University of Rhode Island

DigitalCommons@URI

Open Access Master's Theses

2020

FUNCTIONAL ANALYSIS OF THE CSLD GENES IN PHYSCOMITRELLA PATENS BY CRISPR/CAS9 GENE KNOCKOUT

Kai Yuan

University of Rhode Island, kaiyuan635@gmail.com

Follow this and additional works at: <https://digitalcommons.uri.edu/theses>

Recommended Citation

Yuan, Kai, "FUNCTIONAL ANALYSIS OF THE CSLD GENES IN PHYSCOMITRELLA PATENS BY CRISPR/CAS9 GENE KNOCKOUT" (2020). *Open Access Master's Theses*. Paper 1858.
<https://digitalcommons.uri.edu/theses/1858>

This Thesis is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu.

FUNCTIONAL ANALYSIS OF THE CSLD GENES IN
PHYSCOMITRELLA PATENS BY CRISPR/CAS9 GENE
KNOCKOUT

BY
KAI YUAN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
BIOLOGICAL AND ENVIRONMENTAL SCIENCES

UNIVERSITY OF RHODE ISLAND

2020

MASTER OF SCIENCE THESIS

OF

KAI YUAN

APPROVED:

Thesis Committee:

Major Professor Alison Roberts

Gongqin Sun

Joel Chandlee

Nasser H. Zawia
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2020

ABSTRACT

Plant cellulose synthase catalytic subunits (CESAs) are known to synthesize the β -1,4-glucan chains that comprise cellulose microfibrils. These CESA subunits associate to form plasma membrane localized Cellulose Synthesis Complexes (CSCs) that facilitate microfibril assembly in concert with β -1,4-glucan polymerization. Cellulose synthase-like D proteins have also been suggested to be involved in cellulose biosynthesis. However, it is not known whether CSLDs form CSCs or produce cellulose in microfibrillar form. This has been difficult to evaluate due to interfering CESA activity being present in the same cells. Whereas CESA null mutations are lethal in vascular plants, some evidence suggests that CESAs are not essential for viability in the moss *Physcomitrella patens*. In order to study whether CSLDs form CSCs and produce microfibrillar cellulose, production of a *P. patens* genetic background in which all seven *PpCESAs* are rendered non-functional was attempted through the use of CRISPR/Cas9 mutagenesis. Knocking out the *PpCESA4* and *PpCESA10* in the *ppcesa3/8/6/7* knockout (KO) background produced viable sextuple *ppcesa* KO lines and resulted in a new phenotype characterized by abnormal gametophore morphology. Subsequent attempts to knockout *PpCESA5* in the sextuple KO generated lines with the expected ‘no gametophore’ *ppcesa5* KO phenotype. Although these lines remain to be genotyped, these results suggest that all CESAs can be eliminated in wild type *P. patens* with no consequence to viability. A role for CSLDs in the development of moss tissues has been demonstrated previously through global knockdown of the *P. patens* CSLD family by RNA interference. In order to study the roles of specific CSLDs in *P. patens* development, double KO mutants for *PpCSLD1/7* were generated. No obvious rhizoid growth and cellulose content differences were observed between wild type *P. patens* and *ppcsld1/7* mutants, possibly due to gene redundancy and/or background CESA

interference. These results also confirmed CRISPR/Cas9 as an efficient gene editing tool in *P. patens*.

ACKNOWLEDGMENTS

I would like to thank my mentor, Dr. Alison Roberts for giving me an opportunity to join her lab and training me throughout the project. As an international student, I really appreciate her help with my proposal and thesis writing work. Also, I was encouraged to pursue my doctoral degree by her patience and support. I would also like to thank my committee members, Dr. Gongqin Sun and Dr. Joel Chandlee, for their suggestions and support for my project and Ph.D. application. I would like to thank the lab members, Arielle Chaves, Rongrong Li, and Jacob Beane for teaching me various lab techniques and helping me with protocol trouble-shooting. Finally, I would like to thank my friends and family for their trust and unconditional support throughout these years of study in the United States.

This research was supported by a URI Division of Research and Economic Development Bridge Grant.

TABLE OF CONTENTS

Abstract.....	ii
Acknowledgement.....	iv
Table of Contents.....	v
List of Tables.....	vi
List of Figures.....	vii
Introduction.....	1
Materials and Methods.....	15
Results.....	25
Discussion.....	39
Appendix.....	46
Bibliography.....	115

LIST OF TABLES

Table 1. Synthesized oligonucleotides for protospacer annealing.....	20
Table 2. Primers used to amplify on target region for genotype analysis.....	22
Table 3. Primers used to amplify off-target region for genotype analysis.....	23

LIST OF FIGURES

Figure 1. CRISPR/Cas 9 engineered nuclease.....	9
Figure 2. Entry clone construction	12
Figure 3. Construction of expression vector	13
Figure 4. Gene editing by the CRISPR/Cas9 system	14
Figure 5. PCR target and off target genotype analysis.....	19
Figure 6. The flowchart of gene knockout for genetic background	25
Figure 7. PCR genotype analysis of transient antibiotic resistant colonies resulting from transformation of <i>ppcesa3/8/6/7</i> KO mutant with PpCESA10 (A) and PpCESA4 KO (B, C) vectors.....	31
Figure 8. The phenotype of <i>ppcesa3/8/6/7/10/4</i> KO mutants and wild-type <i>P. patens</i> ..	32
Figure 9 A schematic showing on target edits between two protospacers.....	32
Figure 10. PCR genotype analysis of transient antibiotic resistant colonies resulting from transformation of <i>ppcesa3/8/6/7</i> KO mutant with PpCESA5 (A) and PpCESA10 KO (B) vectors.....	33
Figure 11. PCR genotype analysis of transient antibiotic resistant colonies resulting from transformation of wild type <i>P. Patens</i> (GD 11) with PpCSLD1 (A) and PpCSLD7 KO (B, C) vectors.....	34
Figure 12. PCR genotype analysis of transient antibiotic resistant colonies resulting from transformation of wild type <i>P. Patens</i> (GD 11) with PpCSLD3 (A) and PpCSLD4 KO (B) vectors.....	35
Figure 13. Amino acid sequence of CSLD4 off-targets region	36
Figure 14. Dark field images of wild type and <i>ppclsdl1/7</i> mutants leafless gametophore (arrows) with rhizoids (arrowheads).....	37
Figure 15. Images of fluorescence microscopy.....	38

INTRODUCTION

Overview

Cellulose Synthase (CESA) proteins are well-characterized cellulose synthases in plants (McFarlane et al. 2014). Recent studies suggest other enzymes are also capable of synthesizing cellulose in plants, particularly in cells that grow by tip growth such as root hairs (Bernal et al. 2008; Favery et al. 2001; Galway et al. 2011; Kim et al. 2007; Li et al. 2016; Park et al. 2011; Wang et al. 2001) and pollen tubes (Bernal et al. 2008; Doblin et al. 2001). Among the nine other families in the CESA superfamily, the CESA-like D (CSLD) proteins have the most similar amino acid sequences to the CESAs. *Physcomitrella patens* is a moss containing tip-growing protonemal filaments and rhizoids and it has both CESAs and CSLDs. According to the *Physcomitrella* eFP Browser (http://bar.utoronto.ca/efp_physcomitrella/cgi-bin/efpWeb.cgi), some *Physcomitrella patens* CSLDs (*PpCSLDs*) are highly expressed in protonema and rhizoids. Also, a previous study using global knockdown of the entire *PpCSLD* family by RNA interference (RNAi) showed a decrease in protonemal tip growth, indicating that *PpCSLDs* play a role in protonemal growth (Dimos 2010). However, no obvious phenotype was observed in *ppcsld1*, *ppcsld2*, *ppcsld8* single knockout (KO) mutants, suggesting functional redundancy. These results suggest that at least some of the eight *CSLD* genes in *P. patens* are related to tip growth and rhizoid development, but the individual roles for each of the *PpCSLD* family members remain unknown. This project generated *CSLD* double knockout lines by using CRISPR/Cas9 to define the specific roles of *PpCSLD* family members in protonemal growth and rhizoid development. CRISPR/Cas9 was also used in an attempt to delete all seven members of the *P. patens* *CESA* gene family to create a background line for investigating *CSLD* activity without interference from *CESA* activity.

Cellulose and Cellulose Synthases

As a major component of paper products, cellulose has great economic and practical value. Cellulose is the microfibrillar component of plant cell walls and is composed of hydrogen-bonded β -1,4-glucan chains. β -1,4-glucan chains are synthesized by cellulose synthase enzymes, which are members of the glycosyltransferase-2 protein family (GT-2) (Coutinho et al. 2003).

Cellulose synthases, originally studied in *Acetobacter xylinum* (Delmer 1999), are characterized by an eight transmembrane helix topology and conserved cytosolic substrate binding and catalytic domain containing a D, D, D and QXXRW motif (McFarlane et al. 2014). Besides these conserved domains found in all cellulose synthases, land plant cellulose synthases (called CESAs) have additional domains, including an extended N-terminal Zn-binding RING domain, and a plant-conserved region (P-CR) and class-specific region (CSR) within the loop between transmembrane helix 2 and 3 (Pear et al. 1996). These specific domains in higher plants are hypothesized to function in CESA subunits interaction (Somerville 2006). The eight transmembrane helices interact to form a pore in the membrane. The catalytic region in the cytoplasm binds the substrate UDP-glucose, then the glucan chain is secreted through the transmembrane pore to the cell wall (Morgan et al. 2013).

Cellulose microfibrils are synthesized by membrane-bound complexes called Cellulose Synthesis Complexes (CSCs) (Delmer 1999; McFarlane et al. 2014; Somerville 2006). The complexes can be visualized by freeze fracture electron microscopy (FF-EM) (Mueller and Brown 1980; Kimura et al. 1999; Nixon et al. 2016). CSCs contain CESAs based on FF-EM immunolabeling (Kimura et al. 1999), and 18 CESA subunits make up CSCs (Nixon et al. 2016; Vandavasi et al. 2016; Jarvis 2018;

Newman et al. 2013).

CSLDs and *Physcomitrella patens*

Recent studies indicate not only CESAs, but also other enzymes in land plants might synthesize cellulose. These include proteins encoded by some of the *CESA-like* (CSL) genes: *CSLA-J* (Cutler and Somerville 1997; Richmond 2000; Richmond and Somerville 2000, 2001; Hazen et al. 2002). Among the other nine protein families encoded by the *CESA* gene superfamily, CSLDs have higher sequence similarity to CESAs than other CSL proteins. All CESAs and CSLs are members of the GT-2 protein family containing catalytic residues D, D, D, and QXXRW (Coutinho et al. 2003). Unlike the other CSLs, the CSLDs share the same membrane topology as CESAs (Richmond and Somerville 2000). The sequences between catalytic D and QXXRW residues in CLSDs are highly similar to the CESAs (Doblin et al. 2010). Also the CSLDs and CESAs have Zn-binding RING domains which are not found in CSLB, E, G and H families (Doblin et al. 2001; Roberts and Bushoven 2007).

A significant amount of research has been directed at evaluating the CSLDs of vascular plants. In *Arabidopsis*, CSLD mutants have defects in pollen tube development (Doblin et al. 2001; Bernal et al. 2008) and root hair development (Bernal et al. 2008; Favery et al. 2001; Galway et al. 2011; Kim et al. 2007; Li et al. 2016; Park et al. 2011; Wang et al. 2001). Cellulose is one of the main components of root hair tips and cellulase can cause tip rupture in root hairs (Dumais et al. 2004; Park et al. 2011). However, *CESA* mutations have no obvious effect on root hair tip growth while *CSLD* mutations show root hair rupture (Bernal et al. 2008; Favery et al. 2001; Galway 2006; Galway et al. 2011; Wang et al. 2001). These observations suggest that the CSLDs are involved in cellulose synthesis. The hypothesis that CSLDs make cellulose or cellulose-

like β -1,4-glucan chains is further supported by several cell wall assembly studies in which inhibitors of cellulose biosynthesis are used as tools. One of these studies showed that β -1,4-glucan is produced by isoxaben-habituated *Arabidopsis* cells, while upregulation of *AtCSLD5* is observed by global transcript profiling (Manfield et al. 2004). Functional analysis of cell-wall synthesis in tip-growing cells using inhibitors also indicates involvement of a specific CSLDs in cellulose or a cellulose-like polysaccharide synthesis. Isoxaben inhibits CESAs activity, but has no effect on tip growth or localization of CSLD3 (Park et al. 2011; Favery et al. 2001). In contrast, a range of concentrations of 2, 6-dichlorobenzonitrile (DCB), which inhibits both CESAs and CSLDs, causes rupture in tips (Park et al. 2011). Some CESAs and CSLD3 localizations are changed after CGA 325'615 (CGA) treatment, indicating that this inhibitor affects both CESAs and CSLDs (Park et al. 2011).

Several studies in *P. patens* indicate that CSLDs play a role in cellulose biosynthesis. Cellulose has been found in developing protonemal filaments (Berry et al. 2016). Based on open microarray data analysis with Genevestigator (Hiss et al. 2014), six *PpCSLDs* are highly expressed in protonema or cells in the process of initiating protonemal growth. The cellulose synthesis inhibitors isoxaben and DCB have no effect on protonemal tip growth rate, which is consistent with the study in a close relative of *P. patens*, *Funaria hygrometrica* (Rudolph et al. 1989). However, only DCB but not isoxaben has an effect on protonemal cell integrity (Tran et al. 2018). Global silencing of the entire *PpCSLD* family by RNAi leads to inhibited protonemal tip growth, although no obvious phenotype was observed in *ppcsld1*, *ppcsld2*, *ppcsld8* single knockout mutants (Dimos 2010). In addition, a previous study indicates quadruple *ppcesa4/6/7/10* KO and triple *ppcesa3/5/8* KO mutants show only a slight reduction in protonemal growth (Li 2017). Other *ppcesa* KO mutants only show defects in

gametophore development (Goss et al. 2012; Norris et al. 2017). This evidence suggests that the CSLDs play an essential role in protonema development.

Although the developmental roles of CSLDs have been well studied in root hairs, the methodology for investigating CSLD biochemical function in these cells is difficult. These tip growing cells are difficult to culture and harvest in large quantities. Due to the fact that several *CESA* null mutations in vascular plants are lethal (Desprez et al. 2007; Persson et al. 2005), it has only been possible to study CSLD function in the presence of background *CESA* activity, confounding any direct analysis. CSLD-catalyzed cellulose synthesis has been difficult to prove because β -1,4-glucan chains, potentially produced by CSLD proteins, are difficult to distinguish from cellulose microfibrils synthesized by *CESA* proteins. The Zn-binding RING domains in *CESAs* are thought to be required for CSCs formation (Kurek et al. 2002), so similar CSCs could be formed from CSLDs and synthesize microfibrillar cellulose (Park et al. 2011; Doblin et al. 2010). But it is difficult to determine whether CSLDs form CSCs since CSLD-containing CSCs cannot be distinguished from *CESA*-containing CSCs by FF-EM. If all *CESAs* could be deleted, CSLD activity could be studied directly without any *CESA* activity interference. Thus, the question of whether CSLDs form CSCs in the plasma membrane and produce microfibrillar cellulose could also be examined in all-*CESA* knockout lines.

P. patens is a potential model system to overcome some of the disadvantages of *Arabidopsis* for investigating CSLD function. It is a small plant and typical of many mosses with a dominant haploid phase. In a short life cycle, it first germinates from a haploid spore and then develops into a protonema by polarized tip growth. Chloronema and caulonema are two different cells in the protonema. Compared with chloronema, caulonema has a higher growth rate and fewer chloroplasts. In terms of function,

caulonema works in scavenging nutrients, while chloronema functions primarily in photosynthesis (Schumaker and Dietrich 1998). Buds are formed from protonema after about two to three weeks, and then give rise to leafy gametophores. Contrary to protonema, leafy gametophores do not grow by polarized tip growth and can form rhizoids on the stems. Rhizoids show functionality similar to root hairs in vascular plants and can be stimulated by auxin (Sakakibara et al. 2003). At the end, antheridia and archegonia (reproductive organs) are formed from leafy gametophores. Sperm produced from antheridia can get access to archegonia and fertilize an egg when water is present (Schumaker and Dietrich 1998). A diploid sporophyte consisting of a stalk and capsule is formed from the zygote. Haploid spores then can disperse after meiosis in the capsule (Schumaker and Dietrich 1998). The diploid phase is very rare in culture (Cove 2000). As a non-vascular plant containing tip growing protonemal filaments and rhizoids, *P. patens* could be used as an experimental system to investigate CSLD function because of its advantages. It has a fully sequenced genome (Rensing et al. 2008; Zimmer et al. 2013) with seven *CESA* genes and eight *CSLD* genes (Roberts and Bushoven 2007). Large quantities of tip growing protonemal filaments can be easily cultured for study (Cove et al. 2009). Also, *CESA* activity may not be required for viability which may make it possible to generate viable lines with no *CESAs* to investigate *PpCSLDs* function without *CESA* activity interference, which it is not possible in *Arabidopsis* (Desprez et al. 2007; Persson et al. 2005).

CRISPR/Cas9

Gene editing is a form of genetic engineering that can potentially be used to precisely modify any organism's genome by DNA insertion, deletion or replacement at specific targeted sites. So far, several synthetic nucleases have been used in gene editing

strategies. They are zinc finger nucleases (ZFNs) (Gaj et al. 2013), transcription activator-like effector nucleases (TALENs) (Joung and Sander 2013), mega-nucleases (Silva et al. 2011), and the RNA guide CRISPR/Cas9 system which has become widely used in recent years (Sander and Joung 2014). All of these strategies use the engineered nucleases to introduce genetic modifications into genomes (Gaj et al. 2013). Gene editing is a powerful tool that can be applied in basic studies of gene function in plants and animals. It can also be applied for crop improvement through trait modification and for gene therapy to treat human diseases. Compared to classic transgenesis, gene editing is a better approach for targeted genetic modifications since the strategy is more straightforward, and it can be adapted in most organisms.

CRISPR is short for Clustered Regularly Interspaced Short Palindromic Repeats, and Cas9 is an enzyme that works with a guide RNA to cut targeted DNA sequences. CRISPR/Cas9 gene editing technology is derived from a natural acquired immunity system that bacteria use to fight off viral infections before they cause damage (van der Oost et al. 2009; Marraffini 2015). Generally speaking, CRISPR allows bacteria to generate specific memory of invaders. This acquired memory allows bacteria to recognize past invaders and provides an ability to resist future infection by similar foreign viral invaders. The natural CRISPR system is characterized by three stages, the adaptation, expression, and interference stages (van der Oost et al. 2014; Garneau et al. 2010; Makarova et al. 2011; Gasiunas et al. 2014; Doudna and Charpentier 2014). In the adaptation stage, bacteria can incorporate small pieces of the viral genetic material into the CRISPR locus of their own genomes (Hsu et al. 2014; Yosef et al. 2012). A highly variable locus from different *S. thermophilus* strains was identified through comparative genomics analysis (Bhaya et al. 2011). In this highly variable region, variable sequences known as spacers and many non-contiguous repeats were found.

During the expression stage of the type II CRISPR system, Cas9 and two RNA components are formed. First, foreign DNA fragments working as molecular memory of previous invaders within the repeat CRISPR array are transcribed with CRISPR repeats as CRISPR RNAs (crRNAs) (Brouns et al. 2008). Another RNA component called trans-activating CRISPR RNA (tracrRNA) is transcribed from upstream of the CRISPR locus (Deltcheva et al. 2011). A small region of tracrRNA is complementary to the repeat CRISPR array in crRNA, a double strand RNA is formed. Then the complex is cleaved by RNase III, and a crRNA:tracrRNA complex with one spacer is formed. This complex is loaded onto Cas9 to form an active ribonucleoprotein (RNP) complex. In the interference stage, the crRNA:tracrRNA complex guides Cas9 to target sequences based on complementary base pairing, then Cas9 cuts the binding DNA. The specificity of Cas9 relies on a short DNA sequence next to target sequence termed Proto-spacer Adjacent Motifs (PAMs) which enables Cas9 to distinguish self from non-self DNA (Brouns et al. 2008; Garneau et al. 2010; Marraffini and Sontheimer 2008). In the type II CRISPR/Cas9 system, the complex causes a double strand break by binding to the target sequence and prevents the foreign DNA from propagating (Brouns et al. 2008).

The type II CRISPR/Cas9 system is the most popular CRISPR gene-editing system due to its simplicity requiring one Cas protein, and this artificial CRISPR/Cas9 system contains two important parts. One is the guide RNA, which is a synthetic RNA targeting a specific sequence in DNA, and the other is the Cas9 enzyme, which acts as an endonuclease. The single guide RNA (sgRNA) contains crRNA and tracrRNA. The endonuclease Cas9 forms a complex with the 3' end of the sgRNA, and the 5' end of the sgRNA recognizes a predefined 20 bp protospacer target sequence in the genome based on complementary base pairing. The specific target ensures that Cas9 cuts at the

selected site in a DNA sequence. There is a limitation to protospacer targeting in the genome because PAM is required next to the protospacer for Cas9 recognition and cleavage (Figure 1). Different Cas9 proteins from different species of bacteria recognize different PAM sites. The most widely used is SpCas9, isolated from *Streptococcus pyogenes*. “NGG” is the PAM recognized by this endonuclease. With the guide RNA-target DNA complementarity and PAM-based DNA recognition, two nuclease domains, RuvC and HNH, in the Cas9 protein cut the DNA sequence three nucleotides upstream of the PAM site and cause a blunt double-strand DNA (dsDNA) break (DSB) with high specificity (Garneau et al. 2010). After that, two DNA repair pathways, the Non-Homologous End Joining (NHEJ) pathway and the Homology Directed Repair (HDR) pathway, will be initiated (Gaj et al. 2013). In the NHEJ repair pathway, an insertion/deletion (InDel) is generated in the target site, and a frameshift or premature stop codons is introduced into the open reading frames (ORFs) of target genes. Compared with NHEJ, HDR is less effective in creating InDel mutations, but very useful for inserting tags. The HDR pathway uses a donor DNA template to repair the DSB, and specific DNA sequences are inserted into the target site.

Figure 1. CRISPR/Cas 9 engineered nuclease. The sgRNA works with the Cas9 endonuclease to ensure target-specific modification of the genome.

The goal of this project was to investigate the functional roles of the *CSLD* genes. To achieve this goal, two different strategies were employed. First, attempts were made to produce a valuable genetic background in *P. patens* in which all *CESAs* are knocked out to investigate *CSLD* activity without interference from *CESA* activity. Starting with an existing quadruple *ppcesa3/8/6/7* KO line (gift of Arielle Chaves), CRISPR/Cas9 KO vectors for the remaining *PpCESAs* were designed and applied. In addition, studies were also performed to analyze *CSLDs* in the wild type background by double knockout. *P. patens* has undergone two ancestral whole genome duplication events (Lang et al. 2018). Following the most recent duplication, one ancestral *CSLD* gene duplicated to form a tandem pair on chromosome 2, *CSLD1* and *CSLD7*, both of which are highly expressed in rhizoids. Another ancestral *CSLD* gene duplicated to form an inverted pair on chromosome 1, *CSLD3* and *CSLD4*, both of which are expressed in protonema. The two members of these pairs are likely to have redundant functions. So CRISPR/Cas9 vectors for *PpCSLD1*, *PpCSLD3*, *PpCSLD4*, *PpCSLD7* were made and used to generate *ppcsld1/7* and *ppcsld3/4* double knockouts.

Dual Cas9-sgRNA expression vectors were designed to target *PpCSLDs* and *PpCESAs*. The CRISPOR web tool (<http://crispor.tefor.net/>) (Haeussler et al. 2016) was used to identify targeted CRISPR sequences in the *P. patens* genome. For each targeted *CESA* and *CSLD* gene, two protospacers were identified. Using the CRISPR/Cas9 vector system developed by the Bezanilla Lab at Dartmouth College (Mallett et al. 2019), annealed protospacers were cloned to pENTR Multisite Gateway plasmids (pENTR-PpU6p-sgRNA) (Figure 2A). For this step, the Roberts' lab (URI) has developed a rapid cloning procedure using NEB Golden Gate Assembly. Then the sgRNA expression cassettes (two for each target gene) were moved to a destination vector (Figure 3A), containing the Cas9 gene and one of three antibiotic resistance

cassettes, by Multisite Gateway Pro Technology (Invitrogen, Carlsbad, CA). The collection of CRISPR/Cas9 KO vectors was used to knock out the *PpCESAs* in an existing quadruple *ppcesa3/8/6/7* KO line and *PpCSLDs* in wild type *P. patens*. Protoplasts were transformed with vectors that express dual sgRNAs (Figure 4), then screened by antibiotic selection. After regenerating colonies, potential KO colonies were screened with PCR by looking deletions, then sequenced to identify targeted and off-targeted mutations. Finally, the roles of the *PpCSLDs* family members in KO derivatives were examined for phenotypes in *PpCSLD* double knockouts.

Figure 2. Entry clone construction. (A) Plasmid maps of pENTR-Ppu6p-sgRNA-L1R5 and pENTR-Ppu6p-sgRNA-L5L with modified att sites. (B) Detail of the BsaI recognition (red) and cut sites (arrows). BsaI cutting removes the recognition sites and generates overhangs allowing for directional ligation of custom protospacers.

A:

B:

C:

Figure 3. Construction of expression vector. (A) Plasmid maps of destination vectors with Gateway® cassette, Cas9 expression cassette, and antibiotic selection cassettes (hygromycin, G418, and zeocin resistance). (B) Two-fragment LR recombination reaction for insertion of sgRNAs expression cassettes into destination vectors. (C) Final plasmid map with 2 sgRNA expression cassettes, Cas9 expression cassette, and antibiotic selection cassettes (hygromycin, G418, and zeocin resistance).

Figure 4. Gene editing by the CRISPR/Cas9 system. A CRISPR/Cas9 plasmid vector is used for PEG-mediated protoplast transformation. Inside the protoplast, the vector drives expression of the Cas9 protein and sgRNA, and the Cas9/sgRNA complex targets the moss genome for editing.

MATERIALS AND METHODS

Vector Construction

All gene KO vectors were constructed as previously described (Mallett et al. 2019). The publicly available CRISPOR online software (<http://crispor.tefor.net/>) (Haeussler et al. 2016) was used to identify candidate protospacers in the *P. patens* genome. Protospacers were designed to target individual genes: *PpCSLD1* (Pp3c2_1280V1.1), *PpCSLD3* (Pp3c1_41250V1.1), *PpCSLD4* (Pp3c1_41400V1.1), *PpCSLD7* (Pp3c2_1330V1.1), *PpCESA4* (Pp3c9_2550V1.1), *PpCESA5* (Pp3c2_13330V1.1), *PpCESA10* (Pp3c9_2670V1.1). For each targeted gene, exon sequences were submitted to CRISPOR and protospacer candidates were generated. Two protospacers that were at least 300 bp apart were chosen to target each gene. The protospacers were chosen based on high value for out-of-frame and few off-target sites. By using two protospacers, genotyping was greatly simplified because deletions were easily identified by PCR genotype analysis. All designed protospacers are listed in Table 1.

To make double-stranded protospacers for cloning purposes, two complementary synthesized oligonucleotides (50 μ M each) were annealed using a thermocycler (98°C for 3 min, slow ramp (1%) to primer T_m, hold at T_m for 10 min, slow ramp (1%) to 25°C and hold). sgRNA expression cassettes were constructed by cloning annealed protospacers into pENTR-PpU6p-sgRNA entry vectors (Figure 2A) using the Golden Gate Assembly method (Figure 2B), in which BsaI cleavage and T4 ligation are performed simultaneously (New England Biolabs, Ipswich, MA). The two protospacers targeting each gene were cloned into two different entry vectors, pENTR-Ppu6p-sgRNA-L1R5 and pENTR-Ppu6p-sgRNA-L5L2 (Figure 2A), so that both sgRNA expression cassettes could be cloned into the same destination vector. In each

10 µl Golden Gate Assembly reactions, 37.5 ng entry vector and 1.1 µl 1000-fold annealed protospacer were incubated at 37°C for 1 h to promote BsaI cutting and the ligation of the protospacer, then at 60°C for 5 min to favor BsaI cutting without DNA ligation. The assembly reaction products were transformed into TOP10, *E. coli* cells (Invitrogen Life Technologies, Carlsbad, CA), and kanamycin resistant colonies were picked for overnight culture. Plasmid DNA was extracted and purified from antibiotic resistant colonies using a QIAGEN QIAprep Spin Miniprep kit according to manufacturer's instructions (QIAGEN Inc, Valencia, CA, USA). The resulting plasmids were screened by restriction digestion with BsaI and PvuI or ApaLI, then confirmed by sequencing on the Applied Biosystems BigDye® Terminator v3.1 chemistry (Life Technologies).

Three different Gateway-enabled destination vectors, pZeo-Cas9-gate, pMH-Cas9-gate, and pMK-Cas9-gate (Figure 3A), conferring zeocin, hygromycin, and G418 resistance respectively, were used for dual Cas9-sgRNA expression vector construction. All three vectors include a maize ubiquitin promoter driving the expression of Cas9 (Figure 3A). MultiSite Gateway® Pro LR recombination was used to transfer two sgRNA expression cassettes targeting the same gene to one of the three destination vectors (Figure 3B, C). For the 5 µl LR reaction, 10 ng of each entry clones and about 90 ng destination vector were mixed with 1 µl LR Clonase™ II Plus enzyme mix. After 16 h incubation at room temperature, 0.5 µl Proteinase K solution was added to terminate the reaction. Then the LR reaction products were transformed into *E. coli* cells as described for entry clone transformation. Plasmids from ampicillin resistant colonies were purified and the resulting potential KO constructs were first screened by restriction digestion with EcoRI or EcoRV, then confirmed by sequencing.

Moss tissue culture and protoplast transformation

Moss Subculture

Protonemal filaments of *P. patens* Gransden wild type strain (Ashton and Cove, 1977) and quadruple *ppcesa3/6/7/8* KO line (gift of Arielle Chaves) were cultured as described previously (Roberts et al. 2011). About 150 mg of tissue from the starter culture were scraped into a mound with sterile forceps, then transferred to a sterile 15 ml centrifuge tube with 6-8 ml sterile water. The tissue was broken up using a rotor-stator type tissue homogenizer with a sterile hard tissue tip for 10-20 s on medium speed (Omni International, Kennesaw, GA, USA), and about 2 ml of culture suspension was inoculated onto a cellophane-overlain BCDAT (Roberts et al. 2011) plate. Protonemal tissue cultured at 25°C with constant illumination at 50-80 $\mu\text{mol m}^{-2} \text{s}^{-1}$ at 6-day intervals (at least two rounds) was used for protoplast transformation.

Moss Transformation

Protoplasts were isolated from *P. patens* and transformed as described by Roberts et al. (2011). Briefly, protonemal tissue from two rounds of subculture was digested with driselase for 60 min (wild type) or 30 min (quadruple, quintuple, sextuple KO) to remove the cell wall. Protoplasts were collected by centrifugation, washed three times with 8.5% mannitol, and resuspended in 3M solution. Protoplast suspension and about 10 μg purified KO plasmid were mixed in polyethylene glycol (PEG) solution and heat-shocked (Figure 4). Then protoplasts were resuspended in 3 ml PRML (Roberts et al. 2011) and about 1 ml suspension was spread on cellophane-overlain PRMB (Roberts et al. 2011) plates for 4 days. Transiently transformed protoplasts were selected with 50 $\mu\text{g mL}^{-1}$ G418 (CSLD7 KO and CESA10 KO vectors) or 15 $\mu\text{g mL}^{-1}$ hygromycin (CSLD4 KO and CESA4 KO vectors) or 50 $\mu\text{g mL}^{-1}$ zeocin (CSLD1 KO,

CSLD3KO, and CESA5 KO vectors) for 7 days on BCDAT plates. Regenerating protonemal colonies were moved to BCDAT without antibiotics after 7 days selection for maximal growth.

Genotype Analysis

Primer Design

PCR genotyping primers were designed to amplify a region approximately 300 bp upstream of the first protospacer to 300 bp downstream of the second protospacer using the Primer 3 web tool (<http://bioinfo.ut.ee/primer3-0.4.0/>) (Figure 5A). The designed genotyping primers (Table 2) were used for gene deletion analysis.

Off-target primer information can be found in the output of the desired protospacer page, with sites of CFD scores < 0.02 ignored. The remaining off-target amplicon sequences were used to blast against the COGE database (<https://genomevolution.org/coge/CoGeBlast.pl>), and the sequence for the gene that contains the off-target site was downloaded and saved for off-target primer design. Primers were designed to amplify 500-800 bp with the off-target site in the middle of amplicon sequence using Primer 3 (Figure 5B). All designed primers are listed in Table 3.

A:

B:

Figure 5. PCR target and off target genotype analysis. (A) Forward primer anneals about 300 bp upstream of the first protospacer and reverse primer anneals about 300 bp downstream of the second protospacer. (B) Off-target genotyping primers are designed to amplify off-target sites.

Table 1. Synthesized oligonucleotides for protospacer annealing

Oligonucleotides	Sequence	Out-of-Frame	Number of off-targets	T _m
CSLD1_gRna289f_s	ccatCGTAGTTCTAGGGGGGCTGC	73	0	68° C
CSLD1_gRna289f_a	aaacGCAGCCCCCCTAGAACTACG			
CSLD1_gRna922r_s	ccatAACCGTAAGTGCCCTTAGTC	78	2	62.9 °C
CSLD1_gRna922r_a	aaacGACTAAGGGCACTTACGGTT			
CSLD3_gRna277f_s	ccatTCGGTAAGCAGGGTAATGCG	61	0	64.6 °C
CSLD3_gRna277f_a	aaacCGCATTACCCTGCTTACCGA			
CSLD3_gRna847f_s	ccatATGTTGCGCCGCGATGACTTG	65	1	64.6 °C
CSLD3_gRna847f_a	aaacCAAGTCATCGCGGCGAACAT			
CSLD4_gRna220r_s	ccatGAAGGTCAGACCTACCACTG	66	2	64.6 °C
CSLD4_gRna220r_a	aaacCAGTGGTAGGTCTGACCTTC			
CSLD4_gRna1271r_s	ccatCTACGTTACGGAACCTAGG	62	1	64.6 °C
CSLD4_gRna1271r_a	aaacCCTAGGTTCCGTGAACGTAG			
CSLD7_gRna274f_s	ccatTCAGCCGGGTAACATGGGAC	69	0	66.3 °C
CSLD7_gRna274f_a	aaacGTCCCATGTTACCCGGCTGA			
CSLD7_gRna1027r_s	ccatACCACCGTCATTGCTACCGT	70	2	64.6 °C
CSLD7_gRna1027r_a	aaacACGGTAGCAATGACGGTGGT			
CES5_gRna789r_s	ccatTCGAGACGTACACTAGCCTG	64	1	64.6 °C
CES5_gRna789r_a	aaacCAGGCTAGTGTACGTCTCGA			
CES5_gRna1819r_s	ccatCTACGGAACCCAAAGCCAC	76	4	66.3 °C
CES5_gRna1819r_a	aaacGTGGGCTTTGGGTTCCGTAG			
CesA4_gRna27f_s	ccatTGGATCCGGCGAAAGATTTG	74	2	62.9 °C
CesA4_gRna27f_a	aaacCAAATCTTTCGCCGGATCCA			
CesA4_gRna17r_s	ccatTCCACCGGCGCCAACTGAGA	58	2	68° C
CesA4_gRna17r_a	aaacTCTCAGTTGGCGCCGGTGGA			
CesA10_gRna139f_s	ccatCAGCCAGATGGCTTATGGTC	67	1	64.6 °C
CesA10_gRna139f_a	aaacGACCATAAGCCATCTGGCTG			
CesA10_gRna42f_s	ccatATGCGGGCTCGGATGGGTAT	75	2	66.3 °C
CesA10_gRna42f_a	aaacATACCCATCCGAGCCCGCAT			

PCR screening for gene knockout

Tissue was collected from colonies selected for transient antibiotic resistance and genomic DNA was extracted from potential KO mutants as described previously (Roberts et al. 2011). DNA extracted from each potential KO line (1-2μL) was

amplified with genotyping primers with *Taq* DNA Polymerase in 25 µl reactions according to the manufacturer's protocol (New England Biolabs, Ipswich, MA). Colonies were screened for deletions in the target gene by gel electrophoresis of the PCR products. A 50 µl PCR reaction was conducted for colonies showing small products (as expected for deletions), and 5 µl of product was used for gel electrophoresis to confirm successful amplification. The remaining 45 µl of product was cleaned up by Monarch® PCR & DNA Cleanup Kit according to manufacturer's instructions (New England Biolabs, Ipswich, MA), and sequenced. In some cases, PCR products with no apparent deletion were amplified and sequenced as described above.

Sequence Analysis

Sequencing results were analyzed with CodonCode Aligner (CodonCode Corporation, Centerville, MA). For on-target genotyping, the actual sequence and edited sequences were assembled, and two protospacers were searched and labeled in edited sequences. Then the edited sequence was assembled with its own sequence without trimming. The translated nucleotide sequence was analyzed around two protospacer sites by comparing with the actual amino acid sequence.

Analysis of off-target gene modification

The resulting lines with frameshifts were screened for off-target edits by PCR amplification with off-target primers and sequencing of the PCR products as described above. For off-target genotyping, sequencing results were assembled with the actual sequence. The off-target site was highlighted in the actual sequence and compared to the mutants' sequence.

Table 2. Primers used to amplify on target region for genotype analysis

Primer	Sequence	Annealing temperature
CSLD1_Crdel-F	AAAACACTCAGGCAATCGGG	53°C
CSLD1_Crdel-R	CCATCTTCACCAGGCCAAAC	
CSLD3_Crdel-F	CGACAAGGAGAAGGAGGAGG	53°C
CSLD3_Crdel-R	GCGTGGTCAAAATCCGTAGT	
CSLD4_Crdel-F	TGGTTCGCTTTCTCCTGGAT	53°C
CSLD4_Crdel-R	GTCAACCCTGTTACCTCCGA	
CSLD4_Crdel-F2	CGATGCTATATGGTTGTGGGG	53°C
CSLD4_Crdel-R2	TCGTAGAAGCAGGACACCAC	
CSLD7_Crdel-F	TTCAGCGATCCGACTTTTGC	53°C
CSLD7_Crdel-R	GACAAAGGACGGCGAGATTT	
CSLD7_Crdel-F2	TGGACTGGTAGTGTGGCATT	53°C
CSLD7_Crdel-R2	CCAGAATCATGCTACGAAGGG	
CESA4_Crdel-F	GGGCGGGAAAGGATCTCATA	53°C
CESA4_Crdel-R	CCAGAATCGACAGCACAGTG	
CESA4_Crdel-F2	ACATCCCCAGATCATCAAGCT	52°C
CESA4_Crdel-R2	GGGGCTCGATGTTGAACTT	
CESA4_Crdel-F3	GTGTCAAGCTGCGAAATGGA	53°C
CESA4_Crdel-R3	AGCTCATGTACCTCACCTT	
CESA5_Crdel-F	CAACCGCGAGACATACTTGG	53°C
CESA5_Crdel-R	TGCTTGTTTGGAAGTGACGG	
CESA10_Crdel-F	ACTCCGACGACCTAGACAAC	54°C
CESA10_Crdel-R	CTCCCTCTCCACTTGCTTGA	
CESA10_Crdel-F2	TGAAGAAGAGGACTCCGACG	54°C
CESA10_Crdel-R2	CCTAGTGCTTTGCGACATCC	
CESA10_Crdel-F3	GGTTGAGGGATGTGCATGTG	53°C
CESA10_Crdel-R3	TGCTCGATGGGAATGGAAC	

Table 3. Primers used to amplify off-target region for genotype analysis

Primer	Sequence	Annealing temperature
CSLD1_OffT_1F	TCCCATCTGTGCTGTCTGAAG	52°C
CSLD1_OffT_1R	CGAAGGGAAGCTCATTGAAGT	
CSLD1_OffT_2F	CAGAGCGAGCACAAACCATAC	53°C
CSLD1_OffT_2R	CAAGGCGATGAGGGACGA	
CSLD3_OffT_1F	AGACCAACATCTCCAGTGCA	53°C
CSLD3_OffT_1R	ATGCCTGGAAGTGGGTCAAT	
CSLD4_OffT_1F	AGAAGAAGAAGCGCAGAGGT	53°C
CSLD4_OffT_1R	GCATCCGAACTGACTGCAA	
CSLD4_OffT_1F2	ATGGAACGAGATGGCAATC	52°C
CSLD4_OffT_1R2	GCGCTGTTCTCACGAAGC	
CSLD4_OffT_2F	AGCATTGTACGCATCGGAAC	53°C
CSLD4_OffT_2R	CAAGGAATGAAGCTGTGGGG	
CSLD4_OffT_3F	ACTCTCTCTTGACCTTGCGG	53°C
CSLD4_OffT_3R	TGTGTCCACGCTAGCAATTG	
CSLD4_OffT_3F2	ACTCTCTCTTGACCTTGCGG	52°C
CSLD4_OffT_3R2	TCGCAAAGTCAGCATTTCTGA	
CSLD7_OffT_1F	CGCTCAGAGTATTGTTGCTTCA	52°C
CSLD7_OffT_1R	TGCCTCCATCTTCTCCACTT	
CSLD7_OffT_2F	CAAAGGCATACGGAGCTGTC	53°C
CSLD7_OffT_2R	GCCCACGCTAAAATCAGTTCA	
CESA4_OffT_1F	AGTTCGTTCTCACTGCACT	53°C
CESA4_OffT_1R	CACGCCAAGTCACTCCAAAG	
CESA4_OffT_2F	AGCTCATGTCACCTCACCTT	53°C
CESA4_OffT_2R	CGACTTTCCTGAGGTACGT	
CESA4_OffT_3F	GGTGGTGATTGTTACGGTGG	53°C
CESA4_OffT_3R	TGGCCATGAGAACTGAAGC	
CESA10_OffT_1F	GTGTCAAGCTGCGAAATGGA	53°C
CESA10_OffT_1R	GAGCTTCGTTTCCTGCAGAG	
CESA10_OffT_1F2	GTGTCAAGCTGCGAAATGGA	53°C
CESA10_OffT_1R2	CAACCAACACACCTCACACC	
CESA10_OffT_2F	GGGGAGGGGTATGAAGGAGA	55°C
CESA10_OffT_2R	ATCTGCTCGCTATGGTGGG	
CESA5_OffT-1F	GGTCAGGGTCACTTGGATCA	52°C
CESA5_OffT-1R	TTGTGCGCATGCTTTGGAAA	
CESA5_OffT_2F	AAGACAGACTCGGGACAAGG	51°C
CESA5_OffT_2R	TCAACTGCCATTACTCTGCA	
CESA5_OffT_3F	CAACGCAATGCAGTCTCAGA	53°C
CESA5_OffT_3R	AAGACATTCCAGGGGCAGC	
CESA5_OffT_4F	ATGCGACAGGGGAGAGTATG	53°C
CESA5_OffT_4R	TTTCTCGTGGTGTGCTGTG	
CESA5_OffT-5F	CCAAGTGCCGGCAGTATTAC	54°C
CESA5_OffT-5R	GAGGACGTTGACAGTGGAGA	

Rhizoid Assay

The rhizoid assay was carried out as described previously (Roberts et al. 2011). Wild type *P. patens* and *ppcsld1/7* KO lines were first cultured on BCDAT for 7 days, then approximately 0.5 cm² of 7-day-old chloronemal tissues were spotted on BCD medium (Roberts et al. 2011) and BCD medium with 1 µm l⁻¹ naphthaleneacetic acid, a synthetic plant hormone in the auxin family (Sakakibara et al. 2003), for two weeks. Photographs were taken for wild type and KO lines to document rhizoid growth after two weeks.

Cellulose Assay

After completing the rhizoid assay, for each individual colony, 3-4 rhizoids were isolated from wild type *P. patens* and *ppcsld1/7* KO lines, then treated with IMFL fixative (12.5 mM PIPES, 0.625 mM MgSO₄, 0.5 mM EGTA, 3.7% Formaldehyde, 1 µl Triton X-10 in water) containing 0.01 mg ml⁻¹ calcofluor. Stained samples were examined with fluorescence microscopy (UV excitation). Photographs were taken to document cellulose content of wild type and KO lines.

RESULTS

To facilitate direct analysis of PpCSLD activity, attempts were made to generate a viable genetic background line with all *PpCESAs* rendered non-functional. Since the combination of *PpCESAs* activities that could be eliminated while still retaining viability was unknown, a stepwise approach was implemented. A quadruple *ppcesa3/8/6/7* KO line was first transformed with the individual CESA4 KO, CESA5 KO, or CESA10 KO vectors (Figure 6). After PEG-mediated transformation and antibiotic selection, putative KO lines were genotyped by PCR genotype analysis, then confirmed by sequencing.

Figure 6. Flowchart of the strategy used to recover a viable genetic background non-functional for all seven *PpCESAs*. CESA4 KO, CESA5 KO, and CESA10 KO vectors were used to knockout the respective genes.

Confirmation of *ppcesa3/8/6/7/10* KO lines

Following transformation of *ppcesa3/8/6/7* KO with the CESA10 KO vector, 16 lines selected for transient antibiotic resistance were screened by PCR with the appropriate genotyping primers listed in Table 2 (Figure 7A). PCR products from 8

lines identified with large gene deletions were sequenced. Three mutants, lines 3, 6, and 10, had frameshifts (Figure S3, 4, 7), and no mutations were found in two off-target sites (Figure S9, 10). Line 10 was chosen from the three confirmed *ppcesa3/8/6/7/10* KO mutants to knockout *PpCESA4* as the next targeted gene.

Confirmation of *ppcesa3/8/6/7/10/4* KO lines

Following two transformations of *ppcesa3/8/6/7/10* KO with the *CESA4* KO vector, 96 lines selected for transient antibiotic resistance were screened by PCR with the appropriate genotyping primers listed in Table 2 (Figure 7B, C). A new phenotype was observed in all large gene deletions mutants. All of these mutants showed abnormal gametophore morphology. The gametophore leaves of mutants were shorter and rounded and the entire structure was smaller (Figure 8). Besides these mutants, about 80% of the transient antibiotic resistant lines without apparent large gene deletions also showed defects in gametophore. To examine the relationship between phenotype and gene disruption, PCR products from lines showing small PCR products and abnormal gametophores, wild type-like PCR products and abnormal gametophores, and wild type-like PCR products and normal gametophores were sequenced (Figure 9). These sequencing results confirmed that the abnormal gametophore phenotype is consistently associated with *PpCESA4* disruption. All lines showing the new phenotype had been gene edited with introduction of a frameshift or a premature stop codon. In all, eight mutants were identified with frameshifts (Figure S11-15, S18, 19, 21), and four mutants with large gene deletions. Lines 6, 28, 37, and 41, were used for off-target screening. No off-target edits were found in off-target site 1 and 3 (Figure S24, 26). But in off-target 2 screening, mutations were found in lines 28 (Figure S25). Since off-target 2 targets *PpCESA10*, which was knocked out previously, these mutants were still useful

for downstream work. Four *ppcesa3/8/6/7/10/4* KO mutants, lines 6, 28, 37, and 41, were generated from this transformation, lines 28 and 41 were used for following transformation. So far, about 70% of the transient antibiotic resistant colonies were identified with no gametophores in transformation of *ppcesa3/8/6/7/10/4* KO with the CESA5 KO vector.

Confirmation of *ppcesa3/8/6/7/5* KO lines

Following transformation of *ppcesa3/8/6/7* KO with the CESA5 KO vector, 40 lines selected for transient antibiotic resistance were first pre-screened for the *ppcesa5* KO phenotype, which is characterized by lack of gametophores (Goss et al. 2012). Then the resulting 15 lines were screened by PCR with the appropriate genotyping primers listed in Table 2 (Figure 10A). PCR products from 4 lines identified with large gene deletions were sequenced. Sequencing results analyzed by CodonCode Aligner show that frameshifts were introduced in lines 10, 19, and 31 (Figure S27, 28, 30). No off-targets edits were found in these lines (Figure S31-35). Three *ppcesa3/8/6/7/5* KO mutants were generated from this transformation, and line 10 was used for downstream transformation.

Confirmation of *ppcesa3/8/6/7/5/10* KO lines

Following transformation of *ppcesa3/8/6/7/5* KO with the CESA10 KO vector, 48 lines selected for transient antibiotic resistance were screened by PCR with the appropriate genotyping primers (Table 2). For this knockout, about 50% of the colonies were identified with large gene deletions (Figure 10B). PCR products from 8 mutants, which are selected based on different sizes of PCR products, were sequenced. Six mutants, lines 5, 6, 9, 10, 11, and 30, were identified with frameshifts (Figure S36, 37,

39, 40, 41, 44) and none of them had mutations in off target sites (Figure S45, 46). Line 6 was selected from six resulting *ppcesa3/8/6/7/5/10* mutants for the last gene knockout using *PpCESA4* as the next targeted gene.

Confirmation of *ppcsld1* KO and *ppcsld7* KO lines

To investigate the functional roles of CSLDs, wild-type *P. patens* (GD11) was used for transformation. To generate *ppcsld1/7* double knockouts, protoplasts from GD 11 were transformed with the CSLD1 KO vector. 32 lines selected for transient antibiotic resistance were screened by PCR with the appropriate genotyping primers listed in Table 2 (Figure 11A). PCR products from four mutants identified with gene deletions were sequenced. Two of them were identified with frameshifts (Figure S47, 48), and no off-targets edits were found in these mutants (Figure S51, 52). Two *ppcsld1* KO mutants, lines 4 and 13, were generated in this transformation, and line 4 was chosen to make the double knockout through transformation with the CSLD7 KO vector.

Two individual transformations were conducted to knock out *PpCSLD7* in the *ppcsld1* KO line. Following transformation with the CSLD7 KO vector, 96 lines selected for transient antibiotic resistance were screened by PCR with the appropriate genotyping primers listed in Table 2 (Figure 11B, C). PCR products from four lines were identified with gene deletions, and only line 10 from transformation 2 met the criteria for off-target screening (Figure S60). In this case, additional PCR products from six lines with no apparent large gene deletions were sequenced and all of them were identified with small deletions near the target site (Figure S55-59, 61). Four lines identified with frameshifts (Figure S55, 57, 59, 61) together with line 10 (Figure S60) were used for off-target screening and no mutations were found in two off-target sites

(Figure S63, 64). The resulting five *ppcsld1/7* KO lines, lines 3, 5, 9, 10, and 15, were used for the rhizoid assay and cellulose assay.

Confirmation of *ppcsld3* KO and *ppcsld4* KO lines

To generate *ppcsld3/4* KO mutants, single *ppcsld3* KO mutants were first made in wild type *P. patens*. Following transformation with the CSLD3 KO vector, 41 lines selected for antibiotic resistance were screened by PCR with the appropriate genotyping primers listed in Table 2 (Figure 12A). PCR products from eight mutants identified with large gene deletions were sequenced. Five mutants, lines 4, 5, 20, 28, and 31, were identified with frameshifts (Figure S65, 66, 69, 70, 71), and none of them had mutations in one off-target site (Figure S72). Five *ppcsld3* KO mutants were generated and line 20 was chosen for subsequent transformation with the CSLD4 KO vector.

Following transformation of *ppcsld3* KO line 20 with the CSLD4 KO vector, 48 lines selected for transient antibiotic resistance were screened by PCR with the appropriate genotyping primers listed in Table 2 (Figure 12B). PCR products from seven mutants identified with large gene deletions were sequenced. In off-target screening, mutations were found in two off-target sites (Figure S78, 79). These mutations were located upstream of the catalytic domains of *PpCSLD5* and *PpCSLD8* (Figure 13).

Rhizoid & Cellulose Assay

After PCR genotype analysis, all confirmed *ppcsld1/7* KO lines were used for phenotype analysis which included a rhizoid growth assay and a fluorescent cytochemical cellulose assay. To identify the differences in rhizoid growth between wild type *P. patens* and *ppcsld1/7* KO mutants, lines 3, 5, 9, 10, and 15 of *ppcsld1/7*

KO and wild type *P. patens* were grown on BCD medium with 1 μM 1^{-1} naphthaleneacetic acid for two weeks. No obvious difference was observed between line 3 (Figure 14b, e), line 5 (Figure 14c, f), line 9 (Figure 14g, j), line 10 (Figure 14h, k), line 15 (Figure 14i, l) of *ppcsld1/7* KO and wild type *P. patens* (Figure 14a, d) in rhizoid growth.

As a result, a fluorescent cytochemical cellulose assay was performed to identify differences in cellulose content between *ppcsld1/7* KO mutants and wild type *P. patens*. Isolated rhizoids from wild type and mutants were stained with calcofluor and examined with fluorescence microscopy. No differences were observed when wild type *P. patens* (Figure 15a) was compared to mutants (Figure 15b-f).

Figure 7. PCR genotype analysis of transient antibiotic resistant colonies resulting from transformation of *ppcesa3/8/6/7* KO mutant with CESA10 KO (A) and CESA4 KO (B, C) vectors. (A) PCR products from amplification of DNA extracted from 11 independent lines with primers CESA10_Crdel-F and CESA10_Crdel-R. (B) PCR products from amplification of DNA extracted from 48 independent lines from transformation 1 with primers CESA4_Crdel-F2 and CESA4_Crdel-R2. (C) PCR products from amplification of DNA extracted from 48 independent lines from transformation 2 with primers CESA4_Crdel-F2 and CESA4_Crdel-R2.

Figure 8. The phenotype of *ppcesa3/8/6/7/10/4* KO mutants (a, b, c, e, f) and wild-type *P. patens* (d). *ppcesa3/8/6/7/10/4* KO lines cultured from protonemal tissue on BCDAT medium, showing defect leafy gametophores. Bars=2 mm (a-c), bars=0.5 mm (d-f).

Figure 9. A schematic showing on target edits between two protospacers. *ppcesa3/8/6/7/10/4* KO mutants were made from the transformation of *ppcesa3/8/6/7/10* KO with the CESA4 KO vector, gene deletions (red) or insertions (green) were found between two target sites. Line 6 from transformation 1 was not included due to the complexity of the DNA rearrangement.

Figure 10. PCR genotype analysis of transient antibiotic resistant colonies resulting from transformation of the *ppcesa3/8/6/7* KO mutant with PpCESA5 (A) and PpCESA10 (B) KO vectors. (A) PCR products from amplification of DNA extracted from 15 independent lines with primers CESA5_Crdel-F and CESA5_Crdel-R. (B) PCR products from amplification of DNA extracted from 48 independent lines with primers CESA10_Crdel-F2 and CESA10_Crdel-R2.

Figure 11. PCR genotype analysis of transient antibiotic resistant colonies resulting from transformation of wild type *P. Patens* (GD 11) with PpCSLD1 (A) and PpCSLD7 (B, C) KO vectors. (A) PCR products from amplification of DNA extracted from 32 independent lines with primers CSLD1_Crdel-F and CSLD1_Crdel-R. (B) PCR products from amplification of DNA extracted from 48 independent lines from transformation 1 with primers CSLD7_Crdel-F2 and CSLD7_Crdel-R2. (C) PCR products from amplification of DNA extracted from 48 independent lines from transformation 2 with primers CSLD7_Crdel-F2 and CSLD7_Crdel-R2.

Figure 12. PCR genotype analysis of transient antibiotic resistant colonies resulting from transformation of wild type *P. patens* (GD 11) with PpCSLD3 (A) and PpCSLD4 (B) KO vectors. (A) PCR products from amplification of DNA extracted from 41 independent lines with primers CSLD3_Crdel-F and CSLD3_Crdel-R. (B) PCR products from amplification of DNA extracted from 48 independent lines with primers CSLD4_Crdel-F and CSLD4_Crdel-R.

Figure 13. Amino acid sequence of CSLD4 off-targets region. UDPG-Binding Domain is highlighted with red, Plant Conserved Region (CR-P) is highlighted with green, and mutation is highlighted with underline

Figure 14. Dark field images of wild type and *ppclsd1/7* mutants leafless gametophore (arrows) with rhizoids (arrowheads). Wild type (a, d) and *ppclsd1/7* KO lines 3 (b, e), 5 (c, f), 9 (g, j), 10 (h, k), 15 (i, l), grown on BCD/auxin ($1 \mu\text{m l}^{-1}$) medium for two weeks at 4X magnification or 7X magnification.

Figure 15. Images of fluorescence microscopy. Isolated rhizoids from wild type (a), *ppcsl1/7* KO lines 3 (b), 5 (c), 9 (d), 10 (e), 15 (f), stained with calcofluor.

DISCUSSION

Efficiency of CRISPR/Cas 9 gene knockout in *P. patens*

In this study, the CRISPR/Cas9 system was used to knockout genes of *P. patens* and investigate the possible roles of CSLDs in plant growth. To achieve this goal, two different strategies were applied. The first strategy was to generate an all CESA knockout mutant line (7 genes), which could be used for functional studies of CSLDs in a CESA null background. Also, two *PpCSLD* double knockout mutant lines were generated from wild type to study the developmental functions of CSLDs. The basic approach used was to knock out targeted genes using a dual Cas9-sgRNA expression vector system, which can generate large deletions between two protospacers.

For the two CSLD double knockouts (1/7 and 3/4), attempts were made to transform wild-type *P. patens* with expression vectors targeting two different genes simultaneously and select with two different antibiotics. Compared to the results from transformations using one expression vector, fewer putative KO lines were generated after a week on antibiotic selection. Only a few mutants showed large gene deletions and then only for one targeted gene as shown by PCR genotype analysis. No mutants showed large gene deletions on both genes. Based on this, subsequent experiments for producing CESA null genetic background mutants and *PpCSLD* double KO mutants were designed to knock out individual genes sequentially. For each gene knockout, mutants with large gene deletions were first sequenced, and mutants with frameshift or nonsense mutations were used for off-target screening to verify a desired mutant genotype. Then a mutant with a phenotype similar to the transformed line or consistent among all knockout mutants was chosen to knockout additional genes in a step wise manner.

While transformation with two different expression vectors was unsuccessfully for large gene deletions on both targeted genes in this study, several lines of evidence indicate double knockout lines can be generated by co-transformation with two vectors. For one thing, transformations with two different expression vectors were attempted before I became proficient in moss transformation. In the case of the *ppcsld7* KO screening, one mutant was identified with large gene deletions and frameshifts among 96 colonies screened. Further sequencing work was done on PCR products from lines with no apparent large gene deletions, which showed that most of these lines had been successfully edited. This was also found in *ppcesa4* KO screening (Figure 9). In addition, the off-target screening of lines resulting from transformation of *ppcesa3/8/6/7/10* with the CESA4 KO vector had a high frequency of mutations in *PpCESA10*. All these observations indicate the high efficiency of CRISPR/Cas9 knockout system and that double knockouts may have been generated in initial transformations using two different expression vectors.

Generating a CESA null genetic background in *P. patens*

First round transformations

Since it was impossible to predict which combinations of *PpCESAs* could be knocked out while maintaining viability, the remaining three genes, *PpCESA4*, *PpCESA5*, *PpCESA10*, were knocked out in an existing *ppcesa3/8/6/7* KO line in different sequential orders in an attempt to generate a viable *P. patens* genetic background line with no functional CESAs (Figure 6). Viable knockout lines were generated from all first round transformations. New phenotypes were observed when *ppcesa3/8/6/7* KO was transformed with CESA5 KO and CESA4 KO, but not with the CESA10 KO vector. Following transformation of *ppcesa3/8/6/7* KO with the CESA10

KO vector, *ppcesa3/8/6/7/10* KO mutants were generated. The resulting mutants showed good protonemal growth and normal gametophores similar to *ppcesa3/8/6/7* KO. This is consistent with previous results showing *PpCESA4* and *PpCESA10* are redundant (Li 2017).

PpCESA5 is expressed in gametophores, and *ppcesa5* KO mutants cannot form leafy gametophores (Goss et al. 2012). Following transformation of *ppcesa3/8/6/7* KO with the CESA5 KO vector, about 40% of the resulting colonies lacked gametophores, as expected, and had no other obvious growth defects.

Following transformation of *ppcesa3/8/6/7* KO with the CESA4 KO vector, the resulting colonies showed good protonemal growth, and a new abnormal gametophore phenotype was observed. This is surprising since it was shown previously with *PpCESA4* and *PpCESA10* are functionally redundant (Li 2017). Colonies selected for transient antibiotic resistance from this transformation were lost due to contamination. However, this transformation was repeated later when a similar phenotype was observed in *ppcesa3/8/6/7/10/4* KO mutants (see below).

Second round transformations

In the second round of transformations, some *ppcesa3/8/6/7/10/4* KO mutants made by using the CESA4 KO vector were confirmed base on large deletions and frameshifts in *PpCESA4*, and an unexpected phenotype was observed. Compared to wild type *P. patens*, gametophore leaves of these mutants are shorter and rounded, and the entire structure is smaller (Figure 8). The leaf cells were rounded instead of elongated, which is consistent with cellulose deficiency phenotypes observed in other plants (McFarlane et al. 2014). The same phenotype was also observed in lines for which PCR indicated no large deletions in *PpCESA4*, and some of these PCR products

from these lines were sequenced later. All sequencing results indicated that these lines had been gene edited at the target site with small deletions (Figure 9), and frameshift or nonsense mutations were introduced to the ORF of *PpCESA4* (Figure S11-15). Furthermore, mutant line 13 from transformation 1, with a normal *ppcesa3/8/6/7/10* phenotype, was sequenced to determine the genotype. A small deletion was identified (Figure 9), but without frameshifts (Figure S17). In addition, *ppcesa3/8/6/7/10* KO used for this transformation has normal gametophores similar to *ppcesa3/8/6/7* KO. All evidence indicates the new phenotype results from loss of *PpCESA4* activity. Since mutants have only *PpCESA5*, functional loss of six *PpCESAs* may lead to the strong loss of cellulose synthesis and result in this phenotype.

Interestingly, the *ppcesa3/8/6/7/10/4* phenotype is similar to a phenotype observed in a transformation of *ppcesa3/8/6/7* with the CESA4 KO vector. As stated previously, there are three off-target sites for the CESA4 KO vector, and one of them is in *PpCESA10*. It is also known that *PpCESA4* and *PpCESA10* function redundantly in protonemal growth (Li 2017). This indicates that the phenotype observed in this transformation might be caused by disrupting two genes. In fact, half of *ppcesa3/8/6/7/10/4* KO mutants were identified with off-target edits in *PpCESA10*, indicating that the CESA4 KO vector can cause mutations in *PpCESA10* with a high frequency. Combining this information, the new phenotype could be caused by the disruption of *PpCESA4* or *PpCESA4* and *PpCESA10* together. Transformation of *ppcesa3/8/6/7* KO with the CESA4 KO vector was repeated. Genotyping analysis of colonies selected for transient antibiotic resistance from this transformation, especially off-target screening, could answer the question that whether *PpCESA4* or *PpCESA4* and *PpCESA10* together result in this phenotype.

One of the *ppcesa3/8/6/7/5* KO mutants from first round transformations was used to knock out a sixth gene. Since the CESA4 KO vector was known to cause off-target edits in *PpCESA10*, the CESA10 KO vector was used in second round transformations. Following transformation of *ppcesa3/8/6/7/5* KO with the CESA10 KO vector, six *ppcesa3/8/6/7/5/10* mutants were identified with large gene deletions and frameshifts in *PpCESA10* and all of them showed a consistent phenotype which is lack of gametophores.

Third round transformations

Finally, a sextuple *ppcesa3/8/6/7/10/4* KO was used to knock out the last gene in the third round transformations. Colonies selected for transient antibiotic resistance were pre-screened for the *ppcesa5* KO phenotype, and a large number of colonies were identified with no gametophores, presumably due to *PpCESA5* disruption. These pre-screened colonies are currently being evaluated. Based on these preliminary results, it seems possible that all seven CESAs can be knocked out in wild type *P. patens* and the resultant genotype remains viable.

For another sextuple *ppcesa3/8/6/7/5/10* KO, the CESA4 KO vector was used to create an all *PpCESA* knockout. Following transformation of *ppcesa3/8/6/7/5/10* KO with the CESA4 KO vector, many colonies selected for transient antibiotic resistance were generated without obvious growth defects. Genotyping work needs to be completed with this transformation to verify mutant genotypes. Based on the successful *PpCESAs* KO in different lines, it seems that all CESAs can be knocked out, and genetic background can be generated without compromising viability.

CSLD double knockout

Based on an analysis of *PpCSLD* chromosomal location and synteny (Lang et al. 2018), *PpCSLD1* and *PpCSLD7* are a tandem pair and *PpCSLD3* and *PpCSLD4* are an inverted pair, and both pairs arose from a single ancestral gene as a result of a whole genome duplication. Gene expression data summarized in *P. patens* eFP browser (http://bar.utoronto.ca/efp_physcomitrella/cgi-bin/efpWeb.cgi) showed that the two members of these pairs are likely to have redundant functions in rhizoids and protonemal development, respectively. Therefore, rhizoid growth and protonemal tip growth are expected to be inhibited if the CSLDs are involved in tip growth and not redundant with other *PpCSLDs*. CSLD1, 3, 4, and 7 KO vectors were used to produce *ppcsld1/7* KO, *ppcsld3/4* KO mutants.

Genotyping analysis by PCR and sequencing confirmed the successful generation of 5 *ppcsld1/7* KO lines (Figure 11, Figure S55, 57, 59-61, 63, 64). Rhizoid growth and cellulose assays were performed to test the hypothesis that *PpCSLD1* and *PpCSLD7* are critical for the development of rhizoids or cellulose synthesis in rhizoids. However, no obvious differences were observed when wild type *P. patens* and *ppcsld1/7* KO mutants were compared in both assays (Figure 11, 12). The explanation for the rhizoid assay results may be functional redundancy. Other genes in the *PpCSLD* family may compensate for the loss of *PpCSLD1* and *PpCSLD7*. Thus other *PpCSLDs* need to be knocked out to identify their function in rhizoids. The results of the cellulose assay may also be explained by the presence of background CESAs activity. The CESA null genetic background line could be a good material for cellulose synthesis analysis of CSLDs.

Genotype analysis by PCR and sequencing indicate all of the *ppcsld3/4* KO lines that were generated had off-target mutations (Figure S78, 79). Tip growth assay on *ppcsld3/4* KO mutants were planned to test the hypothesis that *PpCSLD3* and

PpCSLD4 are essential for protonemal tip growth, but the relationship between genotype and phenotype could not be analyzed in these mutants due to the presence of off-target mutations. Mutations were found in off-target site 2 which is within *PpCSLD8* and off-target site 3 which is within *PpCSLD5*, and the mutations were located upstream of the catalytic domain of each gene (Figure 11). For most mutants, a few amino acids were inserted or deleted in two off-target sites, and no frameshifts were introduced (Figure 11). The effects of these mutations on *PpCSLD5* and *PpCSLD8* were unclear. Therefore, the existing lines are inconclusive. To test this hypothesis, the *CSLD4* KO vector could be redesigned, avoiding off-target edits in *PpCSLD5* and *PpCSLD8*.

CONCLUSION.

The production of verified sextuple *ppcesa3/8/6/7/10/4* KO lines, along with partially characterized septuple *ppcesa3/8/6/7/10/4/5* KO lines that lack gametophores but show vigorous protonemal growth, suggest that *PpCESAs* play a minor role in protonemal development and can be knocked out without compromising viability. The *ppcesa3/8/6/7/10* KO lines showed a similar phenotype to *ppcesa3/8/6/7* KO lines, whereas *ppcesa3/8/6/7/10/4* KO lines had defective gametophores. This observation supports our previous results that *PpCESA4* and *PpCESA10* are functionally redundant with *ppcesa4/10* KOs showing a slight reduction in protonemal growth that is not evident in *ppcesa4* KO or *ppcesa10* KO lines (Li 2017). Previous results showed that global silencing of the entire *PpCSLD* gene family by RNAi resulted in inhibition of protonemal tip growth (Dimos 2010). Results reported here showed that *ppcsld1/7* double knockouts have no defects in rhizoid growth. Results from knocking out *PpCSLD3* and *PpCSLD4* were inconclusive due to off-target editing. Thus the individual developmental roles of the *PpCSLD* family members remain unresolved.

APPENDIX 1: Alignment of knockout sequencing results

Figure S1. CodonCode Aligner display of sequencing result for *ppcsa10* KO line 1 generated from *ppcsa3/8/6/7* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

Figure S2. CodonCode Aligner display of sequencing result for *ppcsa10* KO line 2 generated from *ppcsa3/8/6/7* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

Figure S3. CodonCode Aligner display of sequencing result for *ppcesa10* KO line 3 generated from *ppcesa3/8/6/7* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 95-97.

Figure S4. CodonCode Aligner display of sequencing result for *ppcesa10* KO line 6 generated from *ppcesa3/8/6/7* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 113-115.

Figure S5. CodonCode Aligner display of sequencing result for *ppcesa10* KO line 7 generated from *ppcesa3/8/6/7* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

51

Figure S9. CodonCode Aligner display comparing sequencing results for off-target sites 1 in *ppcsa10* KO lines 3 (KY48I_191001), 6 (KY48J_191001), and 10 (KY48K_191001) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 1 in lines 3, 6, and 10.

52 Figure S10. CodonCode Aligner display comparing sequencing results for off-target sites 2 in *ppcsa10* KO lines 3 (KY48R_191008), 6 (KY48S_191008), and 10 (KY48T_191008) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 2 in lines 3, 6, and 10.

Figure S11. CodonCode Aligner display of sequencing result for *ppcesa4* KO line 1 from transformation 1 generated from *ppcesa3/8/6/7/10* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon at position 170-172.

Figure S12. CodonCode Aligner display of sequencing result for *ppcesa4* KO line 2 from transformation 1 generated from *ppcesa3/8/6/7/10* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon at position 173-175.

Figure S13. CodonCode Aligner display of sequencing result for *ppcesa4* KO line 6 from transformation 1 generated from *ppcesa3/8/6/7/10* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon at position 238-240.

Figure S14. CodonCode Aligner display of sequencing result for *ppcesa4* KO line 8 from transformation 1 generated from *ppcesa3/8/6/7/10* KO. Edited sites are highlighted in red. The first row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon at position 235-237.

Figure S15. CodonCode Aligner display of sequencing result for *ppcsa4* KO line 9 from transformation 1 generated from *ppcsa3/8/6/7/10* KO. Edited sites are highlighted in red. The first row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon.

58

Figure S16. CodonCode Aligner display of sequencing result for *ppcesa4* KO line 10 from transformation 1 generated from *ppcesa3/8/6/7/10* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

Figure S17. CodonCode Aligner display of sequencing result for *ppcesa4* KO line 13 from transformation 1 generated from *ppcesa3/8/6/7/10* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. No frameshift was introduced at the first edited site.

Figure S18. CodonCode Aligner display of sequencing result for *ppcsa4* KO line 28 from transformation 2 generated from *ppcsa3/8/6/7/10* KO. Edited sites are highlighted in red. The first row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 148-150.

Figure S19. CodonCode Aligner display of sequencing result for *ppcesa4* KO line 37 from transformation 2 generated from *ppcesa3/8/6/7/10* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 224-226.

Figure S20. CodonCode Aligner display of sequencing result for *ppcesa4* KO line 40 from transformation 2 generated from *ppcesa3*/8/6/7/10 KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

Figure S21. CodonCode Aligner display of sequencing result for *ppcsa4* KO line 41 from transformation 2 generated from *ppcsa3/8/6/7/10* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon at position 237-239.

Figure S22. CodonCode Aligner display of sequencing result for *ppcesa4* KO line 43 from transformation 2 generated from *ppcesa3/8/6/7/10* KO. Edited sites are highlighted in red. The first row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

Figure S23. CodonCode Aligner display of sequencing result for *ppcesa4* KO line 45 from transformation 2 generated from *ppcesa3/8/6/7/10* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. No frameshift was introduced at the first edited site.

Figure S24. CodonCode Aligner display comparing sequencing results for off-target sites 1 in *ppcsa4* KO lines 6 (KY50K_200107A) and 28 (KY50L_200107A), 37 (KY50M_200107A), 41 (KY50O_200107A) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 1 in lines 6, 28, 37 and 41.

67 Figure S26. CodonCode Aligner display comparing sequencing results for off-target sites 3 in *ppcesa4* KO lines 6 (KY50S_200107A) and 28 (KY50T_200107A), 37 (KY50U_200107A), 41 (KY50W_200107A) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 3 in lines 6, 28, 37 and 41.

Figure S27. CodonCode Aligner display of sequencing result for *ppcesa5* KO line 10 generated from *ppcesa3/8/6/7* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon.

Figure S28. CodonCode Aligner display of sequencing result for *ppcesa5* KO line 19 generated from *ppcesa3/8/6/7* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon.

69

Figure S30. CodonCode Aligner display of sequencing result for *ppcsa5* KO line 31 generated from *ppcsa3/8/6/7* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon.

Figure S31. CodonCode Aligner display comparing sequencing results for off-target sites 1 in *ppcesa5* KO lines 10 (KY47E_190924A), 19 (KY47F_190924A), and 31 (KY47G_190924A) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 1 in lines 10, 19, and 31.

71 Figure S32. CodonCode Aligner display comparing sequencing results for off-target sites 2 in *ppcsa5* KO lines 10 (KY47H_190924A), 19 (KY47I_190924A), and 31 (KY47J_190924A) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 2 in lines 10, 19, and 31.

Figure S33. CodonCode Aligner display comparing sequencing results for off-target sites 3 in *ppcesa5* KO lines 10 (KY47K_190924A), 19 (KY47L_190924A), and 31 (KY47M_190924A) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 3 in lines 10, 19, and 31.

Figure S34. CodonCode Aligner display comparing sequencing results for off-target sites 4 in *ppcsa5* KO lines 10 (KY47N_190924A), 19 (KY47O_190924A), and 31 (KY47P_190924A) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 4 in lines 10, 19, and 31.

Figure S35. CodonCode Aligner display comparing sequencing results for off-target sites 5 in *ppcsa5* KO lines 10 (KY47Q_190924A), 19 (KY47R_190924A), and 31 (KY47S_190924A) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 5 in lines 10, 19, and 31.

75

Figure S36. CodonCode Aligner display of sequencing result for *ppcesa10* KO line 5 generated from *ppcesa3/8/6/7/5* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 170-172.

76

Figure S37. CodonCode Aligner display of sequencing result for *ppcesa10* KO line 6 generated from *ppcesa3/8/6/7/5* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon at position 276-278.

77

Figure S38. CodonCode Aligner display of sequencing result for *ppcesa10* KO line 8 generated from *ppcesa3/8/6/7/5* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

79

Figure S40. CodonCode Aligner display of sequencing result for *ppcesa10* KO line 10 generated from *ppcesa3/8/6/7/5* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 243-245.

80

Figure S41. CodonCode Aligner display of sequencing result for *ppcesa10* KO line 11 generated from *ppcesa3/8/6/7/5* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon at position 234-236.

Figure S42. CodonCode Aligner display of sequencing result for *ppcesa10* KO line 14 generated from *ppcesa3/8/6/7/5* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon.

Figure S43. CodonCode Aligner display of sequencing result for *ppcesa10* KO line 14 generated from *ppcesa3/8/6/7/5* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

83

Figure S44. CodonCode Aligner display of sequencing result for *ppcesa10* KO line 30 generated from *ppcesa3/8/6/7/5* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the first edited site resulting in a premature stop codon at position 168-170.

Figure S45. CodonCode Aligner display comparing sequencing results for off-target sites 1 in *ppcsa10* KO lines 5 (KY55K_200121), 6 (KY55L_200121), 9 (KY55M_200121), 10 (KY55N_200121), 11 (KY55O_200121), and 30 (KY55P_200121) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 1 in lines 5, 6, 9, 10, 11, and 30.

Figure S46. CodonCode Aligner display comparing sequencing results for off-target sites 2 in *ppcsa10* KO lines 5 (KY55Q_200121), 6 (KY51R_200121), 9 (KY51S1_200123), 10 (KY55T_200121), 11 (KY51U1_200123), and 30 (KY51V1_200123) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 2 in lines 5, 6, 9, 10, 11, and 30.

Figure S47. CodonCode Aligner display of sequencing result for *ppcsld1* KO line 4 generated from wild type *P. patens*. Edited sites are highlighted in red. The first row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 586-588.

87 Figure S48. CodonCode Aligner display of sequencing result for *ppcsld1* KO line 13 generated from wild type *P. patens*. Edited sites are highlighted in red. The first row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site.

Figure S49. CodonCode Aligner display of sequencing result for *ppcsld1* KO line 14 generated from wild type *P. patens*. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

Figure S50. CodonCode Aligner display of sequencing result for *ppcsld1* KO line 16 generated from wild type *P. patens*. Edited sites are highlighted in red. The first row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

Figure S51. CodonCode Aligner display comparing sequencing results for off-target sites 1 in *ppcsld1* KO lines 4 (KY45H_190917B) and 13 (KY45J_1909268) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 1 in line 4 and 13.

Figure S52. CodonCode Aligner display comparing sequencing results for off-target sites 2 in *ppcsld1* KO lines 4 (KY45I_190926B) and 13 (KY45G_190917B) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 2 in line 4 and 13.

Figure S53. CodonCode Aligner display of sequencing result for *ppcsld7* KO line 4 from transformation 1 generated from *ppcsld1* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

Figure S54. CodonCode Aligner display of sequencing result for *ppsld7* KO line 24 from transformation 1 generated from *ppsld1* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

Figure S57. CodonCode Aligner display of sequencing result for *ppcsld7* KO line 5 from transformation 2 generated from *ppcsld1* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 990-992.

Figure S58. CodonCode Aligner display of sequencing result for *ppcsld7* KO line 8 from transformation 2 generated from *ppcsld1* KO. Edited sites are highlighted in red. The first row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

Figure S59. CodonCode Aligner display of sequencing result for *ppclsld7* KO line 9 from transformation 2 generated from *ppclsld1* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 965-967.

Figure S60. CodonCode Aligner display of sequencing result for *ppcsld7* KO line 10 from transformation 2 generated from *ppcsld1* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 317-319.

Figure S61. CodonCode Aligner display of sequencing result for *ppcsld7* KO line 15 from transformation 2 generated from *ppcsld1* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 948-950.

100

Figure S63. CodonCode Aligner display comparing sequencing results for off-target sites 1 in *ppcsld7* KO lines 3 (KY56A1_200204), 5 (KY56B1_200204), 9 (KY56C1_200204), 10 (KY56D1_200204), and 15 (KY56E1_200204) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 1 in lines 3, 5, 9, 10, and 15.

101

Figure S64. CodonCode Aligner display comparing sequencing results for off-target sites 2 in *ppcsld7* KO lines 3 (KY56F_200128), 5 (KY56G_200128), 9 (KY56H_200128), 10 (KY56I_200128), and 15 (KY56J_200128) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 2 in lines 3, 5, 9, 10, and 15.

Figure S65. CodonCode Aligner display of sequencing result for *ppcsld3* KO line 4 generated from wild type *P. patens*. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 209-211.

Figure S66. CodonCode Aligner display of sequencing result for *ppcsld3* KO line 5 generated from wild type *P. patens*. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 206-208.

Figure S70. CodonCode Aligner display of sequencing result for *ppcslD3* KO line 28 generated from wild type *P. patens*. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 125-127.

Figure S72. CodonCode Aligner display comparing sequencing results for off-target sites 1 in *ppcsld3* KO lines 4 (KY46H_190917B), 5 (KY46I_190919), 20 (KY46J_190919), 28 (KY46K_190919), 31 (KY46L_190919) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 1 in lines 4, 5, 20, 28, and 31.

Figure S73. CodonCode Aligner display of sequencing result for *ppscld4* KO line 4 generated from *ppscld3* KO. Edited sites are highlighted in red. The second row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 383-385.

Figure S74. CodonCode Aligner display of sequencing result for *ppsld4* KO line 10 generated from *ppsld3* KO. Edited sites are highlighted in red. The first row of the translation is in-frame with the start codon. No frameshift was introduced at the second edited site.

Figure S75. CodonCode Aligner display of sequencing result for *ppsld4* KO line 21 generated from *ppsld3* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 210-212.

Figure S76. CodonCode Aligner display of sequencing result for *ppcsld4* KO line 31 generated from *ppcsld3* KO. Edited sites are highlighted in red. The third row of the translation is in-frame with the start codon. A frameshift was introduced at the second edited site resulting in a premature stop codon at position 942-944.

112 Figure S77. CodonCode Aligner display comparing sequencing results for off-target sites 1 in *ppcsld4* KO lines 4 (KY57D_200211A) and 13 (KY57E_200211A), 21 (KY57G_200211A), 31 (KY57H_200211A) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. No mutations were found in off-target site 1 in lines 4, 13, 21, and 31.

113 Figure S78. CodonCode Aligner display comparing sequencing results for off-target sites 2 in *ppcsld4* KO lines 4 (KY57I_200211A) and 13 (KY57J_200211A), 21 (KY57L_200211A), 31 (KY57M_200211A) with off-target site in the *P. patens* genome. The off-target site is highlighted in red. Mutations were found in off-target site 2 in lines 4, 13, 21, and 31.

BIBLIOGRAPHY

- Bernal AJ, Yoo CM, Mutwil M, Jensen JK, Hou G, Blaukopf C, Sorensen I, Blancaflor EB, Scheller HV, Willats WG (2008) Functional analysis of the cellulose synthase-like genes CSLD1, CSLD2, and CSLD4 in tip-growing *Arabidopsis* cells. *Plant Physiol* 148 (3):1238-1253.
doi:10.1104/pp.108.121939
- Berry EA, Tran ML, Dimos CS, Budziszcz MJ, Jr., Scavuzzo-Duggan TR, Roberts AW (2016) Immuno and affinity cytochemical analysis of cell wall composition in the moss *Physcomitrella patens*. *Front Plant Sci* 7:248.
doi:10.3389/fpls.2016.00248
- Bhaya D, Davison M, Barrangou R (2011) CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet* 45:273-297. doi:10.1146/annurev-genet-110410-132430
- Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV, van der Oost J (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321 (5891):960-964.
doi:10.1126/science.1159689
- Coutinho PM, Deleury E, Davies GJ, Henrissat B (2003) An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol* 328 (2):307-317.
doi:10.1016/s0022-2836(03)00307-3
- Cove DJ (2000) The moss, *Physcomitrella patens*. *Journal of Plant Growth Regulation* 19(3):275-283
- Cove DJ, Perroud PF, Charron AJ, McDaniel SF, Khandelwal A, Quatrano RS (2009) The moss *Physcomitrella patens*: a novel model system for plant development and genomic studies. *Cold Spring Harb Protoc* 2009 (2):pdb emo115.
doi:10.1101/pdb.emo115
- Cutler S, Somerville C (1997) Cloning in silico. *Curr Biol* 7 (2):R108-111.
doi:10.1016/s0960-9822(06)00050-9

- Delmer DP (1999) Cellulose biosynthesis: exciting times for a difficult field of study. *Annu Rev Plant Physiol Plant Mol Biol* 50:245-276. doi:10.1146/annurev.arplant.50.1.245
- Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471 (7340):602-607. doi:10.1038/nature09886
- Desprez T, Juraniec M, Crowell EF, Jouy H, Pochylova Z, Parcy F, Hofte H, Gonneau M, Vernhettes S (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 104 (39):15572-15577. doi:10.1073/pnas.0706569104
- Dimos CS (2010) Functional analysis of the *cellulose synthase-like D (CSLD)* gene family in *Physcomitrella Patens*.
- Doblin MS, De Melis L, Newbigin E, Bacic A, Read SM (2001) Pollen tubes of *Nicotiana glauca* express two genes from different beta-glucan synthase families. *Plant Physiol* 125 (4):2040-2052. doi:10.1104/pp.125.4.2040
- Doblin MS, Pettolino F, Bacic A (2010) Plant cell walls: the skeleton of the plant world. *Functional Plant Biology* 37:357-381
- Doudna JA, Charpentier E (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346 (6213):1258096. doi:10.1126/science.1258096
- Dumais J, Long SR, Shaw SL (2004) The mechanics of surface expansion anisotropy in *Medicago truncatula* root hairs. *Plant Physiol* 136 (2):3266-3275. doi:10.1104/pp.104.043752
- Favery B, Ryan E, Foreman J, Linstead P, Boudonck K, Steer M, Shaw P, Dolan L (2001) KOJAK encodes a cellulose synthase-like protein required for root hair cell morphogenesis in *Arabidopsis*. *Genes Dev* 15 (1):79-89. doi:10.1101/gad.188801

- Gaj T, Gersbach CA, Barbas CF, 3rd (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* 31 (7):397-405. doi:10.1016/j.tibtech.2013.04.004
- Galway ME (2006) Root hair cell walls: filling in the framework. *Canadian Journal of Botany* 84:613-621
- Galway ME, Eng RC, Schiefelbein JW, Wasteneys GO (2011) Root hair-specific disruption of cellulose and xyloglucan in AtCSLD3 mutants, and factors affecting the post-rupture resumption of mutant root hair growth. *Planta* 233 (5):985-999. doi:10.1007/s00425-011-1355-6
- Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadan AH, Moineau S (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468 (7320):67-71. doi:10.1038/nature09523
- Gasiunas G, Sinkunas T, Siksnys V (2014) Molecular mechanisms of CRISPR-mediated microbial immunity. *Cell Mol Life Sci* 71 (3):449-465. doi:10.1007/s00018-013-1438-6
- Goss CA, Brockmann DJ, Bushoven JT, Roberts AW (2012) A *CELLULOSE SYNTHASE (CESA)* gene essential for gametophore morphogenesis in the moss *Physcomitrella patens*. *Planta* 235 (6):1355-1367. doi:10.1007/s00425-011-1579-5
- Haeussler M, Schonig K, Eckert H, Eschstruth A, Mianne J, Renaud JB, Schneider-Maunoury S, Shkumatava A, Teboul L, Kent J, Joly JS, Concordet JP (2016) Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* 17 (1):148. doi:10.1186/s13059-016-1012-2
- Hazen SP, Scott-Craig JS, Walton JD (2002) Cellulose synthase-like genes of rice. *Plant Physiol* 128 (2):336-340. doi:10.1104/pp.010875
- Hiss M, Laule O, Meskauskienė RM, Arif MA, Decker EL, Erxleben A, Frank W, Hanke ST, Lang D, Martin A, Neu C, Reski R, Richardt S, Schallenberg-Rudinger M, Szovenyi P, Tiko T, Wiedemann G, Wolf L, Zimmermann P, Rensing SA (2014) Large-scale gene expression profiling data for the model moss *Physcomitrella patens* aid understanding of developmental progression, culture and stress conditions. *Plant J* 79 (3):530-539. doi:10.1111/tpj.12572

- Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157 (6):1262-1278. doi:10.1016/j.cell.2014.05.010
- Jarvis MC (2018) Structure of native cellulose microfibrils, the starting point for nanocellulose manufacture. *Philos Trans A Math Phys Eng Sci* 376 (2112). doi:10.1098/rsta.2017.0045
- Joung JK, Sander JD (2013) TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol* 14 (1):49-55. doi:10.1038/nrm3486
- Kim CM, Park SH, Je BI, Park SH, Park SJ, Piao HL, Eun MY, Dolan L, Han CD (2007) *OsCSLD1*, a *cellulose synthase-like D1* gene, is required for root hair morphogenesis in rice. *Plant Physiol* 143 (3):1220-1230. doi:10.1104/pp.106.091546
- Kimura S, Laosinchai W, Itoh T, Cui X, Linder CR, Brown RM, Jr. (1999) Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant *vigna angularis*. *Plant Cell* 11 (11):2075-2086. doi:10.1105/tpc.11.11.2075
- Kurek I, Kawagoe Y, Jacob-Wilk D, Doblin M, Delmer D (2002) Dimerization of cotton fiber cellulose synthase catalytic subunits occurs via oxidation of the zinc-binding domains. *Proc Natl Acad Sci U S A* 99 (17):11109-11114. doi:10.1073/pnas.162077099
- Lang D, Ullrich KK, Murat F, Fuchs J, Jenkins J, Haas FB, Piednoel M, Gundlach H, Van Bel M, Meyberg R, Vives C, Morata J, Symeonidi A, Hiss M, Muchero W, Kamisugi Y, Saleh O, Blanc G, Decker EL, van Gessel N, Grimwood J, Hayes RD, Graham SW, Gunter LE, McDaniel SF, Hoernstein SNW, Larsson A, Li FW, Perroud PF, Phillips J, Ranjan P, Rokshar DS, Rothfels CJ, Schneider L, Shu S, Stevenson DW, Thummler F, Tillich M, Villarreal Aguilar JC, Widiez T, Wong GK, Wymore A, Zhang Y, Zimmer AD, Quatrano RS, Mayer KFX, Goodstein D, Casacuberta JM, Vandepoele K, Reski R, Cuming AC, Tuskan GA, Maumus F, Salse J, Schmutz J, Rensing SA (2018) The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J* 93 (3):515-533. doi:10.1111/tpj.13801
- Li L, Hey S, Liu S, Liu Q, McNinch C, Hu HC, Wen TJ, Marcon C, Paschold A, Bruce W, Schnable PS, Hochholdinger F (2016) Characterization of maize roothairless6 which encodes a D-type cellulose synthase and controls the

switch from bulge formation to tip growth. *Sci Rep* 6:34395.
doi:10.1038/srep34395

Li X (2017) Characterization of cellulose synthesis complexes in *Physcomitrella patens*. University of Rhode Island,

Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, van der Oost J, Koonin EV (2011) Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 9 (6):467-477. doi:10.1038/nrmicro2577

Mallett DR, Chang M, Cheng X, Bezanilla M (2019) Efficient and modular CRISPR-Cas9 vector system for *Physcomitrella patens*. *Plant Direct* 3 (9):e00168.
doi:10.1002/pld3.168

Manfield IW, Orfila C, McCartney L, Harholt J, Bernal AJ, Scheller HV, Gilmartin PM, Mikkelsen JD, Paul Knox J, Willats WG (2004) Novel cell wall architecture of isoxaben-habituated *Arabidopsis* suspension-cultured cells: global transcript profiling and cellular analysis. *Plant J* 40 (2):260-275.
doi:10.1111/j.1365-3113.2004.02208.x

Marraffini LA (2015) CRISPR-Cas immunity in prokaryotes. *Nature* 526 (7571):55-61. doi:10.1038/nature15386

Marraffini LA, Sontheimer EJ (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322 (5909):1843-1845.
doi:10.1126/science.1165771

McFarlane HE, Doring A, Persson S (2014) The cell biology of cellulose synthesis. *Annu Rev Plant Biol* 65:69-94. doi:10.1146/annurev-arplant-050213-040240

Morgan JL, Strumillo J, Zimmer J (2013) Crystallographic snapshot of cellulose synthesis and membrane translocation. *Nature* 493 (7431):181-186.
doi:10.1038/nature11744

Mueller SC, Brown RM, Jr. (1980) Evidence for an intramembrane component associated with a cellulose microfibril-synthesizing complex in higher plants. *J Cell Biol* 84 (2):315-326. doi:10.1083/jcb.84.2.315

- Newman RH, Hill SJ, Harris PJ (2013) Wide-angle x-ray scattering and solid-state nuclear magnetic resonance data combined to test models for cellulose microfibrils in mung bean cell walls. *Plant Physiol* 163 (4):1558-1567. doi:10.1104/pp.113.228262
- Nixon BT, Mansouri K, Singh A, Du J, Davis JK, Lee JG, Slabaugh E, Vandavasi VG, O'Neill H, Roberts EM, Roberts AW, Yingling YG, Haigler CH (2016) Comparative structural and computational analysis supports eighteen cellulose synthases in the plant cellulose synthesis complex. *Sci Rep* 6:28696. doi:10.1038/srep28696
- Norris JH, Li X, Huang S, Van de Meene AML, Tran ML, Killeavy E, Chaves AM, Mallon B, Mercure D, Tan HT, Burton RA, Doblin MS, Kim SH, Roberts AW (2017) Functional specialization of cellulose synthase isoforms in a moss shows parallels with seed plants. *Plant Physiol* 175 (1):210-222. doi:10.1104/pp.17.00885
- Park S, Szumlanski AL, Gu F, Guo F, Nielsen E (2011) A role for CSLD3 during cell-wall synthesis in apical plasma membranes of tip-growing root-hair cells. *Nat Cell Biol* 13 (8):973-980. doi:10.1038/ncb2294
- Pear JR, Kawagoe Y, Schreckengost WE, Delmer DP, Stalker DM (1996) Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase. *Proc Natl Acad Sci U S A* 93 (22):12637-12642. doi:10.1073/pnas.93.22.12637
- Persson S, Wei H, Milne J, Page GP, Somerville CR (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc Natl Acad Sci U S A* 102 (24):8633-8638. doi:10.1073/pnas.0503392102
- Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud PF, Lindquist EA, Kamisugi Y, Tanahashi T, Sakakibara K, Fujita T, Oishi K, Shin IT, Kuroki Y, Toyoda A, Suzuki Y, Hashimoto S, Yamaguchi K, Sugano S, Kohara Y, Fujiyama A, Anterola A, Aoki S, Ashton N, Barbazuk WB, Barker E, Bennetzen JL, Blankenship R, Cho SH, Dutcher SK, Estelle M, Fawcett JA, Gundlach H, Hanada K, Heyl A, Hicks KA, Hughes J, Lohr M, Mayer K, Melkozernov A, Murata T, Nelson DR, Pils B, Prigge M, Reiss B, Renner T, Rombauts S, Rushton PJ, Sanderfoot A, Schween G, Shiu SH, Stueber K, Theodoulou FL, Tu H, Van de Peer Y, Verrier PJ, Waters E, Wood A, Yang L, Cove D, Cuming AC, Hasebe M, Lucas S, Mishler BD, Reski R, Grigoriev IV, Quatrano RS, Boore JL (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319 (5859):64-69. doi:10.1126/science.1150646

- Richmond T (2000) Higher plant cellulose synthases. *Genome Biol* 1 (4):REVIEWS3001. doi:10.1186/gb-2000-1-4-reviews3001
- Richmond TA, Somerville CR (2000) The cellulose synthase superfamily. *Plant Physiol* 124 (2):495-498. doi:10.1104/pp.124.2.495
- Richmond TA, Somerville CR (2001) Integrative approaches to determining Csl function. *Plant Mol Biol* 47 (1-2):131-143
- Roberts AW, Bushoven JT (2007) The cellulose synthase (*CESA*) gene superfamily of the moss *Physcomitrella patens*. *Plant Mol Biol* 63 (2):207-219. doi:10.1007/s11103-006-9083-1
- Roberts AW, Dimos CS, Budziszek MJ, Jr., Goss CA, Lai V (2011) Knocking out the wall: protocols for gene targeting in *Physcomitrella patens*. *Methods Mol Biol* 715:273-290. doi:10.1007/978-1-61779-008-9_19
- Rudolph U, Gross H, Schnepf E (1989) Investigations of the turnover of the putative cellulose-synthesizing particle “rosettes” within the plasma membrane of *Funaria hygrometrica* protonema cells. *Protoplasma* 148:57-69
- Sakakibara K, Nishiyama T, Sumikawa N, Kofuji R, Murata T, Hasebe M (2003) Involvement of auxin and a homeodomain-leucine zipper I gene in rhizoid development of the moss *Physcomitrella patens*. *Development* 130 (20):4835-4846. doi:10.1242/dev.00644
- Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* 32 (4):347-355. doi:10.1038/nbt.2842
- Schumaker KS, Dietrich MA (1998) Hormone-induced signaling during moss development. *Annu Rev Plant Physiol Plant Mol Biol* 49:501-523. doi:10.1146/annurev.arplant.49.1.501
- Silva G, Poirot L, Galetto R, Smith J, Montoya G, Duchateau P, Paques F (2011) Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr Gene Ther* 11 (1):11-27. doi:10.2174/156652311794520111

- Somerville C (2006) Cellulose synthesis in higher plants. *Annu Rev Cell Dev Biol* 22:53-78. doi:10.1146/annurev.cellbio.22.022206.160206
- Tran ML, McCarthy TW, Sun H, Wu SZ, Norris JH, Bezanilla M, Vidali L, Anderson CT, Roberts AW (2018) Direct observation of the effects of cellulose synthesis inhibitors using live cell imaging of *Cellulose Synthase (CESA)* in *Physcomitrella patens*. *Sci Rep* 8 (1):735. doi:10.1038/s41598-017-18994-4
- van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJ (2009) CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem Sci* 34 (8):401-407. doi:10.1016/j.tibs.2009.05.002
- van der Oost J, Westra ER, Jackson RN, Wiedenheft B (2014) Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat Rev Microbiol* 12 (7):479-492. doi:10.1038/nrmicro3279
- Vandavasi VG, Putnam DK, Zhang Q, Petridis L, Heller WT, Nixon BT, Haigler CH, Kalluri U, Coates L, Langan P, Smith JC, Meiler J, O'Neill H (2016) A structural study of CESA1 catalytic domain of *Arabidopsis* cellulose synthesis complex: evidence for CESA trimers. *Plant Physiol* 170 (1):123-135. doi:10.1104/pp.15.01356
- Wang X, Cnops G, Vanderhaeghen R, De Block S, Van Montagu M, Van Lijsebettens M (2001) *AtCSLD3*, a cellulose synthase-like gene important for root hair growth in *Arabidopsis*. *Plant Physiol* 126 (2):575-586. doi:10.1104/pp.126.2.575
- Yosef I, Goren MG, Qimron U (2012) Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res* 40 (12):5569-5576. doi:10.1093/nar/gks216
- Zimmer AD, Lang D, Buchta K, Rombauts S, Nishiyama T, Hasebe M, Van de Peer Y, Rensing SA, Reski R (2013) Reannotation and extended community resources for the genome of the non-seed plant *Physcomitrella patens* provide insights into the evolution of plant gene structures and functions. *BMC Genomics* 14:498. doi:10.1186/1471-2164-14-498