University of New Hampshire
University of New Hampshire Scholars' Repository

Honors Theses and Capstones | Student Scholarship

Spring 2023

Investigation of the Function of Protein Acyl Transferases using CRISPR-Cas9 to Create Null Mutants

Claire E. Christopher
University of New Hampshire, Durham

Follow this and additional works at: https://scholars.unh.edu/honors

 Part of the Biology Commons, Genetics Commons, and the Plant Biology Commons

Recommended Citation

Christopher, Claire E., "Investigation of the Function of Protein Acyl Transferases using CRISPR-Cas9 to Create Null Mutants" (2023). *Honors Theses and Capstones*. 738.
https://scholars.unh.edu/honors/738

This Senior Honors Thesis is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Honors Theses and Capstones by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact Scholarly.Communication@unh.edu.

Investigation of the Function of Protein Acyl Transferases using CRISPR-Cas9 to Create Null Mutants

Claire Christopher

PI: Dr. Estelle Hrabak

University of New Hampshire, Durham, NH

Genetics Honors Thesis

Abstract

24 Protein Acyl Transferases (PATs) have been identified in the model organism *Arabidopsis thaliana*. Despite knowing the enzymatic function of these genes (palmitoylation), the specific subcellular pathways, protein targets, and cellular activities of these proteins remain unknown. To investigate the pathways that *PAT* genes are involved in, deletion mutations were created in several genes and the phenotype of the organisms carrying homozygous mutations was observed. To introduce mutations, the CRISPR/Cas9 system was inserted into the plant's genome using *Agrobacterium tumefaciens* to cause deletions in *PAT* genes of interest. To date, no CRISPR-induced *pat* mutants have a noticeable phenotype. These results will focus on *PAT14*; identification of mutants in other *PAT* genes is underway but will not be presented in detail here.

Introduction

Proteins can be modified by the addition of various types of molecules such as lipids, sugars, phosphate groups, or methyl groups. Protein modifications can result in the protein becoming activated, inactivated, targeted to new locations in the cell, destabilized, or affected in the ability to interact with other molecules (Ramazi and Zahiri, 2021). Protein modifications are vital to the regulation of cellular processes and the generation of protein diversity. Many protein modification processes such as phosphorylation and glycosylation are ubiquitous across all three domains of life (Beltrao et al., 2013) making the mechanism and impacts of protein modification of particular interest for study.

The addition of lipid molecules to proteins occurs via one of three processes: myristoylation, prenylation, or palmitoylation (S-acylation). In palmitoylation, which is the focus of this study, the hydrophobic 16-carbon fatty acid palmitate is covalently attached to cysteine residues (Jiang et al., 2018). The addition of palmitate causes cytosolic proteins to be targeted to lipid membranes in the cell including the plasma membrane, endoplasmic reticulum, or Golgi apparatus (Jiang et al., 2018). Palmitate can also be added to integral membrane proteins resulting in functional alteration of that protein. Palmitoylation is the only post-translational lipid modification process that is fully reversible which makes it particularly useful as a regulatory modification of proteins (Guan and Fierke, 2011).

Palmitoylation is catalyzed by a group of proteins referred to as Protein Acyl Transferases or PATs in plants. In other organisms such as humans, these proteins are commonly referred to as DHHC proteins due to the fact that the catalytic domain of palmitoyl transferases usually contains the four amino acids Asp-His-His-Cys (Jiang et al., 2018). Despite variations in naming, these enzymes perform the same function: palmitoylation. Protein Acyl Transferases

usually have four transmembrane domains and reside in membranes with the catalytic DHHC domain facing the cytosol. The lipid membranes in which PAT proteins reside vary depending on the specific PAT protein (Batistic 2012; Jiang et al., 2018).

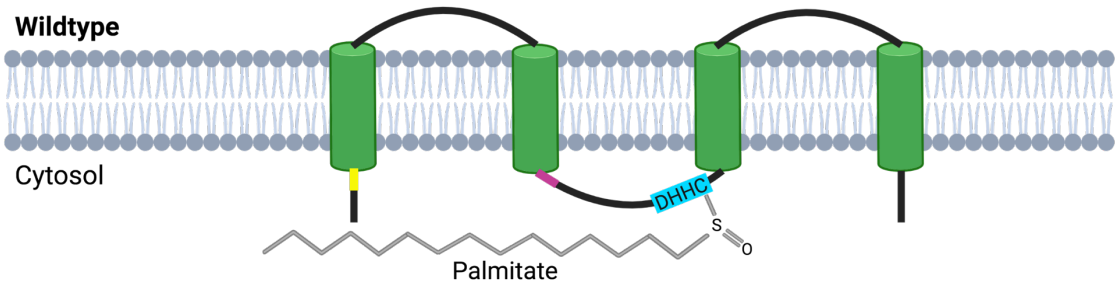

Figure 1. ***Protein acyl transferase primed with palmitate and ready to interact with a substrate (not shown)***. (created at Biorender.com)

In the model organism *Arabidopsis thaliana*, 24 *PAT* genes have been identified (Batistic, 2012). Despite knowing that palmitoylation is the general function of PAT proteins, the specific protein targets, subcellular pathways, and cellular activities that individual PAT proteins are involved in is unknown. One approach to determine which pathways PAT proteins are involved in is to observe the phenotype of organisms with mutations in a particular *PAT* gene of interest. Using this method, other research groups have identified potential functions for *PAT4*, *PAT10*, *PAT13*, *PAT14* and *PAT21*. Mutations in *PAT4* resulted in stunted root hair growth (Wan et al. 2017). Mutations in *PAT10* caused reduced fertility, reduced salt tolerance, and developmental defects of the vascular system (Qi et al. 2013; Zhou et al., 2013). Mutations in *PAT21* led to defects in root and pollen tube growth (Schiefelbein et al., 1993; Ryan et al., 1998; Hemsley et al., 2005). Mutations in *PAT14* resulted in early leaf senescence (Li et al. 2016; Zhao et al. 2016). Interestingly, a phenotype for *pat13* mutants was only observed when a mutation in the

PAT14 gene was also present indicating possible functional redundancy (Lai et al. 2015). To date, there is very little research on the other 19 *PAT* genes.

Expression of most *PAT* genes is relatively low during most of *A. thaliana*'s life cycle, but particular *PAT*s (*PAT1* and *PAT8*) are expressed at high levels in pollen (Batistic et al., 2012; Yuan et al., 2013). Other *PAT*s with high expression in pollen are *PAT2, PAT3, PAT4, PAT13, PAT16*, and *PAT24* (Batistic et al., 2012; Yuan et al., 2013). Of particular interest is *PAT3* which has been reported to be expressed exclusively in pollen (Batistic, 2012). Although multiple *PAT*s have high expression in pollen, a pollen phenotype has only been reported in a *pat24* mutant (Schiefelbein et al., 1993; Ryan et al., 1998).

T-DNA-induced mutants for many of the pollen-expressed *PAT* genes were obtained from the Arabidopsis Stock Center or other sources. Interestingly, none of the homozygous *pat* mutants exhibited any mutant phenotype (E. Hrabak, personal communication). One possible explanation for this result is that the mutations in these genes do not fully abolish gene function based on the location of the T-DNA in the gene. An additional concern was that the T-DNAs used to make the mutants all contain outward-facing promoters meaning that the T-DNAs weren't merely "inert" mutagens but had the potential to drive transcription from the point of insertion (Ulker et al., 2008). Previous research in the Hrabak lab identified RNA transcripts from the *PAT* genes containing T-DNA insertions in some of the mutants. Ideally, RNA transcripts would be absent in a null mutant indicating that the entire gene was rendered nonfunctional. If the T-DNA-mutated gene was able to make an altered but still functional gene product, then the mutants might retain a wildtype phenotype and prove non-helpful in understanding gene function.

We decided to take a different approach to creating bona fide null mutants using CRISPR/Cas9 (Clustered Repetitive Short Interspaced Palindromic Repeats) technology. Cas9 is a naturally occurring endonuclease native to bacteria that functions in conjunction with a CRISPR array in the bacterial genome as an immune defense mechanism against viruses (Arroyo-Olarte et al., 2021). In bacteria, Cas9 recognizes and destroys double-stranded viral DNA that enters the bacterial cell if the cell had previously survived infection with that virus. Cas9 is guided to the invading viral DNA by a tracrRNA and a crRNA created from the CRISPR array of viral DNA fragments stored in the bacterial genome.

The CRISPR/Cas9 mechanism has been modified for use in eukaryotes in several ways. First, no genomic CRISPR array is required. Second, the crRNA and tracrRNA have been combined into a single guide RNA (Figure 2). This powerful technology can now be used to create deletion mutations in a variety of organisms. Cas9 specifically makes double-stranded DNA breaks near the PAM site adjacent to the guide RNA sequence within the host's genome. PAM sites for Cas9 are simply the nucleotide sequence 5'NGG (where N could be any of the 4 possible nucleotides in DNA).

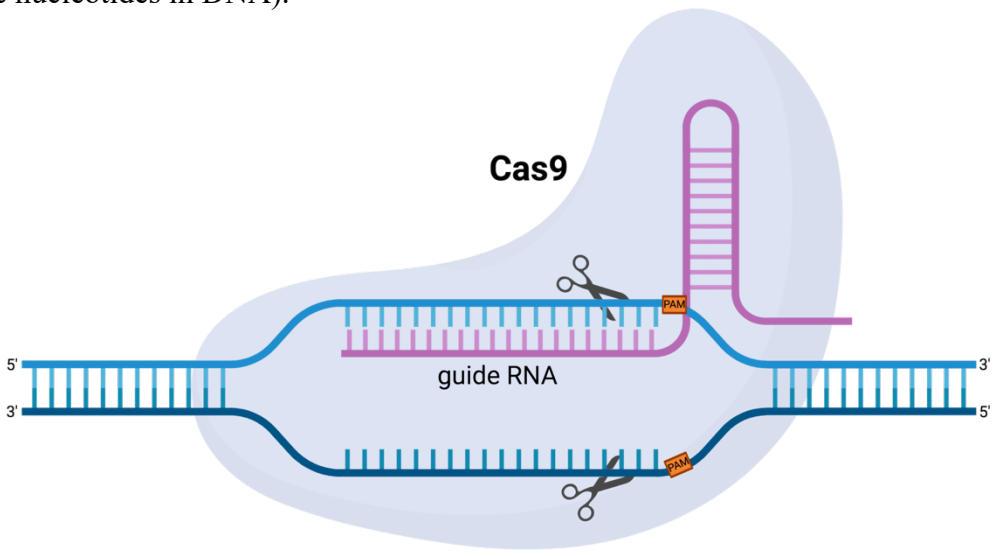

Figure 2. ***Components of the CRISPR/Cas9 system used in eukaryotes.*** Gray: Cas9 endonuclease. Pink: guide RNA. Blue: plant genomic DNA. (Created at Biorender.com)

There are many potential PAM sites within the *A. thaliana* genome. The researcher's task is to find "unique" 20 basepair sequences immediately adjacent to a PAM site as this sequence will be used to create the guide RNA that directs the Cas9 enzyme to its PAM-adjacent cleavage site. Web-based programs are available to identify guide sequences adjacent to PAM sites although finding truly unique sequences is usually a challenge, leaving open the possibility of off-target break sites. When double-stranded breaks are made, the predominant DNA repair mechanism is non-homologous end-joining (NHEJ). NHEJ usually causes highly localized short deletions and base changes (Cannan and Pederson, 2016). When Cas9 creates a double-stranded break, the resulting small deletions and/or base changes caused by NHEJ may not be enough to disrupt gene function. However, if two double-stranded breaks are made close to each other in a DNA sequence, several types of mutations may occur: loss of all or most nucleotides between the breaks, inversion and re-insertion of the DNA fragment between breaks, or larger insertion/deletion mutations that extend past the break sites. Therefore, two nearby double-stranded breaks are more likely to result in large deletions that cause loss of gene function (a null mutant).

Agrobacterium tumefaciens was used for plant transformation. In nature, *A. tumefaciens* is a bacterial pathogen that infects plants by inserting a DNA fragment containing genes necessary for pathogenicity (the T-DNA or transferred DNA) into the plant's genomic DNA. Genes on the T-DNA direct plant cells to produce cytokinin and auxin, which cause unregulated cell division and the development of tumor-like galls ("*Agrobacterium*", 2018). The T-DNA also directs plant cells to produce opines which serve as primary energy source for *A. tumefaciens* ("*Agrobacterium*", 2018). This process was been modified by plant researchers to remove the

cytokinin, auxin, and opine biosynthetic genes and replace them with a researcher's gene(s) of interest that will be transferred into a plant's genome.

In our lab, we decided to first attempt to make a deletion mutation in *PAT14* using the CRISPR/Ca9 system. As previously mentioned, other researchers identified T-DNA-induced mutations in *PAT14* that result in an early senescence phenotype (Li et al. 2016; Zhao et al. 2016). This phenotype is well-characterized and easily identifiable. Therefore, if a large deletion mutation in *PAT14* is created by the CRISPR/Cas9 system, we expect to observe the same early senescence phenotype identified by other researchers. If an early senescence phenotype is observed for our *pat14* CRIPSR mutants, we can be confident that the CRISPR/Cas9 system is working well in our experiments. Additionally, these experiments provide insight to the types of deletions to expect from CRISPR/Cas9 mutagenesis.

Results

Creation of CRISPR Deletion Construct (done by Ben Kaphan)

Two target (PAM) sites were chosen within the *PAT14* gene (about 682 bp apart) and coding strand and non-coding strand guide sequences were designed for each site. One guide should create a double-strand break near translation start codon in exon 1 and the other guide should create a double-strand break in exon 3. Double-stranded guide fragments were created and cloned into either pEn_Comaira.1or pEn_Comaira.2 using BsaI (Figure 3; Lynagh et al., 2018). Plasmid structure was confirmed by digestion with BsaI (no digestion if guide fragment inserted) and DpnI (guide sequences contain DpnI site). Insert orientation was confirmed by sequencing with the M13 primer in the plasmid backbone.

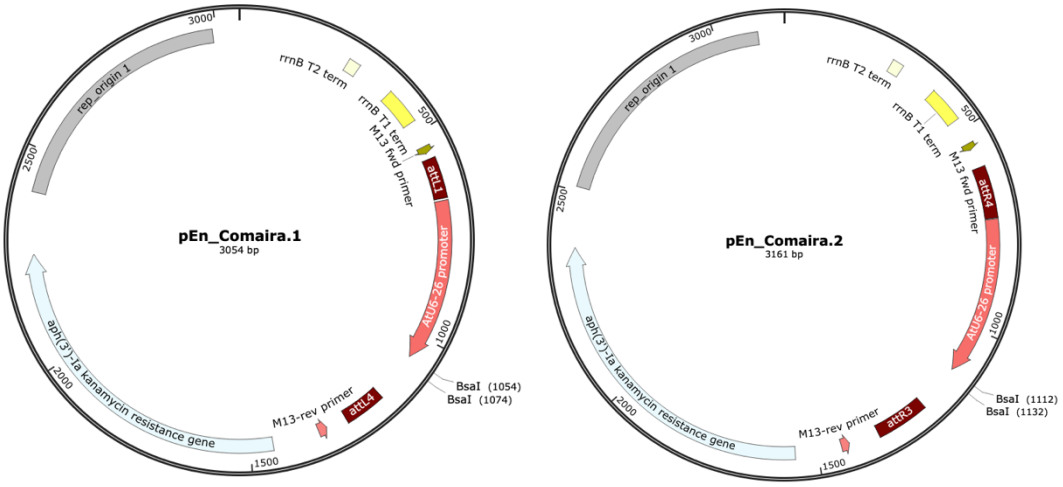

Figure 3. ***Structures of pEn_Comaira.1 and pEn_Comaira.2.*** BsaI was used to clone double-stranded guide RNA fragments into these plasmids. Note two unique *att* sites for Gateway cloning on each plasmid.

A multi-site Gateway recombination reaction was used to create a binary plasmid for plant transformation. The reaction contained both pEn_Comaira plasmids, pEn-RC9 (Figure 4) containing the Cas9 gene, and the binary vector pEarlyGate302-v2 (Figure 5). The final plant

binary vector (Figure 6) was confirmed by restriction digestion and electroporated into *A. tumefaciens* for plant transformation.

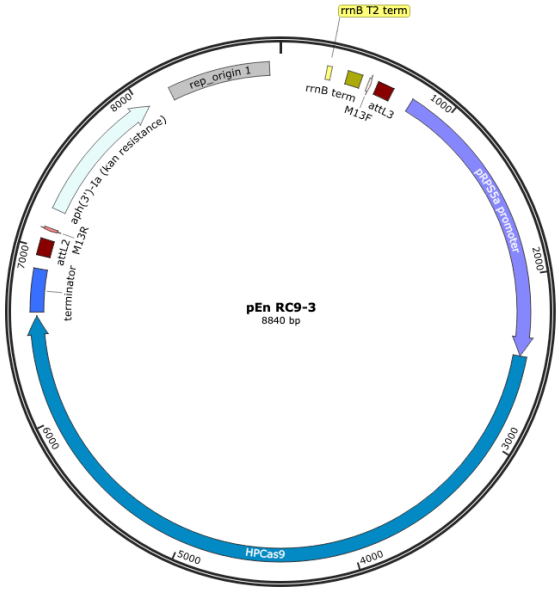

Figure 4. ***Plasmid with a plant ribosomal protein promoter driving expression of Cas9 endonuclease.*** The promoter is highly expressed in somatic and germ cells, including the zygote and early embryo (Lynagh et al., 2018; Oh and Kim, 2021).

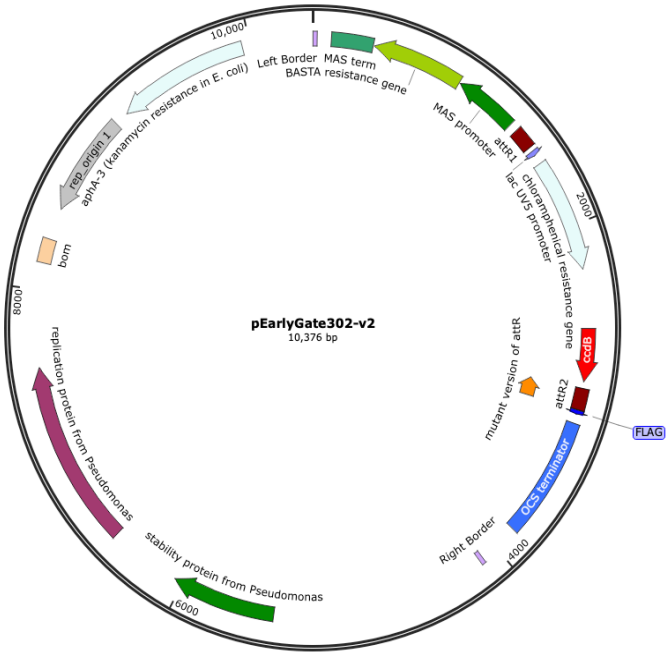

Figure 5. ***Binary plasmid prior to Gateway reaction.*** (provided by E. H. Tan, University of Maine). DNA between *attR1* and *attR2* is replaced by DNA from pEn_Comaira.1 and pEn-Comaira.2 constructs and from pEN_RC9-3.

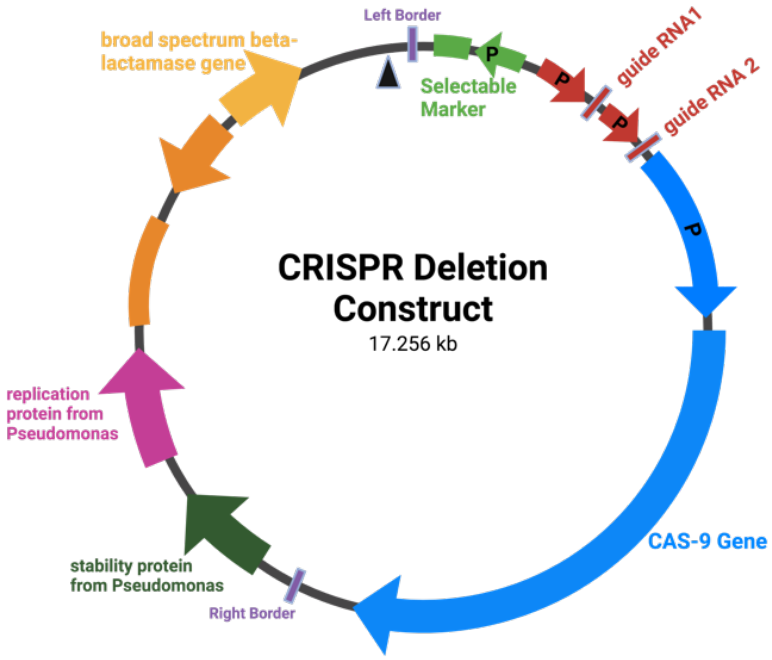

Figure 6. ***Binary plasmid used to create PAT gene deletions.*** The four genes that will be delivered into *A. thaliana* extend from the Left Border clockwise to the Right Border. selectable marker = BASTA-resistance gene (*bar*). Green, red, and blue arrows represent eukaryotic promoters for the associated genes. The regions on the left side of the plasmid function only in *E. coli* or *A. tumefaciens* and are not transformed into the plant. (Created at Biorender.com)

Arabidopsis thaliana Transformation with PAT14-CRISPR Construct (done by Ben Kaphan)

Flowers of wildtype *A. thaliana* were transformed by the floral dip method, and seeds collected from these T_0 plants were sown at high density and selected with BASTA herbicide for plants that received the T-DNA insertion. Four BASTA resistant T_1 plants were obtained; however, two of these BASTA resistant plants did not survive and seeds were only collected from plant 3-3 and plant 2-1 (Figure 7). Amplification of the region around the *PAT14* gene from wildtype *A. thaliana* genomic DNA was expected to make a product of 455. The PCR products produced by amplification of DNA from putative mutants would be smaller if a deletion had occurred. Both plants 3-3 and 2-1 generated products that appeared to be ~455 bp (data not shown) but were sequenced anyway to look for small deletions. Analysis of sequence generated from plant 2-1 revealed a deletion of 24 bp in the expected location on one chromosome.

Unfortunately, the other chromosome still had the wildtype allele of *PAT14*. Sequencing of plant 3-3 was inconclusive.

Identification of Deletions in PAT14 Gene

Multiple pots were planted with T_2 seeds collected from *pat14* CRISPR Plant 2-1 and T_2 seedlings were BASTA sprayed to select for plants containing the herbicide-resistance gene. Five BASTA resistant plants were obtained from this second selection (see row labeled *pat14* CRIPSR mut 2-1 Pot# Plant# T_2 in Figure 7) and T_3 seeds were collected. No PCR was performed on these plants.

For the first round of sequencing post-pandemic, T_3 seeds from three different lineages all originating from *pat14* CRISPR mut 2-1 T_2 were planted : Pot 5 Plant 1, Pot 5 Plant 2, and Pot 6 Plant 1. Crude genomic DNA was isolated from young T_3 plants and PCR amplified to identify any large deletions in *PAT14.* If a large deletion was created, we would expect the amplified DNA fragment to be smaller than the DNA fragment from wildtype *PAT14* genomic DNA. As mentioned previously, gel electrophoresis of PCR amplified DNA fragments did not identify any obvious large deletions in the *PAT14* gene (data not shown). To determine why no large deletions were made, all previous steps and reaction components were checked. Upon close observation, it was determined that one of the guide DNAs (gDNA 1662) designed by a previous lab member contained an extra nucleotide not present in the *PAT14* sequence. This error results in the guide RNA being unable to bind properly meaning that Cas9 system will be unable to make a double-strand break at that site. However, a single double-stranded break near the site of gDNA 960 could still result in small deletions that are disruptive if they result in out-of-frame

deletions or affect regulatory regions. PCR amplified crude genomic DNA from 18 of the T_3 plants was sequenced (Figure 7).

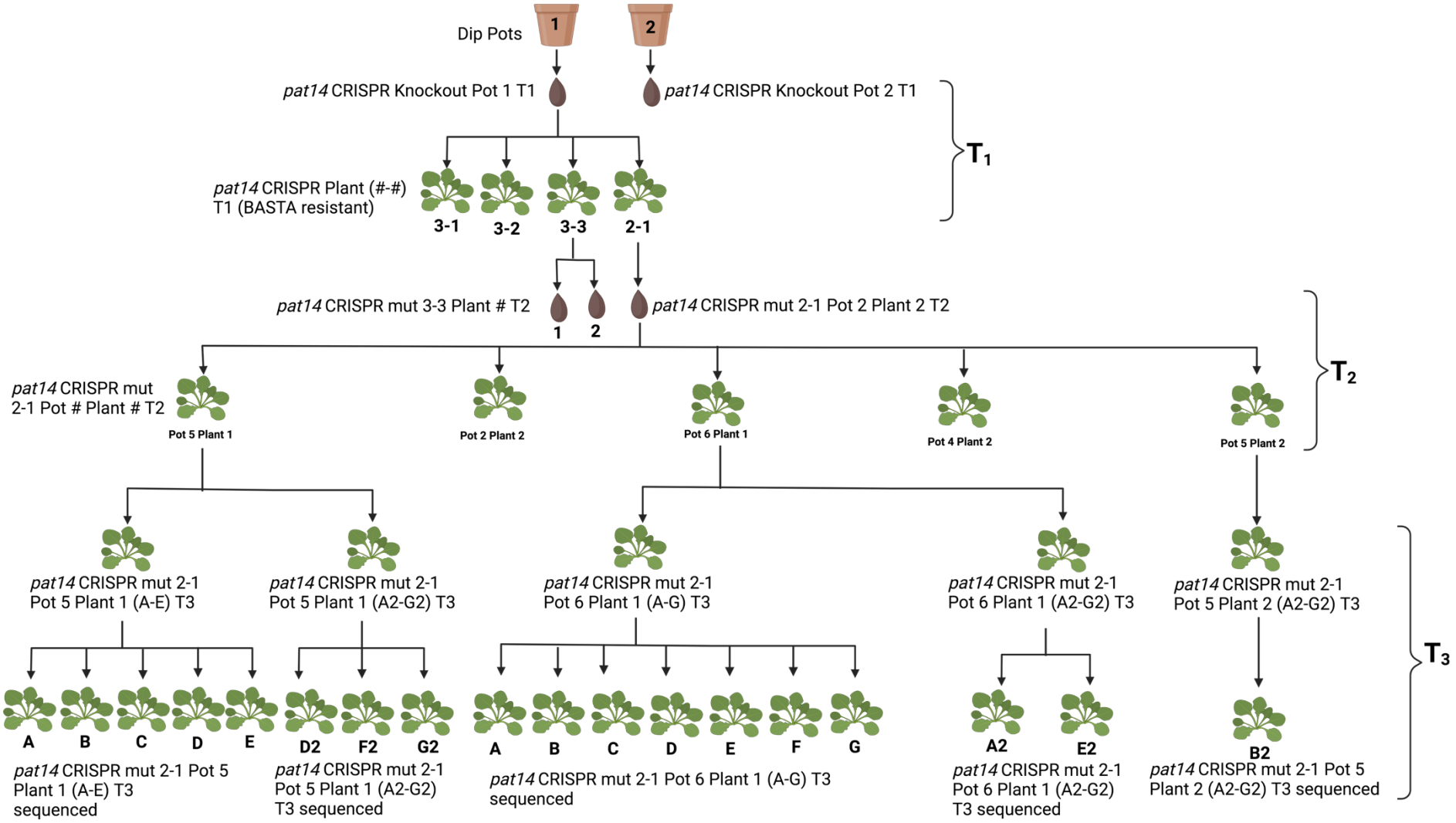

Figure 7. ***Lineage of pat14 CRISPR/cas9 induced mutant plants.*** (Created at Biorender.com)

Genotypic Analysis

Sequencing of mutant CRISPR-induced *pat14* plants revealed six unique mutations: five deletions and one insertion (Figure 8). Deletions ranged from 5 bp to 24 bp and the insertion was 2 bp. Mutation #1 was found in two plants; one plant was homozygous for the mutation and one plant was heterozygous. Mutation #2 was found in five plants; three plants were homozygous for the mutation and two plants were heterozygous. All plants containing mutations #1 and #2 are descendents of *pat14* CRISPR mut 2-1 Pot 5 Plant 1 T_2. Mutation #3 was found in two plants; both plants were heterozygous for the mutation (Figure 8). Mutations #4, #5 and #6 were each found in one plant that was heterozygous for the mutation. All plants containing mutations #3, #4, #5, and #6 are descendents of *pat14* CRISPR mut 2-1 Pot 6 Plant 1 T_2. All other T_3 plants in Figure 7 were determined to be homozygous for the wildtype *PAT14* allele based on sequencing.

Mutations #1, #2, and #3 resulted in in-frame deletions while mutations #4, #5, and #6 resulted in out-of-frame deletions. Of the eighteen T_3 plants that were sequenced, twelve plants contained mutations likely created by Cas9 as expected. However six plants were determined to be homozygous for the wildtype *PAT14* allele.

Figure 8. ***Nucleic acid sequences of PAT14 wildtype (WT) allele compared to sequences of pat14 deletion or insertion alleles likely caused by CRISPR/Cas9 editing***. **Mutation #1** was found in Plants 51D (WT/M) and 51E (M/M). **Mutation #2** was found in Plants 51A (M/M), 51B (WT/M), 51D2 (M/M), 51F2 (M/M), and 51G2 (WT/M). **Mutation #3** was found in Plants 61C (WT/M) and 61D (WT/M). **Mutation #4** was found in Plant 61G (WT/M). **Mutation #5** was found in Plant 61B (WT/M). **Mutation #6** was found in Plant 61A2 (WT/M). Missing nucleotides are indicated with a dash and inserted nucleotides are shown in red.

Nucleic acid sequences of *pat14* mutant alleles were translated into amino acid sequences and the potential PAT14 protein was modeled (Figure 9). To determine the protein products of mutant *pat14* alleles with deletions of the original translation start site (methionine), an algorithm called Translation Initiation Sites (Gleason et al., 2022) was used to predict the most likely downstream translation start site. In mutations #1 and #2, a downstream methionine at amino acid 25 was the most likely alternate translation start site and resulted in identical protein products for both mutations. Since small deletions in transmembrane domains may not affect protein threading (Hernandez et al., 2003; Renaud et al., 1991), we predicted that mutations #1 and #2 could result in properly-oriented and functional PAT14. In mutation #3, translation initiates at the original methionine, but the predicted protein has an internal deletion of amino acids 5-8. As expected, out-of-frame deletions (mutations #4-#6) resulted in severely truncated proteins containing only 24 to 42 amino acids.

PAT 14 WT:
MHRSGTTMAWNVFKFCTALRGLGSIMILLVLGVVGVTYYAVVLTNYGPALSQGGLDSLAALTILILFHFLLAMLLWS
YFSVVFTDPGVVPPNWRPSTDEERGESDPLNSLDFVGLQSDSSSSNPRVRFCRKCNQLKPSRCHHCSVCGRCVLKMD
HHCVWVVNCVGALNYKYFLLFLFYTFLETTLVTLVLMPHFIAFFSDEEIPGTPGTLATTFLAFVLNLAFALSVMGFL
LMHISLVAGNTTTIEAYEKKTTTKWRYDLGKKKNFEQVFGMDKRYWLIPGYTEEDLRRMPELQGLEYPSKPDFDSQ.

Mutation #1:
MILLVLGVVGVTYYAVVLTNYGPALSQGGLDSLAALTILILFHFLLAMLLWSYFSVVFTDPGVVPPNWRPSTDEERG
ESDPLNSLDFVGLQSDSSSSNPRVRFCRKCNQLKPSRCHHCSVCGRCVLKMDHHCVWVVNCVGALNYKYFLLFLFYT
FLETTLVTLVLMPHFIAFFSDEEIPGTPGTLATTFLAFVLNLAFALSVMGFLIMHISLVAGNTTTIEAYEKKTTTKW
RYDLGKKKNFEQVFGMDKRYWLIPGYTEEDLRRMPELQGLEYPSKPDFDSQ.

Mutation #2:
MILLVLGVVGVTYYAVVLTNYGPALSQGGLDSLAALTILILFHFLLAMLLWSYFSVVFTDPGVVPPNWRPSTDEERG
ESDPLNSLDFVGLQSDSSSSNPRVRFCRKCNQLKPSRCHHCSVCGRCVLKMDHHCVWVVNCVGALNYKYFLLFLFYT
FLETTLVTLVLMPHFIAFFSDEEIPGTPGTLATTFLAFVLNLAFALSVMGFLIMHISLVAGNTTTIEAYEKKTTTKW
RYDLGKKKNFEQVFGMDKRYWLIPGYTEEDLRRMPELQGLEYPSKPDFDSQ.

Mutation #3:
MHRSAWNVFKFCTALRGLGSIMILLVLGVVGVTYYAVVLTNYGPALSQGGLDSLAALTILILFHFLLAMLLWSYFSV
VFTDPGVVPPNWRPSTDEERGESDPLNSLDFVGLQSDSSSSNPRVRFCRKCNQLKPSRCHHCSVCGRCVLKMDHHCV
WVVNCVGALNYKYFLLFLFYTFLETTLVTLVLMPHFIAFFSDEEIPGTPGTLATTFLAFVLNLAFALSVMGFLIMHI
SLVAGNTTTIEAYEKKTTTKWRYDLGKKKNFEQVFGMDKRYWLIPGYTEEDLRRMPELQGLEYPSKPDFDSQ.

Mutation #4:
MHRSGTTMWRGTYSSSVLPYEVLDRS.

Mutation #5:
MHRSGTTMERIQVLYCLTRSWIDHDPFGSWCCRCYLLRCRFD.

Mutation #6:
MHRSGTTMRGTYSSVLPYEVLDRS.

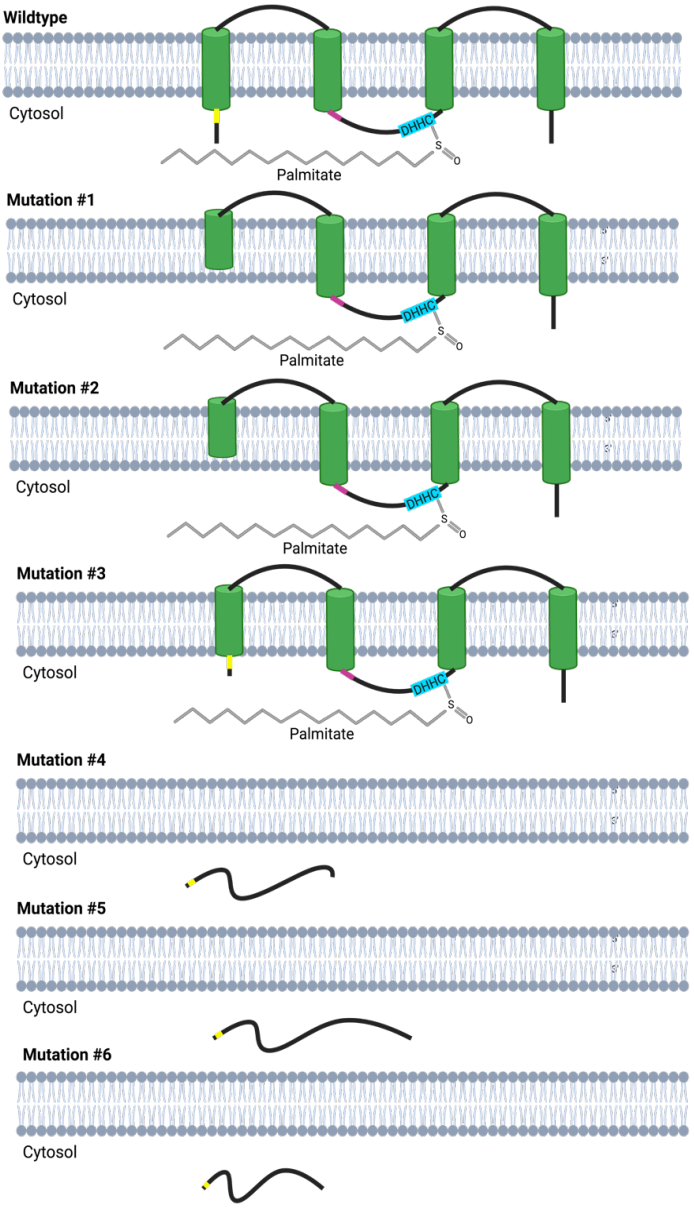

Figure 9. ***Amino acid sequences of PAT14 alleles and corresponding protein models.*** **Mutation #1** was found in Plants 51D (WT/M) and 51E (M/M). **Mutation #2** was found in Plants 51A (M/M), 51B (WT/M), 51D2 (M/M), 51F2 (M/M), and 51G2 (WT/M). **Mutation #3** was found in Plants 61C (WT/M) and 61D (WT/M). **Mutation #4** was found in Plant 61G (WT/M). **Mutation #5** was found in Plant 61B (WT/M). **Mutation #6** was found in Plant 61A2 (WT/M). Transmembrane domains are shown in green, the active site is shown in blue, the conserved BG region is shown in yellow, the conserved DPG region is shown in purple, and incorrect amino acids are shown in red. (Created at Biorender.com)

Phenotypic analysis

The phenotypes of BASTA resistant plants were monitored throughout the plant's life cycle and seeds were collected from all BASTA resistant plants for future generational analyses. Ideally, stable mutants would no longer contain the *Cas9* endonuclease gene as it could continue to create genome modifications. To determine if the *Cas9* gene had been eliminated through normal meiotic segregation, extracted DNA from T_3 plants (not BASTA selected) was PCR amplified (preliminary, data not shown), but the results were inconclusive.

No CRISPR induced *pat14* mutants had an obvious phenotype (Figure 10). Wildtype plants are dark green and exhibit robust growth (Figure 10 A). *pat14-1* mutants created by T-DNA insertion exhibit early leaf senescence, are pale green/yellow, and are small in size (Figure 10 B; Li et al. 2016; Zhao et al. 2016). *pat14* mutants created by CRISPR/Cas9 system were indistinguishable from the wildtype *PAT14* plants (Figure 10 C).

A **B** **C**

Figure 10. ***Phenotypic comparison of wildtype A. thaliana to pat14 mutants caused by T-DNA insertion (showing early senescence phenotype) or by CRISPR deletion.*** **A)** wildtype *Arabidopsis thaliana* **B)** *pat 14-1* mutant created by T-DNA insertion (Li et al. 2016; Zhao et al. 2016) **C)** a representative *pat14* mutant created by CRISPR/Cas9 system

Discussion/ Conclusions

PAT14

Previously, sequencing of *pat14* CRISPR plant 2-1 T_1 identified a 24 bp deletion starting 1 bp upstream of the start codon (Figure 7). This T_1 plant was heterozygous. No sequencing was done for any T_2 plants so it is unknown if there are any remaining wildtype alleles or if any of these plants are transheterozygotes or homozygous mutants. Post-pandemic sequencing of T_3 plants, descended from *pat14* CRISPR plant 2-1 T_1, identified the same 24 bp mutation (mutation #1; Figure 8). However, we also identified multiple new mutations (mutations #2-#6) in T_3 plants descended from *pat14* CRISPR plant 2-1 T_1.

If Cas9 is efficient in creating double-stranded breaks at sites specified by a guide sequence, we would not expect any wildtype alleles to be present in these plants. However, we identified six T_3 plants that were BASTA resistant yet homozygous for the *PAT14* allele. Since Cas9 appears to not be efficient at making double-stranded breaks at the chosen guide sites, it may take more generations to obtain stable homozygous mutants.

Although the genotypes of the T_2 plants were not determined, we expected that if the T_2 plants were heterozygous for a deletion allele, multiple homozygous T_3 plants for that deletion would be detected in subsequent generations because *A. thaliana* self-pollinates. All descendants of *pat14* T_2 CRISPR mut 2-1 Pot 5 Plant 1 contain either mutation #1 or mutation #2. Additionally, some of the descendants of this T_2 plant contain a wiltype *PAT14* allele. Mutations #1 and #2 are almost identical except mutation #2 has an insertion of 2 basepairs which could have resulted from DNA repair mechanisms or a de novo mutation. This suggests that *pat14*

CRISPR mut 2-1 Pot 5 Plant 1 T_2 was heterozygous for mutation #1 and that mutation #2 likely arose in the meiocytes or elsewhere in the germline.

Four other *pat14* mutant alleles (mutations #3-#6) were identified in descendants of *pat14* CRISPR mut 2-1 Pot 6 Plant 1 T_2 although again, other T_3 descendants were homozygous for the wildtype *PAT14* allele. The simplest explanation as to why four new and unique mutant alleles were present in T_3 descendants is that *pat14* CRISPR mut 2-1 Pot 6 Plant 1 T_2 was homozygous for *PAT14*. The new mutations likely were created by Cas9 de novo somewhere in the germline.

As expected, all six of the mutations were near the start codon because the one functional guide RNA designed for *PAT14* targeted that area. Some of the mutations affect regions of 24 bp or less and are homozygous and in-frame (mutations #1-#3) . Mutations #1 and #2 may utilize a downstream in-frame methionine (normally the 25th amino acid of PAT14) embedded in a Kozak sequence as an alternate translation start codon, resulting in a loss of the short cytosolic amino-terminus and a few amino acids of the first transmembrane domain. As previously mentioned, it is likely that small deletions at the start of the 1st transmembrane domain don't affect threading of the protein through hydrophobic membranes. Based on the protein modeling analysis, it is possible that mutations #1-#3 have minimal effect on function of the PAT14 protein. Mutations #4-#6 appear to have resulted in out-of-frame-deletions leading to frameshifts which would most likely result in a nonfunctional *PAT14* gene product. To date, all identified out-of-frame deletions or insertions are heterozygous.

Despite identification of multiple deletion mutations, no mutant phenotype has been observed for any of the CRISPR-induced *pat14 A. thaliana* plants. There are several reasons for

the lack of a mutant phenotype. First, the in-frame deletions are not large enough to destroy PAT14 function. Second, although out-of-frame *pat14* mutant alleles have been identified, all of them are heterozygous, indicating that the one wildtype allele that remains could produce enough functional PAT14 to maintain a wildtype phenotype. Several other labs have observed an early senescence phenotype for *pat14* mutants made using T-DNA insertions (Li et al. 2016; Zhao et al. 2016), and this phenotype could be complemented by transformation with a wildtype *PAT14* allele.

In the future, more genetic screening of mutant *pat14 A. thaliana* plants and phenotypic observation will need to be performed to determine if mutations in *PAT14* lead to a mutant phenotype.

Methods

General Molecular Biology Techniques

i. Polymerase Chain Reaction

All PCR components were combined in thin-walled Eppendorf tubes at final working concentrations of 0.2 mM dNTPs, 1X Ex-Taq Buffer, and 0.2 uM of each primer. Other reaction components include 0.3 μL ExTaq Polymerase (TaKaRa Hot Start Cat. RR006A) and 1 μL DNA template. Primers used in this research are in Table 1. All primers were designed using PrimerSelect (DNAStar, Madison, WI) and ordered from IDT (Coralville, IA). Ideally, primers are approximately 20 bp in length with minimal internal complementarity, have similar annealing temperatures, do not have any off target binding sites, and are not be able to form primer dimers at the chosen annealing temperature. Annealing temperatures for primer pairs were determined using manufacturer's recommendations and by performing PCR primer tests on *A. thaliana* genomic DNA at a range of temperatures. Reaction tubes were placed into a thermocycler (PTC-100, MJ Research Inc. Watertown, MA, SN16368) for PCR amplification (Figure 7).

Table 1. Primers used in this work. Red text indicates bases not found in *A. thaliana* genomic sequence

Primer Name	Sequence	Annealing temperature	Purpose
M13F	5'TGTAAAACGACGGCCAGT	55°C	Verification of entry clones; used with a guide primer
gDNA 960	5'ATTGGATCTGGTACAACAATGGCG 5'AAACCGCCATTGTTGTACCAGATC	57.7°C	Upstream *pat14* target site in *A. thaliana* genome
gDNA 1662	5'ATTGGATGGATCACCATTGGTGTT 5'AAACAAACACAATGGTGATCCATC	57.0°C	Downstream *pat14* target site in *A. thaliana* genome
gDNA 327	5'ATTGGTACCTAACCTCCTTCTTGA 5'AAACATCAAGAAGGAGGTTAGGTA	55.4°C	Upstream *pat3* target site in *A. thaliana* genome
gDNA 1281	5'ATTGGCAACTAGATCTCAGAGCAA 5'AAACTTGCTCTGAGATCTAGTTGC	56.2°C	Downstream *pat3* target site in *A. thaliana* genome
DHC8-5'a	5'CCTCAAGCTCCGCATCATTCCTCCTTC	54.0°C	Upstream of CRISPR deletion site in *PAT14* gene
8-3R	5'CGAGGAAATGGAAGAGGATTAAGAT	54.0°C	Downstream of CRISPR deletion site in *PAT14* gene
Cas9 FWD	5'AAATATCAACCCTGGCAACAAAAC	52.0°C	Amplification of *Cas9*
Cas9 REV	5'GCTAGAGGAAGGGGAATAACAGAT	52.0°C	Amplification of *Cas9*

Segment 1 Initial Denaturation

1. 94°C for 2 minutes

Segment 2 Temperature Cycling

2. Denaturation 94°C for 30 seconds
3. Annealing 50-65°C for 30 seconds
4. Elongation 72°C for 1-5 minutes
5. Repeat steps 2-4 for a total of 35 cycles

Segment 3 Final Elongation

6. 72°C for 5 minutes

Segment 4 Refrigeration

7. 4°C for ever

Figure 7. ***General Polymerase Chain Reaction thermocycler steps***

ii. Gel Electrophoresis

A 0.4 cm thick 1% (w/v) agarose gel was made using quick dissolve agarose (Apex BioReasearch Products Cat.20-102QD) and 1X TAE (Tris-acetate-EDTA buffer). All gels were run at 7 Volts/cm using an EC135 power supply (E-C Apparatus Corporation, St. Petersburg ,Florida, SN:BK100974) until the tracking dye was about halfway across the gel. DNA was visualized with ethidium bromide under ultraviolet light.

iii. Plasmid Isolation

Plasmids were isolated using the Monarch Plasmid Miniprep kit (New England Biolabs, Ipswich, MA, #T1010L) following manufacturer's protocol with the exception that DNA samples were eluted with distilled water instead of DNA elution buffer. pEG_302v2,

pEN_Comaira.1, pEN_Comaira.2, and pEn_RC9 were transformed into electrocompetent *E. coli* TOP10 and transformants were selected on LB medium containing 50 mg/L kanamycin. pEG_302v2 was isolated from *E. coli* and diluted to 50 ng/μL. pEN_Comaira.1, pEN_Comaira.2, and pEn_RC9 were isolated from *E. coli* and diluted to 20 ng/μL.

iv. DNA Purification

PCR products were purified using the Monarch PCR DNA Cleanup Kit (New England Biolabs, Ipswich, MA, #T1030S) following manufacturer's protocol with the exception that DNA samples were eluted with distilled water instead of DNA elution buffer.

v. DNA Quantification

All DNA quantification was done using a fluorometer (Quibit, Invitrogen Cat.Q32857) following manufacturer's instructions.

vi. Bacterial Transformation

All bacterial transformations were done by electroporation (Eppendorf Electroporator 2510, SN:EC0433) following manufacturer's recommendations.

vii. *Bacterial growth medium*

All bacteria were grown in LB medium (1% tryptone, 0.5% yeast extract, 171 mM NaCl) with addition of 1.3% agar when needed. *Escherichia coli* were grown at 37°C and *Agrobacterium tumefaciens* were grown at 28°C.

Creation of Binary Plasmids for Plant Transformation

i. gRNA design, synthesis, and preparation

Two CRISPR target sites (PAM) were identified in the *pat* gene of interest; these sites are intended to mark the beginning and end of the desired deletion. The sites were chosen using the programs CRISPR-P (http://crispr.hzau.edu.cn/CRISPR2/) (Liu et. al, 2017) and Benchling [Biology Software] (2021) (https://benchling.com). Final CRISPR guide RNA sequences were chosen based on parameters including high on target efficiency, low predicted off-target binding, target location withing *pat* gene (avoided sites close stop codons and sites containing a high percentage of nucleotide polymorphisms), and total length around 20 bp (Mormile, 2020). When ordering primers, 5'ATTG was added to the coding strand primer and 5'AAAC was added to the template strand primer to enable ligation to the Comaira plasmids. Complementary single-stranded DNA oligonucleotides for the chosen target sequence (Table 1) were ordered from IDT (Coralville, IA). The complementary oligos were rehydrated to a concentration of 25 μM, equal volumes were incubated at 95°C for 5 minutes to create double-stranded DNA with four nucleotide 5' overhangs, and annealed by slowly cooling to room temperature for 20 minutes to create double-stranded guide DNAs (gDNA) with four nucleotide 5' overhangs.

ii. Cloning guide sequences into into pEn_Comaria.1 and pEn_Comaria.2

A. Digestion of Entry Vectors

500 ng of pEn_Comaira.1 or pEn_Comaira.2 were digested separately digested with BsaI-HF (NEB Cat. No. R3535S) overnight following manufacturer's recommendations with the exception of digestion duration. The enzyme was inactivated at 65°C for 20 minutes. Plasmid samples were adjusted to 25 ng/μL.

B. Ligation of gRNA to Entry Clones

For each CRISPR gDNA, 1 μL of a digested pEn_Comaira, 1 μL of an annealed gDNA, 1 μL T4 DNA ligase (New England Biolabs Cat.M0202S), and 5 μL of 2x T4 DNA ligase buffer were combined in a 10 μL reaction and incubated overnight at room temperature. Part of the ligation mixture was transformed into high-competency *E. coli* TOP10 and selection was performed on LB plates containing kanamycin at 50 mg/L. PCR using the M13F primer and the appropriate gRNA oligo primers (Table 1) followed by gel electrophoresis was used to confirm the presence of the gDNA in the pEn_Comaira plasmids.

iii. Multisite Gateway cloning reaction

Plasmids (pEG_302v2 , pEn_Comaira.1 containing gDNA 1, pEn_Comaira.2 containing gDNA 2, and pEn_RC9) were combined into the final "deletion construct" plasmid using a Gateway reaction. 1 μL of pEG_302v2 (50 ng/ μL), 1 μL of pEn_Comaira.1 containing gDNA 1 (20 ng/ μL), 1 μL of pEn_Comaira.2 (20 ng/ μL) containing gDNA 2, 1 μL of pEn_RC9 (20 ng/ μL), and 1 μL of 5X LR Clonase II Plus enzyme mix (Invitrogen Corp., Carlsbad, CA Ref 11791-020) were added to a 1.5 mL microfuge tube and incubated overnight at room temperature. 1 μL of Proteinase K (Invitrogen P/N 59895) was added and the reaction was incubated at 37°C for 10 minutes to digest the the LR Clonase II. The cloning mixture was electroporated into TOP 10 high-competency *E. coli* and selection was performed on LB plates containing carbenicillin at 100 μg/mL.

iv. verification of assembly

The deletion construct plasmid was isolated from high-competency *E.coli* TOP10 and verified via restriction enzyme digestion. To perform the restriction enzyme digest, enzymes

were chosen to create product sizes between 500-3,000 bp and have good size differentiation. Gel electrophoresis was performed on the products to verify that they are at the correct product size.

Transformation of Agrobacterium tumefaciens

A. tumefaciens (strain GV3101) was transformed with the deletion construct and selection was performed on LB plates containing 50 mg/L kanamycin.

Growth of Arabidopsis thaliana

A. thaliana seeds were planted in moist potting mix (equal parts peat-based potting mix and pearlite) and grown at 21°C with a 16 hour photoperiod and photosynthetic photon flux of 100 umol sec^{-1}. Plants were fertilized with MiracleGro fertilizer at each watering.

Transformation of Arabidopsis thaliana

A. tumefaciens were growth in LB broth containing 50 mg/L carbenicillin overnight. Cells were collected by centrifugation (3000 x g for 5 minutes) and resuspended in 5% sucrose and 0.05% silwet at an optical density between 0.600 and 1.00 at 600 nm. Flowers of intact wildtype *A. thaliana* plants (Col-0 ecotype) at approximately 6 weeks old were dipped into the bacterial suspension (Clough & Bent, 1998). These T_0 plants, referred to as "dipped plants", were incubated overnight in low light and high humidity before plants were rinsed with water to remove excess sucrose. T_1 seeds were harvested when siliques were mature.

Selection of transformed Arabidopsis thaliana

T_1 seeds were planted at a high density in pots filled with moist potting mix. After approximately 2 weeks, plants were sprayed with 0.05% BASTA (glufosinate) to select for any young plants that appear to contain the T-DNA region from the *PAT* deletion construct. BASTA resistant plants were then transplanted to one per pot and grown to full maturity. Seeds were harvested when siliques were mature.

Isolation of crude plant genomic DNA

About 2 weeks after selection, small leaf samples were collected from BASTA resistant plants and placed into separate 1.5 mL locking tubes. 40 μL of NaOH (0.25M) was added to each tube and the leaf sample was mechanically damaged with a pipette tip. Samples were boiled for 30 seconds. 40 μL of HCl (0.25M) was immediately added, followed by 20 μL of 0.5M Tris-HCl_2 pH8. Samples were boiled for 2 minutes and then stored frozen until use as DNA template in PCR (Klimyuk et al., 1993).

DNA Sequencing and Analysis

All sequencing was done through Azenta Life Sciences (South Plainfield, NJ).

Arabidopsis thaliana lineage tracking and naming

To keep track of *pat* mutants, each original BASTA resistant plant was given a unique name; for example, "*pat3* mutant #X). To denote the generation, the wildtype *A. thaliana* plant transfected with *A. tumefaciens* "dipped plants" were denoted as T_0 plants and seeds produced were denoted as T_1. T_1 seeds yield T_1 plants. Any seeds collected from T_1 plants are denoted as T_2.

Acknowledgements

pEG_302v2, pEn_Comaira.1, pEn_Comaira.2, and pEn_RC9 were the gift of Ek Han Tran at the University of Maine. Ben Kaphan designed the *PAT14* guide RNAs and created the T_1 generation of BASTA resistant plants for *PAT14*. The Hammel Center for Undergraduate research provided funding support.

Appendix

PAT3

After transformation of wildtype *A. thaliana* plants with *A. tumefaciens* containing a *PAT3* CRISPR deletion construct that I made, no BASTA resistant plants were obtained despite multiple independent transformation attempts and re-electroporation of the deletion construct into *A. tumefaciens*. We hypothesized that plants might require two functional copies of the *pat3* gene and so any deletion that creates a nonfunctional PAT3 protein would be detrimental. Because *PAT3* is expressed in pollen, perhaps it is essential for pollen viability or for seed development. We have been able to obtain BASTA resistant plants for other *pat* mutants so a lack of successful transformation using *A. tumefaciens* is unlikely.

In an attempt to make a smaller deletion in the *PAT3* gene region, the *PAT3* deletion construct was remade containing only the CRISPR guide RNA called "g327". If the deletion construct only contains one of the CRISPR guides, then the CRISPR/Cas9 complex should only cut at one site in *PAT3* resulting in a much smaller deletion. After transformation of wildtype *A. thaliana* plants with *A. tumefaciens* containing this new deletion construct, no BASTA resistant plants were obtained.

Since creating a double-strand break at only one site in *PAT3* was also unsuccessful in generating viable BASTA-resistant seedlings, it is possible that the guide RNA (g327) is binding to one or more additional other regions of *A. thaliana's* genomic DNA causing lethal effects. This theory seems unlikely as the guide RNAs are small (~20 bp) and no other lethal effects of CRISPR/Cas9 guide RNAs has been reported by other researchers using this method. Future research will need to be done to determine if mutations in *PAT3* are lethal, what types of mutations are lethal, and what specific subcellular pathways *PAT3* is involved in.

PAT1

Prior to the pandemic, **o**nly 1 BASTA resistant plant was obtained in the T_1 generation. No sequencing of any putative mutants has been performed (Figure 11). Crude genomic DNA for T_2 and T_3 generation plants is available for future analysis. During growth of T_2 plants A2-J2, we observed that 2 of the 10 plants died suddenly between 4-6 weeks of age (plants G and F). It would be interesting to know if these plants contain any mutations in *PAT1*.

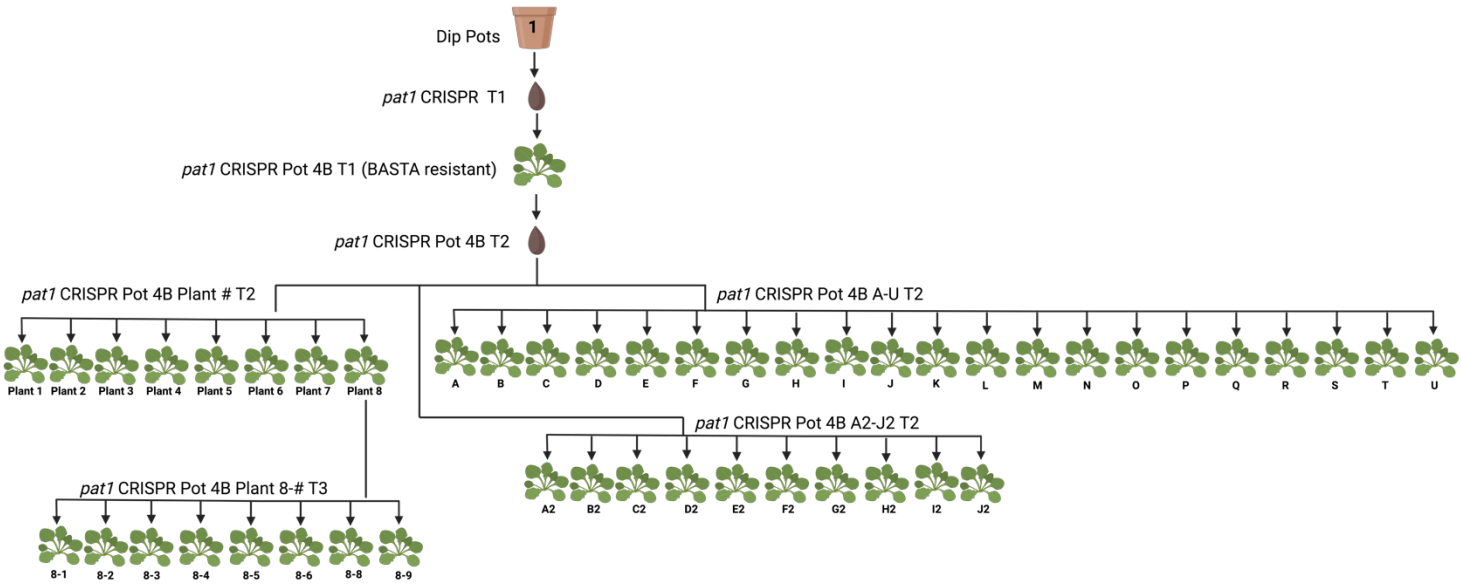

Figure 11. ***Lineage of pat1 CRISPR/Cas9 induced mutant plants.*** (Created at Biorender.com)

References

Agrobacterium. (2018, May 13). microbewiki.kenyon.edu. Retrieved January 11, 2023, from https://microbewiki.kenyon.edu/index.php/ Agrobacterium_tumefaciens

Arroyo-Olarte, R.D., Rodríguez, R.B., Morales-Ríos, E. (2021). Genome Editing in Bacteria: CRISPR-Cas and Beyond. Microorganisms. 9(4):844.

Batistic, O. (2012). Genomics and localization of the Arabidopsis DHHC-cysteine-rich domain S- acyltransferase protein family. Plant Physiol. 160: 1597-1612.

Beltrao, P., Bork, P., Krogan, N.J., van Noort, V. (2013). Evolution and functional cross-talk of protein post-translational modifications. Mol Syst Biol., 9(714).

Cannan, W.J., Pederson, D.S. (2016). Mechanisms and consequences of double-strand DNA break formation in chromatin. J Cell Physiol. 231:3-14.

Clough, S.J., Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. 16:735-743.

Gleason, A.C., Ghadge, G.,Chen, J., Sonobe, Y. Roos, R.P. (2022). Machine learning predicts translation initiation sites in neurologic diseases with nucleotide repeat expansions. PloS One. 17(6):e0256411.

Guan, X., Fierke, C.A. (2011) Understanding protein palmitoylation: biological significance and enzymology. Sci China Chem. 54:1888-1897.

Hemsley, P.A., Kemp, A.C., Grierson, C.S. (2005). The TIP GROWTH DEFECTIVE1 S-acyl transferase regulates plant cell growth in Arabidopsis. Plant Cell 17: 2554-2563

Hernandez, R., Sinodis, C., Horton, M., Ferreira, D., Yang, C., Brown, D.T. (2003). Deletions in the transmembrane domain of a sindbis virus glycoprotein alter virus infectivity, stability, and host range. J Virol. 77(23):12710-9.

Jiang, H., Zhang, X., Chen, X., Aramsangtienchai, P., Tong, Z. Lin, H. (2018). Protein lipidation: occurrence, mechanisms, biological functions, and enabling technologies. Chem Rev. 118: 919–988.

Klimyuk,V.I., Carroll, B.J., Thomas, C.M., Jones, J.D. (1993). Alkali treatment for rapid preparation of plant material for reliable PCR analysis. Plant J. 3(3):493-4.

Lai, J., Yu, B., Cao, Z., Chen, Y., Wu, Q., Huang, J., Yang, C. (2015). Two homologous protein S-acyltransferases, PAT13 and PAT14, cooperatively regulate leaf senescence in Arabidopsis. J Exp Bot. 66: 6345-6353.

Lynagh, P.G., Inagaki, S., Amundson, K.R., Marimuthu, M.P.A., Pike, B. R., Henry, I.M., Tan, E.H., Comai, L. (2018). Precise translocation and duplication using CRISPR/Cas9 in *Arabidopsis thaliana*. BioRXiv.

Li, Y., Scott, R., Doughty, J., Grant, M., Qi, B. (2016). Protein S-acyltransferase 14: a specific role for palmitoylation in leaf senescence in Arabidopsis. Plant Physiol. 170: 415-428.

Liu H, Ding Y, Zhou Y, Jin W, Xie K, Chen LL. (2017). CRISPR-P 2.0: An Improved CRISPR-Cas9 Tool for Genome Editing in Plants. Mol Plant. 10(3):530-532.

Mormile, B.W. (2020). Fundamentals of genome editing using CRISPR/Cas9. JRBiotek virtual 'Reach and Teach' Scientific Lecture Series Online. Retrieved from: https://www.youtube.com/watch?v=6_6lGcy258o

Oh, Y., Kim, S. (2021). RPS5A Promoter-Driven Cas9 Produces Heritable Virus-Induced Genome Editing in Nicotiana attenuate. Mol Cells 44:911-919

Qi, B., Doughty, J., Hooley, R. (2013). A Golgi and tonoplast localized S-acyl transferase is involved in cell expansion, cell division, vascular patterning and fertility in Arabidopsis. New Phytol. 200: 444-456.

Ramazi, S., Zahiri, J. (2021). Posttranslational modifications in proteins: resources, tools and prediction methods. Database (Oxford). doi: 10.1093/database/baab012

Renaud, K.J., Inman, E.M., Fambrough, D.M. (1991). Cytoplasmic and transmembrane domain deletions of Na,K-ATPase beta-subunit. Effects on subunit assembly and intracellular transport. J Biol Chem. 266(30):20491-7.

Ryan, E., Grierson, C.S., Cavell, A., Steer, M., Dolan, L. (1998). TIP1 is required for both tip growth and non-tip growth in Arabidopsis. New Phytol. 138: 49-58.

Schiefelbein, J., Galway, M., Masucci, J., Ford, S. (1993). Pollen tube and root-hair tip growth is disrupted in a mutant of Arabidopsis thaliana. Plant Physiol.103: 979-985.

Ulker, B., Peiter, E., Dixon, D.P., Moffat, C., Capper, R., Bouché, N., Edwards, R., Sanders, D., Knight, H., Knight, M.R. (2008). Getting the most out of publicly available T-DNA insertion lines. Plant J. 56(4):665-77.

Wan, Z.Y., Chai, S., Ge, F.R., Feng, Q.N., Zhang, Y., Li, S. (2017). Arabidopsis PROTEIN S-ACYL TRANSFERASE4 mediates root hair growth. Plant J. 90: 249-260.

Yuan, X., Zhang, S., Sun, M., Liu, S., Zi, B., Li, X. (2013). Putative DHHC-Cysteine-rich domain S-acyltransferase in plants. PloS One 8: e75985.

Zhao, X.Y., Wang, J.G., Song, S.J., Wang, Q., Kang, H., Zhang, Y., and Li, S. (2016). Precocious leaf senescence by functional loss of PROTEIN S-ACYL TRANSFERASE14 involves the NPR1-dependent salicylic acid signaling. Sci Reports 6: 20309.

Zhou, L.Z., Li, S., Feng, Q.N., Zhang, Y.L., Zhao, X., Zeng ,Y.L., Wang, H., Jiang, L., Zhang, Y. (2013). Protein S-Acyl Transferase10 is critical for development and salt tolerance in Arabidopsis. Plant Cell 25: 1093-1107.