potential function.**

## 2. Functional validation of the first repeat of moss SHLF in the mutant *shlf* background

The structure of the SHLF protein consists of an N terminal with a signal peptide potentially targeting the protein to the Endoplasmic Reticulum. This SHLF gene is 2.1 Kb consists of an N terminal region followed by 4 repeats and the C terminal region.The 2$^{nd}$, 3$^{rd}$ and the 4$^{th}$ repeat, each of 513 bp, are 99% identical whereas the 1$^{st}$ repeat shares a 90% similarity with the repeats (Mohanasundaram et al, 2019, under revision). There is an 8 amino acid ER localisation signal at the N terminal region. These repeats form 171 amino acid long Tandem Direct Repeats. Such repeat containing proteins are found in nature including Armadillo repeat protein which consists of a 40 amino acid long region which folds up to form a solenoid structure. It forms a part of many proteins for eg. β catenin, α importin etc. (Hatzfeld, 1999)

The Tnt1 insertion is located interspersed between the 4$^{th}$ repeat and the C terminal region. This might result in the formation of a non-functional protein or a truncated protein.


Figure 2: A pictorial representation of the *SHLF* gene structure- containing the N terminal region followed by the four repeats and the C terminal region. The Tnt1 insertion was present between the fourth repeat and the C terminal region

The strikingly high levels of similarities between the repeats allowed us to question whether each of these repeats, individually, had the capacity to perform the function of the entire protein. The idea was to understand whether a functional protein could be formed by the N terminal, a single repeat and the C terminal region, and whether this could rescue the mutant phenotype by reverting it to the Wild type condition.

**PROPOSED HYPOTHESIS: Whether a single repeat of the gene is enough to rescue the mutant phenotype**.

## 3. Overexpression of a high expressing SHLF paralog from *Marchantia polymorpha* in *shlf* mutant background

The SHLF family of protein was found to be highly unique to the Bryophyte lineage. Its homologs were found conserved mainly in the lower Streptophytes, with some degree of conservation across liverworts and algae such as *Klebsormidium*.(Mohanasundaram et al, 2019, Under Revision).

Homologs of this particular gene were also found in *Marchantia polymorpha*. There were however 5 paralogs to *SHLF* with varying levels of gene expression. The paralog Mapoly0112s0050 showed the highest level of expression when analysed via RNA Seq Analysis. This particular paralog showed the presence of five repeats as opposed to the four that are seen in *SHLF* of moss.


Figure 3: A pictorial representation of the MpHP50 (Mapoly0112s0050) gene structure- containing five repeats

*Marchantia polymorpha* is a liverwort which is dioecious. It bears separate thalli for both the male and female. The thalli bear gemma cups which containing the asexual reproductive structures known as gemma. These gemma, under natural conditions are displaced from the gemma cups under the influence of rain and land on the soil. Under appropriate conditions of light, moisture and nutrients these gemma sprout to give rise to the thallus all over again. The haploid phase of the *Marchantia* is predominant in its life cycle (Shimamura,2016; Bowman, 2016)

The presence of conserved homologs across species allows us to question whether a cross-species complementation event could result in the rescue of the mutant phenotype. And in doing so questions regarding the functional conservation of the gene across the Bryophyte lineage could be addressed.

a  b 

Figure4: a) Tak2 (female) and b) Tak1 (male) *Marchantia polymorpha*

**PROPOSED HYPOTHESIS: To understand whether a paralog of the SHLF protein could rescue the mutant phenotype in order to allow us to understand the functional conservation of its function across the Bryophyte lineage.**

## 4. Understanding the role of *SCARECROW* in *Physcomitrella patens* for leaf development.

Reverse genetics is an alternate approach to understanding the function of a gene. The approach involves genetically modifying a known region of the genome, followed by looking at the phenotypic changes that occur within the system in order to gain some insight into the function of the gene of interest. (Gilchrist et al, 2010)

*Scarecrow* is a transcription factor involved in maintaining the cell division pattern in the quiescent centre cells of the root and maintaining the surrounding stem cells in the root. In *Arabidopsis* it has been seen that the root contains four layers namely- the epidermis, the cortex, the endodermis and the pericycle. Each layer arises from a set of four individual initial germ cell namely- endodermal/cortical initial, endodermal/root cap initial, columella root cap

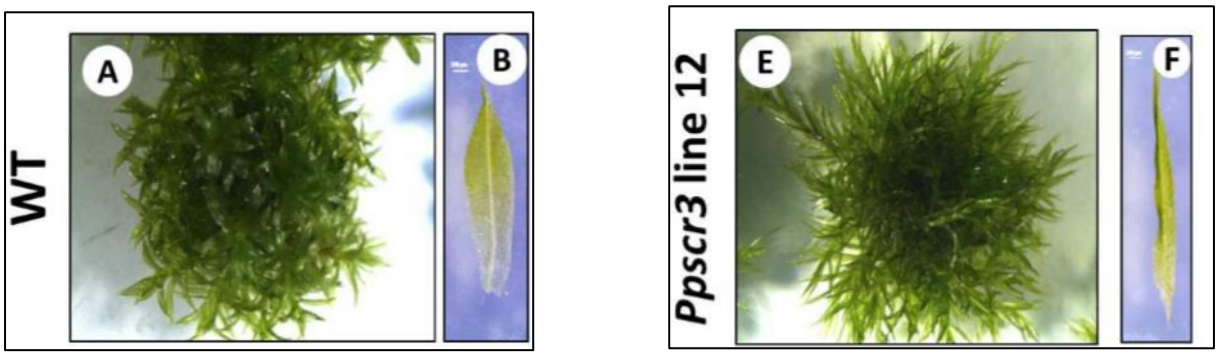



Figure 5: Comparison of the phenotypic differences between the Wild Type *Physcomitrella patens* and the *scr* mutant

initial and quiescent centre. This gene is involved in regulating an asymmetric cell division pattern in the endodermal/cortical initial cell. This initial cell first undergoes an anticlinal division followed by a periclinal cell division. This second division results in the differentiation of the cortical and the endodermal cell layers. Mutation in the *SCARECROW* gene results in the abolishment of the second periclinal division. As a result of this there is no division of the fates of the cortical and the endodermal cell layer and instead in its place there lies a mutated single layer displaying the properties of both the endodermis as well as the cortex (Laurenzio et al, 1996). *SCARECROW* and *SHORTROOT* are involved in maintaining the cell division pattern in the leaves of *Arabidopsis* (Dhondt et al, 2010). The do so by regulating the duration of the proliferation phase in the cell cycle of the leaves. Mutating these, result in hampered leaf development which is not primarily due to a disorganised root but directly a result of an early exit from the S phase of the cell cycle. These genes have been reported to be involved in normal shoot gravitropism. In order to understand

whether gene regulatory networks are conserved between *Arabidopsis* and moss, a reverse genetics screen was adopted targeting Moss *Scarecrow* homologs.

Three *SCR* paralogs namely -*PpSCR1, PpSCR2* and *PpSCR3* were detected with only *PpSCR3* displaying the phenotype of the slender leaf. In order to understand the spatio-temporal expression pattern of the *SCR* gene it is essential to understand and visualise the activity of the promoter. The objective of the experiments involved cloning the promoter of *PpSCR3* and *PpSCR2* along a fluorescent protein, which will then be expressed in the WT *P. patens* background.

**RATIONALE: To understand the gene regulatory pattern of *Scarecrow* in *Physcomitrella patens***

# 2.METHODS

## 1. Generation of GFP fusion lines of SHLF to study its sub-cellular localization in the moss, *Physcomitrella patens.*

### 1.1 Culturing of *Physcomitrella patens*

*Physcomitrella patens* (ecotype -Gransden ) was incubated at 16:8 ratio of light:dark cycle in the incubator which is maintained at 24°C for all experiments. (Cove et al 2009). The tissue was cultured on a BCDAT media, in 10cm petri dishes.

The following table gives the composition of the BCDAT media (CSH)

| Solution Name | Final concentration for 1Litre |
|---|---|
| B | 1mM MgSO4 |
| C | 1.84mM KH2PO4 |
| D | 10mM KNO3 |

| | |
|---|---|
| A | 5mM Diammonium Tartarate |
| T | TES (Hogland's trace elements) |

Composition of TES

| Reagent | Quantity for 1L | Final Concentration |
|---|---|---|
| Al2(SO4)3.K2SO4.24H2O | 55mg | 0.006%(W/V) |
| CoCl2.6H20 | 55mg | 0.006%(W/V) |
| CuSO4.5H2O | 55mg | 0.006%(W/V) |
| H3BO3 | 614mg | 0.061%(W/V) |
| KBr | 28mg | 0.003%(W/V) |
| KI | 28mg | 0.003%(W/V) |
| LiCl | 28mg | 0.003%(W/V) |
| MnCl2.4H2O | 389mg | 0.039%(W/V) |
| SnCl2.2H2O | 28mg | 0.003%(W/V) |
| ZnSO4.7H2O | 55mg | 0.006%(W/V) |
| H2O | Make it upto 1 litre | |

The protonema of moss was generated using a homogeniser (blade and shaft) and propagated on a cellophane sheet placed onto a BCDAT media containing petri plate. In order to maintain protonema for transformation, it was homogenised at a frequency of at least 5 days.

1.2 Amplification of *SHLF* from vector template pTFH15.3

Given below is a table containing the list of the primers involved in amplification of SHLF from the vector

| Primer name | Primer sequence |
|---|---|
| UHG attB1 HP F | GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACATGGCGTCCAGCTCCAGGGCCT |
| UHG attB5r HP R | GGGGACAACTTTTGTATACAAAGTTGTAGCGGATATCGCAATTTCTTTCATCG |
| UGH attB5 HP F | GGGGACAACTTTGTATACAAAAGTTGTAATGGCGTCCAGCTCCAGGGCCT |
| UGH attB2 HP R | GGGGACCACTTTGTACAAGAAAGCTGGGTTCTAAGCGGATATCGCAATTTCTTTC |

Table 1 : List of primers involved in the amplification of *SHLF* from pTFH15.3_*SHLF*

The following reaction conditions were employed in order to amplify the gene of interest-

For amplification the master mix contained

Buffer for Phusion polymerase (5X HF buffer)

Phusion Polymerase

Forward primer (with att sites) (5µm)

Reverse primer (with att sites) (5µm)

Vector template

dNTP's (10mM)

Milli Q water

Since the technique being employed was that of gateway cloning, there were two sets of conditions employed to amplify the gene of interest. One was for the C terminal GFP fusion construct -given the abbreviation UHG (Ubiquitin maize promoter- SHLF protein-GFP). This construct acts as a test to study the subcellular localisation of the protein. The other set of conditions would again amplify the same gene but the PCR product would be involved in the generation of the UGH or N terminal fusion product- given the abbreviation UGH (Ubiquitin maize promoter-GFP-SHLF). This construct would act as a control since the ER localisation signal present would no longer function as a signal peptide.

PCR conditions for UGH

| Temperature | Time |
|---|---|
| 98°C | 30" |
| 98°C | 10" |
| 60.9°C | 30" |
| 72°C | 1'25" |
| Go to step 2 | 5 times |
| 98°C | 10" |
| 72°C | 30" |
| 72°C | 1'25" |
| Go to 6 | 30 times |
| 72°C | 3' |
| 4°C | Forever |
| END | |

| Temperature | Time |
|---|---|
| 98°C | 30” |
| 98°C | 10” |
| 62.4°C | 30” |
| 72°C | 1’25” |
| Go to step 2 | 5 times |
| 98°C | 10” |
| 72°C | 30” |
| 72°C | 1’25” |
| Go to 6 | 30 times |
| 72°C | 3’ |
| 4°C | Forever |
| END | |

PCR conditions for UHG

The PCR product so generated was gel eluted from an 0.8%agarose gel. This gel elute was purified using Promega gel purification kit.

The gel was weighed and an equal amount of MBS (Membrane Binding Solution) was added and the mixture was heated at 70°C for nearly 20-25 mins. Till the gel had melted.

1. The mixture was passed through a binding column.
2. The binding column was washed with membrane washing solution containing ethanol
3. The column was allowed to dry and the DNA was eluted in Milli Q water and heated at 65°C for 2-5’.
4. The gel eluted was then tested on a 0.8%agarose gel.
5. It quality and quantity was tested. The quality should by a good 260/280 ratio of nearly 0.8 and a good 260/230 ratio for measuring the purity of the nucleic acid. This needs to be ensured in order to proceed with further reactions. A low concentration and a poor quality may hamper with the BP raction.

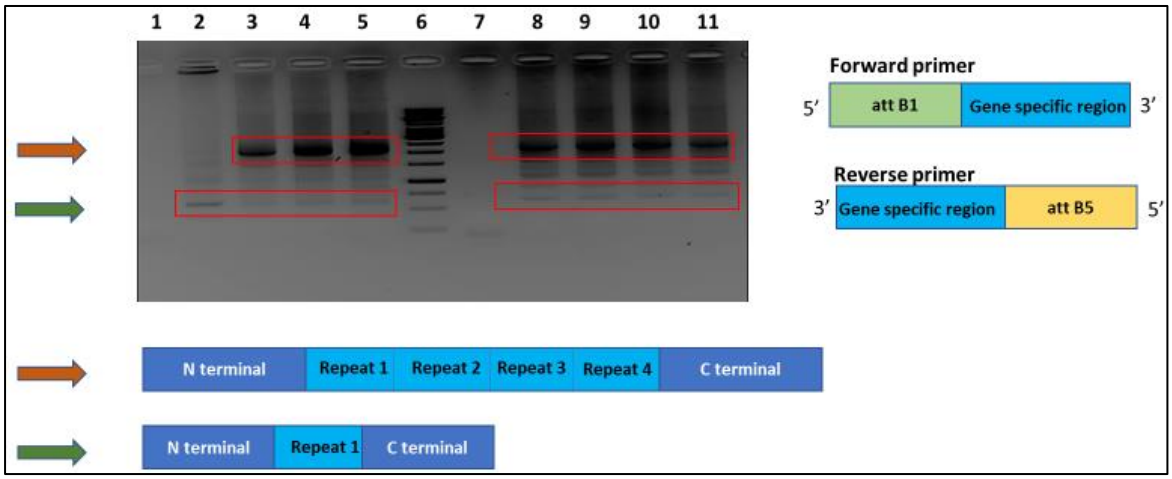



Figure 6: PCR amplification of *shlf* (entire gene and first repeat) using att site containing primers. The Red arrow marks the position of the entire amplicon of *shlf* whereas the arrow in green marks the position of the first repeat. Bothe will be used for proceeding with a BP reaction.

1.3 Using two-fragment gateway cloning approach to generate destination vectors

Gateway cloning is technique employed to insert a gene of interest into a vector without the requirements of the traditional restriction digestion and ligation protocol. This technique reduces a lot of the steps involved in troubleshooting during cloning. This technique employs the use of the mechanism of DNA integration by λ phage into the *E. coli* genome. The genome of the phage contains att P sites and those of the bacteria contain att B sites. Factor like Int (Integrase)- a phage protein and IHF (Integration Host Factor)- a bacterial protein interact and result in the integration of the phage genome. This leads to the generation of att L and att R sites flanking the region that has been integrated. The reverse reaction can be carried out when another phage protein known as Xis (Excisionase) functions alongside the Int and IHF. This reverts the att L and att R sites back to P and B. The BP clonase contains Int, IHF and the LR clonase contains Int,IHF and Xis.(Reece et al 2018, Park et al, 2015)

Given below is an overview of the experimental design involved in executing the following project:[3]

| Aim | To study the sub-cellular localisation | | To understand the sufficiency of single repeat | |
|---|---|---|---|---|
| Constructs | mUBI::SHLF-eGFP | mUBI::eGFP-SHLF | mUBI::SHLF(1R)-mRUBY | mUBI::SHLF(1R)-mEGFP |
| Purpose | Test | Negative ctrl | test | test |
| | Intact ER signal | Disrupted ER signal | A single repeat of SHLF | A single repeat of SHLF |

**Gateway vectors available in the lab (Courtsey- Professor Magdalena Bezanilla)**

| DONR VECTORS | ENTRY VECTORS | DESTINATION VECTORS |
|---|---|---|
| pDONR P1-P5r | pENTR L1R5 mEGFP | pTK-UBI-Gate |
| pDONR P5-P2 | pENTR L1R5 mRuby | pTH-UBI-Gate |
| | pENTR L5L2 mEGFP | |
| | pENTR L5L2 mRuby | |

Table 2 : Table summarising the required constructs and their purpose(mentioned in detail in section 1.2)

## 1.3.1 Using BP clonase to generate ENTRY vectors

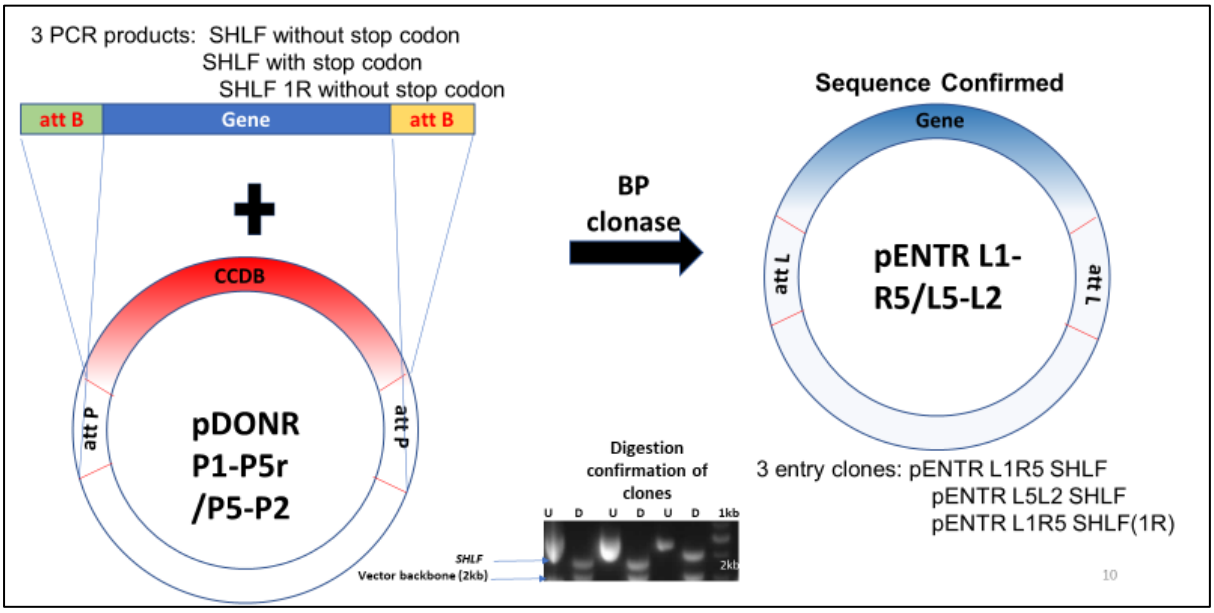


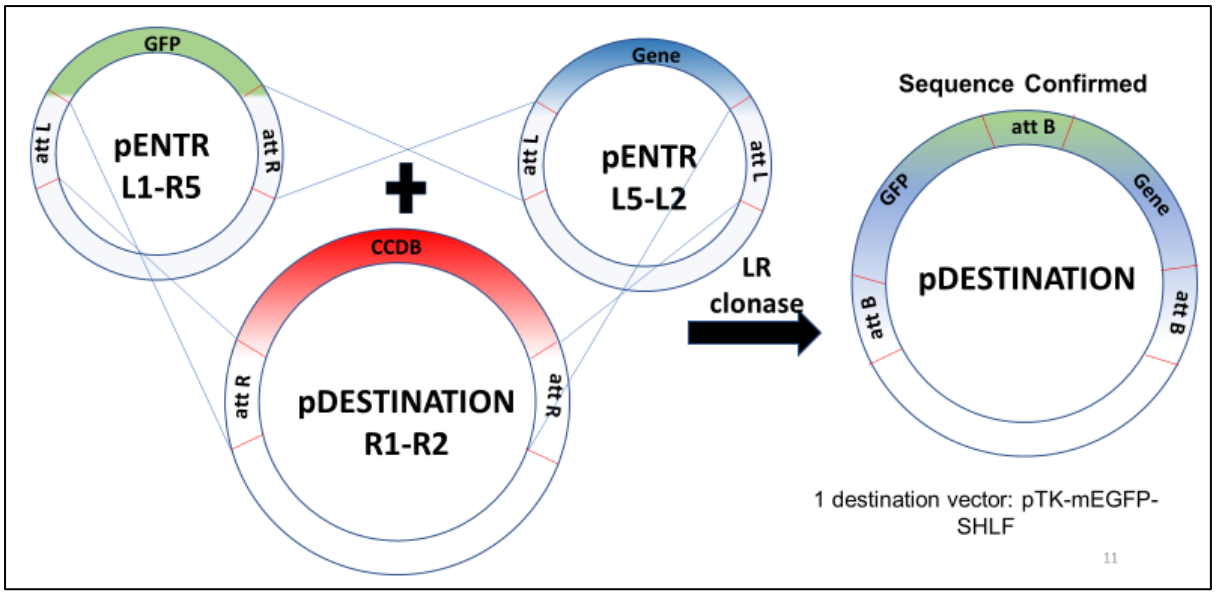

Figure 7: Pictorial representation of the BP reaction

The purified PCR product was then proceeded to be integrated into the DONR vector in order to generate an ENTRY vector. The following reaction gives an overview of the conditions involved in the BP reaction

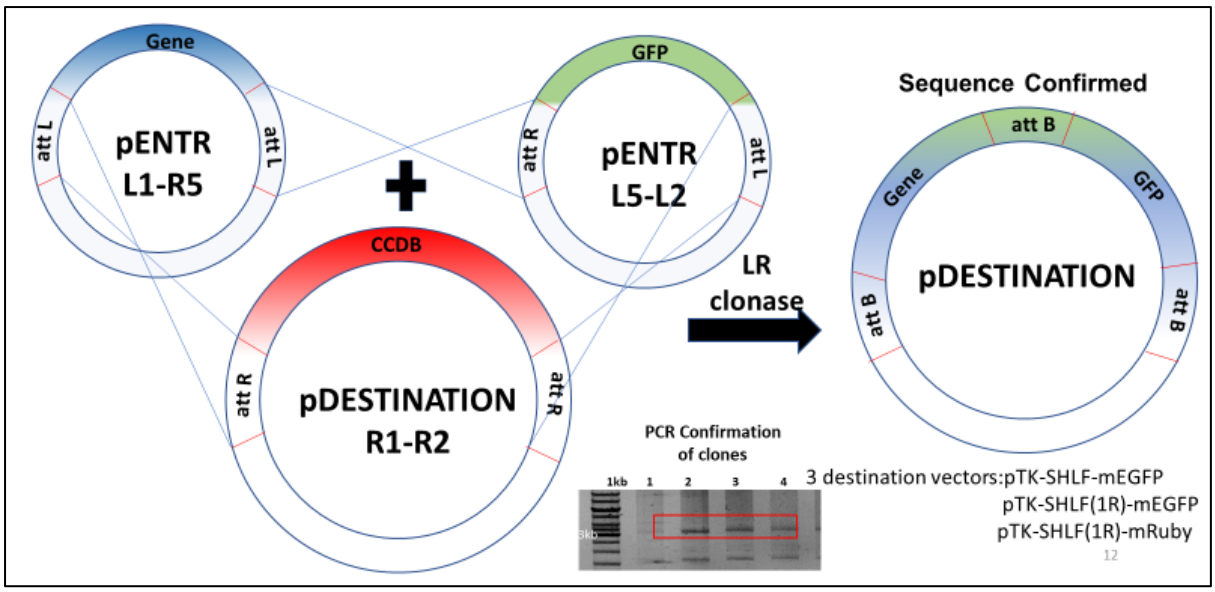
| Components | Amount |
|---|---|
| att B PCR product | 15-150ng |
| pDONR P1P5r/P5P2 | 150-250ng |
| BP Clonase II enzyme mix | 0.5-2 µl |
| TE Buffer pH-8 | Make up the volume |

The following reaction was incubated at 25°C for nearly 18 hours. The reaction was then inactivated by the addition of Proteinase K and incubated at 37°C for 10 minutes. This reaction was then transformed into *E. coli* competent cells. The colonies obtained on a kanamycin selection were screened via colony PCR using M13 universal primers to select for potential ENTRY vectors.

1.3.2Generation of Destination clones using LR clonase and varying combinations of ENTRY vectors




a




b

Figure 8: Pictorial representation of the LR reaction a) N-terminal fusion b) C-terminal fusion

Once and ENTRY vector is generated, the LR reaction along-with the fluorescent protein containing ENTRY vectors and destination vector could be carried out. The following reaction gives an overview of the reaction conditions involved in the LR reaction.

| Components | Amount |
|---|---|
| Entry vector -pENTR L1R5 UHG/pENTR L5L2 UGH | 15-150ng |
| Entry vector – pENTR L5L2 GFP/ pENTR L1R5 GFP | 15-150ng |
| Destination vector-pTK UBI gate | 150ng |
| LR Clonase enzyme mix | 0.5-2 µl |
| TE buffer p.H.8 | Make up the volume |

This reaction was incubated at 25°C for over 18 hours followed by inactivation by proteinase K (1-2 µl) added to the reaction mixture and incubated at 37°C for nearly 10 minutes. The reaction mixture was then transformed in to competent *E. coli* cells with Carbenicillin selection and the colonies were screened for potential clones.

Once a clone was detected it was sent for sequencing and verified for the accuracy of the nucleotides in order to prevent any mutations in the protein to be generated within the plant system. The plasmid was purified in bulk using the Promega MIDI prep kit and digested using PmeI (NEB)- functional in cutsmart buffer to prepare it for transformation.

| COMPONENTS | AMOUNT |
|---|---|
| 10X Cutsmart buffer | 2 µl |
| PmeI | 0.5-1 µl |
| pTK-UBI gate UGH/UHG | 900ng-1µg |
| MQ water | Make up the volume to 20ul |

Digest at 37°C for 4 hours to overnight. Inactivate the reaction by heating at 65°C for 20 minutes. Purify the reaction using Promega gel purification kit.

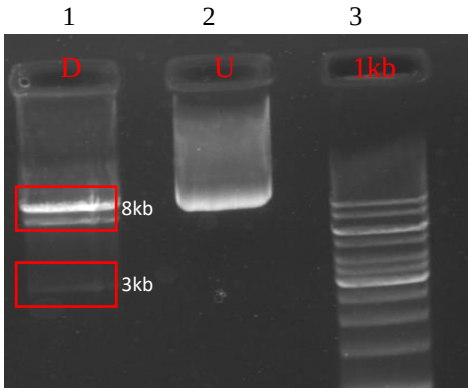



Figure 9: Digestion of pTK UBI gateway vector/UGH/UHG using PmeI

1.4 Transforming *shlf P.patens* protoplast using PEG (PEG mediated transformation) in order to generate and raise stable transgenics.

The process of PEG mediated transformation required the following components (Liu et al, 2011; Frank et al, 2005)

Components of Driselase

Crude powder containing laminarinase, xylanase and cellulase. The enzymes are derived from fungus and are used to digest the cells wall carbohydrates exposing the cytoplasm of the cells allowing us to create protoplasts.

(Components were Filter sterilized using 0.45 microns filter)

PEG-CMS

2g of PEG (Poly ethylene Glycol)

4.5mL 8% Mannitol

500ul $Ca(NO_3)_2$

50ul Tris pH 8.0

(Components were Filter sterilized using 0.2 microns filter)

MaMg

0.405g Mannitol

4.42mL Distilled Water

75ul $MgCl_2$

500ul MES

(Components were Filter sterilized using 0.2 microns filter)

| COMPONENT | PRML(20mL) | PRMB(60mL) | PRMT(30mL) |
|---|---|---|---|
| B | 200ul | 600ul | 300ul |
| C | 200ul | 600ul | 300ul |
| D | 200ul | 600ul | 300ul |
| AT | 200ul | 600ul | 300ul |
| TES | 20ul | 60ul | 30ul |
| Agar | - | 0.336g (0.56%) | 0.12g(0.4%) |
| Mannitol | 1.2g | 3.6g | 2.4g |
| $CaCl_2$ | 200ul | 600ul | 300ul |

1. Over 6-day old protonema of Line 5 mutant moss was added to driselase and kept on the rocker for over an hour to allow the tissue to generate protoplast
2. Using the cell strainer, the protoplast was filtered into a 10ml sterile flask
3. The protoplast was spun down at 800rpm at 21 degrees for 4 minutes at an acceleration of 2 and no brake

4. The first wash step involved the addition f 10ml of 8% Mannitol. This was again spun down at the same conditions.
5. A second wash using 8% mannitol was carried out and the protoplast was spun down again
6. To the spun down protoplast about 1mL of MaMg was added
7. To a fresh sterile 10mL tube 10-15ug of DNA (PmeI digested pTK UBI gate UGH/UHG) was added, to this 300ul of protoplast-MaMg mix was added and finally 300ul of PEG-CMS was added.
8. A heat shock of around 45 degrees was given for around 5 minutes.
9. After exactly 5 minutes the protoplast within the tube were kept in a flask containing Distilled water at room temperature for around 15 minutes.
10. The mix was then diluted using 8%mannitol with successive additions of 300ul, followed by 600ul and finally 1mL of 8% Mannitol at an interval of 3 minutes.
11. The mix was spun down again.
12. And to this PRML was added (2ml)
13. This was kept under dark conditions overnight.
14. The next day PRMB plates were prepared and the protoplast were kept in the incubator exposed to light for nearly an hour. These were then spun down to retrieve the protoplast.
15. The protoplast were resuspended in PRMT at 45 degrees and the mix was poured over the PRMB regeneration medium.
16. After nearly 7 days the protoplast were placed on the selection media containing G418 which allowed the selective growth of protoplast which had incorporated the vector.
17. A single selection is not sufficient as even transient transformants can grow, thus a second even of selection after putting the protoplasts in relaxation will be required.

    The passing of the colonies through two successive selection pressures ensures that false positive do not survive. Once colonies pass the second selection, can

we say with decent confidence that the lines are true. Further molecular tests need to be conducted in order to confirm the positivity of the survivors.

## 2. Functional validation of the first repeat of moss SHLF in the mutant *shlf* background

### 2.1 Culturing of WT *Physcomitrella patens* (strain Gransden) on BCDAT media.

*Physcomitrella patens* wild type (Ecotype- Gransden) and *shlf* mutant were subcultured regularly and maintained according to the procedure mentioned in the section 1.1. The protonema of *shlf* was maintained in order to carry out transformation events to generate transgenics via PEG mediated transformation.

### 2.2 Isolation of the RNA from the WT background.

Rationale- The RNA was isolated so as to prepare cDNA from it to get the desired gene of interest.

Procedure-

1. WT moss gametophores were crushed in a mortar pestle using liquid Nitrogen.
2. 500ul of Trizol was added to the tubes containing the powder of crushed tissue.
3. This mix was vortexed
4. It was incubated at room temperature for 5-10 minutes
5. 200ul of chloroform was added, the mixture was inverted mixed 10-15 times and incubated at room temperature for 10 minutes.
6. This was spun down for 15 minutes at 4 degrees at a speed of 12000rcf

7. The supernatant was transferred to a fresh tube, to this a 500ul of Isopropanol was added, it was invert mixed and incubated at room temperature for 10-15 minutes. The mix was spun down
8. 1mL of 75% ethanol was added and this mix was spun down at 7500rcf,4 degrees for 5 minutes.
9. The pellet so obtained was air dried was 45 minutes
10. To this 30ul of DEPC water was added
11. This mix was boiled at 55 degrees for 10 minutes

2.3 Preparing cDNA from the RNA so isolated

Rationale- The cDNA is prepared to be used as a template during the amplification of the gene of interest during PCR

Primer annealing mix

| COMPONENT | AMOUNT |
|---|---|
| 50uM Oligo dT mix | 1ul |
| 10mM dNTP | 1ul |
| RNA | 2ul |
| DEPC water | 9.5ul |

This annealing mix was spun for a short duration of time and set up in the PCR machine for 5 minutes at 65 degrees

The following mixture was prepared and added to the annealing mix

| COMPONENT | AMOUNT |
|---|---|
| SS4 Buffer | 4ul |
| 100mM DTT | 1ul |

| | |
|---|---|
| RNAse Inhibitor | 1ul |
| SS4-RT | 0.5ul |

Reaction conditions

| TEMPERATURE | DURATION |
|---|---|
| 50 Deg | 50’ |
| 80 Deg | 5’ |
| 4 Deg | Forever |

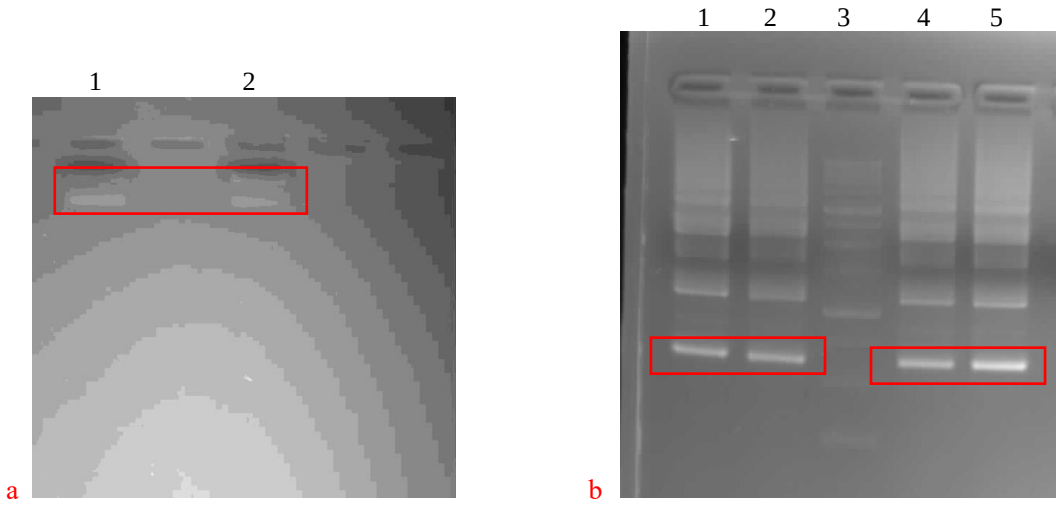



Figure 10: a) RNA isolated from the WT *P. patens* b) Gene amplicon (573bp)

2.4 Amplification of a single repeat of *SHLF* from the cDNA prepared.

For amplification the mix contained-

Buffer for Ex Taq polymerase

Ex Taq Polymerase

Forward primer (Gene specific)

Reverse primer (Gene specific)

dNTP’s

Milli Q water

Conditions for PCR amplification

| TEMPERATURE | TIME |
|---|---|
| 98 Deg | 40'' |
| 98 Deg | 30'' |
| 55/60 Deg | 45'' |
| 72 Deg | 2'10'' |
| 72 Deg | 5' |
| 4 Deg | Forever |

X34

2.5 Cloning the amplicon in the pTFH15.3 vector via the standard restriction digestion and ligation protocol.

Rationale- The restriction sites of AscI and ApaI are present in the forward and the reverse primer respectively hence the amplified gene also contains those sites. The digestion of the vector with the same set of enzymes will allow us to directionally clone the repeat into pTFH15.3.

The digestion reaction was carried out 37°C water bath from time ranging between 4 hours to overnight. The digested product was inactivated at 65°C for 20 mins and purified using Promega Gel purification Kit.

| COMPONENT | AMOUNT |
|---|---|
| 10X Cutsmart Buffer | 5ul |
| AscI | 4ul |
| ApaI | 4ul |
| PCR Product | 40ul |
| MQ | 47ul |

2.6 Transformation of *shlf P. patens* protoplast to raise stable transgenics.

PEG mediated transformation of *shlf* protonema was carried out using 10-12ug od digested DNA. The procedure is similar to that described in detail in section 1.4. The colonies which passed second selection were true transgenics.

## 3. Overexpression of a high expressing SHLF paralog from *Marchantia polymorpha* in *shlf* mutant background

Listed below are some of the main targets which were to be met so as to successfully execute the plan:

3.1 Culturing of healthy thalli of Tak 1(Male) and Tak 2(Female) *Marchantia polymorpha*

*M. polymorpha* was incubated at 16:8 ratio of light:dark cycle in the incubator which is maintained at 24°C for all experiments. The tissue was grown on Gamborgs B5 media in 10cm petri dishes.(Ishizaki et al, 2016)

<1/2 Gamborg's B5>

| 10X Gamborg's B5 stock | 50mL |
|---|---|
| MES | 0.5g |
| Add distilled H2O up to | 1000mL |

For the media preparation, after adding the components listed above the pH was adjusted to 5.5 using 1M KOH. 0.8% agar was added and the media was sterilised in an autoclave.

10X Gamborg's B5 stock

| COMPONENTS | AMOUNT |
|---|---|
| NaH2PO4.2H2O | 1.75g |
| KNO3 | 25g |
| (NH4)2SO4 | 1.34g |
| MgSO4.7H2O | 2.5g |
| CaCl2.2H2O | 1.5g |
| EDTA.NaFe(III) | 400mg |
| Gamborg's B5 micro-elements | 10mL |
| 0.075% KI solution | 10mL |
| Add distilled H2O up to | 1000mL |

Gamborg's B5 Microelements (1000X)

| COMPONENTS | AMOUNT |
|---|---|
| Na2MoO4.2H2O | 25mg |
| 25mg/mL CuSO4.5H2O sol | 0.1mL (2.5mg CuSO4.5H2O) |
| 25mg/mL CoCl2.6H2O sol | 0.1mL (2.5mg of CoCl2.6H2O) |
| ZnSO4.7H2O | 200mg |
| MnSO4.H2O | 1g |
| H3BO3 | 300mg |
| Add distilled water up to | 100mL |

3.2 RNA Sequencing Analysis

Rationale- The rationale behind performing an RNA sequencing analysis was to figure out which paralog of the Hypothetical protein in *Marchantia* expressed the most.

The process of RNA sequencing analysis includes the following steps(Kukurba et al, 2015; Koch et al, 2018)

1. Bowtie- Is an ultrafast short DNA sequence aligner. This software aligns sequences around 35bp reads to the entire genome. Bowtie takes small sets of individual reads as input and gives an alignment as output.
2. Tophat- Is a splice junction mapper. It maps the junctions of Exons-Introns without a reference annotation. This programme takes the input reads, splits them into smaller fragments and individually maps them to the genome. The output result is stitched back together to produce end-to-end read alignments.
3. Cufflinks- This programme assembles transcriptomes from RNA seq data and quantifies their expression.
4. Cuffmerge- When multiple RNA seq libraries are available and the transcriptome of each of these have been assembled. Cuffmerge is brought into picture to generate a master transcriptome.
5. Cuffdiff- This programme understands and predicts the differential expression in the genes. It compares different conditions and understands which genes are up regulated and which are downregulated. It can also figure out which genes are differentially spliced or are undergoing isoform level regulation.

A thorough analysis of the transcriptome of *Marchantia polymorpha* resulted in the detection of five paralogs of the gene.

PARALOGS OF HP IN *Marchantia polymorpha*

1. Mapoly0112s0050 (amplicon size 3.2kb)
2. Mapoly0112s0046 (amplicon size 2.3kb)
3. Mapoly0193s0017 (amplicon size 1.5kb)
4. Mapoly0120s0049
5. Mapoly0318s0001

These paralogs are arranged in the order of the decreasing order of their relative expression wherein the first paralog is expressed the most (Mohanasundaram et al, 2019, under revision).

Primers were designed for the top three paralogs so as to amplify them from the cDNA.

*SHLF* in moss contains four tandem direct repeats which are highly similar, whereas the *SHLF* paralog in *Marchantia polymorpha* contains five repeats. Given below is a diagrammatic overview of the process of RNA Seq analysis.

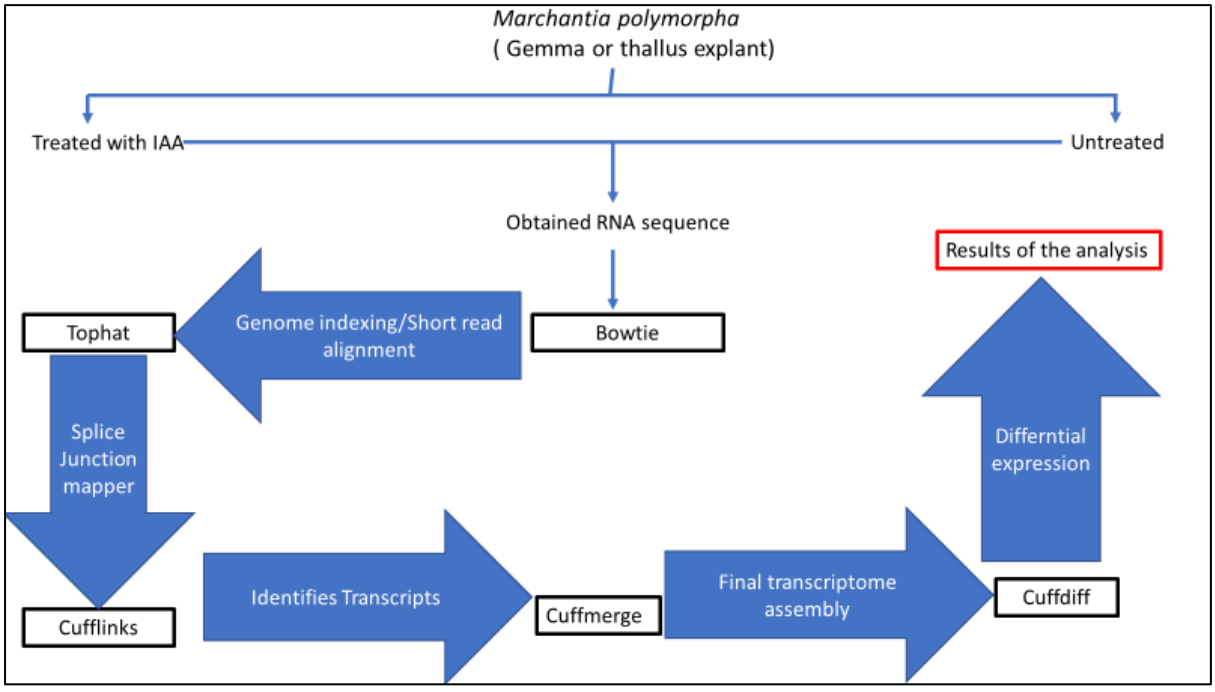



Figure 11: An overview of the RNA Seq analysis procedure

3.3 Isolating the RNA from the thallus of *Marchantia polymorpha* (Tak1)

A 23day old thalli of Tak1 was selected for the purpose of RNA isolation. Refer to section 2.2 for details about the protocol.

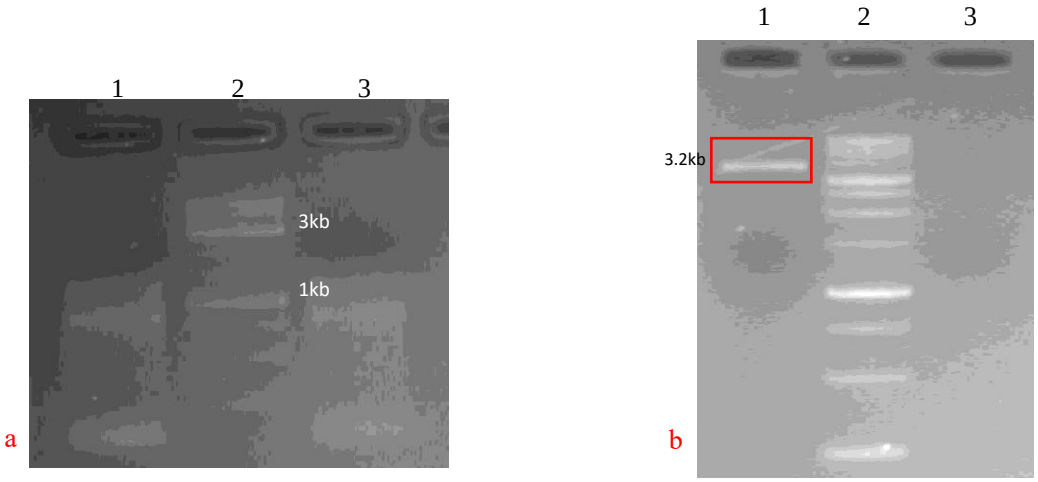



Figure 12: a) RNA isolated from the Tak1 thalli b) Gene amplicon (3.2kb)

3.4 Preparing cDNA from the isolated RNA

Rationale- The cDNA is prepared to be used as a template during the amplification of the gene of interest during PCR

Primer annealing mix

| COMPONENT | AMOUNT |
|---|---|
| 50uM Oligo dT mix | 1ul |
| 10mM dNTP | 1ul |
| RNA | 2ul |
| DEPC water | 9.5ul |

This annealing mix was spun for a short duration of time and set up in the PCR machine for 5 minutes at 65 degrees

The following mixture was prepared and added to the annealing mix

| COMPONENT | AMOUNT |
|---|---|
| SS4 Buffer | 4ul |
| 100mM DTT | 1ul |
| RNAse Inhibitor | 1ul |
| SS4-RT | 0.5ul |

Reaction conditions

| TEMPERATURE | DURATION |
|---|---|
| 50 Deg | 50’ |
| 80 Deg | 5’ |
| 4 Deg | Forever |

3.5 Amplification of MpHP50 from the cDNA

The following primers were used for amplification of the gene of interest

| Primer Name | Primer sequence |
|---|---|
| Forward | *AAAAA*GGCGCGCCATGGCAAGCTCAGTAGCTTCTGG |
| Reverse | *AAAAA*GGGCCCTCAGTACTTGGCAATAATCTTCTG |

Table 3: Primers for the amplification of MpHP50 from cDNA of Tak1

For amplification the mix contained-

Buffer for Ex Taq polymerase

Ex Taq Polymerase

Forward primer

Reverse primer

dNTP’s

Milli Q water

The temperature conditions were as follows

| TEMPERATURE | TIME |
|---|---|
| 98 Deg | 40'' |
| 98 Deg | 30'' |
| 55/60 Deg | 45'' |
| 72 Deg | 3' 20'' |
| 72 Deg | 5' |
| 4 Deg | Forever |

X34

3.6 Restriction digestion of the PCR product using ApaI and AscI

Rationale- The restriction sites of AscI and ApaI were present in the forward and the reverse primer respectively hence the amplified gene also contains those sites. The digestion of the vector with the same set of enzymes would allow us to directionally clone MpHP50 into pTFH15.

Given below are the reaction conditions for a bulk digestion of the PCR product

| COMPONENT | AMOUNT |
|---|---|
| 10X Cutsmart Buffer | 5ul |
| AscI | 4ul |
| ApaI | 4ul |
| PCR Product | 40ul |
| MQ | 47ul |

This PCR product which has been digested in bulk was set up for ligation with the pTFH15.3 digested with the same enzymes.

3.7 Ligation mix was transformed into competet *E. coli* cells

The ligation mix transformed gave rise to colonies on a selection media. The potential colonies were screened using M13 universal primers and those that seemed positive were sent for sequencing

3.8 Digestion of the positive clone using NotI

The positive clone was digested using Not1 in order to linearize it for the purpose of transformation

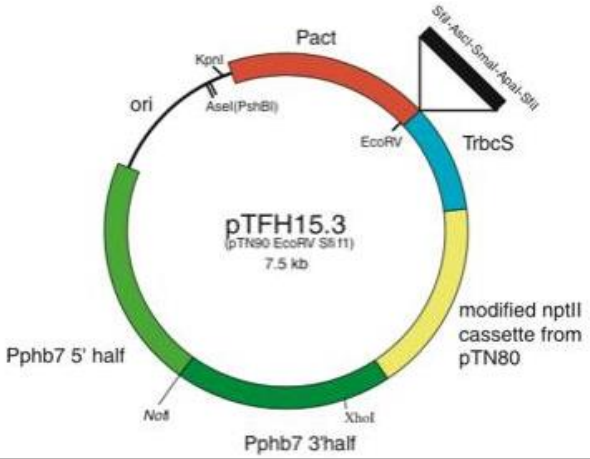



Figure 13: Vector map of pTFH15.3 driven by a Rice Actin Promoter

3.9 Transformation of the protonema of *shlf* mutant moss via PEG mediated transformation

Around a 5-6day old *shlf* protonema was used for the purpose of transformation. For details regarding the procedure refer to section 1.4. Only once the colonies passed second selection could potentially positive clones be selected for.

3.10 Morphological studies of the lines generated post second selection

Morphological studies include measurement of the width of the phyllid, length of the phyllid, internodal distance, PDAC count.

The ninth leaf from the top of the gametophore was selected for measurements since it is after this leaf that the variations in the shape and size are less.

PDAC count was measured by staining the leaves using aniline blue. The leaf was immersed in a solution of aniline blue and placed in a vacuum

chamber. The aniline blue stain images were taken using a confocal microscope and the spots on the images were counted. (Zavaleiv et al, 2015)

## 4. Understanding the role of *SCARECROW* in *Physcomitrella patens* for leaf development.

### 4.1 Promoter analysis of *SCR2* and *SCR3*

The gene id for both *SCR2* and *SCR3* are listed below

1. Pp3c22_13060V3 (*SCR2*)
2. Pp3c19_18560 (*SCR3*)

Promoter regions for both these gene ID's were obtained on "Phytozome". Nearly 5kb region upstream of the gene was selected.

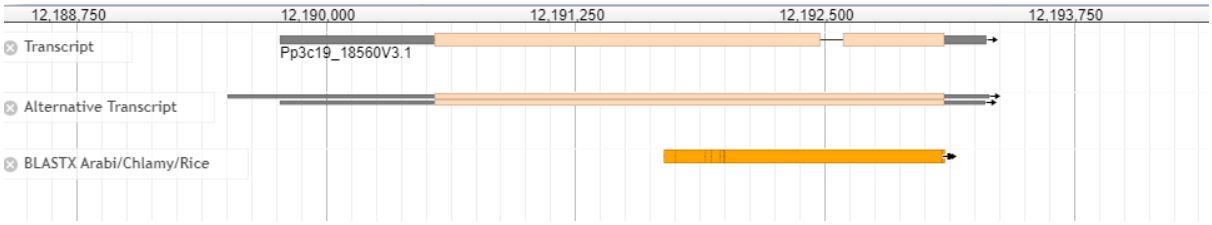



Figure 14: Genome sequence of *SCR3* located on Phytozome

An expression analysis study was then carried out on Phytozome in order to understand what conditions or a set of chemical substances could affect the levels of the transcripts of the genes in question(Nanjareddy et al,2017). The figure below gives an overview of what all conditions could potentially affect the rate of transcription of these genes. For eg. In case of *SCR2* red light, strigolactone, water stress and heat stress act to inhibit the gene. Whereas strigolactone and red light upregulate the levels of *SCR3*.

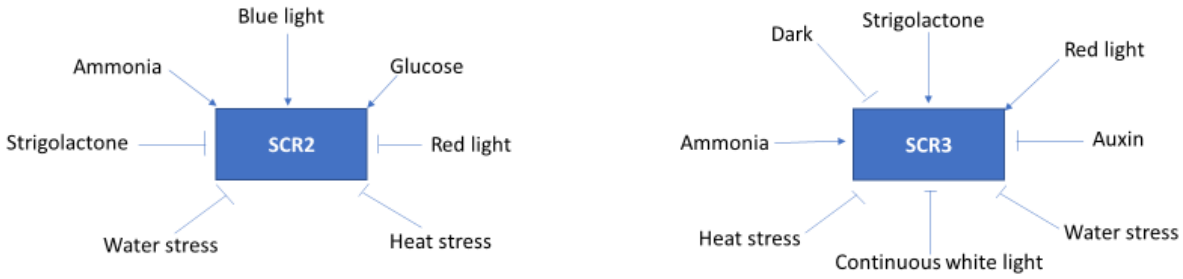



Figure 15: Expression analysis of SCR2 and SCR3 gene (PLACE)

An expression analysis was followed by a promoter analysis on a software called PLACE. The 5kb upstream region was fed into the system and all the potential target binding sites for multiple transcription factors were listed. The image below highlights the various sites for the binding sites.

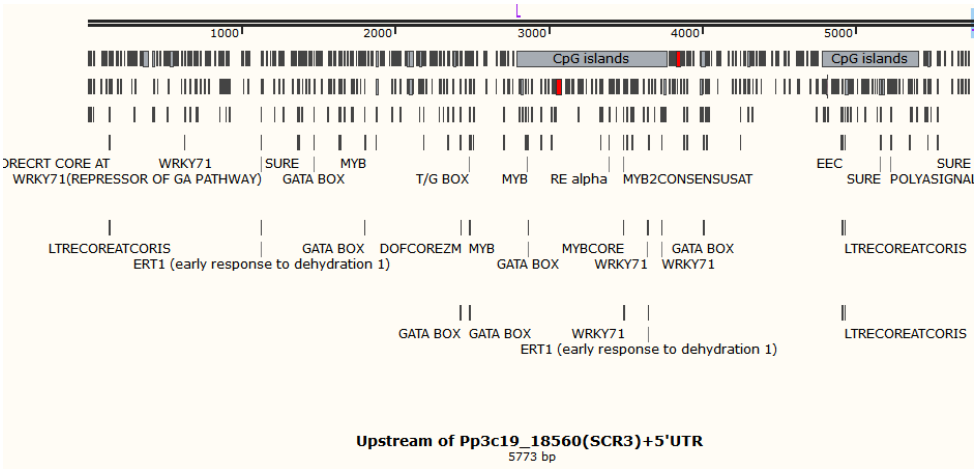



Figure 16: Annotating the potential transcription factor binding target sites in the promoter region of *SCR3*

Factors such as GATA box, TATA box, PolyA signal were detected which are highly essential for the binding of RNA polymerase. ERT1 is involved in water stress. Once the entire 5kb region upstream was annotated the CpG islands were included in the promoter region and a 3kb region was selected as a promoter for *SCR3* whereas a 1.5kb region was selected as the promoter for *SCR2*.

4.2 Designing att site containing primers

For both SCR2 and SCR3 promoter region as well as for tdTomato (fluorescent protein).

The following primers were designed along-with att B sites in order to proceed with a Gateway Cloning strategy.

| Primer name | Sequence |
|---|---|
| attB1SCR3prom_F | 5' GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGTAAGTGTAAGTGTGCAGAGTTTCTCAC 3' |
| attB5rSCR3prom_R | 5' GGGGACAACTTTTGTATACAAAGTTGTACTCACCGTATTGCCTGCTCTCC3' |
| attB1SCR2prom_F | 5'GGGGACAAGTTTGTACAAAAAAGCAGGCTTACGTCAAACTATCCTTCTTCTCTACTGATTG 3' |
| attB5rSCR2prom_R | 5' GGGGACAACTTTTGTATACAAAGTTGTTGCTTGGGGGCTGCTGGAAAAG 3' |

Table 4: Primer sequences for the amplification of *SCR2* promoter and *SCR3* promoter from g DNA of WT *P. patens*

- Isolation of gDNA from WT *P. patens*.

  Genomic DNA was isolated from the WT moss. (Doyle, 1990)

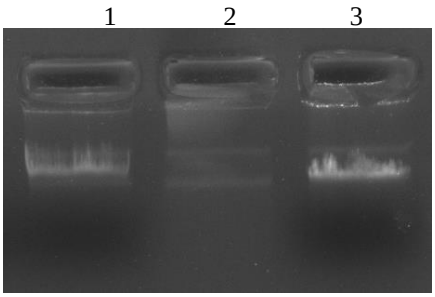



Figure 17: gDNA isolated from WT *P. patens* background (lane 1- isolated gDNA, lane 3- referenceλ DNA)

Composition of CTAB buffer for the purpose of gDNA isolation

1.2% CTAB [Sigma H-5882]

2. 1.4 M NaCI

3. 0.2% 2-mercaptoethanol

4. 20 mM EDTA,

5. 100 mM Tris-HCI, pH 8.0

Given below is an overview of the protocol for gDNA isolation

1. Preheat the CTAB buffer in a water bath for around 30 minutes at 60°C.
2. Crush the WT gametophore tissue in liquid nitrogen using mortar pestle.
3. Treat around 0.1g of the crushed sample in a 1.5mL centrifuge tube with the CTAB buffer in a water bath set at 60°C for around 45 minutes.
4. Spin down the tube for 10 minutes at 14,000 rpm.
5. Collect the supernatant and extract once with Chloroform IAA (24:1). Mix gently and spin down at room temperature. Collect the supernatant in a fresh tube.
6. Add isopropanol in equal volumes, invert mix and leave the tube undisturbed at -20°C for 45 min-1hour.
7. Spin down the tube at room temperature.
8. Remove the supernatant and look for the pellet.
9. Add 70%Ethanol and wash the pellet.
10. Spin the tube and remove the supernatant, dry the pellet.
11. Resuspend the pellet in Milli Q pre heated at 65°C. Load the gDNA on the gel to check for quality and concentration.

4.4 Amplification of *SCR3* promoter and *SCR2* promoter from the gDNA

A two-step touch up PCR technique was employed to amplify the promoter region from the gDNA. Advantage polymerase was used for this purpose.

The reaction mixture contained.

gDNA template – 150ng

Forward primer- 5µM

Reverse primer- 5µM

Advantage polymerase

dNTP- 10mM

1. For the amplification of SCR3 promoter region from the gDNA of WT *P. patens.* The following steps were taken as precautionary measure and troubleshooting:
   1.1 The gDNA quality was tested with each isolation by amplifying some known region of the genome from it.
   1.2 The amplification of a selected 3kb region upstream of the SCR3 gene proved to be very vague with positive results in some attempts and negative results for the same set of conditions in other attempts. Thus, the amplification of the 3kb region could not be replicated despite several attempts. Varying conditions of amplification were employed including trying out various enzymes, different types of PCR strategies (two step touch-up, touchdown PCR, Gradient PCR)[10].
   1.3 Finally, the size of the amplicon was reduced to 2.3kb and 1.9 kb. Amplification of the 1.9kb region was successful.

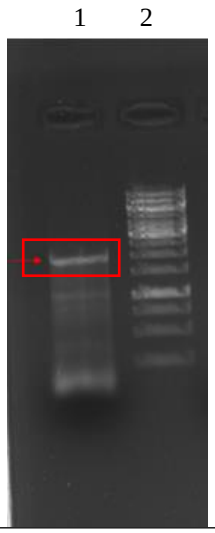



Figure 18: Amplification of 1.9kb of SCR3 promoter from WT- gDNA of *P. patens*

For *SCR2* promoter the amplification worked in the first attempt.

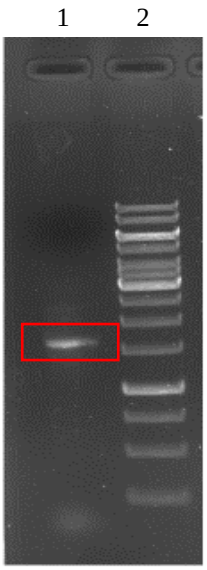



Figure 19: Amplification of SCR2 promoter using att sites containing primers {Amplicon size(1504bp)}

## 4.5 Amplification of tdTomato from a vector template

Dr5V2-d containing sequences for 3XGFP and tdTomato was used as a template for the amplification of tdTomato.

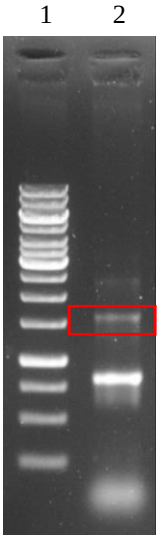



Figure 20: Amplification of tdTomato using att sites containing primers {Amplicon size(1.5kbp)}

## 4.6 Generating ENTRY vectors from DONR vectors via BP reaction

pDONR P1P5r was the DONR vector which was used for setting up a BP reaction with the purified PCR products of the promoter. pDONRP5P2 was the DONR vector used for setting up a BP reaction with the purified PCR product of the tdTomato. For details regarding the BP reaction refer to section 1.3.1.

## 4.7 Modification of pTK Ubi gateway vector by removing the Maize ubiquitin promoter

By removing the maize ubiquitin promoter via restriction digestion followed by blunting of sticky ends, self-ligation and finally screening of potential clones.

1.1. The Maize Ubiquitin promoter was digested out using two enzymes namely Ale1-v2 and Sgr-D1.

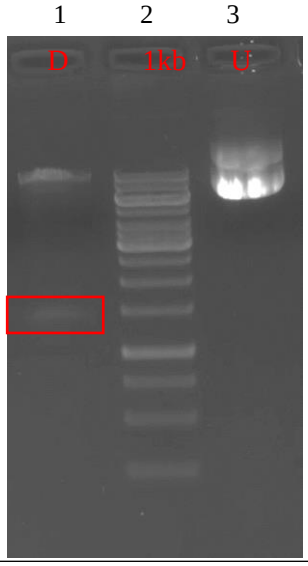



Figure 21: Digestion of the pTK-Ubi-gateway vector resulting in the removal of the Maize ubiquitin promoter (indicated in red)

1.2 The pTK backbone digested out had one sticky end cut and one blunt end cut. The sticky end was blunted using Phusion polymerase enzyme as well as the Klenow fragment. This was followed by a self-ligation and

transformation into the CCDB resistant *E. coli* bacterial cells. The CCDB resistant cells were prepared fresh every time a transformation event was carried out- via both chemical mediated preparatory technique as well as electrocompetent cell preparatory technique.

1.2.1 Nearly 15 different individual attempts (with different conditions) were made to self-ligate the vector but this proved to be unsuccessful. Multiple vectors of varying sized were obtained post transformation into CCDB resistant competent cells, but none of these were of the expected size.

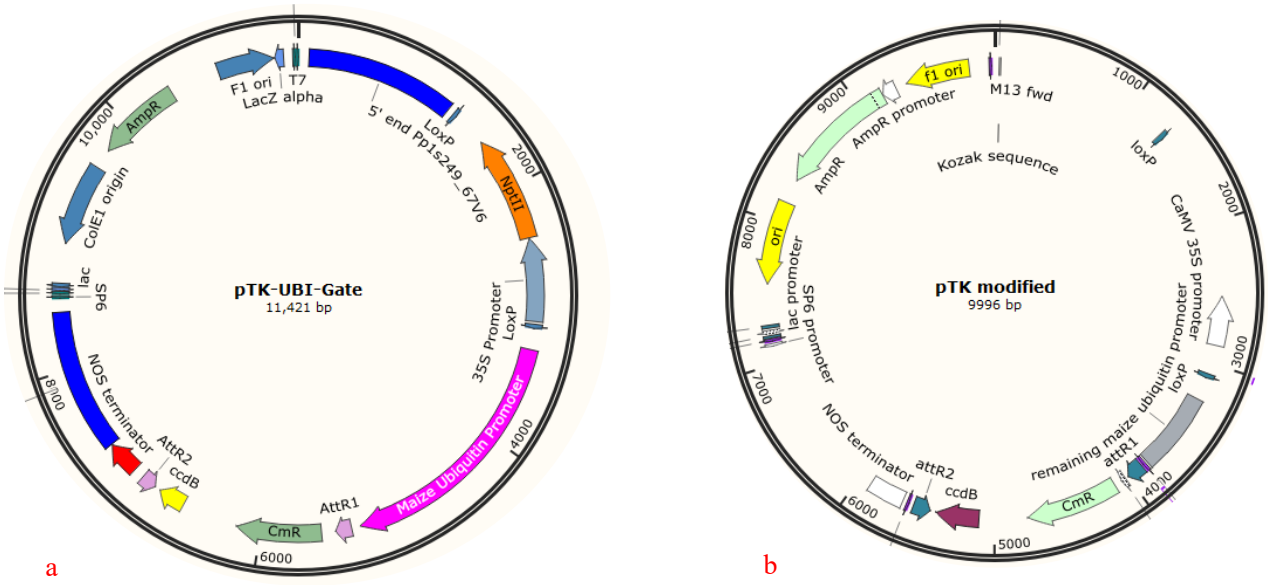



Figure 22: a) pTK UBI gate- original structure containing the maize ubiquitin promoter (shown in pink)

b) pTK modified gate- modified vector after the removal of the maize ubiquitin promoter, the remaining promoter is depicted in grey

4.8 <u>Generation of destination clones via LR reaction followed by PEG mediated transformation of WT *P. patens* protoplast in order to generate stable transgenics. –</u>

The experiment is currently at hold at this step since the destination vector has not been prepared.

# 3. RESULTS AND DISCUSSIONS

## 1. Generation of GFP fusion lines of SHLF to study its sub-cellular localization in the moss, *Physcomitrella patens*

Given below is a table containing the list of all the constructs that were prepared during the cloning experiments. There were a total of 4 ENTRY vector and 4 destination vectors that were generated.

| ENTRY CLONES | DESTINATION CLONES |
|---|---|
| pENTR UHG(1R) L1R5 | pTK UHG(1R) mRUBY |
| pENTR UGH(1R) L5L2 | pTK UHG(1R) mEGFP |
| pENTR UHG(HP) L1R5 | pTK UHG(HP) mEGFP |
| pENTR UGH(HP) L5L2 | pTK UGH(HP) mEGFP |

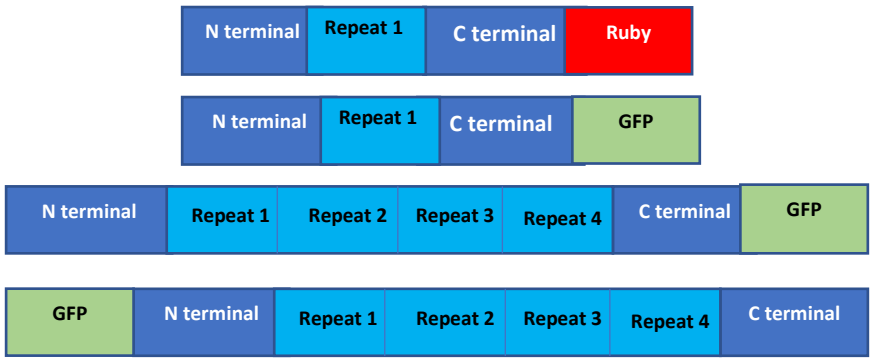



Figure 23: A list of the contructs generated and a cartoon showing the arrangement of the fluorescent protein with respect to the whole gene (entire *SHLF*) or a single repeat

The transformants that were generated and passed the second selection showed a WT like phenotype for both the constructs- N terminal GFP fusion (UGH#89-

control) as well as the C terminal GFP fusion (UHG#133- test). (Mohanasundaram et al, 2019, under revision).

The Control line or N terminal fusion UGH#89 displayed a diffused and random localisation of the GFP which was also seen within the nucleus. This localisation was similar to that seen in the eGFP lines of the *P. patens.* In case of this control line the ER localisation signal fails to act as a peptide signal due to the presence of a GFP upstream. Since the signal is no longer translated towards the N terminal, the GFP -SHLF fusion protein acts like GFP and localises everywhere in the cell including the nucleus.

The test line or C-terminal fusion UHG#133 displayed a very specific GFP localisation which was not detected within the nucleus. The localisation seen within the C terminal fusion line displayed a very typical cortical- network like patterning when observed under the confocal microscope. This signal when compared with the ER-GFP lines showed a striking similarity, indicating that the protein might be getting localised to the ER owing to its localisation signal present within the N terminal. This also indicated that within the test line the ER localisation signal is being translated towards the N terminal allowing it to function as a localisation signal. It could also potentially indicate that the protein is being secreted by the traditional secretory pathway.

## 2. Functional validation of the first repeat of moss SHLF in the mutant *shlf* background

Transformation of the *shlf* mutant protoplasts via PEG resulted in the generation of stable transgenics which passed the second selection. The transgenics displayed a *shlf* phenotype indicating that either a single repeat is

not enough to recover the phenotype or maybe the definition of a functional repeat is not very well understood.

Another reason for the failure in the reversal of the phenotype could be the following-

The PCR reaction carried out to amplify the single repeat was actually specific to the N and C terminal. Each PCR leads to the amplification of the N terminal a single repeat and a C terminal region, followed by the N and C terminal with two, three and four repeats.

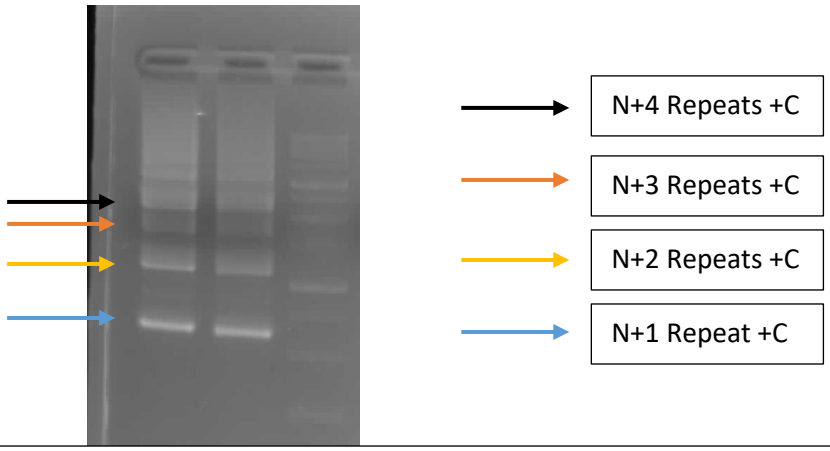



Figure 24: Amplification of *SHLF* from cDNA gives all the repeats with the N and C terminal in each case

The amplification of the single repeat with N and C terminal was not straightforward. On sequencing it was seen that the single repeat was a hybrid between certain segments of the first repeat and the fourth repeat.


Figure 25: Amplification gives rise to a single repeat which is a hybrid between the 1st and the4th repeat

The main issue here is that a single repeat has been defined bioinformatically and not functionally. This could be one of the reasons why the single repeat

failed to rescue the mutant phenotype. More studies would have to be conducted in order to understand the tandem direct repeat containing nature of the protein.

## 3. Overexpression of a high expressing SHLF paralog from *Marchantia polymorpha* in *shlf* mutant background

PEG mediated transformation of *shlf* protoplasts followed by two rounds of selection lead to the generation of stable transgenics.

After a second round of selection there were three surviving transgenic lines labelled – 1A, 6 and 9.(Mohanasundaram et al, 2019, under revision)

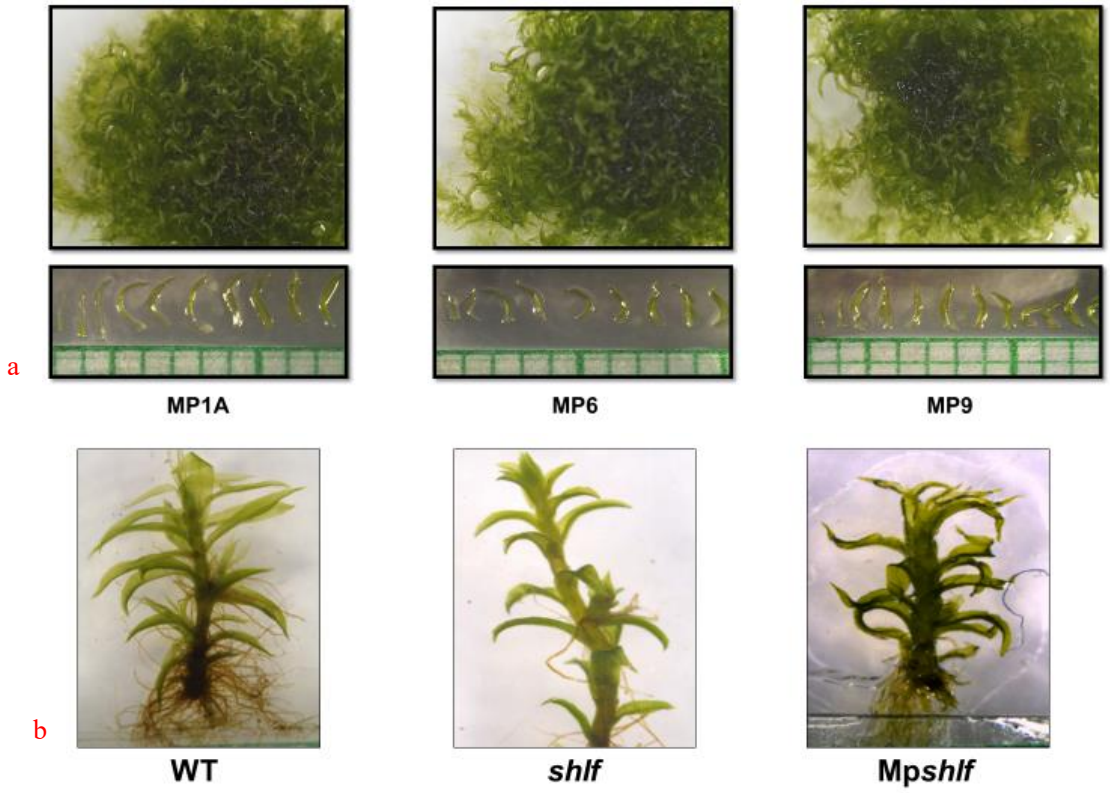



Figure 26: a) Colony images and leaf array of the three surviving transgenic lines b) Comparing gametophores between WT, *shlf* and the transgenic lines *shlf*;Mp_OE

The lines were confirmed via qRT-PCR to detect the levels of the transcripts, line 6 showed the highest levels of the transcript whereas line 9 had the least.

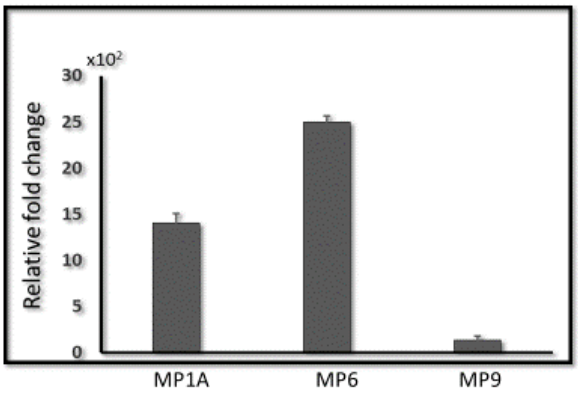



Figure 27: Transcript levels of the *shlf*;Mp_OE in the stable transgenics(Mohanasundaram et al, 2019 , under revision)

The transgenics showed a characteristic twisted phyllid phenotype. In order to understand whether the *Marchantia* ortholog of *SHLF* had managed to rescue the mutant phenotype a large array of phenotypic data was collected allowing us to measure traits like – phyllid length, phyllid width (at the base and the middle), internodal distance and PDAC (plasmodesmata associated callose).

The data for phyllid length showed that there was some partial recovery of the phenotype. The 9$^{th}$ leaf from the top of a gametophore was chosen for all phenotypic analysis, since it was after this phyllid, that the shape and size of the subsequent leaves were more or less stable. The phyllid length of the transgenics were significantly greater than the *shlf* mutant and the WT and line 6 phyllid lengths were comparable. There seems to have been a partial recovery.

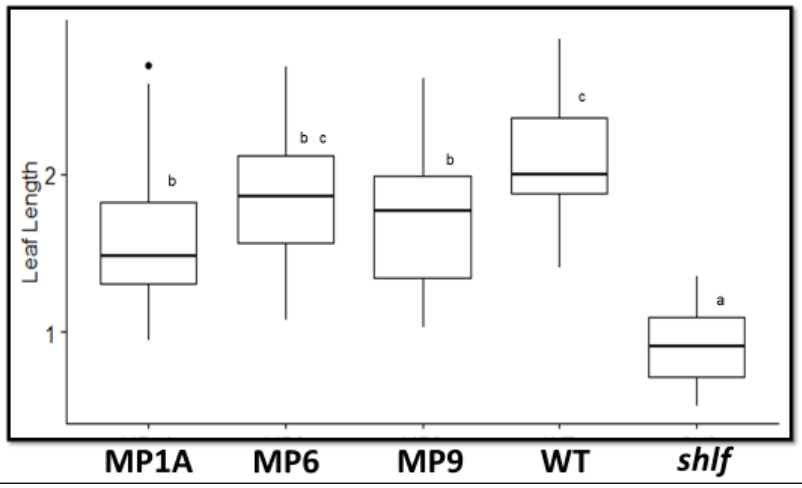
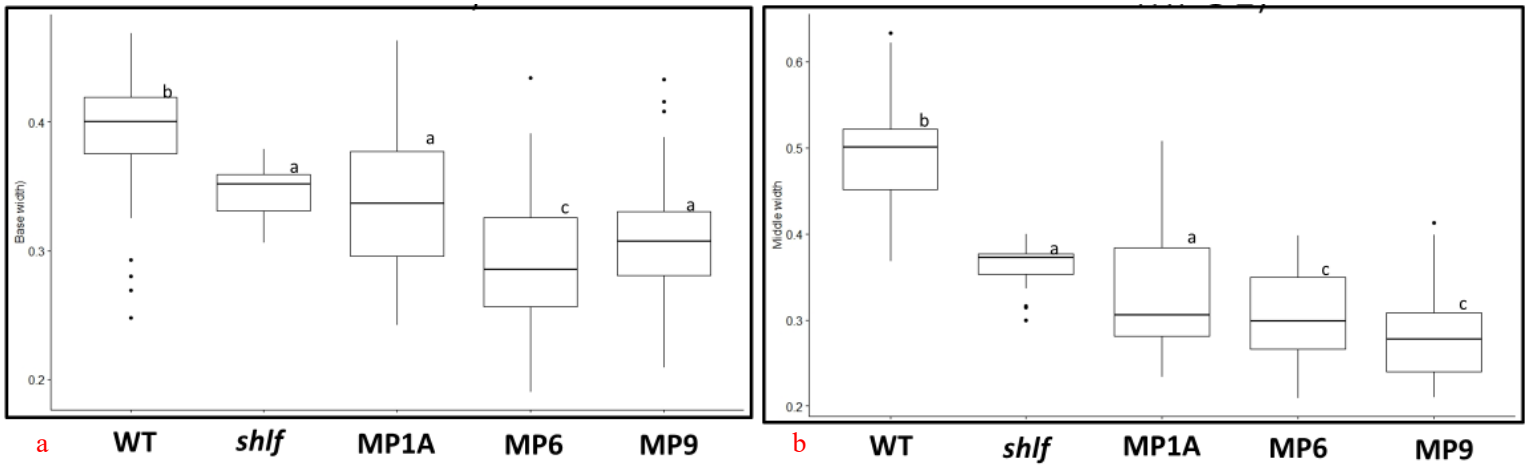



Figure 28: Comparing phyllid length of the *shlf*;Mp_OE ,WT and *shlf* mutant (n=30 individual phyllids)




a b

Figure 29: Comparing phyllid width of the *shlf*;Mp_OE ,WT and *shlf* mutant a)width at the base of the phyllid b) width at the middle of the phyllid (n=30) (Mohanasundaram et al, 2019 , under revision)

Comparing the phyllid width at the base and the middle showed that there was not a significant change in the shape of the transgenic leaf when compared to the *shlf* mutant. The *shlf* phyllid displays a lanceolate shape which has nearly the same thickness at the base and at the middle of the phyllid. The WT has an oblanceolate shape wherein the widest part of the phyllid is at the middle. From the data it seems like there has not been a significant change in the shape of the phyllid and it is more or less lanceolate.

PDAC (plasmodesmata associated callose) is an indirect way of measuring the number of plasmodesmatal connections between the cells of the phyllid. The callose deposits in the plasmodesmata are stained using aniline blue and visualised using a confocal microscope.

The PDAC count for the WT and the *shlf* mutant indicates a decreased amount of the plasmodesmata in the mutant background. The PDAC count was also carried out for the *shlf*;Mp_OE (lines 6 and 9) and the results showed that the PDAC/ length did not show a complete recovery when compared to the WT. Though line 9 did show some partial recovery.[5]

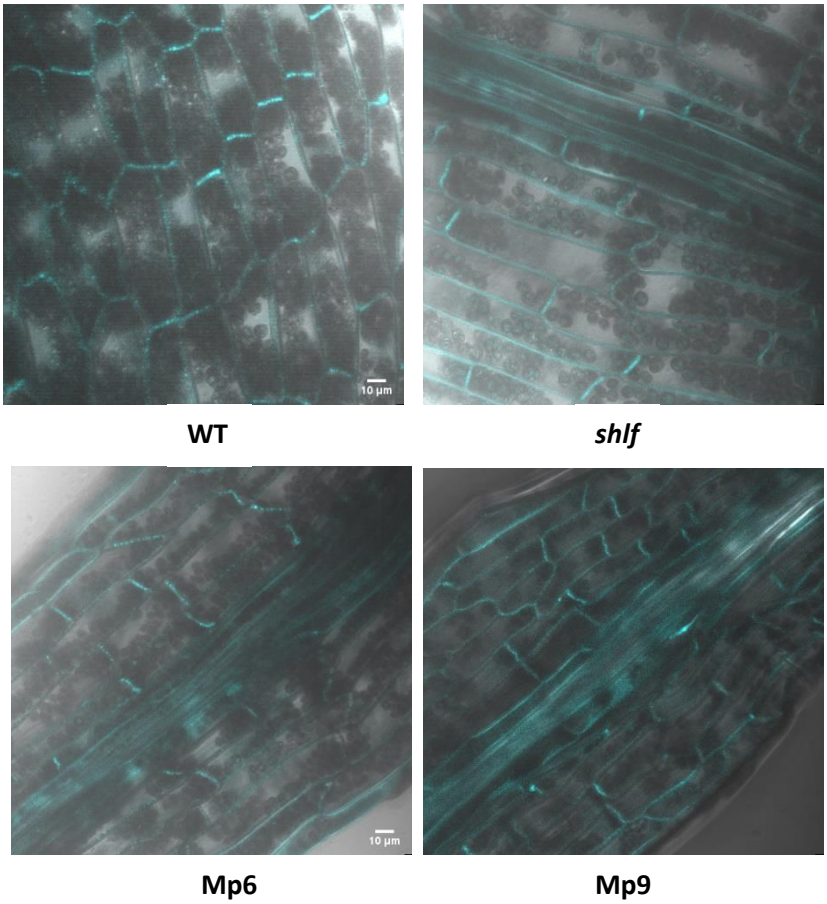



Figure 30: Representative images depicting the aniline blue staining of the *shlf*;Mp_OE ,WT and *shlf* mutant

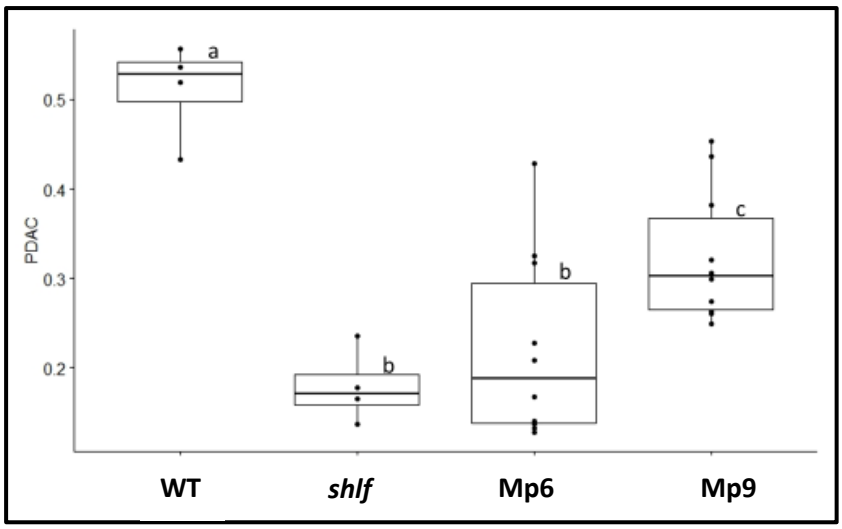



Figure 31: Comparing PDAC/cell wall length of the *shlf*;Mp_OE ,WT and *shlf* mutant (n=30 individual phyllids) (Mohanasundaram et al, 2019 , under revision)

Plasmodesmatal connections between the cell walls of the phyllids in the WT were seen to be much greater than those present in the *shlf* mutant. The phenotype of the mutant might also be attributed to the reduced number of plasmodesmatal connections since this might affect the flow of auxin throughout the plant. (Mohanasundaram et al, 2019 , under revision)

The internodal distance measured and compared between the *shlf* mutant and the WT showed that the it is greatly increased in the mutant. In order to understand whether the *SHLF* ortholog is able to recover the phenotype the internodal distance was measured for two lines (Mp6, Mp9) and compared to WT and the *shlf* mutant.

The data indicated that there was a recovery in the phenotype.

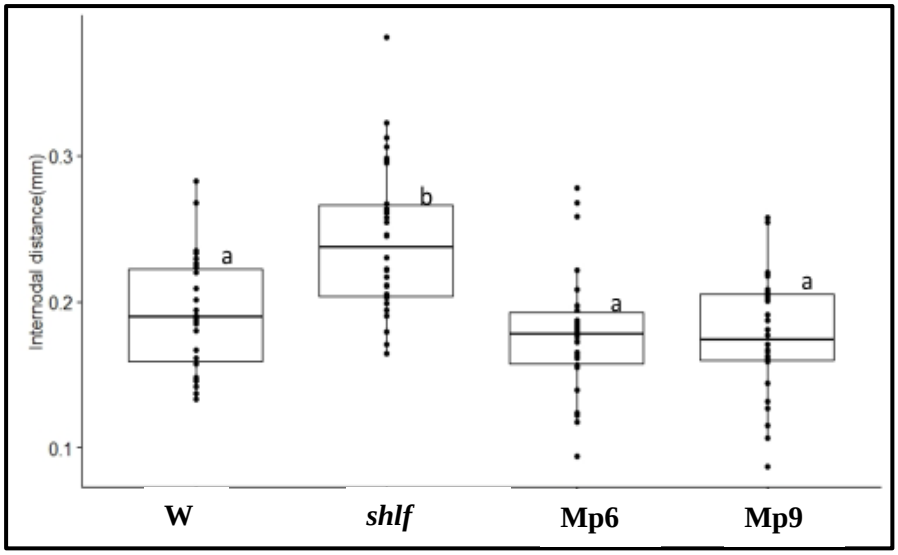


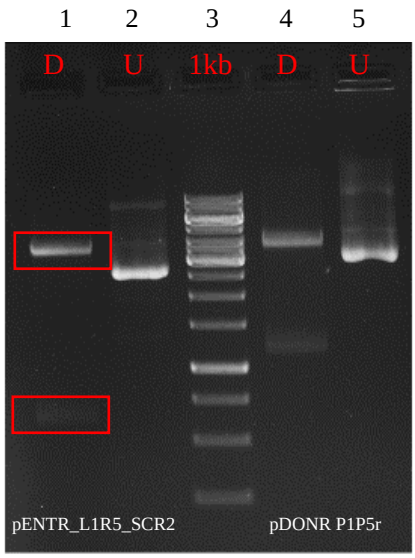

Figure 32: Comparing internodal distance of the *shlf*;Mp_OE ,WT and *shlf* mutant (n=30 individual phyllids) (Mohanasundaram et al, 2019 , under revision)

These results indicate that there has been some partial recovery of the phenotype in the *shlf*;Mp_OE lines. These results are not enough to conclude a complete conservation of the role of the *SHLF* protein across the Bryophyte lineage though.

## 4. Understanding the role of *SCARECROW* in *Physcomitrella patens* for leaf development.

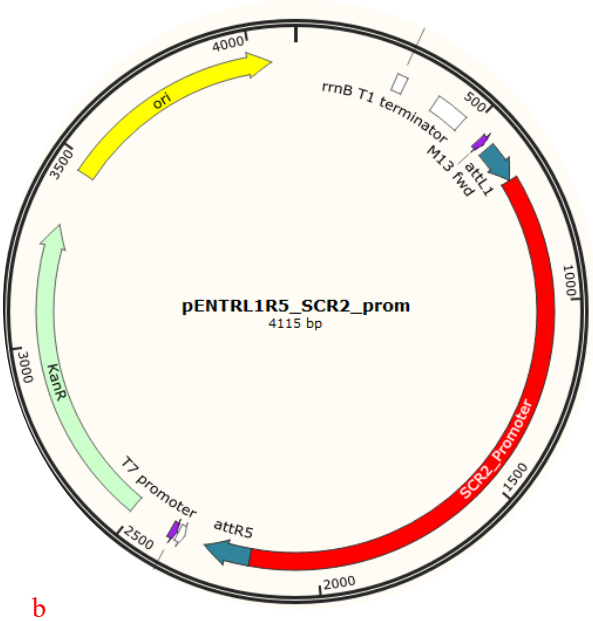
Although this particular project is ongoing, many targets have been met;

1. Successful amplification of SCR2 promoter from gDNA of WT *P. patens* using primers containing att sites. Followed by successful integration of the promoter region into the pDONRP1P5r vector to generate pENTRL1R5 vector. This clone was confirmed by digestion and sequencing.




a

**Expected band-**
**pENTR_L1R5 _SCR2 promoter- 620+3495**
**pDONR P1P5r- 1170+3595**




b

Figure 33: a) Digestion confirmation of pENTR_L1R5_SCR2 promoter b) Vector map of pENTR L1R5 SCR2 promoter

2 Successful amplification of tdTomato (fluorescent protein) from the template vector Dr5V2-D using att site containing primers. Followed by successful integration of this amplicon into the pDONRP5P2 via BP reaction leading to the generation of an ENTRY clone pENTRL5L2_tdTomato. The clone was confirmed via sequencing.

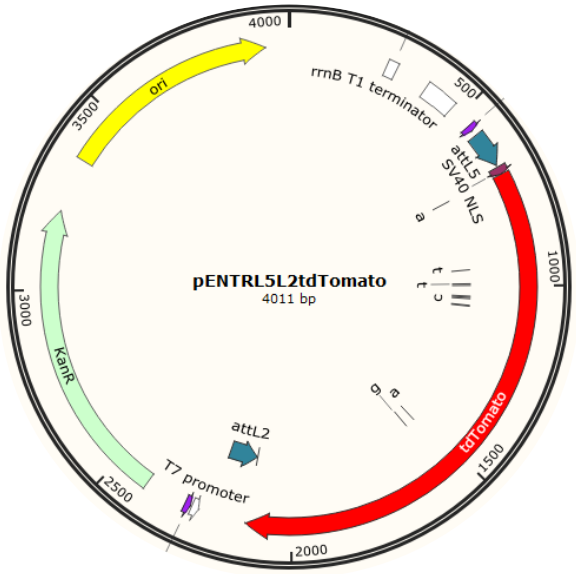



Figure 34: Vector map of pENTRL5L2_tdTomato (region indicated in red)

2. Successful amplification of the 1.9kb region of the promoter of *SCR3* after multiple failed attempts at the amplification of the 3 kb region. The change in strategy allowed for this PCR to work successfully.

The following targets are posing a serious challenge:

1. The BP reaction of the amplicon with the pDONRP1P5r posed a serious challenge (as it did with SCR2 cloning). None of the ENTRY clones generated post BP reaction seemed to be positive. This reaction was carried out multiple times only to yield negative results. The problem seemed to be with the DONR vector which seemed to be recombining with itself once exposed to the BP clonase. It seems that the P1 and P5r regions of the pDONRP1P5r were recombining with each other. On sequencing on potentially positive clone it was seen that the CCDB gene was intact but present in the reverse orientation. What this might have resulted in was the inactivation of the toxin thereby allowing the competent Dh5α cell to grow and thereby form a colony. Thereby, this reaction when transformed into the *E. coli* bacterial cells for the purpose of screening also allowed the self -recombined pDONRP1P5r vector to propagate, hence maybe preventing the integration of the amplicon of the SCR3 promoter into the DONR vector. Multiple steps were taken to prevent the recombination of the pDONRP1P5r- which included, fresh isolation of the vector before each BP reaction, growing the cells at 28°C instead of the standard 37°C. None of these preventative measures yielded any positive results.

2. Modification of pTK Ubi gateway vector by removing the Maize ubiquitin promoter

   2.1 The Maize Ubiquitin promoter was digested out using two enzymes namely Ale1-v2 and Sgr-D1.

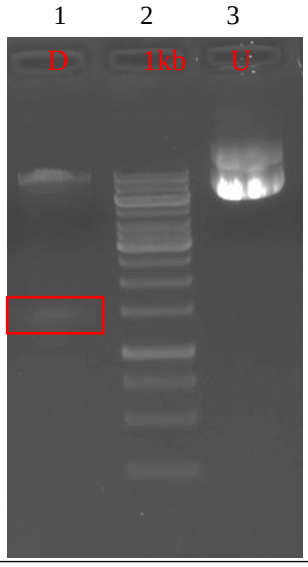



Figure 35: Digestion of the pTK-Ubi-gateway vector resulting in the removal of the Maize ubiquitin promoter (indicated in red)

2.2 The pTK backbone digested out had one sticky end cut and one blunt end cut. The sticky end was blunted using Phusion polymerase enzyme as well as the Klenow fragment. This was followed by a self-ligation and transformation into the CCDB resistant *E. coli* bacterial cells. The CCDB resistant cells were prepared fresh every time a transformation event was carried out- via both chemical mediated preparatory technique as well as electrocompetent cell preparatory technique.

2.3 Nearly 15 different individual attempts (with different conditions) were made to self-ligate the vector but this proved to be unsuccessful. Multiple vectors of varying sized were obtained post transformation into CCDB resistant competent cells, but none of these were of the expected size.

# References


Bowman JL. A Brief History of Marchantia from Greece to Genomics. Plant Cell Physiol. 2016;57(2):210-229. doi:10.1093/pcp/pcv044

Cove, D.J., Perroud, P.-F., Charron, A.J., McDaniel, S.F., Khandelwal, A. and Quatrano, R.S. (2009). Culturing the Moss Physcomitrella patens. Cold Spring Harbor Protocols 2009, pdb.prot5136.

Doyle, J.J. (1990). Isolation of plant DNA from fresh tissue. Focus 12, 13–15.

Erin Gilchrist, George Haughn, Reverse genetics techniques: engineering loss and gain of gene function in plants, Briefings in Functional Genomics, Volume 9, Issue 2, March 2010, Pages 103–110,

Frank W, Decker EL, Reski R. Molecular tools to study Physcomitrella patens. Plant Biol (Stuttg). 2005;7(3):220-227. doi:10.1055/s-2005-865645

Harrison, C.J., Roeder, A.H., Meyerowitz, E.M. and Langdale, J.A. (2009). Local Cues and Asymmetric Cell Divisions Underpin Body Plan Transitions in the Moss Physcomitrella Patens. Current Biology 19, 461–47

Hatzfeld M. The armadillo family of structural proteins. Int Rev Cytol. 1999;186:179-224. doi:10.1016/s0074-7696(08)61054-2

JenniferL. Morris, MarkN. Puttick, James,W. Clark, Dianne Edwards, Paul Kenrick, Silvia Pressel, CharlesH. Wellman, Ziheng Yang, Harald Schneider , Philip C. J. Donoghue “Timescale of early land plant evolution” Proceedings of the National Academy of Sciences Mar 2018, 115 (10) E2274-E2283; DOI: 10.1073/pnas.1719588115

Kimitsune Ishizaki, Ryuichi Nishihama, Katsuyuki T. Yamato, Takayuki Kohchi, Molecular Genetic Tools and Techniques for Marchantia

polymorpha Research , Plant and Cell Physiology, Volume 57, Issue 2, February 2016, Pages 262–270,

Koch CM, Chiu SF, Akbarpour M, et al. A Beginner's Guide to Analysis of RNA Sequencing Data. Am J Respir Cell Mol Biol. 2018;59(2):145-157. doi:10.1165/rcmb.2017-0430TR

Kukurba KR, Montgomery SB. RNA Sequencing and Analysis. Cold Spring Harb Protoc. 2015;2015(11):951-969. Published 2015 Apr 13. doi:10.1101/pdb.top084970

Di Laurenzio L, Wysocka-Diller J, Malamy JE, et al. The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root. Cell. 1996;86(3):423-433. doi:10.1016/s0092-8674(00)80115-4

Ligrone, R., Duckett, J. G., & Renzaglia, K. S. (2012). Major transitions in the evolution of early land plants: A bryological perspective. Annals of Botany, 109(5), 851–871.

Liu YC, Vidali L. Efficient polyethylene glycol (PEG) mediated transformation of the moss Physcomitrella patens. J Vis Exp. 2011;(50):2560. Published 2011 Apr 19. doi:10.3791/2560

Mohanasundaram, B., Rajmane, V.B., Jogdand, S.V. et al. Agrobacterium-mediated Tnt1 mutagenesis of moss protonemal filaments and generation of stable mutants with impaired gametophyte. Mol Genet Genomics 294, 583–596 (2019).

Moresco EM, Li X, Beutler B. Going forward with genetics: recent technological advances and forward genetics in mice. Am J Pathol. 2013;182(5):1462-1473. doi:10.1016/j.ajpath.2013.02.002

Nanjareddy K, Arthikala MK, Aguirre AL, Gómez BM, Lara M. Plant Promoter Analysis: Identification and Characterization of Root Nodule Specific Promoter in the Common Bean. J Vis Exp. 2017;(130):56140. Published 2017 Dec 23. doi:10.3791/56140

Park J, Throop AL, LaBaer J. Site-specific recombinational cloning using gateway and in-fusion cloning schemes. Curr Protoc Mol Biol. 2015;110:3.20.1-3.20.23. Published 2015 Apr 1. doi:10.1002/0471142727.mb0320s110

Reece-Hoyes JS, Walhout AJM. Gateway Recombinational Cloning. Cold Spring Harb Protoc. 2018;2018(1):pdb.top094912. Published 2018 Jan 2. doi:10.1101/pdb.top094912

Shimamura M. Marchantia polymorpha: Taxonomy, Phylogeny and Morphology of a Model System. Plant Cell Physiol. 2016;57(2):230-256. doi:10.1093/pcp/pcv192

Stefan A. Rensing, Bernard Goffinet, Rabea Meyberg, Shu-Zon Wu, Magdalena Bezanilla "The Moss Physcomitrium (Physcomitrella) patens: A Model Organism for Non-Seed Plants". The Plant Cell May 2020, 32 (5) 1361-1376; DOI: 10.1105/tpc.19.00828

Stijn Dhondt, Frederik Coppens, Freya De Winter, Kamal Swarup, Roeland M.H. Merks, Dirk Inzé, Malcolm J. Bennett, Gerrit T.S. Beemster "SHORT-ROOT and SCARECROW Regulate Leaf Growth in Arabidopsis by Stimulating S-Phase Progression of the Cell Cycle" .Plant Physiology Nov 2010, 154 (3) 1183-1195; DOI: 10.1104/pp.110.158857

Tomescu, A. M. F. (2009). Megaphylls, microphylls and the evolution of leaf development. Trends in Plant Science, 14(1), 5–12.

Zavaliev R, Epel BL. Imaging callose at plasmodesmata using aniline blue: quantitative confocal microscopy. Methods Mol Biol. 2015;1217:105-119. doi:10.1007/978-1-4939-1523-1_7