Design and Preparation of CRISPR Constructs Targeting *Arabidopsis thaliana Receptor Like Kinase* (*RLK*) and *S-locus Cysteine Rich Like* (*SCRL*) Genes.

Kilner, C.,[1] Long, S.,[1] Biondo, A.,[1] Conway, C.,[1] DiLeo, A.,[1] Frederick, N.,[1] McAndrew, N.,[1] McGraw, D.,[1] Musto, C.,[1] Ng, G.,[1] O'Neill, G.,[1] Sacks, C.,[1] Sarno, E.,[1] Sulzinski, M.,[1] Villa, A.,[1] Westawski, J.,[1] Nasrallah, J.,[2] Nasrallah, M.,[2] Dwyer, K.[1,3] The University of Scranton, 2016.

[1] Fall 2015 Molecular Biology II Laboratory Students, The University of Scranton, Scranton, Pennsylvania, 18510
[2] Section of Plant Biology, School of Integrative Plant Science, Cornell University, Ithaca, New York, 14850
[3] Biology Department, College of Arts and Sciences, The University of Scranton, Scranton, Pennsylvania, 18510

Abstract: In *Arabidopsis thaliana* plants, the *Receptor Like Kinase* (*RLK*) and *S-locus Cysteine Rich Like* (*SCRL*) genes encode proteins proposed to initiate various signal transduction pathways regulating important plant processes such as development, reproduction and defense. To ascertain if this is so and to characterize the function of targeted *RLK* and *SCRL* proteins, CRISPR constructs were generated to "knockout" the corresponding *RLK* and *SCRL* genes.

Introduction: The *Arabidopsis* genus belongs to the Brassicacaceae family, all angiosperms (flowering plants), which includes agriculturally important plants such as kale, broccoli, Brussel sprout, and cauliflower.

Arabidopsis lyrata and *Arabidopsis haleri* plants demonstrate self-incompatibility (SI), a trait that prevents genetic inbreeding and reduces homozygosity. Unlike these two plants, *Arabidopsis thaliana* is self-fertile due to the degeneration of the *SRK* (*S-locus Receptor Kinases*) and/or *SCR* (*S-locus Cysteine Rich*) genes normally encoding the necessary SI self-recognition receptor and ligand proteins, respectively. In *A. thaliana* the *SRK* and *SCR* pseudogenes belong to gene families encompassing 40 *SD-RLK* (*S-Domain Receptor Like Kinases*) and 28 *SCRL* (*SCR-like*) homologous genes. The functions of these *SD-RLK* and *SCRL* genes,while not yet delineated, are hypothesized to play an important role in the cell signaling involved in regulating plant development, reproduction and/or plant defense.

We initiated the use of CRISPR/Cas9 genetic engineering technology for the functional analysis of selected *SD-RLK* and *SCRL* genes. The CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) system, occurring naturally in some prokaryotes, involves use of a RNA-guided endonuclease (Horvath and Barrangou, 2010). In prokaryotes, it specifically binds invading viral DNA via complementary basepairing of the guide RNA, thereby anchoring the associated Cas9 endonuclease protein to make a double-stranded cut in the target DNA (Figure 1). This incapacitates the virus. Thus the system is nicknamed "the immune system of bacteria."

Researchers have adapted the CRISPR system to specifically "knockout" any eukaryotic gene facilitating functional analysis for that gene (Jinek et al., 2012; Belhaj et al., 2013). This is done by characterizing the phenotypic changes that occur in that organism with the loss of that gene, suggesting what processes require a functional version of that gene. Thus, we will use CRISPR technology to specifically incapacitate certain *SD-RLK* or *SCRL* genes and then characterize the phenotypic effect this has on the resulting transgenic plants.

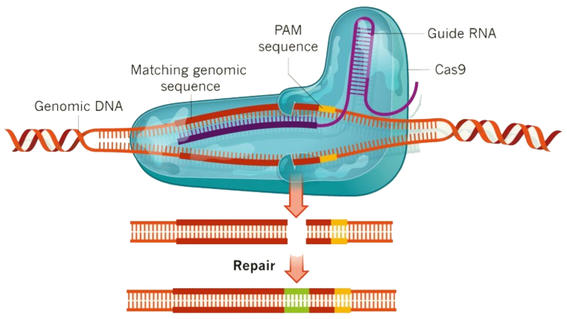


Figure 1. Pictured above is a Cas9 CRISPR RNA-guided endonuclease protein capable of binding and digesting the chosen target DNA. Subsequent repair of the double-stranded break in the genomic DNA is imprecise, resulting in small insertions or deletions of nucleotides (represented by green DNA). Consequentially, with a shift in the reading frame, a null mutation of the targeted gene is generated. (Image: Yang, 2014).

References

Hovrath, P., Barrangou, R. (2010) CRISPR/Cas, the immune system of Bacteria and Archaea. *Science* 327:5692, pp. 167-170.
Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J., Charpetnier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816.
Sander, J., Joung, J. K. (2014) Naturally occurring and engineered CRISPR-Cas systems. *Nature Biotechnology* 32, 347-335. doi:10.1038/nbt.2842
Yang Lei, Li Lu, Hai-Yang Liu, Sen L, Feng Xing, Ling-Ling Chen*. CRISPR-P: A web tool for synthetic single-guide RNA design of CRISPR-system in plants. *Mol Plant*, 2014, 7(9): 1494-1496. doi: 10.1093/mp/ssu044.

Results: The sequence for the gene-specific RNA-guide for the Cas9 endonuclease protein (designated sgRNA) was designed according to standard, accepted requirements (Jinek, et. al.). Following delineation of the seven exons and six introns of *SD-RLK* gene *At1g11300*, the first exon of the gene was chosen for sgRNA targeting in hopes of maximum "knockout" efficiency. The exon sequence was entered into the CRISPR-P program (Yang, 2014), and a target sequence, which had reduced off-target homology, was selected (Figure 2).

Table 1. Nucleotide sequence of forward primer CR11300FP1, reverse primer CRRP1, and the final sgRNA gene PCR product these primers generate from plasmid template pICH86966::sgRNA PDS. Red indicates the 20 nucleotide guide sequence that will recognize the complementary sequence in the *At1g11300* gene, targeting Cas9 endonuclease digestion. Blue indicates the BsaI recognition sites used in the first cut-ligation cloning. Black sequences encode the portion of the sgRNA needed for Cas9 interactions. Finally, all bolded bases in the PCR product are of primer origin.

Sequence Name	Sequence
CR11300FP1 (Primer)	5' TGTGGTCTCAATTGGAGTTGTCTGCTAGTGGATGTTTTAGAGCTAGAAATAGCAAG 3'
CRRP1 (Primer)	5' TGTGGTCTCAAGCGTAATGCCAACTTTGTAC 3'
At1g11300-1	**TGTGGTCTCAATT**GGAGTTGTCTGCTAGTGGAT**GTTTTAGAGCTAGAAATAGCAAG**TTAAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTCTAGACCCAGCTTTCTT**GTACAAAGTTGGCATTACGCTTGAGACCACA**

To generate of the *At1g11300* sgRNA-encoding DNA, a polymerase chain reaction was conducted using target DNA containing forward PCR primer CR11300FP1, reverse primer CRRP1, and template plasmid pICH86966::sgRNA PDS (Table 1).

The resulting the *At1g11300* sgRNA-encoding DNA was then placed under the regulatory control of the *A. thaliana* U6 promoter (for initiation of RNApIII transcription) in plasmid pICH47751 by the first cut-ligation reaction using restriction enzyme BsaI (Figure 3).

Subsequently, (1) the AtU6 promoter::sgRNA gene region of pCK300-1sgRNA was excised using restriction enzyme BbsI and combined with (2) plasmid pICH47732's NPTII gene (necessary for selection of transgenic plants via kanamycin resistance), (3) plasmid pICH47742's gene encoding the Cas9 enzyme of the CRISPR system and (4) plasmid pICH41766's linker DNA into a Ti plasmid pAGM4723 by a second cut-ligation reaction (Figure 4). Portions of the resultant pCK300-1CR were sequenced to ensure proper assembly.

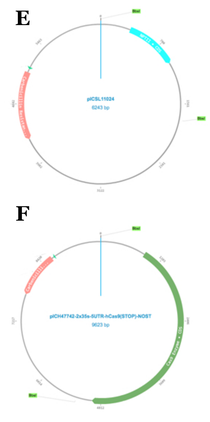
The development of pCK300-1sgRNA and pCK300-1CR constructs relied on the BsaI and BbsI restriction enzymes, which digest DNA outside of their recognition site, producing unique, designed single-stranded DNA overhangs. With these specific sticky ends, the DNA cassettes listed above could be annealed in the expected orientations found in the pCK300-1sgRNA and pCK300-1CR constructs.

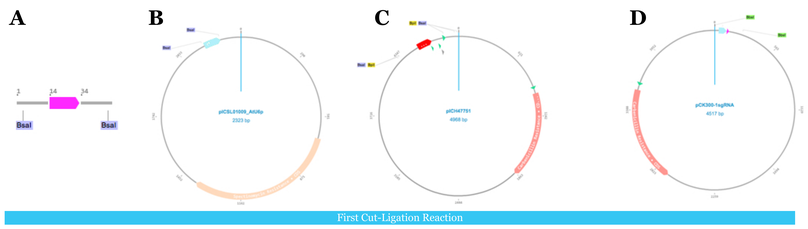
The above methodology was performed by each of the sixteen students in the Molecular Biology II Teaching Laboratory to generate one to two CRISPR knockout plasmids for five *SD-RLK* and four *SCRL* genes (Table 2).

Table 2. List of A. thaliana target genes, names of CRISPR constructs generated, targeted sequence in the sgRNA, and names of CRISPR construct creators.

A. thaliana Target Gene	Name of CRISPR Construct	Target Sequence in sgRNA	CRISPR Construct Creator
At1g11300 (RLK)	pCK300-1CR pSL300-2CR	GGAGTTGTCTGCTAGTGGAT GTTGTTCATAATGAAGAGCTC	Christopher Kilner Stephen Long
At1g11330 (RLK)	pGN330-1CR pJW330-2CR	GAGACCCTTCTATGCAAAAG GCAAGTAGCATTTGGAGCTAC	Gary Ng John Wetawski
At1g61380 (RLK)	pAV380-1CR	GAGAACCCATGTGATCTATA	Ashley Villa & Christopher Conway
At1g61480 (RLK)	pCS480-1CR pAB480-2CR	GCCTCTCATAGTACACGCCT GACGCAGTGCAATTTTCTGCA	Cody Sacks Alyssa Biondo
At1g61500 (RLK)	pES500-1CR pMS500-2CR	GTTCACTGGGATACCGTTTA GAGGGAATAGGGTGCCTAGTG	Erika Sarno Michelle Sulzinski
At1g60983 (SCRL8)	pDM983-1CR pNM983-2CR	GGAGTGCAAGTGCAATCGGT GCCATTGCGGAGAAAAGGCCG	Daniel McGraw Nicole McAndrew
At1g60987 (SCRL5)	pAD987-1CR	GATCAAGGATGAATATTAATG	Anthony DiLeo & Norman Frederick
At1g45875 (SCRL27)	pGO875-1CR	GTATTCGAACAAAATTGTGGG	Grace O'Neill & Christa Musto
At4g22105 (SCRL26)	pKD105-1CR	GGCTAACCTACGCGTCCGGA	Kathleen Dwyer

Figure 2. Selection of target sequence via CRISPR-P Program (http://cbi.hzau.edu.cn/crispr/). CRISPR sgRNA design requires selection of 20 base pair nucleotide sequence within the target gene of interest (Jinek, et. al. 2012). The N20 sequence must be selected near a consensus PAM NGG sequence, as the NGG sequence at the 3' end is needed for the Cas9 endonuclease to bind. Along with this NGG 3' requirement, it should be noted that the last 8-10 basepairs of the N20 sequence are most important for Watson-Crick basepair binding of the target gene. With utilizationof the U6 promoter for RNApIII, the selection of the sgRNA must take into account a GN19 sequence; when added to the previous requirements, the sgRNA target sequence will be 5' GGAGTTGTCTGCTAGTGGAT 3' followed immediately by TGG (not part of the sgRNA itself). The target sequence for the sgRNA can be chosen from either strand, as the Cas9 endonuclease induces a double-stranded cut. To reduce indels occurring at off target homologous sequences, and thus generating phenotypic results later on that are not related to the function of the gene of interest, an effort was made to choose a target sequence which had the least homology with the rest of the *A. thaliana* genome.





First Cut-Ligation Reaction

Figure 3. Summary of the First-Cut Ligation Reaction. BsaI is a restriction enzyme that produces staggered, double-stranded breaks in the DNA outside of the recognition sites generating specifically chosen sticky ends. Thus, the desired DNA cassettes in a cut-ligation reaction can be designed, via these specifically chosen sticky ends, to assemble in only one order. In this instance that means (A) the designed sgRNA gene PCR product (magenta) will be preceded and regulated by (B) the *Arabidopsis thaliana* U6 promoter (sky blue) from plasmid pICSL01009::AtU6p when ligated into (C) the pICH47751 plasmid, thus forming (D) the final product pCK300-1sgRNA, containing the new *AtU6p::sgRNA* gene. This process also excised a portion of the *LacZ* gene (red) in pICH47751, enabling blue/white selection of cells transformed with pCK300-1sgRNA.

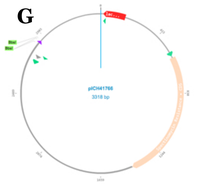
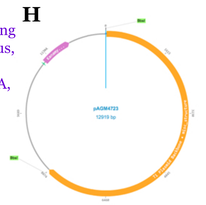


Figure 4. Summary of the Second-Cut Ligation Reaction. Similarly, the second-cut ligation reaction employed restriction enzyme BbsI to produce specifically chosen sticky ends flanking each DNA cassette to be assembled in the proper order into the final CRISPR construct. Thus, the *NPTII* kanamycin resistance gene (light blue) of (E) pICSL11024, the *Cas9* gene (green) of (F) pICH47742, the new *AtU6p::sgRNA* gene (sky blue/magenta) of (D) pCK300-1sgRNA, and the linker DNA of (G) pICH41766 (light purple) were ligated in this order into (H) Ti plasmid pAGM4723, producing the final CRISPR construct (I) designated pCK300-1CR. This process excised the gene encoding a red pigment present in the unmodified pAGM4723 vector, enabling red/white selection of cells transformed with pCK300-1CR.



Conclusions: The potential use of the CRISPR/Cas9 system as a teaching model has also been demonstrated, as the construction of pCK300-1CR and other CRISPR plasmids (Table 2) occurred over the span of one semester—with a lab period of three hours, once a week, for fourteen weeks.

In the near future, *A. thaliana* plants will be transformed by the CRISPR constructs via *Agrobacterium tumefaciens* infection (a process that facilitates the random integration of the CRISPR related T-DNA into plant chromosomes). Following transformation, the transgenic plants will undergo further genetic and phenotypic assays to characterize the function (or lack thereof) of the targeted genes. These analyses will determine the efficacy of the CRISPR/Cas9 system for "knocking out" specific *SD-RLK* or *SCRL* genes.

Second Cut-Ligation Reaction (Golden Gate Ligation)