University of Stavanger

Faculty of Science and Technology

MASTER'S THESIS

Study program/ Specialization: Biological Chemistry	Spring semester, 2017 Open access
Writer: Yvonne Sletthaug	…… (Writer's signature)
Faculty supervisor: Amr Ramzy Abass Kataya	
Thesis title: Targeted mutagenesis of peroxisomal protein phosphatases using CRISPR-Cas9 approach	
Credits (ECTS): 60	
Key words: Peroxisomal protein phosphatases, CRISPR/Cas9, *Arabidopsis, S. lycopersicum*, PAP7, PLL2, PLL3, PSY2L	Pages: 73 + enclosure: 11 Stavanger, 12.06/2017 Date/year

Front page for master thesis
Faculty of Science and Technology
Decision made by the Dean October 30th 2009

Acknowledgment

First, I would like to thank my supervisor, Dr. Amr R.A Kataya, for the patient guidance throughout this year. Thank you for always keeping your door open and for giving me the confidence I needed, when I needed it.

I would also like to thank Prof. Cathrine Lillo for letting me work in her lab and the other members at the lab, for contributing to a fun and pleasant environment. Thanks to my fellow students, for "suffering" alongside me in long days of writing and for never saying no to lunch.

Last but not least, I would like to thank my friends and family for their support and for the understanding of the time and effort that has been put down in writing this thesis.

Abstract

The newest development in the techniques of gene editing is the clustered regularly interspaced short palindromic repeat - associated protein9 (CRISPR-Cas9). This is a complex consisting of single-guide RNA (sgRNA) and a Cas9 protein. The sgRNA can, based on designing of specific spacers, recognize specific target sites in the DNA leading to double stranded breaks (DBS) through cutting by Cas9. In this study, the establishment of CRSISPR/Cas9 editing for stable plant transgenic lines was aimed. Selection of suitable plasmids harboring Cas9 encoding gene, and genes encoding sgRNA components was accomplished. The selection of vectors depended also on choosing of optimum promoter, for example using the egg cell specific promotor EC1.1 promoter fused with an EC1.2 enhancer in *Arabidopsis thaliana (Arabidopsis)* and the *Cauliflower mosaic virus* (CaMV) 35S promotor in *Solanum lycopersicum (S. lycopersicum).*

The CRISPR/Cas9 method was used in order to knockout peroxisomal protein phosphatases in *Arabidopsis*, and the putative regulator of protein phosphatase 4 (PSY2L) in tomato. Two predicted spacers were used for generating two sgRNAs in order to direct Cas9 to two different targets in the genome of *Arabidopsis* purple acid phosphatase 7 (PAP7), Pol-like phosphatase (PLL2, PLL3, PLL3/PLL2 (two different combinations)), and two variants for putative PSY2L in tomato. Golden gate cloning was used to clone the selected spacers, and the obtained pCAMBIA-based binary vectors including cloned Cas9 and constructed sgRNAs were transformed to *Agrobacterium tumefaciens (Agrobacterium)* GV3101. Subsequently, *Arabidopsis* plants were transformed with these vectors, and seeds were screened for successful transformations. Genomes for the T1 generated plants were analyzed by T7 endonuclease 1 assay (T7E1), and only heterozygote plants could be detected. We were able to obtain 10/10 heterozygote PAP7 mutants, 3/5 heterozygote PLL3 mutants, 2/5 heterozygote PLL2 mutants and 1/10 PLL3/PLL2 double mutants for *Arabidopsis*. T2 generations was also screened on selectable markers, and representative plants were examined phenotypically. Further analysis will be needed to distinguish between wild type and homozygote plants using T7E1 assay and/or by sequencing. Moreover, the constructed pCAMBIA-based binary vectors and pGreen-based binary vectors for editing PSY2L in tomato, and peroxisomal phosphatases in *Arabidopsis* are successfully cloned.

Abbreviations

Cas9	Associated Protein 9
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeat
crRNA	CRISPR RNA
DSB	Double Strand Break
HDR	Homology-Direct Repair
LB	Luria-Bertani
MS	Murashige & Skoog
NHEJ	Non-Homologous End Joining
PAM	Protospacer Adjacent Motif
PAP7	Purple Acid Phosphatase 7
PCR	Polymerase Chain Reaction
PLL2	Pol-like 2
PLL3	Pol-like 3
PP4	Protein Phosphatase 4
pre-crRNA	precursor CRISPR RNA
PSY2L	PSY2-like
PTS1	Peroxisome Targeting Signal 1
PTS2	Peroxisome Targeting Signal 2
SSN	Sequence Specific Nuclease
sgRNA	single-guide RNA
TALEN	Transcription Activator-Like Effector Nuclease
T-DNA	Transfer DNA
tracrRNA	Trans activating RNA
T7E1	T7 Endonuclease 1
WT	Wild Type
ZFN	Zinc Finger Nuclease

Contents

List of tables

List of figures

1 Introduction

1.1 CRISPR/Cas9

1.1.1 History of gene manipulation

In the past decade, there has been a great development in the ability and techniques for studying genes and to improve crops. In the early studies, natural mutants were used to characterize important biological mechanisms. Over the years, large mutant libraries have been constructed using biological, physical and chemical mutagenesis. Since random mutagenesis requires large scale screening, it is both time consuming and costly. In addition, random mutagenesis can cause undesirable mutations and rearrangements (Ma et al. 2016). Other strategies used for the study of plant genes and their functions are antisense RNA (Mol et al. 1990), virus induced gene-silencing (Baulcombe 1999) and RNA interference (Smith et al. 2000). These methods rely on suppressing specific genes by repressing their corresponding mRNAs. Usually, only a partial repression of these mRNAs is achieved (Ma et al. 2016).

Not until the emergence of programmable sequence-specific nucleases (SSNs), scientists could perform targeted mutagenesis. SSNs induce double-stranded breaks (DSBs) in specific chromosomal sites of the DNA. These breaks can be repaired by one of two pathways; the non-homologous end-joining (NHEJ) pathway or the homology-directed repair (HDR). For the HDR pathway to be able to repair the breaks, it requires homologous donor templates present in the moment of the DSB. The error-prone NHEJ pathway does not require such templates and the breaks are fixed by insertions, deletions or substitutions. This makes the NHEJ pathway the most likely pathway to repair the breaks and cause mutations. Two well-known programmable SSNs are the zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) (Ma et al. 2016). The ZFNs were the first SSNs to be used for genome editing (Bibikova et al. 2003) and has been widely used to edit genes in plants. Due to difficulties in constructing these ZFNs and the high expenses that came with to do so, they were quickly replaced in favor of the TALENs (Boch et al. 2009; Moscou and Bogdanove 2009). Even though TALENs were easier to use than the ZFNs, complicated tandem repeat domains in the TAL proteins still had to be constructed (Ma et al. 2016).

The latest break-through in the technology of targeted mutagenesis is the clustered regularly interspaced short palindromic repeat (CRISPR)- associated protein9 (Cas9). This system is a bacterial immune defense system that has been adapted from the type II CRISPR system of the bacterium *Streptococcus pyogenes* (Ma et al. 2016). Shortly after exposure of invasive genetic elements from bacteriophages or plasmids, the immune system of the bacteria integrates short fragments of the foreign DNA into its own chromosome. If the bacteria gets invaded a second time by the same

invader, the genetic record enables the host to quickly recognize and destroy the invasive elements (Jiang and Doudna 2017). The systems have been divided into six different types; type I-VI, each employing its own sets of Cas proteins. Compared to the type I and type III systems, which depend on large Cas protein complexes for cleavage of foreign DNA, the type II system only requires one Cas protein (Jiang and Doudna 2017).

Due to it's simple, efficient and cost-effective ways of making both single and multiple mutations in several different organisms, CRISPR is becoming the most popular approach for genome engineering (Ma et al. 2016). Stable transgenic lines have been reported for both *Arabidopsis thaliana (Arabidopsis)* and rice (Feng et al. 2014; Zhang et al. 2014).

1.1.2 Structure and mechanism of the CRISPR/Cas9 type II system.

When foreign DNA enters the bacteria, the fragments of the foreign DNA is integrated into the host chromosome at the proximal end of a CRISPR array (Jiang and Doudna 2017). This is known as the adaptive phase (Terns and Terns 2011). In the phase of CRISPR RNA biogenesis, numerous precursor CRISPR RNAs (pre-crRNAs) are transcribed from the CRISPR array and further processed to yield CRISPR RNAs (crRNAs)(Terns and Terns 2011; Wiedenheft et al. 2012). In the type I and type III CRISPR systems, multiple Cas proteins form complexes with each crRNA before scavenging the intracellular environment for foreign DNA. As mentioned above, the type II system only requires one Cas protein, known as Cas9. Upon a second infection, the crRNAs of the type II system undergo a maturation by binding of a trans activating crRNA (tracrRNA) forming a dual RNA structure. This structure then directs Cas9 for detection and cleavage of foreign DNA (Jiang and Doudna 2017; Wiedenheft et al. 2012).

In order to use this system in genome engineering, it has been simplified by the creation of a chimeric single guide RNA (sgRNA). This mimics the original crRNA-tracrRNA duplex formed in the bacteria (Jinek et al. 2012). The system is built up of two main components; the Cas9 protein and the sgRNA, and is based on the pairing of this sgRNA to a specific DNA target site and cleavage by Cas9 to induce mutations (Figure 1) (Jinek et al. 2012).

The 5’ terminal part of the sgRNA contains a 19-nucleotide target sequence called a spacer. By designing a spacer sequence complementary to a DNA target site sequence, the nuclease complex formed by the Cas9 protein and the sgRNA can recognize this specific target site and make a DSB (Ma et al. 2016). These features makes the CRISPR/Cas9 system much easier to manipulate than both ZFNs and TALENs, which requires advanced and time-consuming protein engineering of DNA-recognition domains for each target site (Jiang and Doudna 2017).

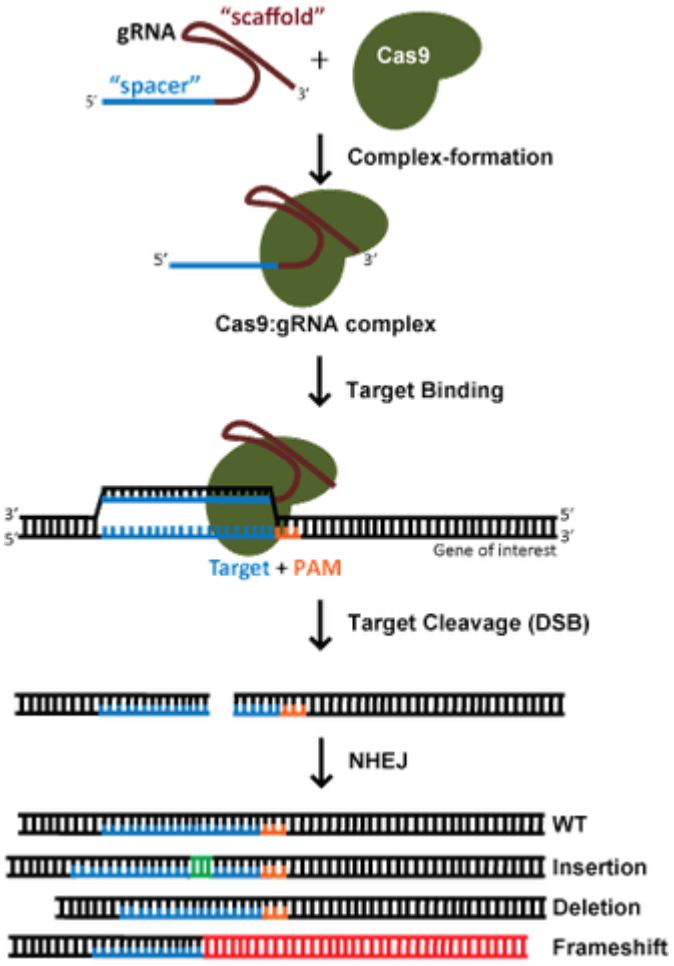


Figure 1: Structure and mechanism of CRISPR/Cas9. The sgRNA is built up by a spacer sequence and sgRNA “scaffold”. By forming a complex with a Cas9 protein, the complex can detect target sequences in the DNA that is complementary to the spacer sequence. If the target sequence is located directly upstream a PAM, the spacer sequence binds the target sequence leading to DBS cut by Cas9. The figure is taken from Addgenes CRISPR/Cas9 guide (Addgene 2017)

For effective target recognition by the CRISPR/Cas9 nuclease complex, the target sequence needs to be localized directly upstream a protospacer adjacent motif (PAM) (Jinek et al. 2012; Sternberg et al. 2014). When the complex has localized the target site, the Cas9 protein melts the target sequence and separates the strands, making it possible for the spacer part of the sgRNA to bind to the complementary strand (Ma et al. 2016). A so-called seed-sequence within the spacer has shown to be important for target specificity. In the type II CRISPR system, this is a 10-12 nucleotide long sequence located in the 3’ end of the spacer. Mismatches in this region impairs or prevents DNA binding and cleavage. Close homology between the spacer and the target DNA in the seed region can, even with large amount of mismatches elsewhere in the spacer, lead to off-target binding and cleavage (Jiang and Doudna 2017). When the spacer has bound the target sequence, the RuvC and the HNH domains of the Cas9 protein cut both strands in the target site approximately three nucleotides upstream of the PAM (Ma et al. 2016). The HNH domain cuts the DNA strand complementary to the spacer, while the RuvC domain cleaves the non-complementary strand (Jiang and Doudna 2017). The resulting DSB is usually

repaired by the NHEJ pathway (unless a homologous donor template is present) causing mutations in the DNA target site (Ma et al. 2016).

1.1.3 CRISPR/Cas9 Vector systems in plants

Early efforts in editing genes in plants by the CRISPR/Cas9 system have been successful, but not very efficient (Li et al. 2013). To increase the efficiency, several vector systems have been developed (Ma et al. 2016; Wang et al. 2015; Xing et al. 2014; Zhang et al. 2016).

In plants, the expression of sgRNAs are usually driven by U3 or U6 promoters, while the expression of *Cas9* is driven by constitutive promotors like those of Ubiquitin gene of maize, rice and *Arabidopsis*, as well as the *Cauliflower* mosaic virus (CaMV) 35S promoter (Ma et al. 2016). A third *Cas9* promoter, the promoter of the egg-cell specific *EC1.2* gene, has shown to further enhance the editing efficiency CRISPR/Cas9 in *Arabidopsis* (Wang et al. 2015).

1.1.4 Delivery of expression cassettes into plants

The *Cas9*- and sgRNA expression cassettes can be arranged in separate or single vector constructs before delivery into the plant cells. In the first attempts to test the efficiency of the CRISPR/Cas9 system, scientists delivered plasmids carrying the *Cas9* and sgRNA expression cassettes directly into the protoplasts or leaves of the plants. By delivery into protoplasts, obtaining stable transgenic lines showed to be difficult for several plant species (Jiang et al. 2013). An efficient and widely used method for delivering plasmids containing *Cas9* and sgRNA expression cassettes into plants is by *Agrobacterium tumefaciens (Agrobacterium)* mediated transformation (Ma et al. 2016). *Arabidopsis* is ususally transformed by *Agrobacterium* mediated floral dipping (Ma et al. 2015; Wang et al. 2015; Xing et al. 2014). For rice, maize, tomato and other mono- and dicot plants, *Agrobacterium* mediated transformation of callus, immature embryos or other tissues have been used (Ma et al. 2016). For *Solanum lycopersicum* (*S. lycopersicum), Agrobacterium* transformation of callus has proven to be an efficient way of editing genes in tomato (Brooks et al. 2014).

Biolistic transformation of callus or immature embryos have also been used to integrate Cas9 and sgRNA expression constructs (Li et al. 2015; Shan et al. 2013; Svitashev et al. 2015). Even though this transformation has shown to be successful in producing heritable mutations, it requires expensive equipment such as a gene gun.

In an attempt to ease the public concerns about gene modified organisms (GMOs), a DNA free strategy for editing plant genomes has been developed (Woo et al. 2015). Here, pre-assembled

complexes of purified Cas9 protein and synthesized sgRNA were transfectively delivered into protoplasts (Ma et al. 2016; Woo et al. 2015).

1.1.5 Multiplex genome editing by CRISPR/Cas9 system

To be able to study complex protein families and related genes, and to analyze epistatic relationships in genetic pathways, multiple genes needs to be simultanoulsy edited. Several strategies have been developed to meet the need of multiplex genome editing. As the *Agrobacterium* co-transformation of several T-DNAs from separate binary vectors in plants have shown to be uncontrollable, different strategies have been developed to assemble multiple sgRNA expression cassettes into one CRISPR/Cas9 binary vector (Ma et al. 2016).

Sequential rounds of regular cloning or multiple restiction enzymes, yeilding sequential compatible palindromic sticky ends, can be used to insert different sgRNA expression cassettes into one binary vector (Zhang et al. 2016; Zhou et al. 2014). The drawback of these methods is that only a few (up to three) sgRNA expression cassettes can be inserted at the time, making the construction of vectors with multiple sgRNA expression cassettes time-consuming (Ma et al. 2016).

By using Golden Gate cloning, scientists have overcome this problem. Here, the DNA fragments and the vector are cut and ligated together in the same reaction. The cutting is performed by help of type II restriction enzymes creating sequential, non-palindromic sticky ends in the fragments of DNA (Engler et al. 2008; Ma et al. 2016). Based on this strategy, Xing et al. (2014) and Ma et al. (2015), have developed CRISPR/Cas9 vector systems in which CRISPR/Cas9 binary constructs containing multiple sgRNA expression cassettes can be prepared in a single round of cloning.

1.2 Protein phosphatases

1.2.1 Protein phosphorylation and de-phosphorylation

Proteins are important molecules in regulating mechanisms in cells. By phosphorylation and de-phosphorylation, the protein activity is adjusted. This influences reaction rates, cellular localization, stability and the ability for the protein to interact with other proteins. Phosphorylation of proteins is performed by protein kinases. When a protein kinase phosphorylates another protein, it transfers a phosphate group from ATP to the hydroxyl group of a Serine, Threonine or Tyrosine residue of this protein. The less studied protein phosphatases regulate protein activity by de-phosphorylating proteins. This occurs by freeing the phosphate groups from the protein by hydrolyzing the phosphoester bonds (Lillo et al. 2014).

1.2.2 Peroxisomes

Peroxisomes are small membrane-enclosed organelles with a size of 0.1-1 µm in diameter. These organelles are involved in several biological processes. First they were discovered to remove hydrogen atoms from different substrates and transferring them to oxygen, producing hydrogen-peroxide, which again was removed by catalase (De Duve and Baudhuin 1966). Later, they have also been discovered to be involved in many other processes e.g. fatty acid β-oxidation, photorespiration, jasmonate biosynthesis, polyamine catabolism, branched-chain amino acid metabolism, ureide pathway and salicylic acid biosynthesis (Kaur et al. 2009).

Most peroxisomal matrix proteins identified, harbors peroxisome targeting signal type 1 (PTS1). In some of the identified proteins, PTS1 is replaced by a peroxisome targeting signal type 2 (PTS2). By recognition of PTS1 and PTS2 peptides by soluble receptors located in the cytosol, the proteins harboring them are guided to specific docking sites at the membrane of the peroxisome (Kaur et al. 2009).

1.2.3 Protein phosphatases

Phosphatases are considered highly specific towards their protein substrates and are regulated in complex manners. Since many of the them only function when being part of complexes with one or several regulatory subunits, they are difficult to study. Protein phosphatase complexes often consist of several regulatory subunits. They are therefore believed to rival protein kinases in binding substrates (Lillo et al. 2014).

Little information is found about peroxisomal protein phosphatases. Peroxisomal targeting of the protein phosphatase 2A (PP2A) regulatory subunit B'θ has been reported by Matre et al. (2009). Later, Kataya et al. (2015a) showed that the peroxisomal import of some catalytic- and scaffolding subunits of PP2A depends on B'θ. They also showed that B'θ knock out mutants were impaired in peroxisomal β-oxidation. Another protein phosphatase, the MAP kinase phosphatase 1, has by Kataya et al. (2015b) been shown to target peroxisomes after being exposed to different biotic and abiotic stresses.

The protein phosphatases are divided into four gene families; Serine/threonine-specific phosphor-protein phosphatases (PPP), Mg^{2+}-dependent protein phosphatases (PPM/PP2C), Asp based protein phosphatases and Phospho-tyrosine phosphatases (PTP) (Lillo et al. 2014).

1.2.3.1 PSY2L

The PPPs are considered the most highly conserved proteins across eukaryotic cells and accounts for more than 80% of the protein phosphatase activity. They can be further divided into subgroups: PP1, PP2/PP2A. PP3/PP2B, PP4, PP5, PP6, PP7, PPKL/Kech and bacterial like protein phosphatases (Lillo et al. 2014).

A common ancestor of PP2A, PP6 and PP4 is suggested by the formation of these three phosphatases in a separate cluster (Uhrig et al. 2013).

In the yeast specie *Saccharomyces cerevisiae (S. cerevisiae)*, PP4 seem to have overlapping functions with PP2A and is lethal when both PP2A catalytic subunits PPH21, PPH22 and the catalytic subunit of PP4 is removed (Stark 1996). The catalytic subunit of PP4 in *S. cerevisiae*, PPH3, is found in complexes with two regulators named YBL1046W and PSY2 (Lillo et al. 2014).

In mammals, PP4 seems to be essential in the development of thymocytes and pre-T-cell receptor signaling and is lethal in the early embryonic stages when removed (Shui et al. 2007). The mammalian homologs of YBL1046W and PSY2 are PP4R2 and PP4R3. In addition, mammals possess two regulatory subunits named PP4R1 and PP4R4 (Cohen et al. 2005).

In *Arabidopsis*, little is known about the functions of PP4. The *Arabidopsis* PP4 is a complex of two catalytic-, PP4-1 and PP4-2, and two putative regulatory subunits, PP4R2L and PSY2L (AT3G06670). *Arabidopsis* PP4R2L and PSY2L are considered homologs of *S. cerevisiae* YBL1046W and PSY2, and mammalian PP4R2 and PP4R3 (Lillo et al. 2014). According to Lillo et al. (2014), the *Arabidopsis* PSY2-like subunit show 30% identity with human PP4R3 and *S. cerevisiae* PSY2, while the *Arabidopsis* PP4R2L subunit show 32% identity with the human PP4R2.

The function and the subcellular localization of PP4 in plants have not yet been fully investigated. The catalytic subunits of PP4 in *Arabidopsis* have been found to be localized to the cytoplasm and suggested to target cytosol and nucleus (Lillo et al. 2014). Phenotyping of homozygous *Arabidopsis PSY2L* knock out plants obtained by inserting T-DNA in the *PSY2L* gene showed clear dwarfism. They also showed delayed growth and had an extended life-span (Napitupulu 2016). Due to high expression of the genes for PP4 regulatory- and catalytic subunits in seeds and embryos, it seems like the protein might be involved in processes during stages of development (Lillo et al. 2014).

1.2.3.2 PLL3 and PLL2

The PPM/PP2C phosphatases are, in *Arabidopsis*, divided into 13 subfamilies (Xue et al. 2008). By screening the Arabidopsis genome for protein phosphatases harboring PTS1, Kataya et al. (2016)

identified two members of the subfamily C; PLL3 and PLL2. According to the SUBA database, PLL3 and PLL2 show 62% identity. The same database also predicted the two proteins to target nucleus (Kataya et al. 2016; Hooper et al. 2014). By generating fusion proteins of PLL3 and PLL2 tagged with enhanced yellow fluorescent protein (EYFP), Kataya et al. (2016), showed that both PLL3 and PLL2 targets peroxisomes. In addition, PLL2 also showed targeting of nucleus and nucleolus.

Even though the functions of many other members of this family has been revealed, such as regulation of meristem development and leaf-development, no functions have been ascribed to PLL3 and PLL2 (Song and Clark 2005; Yu et al. 2003). The effects of sucrose were tested on PLL3 and PLL2 mutants by Kataya et al. (2016). Only PLL3 mutants showed any form of sugar-dependency, but the results could not be confirmed due to lack of other knock out lines. No developmental phenotypes have been observed in homozygous PLL3 and PLL2 mutants (Kataya et al. 2016; Song and Clark 2005).

1.2.3.3 PAP7

Another protein shown to target peroxisomes is the purple-acid phosphatase 7 (Kataya et al. 2016). This protein belongs to a family of di-nuclear metallohydrolases. Common for the proteins in this family is that they contain a di-nuclear center built by two closely spaced metal ions. This center is required for the proteins to be able to carry out hydrolytic reactions (Schenk et al. 2013). PAPs are widespread in eukaryotes and have been identified and characterized from a wide range of organisms, such as bacteria, mammals, fungi, and plants (Flanagan et al. 2006; Schenk et al. 2000a; Schenk et al. 2000b).

In mammals, PAPs are also known as tartrate resistant acid phosphatases and are believed to be involved in bone resorption and bone metabolism, iron-transport (in pigs during pregnancy) and the generation of reactive oxygen species (ROS) in immune-response. These beliefs have been supported by the findings of high expressions of PAPs in osteoclasts, activated macrophages and dendritic cells. Mammalian PAPs have also shown to be bifunctional and can catalyze both hydrolytic reactions and peroxidation (Schenk et al. 2013).

In most plant PAPs, the di-nuclear center is formed by the connection of a metal ion with a zinc or magnesium ion instead of another metal ion (Durmus et al. 1999; Schenk et al. 1999). Due to upregulation of PAP expression in phosphor-starving *Arabidopsis (Del Pozo et al. 1999; Veljanovski et al. 2006; Wang et al. 2011)*, tomato (Bozzo et al. 2002, 2004), rice (Zhang et al. 2011) and many other plant species, some of the PAPs in plants are believed to mobilize inorganic phosphate from organophosphates in soil (Schenk et al. 2013). 29 putative genes have been identified in Arabidopsis (Tran et al. 2010). Until now, AtPAP10, AtPAP12, AtPAP17, AtPAP25 and AtPAP26 have been

identified to be involved in phosphor-starvation (Del Pozo et al. 1999; Del Vecchio et al. 2014; Tran et al. 2010; Veljanovski et al. 2006; Wang et al. 2011).

PAP7 is one of the less studied PAPs and little information can be found about this protein. By testing the transcriptional response in phosphor-deprived *Arabidopsis* plants, Li et al. (2002) could not observe any obvious changes in AtPAP7 expression.

1.3 Aim and objectives of the study

In order to continue the study of the newly discovered peroxisomal protein phosphatases, PAP7, PLL3 and PLL2 (Kataya et al. 2016), *Arabidopsis* knock out lines have to be obtained. In addition to knocking out peroxisomal protein phosphatases in *Arabidopsis,* we also wanted to knock out the regulatory subunit PSY2L of PP4 in tomato to see if the effect observed in *Arabidopsis* could be similar in tomato (Napitupulu 2016).

The aim of the study was therefore to generate stable transgenic plant lines through CRISPR/Cas9 editing. By establishing an effective and inexpensive way to do so, stable transgenic lines with knock outs in several other genes can be obtained for future research.

2 Materials

2.1 Kits

The kits used in this study are listed in table 1.

Table 1: Kits used for plasmid isolation and gel extraction.

Kit	Purpose	Supplier
NucleoSpin® Plasmid (NoLid)	Plasmid isolation	Macherey-Nagel
GenElute™ Gel Extraction Kit	Gel extraction	Sigma-Aldrich
NucleoSpin® Gel and PCR Clean-up		Macherey-Nagel
Monarch DNA Gel Extraction Kit		New England Biolabs

2.2 Plants

Knocking out the genes of protein phosphatases to generate stable transgenic plants were aimed for *Arabidopsis thaliana* and tomato *Solanum lycopersicum.* The plants were provided by Prof. Cathrine Lillo (Table 2).

Table 2: Plants used for transformation

Plant specie	Strain/ecotype	Purpose
Arabidopsis thaliana	Columbia	Transformation by flowerdipping
Solanum lycopersicum	Heinz	Transformation of tomato cotyledons

2.3 Bacteria

Two bacterial strains were used in this study; one for cloning and one for transformation (Table 3). Both were provided by Prof. Cathrine Lillo.

Table 3: Bacteria for cloning and transformation

Bacteria specie	Strain	Purpose
Escherichia coli	JM 109	Cloning of plasmids with expression cassettes
Agrobacterium tumefacien	Gv3 101	Transformation of *Arabidopsis* and *S. lycopersicum*

2.4 Plasmids

Four different plasmids were used in the process of making stable plant transgenic lines; pCBC-DT1T2, pHSE401, pHEE401 and pHSN401 (Table 4). Plasmid map figures are listed in Appendix A-1.

Table 4: Plasmids for making constructs with sgRNAs

Plasmid name	Purpose
pCBC-DT1T2 Addgene Plasmid # 50590	Template for making expression cassettes
pHSE401 Addgene Plasmid #62201	Assembly of recombinant vectors containing sgRNAs
pHEE401 Addgene Plasmid #71286	
pHSN401 Addgene Plasmid #50588	

All plasmids were provided as a gift from Qi-Jun Chen (Wang et al. 2015; Xing et al. 2014)

3 Methods

3.1 CRISPR/Cas9 editing of *Arabidopsis thaliana* PAP7, PLL3 and PLL2 and tomato *Solanum lycopersicum* PSY2L by *Agrobacterium* transformation

3.1.1 Expression cassettes for single guide RNAs

The methodology article, "A CRIPSR/Cas9 toolkit for multiplex genome editing in plants" by Xing et al. (2014) was used as a starting point in making expression cassettes. Two cassettes were made to target genes in *S. lycopersicum*; one with two sgRNAs targeting *PSY2L 1* and one with two sgRNAs dual targeting *PSY2L-1* and *PSY2L-2*. For *Arabidopsis*, five cassettes were made: one with two sgRNAs targeting *PAP7*, one with two sgRNAs targeting *PLL3*, one with two sgRNAs targeting *PLL2* and two cassettes with two sgRNAs dual targeting *PLL3* and *PLL2*.

3.1.1.1 Identification of spacers

The identification of spacers for each sgRNA is crucial for the sgRNAs to be able to identify and target specific genes in *Arabidopsis* and *S. lycopresicum.* Two spacers were identified for targeting *S. lycopersicum PSY2L-1*, two for dual targeting *PSY2L-1* and *PSY2L-2*, two for targeting *Arabidopsis PAP7*, two for targeting *PLL3* and two for targeting *PLL2* (Figure 2)

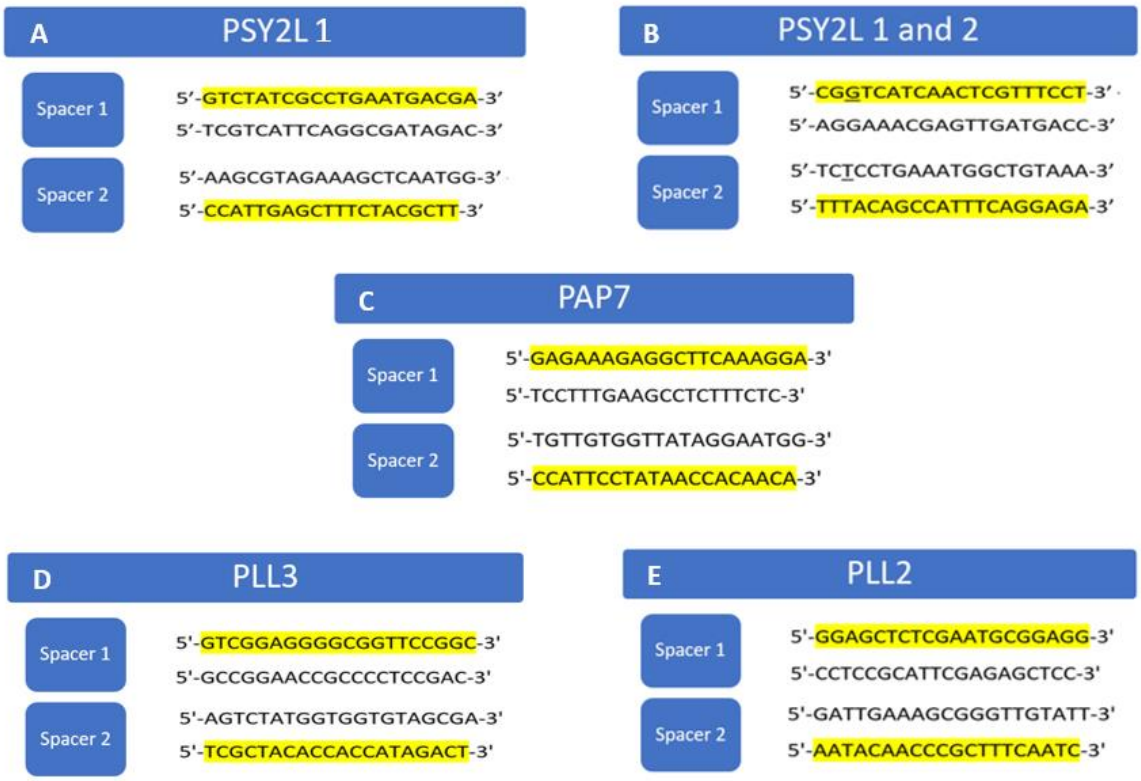


Figure 2: Predicted spacers (Yellow) for sgRNAs targeting PSY2L (A), dual targeting PSY2L 1 and 2 (B), PAP7 (C), PLL3 (D) and PLL2 (E).

PSY2L homologs in tomato were found by aligning Arabidopsis *PSY2L* (AT3G06670.1) towards *S. lycopersicum* (taxid: 4081) in NCBI. Four tomato and one human homolog appeared. The tomato homolog (Solyc12g099320.1.1) with the highest percent of homology (64%) were chosen to generate sgRNAs.

Two loci were found when searching for the annotation of *PSY2L*. The first locus (Solyc12g099320.1.1, start: 2386, end: 2455) was named *PSY2L-1* and the second locus (Solyc01g060080.2.1, start: 1043, end:1112) was named *PSY2L-2*. Alignment of *PSY2L-1* and *PSY2L-2* showed 58.8% identity and only the three variants of *PSY2L-1* came up when *PSY2L-2* was aligned with *S. lycopersicum* (taxid: 4081).

CRISPR-P (Lei et al. 2014) and CRISPR RGEN Tools (Park et al. 2016) were used to identify the first spacer of PSY2L-1 (Figure 2A). The spacer was located at the negative strand of an exon. When using mRNA, the same spacer was also identified by CCTOP (Stemmer et al. 2015). CCTOP was also used to identify the second spacer of PSY2L-1. This spacer was located at the positive strand of an exon, but since the second spacer should be reversely inserted into the expression cassette, the reverse complementary sequence of the spacer was used to design the primer. The same spacer was also identified by WustL.

For identification of possible dual targets of *PSY2L-1* and *PSY2L-2*, CCTOP was used (Figure 2B). The spacer would target *PSY2L-1* and was located at the positive strand of an exon. For *PSY2L-2* the third nucleotide of the spacer (read 5' → 3'), Guanine (G), was replaced by Cytosine (C) which gave only one mismatch between the spacers. Therefore, the spacer targeting *PSY2L-1* would also target *PSY2L-2* and was chosen to be used in the designing of the first primer. The same spacer was identified by CRISPR-P. The second spacer identified for sgRNA dual targeting *PSY2L-1* and *PSY2L-2* was located on the negative strand of an exon. Since the second spacer was going to be reversely inserted into the expression cassette, the reverse complementary sequence was used to design the primer. This spacer would target *PSY2L-2*. For *PSY2L-1* the third nucleotide of the spacer (read 5'→3'), Thymine (T), was replaced by Guanine (G) which gave only one mismatch between the spacers. Therefore, the spacer targeting *PSY2L-2* would also target *PSY2L-1* and chosen to be used in the designing of the second primer.

Spacers for sgRNAs targeting *Arabidopsis PAP7* (AT2G01880) was identified by using CHOPCHOP (Labun et al. 2016) and CCTOP (Figure 2C). The first spacer was located at the positive strand of an exon, while the second spacer was located at the negative strand of an exon. The reverse complementary sequence of the second spacer was used to design the primer.

Spacers for sgRNAs targeting *Arabidopsis PLL3* (AT3G09400.1, isoform 1) was also identified using CHOPCHOP and CCTOP (Figure 2D). Both spacers were located at the positive strand of an exon. For the second spacer, the reverse complementary strand was used to design the primer.

CHOPCHOP was also used to identify pacers for sgRNAs targeting *Arabidopsis PLL2* (AT5G02400) (Figure 2E). The first spacer was located at the negative strand of an exon, while the second spacer was located at the positive strand of an exon. To design the primer for the second spacer, the reverse complementary sequence was used.

3.1.1.2 Assembly of expression cassettes by PCR

By PCR, a pCBC-DT1DT2 plasmid (Addgene plasmid # 50590), was used as a template for making sgRNA expression cassettes. The PCR was set up by mixing the protocol from the article of Xing et al. (2014) with the protocol for the Pfu DNA Polymerase (Thermo Scientifics) (Table 5). The cycling conditions was set to fit the enzyme (Table 6).

Table 5: Setup for PCR reaction when making expression cassettes.

Component:	**Volume (µl):**
10X *Pfu* buffer with $MgSO_4$	5
dNTP mix	5
T1-Bs Forward primer (10 µM)	2
T1-F0 (1 µM)	1
T2-R0 (1 µM)	1
T2-Bs Reverse Primer (10 µM)	2
pCBC-DT1DT2 (50 ng/µl)	1
Pfu DNA Polymerase (2,5 U/µl)	1
Water, nuclease free	32
	Total volume: 50

Table 6: Cycling conditions for PCR reaction when making expression cassettes.

Step:	**Temperature:**	**Time:**	**Number of cycles:**
Initial denaturation	95	3 min	1
Denaturation	95	30 s	35
Annealing	60	30 s	
Extension	72	2 min	
Final extension	72	15 min	1

Seven PCR reactions were made. In each reaction, a specific set of primers were used (Table 7). The first reaction was designed to make an expression cassette containing two sgRNAs targeting *PSY2L 1*. The second reaction to dual target *PSY2L 1* and *2*, the third reaction to target *PAP7*, the fourth reaction to target *PLL3* and the fifth to target *PLL2*. Another two cassettes were made to dual target *PLL3* and *PLL2* by mixing the primers of reaction four and five. In sample six primers for PLL3's spacer

one were mixed with primers for PLL2's spacer two. In sample seven, primers for PLL3's spacer two were mixed with primers for PLL2's spacer one.

Table 7: Primers used in making expression cassettes.
The table shows an overview with target, primer name and primer sequence.

Reaction number:	Target:	Primer name:		Primer Sequence:
1	PSY2L 1	CRISPR_9	T1-BsF	AATAATGGTCTCTATTGTCTATCGCCTGAATGACGAGTT
		CRISPR_10	T1-F0	TGTCTATCGCCTGAATGACGAGTTTTAGAGCTAGAAATAGC
		CRISPR_11	T2-R0	AACCCATTGAGCTTTCTACGCTCAATCTCTTAGTCGACTCTAC
		CRISPR_12	T2-BsR	ATTATTGGTCTCTAAACCCATTGAGCTTTCTACGCTC
2	PSY2L 1/2	CRISPR_13	T1-BsF	AATAATGGTCTCTATTGGGTCATCAACTCGTTTCCTGTT
		CRISPR_14	T1-F0	TGGGTCATCAACTCGTTTCCTGTTTTAGAGCTAGAAATAGC
		CRISPR_15	T2-R0	AACTTTACAGCCATTTCAGGAGCAATCTCTTAGTCGACTCTAC
		CRISPR_16	T2-BsR	ATTATTGGTCTCTAAACTTTACAGCCATTTCAGGAGC
3	PAP7	CRISPR_17	T1-BsF	AATAATGGTCTCTATTGAGAAAGAGGCTTCAAAGGAGTT
		CRISPR_18	T1-F0	TGAGAAAGAGGCTTCAAAGGAGTTTTAGAGCTAGAAATAGC
		CRISPR_19	T2-R0	AACCCATTCCTATAACCACAACCAATCTCTTAGTCGACTCTAC
		CRISPR_20	T2-BsR	ATTATTGGTCTCTAAACCCATTCCTATAACCACAACC
4	PLL3	CRISPR_21	T1-BsF	AATAATGGTCTCTATTGTCGGAGGGGCGGTTCCGGCGTT
		CRISPR_22	T1-F0	TGTCGGAGGGGCGGTTCCGGCGTTTTAGAGCTAGAAATAGC
		CRISPR_23	T2-R0	AACTCGCTACACCACCATAGACCAATCTCTTAGTCGACTCTAC
		CRISPR_24	T2-BsR	ATTATTGGTCTCTAAACTCGCTACACCACCATAGACC
5	PLL2	CRISPR_25	T1-BsF	AATAATGGTCTCTATTGGAGCTCTCGAATGCGGAGGGTT
		CRISPR_26	T1-F0	TGGAGCTCTCGAATGCGGAGGGTTTTAGAGCTAGAAATAGC
		CRISPR_27	T2-R0	AACAATACAACCCGCTTTCAATCAATCTCTTAGTCGACTCTAC
		CRISPR_28	T2-BsR	ATTATTGGTCTCTAAACAATACAACCCGCTTTCAATC
6	PLL3/ PLL2	CRISPR_21	T1-BsF	AATAATGGTCTCTATTGTCGGAGGGGCGGTTCCGGCGTT
		CRISPR_22	T1-F0	TGTCGGAGGGGCGGTTCCGGCGTTTTAGAGCTAGAAATAGC
		CRISPR_27	T2-R0	AACAATACAACCCGCTTTCAATCAATCTCTTAGTCGACTCTAC
		CRISPR_28	T2-BsR	ATTATTGGTCTCTAAACAATACAACCCGCTTTCAATC
7	PLL2/ PLL3	CRISPR_25	T1-BsF	AATAATGGTCTCTATTGGAGCTCTCGAATGCGGAGGGTT
		CRISPR_26	T1-F0	TGGAGCTCTCGAATGCGGAGGGTTTTAGAGCTAGAAATAGC
		CRISPR_23	T2-R0	AACTCGCTACACCACCATAGACCAATCTCTTAGTCGACTCTAC
		CRISPR_24	T2-BsR	ATTATTGGTCTCTAAACTCGCTACACCACCATAGACC

To check the size of the products, the PCR products were loaded on a 1% agarose gel in 1x TAE buffer and run alongside a 1kb ladder (Hyperladder 1, Bioline).

Before fusing PCR products and plasmids the samples were run on a 1% gel for gel extraction. The products were extracted by using the Gen Elute™ Gel Extraction Kit (Sigma-Aldrich) and by following the protocol provided by the same kit

3.1.1.3 Digestion and ligation of spacers and plasmids – Golden Gate Reaction

Two different pCambia plasmids were used in the Golden Gate reaction; pHSE401 (Addgene, Plasmid #62201) and pHEE401 (Addgene, Plasmid #71286). pHSE041 was digested and ligated with spacers to make sgRNAs targeting *PSY2L* and dual targeting *PSY2L 1* and *2*, while pHEE401 was used to make sgRNAs targeting *PAP7*, *PLL3* and *PLL2*.

The extracted PCR products and plasmids where digested and ligated in the same reaction by Golden Gate Reaction. The reaction was set up as in Xing et al. (2014) with some minor modifications (Table 8).

Table 8: Golden Gate reaction setup for assembly of expression cassettes with plasmids.

Component:	**Volume (µl):**	**Reaction conditions:**
Purified PCR fragment (~100 ng/µl)	2	3 hours at 37°C 5 minutes at 50°C 10 minutes at 80°C
Plasmid: pHSE401 or pHEE401 (~100 ng/µl)	2	
10X T4 DNA Ligase Buffer (Promega)	1.5	
10X BSA	1.5	
Bsa1 – HF (NEB)	1	
T4 DNA Ligase (Promega)	1	
Water, nuclease free	6	
	Total volume: 15	

3.1.2 Transformation of competent *E. coli* JM 109 cells

Competent *E. coli* JM109 cells were transformed by following the protocol "One-step preparation of competent *E. coli*: transformation and storage of bacterial cells in the same solution. " (Chung et al. 1989).

Approximately ⅓ of the product from the Golden Gate reaction was added and mixed into ice cold competent *E. coli* JM109 cells (~150-200 µl). The cells were incubated on ice for 30 minutes before given a heat shock at 42°C for 50 seconds. After heat shocking, the cells were again incubated on ice for 2 minutes and then LB broth (500 µl) was added. The tubes, containing cells and LB broth, were incubated on a shaker at 37°C for approximately two hours.

After incubation, overnight cultures were made by plating the *E. coli* cells (500 µl) on LB agar plates containing kanamycin (50 µg/ml). The plates were incubated over night at 37°C.

3.1.2.1 Colony PCR of transformed E. coli

Two colonies from each plate were chosen to be used in a colony PCR**.** The colony PCR was set up by using the Dream Taq DNA Polymerase protocol (Thermo Scientific) (Table 9 and 10). The primers are listed in Table 25 in Appendix A-4.

Table 9: Setup for colony PCR of transformed E. coli.

Components:	**Volume (µl):**
10X Dream Taq Buffer (Thermo Scientific)	2
dNTP Mix (Bioline)	2
Forward Primer (CRISPR_39_U6-26p-F (10 µM))	2
Reverse Primer (CRISPR_42_U6-29p-R (10 µM))	2
Bacteria	1 colony
Dream Taq DNA Polymerase (500 U, Thermo Scientific)	0.2
Water, nuclease free	11.8
	Total Volume: 20

Table 10: Cycling conditions for the colony PCR of the transformed E. coli

Step:	**Temperature:**	**Time:**	**Number of cycles:**
Initial denaturation	95	5 min	1
Denaturation	95	30 s	40
Annealing	60	30 s	
Extension	72	1 min and 20 s	
Final extension	72	10 min	1

Two controls were made by replacing the bacteria colony in the protocol with a 1/20 solution of the plasmids (1 µl, ~100 ng/µl) pHSE401 and pHEE401.

The size of the products was checked by gel electrophorese. The colony PCR products were loaded onto a 1% agarose gel together with a 1 kb ladder (Hyperladder 1, Bioline) and the two controls.

3.1.2.2 Overnight cultures of transformed E. coli

Based on the results of the gel electrophorese, one of the colonies used for colony PCR from each plate was chosen to make overnight cultures. Each colony was added to a tube containing LB broth (5 ml) with kanamycin (50 µg/ml). The tubes containing bacteria was incubated overnight on a shaker at 37°C.

3.1.3 Plasmid isolation

Approximately 4.5 ml of each overnight culture was transferred to Eppendorf-tubes and spun down. The supernatant was removed for each round in the centrifuge.

Plasmids from the pellet were isolated by using a Plasmid DNA purification kit (Macherey-Nagels). The Nucleospin® Plasmid (NoLid) protocol was followed.

3.1.4 Sequencing

To check if the correct spacer-sequences for each target gene had been inserted correctly into the plasmids, the plasmids had to be sent for sequencing.

The isolated plasmids were prepared according to demands of SegLab in Göttingen, Germany, and sent for sequencing.

3.1.5 Transformation of competent *Agrobacterium* Gv3101 cells

Agrobacterium strain Gv3101 was made competent and transformed by following the freeze-thaw protocol of Wang (2006) with minor modifications. To make the cells competent, overnight culture of *Agrobacterium* (4 ml) was transferred to LB broth (100 ml) containing rifampicin (25 µg/ml) and gentamicin (10 µg/ml). The culture was incubated at 28°C for 3-4 hours until OD600 reached 0.5. The culture was then chilled on ice before centrifuged at 2500 RPM in 4°C for 10 minutes. The supernatant was discarded and the cells were gently re-suspended in 2 ml $CaCl_2$ (20 mM).

The competent *Agrobacterium* cells were then transformed by adding plasmids containing sgRNAs (1 µg). Two controls were made by transforming cells with pHSE401 and pHEE401 plasmids without sgRNAs. The mixture of cells and plasmids were immediately frozen in liquid nitrogen. After freezing, the mixtures were given a "heatshock" by thawing them in a 37°C waterbath for 15 minutes. The tubes were then added LB broth (500 µl) and incubated at 28°C for 3.5 hours. After incubation, the cells were spread on LB agar plates containing kanamycin (50 µg/ml), rifampicin (25 µg/ml) and gentamicin (10 µg/ml) and then incubated at 28°C for 48 hours.

Two colonies from each plate were chosen to make overnight cultures of the transformed *Agrobacterium*. Each colony was transferred to LB broth (5 ml) containing the same antibiotics, in the same amounts, as when plated. The overnight cultures were incubated at 28°C.

3.1.5.1 Colony PCR of transformed Agrobacterium

The overnight cultures were prepared for a colony PCR by centrifuging a small amount (150 µl) of each culture (14.8 rpm, 1 min) and throwing away the supernatant. The cells were then re-suspended in nuclease free water (20 µl) and transferred to PCR tubes. The samples were heated (95°C, 10 min) and then spun down (5 min). The supernatant from these samples were used in the colony PCR reaction.

The colony PCR reaction was set up as described in Table 11 and 12.

Table 11: Setup for colony PCR reaction of transformed Agrobacterium.

Components:	**Volume (µl):**
10X Dream Taq Buffer (Thermo Scientific)	2
dNTP Mix (Bioline)	2
Forward Primer (CRISPR_39_U6-26p-F (10 µM))	2
Reverse Primer (CRISPR_42_U6-29p-R (10 µM), CRISPR_40 for the controls)	2
Bacteria (supernatant)	2
Dream Taq DNA Polymerase (500 U, Thermo Scientific)	0.2
Water, nuclease free	9.8
	Total Volume: 20

Table 12: Cycling conditions for the colony PCR of transformed Agrobacterium.

Step:	**Temperature:**	**Time:**	**Number of cycles:**
Initial denaturation	95	5 min	1
Denaturation	95	30 s	40
Annealing	60	30 s	
Extension	72	30 s	
Final extension	72	10 min	1

The size of the colony PCR products was checked by gel electrophorese. The samples were loaded on a 1% agarose gel in addition to a ladder (Hyperladder 1, Bioline)

3.1.5.2 *Glycerol stocks of transformed Agrobacterium*

Each overnight culture was used to make glycerol stocks (1ml overnight culture + 500 µl 70% glycerol) to be stored at -80^{o}C for later use.

3.1.6 Cultivation of *Arabidopsis* plants

Arabidopsis Col-0 seeds were sown directly on soil. Approximately 4-6 seeds were sown in each pot and transferred directly into plantroom (16 h light) to grow. After approximately 6 weeks, when the plants had grown shoots and started flowering, the colts were cut to promote more shoots. The plants were ready for dipping in the early stages of flowering.

3.1.7 *Agrobacterium* mediated transformation of *Arabidopsis* by flower dipping

The glycerol stocks of successfully transformed *Agrobacterium* were further used to transform *Arabidopsis* plants. The plants were transformed as described in Clough and Bent (1998), with minor modifications.

3.1.7.1 Overnight cultures for dipping medium

Overnight cultures containing LB broth (5 ml), glycerol stock (approximately 15 µl) and the antibiotics kanamycin (50 µg/ml), gentamicin (10 µg/ml) and rifampicin (25 µg/ml) were made and set for incubating at 28°C on a shaker. The next day new overnight cultures were made by inoculating the overnight cultures (0.5 ml) from the day before to Erlenmeyer flasks containing LB broth (100 ml) and the same antibiotics, kanamycin (50 µg/ml), gentamicin (6 µg/ml) and rifampicin (25 µg/ml). The cultures were put on a shaker at 28°C overnight.

The next day the cultures were spun down (4000 rpm, 15 min), supernatant removed and the cells were re-suspended in dipping media (Table 13).

Table 13: Dipping medium for transform Arabidopsis plants.

Dipping media (1 L):	
Sucrose	50g
$MgCl_2$ x $6H_2O$	2.03g
Silwet	100 µl (added prior to use)

3.1.7.2 Dipping of plants

Two months old, flowering *Arabidopsis* plants, containing several stems of flowers, were then placed upside down in the dipping-medium containing transformed *Agrobacterium* (Figure 3). The plants were soaked in the dipping-medium for 20 minutes before they were removed, covered with a plastic bag and stored/put on the side overnight. Two plants each were dipped for the solutions containing sgRNAs dual targeting *PSY2L 1* and *2*, targeting *PAP7, PLL3, PLL2* and dual targeting *PLL2* and *PLL3*. The next day, the plastic bags were removed, the plants were placed standing straight and treated as normal.

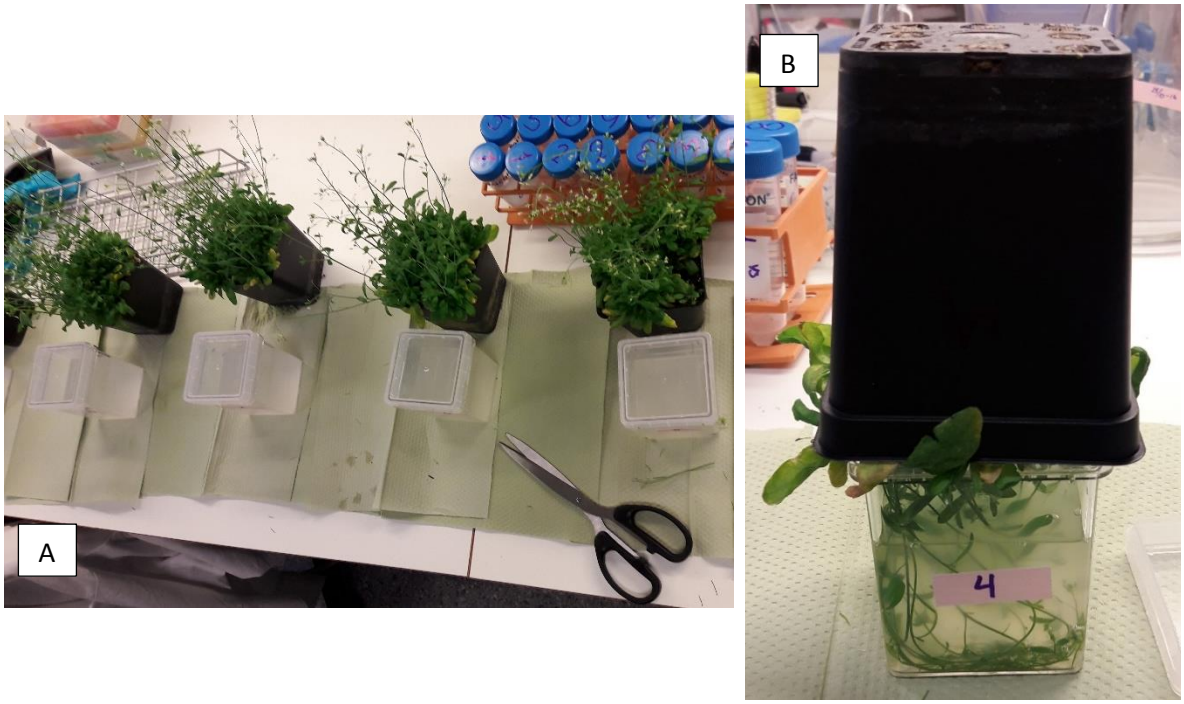


Figure 3: Dipping of Arabidopsis thaliana in dipping media containing transformed Agrobacterium.
Each plant was dipped in separate containers, containing transformed Agrobacterium in dipping-medium (A). When dipping the plant, each plant was turned upside-down and placed in the container in such a manner that as many of the plants flowers were covered with media. The plants were soaked in 20 minutes before they were removed.

3.1.8 Screening for transformed *Arabidopsis* plants

Approximately three weeks after dipping, the first seeds from the plants were harvested. Seeds from each plant were harvested every other day for about a week. After harvesting seeds for about a week, the plants were put in a drying room to dry out and prevent flowering.

For the first screening, the seeds from the first harvest from each plant were used. Seeds from each transformation were sterilized.

For seed sterilization, a 70% ethanol and 0.01% Triton solution (1 ml) solution were added to each Eppendorf tube containing seeds. The tubes were then put on a shaker for 15 minutes. The 70% ethanol, 0.01% Triton solution was removed and replaced with 99.5% ethanol (1 ml). The tubes were then put on a shaker for 10 minutes. The ethanol was removed and replaced with another 1 ml 99.5% ethanol. The tubes were flicked, the ethanol removed and the seeds were left in the sterile hood for drying.

Sterilization protocol (carried out in sterile hood):

1. A solution of 70 % ethanol + 0.01 % Triton (1 ml) was added to each Eppendorf tube containing seeds. The tubes were then put on shaker for 15 minutes.

2. Removed the solution
3. Added 99.5% ethanol to each tube and put in shaker for 10 minutes.
4. Removed solution
5. Repeated step 3-4 without shaking for 10 minutes (only added, flipped and removed)
6. The seed were left for drying in the sterile hood for a couple of hours.

After sterilization, the seeds were sown on separate plates containing ½ Murashige and Skoog medium (MS)(Table 14) (Murashige and Skoog 1962). A pH of 5.8 was used. Since both plasmids (pHSE401 and pHEE401) contained hygromycin resistance, hygromycin (15 µg/ml) was added to the medium to distinguish transformed plants from non-transformed plants.

Table 14: ½ MS medium for screening of transformed A. thaliana plants.
pH was adjusted to 5.8, the agar-agar was added and then the medium was autoclaved

Stock solutions:			**½ MS (1L)**
KNO_3		(95 g/l)	10 ml
NH_4NO_3		(120 g/l)	6.5 ml
$MgSO_4$ x $7H_2O$		(37 g/l)	5 ml
$KH2PO_4$		(17 g/l)	10 ml
$CaCl_2$ x $2H_2O$		(44 g/l)	5 ml
Fe/EDTA (1L)	Na_2 x EDTA	0.373 g	25 ml
	$FeSO_4$ x$7H_2O$	0.278 g	
Minor 1 (1L)	$ZnSO_4$ x $7H_2O$	0.920 g	5 ml
	H_3BO_3	0.620 g	
	$MnSO_4$ x $4H_2O$	2.230 g	
Minor 2 (1L)	Na_2MoO_4 x $2H_2O$	0.025 g	5 ml
	$CuSO_4$ x $5H_2O$	0.003 g	
	$CoCl_2$ x $6H_2O$	0.003 g	
	KI	0.083 g	
Sucrose		0.5 %	5 g
Distilled H_2O			Up to 1L
Agar-Agar			7 g

Three controls were made; one containing seeds from plants transformed with non-recombinant pHSE401 plasmids, one with non-recombinant pHEE401 plasmids and one containing seeds from WT *Arabidopsis*. The plates were then placed in a dark and cold room for 24 hours before placed in plant room (16 h light/8h dark). After approximately two weeks of growing, the plates were screened for transformed plants.

After two weeks of drying, the last seeds were harvested from the transformed plants. Another screen was done with the seeds harvested from dried out plants. The same medium and the same type and amount for antibiotic as for the previous screening were used.

Seeds harvested before drying were mixed and used for a third screening. The same concentration as before was used for the hygromycin.

A fourth screening was made with the seeds harvested after drying. This time the concentration of hygromycin in the medium was adjusted to 25 µg/ml to make it easier to distinguish the positive seedlings from the negative ones.

After approximately two weeks of growing in plant room (16-hour light), the positive seedlings from the screenings were transferred to soil and placed in plant room (16-hour light) to grow leaves big enough for DNA isolation.

3.1.9 T7 endonuclease 1 assay to check for heterozygous plants in the T1 generation

Only five plants from each transformation was chosen for the DNA extraction and the following T7 endonuclease 1 assay. Plants showing some sort of phenotype was chosen over others. The age of the plants used for the extraction varied.

3.1.9.1 DNA extraction of transformed Arabidopsis plants

The DNA extraction of leaves from positive seedlings transferred to soil was done by following Cold Spring Harbor Protocols protocol; "Isolation of Plant DNA for PCR and Genotyping Using Organic Extraction and CTAB" with minor modifications (Springer 2010). In lack of a freeze dryer, the plant material was frozen in liquid nitrogen. Instead of a Retsch 300 matrix mill, each sample was grinded using a manual mill.

The CTAB plant extraction buffer was made as instructed by Cold Spring Harbor Protocols (Springer 2010)

Before running PCR with specifically designed primers, the isolated DNA was run on 1% agarose gel in 1x TAE buffer to check for presence of genomic DNA.

3.1.9.2 PCR of genomic DNA

Primers for the PCR reaction were designed by using CHOPCHOP.

Primers used in the PCR reaction and the expected sizes of the amplicons can be seen in Table 15. The PCR was set up as described in Table 16 and run with the cycling conditions as shown in Table 17.

Table 15: Primers used for T7E1 assay PCR.

Gene	Primer Name		Primer sequence 5'-3'	PCR amplicon size		sgRNA
PAP7	CRISPR_43	F	AAACTGAAAAATTGGCAGATGG	176	Spacer 1	GAGAAAGAGGCTTCAAAGGAAGG
	CRISPR_44	R	CCTGAATACCATTGTTTTTGGAG			
	CRISPR_45	F	ATGGTTTTGTCGCAGATCTTTT	196	Spacer 2	TGTTGTGGTTATAGGAATGGTGG
	CRISPR_46	R	TGTGATCTTCTGGTTCAGTGAAGT			
PLL3	CRISPR_47	F	CGATACGTAACCGGAGTAGGAC	187	Spacer 1	GTCGGAGGGGCGGTTCCGGCGGG
	CRISPR_48	R	TTGAAGTATTGGCGCTAACAGA			
	CRISPR_49	F	AGGCGCTTCTAGAGATGTTCAG	288	Spacer 2	AGTCTATGGTGGTGTAGCGACGG
	CRISPR_50	R	ATCTTAAGTGCGTTTGCGTTTT			
PLL2	CRISPR_51	F	TCTGTCATCGGATTCTGATTGT	246	Spacer 1	GGAGCTCTCGAATGCGGAGGCGG
	CRISPR_52	R	TTCTCTTTCTCCGTCTTCTTCG			
	CRISPR_53	F	ATTTTGCCTCTTTACCACTCCA	247	Spacer 1	GATTGAAAGCGGGTTGTATTCGG
	CRISPR_54	R	TTCGCGAATAGTGTTTTGAATG			

Table 16: Setup for PCR reaction of extracted DNA from plants grown in soil after screening

Components:	**Amounts:**
10X Dream Taq Buffer (Thermo Scientific)	5 µl
dNTP Mix (Bioline)	5 µl
Forward Primer (50 µM))	1 µl
Reverse Primer (50 µM))	1 µl
Template DNA	~100 ng
Dream Taq DNA Polymerase (5 U/µl, Thermo Scientific)	0.25 µl
Water, nuclease free	Up to 50 µl
	Total Volume: 50 µl

Table 17: Cycling conditions for PCR reaction of DNA extracted from screened plants grown in soil

Step:	**Temperature**	**Time**	**Number of cycles**
Initial denaturation	95	5 min	1
Denaturation	95	30 s	40
Annealing	60	30 s	
Extension	72	1 min	
Final extension	72	10 min	1

3.1.9.3 Gel electrophoresis and gel extraction of PCR product

The PCR products (5 µl) were run on a 2% agarose gel in 1x TAE buffer together with a 100 bp ladder (NEB Quick Load) to check size and quality. The rest of the product (25 µl) were then run on a 1 % agarose gel for gel extraction. The bands were cut out and DNA was extracted using a gel extraction and PCR clean up kit (Macherey-Nagel) and by following the gel extraction protocol with minor modification; instead of Buffer NE, nuclease free water was used to elute the samples.

3.1.9.4 T7E1 assay

The T7E1 assay was set up as described in the protocol "Determining Genome Targeting Efficiency using T7 Endonuclease I (M0302)" (NEB 2016) with minor modifications. Approximately 200 ng product from the gel extraction was mixed together with a 10X NEB buffer 2 (2µl) and nuclease free water (up to 19 µl). The mixtures were run through a hybridization reaction in a PCR cycler (Table 18)

Table 18: Hybridization reaction for T7 endonuclease 1 assay.

Temperature:	**Time:**
95°C	5 min
↓	-2°C/s
85°C	
↓	-0.1°C/s
25°C	
↓	
4°C	Infinite

After hybridization, T7 endonuclease 1 (1 µl) was added to each sample. The samples were incubated at 37°C for 15 minutes and loaded directly on a 2% agarose gel after incubation.

A second control was made by extracting DNA from wild type *Arabidopsis*, run a PCR with the PAP7, PLL3 and PLL2 primers, gel extraction and finally the T7 endonuclease 1 assay.

3.1.10 Sterilization, sowing and cultivation of *S. lycopersicum*

For sterilization of seeds, a 1% Ca-hypochlorite solution (25 ml) was made and added one drop Tween 20. The solution was shaken and left to settle. The supernatant (1 ml) of the settled Ca-hypochlorite/Tween 20 solution was then added to 9 ml ethanol (96 %).

The sterilization was performed in a sterile hood. Ten seeds were placed in Eppendorf tubes (five in each tube) and added the ethanol/Ca-hypochlorite/Tween20 solution. The tubes were shaken and then incubated for maximum 5 minutes before the solution was pipetted off. The seeds were then rinsed with 1 ml ethanol (96%). Rinsing with ethanol was repeated three times. The tubes with seeds were then left in the sterile hood overnight for drying.

The seeds were then sown on ½ MS plates without sucrose (Table 19). Two plates were made, with five seeds in each. The plates were sealed and directly placed in plant room (16 h light)

Table 19: ½ MS medium for cultivation of S. lycopersicum seeds.
pH was adjusted to 5.8, agar-agar was added and the medium was autoclaved

Stock solutions:			**½ MS (1L)**
KNO_3		(95 g/l)	10 ml
NH_4NO_3		(120 g/l)	6.5 ml
$MgSO_4$ x $7H_2O$		(37 g/l)	5 ml
KH2PO_4		(17 g/l)	10 ml
$CaCl_2$ x $2H_2O$		(44 g/l)	5 ml
Fe/EDTA (1L)	Na_2 x EDTA	0.373 g	25 ml
	$FeSO_4$ x$7H_2O$	0.278 g	
Minor 1 (1L)	$ZnSO_4$ x $7H_2O$	0.920 g	5 ml
	H_3BO_3	0.620 g	
	$MnSO_4$ x $4H_2O$	2.230 g	
Minor 2 (1L)	Na_2MoO_4 x $2H_2O$	0.025 g	5 ml
	$CuSO_4$ x $5H_2O$	0.003 g	
	$CoCl_2$ x $6H_2O$	0.003 g	
	KI	0.083 g	
pH			5.8
Distilled H_2O			Up to 1L
Agar-Agar			7 g

3.1.11 Transformation of *S. lycopersicum*

Overnight cultures of agrobacterium were made by adding glycerol stocks of transformed *Agrobacterium* (15 µl) containing pHSE401 plasmids with sgRNAs targeting *PSY2L 1*, *PSY2L 1* and *2* to LB broth (5 ml) containing kanamycin (50 µg/ml), gentamycin (10 µg/ml) and rifampicin (25 µg/ml). The cultures were incubated on shaker at 28°C overnight.

The transformation of the tomato cotyledons was done as described in McCormick's "Transformation of tomato with *Agrobacterium tumefaciens*" (McCormick 1997), with minor modifications. No acetosyringone was added to the 20-fold diluted *Agrobacterium* used for transforming the cotyledons. For the selection medium (D1 medium), hygromycin (25µg/ml) were used to select transformed cells, and carbenicillin (100 µg/ml) to kill the *Agrobacterium* without effecting the plants.

3.2 In-vivo CRISPR/Cas9 modifications through PEG transfections of isolated *S. lycopersicum* and *Arabidopsis* protoplasts by recombinant vectors.

3.2.1 Hybridization of complementary oligonucleotides

Complementary oligonucleotides for spacers targeting *PSY2L 1, PSY2L 1/2, PAP7, PLL3* and *PLL2* were made by mixing the primers shown in Table 20.

Two different methods were tested to anneal the complementary oligonucleotides, method A (Ran et al. 2013) and method B (Huang 2014) (Table 21)

Table 20: Overview of reactions with target, primer name and primer sequence. The red nucleotides represent the overhangs on each primer.

Reaction number	Target	Primer name		Primer sequence with oligos
1	PSY2L 1	CRISPR_29	T1/PSY-1/F	ATTGGTCTATCGCCTGAATGACGA
		CRISPR_30	T1/PSY-1/R	AAACTCGTCATTCAGGCGATAGAC
2	PSY2L 1/2	CRISPR_31	T2/PSY-2/1/F	ATTGTCTCCTGAAATGGCTGTAAA
		CRISPR_32	T2/PSY-2/1/R	AAACTTTACAGCCATTTCAGGAGA
3	PAP7	CRISPR_33	T1/PAP7/F	ATTGGAGAAAGAGGCTTCAAAGGA
		CRISPR_34	T1/PAP7/R	AAACTCCTTTGAAGCCTCTTTCTC
4	PLL3	CRISPR_35	T1/PLL3/F	ATTGGTCGGAGGGGCGGTTCCGGC
		CRISPR_36	T1/PLL3/R	AAACGCCGGAACCGCCCCTCCGAC
5	PLL2	CRISPR_37	T2/PLL2/F	ATTGGATTGAAAGCGGGTTGTATT
		CRISPR_38	T2/PLL2/R	AAACAATACAACCCGCTTTCAATC

Table 21: Setup for the annealing of complementary oligonucleotides

Method A		
Components:	**Volume (µl):**	**Reaction conditions**
Forward Primer (100 µM)	1	30 min at 37°C 5 min at 95°C RamP down 5°C/min 25°C
Reverse Primer (100 µM)	1	
10X T4 Ligation Buffer (Promega)	1	
PNK	1	
Water, nuclease free	6	
Total volume:	10	
Method B		
Components	**Volume**	**Reaction conditions**
Forward Primer (100 µM)	1	5 min at 95°C RamP down 5°C/min 25°C
Reverse Primer (100 µM)	1	
*10X annealing buffer	5	
Water, nuclease free	43	
Total volume:	50	

*Tris (10mM, pH 8.0), NaCl (50 mM), EDTA (1 mM)

In the first two reactions, making spacers targeting *PSY2L 1* and dual targeting *PSY2L 1* and *2*, both methods were used to anneal the oligonucleotides. Only method B was used to anneal the oligonucleotides for the last three spacers, targeting *PAP7*, *PLL3* and *PLL2*.

After annealing the oligonucleotides, the oligonucleotides and the pGREEN-like binary vector pHSN401 (Addgene plasmid #50588) were digested and ligated in the same reaction. The setup was similar to the Golden Gate reaction used for digestion and ligation of the expression cassette (Table 22).

Table 22: Setup for the restriction-ligation-reaction of the annealed oligonucleotides with pHSN401 plasmid (Addgene)

Components:	**Volume (µl):**	**Reaction conditions:**
Diluted Oligonucleotides	2	3 hours at 37°C 30 minutes at 70°C
pHSN401 plasmid (100 ng)	1	
10X T4 Ligase Buffer (Promega)	2	
10X BSA (NEB)	2	
T4 DNA Ligase (Promega)	0.75	
BsaI	1	
Water, nucleasefree	11.25	
Total Volume:	20	

Instead of adding purified PCR fragments, the annealed oligonucleotides were added to the reaction. Before adding the oligonucleotides to the reaction, the oligonucleotides made with reaction A were diluted 1/200 and the oligonucleotides made with reaction B was diluted 1/100.

A control containing non-recombinant pHSN401 plasmid was made by adding nuclease-free water (2 µl) instead of diluted oligonucleotides to the reaction.

3.2.2 Transformation of competent *E. coli* JM109 cells.

Competent *E. coli* JM109 cells were transformed as described in methods section 3.1.2. Instead of adding ⅓ of the product from the digestion and ligation step, ½ of the product was added to the competent bacterial cells.

3.2.3 Colony PCR

After the transformed *E. coli* cells had incubated on LB agar plates containing kanamycin (50 µg/ml) at 37°C for 24 hours, a colony PCR was performed. Three colonies from each plate were chosen to make the PCR. The colony PCR was set up by using the Dream Taq DNA Polymerase protocol (Thermo Scientific) (Table 23)

Table 23: Setup for colony PCR of transformed E. coli containing plasmids with spacers

Components:	**Volume (µl):**
10X Dream Taq Buffer (Thermo Scientific)	2
dNTP Mix (Bioline)	2
*Forward Primer (10 µM)	2
Reverse Primer (CRISPR_40_U6-26t-R (10 µM))	2
Bacteria	1 colony
Dream Taq DNA Polymerase (500 U, Thermo Scientific)	0.2
Water, nuclease free	11.8
Total Volume:	20

*For PSY2L 1: CRISPR_29, For PSY2L 1 and 2: CRISPR_31, For PAP7: CRISPR_33, For PLL3: CRISPR_35, For PLL2: CRISPR_37.

The forward primers are listed in table 20, while the reverse primer sequence is listed in Table 25 in Appendix A-4.

A control containing plasmid only was made by adding nuclease-free water instead of the bacterial colony.

To check the size of the colony PCR products the products were run on a 2% agarose gel in 1xTAE - buffer together with a 100 bp DNA ladder (NEB).

3.2.4 Plasmid isolation and sequencing

To check if the correct spacer-sequence for each target gene had been inserted into the plasmids, the plasmids had to be sent for sequencing.

Based on the results of the gel electrophoresis, one of the colonies from each plate used for colony PCR was chosen to make overnight cultures. Each colony was added to a tube containing LB broth (5 ml) and kanamycin (50 μg/ml) and incubated overnight on a shaker at 37°C.

Plasmids were then isolated by using a kit for Plasmid DNA purification (Macherey-Nagel). The Nucleospin® Plasmid (NoLid) protocol was followed. The plasmids were then prepared according to Seqlab's demands and sent for sequencing at Seqlab in Göttingen, Germany.

4 Results

4.1 CRISPR/Cas9 editing of *Arabidopsis thaliana* PAP7, PLL3 and PLL2 and tomato *Solanum lycopersicum* PSY2L by *Agrobacterium* transformation

4.1.1 Expression cassettes with sgRNAs

To check if the expression cassettes for sgRNA targeting Arabidopsis *PAP7, PLL3, PLL2* and tomato *PSY2L*, using pCBC-DT1T2 as template (As described in methods section 3.1.1.2) had been successfully assembled, the PCR products were run on a 1% agarose gel. The size of the cassettes' PCR products was expected to be 626 bp (Xing et al. 2014). All seven PCR products gave bands with a size of approximately 600 bp, indicating successful assembly of expression cassettes containing two sgRNAs (Figure 4). Sample number two, containing spacers dual targeting PSY2L 1 and 2, gave a very week band due to the low concentration (13.8 ng/µl) of DNA. Even though the band was week it still appeared and in the same size as the rest of the products.

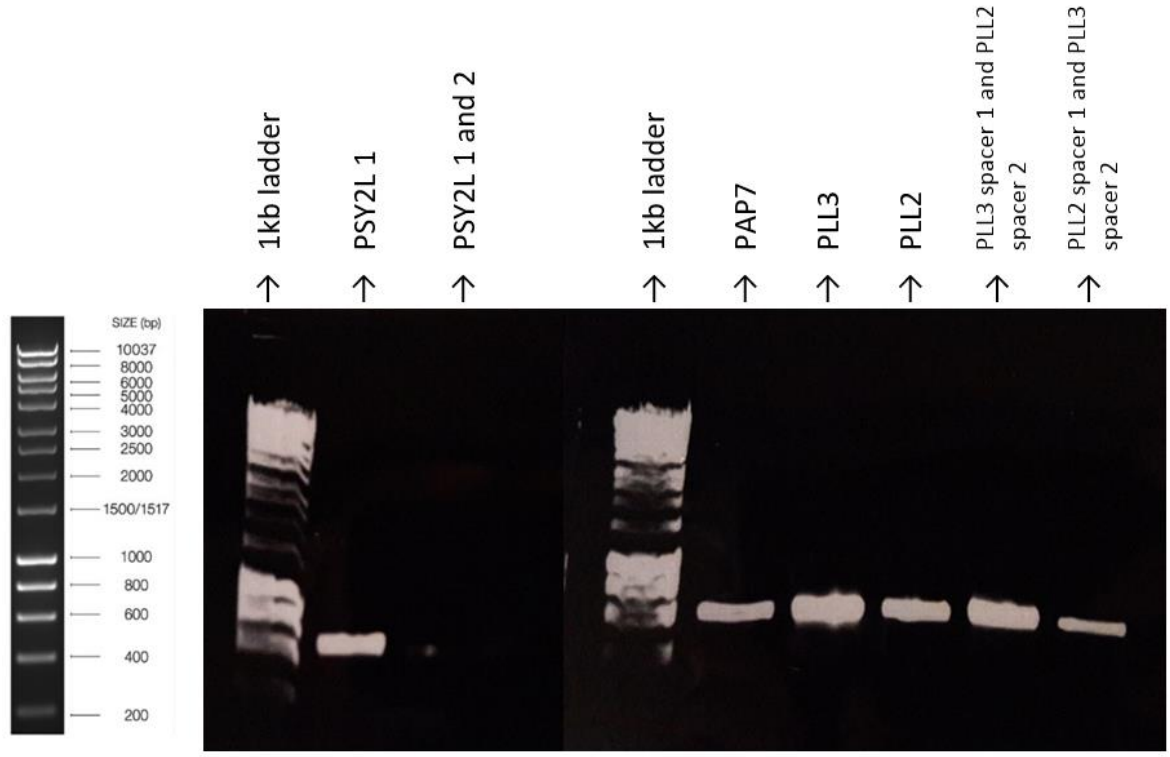


Figure 4: PCR products containing the spacers of two sgRNA run on 1% agarose gel with a 1kb ladder (Hyperladder 1, Bioline).
The products were yielded by using pCBC-DT1DT2 as template and specific primers for each reaction. As expected, each reaction gave a product of approximately 600 bp.

4.1.2 Verification of transformed *E. coli* JM109 cells

The PCR products containing sgRNA expression cassettes were gel extracted and ligated into their plasmids by Golden Gate reaction (as described in methods section 3.1.1.3). The plasmids were then transformed into competent *E. coli* cells for cloning. Since the plasmids were resistant to kanamycin when inserted into bacteria, the transformed *E. coli* cells were spread on LB agar plates containing kanamycin (50 µg/ml) and incubated at 37°C for 24 hours. After incubation, all samples were able to grow bacterial colonies (Figure 5A). This showed that bacterial cells in these colonies had successfully taken up plasmids containing kanamycin resistance.

4.1.2.1 Test on spectinomycin plate

Because the vectors were setup to excise the spectinomycin resistance gene and replace it with the sgRNAs cassettes (Xing et al. 2014), none of the transformed *E. coli* bacteria were able to grow on LB agar plate containing spectinomycin (50 µg/ml). This indicates successful cutting and shows that a sequence has been inserted in this area. In the controls, which only contained plasmids without insert of spacers, the bacteria were able to grow due to a complete spectinomycin resistance gene (Figure 5B).

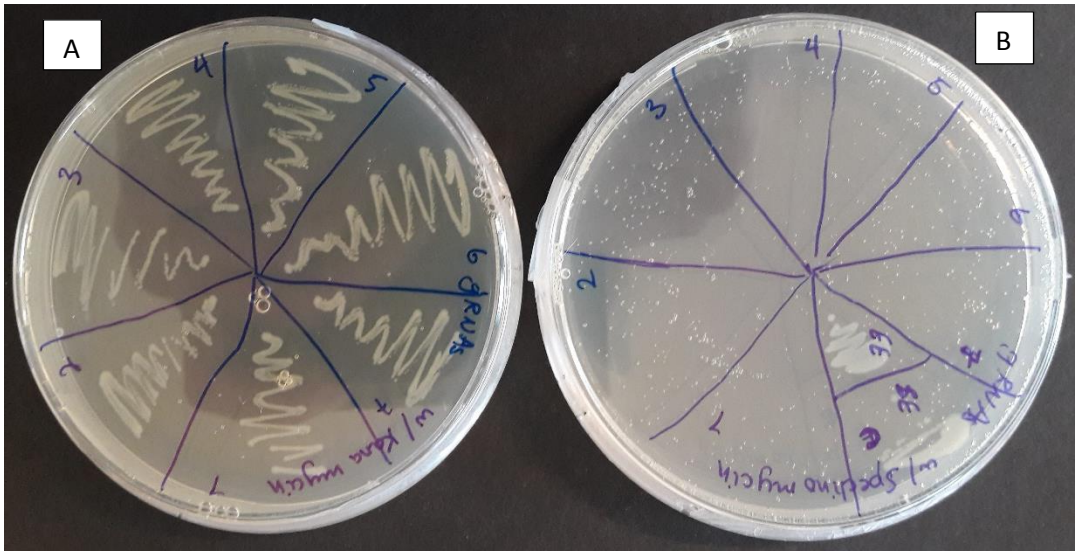


Figure 5:Transformed E. coli on LB agar plates with kanamycin and spectinomycin.
As a test, transformed cells from each reaction was stroked out on one plate containing LB agar with kanamycin (50 µg/ml) (A) and on one plate containing LB agar with spectinomycin (50 µg/ml) (B). As a control, plasmid pHSE401 and pHEE401 was stroked on the LB agar plate containing spectinomycin. As expected and due to kanamycin resistance, all colonies grew on the plate containing kanamycin, but no colonies (except from the controls) were able to grow on the plate containing spectinomycin. (The small dots throughout plate B are bubbles and not bacterial colonies)

4.1.3 Colony PCR products on 1% agarose gel

To be sure that the correct piece of DNA had been inserted in the spectinomycin resistance sequence of the plasmid, a colony PCR of two randomly selected colonies from each plate was made (Figure 6). If the correct piece had been inserted, the size of the products should be approximately 726 bp (Xing et al. 2014). All the PCR products from the colonies gave good and clear bands of approximately 700 bp (Blue arrow). The controls (Shown by red arrows), containing bacteria with plasmid without insert gave, as expected, no bands. Thereby we were able to successfully clone our new sgRNAs expression cassettes (mentioned in results section 4.1.1) into the binary vectors pHSE401 and pHEE401.

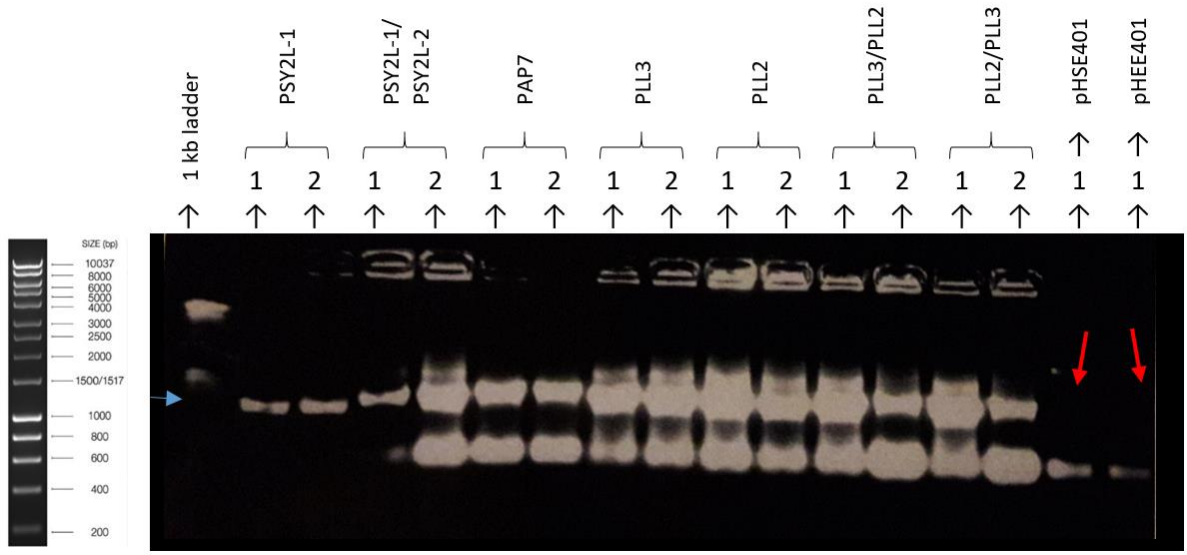


Figure 6: Colony PCR products run on 1% agarose gel.
Two colonies from each plate was used to make colony PCR. Two controls containing plasmid instead of bacteria-colony was also made. The controls had no sgRNA insert. The samples were run alongside a 1kb ladder (Hyperladder 1, Bioline). All of the products gave bands in the size of approximately 700bp indicating successful cloning.

4.1.4 Sequencing of plasmids isolated from colonies of transformed *E. coli*

In order to confirm successful insertion of sgRNA cassettes into the binary vectors (described above), all seven constructs were sent for sequencing. Successfully, sequencing proved correct subcloning and correct dual spacer integrations for all six constructs (Figure 7) but not for gRNA3 – pHEE401 (targeting PAP7).

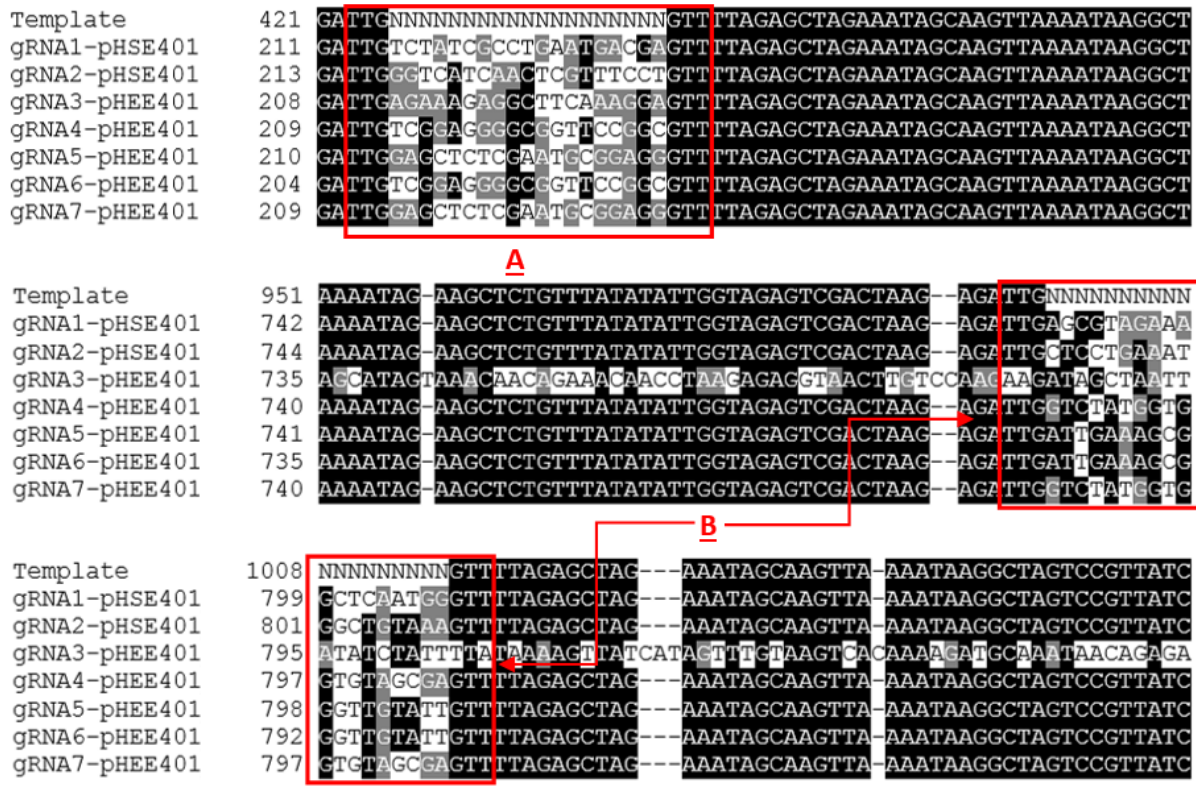


Figure 7: Alignment of sequenced plasmids containing spacers of two guide RNAs.
The sequences were aligned by Clustal W. (A) First spacer. (B) Second spacer. Template: "Sequence of two gRNA expression cassettes for dicots, DT1DT2-PCR + pHSN401 et al. sourced from Xing et al. 2014. gRNA1-pHSE401: sequence of spacers targeting PSY2L 1. gRNA2-pHSE401: sequence of spacers dual targeting PSY2L 1 and 2. gRNA3-pHEE401: sequence of spacers targeting PAP7. gRNA4-pHEE401: sequence of spacers targeting PLL3. gRNA5-pHEE401: sequence of spacers targeting PLL2. gRNA6-pHEE401: sequence of spacers dual targeting PLL3 and PLL2. gRNA7-pHEE401: sequence of spacers dual targeting PLL3 and PLL2.
All samples contained the correct nucleotide sequence of the spacer for gRNA one (A). All samples, except from sample number three contained the correct nucleotide sequence of the spacer for gRNA two (B). The first sequence in the figure is from the template, where N shows the location of the 19 specific targeting nucleotides for the gRNAs.

Since gRNA3-pHEE401 did not contain the correct spacer sequence for sgRNA number two, plasmids were isolated from four new colonies from the same plate. When sequenced as in figure 8, all samples, except from gRNA3-6-pHEE401 showed correct spacer sequence for the first sgRNA. When searching for the spacer sequence of the second sgRNA, only sample gRNA3-5b-pHEE401 and gRNA3-6b-pHEE401 contained the correct sequence.

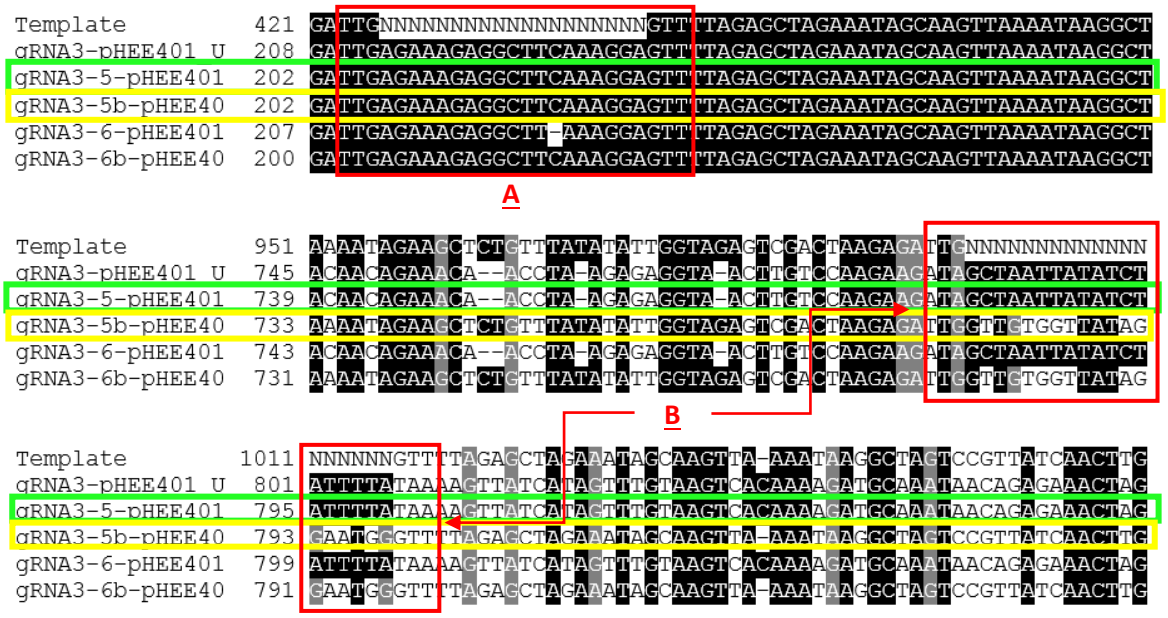


Figure 8: Alignment of sequenced plasmids from sample three.
The alignment was done by Clustal W. (A) Area of the first spacers. (B) Area of the second spacers. All plasmids, except from the plasmids in sample gRNA3-6-pHEE401, contained the correct spacer sequence for the first gRNA. Only the plasmids in sample gRNA3-5b-pHEE401 and gRNA3-6b-pHEE401 had correct spacer sequence for gRNA number two.

Both gRNA3-5b-pHEE401 (yellow in figure 8), which contained both spacers, and gRNA3-5-pHEE401 (green in figure 8), which only contained the spacer for the first sgRNA, were chosen to be further used in the experiment. The reason for using the sample which only contained one spacer was to see if it was possible to knock out *PAP7* with only one sgRNA present.

4.1.5 Transformation of *Agrobacterium* Gv3 101

The successful constructs of plasmids with sgRNAs were transformed into *Agrobacterium* Gv3 101 by the freeze-thaw method (se methods section 3.1.5). For selection of successfully transformed cells, the bacteria were spread on LB agar plates containing kanamycin (50 µg/ml), gentamicin (10 µg/ml), and rifampicin (25 µg/ml).

Large numbers of colonies were able to grow on the plate (Figure 9), showing successful transformation of the bacteria.

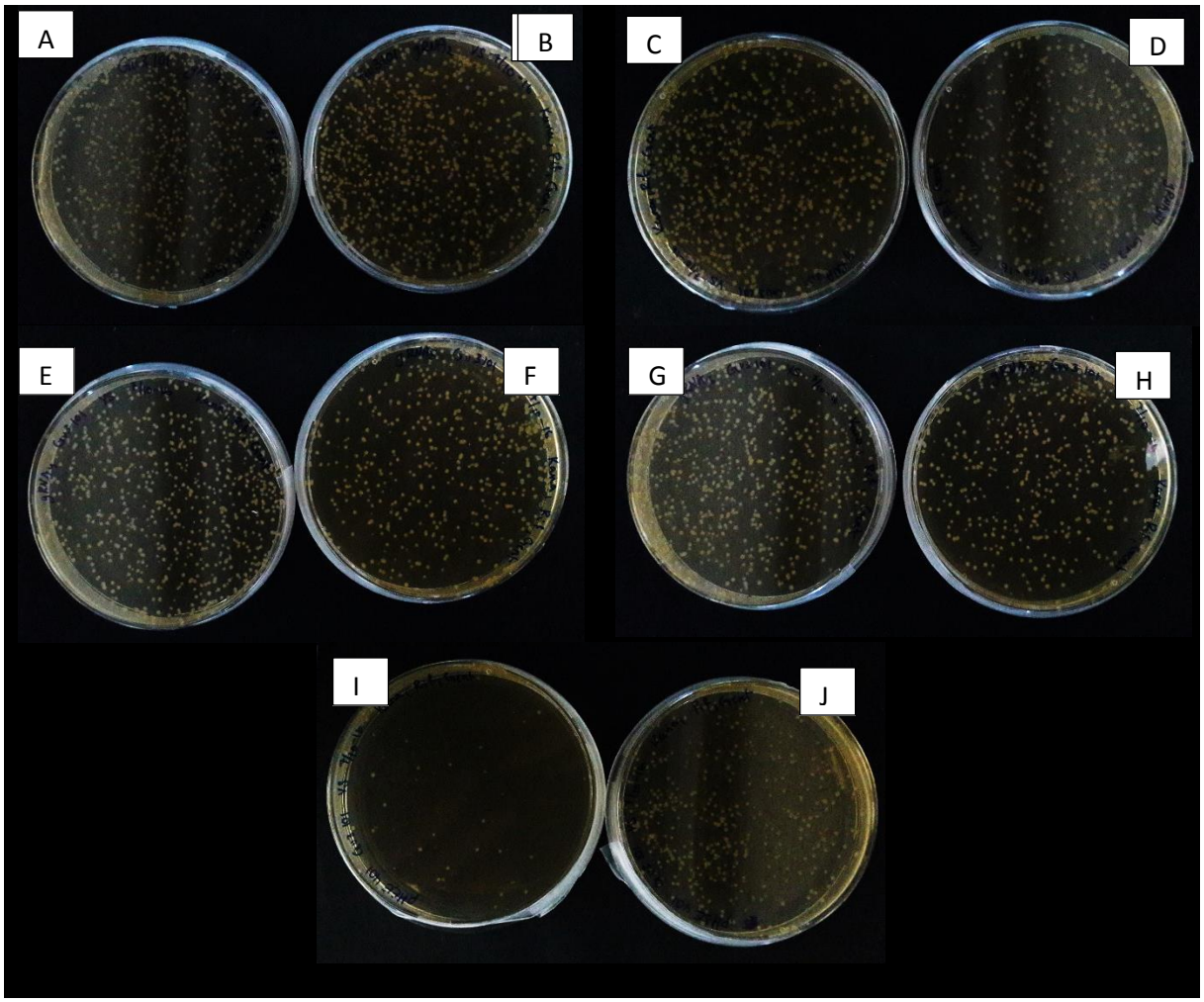


Figure 9: Transformed Agrobacterium Gv3 101 containing plasmids with sgRNAs and antibiotic resistance. Transformed Agrobacterium cells were plated on LB agar plates containing antibiotics, kanamycin (50 µg/ml), Gentamicin (10 µg/ml) and Rifampicin (25µg/ml). (A) Agrobacterium colonies containing two sgRNAs targeting S. Lycopersicum PSY2L 1. (B) Agrobacterium colonies containing two sgRNAs dual targeting S. Lycopersicum PSY2L 1 and 2. (C) Agrobacterium colonies containing two sgRNAs targeting Arabidopsis PAP7. (D) Agrobacterium colonies containing one sgRNA targeting Arabidopsis PAP7. (E) Agrobacterium colonies containing two sgRNAs targeting Arabidopsis PLL3. (F) Agrobacterium colonies containing two sgRNAs targeting Arabidopsis PLL2. (G) Agrobacterium colonies containing two sgRNAs dual targeting Arabidopsis PLL3 and PLL2. (H) Agrobacterium colonies containing two sgRNAs dual targeting Arabidopsis PLL3 and PLL2. (I) Agrobacterium colonies containing pHEE401 plasmid without sgRNAs. (J) Agrobacterium colonies containing pHSE401 plasmid without sgRNA

For more confirmation of successful transformation into *Agrobacterium* strain and to exclude any cross-contamination possibilities, we performed colony PCR using one of the spacer primers for the independent constructs in the reaction setup. When running a colony PCR of selected *Agrobacterium* colonies from each plate, all colony PCR products, except from the one containing only one sgRNA targeting *PAP7* and the controls containing pHSE401 and pHEE401 plasmids with no sgRNA, gave good and clear bands when running the products on gel (Figure 10). The bands showed, as expected, to be right above 700 bp in size (blue arrow). The products containing only one sgRNA targeting *PAP7* (PAP7 (One spacer)) and the controls containing pHSE401- and pHEE401 plasmids without sgRNA gave no bands due to the lack of a U6-29p promotor in the plasmids. Only the plasmids containing

two sgRNAs had this promotor. The band showing for the second colony of pHEE401 is assumed to be a contamination.

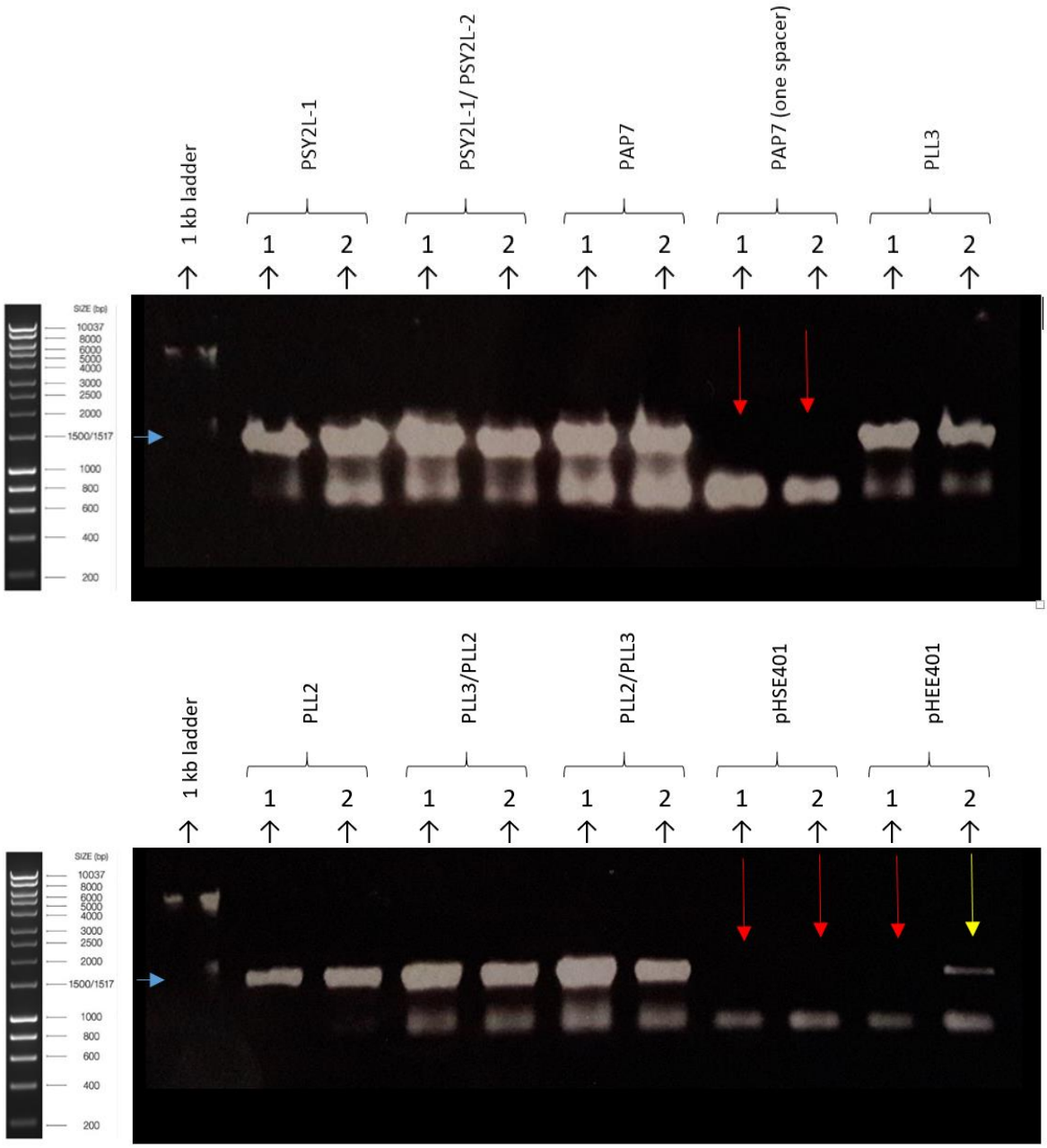


Figure 10: Colony PCR products of transformed Agrobacterium run on a 1% agarose gel.
A 1 kb ladder was run alongside the samples (Hyperladder 1, Bioline). The bacteria which had obtained recombinant plasmids containing two sgRNAs gave bands with the size of approximately 700 bp. The controls and the colonies containing recombinant plasmids with only on sgRNA targeting PAP7 gave no bands. Colony number two of the pHEE401 control showed a band which is considered as a contamination.

Since no bands in the correct size could be observed in the gel electrophoresis of the *Agrobacterium* containing one sgRNA targeting *PAP7* and for the controls, a new colony PCR was made for these samples. This time the U6-29p reverse primer was replaced by the U6-26t-reverse primer to be able

to get a product. This resulted in good and clear bands of approximately 1600 bp (Blue arrows) from the "empty" plasmids still containing the sequence for spectinomycin resistance and approximately 400 bp (Yellow arrows) from the plasmid with only one sgRNA (Figure 11). Thereby, we confirmed successful transformation of the constructs into *Agrobacterium.*

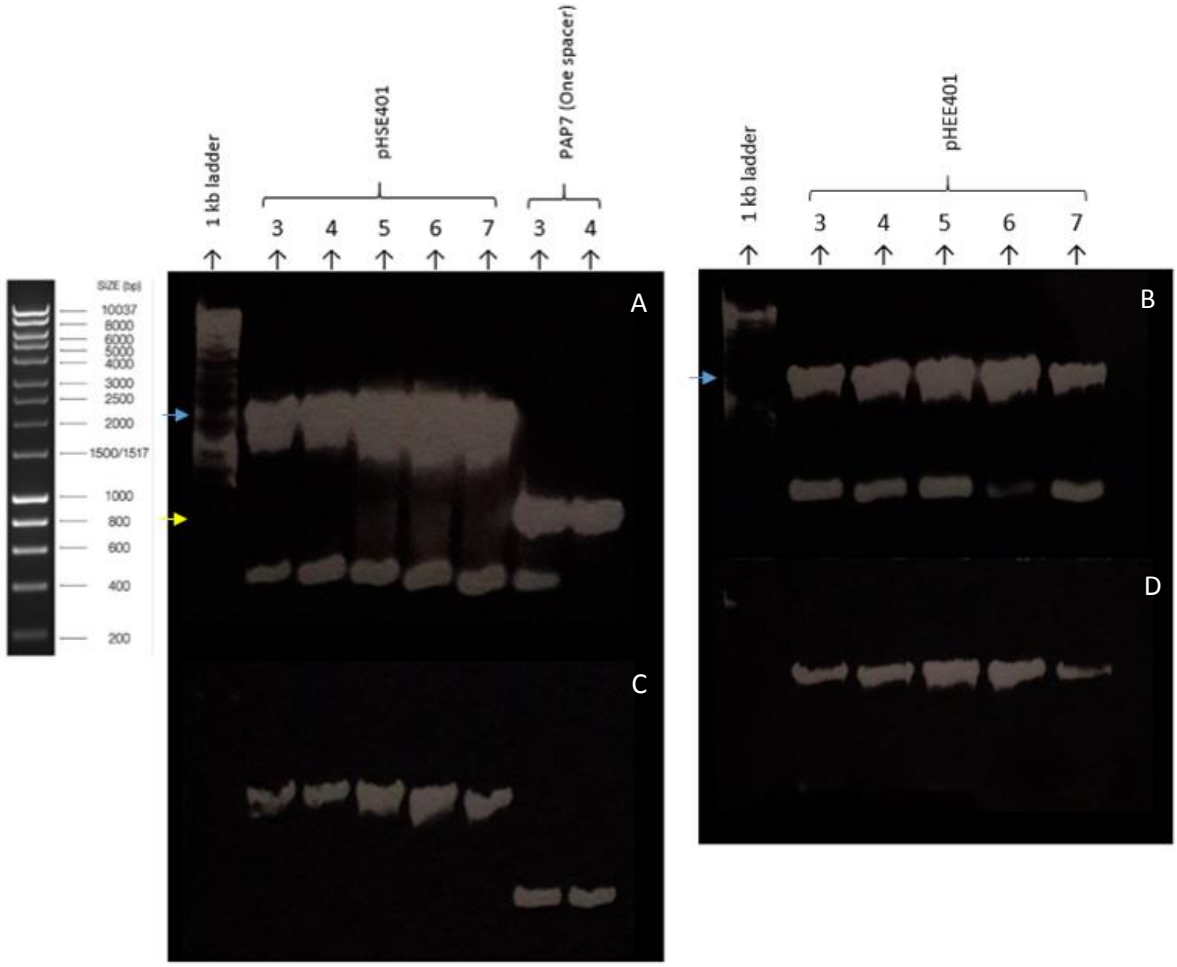


Figure 11: Colony PCR products of new colonies of transformed Agrobacterium containing pHSE401 without sgRNAs, sgRNA targeting PAP7 and pHEE401 without sgRNA.
The products were run on 1% agarose gel alongside a 1kb ladder (Hyperladder 1, Bioline). To be able to show the ladder, the resolution was turned so high that the bands of the samples got smeared (A, B). With lower resolution, the bands of the samples could be observed much clearer, but in expense of the ladder (C, D)

4.1.6 Blasting of *PSY2L* in Tomato *Solanum lycopersicum* versus *Arabidopsis thaliana*

Initially, the transformed *Agrobacterium* containing sgRNAs targeting *PAP7*, *PLL3, PLL2* and dual targeting of *PLL3* and *PLL2* were designed to transform *Arabidopsis*. The transformed *Agrobacterium* containing sgRNAs targeting *PSY2L 1* and dual targeting *PSY2L 1* and *2* were designed to transform

tomato. The *PSY2L* gene in tomato was blasted against the same gene in *Arabidopsis* to see if it might be possible to transform *Arabidopsis* with the same PSY2L spacers designed for transforming tomato.

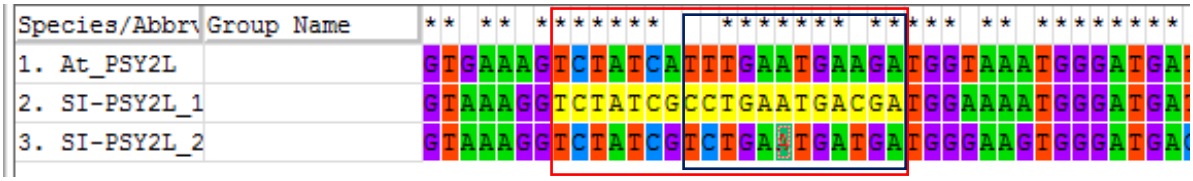
Based on the CRISPR/Cas9 online target prediction tool CCTop, the maximal number of mismatches in the first 12 (read from left to right) nucleotides between the spacer sequences and actual sequence in the gene were two. There could be no more than four mismatches all together.

When searching for the spacer of the first target for *PSY2L 1* in the alignment of *Arabidopsis PSY2L* (At_PSY2L) and tomato *S. lycopersicum PSY2L 1* (SI-PSY2L_1) and *2* (SI-PSY2L_2), three mismatches appeared in the first 12 nucleotides between At_PSY2L and SI-PSY2L_1. In total, four mismatches appeared between these two (Figure 12A). This means that there were too many mismatches for this spacer to be used in transforming *Arabidopsis*. When searching for the spacer of the second target for *PSY2L 1*, only two mismatches were found in the first 12 nucleotides. Five mismatches appeared in total, making also this spacer unsuitable for *Arabidopsis* transformation (Figure 12B).

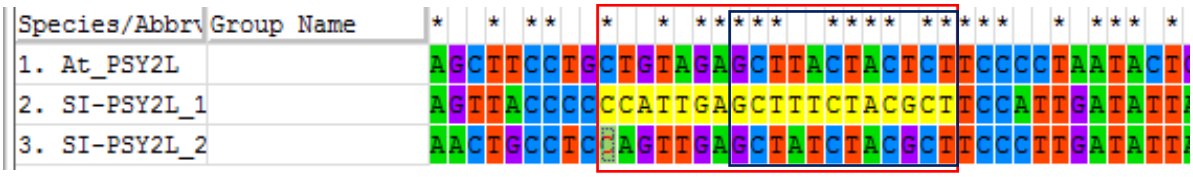
There were also searched for mismatches between the spacers of the sgRNAs dual targeting *PSY2L 1* and *2* and the *Arabidopsis* sequence. The spacer of the first target gave two mismatches in the first 12 nucleotides between At_PSY2L and SI-PSY2L_1, but five mismatches in total. Between At_PSY2L and SI-PSY2L_2, only one mismatch in the first 12 nucleotides appeared. In total, four mismatches between these two appeared (Figure 12C). The spacer of the second target gave no mismatches at all. Neither in the first 12 nucleotides or in total (Figure 12D).

Based on the blasting results, only the *Agrobacterium* containing sgRNAs dual targeting *PSY2L 1* and *2* was used (together with the rest of the sgRNAs designed for *Arabidopsis*) to transform *Arabidopsis* plants.

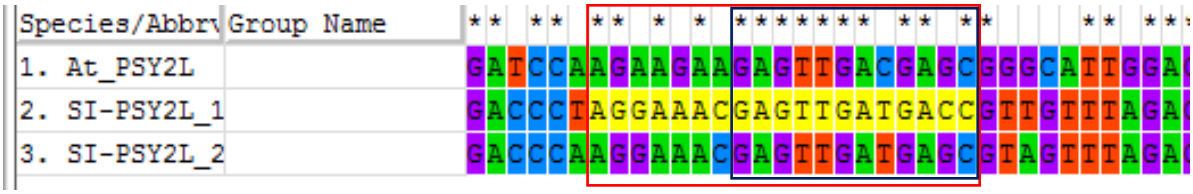
A



B



C



D

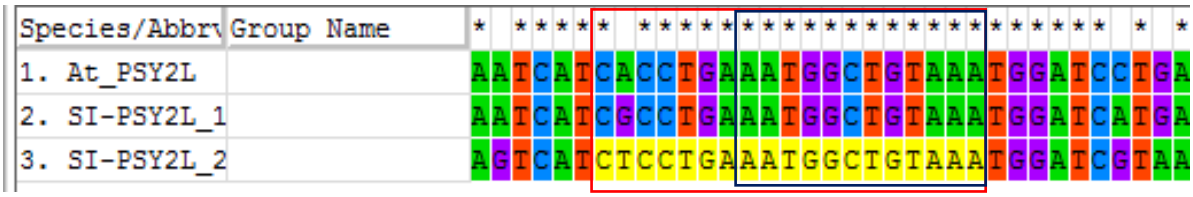


Figure 12: Alignment of Arabidopsis and tomato PSY2L sequences in search of mismatches in sgRNA spacers. The yellow outlines shows the spacer sequences and where it gives a full match. The spacer for gRNA1 targeting PSY2L gave three mismatches in the first 12 nucleotides (blue box) when aligned with PSY2L sequence of A. thaliana (A). It gave four mismatches in total (red box). The spacer for gRNA2 targeting PSY2L gave two mismatches in the first 12 nucleotides (blue box), but five mismatches in total (red box) (B). The spacer for gRNA1 dual targeting PSY2L 1 and 2 gave two mismatches in the first 12 nucleotides (blue box) between SI-PSY2L 1 and At_PSY2L and five mismatches in total (red box). Between SI-PSY2L 2 and At_PSY2L, one mismatch was found in the first 12 nucleotides. Four mismatches appeared in total (C). No mismatches were found in the spacer for gRNA2 in the first 12 nucleotides in any of the aligned sequences. One mismatch was found in total between the three sequences (D).

4.1.7 Screening of T1 generation from transformed Arabidopsis plants

Using colonies from the successfully transformed *Agrobacterium* cells (results, section 4.1.5), *Arabidopsis* plants were transformed by the floral dip method (Clough and Bent 1998) (as described in method section 3.1.7). Seeds collected from transformed plants were screened by sowing them on ½ MS plates containing hygromycin (15 µg/ml). In the first screening, no transformed seedlings were observed. (Figure 13A, B and C). Since the pHSE401 and pHEE401 plasmids contained hygromycin

resistance, transformed seedlings should have shown strong growth with bigger leaves and longer roots than non-transformed seedlings. On all plates, seedlings showed inhibited growth with the same size on leaves and roots. This was also confirmed by comparing the seedlings with the control plate containing wild type seedlings (Figure 13D). The wild type *Arabidopsis* seedlings showed inhibited growth in both cotyledons and roots due to lack of hygromycin resistance. Therefore, no successfully transformed plants were obtained for the T1 generation in the first screening.

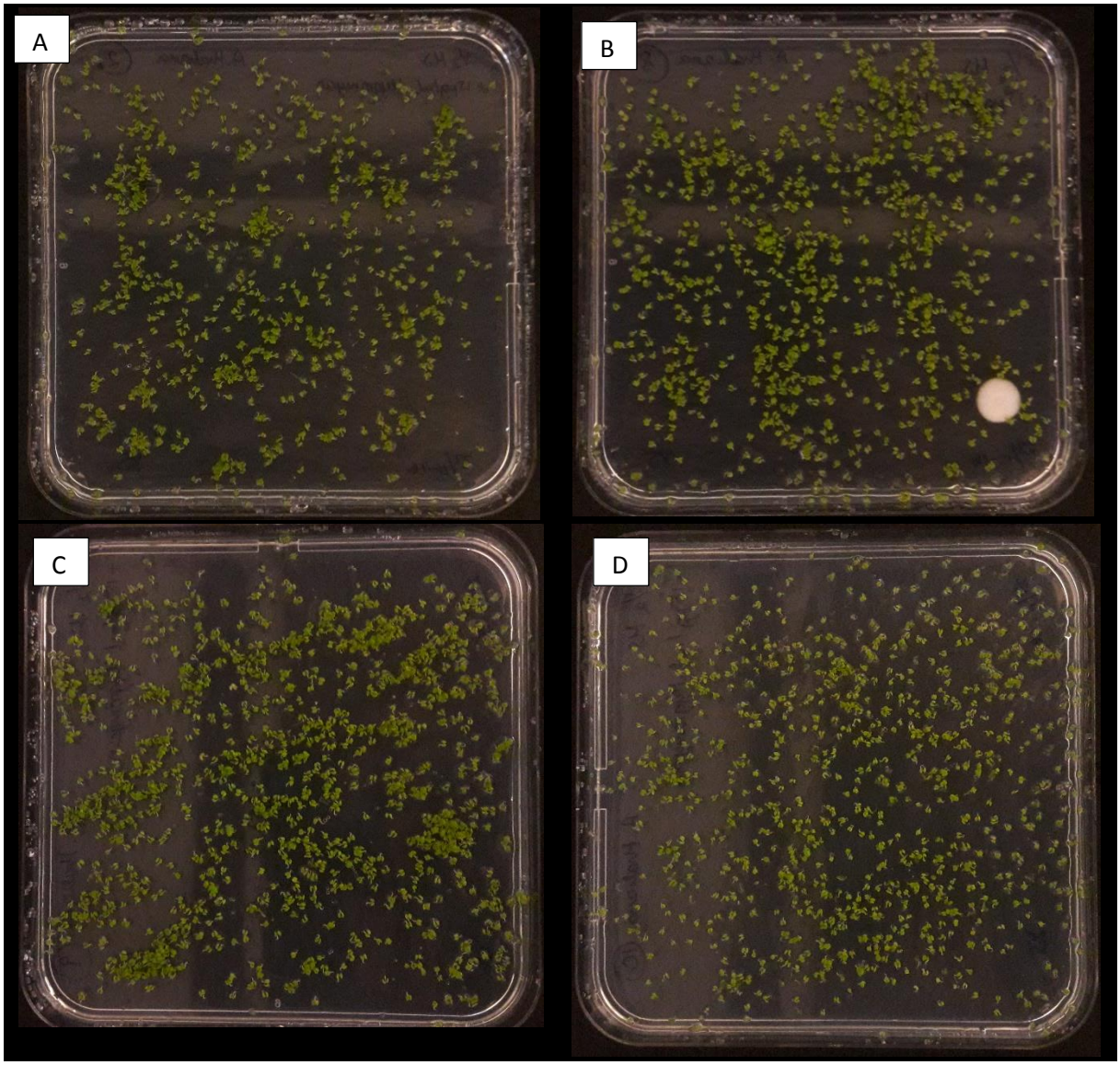


Figure 13: Two weeks old seedlings of seeds harvested from Arabidopsis plants dipped in transformed Agrobacterium. The seeds were sown on ½ MS medium containing sucrose (0.5 %) and Hygromycin (15 µg/ml). No transformed plants were observed in any of the plates. (A) seedlings from plant dipped in Agrobacterium containing two spacers targeting PAP7. (B) seedlings from plant dipped in Agrobacterium containing pHSE401 plasmid without spacers (control). (C) seedlings from plant dipped in Agrobacterium containing pHEE401 plasmid without spacers (control). (D) seedlings from wild type Arabidopsis

When harvesting seeds after drying the plants for two weeks, most of the seeds showed the same inhibited growth as for the first screening (Figure 14A, D and F). Approximately 2-8 seedlings on each plate showed the ability to grow long roots and big cotyledons despite the presence of hygromycin in the medium (Figure 14). All seedlings unaffected by the antibiotics showed the same phenotype and did not differ from each other (Figure 14B and 14E). The only exception was for some of the seedlings (Figure 14C) from plants dipped in agrobacterium containing plasmids with spacers for dual targeting of *PSY2L 1* and *2*. These seedlings were bigger than the ones without plasmid, but not as big as the ones seen in figure 14B and 14E. The control plates with seedlings from plants dipped in *Agrobacterium* containing pHSE401 plasmids without spacers gave only one single seedling unaffected by hygromycin. In the pHEE401 control plates, all seedlings seemed small and affected (Figure 14F).

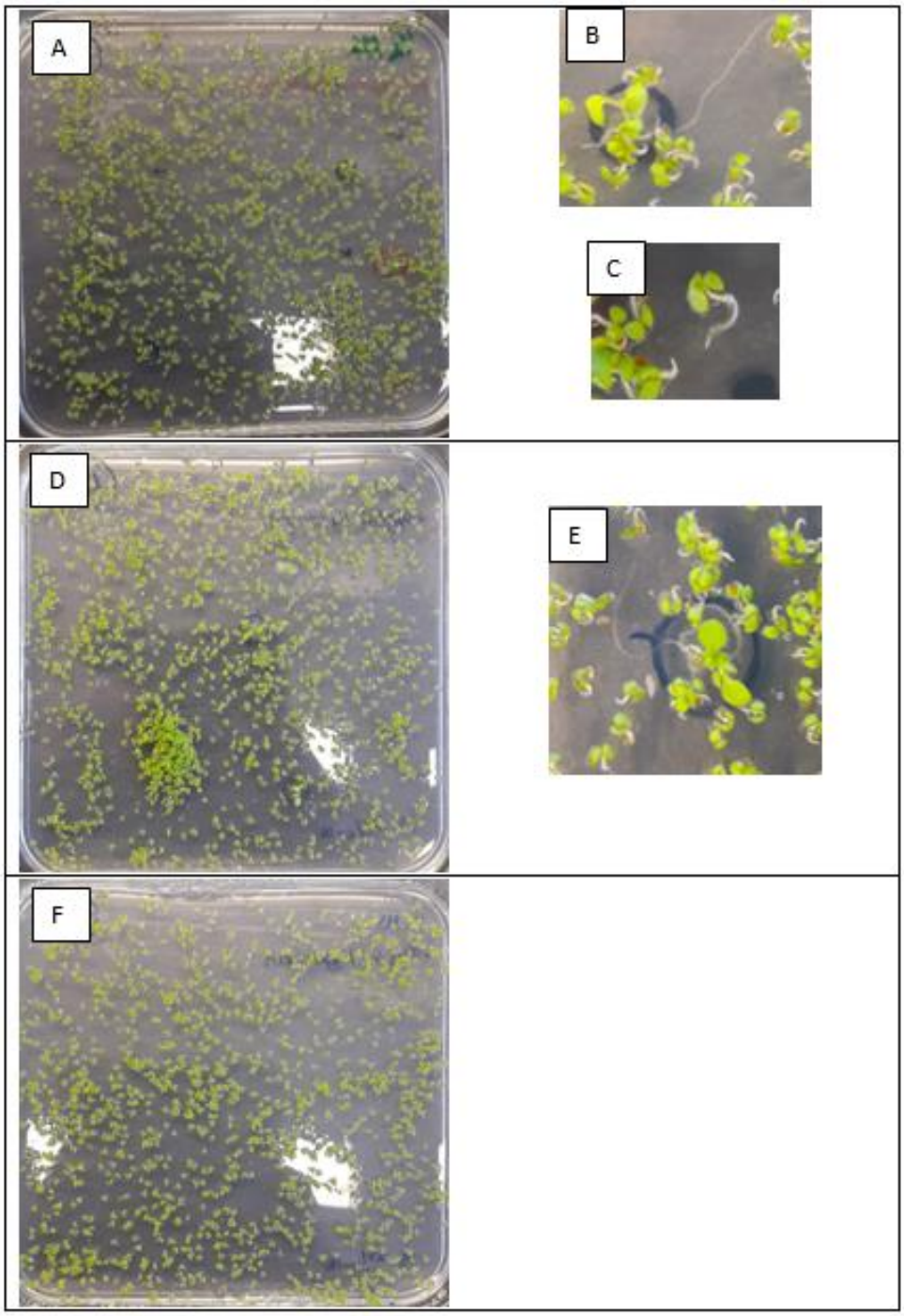


Figure 14: Second screening of seeds harvested from Arabidopsis plants dipped in transformed Agrobacterium containing plasmids with spacers.
The seeds were sown on ½ MS medium containing sucrose (0.5 %) and hygromycin (15 µg/ml). Approximately 2-8 seedlings on each plate was unaffected by the antibiotic in the medium. Seeds harvested from plants dipped in Agrobacterium containing plasmids with spacers dual targeting PSY2L 1 and 2 (A) showed two phenotypes (B and C) when unaffected by the antibiotic. Seeds harvested from plants dipped in Agrobacterium containing plasmids with spacers targeting PLL3 (D), PLL2, PAP7 and dual targeting PLL3 and PLL2 showed only one phenotype (E) when unaffected by the antibiotic. For the controls containing plasmids without spacers, seeds harvested from plants dipped in Agrobacterium with "empty" pHEE401 plasmids (F) gave no transformed seedlings. Seeds harvested from plants dipped in Agrobacterium with "empty" pHSE401 plasmids gave only one transformed seedling.

Due to the unexpected roles of our phosphatases and the possibility to obtain any homozygous plants, we also picked up some seedlings which seemed dwarf but differed from wild type affected by hygromycin. Only the big seedlings with big cotyledons and very long root could survive. Seedlings with the size seen in figure 14C could not handle the change of environment and died short time after transfer.

For the third screening (Figure 15), seeds from the first week of harvesting (seeds harvested before drying the plants), gave mostly seedlings with inhibited growth (Figure 15A and C). Some seedlings differed from others by having slightly longer roots, but still small cotyledons (Figure 15B and E). Between 2-6 seedlings with this phenotype were observed on the PSY2L 1/2 plates (Figure 15A and B) and on the PAP7-, PLL3- and control pHSE401 plates. On the one-spacer PAP7 (Figure 15C, D and E)-, PLL2-, PLL3/PLL2- and control pHEE401 plates, both seedlings with small cotyledons and longer roots and seedlings with big cotyledons and long roots were observed (Figure 15D).

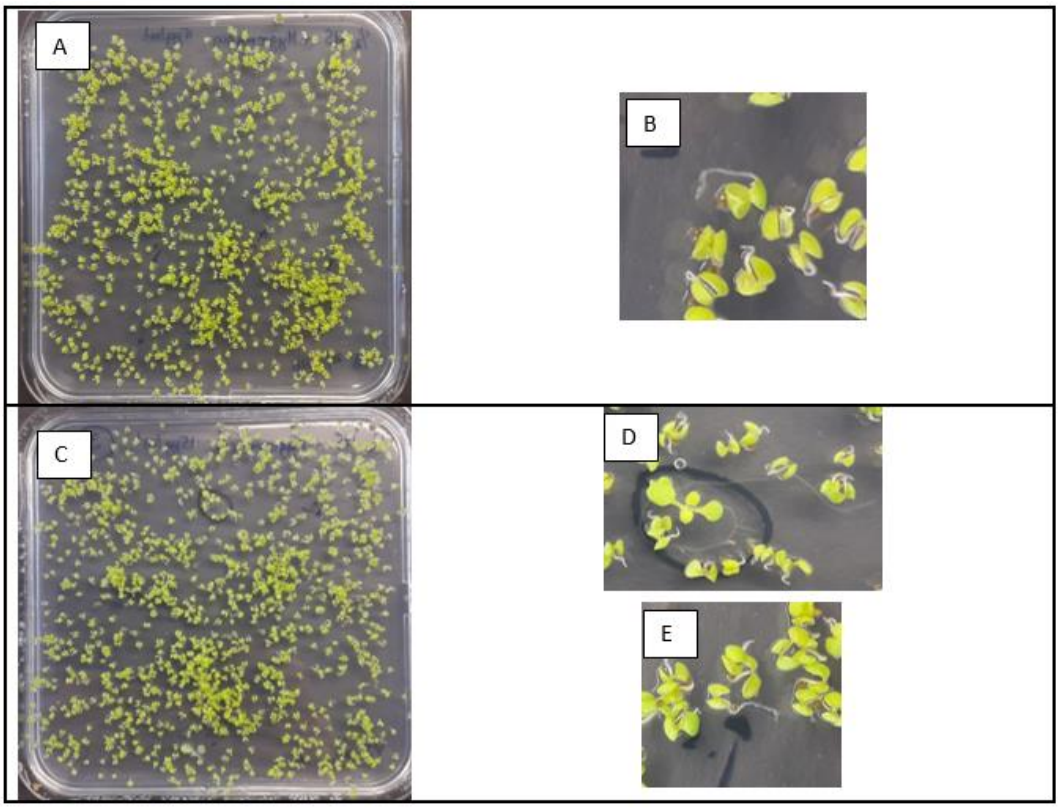


Figure 15: Third screening of seeds harvested from Arabidopsis plants dipped in transformed Agrobacterium containing plasmids with spacers.
The seeds were sown on ½ MS medium containing sucrose (0.5 %) and Hygromycin (15 µg/ml). approximately 2-6 seedlings on each plate seemed unaffected by the antibiotic in the medium. Of the unaffected seedlings, two phenotypes were observed: seedlings with small cotyledons (B and E), but slightly longer roots than the wild type and seedlings with big cotyledons and very long roots (D). Only the first phenotype was observed in the PAP7-, PLL3-, control pHSE401- and PSY2L 1/2 plates (A). In the PLL2-, PLL3/PLL2-, control pHEE401- and the one-spacer-PAP7 plates (C) both phenotypes were observed.

As for the second screening, only seedlings with big cotyledons and long root could handle the change of environment when transferred to soil. The rest of the seedlings died. Since most of the seedlings transferred to soil showed the phenotype with small cotyledons and slightly longer root than wild type, only a small amount from this screening survived.

The fourth and last screening, done by seeds harvested from dried out plants, showed seedlings with the same phenotype as for the second screening (Figure 16). Approximately 5-15 seedlings on each plate seemed unaffected by the antibiotic in the medium and showed a clear phenotype with big cotyledons and very long roots (Figure 16B). In the plates containing the pHEE401 control seedlings, all seedlings were affected by the antibiotic. Only one seedling from the pHSE401 control plates showed a phenotype.

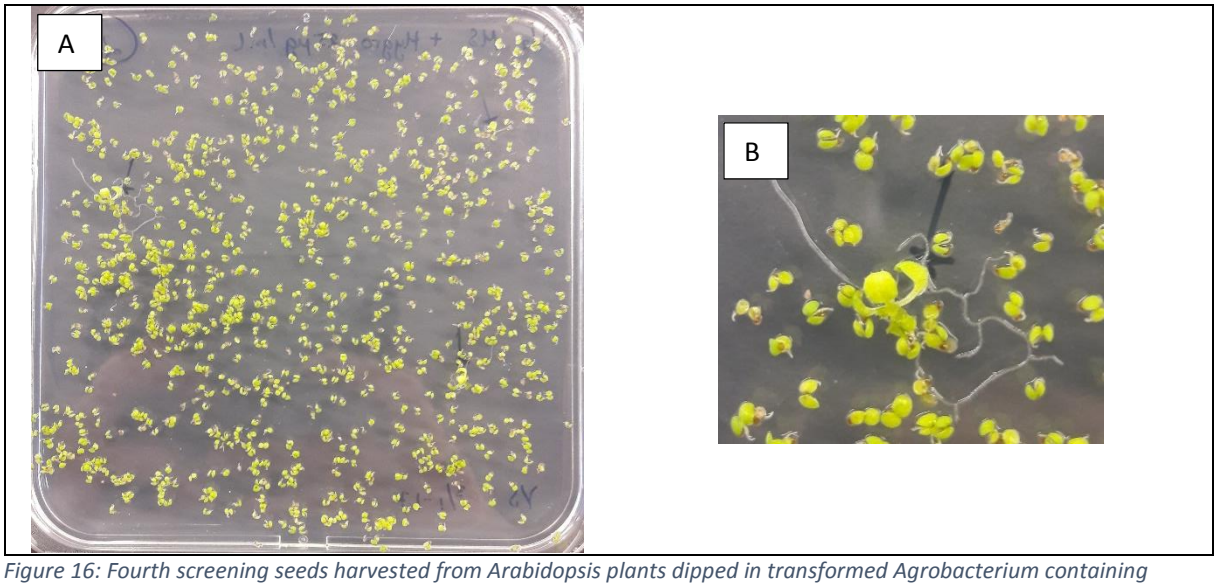


Figure 16: Fourth screening seeds harvested from Arabidopsis plants dipped in transformed Agrobacterium containing plasmids with spacers.
The seeds were sown on ½ MS medium containing sucrose (0.5 %) and hygromycin (25 µg/ml). (A) PAP7 plate with seedlings affected (small) and seedlings unaffected (big) by hygromycin. (B) Unaffected seedling on PAP7 plate.

Since a large number of seedlings from the fourth screening seemed unaffected by the antibiotic and had big cotyledons and long root, only some of the seedlings from each plate were transferred to soil (depended on the number of surviving seedlings from previous screenings). All seedlings survived the transfer.

4.1.8 Verification of cutting by Cas9

4.1.8.1 PCR of extracted DNA

The first step in verifying successful cutting by the Cas9 protein in the target genes was to isolate genomic DNA by DNA extraction (methods, section 3.1.9.1). Based on phenotypes observed in positive seedlings (described in result section 4.1.7), five plants from each transformation were chosen. A PCR was made from the genomic DNA of each plant to amplify around the target site of the target genes.

Based on the designing of the primers for the PCR, different sizes of the products were expected (Appendix A-2). The product size for PAP7 was expected to be 176bp for the first spacer and 196 bp for the second spacer. As seen on the upper bands in figure 17A, B, D and E, all PCR products for PAP7 spacer one and two gave bands in the correct sizes. For PLL3, the expected size of the PCR product for spacer one was 187 bp and 288 bp for the PCR product of spacer two (Figure 17A-E). All PCR samples gave bands in the correct sizes. The same positive results were also obtained for the PLL2 PCR products of spacer one and two. The expected size for the PLL2 PCR product of spacer one was 246 bp, while the expected size for the PLL2 PCR product of spacer two was 247 bp (Figure 17A, C, D and E). The additional controls with DNA extracted from wild type also showed the correct sizes for each spacer (Figure 18)

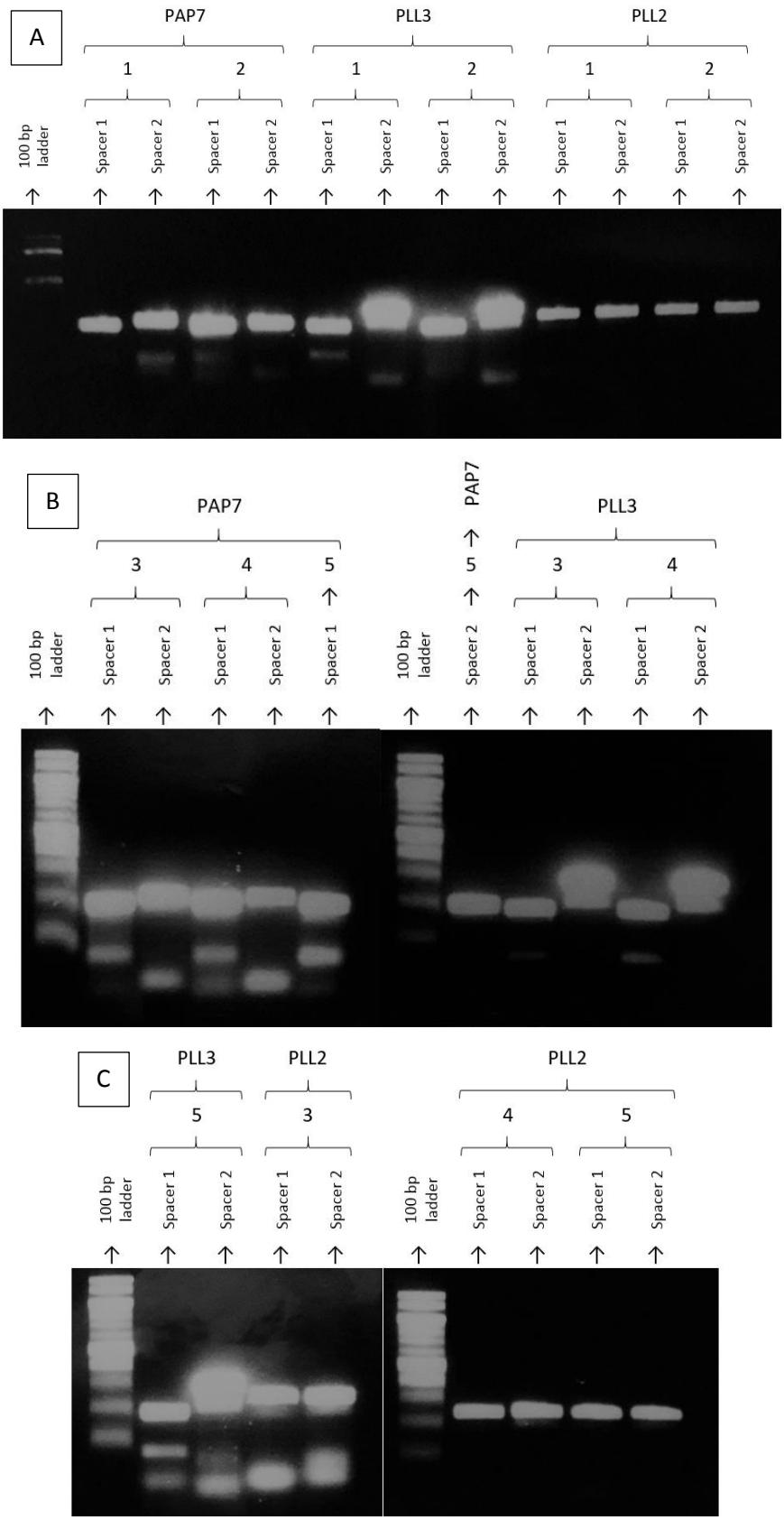


Figure 17: PCR products of DNA extracted from the leaves of the T1 generation of transformed Arabidopsis. The products were run on a 2% agarose gel together with a 100 bp ladder (Bioline, Quick Load). All products gave good bands in the correct size.

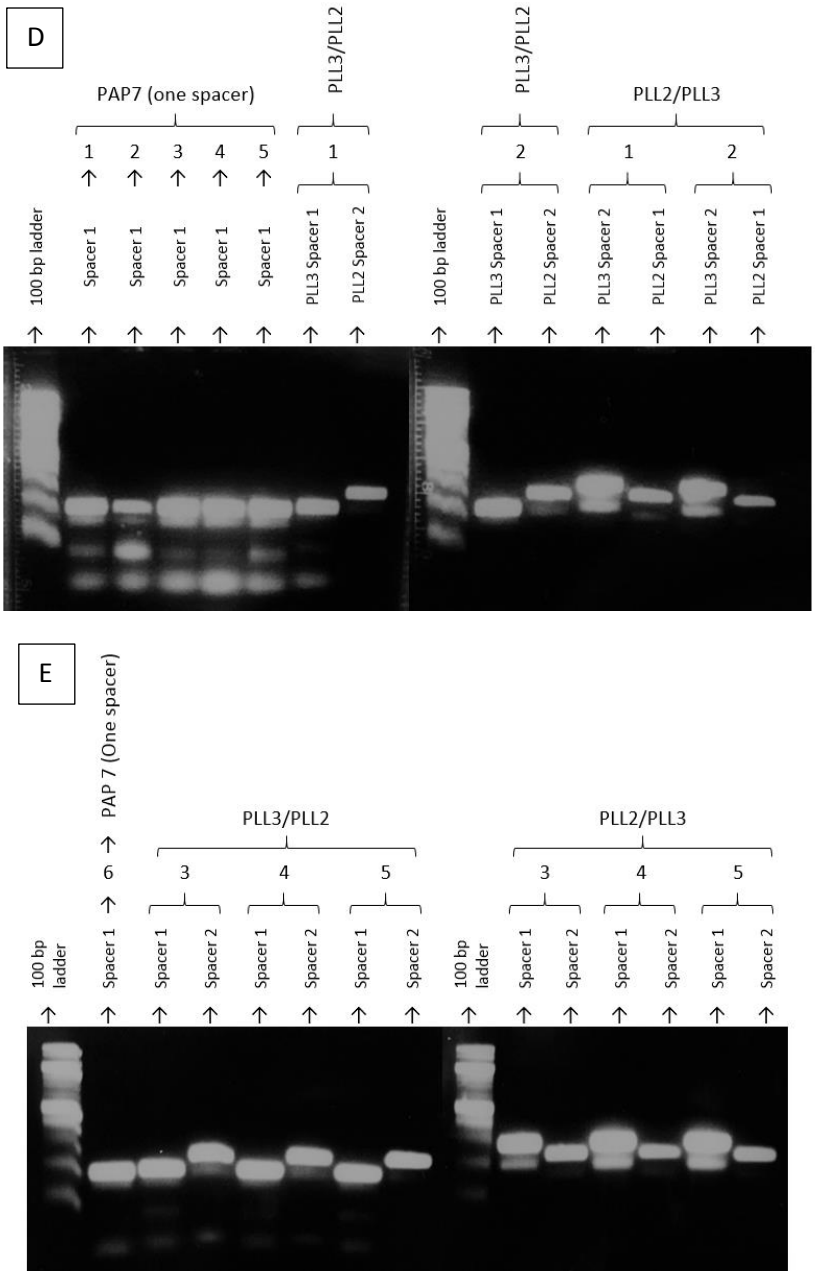


Figure 17: Continued

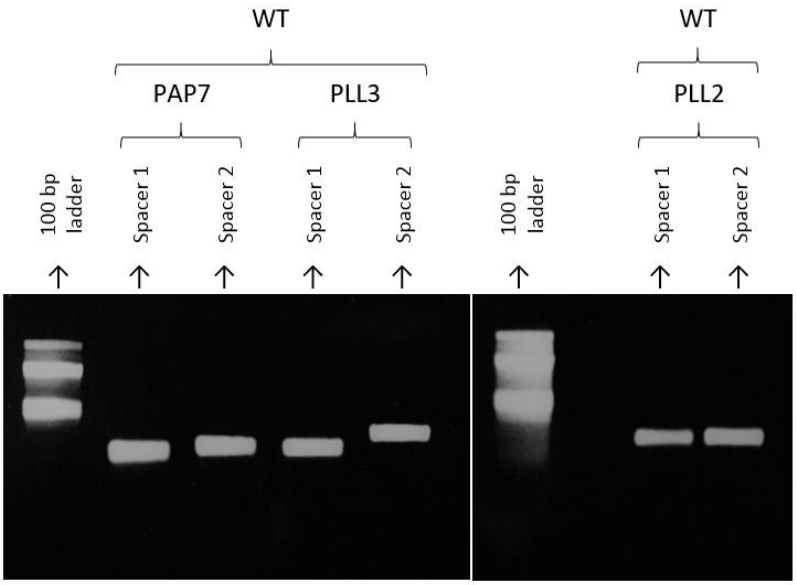


Figure 18: PCR products of DNA extracted from the leaves wild type Arabidopsis.
The products were run on a 2% agarose gel together with a 100 bp ladder (Bioline, Quick Load).

4.1.8.2 T7E1 assay of selected Arabidopsis plants from the T1 generation

After gel-extracting the PCR products (result, section 4.1.8.1), the products were further used in a T7E1 assay (as described in methods section 3.1.9.4) to see if the selected plants could be heterozygous for the mutations. In a hybridization reaction, all DNA strands will separate and then randomly anneal together. If the plants are heterozygous, some strands from the non-mutated allele and strands from the mutated allele will anneal causing a small mismatch between the strands. The mismatch sensitive T7 endonuclease 1 enzyme will then cut the strands, giving two extra DNA fragments. These fragments will then be detected when running the T7E1 assay products on an agarose gel. Figure 19, 20, and 21 show the running of T7E1 assay reaction products on a 2% agarose gel together with a 100 bp Quick Load ladder (NEB). Figure 22 and 23 show the running of T7E1 assay reaction products on a 2% agarose gel together with a 50 bp ladder (Gene Ruler, Thermo Scientific). Bright, strong bands show uncut product and weak bands or smear below the bright bands show product cut by the T7 endonuclease 1 enzyme.

The first two plants containing sgRNAs targeting *PAP7*, showed two weak, but clear bands for spacer number one with sizes of approximately 115 bp and 61 bp (Figure 19). For the second spacer, the figure shows two weak, but clear bands with a size of approximately 126 bp and 70 bp.

For the first two plants containing sgRNAs targeting *PLL3*, the bands were not that clear (Figure 19). For spacer number one, plant number one showed a clear band of approximately 60 bp. The expected size of the bands for this spacer was approximately 98 bp and 89 bp. For spacer one, on the second plant, a very weak band can be seen with the correct size.

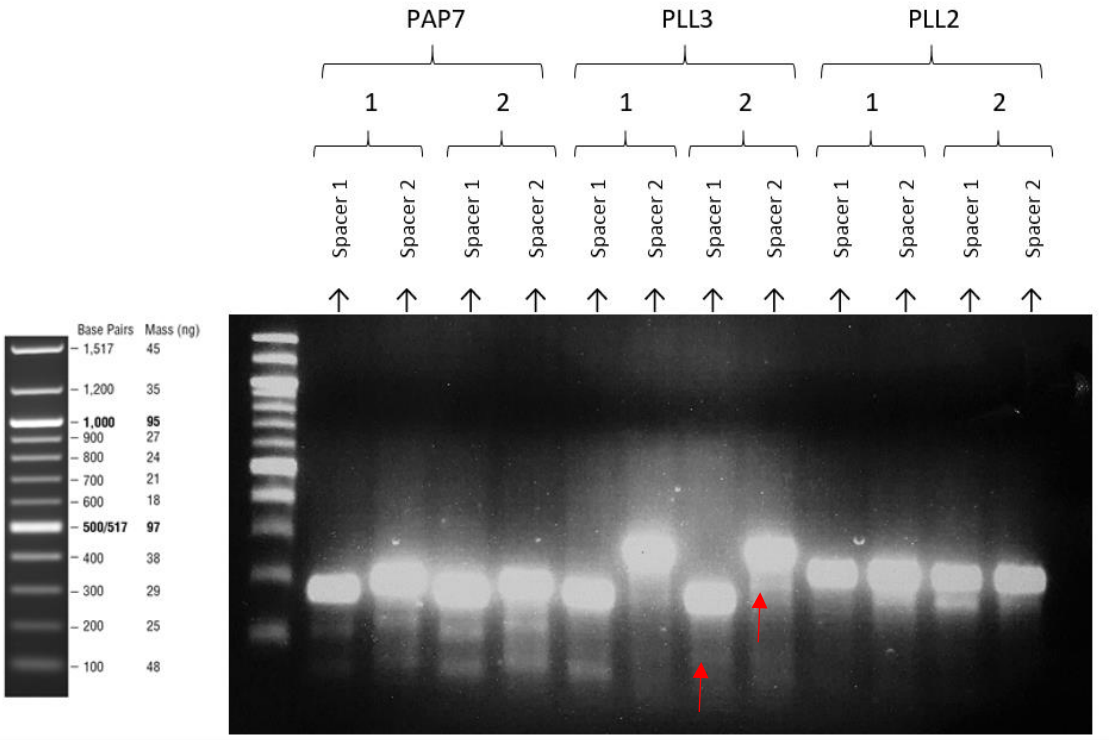


Figure 19: T7E1 assay reaction products run on 2% agarose gel.
Heterozygosity of two plants containing sgRNAs targeting PAP7, two plants with sgRNAs targeting PLL3 and two plants with gRNAs targeting PLL2 were tested. The size of the bands was measured by running a 100bp ladder (Bioline, Quick Load) alongside the samples. The strong bands show uncut product. Weak bands or smear below the strong bands indicates cutting by T7 endonuclease 1 and heterozygote plants.

For spacer number two on the same plants, the expected size for the bands from are 221 bp and 67 bp. Only one band can be seen, and plant number two is the only one showing a clear band off approximately 221 bp. The 67 bp strand cannot be seen in any of the plants.

In the first two plants containing sgRNAs targeting *PLL2*, the expected size for the fragments of spacer number one is 200 bp and 46 bp (Figure 19). The expected size for the fragments of space number two is 134bp and 113 bp. No bands with the size of 46 bp, 113 bp or 134 bp can be seen. But a 200 bp band is seen in both spacers of plant two and in spacer two of plant one. Only a smear can be seen in plant one's spacer number one.

For the last three plants containing sgRNAs targeting *PAP7*, all of them showed two fragments of the first spacer (Figure 20). The clearest bands were obtained for plant number four. The bands for plant number three and five is much weaker, but can still be seen. No fragments can be seen from spacer

number two in plant three and four, but a smear appears from the second spacer in plant number five.

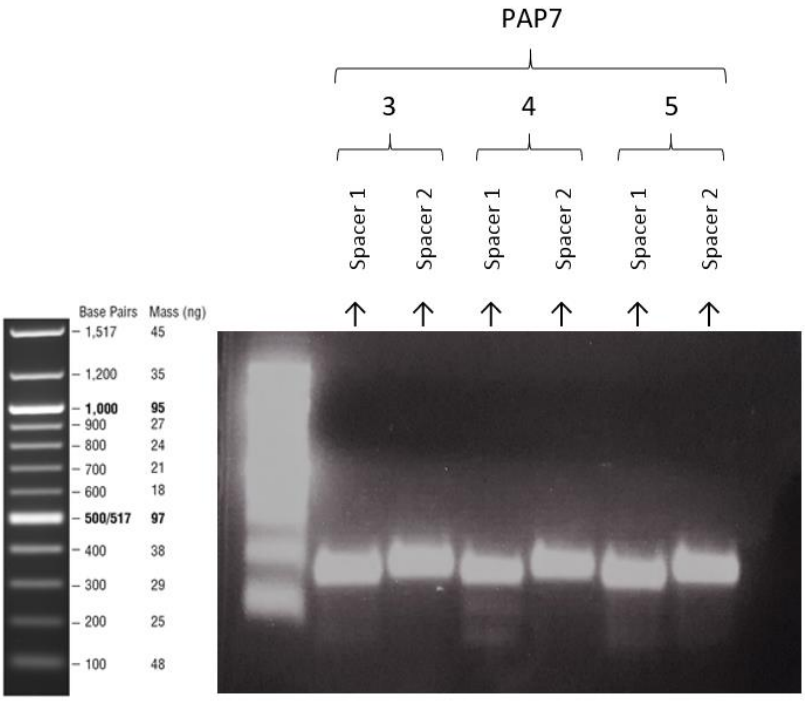


Figure 20: T7E1 assay reaction products run on 2% agarose gel. Heterozygosity of three plants containing sgRNAs targeting PAP7 were tested. The size of the bands was measured by running a 100bp ladder (Bioline, Quick Load) alongside the samples. The strong bands show uncut product. Weak bands or smear below the strong bands indicates cutting by T7 endonuclease 1 and heterozygote plants

For the last three plants containing sgRNAs targeting *PLL3*, no fragments can be observed for the first spacer (Figure 21). For the second spacer, plant number five seem to give two weak and unclear bands. For the last three plants containing sgRNAs targeting *PLL2*, no fragments can be observed in any of the two spacers.

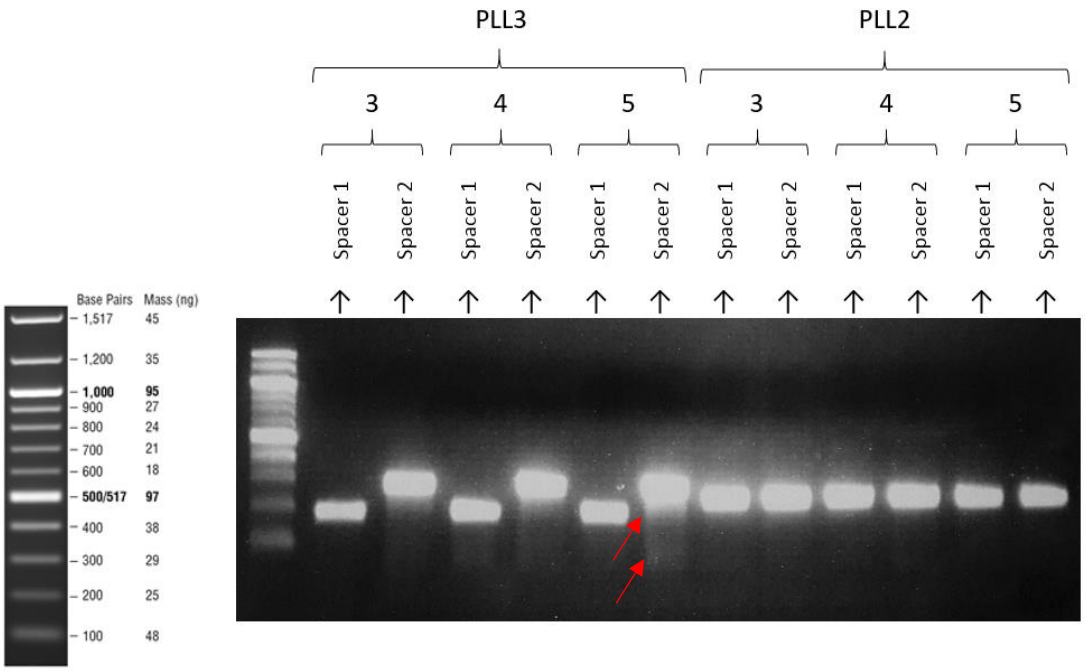


Figure 21: T7E1 assay reaction products run on 2% agarose gel.
Heterozygosity of three plants containing sgRNAs targeting PLL3 and three plants with sgRNAs targeting PLL3 were tested. The size of the bands was measured by running a 100bp ladder (Bioline, Quick Load) alongside the samples. Strong bands show uncut product. Weak bands or smear below the strong bands indicates cutting by T7 endonuclease 1 and heterozygote plants

All five plants containing only one gRNA targeting *PAP7* showed signs of cutting by T7 endonuclease 1 (Figure 22A). Plant four and five gave the clearest bands with sizes just above 100 bp and 50 bp. The fragment bands observed in plant number one, two and three shows the same sizes as for plant number four and five, but the bands are smeared and not clear.

In the plants containing sgRNAs dual targeting *PLL2* and *PLL3*, clear bands can only be seen in the plants containing PLL3's spacer two and PLL2's spacer one. In the first plant, PLL2's spacer one show clear fragment bands of approximately 200 bp and 46 bp. PLL3' spacer two also show bands, but the bands are smeared and without a clear size. For the second plant, PLL3's spacer two is the one showing clear fragment bands. The sizes of the bands are approximately 221 bp and 67 bp. PLL2's spacer one in this plant shows no fragment bands. No fragment bands can be seen in the other sets of plants containing sgRNAs targeting *PLL3* and *PLL2*. PLL3's spacer one and PLL2's spacer two in both plants only show smear below the uncut band.

None of the negative controls (Figure 22B) for these plants showed any bands or smear. This could mean that smearing or fragment bands indicates some sort of cutting by the T7 endonuclease 1 enzyme.

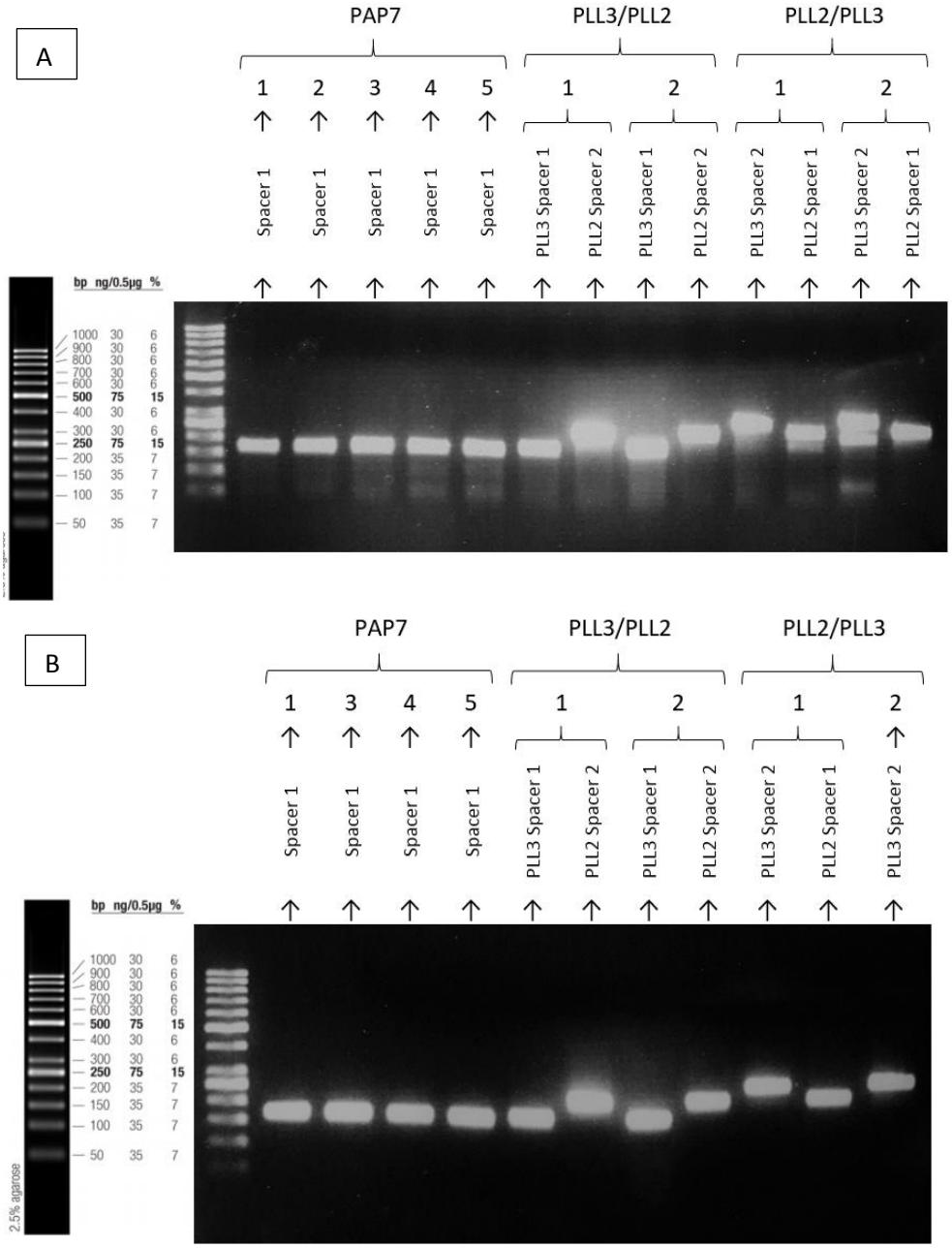


Figure 22: A. T7E1 assay reaction products run on 2% agarose gel.
Heterozygosity of five plants containing one gRNA targeting PAP7, two plants with sgRNAs dual targeting PLL3 and PLL2 and two plants with sgRNAs dual targeting PLL2 and PLL3 were tested. The size of the bands was measured by running a 50 bp ladder (Thermo Scientific, GeneRuler) alongside the samples. The strong bands show uncut product. Weak bands or smear below the strong bands indicates cutting by T7 endonuclease 1 and heterozygote plants B. Negative controls of the T7E1 assay reaction products. The controls were not added enzyme, but treated in the same way as the products tested for heterozygosity.

The last three plants for both sets of dual targeting *PLL3* and *PLL2*, gave no fragment bands (Figure 23A). A very weak smear can be seen below the uncut bands of the PLL2 spacer two from plant five containing gRNAs dual targeting *PLL3* and *PLL2*. Below the uncut bands of the PLL3 spacer two from plant number three and four containing gRNAs dual targeting *PLL2* and *PLL3*.

A sixth plant containing one sgRNA targeting *PAP7* was tested by T7E1 assay. Despite its strong phenotype, no bands can be seen after treatment with the enzyme.

Due to low concentrations after gel extraction, the only samples possible to make controls of were the PLL3's spacer one from plants containing sgRNAs dual targeting *PLL2* and *PLL3* and PAP7's spacer one from the PAP7 plant (Figure 23B).

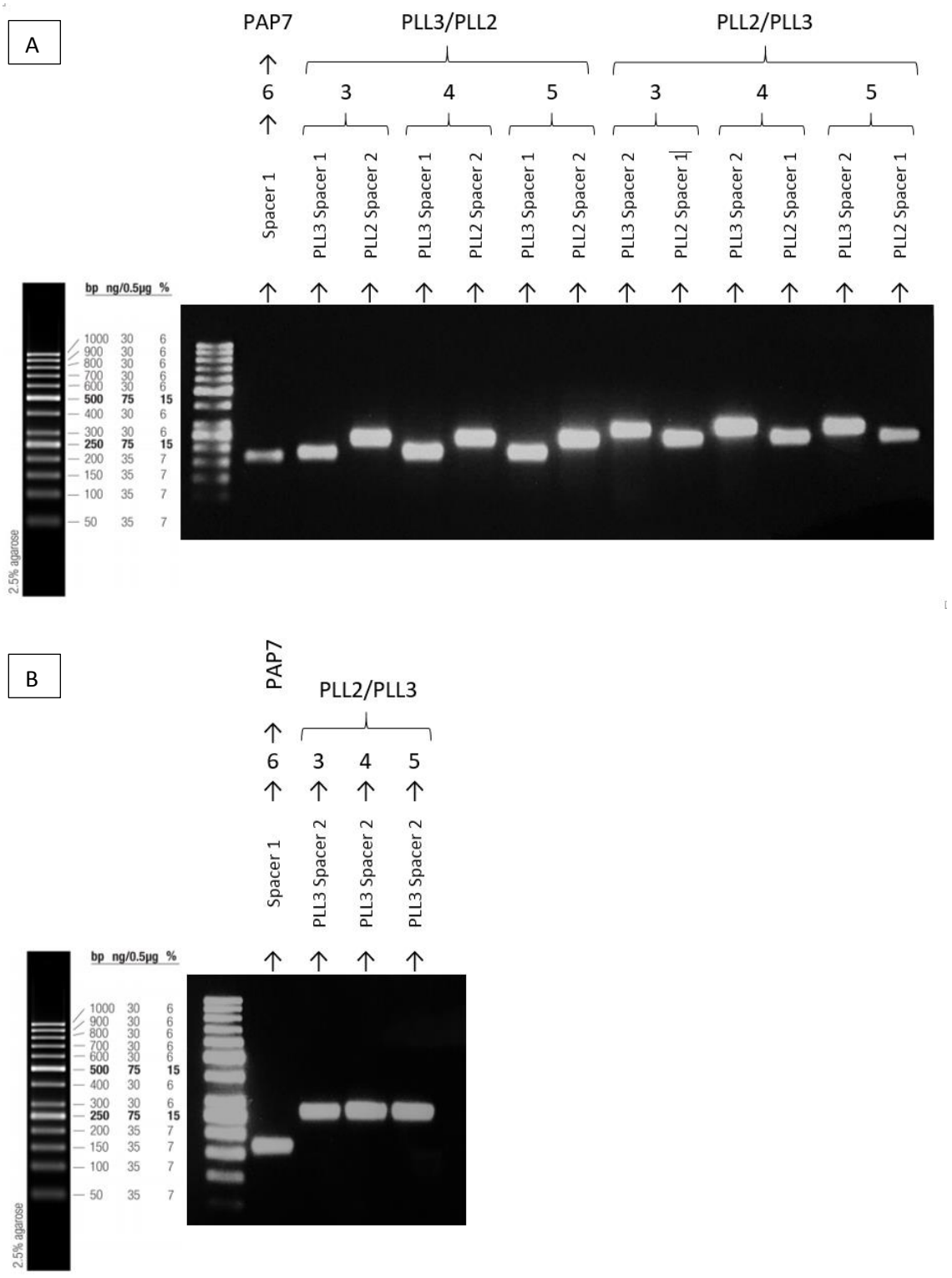


Figure 23: T7E1 assay reaction products run on 2% agarose gel.
Heterozygosity of one plant containing one sgRNA targeting PAP7, three plants with sgRNAs dual targeting PLL3 and PLL2 and three plants with sgRNAs dual targeting PLL2 and PLL3 were tested (A). The size of the bands was measured by running a 50 bp ladder (Thermo Scientific, GeneRuler) alongside the samples. The strong bands show uncut product. Weak bands or smear below the strong bands indicates cutting by T7 endonuclease 1 and heterozygote plants. Negative controls of the T7E1 assay reaction products (B). The controls were not added enzyme, but treated in the same way as the products tested for heterozygosity.

The wild type control for each PAP7-, PLL3- and PLL2 spacer show no cutting by the T7 endonuclease 1 enzyme (Figure 24). No smear or bands can be seen after running the T7E1 assay products on a 2% agarose gel.

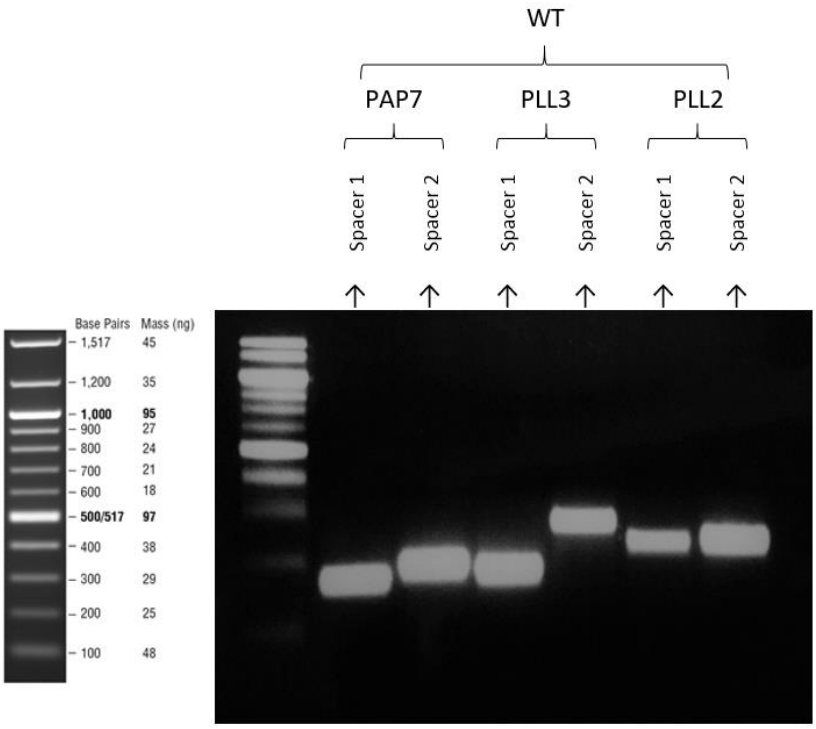


Figure 24: T7E1 assay wild type controls of primers used to check for heterozygote plants.
The wild type DNA was treated the same way as the extracted DNA from the T1 generation of transformed Arabidopsis plants. The products were run on a 2% agarose gel alongside a 100 bp ladder (Bioline, Quick Load)

Table 24: Overview of T7E1 assay results.
Plants which gave fragment bands for spacer one, spacer two or both, indicating cutting by T7 endonuclease 1 are marked with +. Plants which gave only smear or unclear cutting are marked with ? and plants showing no fragment bands or smear are marked with -.

Target gene	**Spacer #**	**Plant #**					
		1	2	3	4	5	(6)
PAP7	1	+	+	+	+	+	
	2	+	+	-	-	?	
PAP7 (one spacer)	1	+	+	+	+	+	-
PLL3	1	+	+	-	?	?	
	2	?	+	?	?	+	
PLL2	1	?	+	-	-	-	
	2	+	+	-	-	-	
Dual PLL3/PLL2	PLL3's 1	?	?	-	-	-	
	PLL2's 2	?	?	-	-	?	
Dual PLL2/PLL3	PLL3's 2	+	+	?	?	-	
	PLL2's 1	+	?	-	-	-	

An overview of plants considered heterozygous based on the T7E1 assay results is provided in Table 24.

Five out of five plants from the T1 generation containing two sgRNAs targeting *PAP7* is, based on the T7E1 assay, considered heterozygous. All plants gave fragment bands for spacer number one when tested by T7E1 assay. Only plant number one and two gave clear fragment bands for spacer number two as well. Spacer two in plant number three and four show no fragment bands, and only gave a smear in plant number five.

In the T1 generation of plants containing only one sgRNA targeting *PAP7*, five out of six plants are considered heterozygous. The first five plants gave fragment bands after the assay, while sixth and last plant gave neither bands or smear.

For plants containing sgRNAs targeting *PLL3*, three out of five plants from the T1 generation tested by T7E1 assay is considered heterozygous. Plant number two is the only one considered to be heterozygous for both spacers. Plant number one gave fragment bands for spacer number one and plant number five for spacer number two.

In the T1 generation of plants containing sgRNAs targeting *PLL2*, two out of five plants are considered heterozygous. Plant number two gave fragment bands for both spacers. While plant number one gave fragment bands for spacer number two and a smear for spacer number one.

Whether any of the five T1 plants containing sgRNAs dual targeting *PLL3* and *PLL2* (spacer number one from PLL3 and spacer number two from PLL2) can be considered heterozygous is unclear. None of the plants gave clear fragment bands in the correct size when tested by T7E1 assay. The first two

plants gave some unclear fragment bands together with some smear, but the size of the bands did not match the size of what was expected.

For the seconds set of T1 plants containing sgRNAs targeting *PLL2* and *PLL3* (PLL3's spacer number two and PLL2's spacer number two), two out of five plants can be considered heterozygous. Plant number one gave fragment bands for both PLL3's spacer two and PLL2's spacer one. Plant number two gave fragment bands for PLL3's spacer two and a smear for PLL2's spacer one.

4.1.9 Screening of the T2 generation from *Arabidopsis* plants tested by the T7E1 assay

For studying phenotypes in the T2 generation of transformed Arabidopsis, seeds from plants showing heterozygosity based on the T7E1 assay were screened on ½ MS medium with hygromycin (25 µg/ml). Seedlings showing specific phenotypes were, after two weeks, transferred to soil. Three phenotypes were picked out and transferred: seedlings with very small cotyledons and no roots, seedlings with big cotyledons and small (barely present) roots, and seedlings with big cotyledons and long roots.

At five weeks old, no obvious or clear phenotypes, except from differences in size, were observed for the T2 generation (Appendix A-3). A small difference was observed in the size of the T2 generation of PLL3 plant number one and PLL3 plant number two. For PLL3 plant number two, the T2 generation was significantly smaller than that of the T2 generation for PLL2 plant number one. Despite the big variety in sizes of the T2 generation of PAP7 (one spacer) plants, the T2 generation of plant number three stood out with extremely big leaves compared to the rest of the plants.

When older and with flowers, no difference in size could be observed between the plants.

4.1.10 Tomato transformation and growth of callus

In an attempt to make mutations in the gene of tomato PSY2L, *Agrobacterium* transformation was also used to insert plasmids (results section 4.1.4) into tomato *S. lycopersicum* (as described in method section 3.1.11)

When tomato cotyledons, dipped and swirled in transformed *Agrobacterium*, were transferred to the selection medium (D1) and placed in the plant room (16 h light), the carbenicillin did not kill or suppress the *Agrobacterium* and the bacteria overgrew the cotyledons. This made it impossible for the cotyledons to grow any callus and whether the transformation was successful or not could not be determined. Most likely the carbenicillin itself wasn't strong enough to kill the bacteria and a much stronger antibiotic should have been used.

Due to time limitations, no further attempts to grow callus were made.

4.2 In-vivo CRISPR/Cas9 modifications through PEG transfections of isolated *S. lycopersicum* and *Arabidopsis* protoplasts by recombinant vectors.

4.2.1 Colony PCR of transformation of *E. coli*

In an attempt to knock out genes in tomato and Arabidopsis by transforming protoplasts, complementary oligonucleotides (spacers) were digested and annealed together with a pHSN401 plasmid (as described in methods section 3.2.1). The recombinant plasmid was then inserted into *E. coli* cells for cloning (methods, section 3.2.2)

The colony PCR run after transforming E. coli showed that the transformation had been successful. All colonies gave band in the expected size of 162 bp, meaning that the bacteria had received recombinant pHSN401 plasmids (Figure 25 and 26).

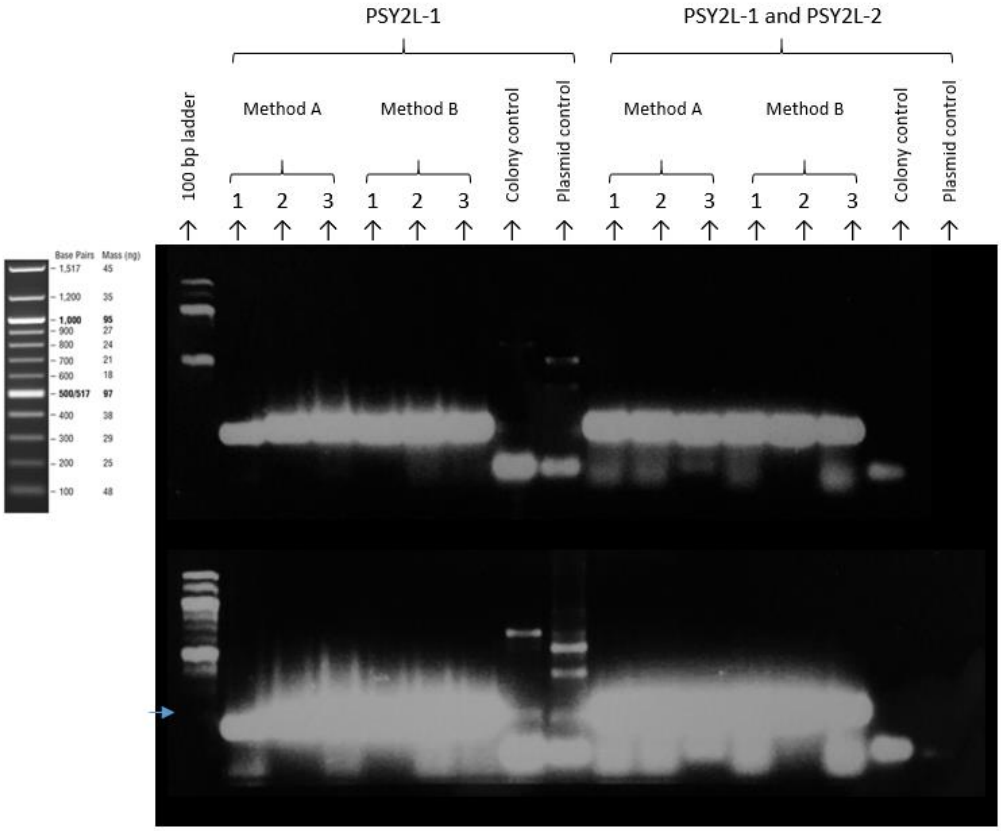


Figure 25: Colony PCR products run on 2% agarose gel in 1x TAE buffer alongside a 100 bp ladder (Quick Load, Bioline). Six different colonies were checked for each gene-target, PSY2L-1 and PSY2L 1 and 2. Three colonies with complementary oligonucleotides annealed by method A and three with complementary oligonucleotides annealed by method B. All colonies gave bands with the size of 162 bp. The controls gave bands with a significantly lower size.

Even though both methods for annealing complementary oligonucleotides seemed to function well, method B was chosen to be further used when annealing the last three pairs of oligonucleotides. By using method B, phosphorylation of the oligonucleotides by the PNK enzyme was avoided.

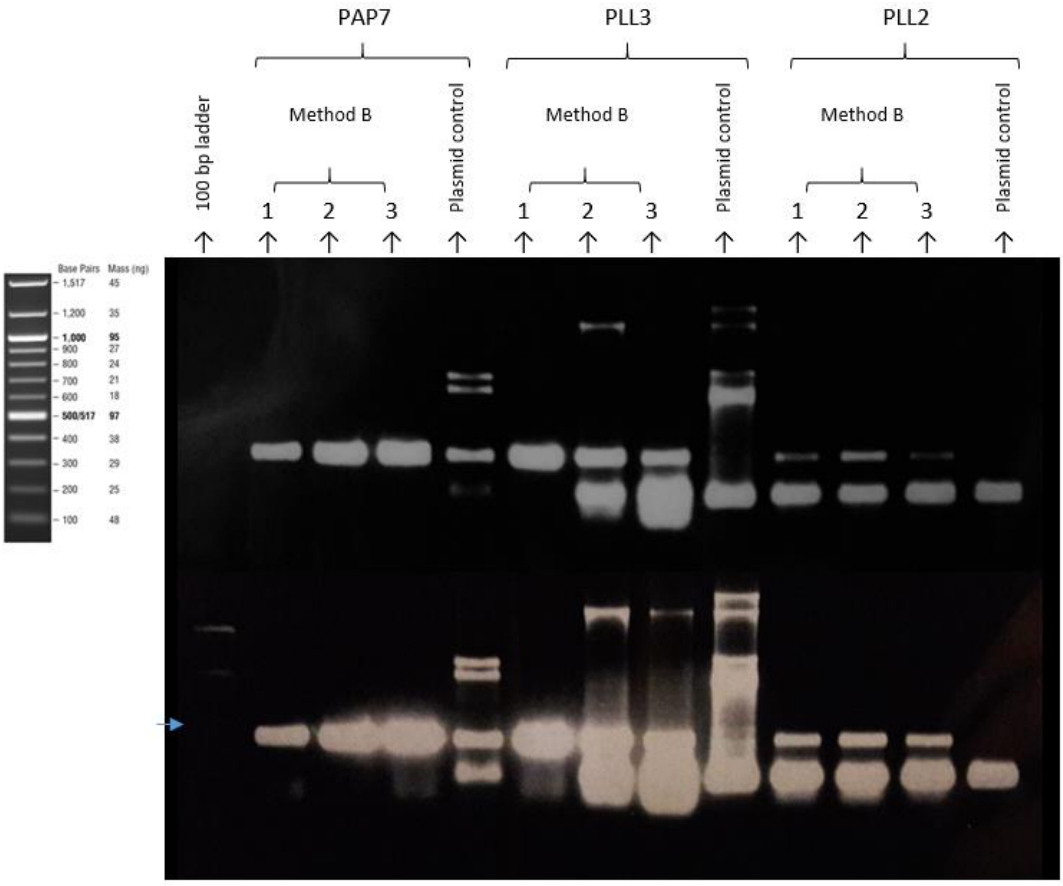


Figure 26: Colony PCR products run on 2% agarose gel in 1x TAE buffer alongside a 100 bp ladder (Quick Load, Bioline). Three different colonies were checked for each gene-target, PAP7, PLL3 and PLL2. All of them with complementary oligonucleotides annealed by method B. All colonies gave bands with the size of 162 bp. The controls gave bands with a significantly lower size.

For sequencing of plasmids containing sgRNAs targeting *PSY2L 1* and *PSY2L 1* and *2*, colonies annealed by both methods was used. Due to the good quality of the bands in the colony PCR, one of the three colonies for each method was selected randomly. For *PSY2L 1*, colony number three from method A and colony number two from method B was selected. For *PSY2L 1* and *2*, colony number three from method A and colony number two from method B was selected.

Since only method B was used to anneal the complementary oligonucleotides for sgRNAs targeting *PAP7*, *PLL3* and *PLL2*, one colony for each gene was selected for plasmid isolation and sequencing. The selected colony for PAP7 was number one, while colony number two and colony number one was selected for PLL3 and PLL2.

4.2.2 Sequencing of plasmids isolated from transformed *E. coli*

In the first attempt of sequencing the plasmids containing sgRNAs targeting *PSY2L 1*, *PSY2L 1* and *2*, *PAP7*, *PLL3* and *PLL2*, the plasmids with sgRNAs targeting *PSY2L 1* (gRNA1a, gRNA1b), *PSY2L 1* and *2* (gRNA2a, gRNA2b) (Figure 27A) and *PAP7* (gRNA3b) (Figure 27B) showed correct insertion of the spacer sequences. The lab was unable to sequence the plasmids containing sgRNAs targeting *PLL3* (gRNA4b) and the plasmids containing gRNAs targeting *PLL2* (gRNA5b) showed incorrect insert of the spacer. Due to the unsuccessful sequencing of these two samples, plasmids from PLL3's colony number one and three and plasmids from PLL2's colony number two and three (Figure 26) was isolated and sent for a second sequencing.

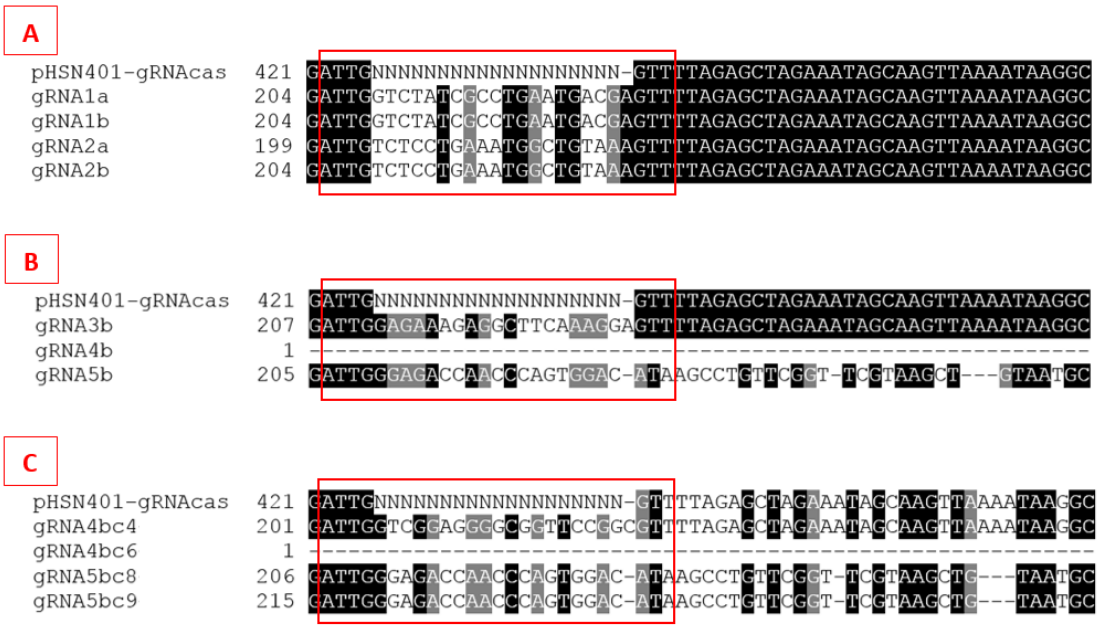


Figure 27: Aligned spacer sequences of sequenced plasmids containing one guide RNA.
The sequences were aligned by Clustal W. "Sequence of one sgRNA expression cassettes for dicots, 23 bp insert + pHSN401 et al. sourced from Xing et al. 2014 was used as template for the alignment, where the 19 N represents the area of spacer inserts. Sequencing of plasmids containing sgRNA targeting PSY2L 1, PSY2L 1 and 2 (A) and sequencing of plasmids containing sgRNA targeting PAP7, PLL3 and PLL2 (B). Due to unsuccessful sequencing of plasmids with sgRNAs targeting PLL3 and PLL2, new plasmids were extracted from new colonies and sent for a second sequencing (C).
pHSN401-gRNA-cas: Template. gRNA1a: spacer targeting PSY2L 1, with oligonucleotides annealed by method A. gRNA1b: spacer targeting PSY2L 1, with oligonucleotides annealed by method B. gRNA2a: spacer dual targeting PSY2L 1 and 2, with oligonucleotides annealed by method A. gRNA2b: spacer dual targeting PSY2L 1 and 2 with, oligonucleotides annealed by method B. gRNA3b: spacer targeting PAP7, with oligonucleotides annealed by method B. gRNA4b: spacer targeting PLL3, with oligonucleotides annealed by method B. gRNA5a: spacer targeting PLL2, with oligonucleotides annealed by method B. gRNA4bc4 and gRNA4bc6: spacers targeting PLL3, oligonucleotides annealed by method B. gRNA5bc8 and gRNA5bc9: spacers targeting PLL2, oligonucleotides annealed by method B.
For the first sequencing, plasmids with sgRNAs made for targeting PSY2L 1, PSY2L 1 and 2 and PAP7 showed correct insert of spacer. The plasmids with gRNAs made for targeting PLL3 and PLL2 was either impossible to sequence or had the wrong insert for the spacer. The second sequencing of the unsuccessful plasmids gave one sample containing the correct spacer for targeting PLL3, but the spacer insert for sgRNA targeting PLL2 was still incorrect.

The results from the second sequencing showed that one of the two plasmid samples containing sgRNA targeting *PLL3* (gRNA4bc4) had the correct insert of the spacer (Figure 27C). The second sample for this gene (gRNA4bc6) was, as for the first sequencing, impossible to sequence. None of the plasmids samples containing sgRNAs targeting *PLL2* (gRNA5bc8 and gRNA5bc9) showed correct spacer sequence after the second sequencing. For some reason, all plasmid samples containing sgRNAs targeting *PLL2* showed the same incorrect sequence in the spacer-area.

Due to time limitations, we were not able to establish protocols for isolation of *Arabidopsis* and tomato protoplasts.

5 Discussion

5.1 Expression cassettes and cloning of recombinant vectors for *Agrobacterium* mediated transformation of *Arabidopsis*

In the attempt of making expression cassettes for sgRNAs targeting tomato *PSY2L 1*, *PSY2L 1* and *2* and *Arabidopsis PAP7*, *PLL3* and *PLL2*, both spacers for each gene were successfully integrated into the cassettes. Verification by gel electrophoresis showed that the PCR products obtained when using the pCBC-DT1T2 as template were of approximately 600 bp in size. According to Xing et al. (2014), this was the size expected for the PCR products. The products were gel extracted and ligated into specific plasmids. Based on previous and successful attempts in gene editing in tomato *S. lycopersicum* by CRISPR/Cas9 systems containing a 35S promotor (Brooks et al. 2014), a pHSE401 plasmid containing the same promotor was chosen to be used in tomato for knocking out PSY2L. As reported by Wang et al. (2015), most attempts using the 35S promotor for driving Cas9 transcription in Arabidopsis only resulted in somatic mutations in the first generation. They reasoned that the lack of homozygous mutants was due to low expression levels of Cas9 in egg cells and one-cell stage embryos. By using egg-cell specific promotors they successfully created T1 homozygous or bi-allelic mutants. When testing different combinations of three promotors and two terminators and thereafter different combinations of two egg-cell specific promotors fused with different enhancers, they showed that a combination of an EC1.2 enhancer fused to a EC1.1 promotor gave the best results in obtaining homozygous/bi-allelic mutants and avoiding mosaic plants in the first and second generations. Therefore, a pHEE401 plasmids containing a EC1.1 promotor fused to an EC1.2 enhancer was chosen to be used in *Arabidopsis*.

Since both plasmids contained genes for kanamycin resistance, *E. coli* transformed with recombinant plasmids were plated and cultured on LB plates with kanamycin for cloning. Colonies were able to grow on all plates, showing successful uptake of plasmids. The plasmids were designed to excise the spectinomycin resistance gene and replace it with sgRNA expression cassettes when digested and ligated by Bsa1. To test correct cutting by Bsa1 and ligation of the sgRNA sequences into the spectinomycin resistance gene of the plasmid, transformed *E. coli* was plated on a LB plate containing spectinomycin. Controls with bacteria transformed with non-recombinant plasmids were also plated on the same plate. Only bacteria from the controls were able to grow colonies, indicating successful cutting by Bsa1 and ligation of expression cassettes into the area of the spectinomycin resistance gene. The successful cloning of the expression cassettes into the plasmids was, as suggested by Xing et al. (2014), further verified by colony PCR. The expected size of 726 bp products were obtained for all the constructs. To confirm correctly insertion of the expression cassettes and correct spacer sequences, the plasmids were isolated and sent for sequencing. All plasmids, except for the one

containing sgRNAs targeting *PAP7*, showed correct insertion of expression cassettes with correct spacer sequences. The second attempt in sequencing plasmids containing sgRNAs targeting *PAP7*, plasmids were isolated from four new colonies. Two of them showed correct insertion of both spacers, one with the correct insertion with spacer number one and one with no correct insertion with any of the spacers. Despite the recommendations for using binary vectors containing two sgRNAs targeting the same gene (Wang et al. 2015), we chose to (in addition to using one of the constructs containing two sgRNAs targeting *PAP7*) test if we were able to knock out *PAP7* with only one sgRNA. This gave us two binary vectors for targeting *PAP7*; one containing a two sgRNA expression cassette and one containing a one sgRNA expression cassette.

5.2 Transformation of *Arabidopsis* by *Agrobacterium* mediated flower dipping

The most widely used and the most successful method in transforming *Arabidopsis* is the *Agrobacterium*-mediated floral dipping (Ma et al. 2016; Clough and Bent 1998). The successful constructs with sgRNAs were transformed into *Agrobacterium* strain Gv3101. Successful transformation of *Agrobacterium* was confirmed by colony PCR. Primers suggested by Xing et al. (2014), were used to amplify around the expression cassette, making PCR products of approximately 726 bp in size. When the PCR products were run on agarose gel, all constructs except from the two controls and the one containing only on sgRNA targeting *PAP7*, gave bands in the 700 bp area. No bands were obtained for the controls or for the construct containing only one sgRNA targeting *PAP7*. The unsuccessful amplification of the construct containing one sgRNA targeting *PAP7* was due to the lack of complementary sequence for the U6-29p reverse primer. As described in Xing et al. (2014), a U6-26t reversed primer should be used to amplify around the sgRNA in a one sgRNA expression cassette. When replacing the U6-29p reverse primer with the U6-26t reversed primer a product of approximately 400 bp in size was obtained for this construct.

Since the phenotype of *Arabidopsis PSY2L* knock outs have been studied and found to be dwarfed compared to WT, we wanted to see if it was possible to knock out *PSY2L* in Arabidopsis with the same spacers designed to knock out *PSY2L* in tomato. Even though no primers were designed for the amplification of *PSY2L* spacers for *Arabidopsis* when testing for cutting by Cas9, successful homozygous mutants would show clear phenotypes. The gene of PSY2L in tomato was blasted against the same gene in Arabidopsis. Based on the blasting results only the spacers designed for dual targeting of *PSY2L 1* and *PSY2L 2* would also target *Arabidopsis PSY2L*.

Arabidopsis plants were transformed by *Agrobacterium* mediated floral dipping. After dipping, the plants were treated as normal to develop siliques.

5.3 Screening for transformed *Arabidopsis* plants and phenotyping of T1 generation

Seeds were harvested both before and after drying the plants. Harvested seeds were screened for transformed plants by sterilizing and sowing them on ½ MS medium containing hygromycin. Since both plasmids contained hygromycin resistance, only seedlings from seeds which had integrated the plasmid into their genome would be able to grow uninhibited. In three out of the four screenings, transformed seedlings were obtained. The first screening was done with seeds harvested before drying and gave no transformed seedlings. All seedlings showed the same inhibited growth as the wild type *Arabidopsis*. This was probably due to too early harvesting. The seeds harvested might have come from flowers that had come too far in the embryogenesis to transform by *Agrobacterium*. Seeds harvested after drying gave large number of transformed seedlings when screened. To see if the difference between transformed and non-transformed seedlings could be even more prominent, the concentration of the antibiotic was adjusted from 15 µg/ml to 25 µg/ml. The last screening gave the highest number of transformed seedlings compared to earlier screenings. It seems like the time of harvesting and the concentration of the antibiotic used affects the possibilities to discover transformed seedlings when screening. We had problems obtaining transformed seedlings in the controls containing non-recombinant plasmids

Approximately 10-15 transformed seedlings from each transformation were transferred to soil to grow plants big enough for isolation of genomic DNA. Throughout the period of growing, no specific or repeating phenotypes were observed in the plants. Some plants differed from others either by growing slower or flowering earlier. Two of the plants from the T1 generation of plants transformed to dual target *PSY2L 1* and *PSY2L 2* were, at 4 weeks old, smaller than the other plants obtained from the same transformation. But as they grew, started flowering and producing seeds these plants could not be distinguished from the rest. The T1 generation of plants transformed to target *PAP7*, *PLL3*, *PLL2* and dual targeting *PLL3* and *PLL2* showed no specific phenotypes. These results are consistent with the results of previous studies of phenotypes in PLL3, PLL2 and PLL3/PLL2 double mutants (Song and Clark 2005). No attempts in studying PAP7 mutant phenotypes have been reported. During stem-development, the T1 generation of plants transformed to knock out PLL2 seemed to develop stems and flowers earlier than the other plants. Also, the stems of the T1 plants transformed to make PLL3/PLL2 double mutants seemed less robustly than the rest.

5.4 Search for heterozygote *Arabidopsis* plants by T7E1 assay

For further studies of PAP7-, PLL3- and PLL2 knock-out mutants, the genotype of the mutants must be determined. In our attempt of making such mutants, the first step in genotyping was to detect cutting by the Cas9 protein. Due to time limitations, only five plants from each T1 generation were chosen. Among the many different methods to do so, such as using reporter genes, endonucleases and PAGE, we chose to detect mutations by performing a T7 endonuclease 1 assay. This is an easy, fast and cheap method, suitable for any targeted mutations (Ma et al. 2016). Genomic DNA was extracted from leaves of the chosen plants and a PCR amplifying around the target site was performed. A specific amount of each PCR product was further used to perform the assay.

This assay is based in detecting and cutting of mismatches in the DNA by a T7 endonuclease 1 enzyme. In presence of a buffer, the DNA strands are separated and reannealed. The annealing happens randomly in such a way that even strands containing a small mutation (deletion, substitution or insertion) can anneal with non-mutated strands forming a mismatch. When adding the enzyme, the enzyme will detect this mismatch and cut both strands. When running the T7E1 assay products through gel electrophoresis, this cutting will be seen by the presence of two extra bands.

In the first round of an T7E1 assay, only heterozygous plants can be detected. The first two plants with sgRNAs targeting *PAP7* showed clear cutting by the enzyme in both spacer areas. This indicates that the plants are heterozygous and that Cas9 has performed successful cutting in the target site, causing a mutation in one allele. In the last three plants, all showed successful cutting in spacer number one. Due to the lack of extra bands for spacer number two in these plants, the Cas9 likely did not cause any mutations in this target site. Since all five plants showed cutting for spacer number one, they can all be considered as heterozygous PAP7 mutants. When testing the plants containing only one spacer targeting *PAP7*, all five plants showed cutting by T7 endonuclease 1. Therefore, these plants can be considered heterozygous PAP7 mutants. When comparing the two spacers used for knocking out PAP7, spacer number one seems to be the most effective in leading the Cas9 protein for cutting.

The only plant showing cutting by the enzyme for both PLL3 spacers is plant number one. This plant is therefore considered as a heterozygous PLL3 mutant. plant number two and five showed cutting in one spacer only; spacer number one in plant two and spacer number two in plant five. Even though only on spacer seemed to function in these plants, both are considered heterozygous PLL3 mutants. Only smearing can be seen on the gels run with T7E1 assay products from plant number three and four. Obviously, some sort of reaction has occurred, but since no clear bands can be visualized, the

plants can only be assumed to be either wild type or homozygous. For knocking out PLL3, both spacers seem equally effective.

For plants containing sgRNAs targeting *PLL2*, only plant number two show cutting in both spacers. Plant number one only show cutting in spacer number two. Since both plants show cutting by the enzyme in either one or both spacers, cutting by the Cas9 protein have been achieved in both plants. Therefore, both are considered as heterozygous PLL2 mutants. The last three plants showed no cutting by the enzyme for any of the spacers and must therefore be considered as either wild type or homozygous mutants. For knocking out PLL2, spacer number one seems most effective

In an attempt of knocking out PLL2 and PLL3 simultaneously, two extra constructs were made. Here, only one sgRNA was used for each gene. The first construct contained a sgRNA expression cassette with PLL3's spacer number one and PLL2's spacer number two. The second construct contained a cassette with PLL3's spacer number two and PLL2's spacer number one. The first two plants, containing PLL3' spacer one and PLL2's spacer two, showed no clear bands when running the T7E1 assay products on agarose gel. Only smearing of the bands can be observed. A smear can also be seen in plant five for PLL2's spacer number two. It can be concluded that some sort of reaction has occurred, but the results can not be used to distinguish heterozygous plants from homozygous- or wild type plants. Plant number three and four gave neither bands or smearing and are considered homozygous or wild type. In the plants containing PLL3's spacer number two and PLL2's spacer number one, only plant number one and two showed cutting by the enzyme. Plant number one showed cutting in both spacers, while plant number two only showed clear cutting in PLL3's spacer number two. Even though smearing appeared in PLL2's spacer number two in plant two, it can not be concluded that any cutting has occurred in this area. Therefor only plant one can be considered as a heterozygous PLL3/PLL2 double mutant. Since only PLL3's spacer one seemed to be cut in plant number two, this plant can only be considered as a heterozygous PLL3 mutant. For PLL3's spacer two, plant three and four showed smearing bands. In the same plants, no bands could be observed PLL2's spacer number one. In plant five, no cutting had occurred for either PLL3's spacer two or PLL2's spacer one. Therefore, these plants are likely wild type or homozygotes.

No conclusions on whether the plants are heterozygous, homozygous or wild type can be drawn from this assay. As mentioned earlier the method is only used to measure the efficiency of editing and will only give an indication of whether the Cas9 has cut or not and the possibility of obtaining heterozygous plants. The first generation of transformed plants were expected to yield mostly wild type or heterozygotes, so the chances of finding a homozygote plant were small. Most likely, the plants in which showed no cutting by the enzyme in the T7E1 assay, are wild type plants. To check if

there could be any homozygotes between the wild types, another T7E1 assay could have been done. By mixing the genomic DNA of these plants with the genomic DNA of a known wild type plant. A mismatch would appear in the hybridization process of the assay. This mismatch would then have been detected by the T7E1 enzyme, in which would have cleaved the strands.

To draw any conclusion in whether the editing of genes had been successful or not, the target sites in the plants would have needed to be sequenced. Due to time limitations, this could not be done.

5.5 Transformation of tomato cotyledons by *Agrobacterium*

In the attempt of transforming tomato *S. lycopersicum* to knock out *PSY2L*, cotyledons were inoculated in transformed *Agrobacterium.* When plated on selection plates containing antibiotics, the carbenicillin were not able to kill the excess *Agrobacterium* leading to overgrowth and prevention of cotyledons to grow any callus. Compared to the concentration of carbenicillin used in McCormick (1997), the amount used for this selection was five times lower. This was probably the cause of overgrowth and could possibly be prevented by higher concentrations. The next step would have been to increase the carbenicillin concentration to see if that would inhibit the growth of *Agrobacterium*. Due to time-limitations no further attempts were made in growing callus.

5.6 Expression cassettes and cloning of recombinant vector for transformation of protoplasts

In the plans of knocking out *PSY2L*, *PAP7*, *PLL3* and *PLL2* in tomato *S. lycopersicum* and *Arabidopsis* in-vivo by PEG transfection of protoplasts, constructs were successfully made for all genes except from *PLL2*. Several attempts with several colonies were used to get a successful construct for *PLL2*, but due to failed attempts of establishing protocols for protoplast isolation and later time-limitations, the rest of this experiment were shelved. This made it impossible to verify if the sgRNAs designed to target the specific genes would function or not.

6 Conclusion

In this study we have successfully assembled and cloned recombinant plasmids containing one or two sgRNA expression cassettes for targeting of peroxisomal protein phosphatases, such as PAP7, PLL3, PLL2 and PLL3/PLL2 in *Arabidopsis* and the putative regulator of protein phosphatase 4 (PSY2L) in tomato.

Heterozygous *Arabidopsis* T1 plants have been obtained by transforming WT *Arabidopsis* using *Agrobacterium* mediated flower dipping. But further testing needs to be performed to verify the results.

Successful transformation of *S. lycopersicum* could not be verified due to problems with the antibiotics in the selection medium. Whether the sgRNAs assembled into pHSN401 vector for transformation of protoplasts would be functional or not could

7 References

Addgene (2017) CRISPR/Cas9 Guide. Addgene. https://www.addgene.org/crispr/guide/. Accessed May 12 2017

Baulcombe DC (1999) Fast forward genetics based on virus-induced gene silencing. Curr Opin Plant Biol 2 (2):109-113. doi:10.1016/s1369-5266(99)80022-3

Bibikova M, Beumer K, Trautman JK, Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. Science 300 (5620):764. doi:10.1126/science.1079512

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326 (5959):1509-1512. doi:10.1126/science.1178811

Bozzo GG, Raghothama KG, Plaxton WC (2002) Purification and characterization of two secreted purple acid phosphatase isozymes from phosphate-starved tomato (Lycopersicon esculentum) cell cultures. European Journal of Biochemistry 269 (24):6278-6286. doi:10.1046/j.1432-1033.2002.03347.x

BOZZO GG, RAGHOTHAMA KG, PLAXTON WC (2004) Structural and kinetic properties of a novel purple acid phosphatase from phosphate-starved tomato (Lycopersicon esculentum) cell cultures. Biochemical Journal 377 (2):419-428. doi:10.1042/bj20030947

Brooks C, Nekrasov V, Lippman ZB, Van Eck J (2014) Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. Plant physiology 166 (3):1292-1297. doi:10.1104/pp.114.247577

Chung CT, Niemela SL, Miller RH (1989) One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. Proceedings of the National Academy of Sciences of the United States of America 86 (7):2172-2175

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant journal : for cell and molecular biology 16 (6):735-743

Cohen PT, Philp A, Vazquez-Martin C (2005) Protein phosphatase 4--from obscurity to vital functions. FEBS Lett 579 (15):3278-3286. doi:10.1016/j.febslet.2005.04.070

De Duve C, Baudhuin P (1966) Peroxisomes (microbodies and related particles). Physiological reviews 46 (2):323-357

Del Pozo JC, Allona I, Rubio V, Leyva A, De La Peña A, Aragoncillo C, Paz-Ares J (1999) A type 5 acid phosphatase gene from Arabidopsis thaliana is induced by phosphate starvation and by some other types of phosphate mobilising/oxidative stress conditions. The Plant Journal 19 (5):579-589. doi:10.1046/j.1365-313X.1999.00562.x

Del Vecchio HA, Ying S, Park J, Knowles VL, Kanno S, Tanoi K, She Y-M, Plaxton WC (2014) The cell wall-targeted purple acid phosphatase AtPAP25 is critical for acclimation of Arabidopsis thaliana to nutritional phosphorus deprivation. The Plant Journal 80 (4):569-581. doi:10.1111/tpj.12663

Durmus A, Eicken C, Horst Sift B, Kratel A, Kappl R, Hüttermann J, Krebs B (1999) The active site of purple acid phosphatase from sweet potatoes (Ipomoea batatas). European Journal of Biochemistry 260 (3):709-716. doi:10.1046/j.1432-1327.1999.00230.x

Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high throughput capability. PLoS One 3 (11):e3647. doi:10.1371/journal.pone.0003647

Feng Z, Mao Y, Xu N, Zhang B, Wei P, Yang DL, Wang Z, Zhang Z, Zheng R, Yang L, Zeng L, Liu X, Zhu JK (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 111 (12):4632-4637. doi:10.1073/pnas.1400822111

Flanagan JU, Cassady AI, Schenk G, Guddat LW, Hume DA (2006) Identification and molecular modeling of a novel, plant-like, human purple acid phosphatase. Gene 377:12-20. doi:10.1016/j.gene.2006.02.031

Hooper CM, Tanz SK, Castleden IR, Vacher MA, Small ID, Millar AH (2014) SUBAcon: a consensus algorithm for unifying the subcellular localization data of the Arabidopsis proteome. Bioinformatics 30 (23):3356-3364. doi:10.1093/bioinformatics/btu550

Huang (2014) CRISPR gRNA Expression Vector Cloning Protocol. Addgene. https://media.addgene.org/data/plasmids/61/61089/61089-attachment_auIxoXmMvBFA.pdf. Accessed May 12 2017

Jiang F, Doudna JA (2017) CRISPR-Cas9 Structures and Mechanisms. Annual review of biophysics. doi:10.1146/annurev-biophys-062215-010822

Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic Acids Res 41 (20):e188. doi:10.1093/nar/gkt780

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337 (6096):816-821. doi:10.1126/science.1225829

Kataya AR, Heidari B, Hagen L, Kommedal R, Slupphaug G, Lillo C (2015a) Protein phosphatase 2A holoenzyme is targeted to peroxisomes by piggybacking and positively affects peroxisomal beta-oxidation. Plant physiology 167 (2):493-506. doi:10.1104/pp.114.254409

Kataya AR, Schei E, Lillo C (2015b) MAP kinase phosphatase 1 harbors a novel PTS1 and is targeted to peroxisomes following stress treatments. Journal of plant physiology 179:12-20. doi:10.1016/j.jplph.2015.03.002

Kataya AR, Schei E, Lillo C (2016) Towards understanding peroxisomal phosphoregulation in Arabidopsis thaliana. Planta 243 (3):699-717. doi:10.1007/s00425-015-2439-5

Kaur N, Reumann S, Hu J (2009) Peroxisome biogenesis and function. The Arabidopsis book 7:e0123. doi:10.1199/tab.0123

Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E (2016) CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic Acids Res 44 (W1):W272-276. doi:10.1093/nar/gkw398

Lei Y, Lu L, Liu HY, Li S, Xing F, Chen LL (2014) CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants. Molecular plant 7 (9):1494-1496. doi:10.1093/mp/ssu044

Li D, Zhu H, Liu K, Liu X, Leggewie G, Udvardi M, Wang D (2002) Purple acid phosphatases of Arabidopsis thaliana. Comparative analysis and differential regulation by phosphate deprivation. J Biol Chem 277 (31):27772-27781. doi:10.1074/jbc.M204183200

Li JF, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J (2013) Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nature biotechnology 31 (8):688-691. doi:10.1038/nbt.2654

Li Z, Liu ZB, Xing A, Moon BP, Koellhoffer JP, Huang L, Ward RT, Clifton E, Falco SC, Cigan AM (2015) Cas9-Guide RNA Directed Genome Editing in Soybean. 169 (2):960-970. doi:10.1104/pp.15.00783

Lillo C, Kataya AR, Heidari B, Creighton MT, Nemie-Feyissa D, Ginbot Z, Jonassen EM (2014) Protein phosphatases PP2A, PP4 and PP6: mediators and regulators in development and responses to environmental cues. Plant, cell & environment 37 (12):2631-2648. doi:10.1111/pce.12364

Ma X, Zhang Q, Zhu Q, Liu W, Chen Y, Qiu R, Wang B, Yang Z, Li H, Lin Y, Xie Y, Shen R, Chen S, Wang Z, Chen Y, Guo J, Chen L, Zhao X, Dong Z, Liu YG (2015) A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. Molecular plant 8 (8):1274-1284. doi:10.1016/j.molp.2015.04.007

Ma X, Zhu Q, Chen Y, Liu YG (2016) CRISPR/Cas9 Platforms for Genome Editing in Plants: Developments and Applications. Molecular plant 9 (7):961-974. doi:10.1016/j.molp.2016.04.009

Matre P, Meyer C, Lillo C (2009) Diversity in subcellular targeting of the PP2A B'eta subfamily members. Planta 230 (5):935-945. doi:10.1007/s00425-009-0998-z

McCormick S (1997) Transformation of tomato with Agrobacterium tumefaciens. In: Lindsey K (ed) Plant Tissue Culture Manual: Supplement 7. Springer Netherlands, Dordrecht, pp 311-319. doi:10.1007/978-94-009-0103-2_17
Mol JN, van der Krol AR, van Tunen AJ, van Blokland R, de Lange P, Stuitje AR (1990) Regulation of plant gene expression by antisense RNA. FEBS Lett 268 (2):427-430
Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326 (5959):1501. doi:10.1126/science.1178817
Murashige T, Skoog F (1962) A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. Physiologia Plantarum 15 (3):473-497. doi:10.1111/j.1399-3054.1962.tb08052.x
Napitupulu TP (2016) Investigation of the Physiological Function of Protein Phosphatase 4 (PP4) in Arabidopsis thaliana. University of Stavanger,
NEB (2016) NEB: Determining Genome Targeting Efficiency using T7 Endonuclease I (M0302). . protocols.io. https://www.protocols.io/view/Determining-Genome-Targeting-Efficiency-using-T7-E-etbbein. Accessed May 2017 2017
Park J, Kim J-S, Bae S (2016) Cas-Database: web-based genome-wide guide RNA library design for gene knockout screens using CRISPR-Cas9. Bioinformatics 32 (13):2017-2023. doi:10.1093/bioinformatics/btw103
Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013) Genome engineering using the CRISPR-Cas9 system. Nature protocols 8 (11):2281-2308. doi:10.1038/nprot.2013.143
Schenk G, Ge Y, Carrington LE, Wynne CJ, Searle IR, Carroll BJ, Hamilton S, de Jersey J (1999) Binuclear Metal Centers in Plant Purple Acid Phosphatases: Fe–Mn in Sweet Potato and Fe–Zn in Soybean. Archives of Biochemistry and Biophysics 370 (2):183-189. doi:http://dx.doi.org/10.1006/abbi.1999.1407
Schenk G, Guddat LW, Ge Y, Carrington LE, Hume DA, Hamilton S, de Jersey J (2000a) Identification of mammalian-like purple acid phosphatases in a wide range of plants. Gene 250 (1-2):117-125
Schenk G, Korsinczky ML, Hume DA, Hamilton S, DeJersey J (2000b) Purple acid phosphatases from bacteria: similarities to mammalian and plant enzymes. Gene 255 (2):419-424
Schenk G, Mitić N, Hanson GR, Comba P (2013) Purple acid phosphatase: A journey into the function and mechanism of a colorful enzyme. Coordination Chemistry Reviews 257 (2):473-482. doi:https://doi.org/10.1016/j.ccr.2012.03.020
Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL, Gao C (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. Nature biotechnology 31 (8):686-688. doi:10.1038/nbt.2650
Shui JW, Hu MC, Tan TH (2007) Conditional knockout mice reveal an essential role of protein phosphatase 4 in thymocyte development and pre-T-cell receptor signaling. Molecular and cellular biology 27 (1):79-91. doi:10.1128/mcb.00799-06
Smith NA, Singh SP, Wang MB, Stoutjesdijk PA, Green AG, Waterhouse PM (2000) Total silencing by intron-spliced hairpin RNAs. Nature 407 (6802):319-320. doi:10.1038/35030305
Song SK, Clark SE (2005) POL and related phosphatases are dosage-sensitive regulators of meristem and organ development in Arabidopsis. Developmental biology 285 (1):272-284. doi:10.1016/j.ydbio.2005.06.020
Springer NM (2010) Isolation of plant DNA for PCR and genotyping using organic extraction and CTAB. Cold Spring Harbor protocols 2010 (11):pdb.prot5515. doi:10.1101/pdb.prot5515
Stark MJ (1996) Yeast protein serine/threonine phosphatases: multiple roles and diverse regulation. Yeast 12 (16):1647-1675. doi:10.1002/(SICI)1097-0061(199612)12:16<1647::AID-YEA71>3.0.CO;2-Q
Stemmer M, Thumberger T, del Sol Keyer M, Wittbrodt J, Mateo JL (2015) CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. PLOS ONE 10 (4):e0124633. doi:10.1371/journal.pone.0124633
Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature 507 (7490):62-67. doi:10.1038/nature13011

Svitashev S, Young JK, Schwartz C, Gao H, Falco SC, Cigan AM (2015) Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. 169 (2):931-945. doi:10.1104/pp.15.00793

Terns MP, Terns RM (2011) CRISPR-based adaptive immune systems. Current opinion in microbiology 14 (3):321-327. doi:10.1016/j.mib.2011.03.005

Tran HT, Hurley BA, Plaxton WC (2010) Feeding hungry plants: The role of purple acid phosphatases in phosphate nutrition. Plant Science 179 (1–2):14-27. doi:https://doi.org/10.1016/j.plantsci.2010.04.005

Uhrig RG, Labandera AM, Moorhead GB (2013) Arabidopsis PPP family of serine/threonine protein phosphatases: many targets but few engines. Trends Plant Sci 18 (9):505-513. doi:10.1016/j.tplants.2013.05.004

Veljanovski V, Vanderbeld B, Knowles VL, Snedden WA, Plaxton WC (2006) Biochemical and Molecular Characterization of AtPAP26, a Vacuolar Purple Acid Phosphatase Up-Regulated in Phosphate-Deprived Arabidopsis Suspension Cells and Seedlings. Plant physiology 142 (3):1282-1293. doi:10.1104/pp.106.087171

Wang K (2006) Agrobacterium protocols, vol 1. Springer,

Wang L, Li Z, Qian W, Guo W, Gao X, Huang L, Wang H, Zhu H, Wu J-W, Wang D, Liu D (2011) The Arabidopsis Purple Acid Phosphatase AtPAP10 Is Predominantly Associated with the Root Surface and Plays an Important Role in Plant Tolerance to Phosphate Limitation. Plant physiology 157 (3):1283-1299. doi:10.1104/pp.111.183723

Wang ZP, Xing HL, Dong L, Zhang HY, Han CY, Wang XC, Chen QJ (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome biology 16:144. doi:10.1186/s13059-015-0715-0

Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. Nature 482 (7385):331-338. doi:10.1038/nature10886

Woo JW, Kim J, Kwon SI, Corvalan C, Cho SW, Kim H, Kim SG, Kim ST, Choe S, Kim JS (2015) DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. Nature biotechnology 33 (11):1162-1164. doi:10.1038/nbt.3389

Xing HL, Dong L, Wang ZP, Zhang HY, Han CY, Liu B, Wang XC, Chen QJ (2014) A CRISPR/Cas9 toolkit for multiplex genome editing in plants. BMC plant biology 14:327. doi:10.1186/s12870-014-0327-y

Xue T, Wang D, Zhang S, Ehlting J, Ni F, Jakab S, Zheng C, Zhong Y (2008) Genome-wide and expression analysis of protein phosphatase 2C in rice and Arabidopsis. BMC Genomics 9 (1):550. doi:10.1186/1471-2164-9-550

Yu LP, Miller AK, Clark SE (2003) POLTERGEIST encodes a protein phosphatase 2C that regulates CLAVATA pathways controlling stem cell identity at Arabidopsis shoot and flower meristems. Current biology : CB 13 (3):179-188

Zhang H, Zhang J, Wei P, Zhang B, Gou F, Feng Z, Mao Y, Yang L, Zhang H, Xu N, Zhu JK (2014) The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. Plant biotechnology journal 12 (6):797-807. doi:10.1111/pbi.12200

Zhang Q, Wang C, Tian J, Li K, Shou H (2011) Identification of rice purple acid phosphatases related to posphate starvation signalling. Plant Biology 13 (1):7-15. doi:10.1111/j.1438-8677.2010.00346.x

Zhang Z, Mao Y, Ha S, Liu W, Botella JR, Zhu JK (2016) A multiplex CRISPR/Cas9 platform for fast and efficient editing of multiple genes in Arabidopsis. Plant cell reports 35 (7):1519-1533. doi:10.1007/s00299-015-1900-z

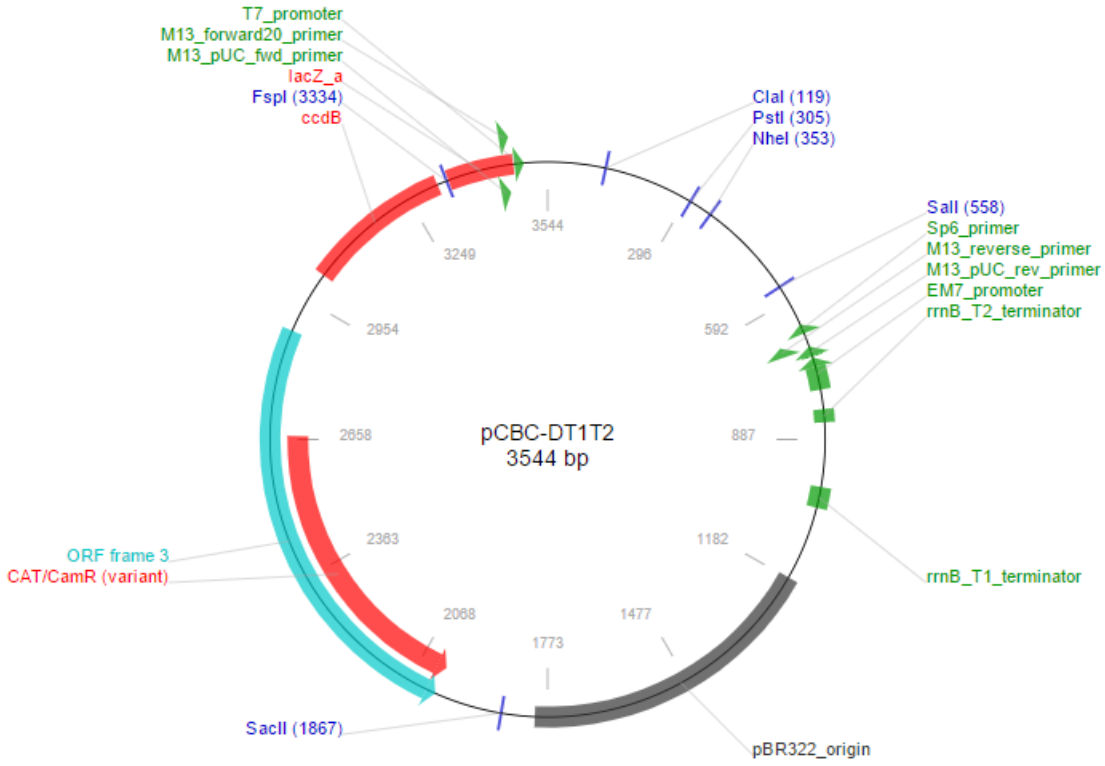
Zhou H, Liu B, Weeks DP, Spalding MH, Yang B (2014) Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. Nucleic Acids Res 42 (17):10903-10914. doi:10.1093/nar/gku806

Appendix

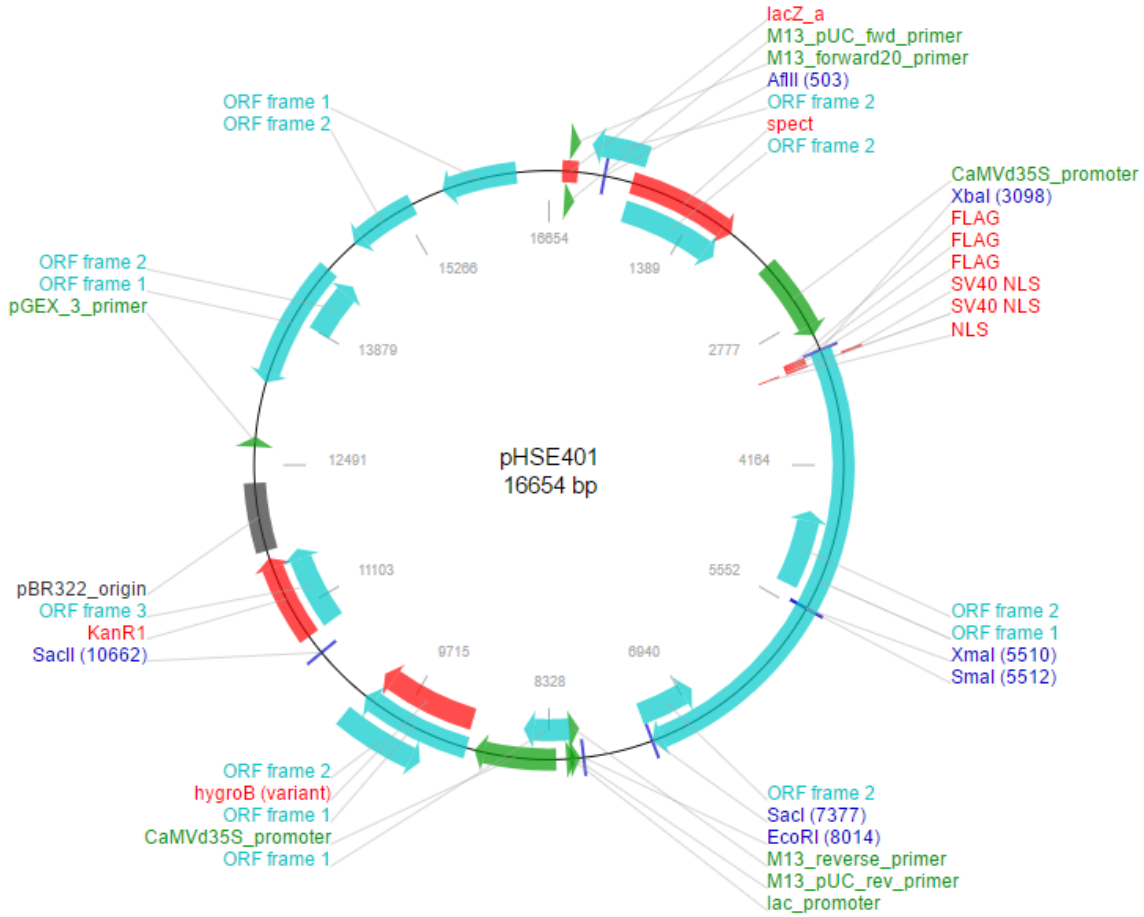
A-1 Plasmid maps

The plasmid maps were created by Addgene from the full sequences supplied by Qi-Jun Chen.

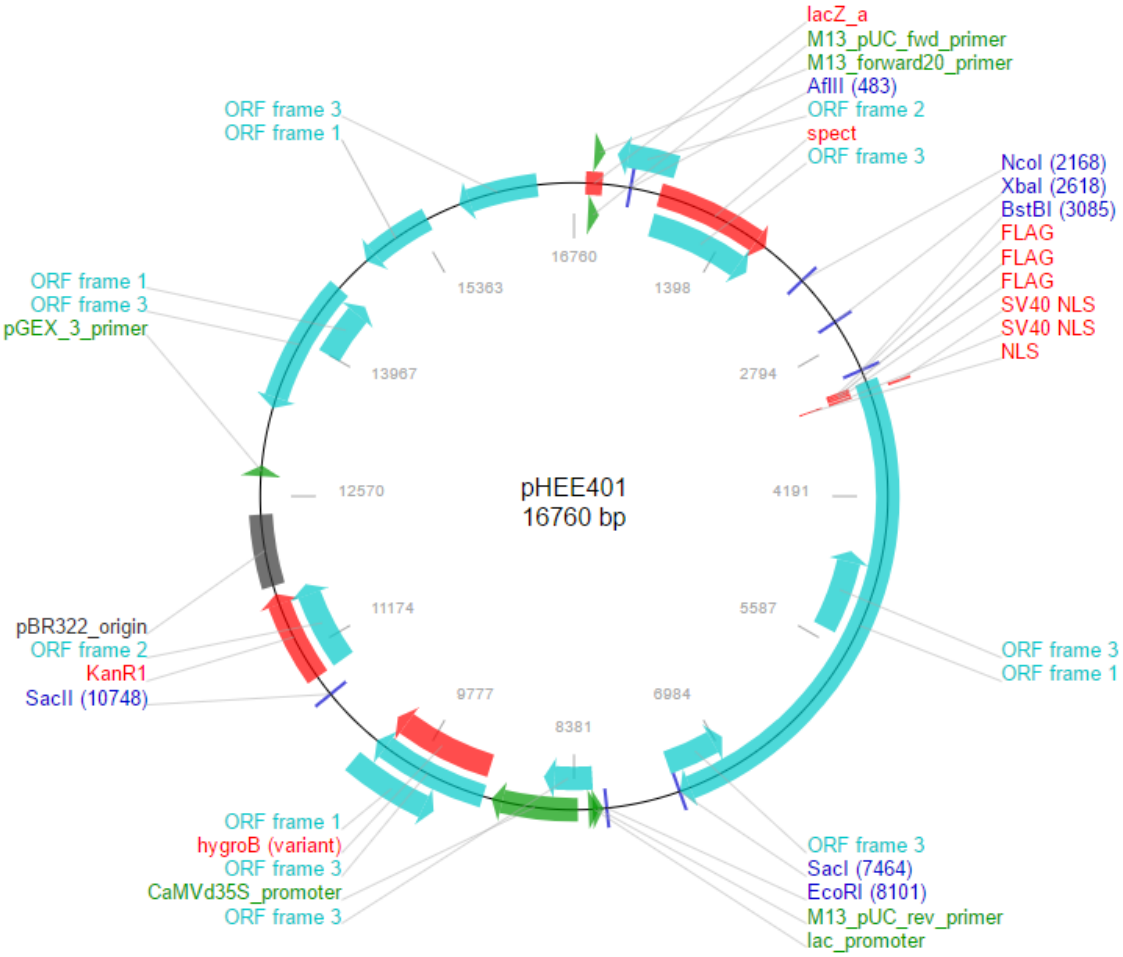
Plasmid map of pCBC-DT1T2



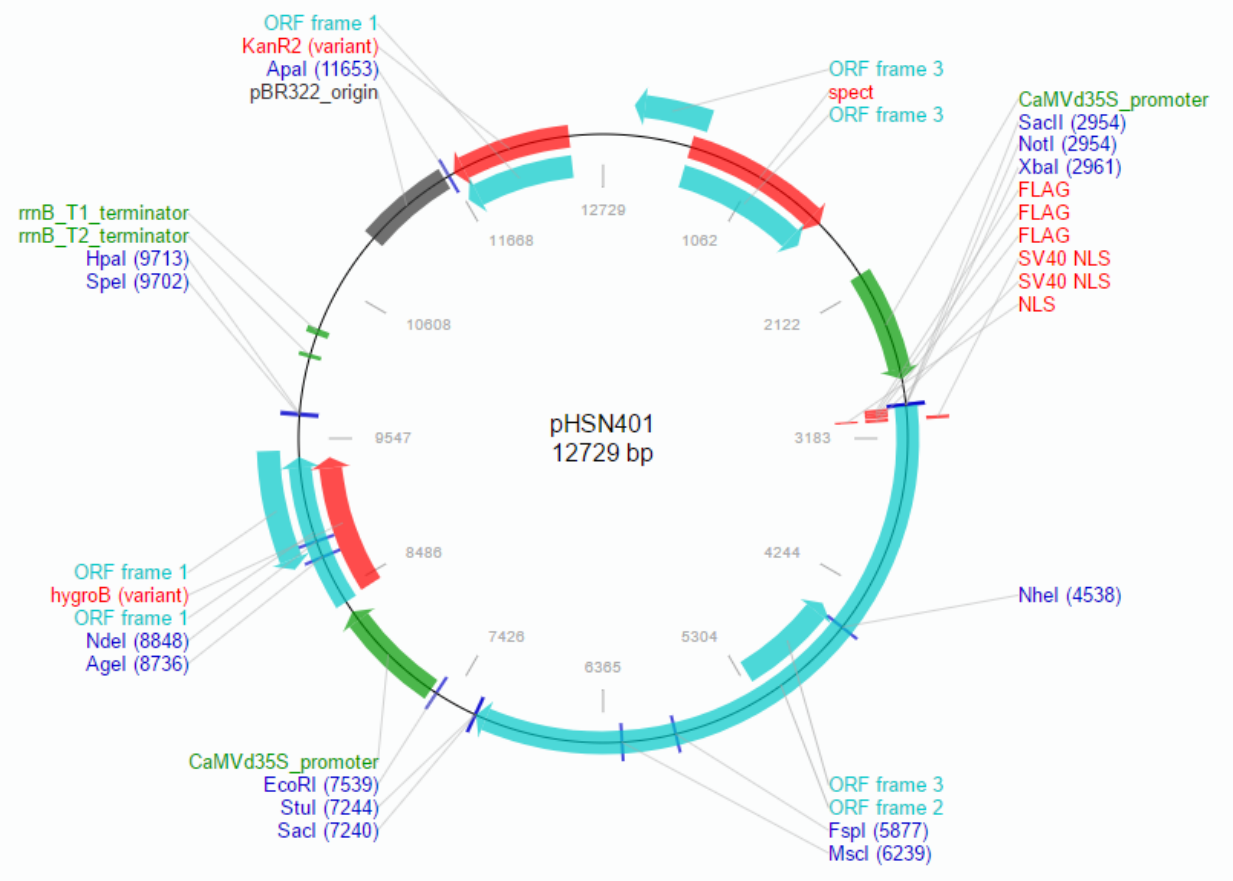
Plasmid map of pHSE401



Plasmid map of pHEE401



Plasmid map of pHSN401

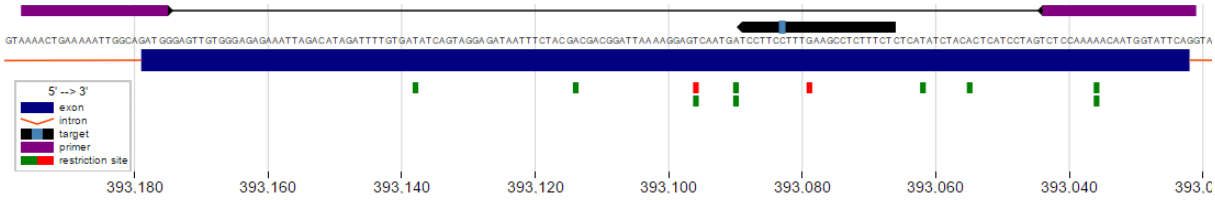


A-2 Primers for amplification around the target site of extracted DNA from T1 plants.

Primers suggested by CHOPCHOP to amplify around the target sites to make products suitable for T7E1 assay. The black area show the sequence complementary to the spacer and the area in which T7 endonuclease will cut.

PAP7 spacer 1

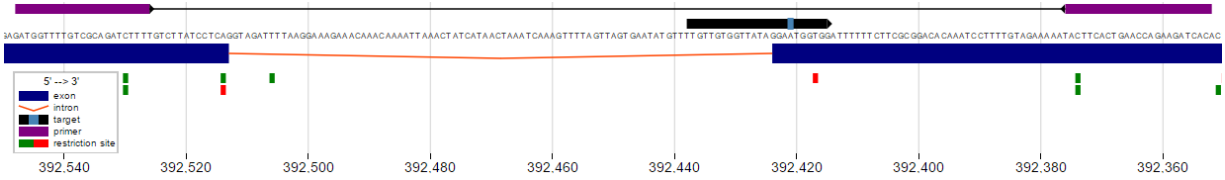
Gene specific part of sgRNA
GAGAAAGAGGCTTCAAAGGAAGG
There are no predicted off-targets for this guide



Pair	Left primer coordinates	Left primer	Left primer Tm	Left primer off-targets	Right primer coordinates	Right primer	Right primer Tm	Right primer off-targets	Pair off-targets	Product size
1	Chr2:393176-393198	AAACTGAAAAATTGGCAGATGG	60.3	0	Chr2:393022-393045	CCTGAATACCATTGTTTTTGGAG	59.8	0	0	176

PAP7 spacer 2

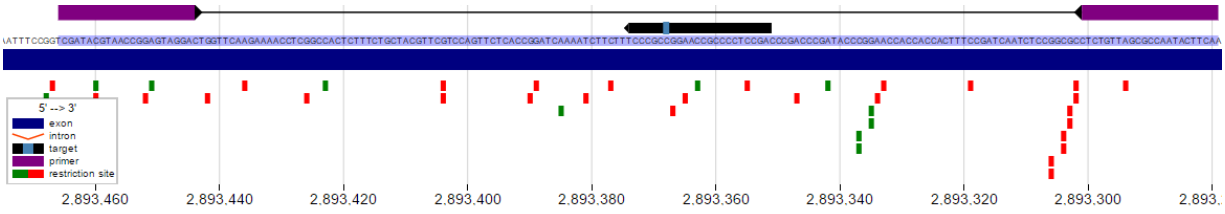
Gene specific part of sgRNA
TGTTGTGGTTATAGGAATGGTGG
There are no predicted off-targets for this guide



Pair	Left primer coordinates	Left primer	Left primer Tm	Left primer off-targets	Right primer coordinates	Right primer	Right primer Tm	Right primer off-targets	Pair off-targets	Product size
1	Chr2:392527-392549	ATGGTTTTGTCGCAGATCTTTT	60.0	0	Chr2:392353-392377	TGTGATCTTCTGGTTCAGTGAAGT	60.2	0	0	196

PLL3 spacer 1

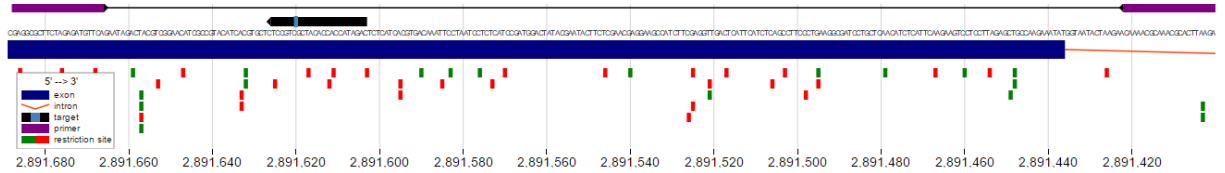
Gene specific part of sgRNA
GTCGGAGGGGCGGTTCCGGCGGG
There are no predicted off-targets for this guide



Pair	Left primer coordinates	Left primer	Left primer Tm	Left primer off-targets	Right primer coordinates	Right primer	Right primer Tm	Right primer off-targets	Pair off-targets	Product size
1	Chr3:2893445-2893467	CGATACGTAACCGGAGTAGGAC	59.9	0	Chr3:2893280-2893302	TTGAAGTATTGGCGCTAACAGA	59.9	0	0	187

PLL3 spacer 2

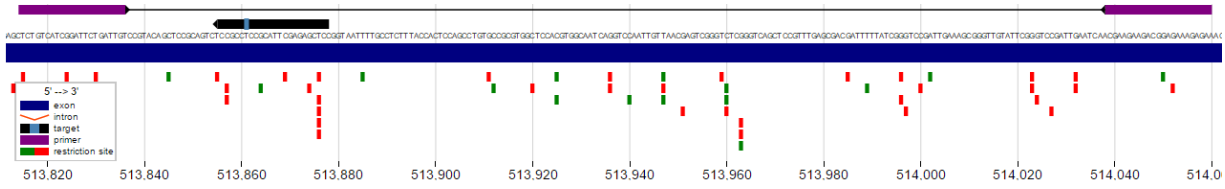
Gene specific part of sgRNA
AGTCTATGGTGGTGTAGCGACGG
There are no predicted off-targets for this guide



Pair	Left primer coordinates	Left primer	Left primer Tm	Left primer off-targets	Right primer coordinates	Right primer	Right primer Tm	Right primer off-targets	Pair off-targets	Product size
1	Chr3:2891667-2891689	AGGCGCTTCTAGAGATGTTCAG	60.2	0	Chr3:2891401-2891423	ATCTTAAGTGCGTTTGCGTTTT	60.2	0	0	288

PLL2 spacer 1

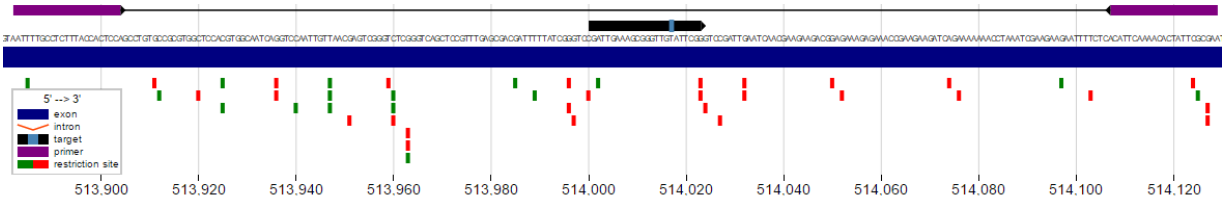
Gene specific part of sgRNA
GGAGCTCTCGAATGCGGAGGCGG
There are no predicted off-targets for this guide



Pair	Left primer coordinates	Left primer	Left primer Tm	Left primer off-targets	Right primer coordinates	Right primer	Right primer Tm	Right primer off-targets	Pair off-targets	Product size
1	Chr5:513815-513837	TCTGTCATCGGATTCTGATTGT	59.6	0	Chr5:514039-514061	TTCTCTTTCTCCGTCTTCTTCG	60.1	0	0	246

PLL2 spacer 2

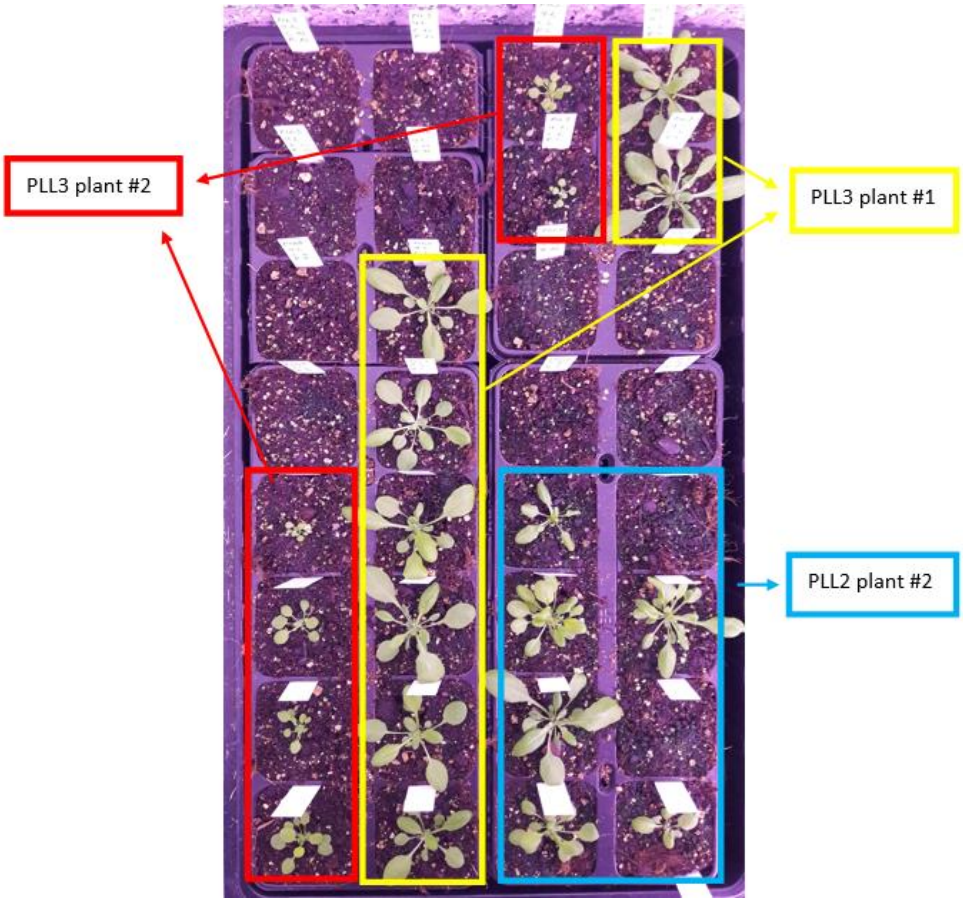
Gene specific part of sgRNA
GATTGAAAGCGGGTTGTATTCGG
There are no predicted off-targets for this guide

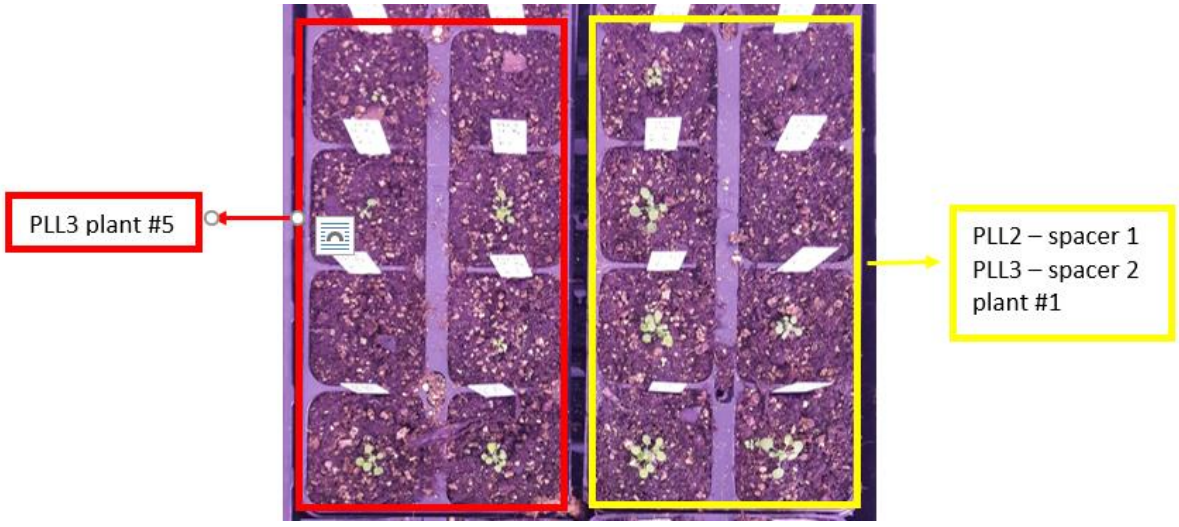


Pair	Left primer coordinates	Left primer	Left primer Tm	Left primer off-targets	Right primer coordinates	Right primer	Right primer Tm	Right primer off-targets	Pair off-targets	Product size
1	Chr5:513883-513905	ATTTTGCCTCTTTACCACTCCA	60.0	0	Chr5:514108-514130	TTCGCGAATAGTGTTTTGAATG	60.1	0	0	247

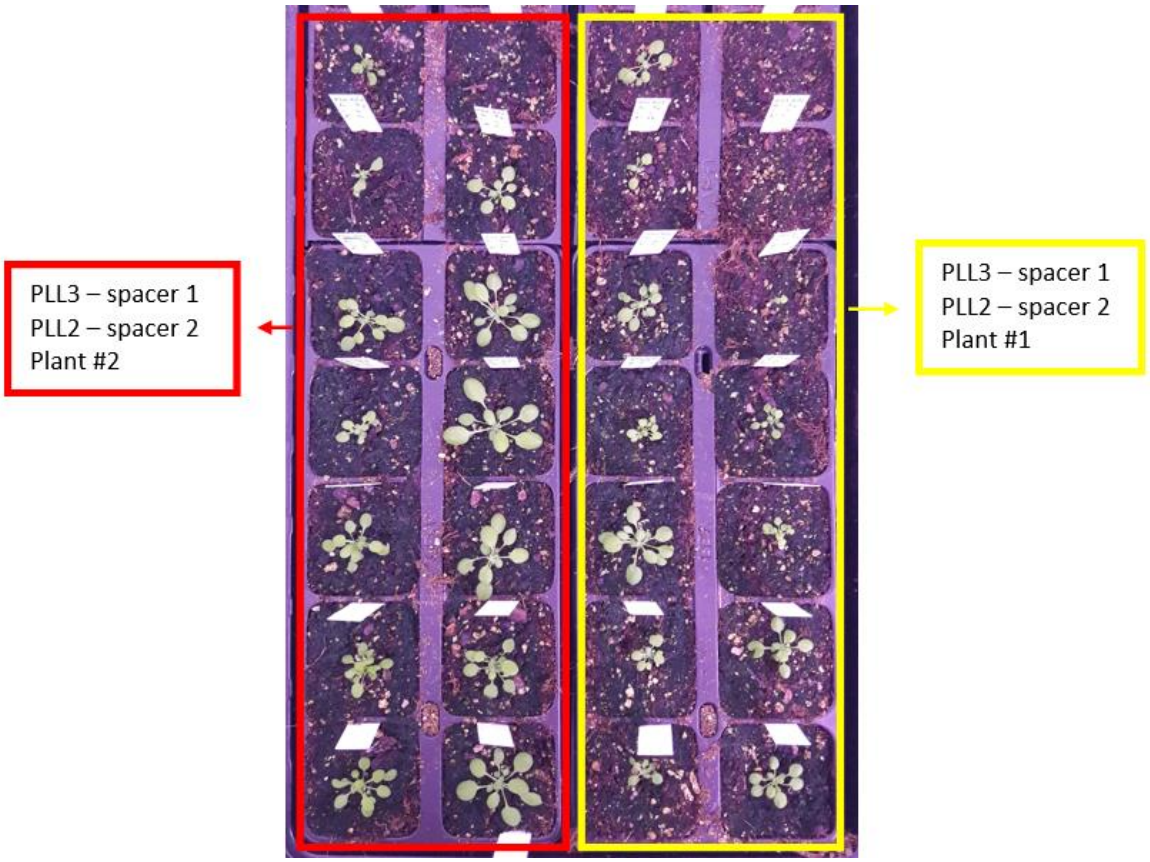
A-3 Phenotyping T2 generation

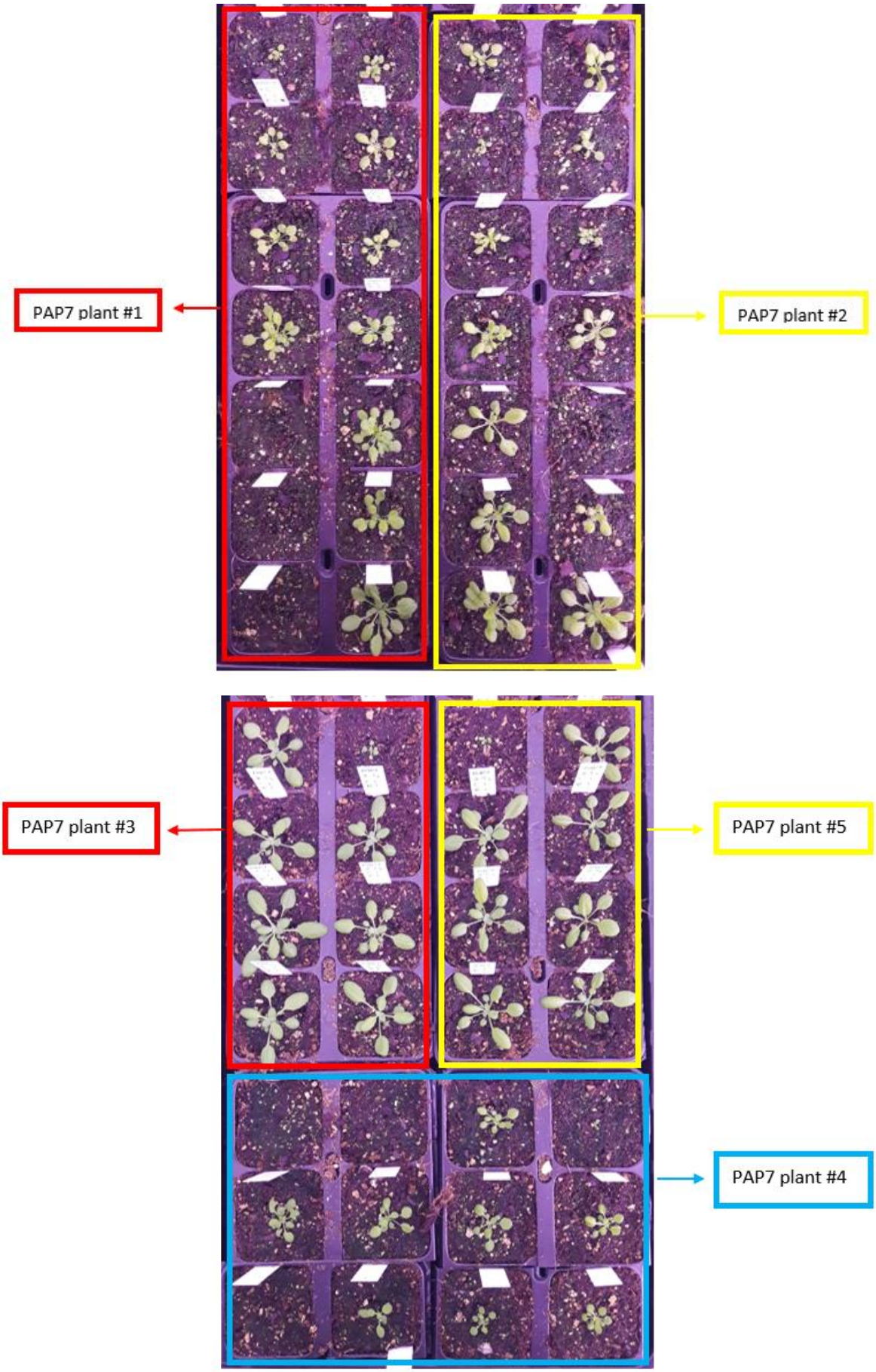
Phenotypes of 4-6 week old T2 plants from plants of the T1 generation which showed cutting by T7 endonuclease 1 in the area of one or both spacers.

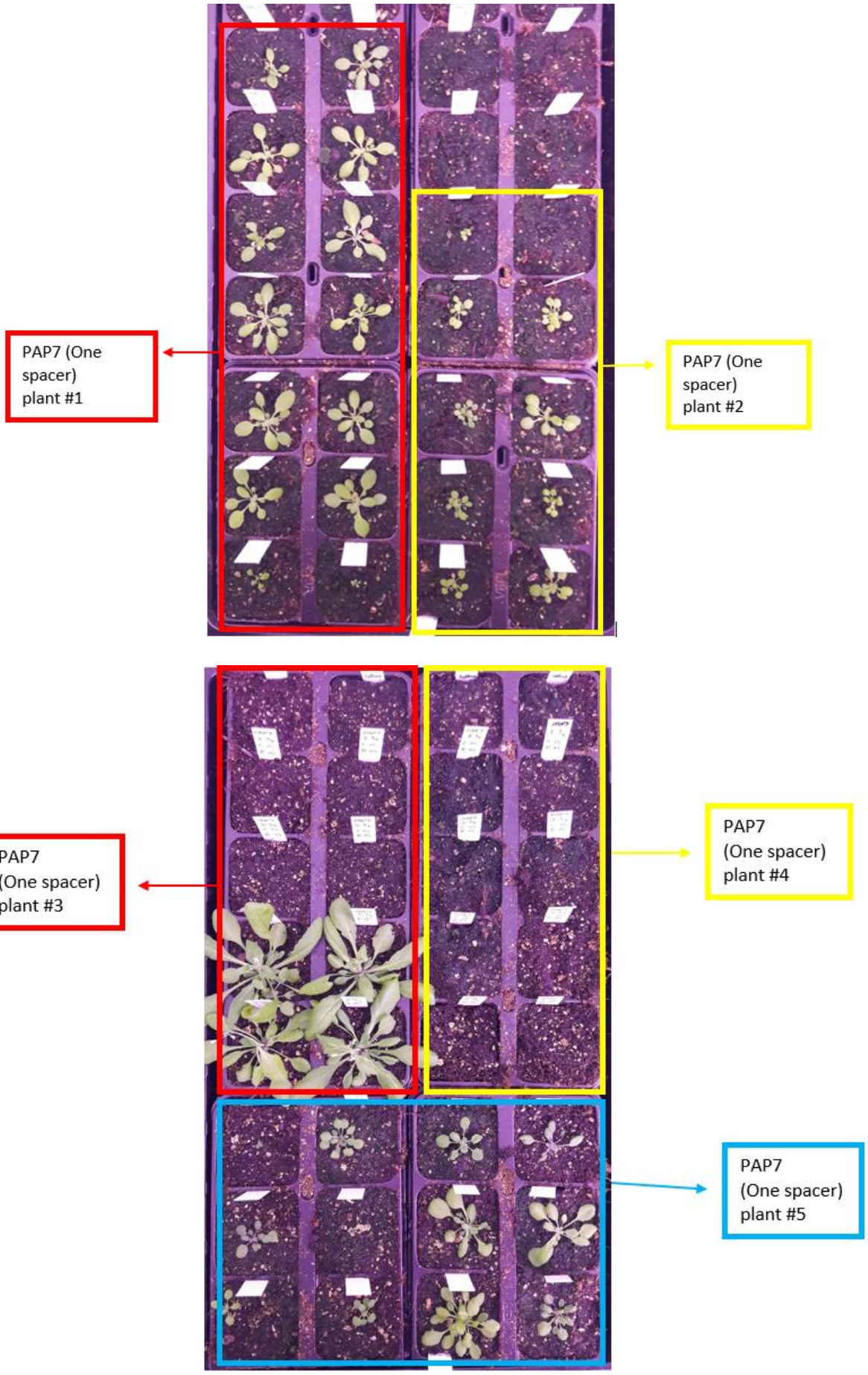





PAP7 (One spacer) plant #6
PLL2 plant # 1


PLL3 – spacer 1
PLL2 – spacer 2
Plant #2
PLL3 – spacer 1
PLL2 – spacer 2
Plant #1


PAP7 plant #1
PAP7 plant #2
PAP7 plant #3
PAP7 plant #5
PAP7 plant #4


PAP7 (One spacer) plant #1
PAP7 (One spacer) plant #2
PAP7 (One spacer) plant #3
PAP7 (One spacer) plant #4
PAP7 (One spacer) plant #5

A-4 Sequencing- and colony PCR primers

Table 25: Primers for sequencing and colony PCR

Primer name		Primer sequence
CRISPR_39	U6-26p-F	TGTCCCAGGATTAGAATGATTAGGC
CRISPR_40	U6-26t-R	CCCCAGAAATTGAACGCCGAAGAAC
CRISPR_41	U6-29p-F	TTAATCCAAACTACTGCAGCCTGAC
CRISPR_42	U6-29p-R	AGCCCTCTTCTTTCGATCCATCAAC