Mutating AT3G08680 Using CRISPR-Cas9 to Determine Functional Redundancy

Vanessa Kwong

Dynamic Genome

12 March 2021

Background

Roots play a significant role in the biological aspect of almost every plant—from providing nutrition for growth and storing resources for development, to giving and maintaining its structure. Depending on the environment, roots may take on different appearances for the survival of each plant. For instance, in an environment where erosion often occurs, a plant may have fibrous roots for stronger anchorage (Van Norman, 2021). The spreading of roots in soil will ground the plant as water and additional nutrients are absorbed. Beneath the surface-level view of a root, there are many components that carry out each of these tasks as a system.

A eudicot seedling's primary root consists of three sections: the meristematic zone for cell division and formation, elongation zone for cell lengthening, and the differentiation zone for cell specialization (Van Norman, 2021). Cell-to-cell communication is crucial to maintain the organized cell structures in roots and carry out the functions needed for survival. Polarity in plant cells can be observed when a protein localizes towards a specific region of a single cell, whether it be the shootward, rootward, inner, or outer areas. In Figure 1, for instance, the region colored in red represents an Inflorescence and Root Apices Receptor Kinase (IRK) protein localized towards the outer plasma membrane of the plant cell (Campos, et al, 2020). Cell polarity often results from evolution and adaptation of the plant to its environment (Yang, 2008), and is usually greatly involved in cell growth, transportation of nutrients, or bringing in stimuli for growth from outside sources (Wallner, 2020).

Arabidopsis thaliana, the species used in this experiment, is a well-known weed plant model organism, particularly with experiments dealing with genetics. The efficient generation and self-reproduction of *A. thaliana* (Koornneef & Meinke, 2010), as well as the extensive research done globally

has made the organism a suitable model for research over time (The Arabidopsis Genome Initiative, 2000). A. thaliana was one of the first plants to be sequenced, meaning that there are a wide variety of resources available for this research to be built upon. Conclusions drawn from studying this model organism can often be applied to related species or used as references when conducting research on other plants.

Polar localized kinases, or PLKs, code for transmembrane receptor kinases, which are an important factor in communication between cells. WALLFLOWER (WFL/PLK3) and KINASE ON THE INSIDE (KOIN/PLK4) are two examples of PLKs. KOIN, predicted to act as a cell division repressor, is localized towards the inner polar domain of endodermal cells in the meristematic and elongation zones because of its kinase domain (Van Norman, 2021). WFL, which likely provides the positioning of gaps that help with hair growth, is localized towards the inner polar domain in epidermal cells of the meristematic and elongation zones, along with the lateral root (Van Norman, 2021). Both PLKs consist of a transmembrane helix domain, leucine-rich repeat N-term domain, several leucine-rich repeats, and kinase domains which are approximately 65% similar (Van Norman, 2021).

The objective of this research project is to mutagenize and "knock-out" various genes related to WFL and KOIN to determine if there is redundancy in functionality. Selected pairs of guide RNAs for the coding sequences of each gene are designed to match the target sequence on the non-complementary strand of DNA. Primers are made for each pair of guide RNAs and are used with the CRISPR-Cas9 system to cleave the genomic sequence, referencing to a protospacer adjacent motif (PAM) sequence. PAM sites contain nucleotides NGG and position nucleases in the right orientation to cut. This results in a large deletion, or indel, of nucleotides in the coding sequence that may impact the proteins coded for. Insertions or deletions of a section in a gene can result in a shifted frame for translation, and therefore, cause different proteins to be made than the original. In this case, the sequence coding for a PLK is altered. It is important to target sequences that code for the kinase domain due to its importance in polarity, and sequences outside the kinase domain. Ultimately, the comparison of *A. thaliana* phenotypes between ones with mutated and wild genetic sequences will determine whether each gene covers the same function as and WFL and KOIN.

Experimental Design

The gene assigned for this experiment is at locus AT3G08680 within the KOIN subgroup in LRR III (Figure 2), which codes for a PLK classified as a leucine-rich repeat receptor-like kinase, or LRR-RLK. It functions primarily in protein serine and threonine kinase activity, kinase activity, and ATP binding, and is expressed mostly during the mature pollen stage.

AT3G08680 consists of 3247 base pairs in its genomic DNA and 1923 base pairs in its coding sequence. In Figure 3, the coding sequences of the gene are represented by a pink color and are within the exons. Two total exons are present in the coding sequence; the first positioned at 1000 to 2282 nucleotides with a total of 1282 base pairs, and the second positioned at 2358 to 2999 nucleotides with a total of 641 base pairs.

The primer selected for the gene is Primer 1, with a total product size of 1545 nucleotides. The forward primer begins at the position of the 1212th nucleotide and ends at 1231st nucleotides. The reverse primer spans from the 2737th nucleotide to the 2756th nucleotide. Its sequences are shown in Figure 4, along with GC percentages. This primer is designed for the coding sequence of the assigned gene, AT3G08680. The alignment of the AT3G08680 coding sequence, the selected primer, and guide RNAs are shown in Figure 5. In the image, it is seen that the primer does not cover the entire span of the short guide RNA pair.

In the phylogenetic tree shown in Figure 6, genes in the LRR III subfamily that are closely related to PLKs KOIN and WFL are shown, out of a total of 35 proteins (Van Norman, 2021). The assigned gene, AT3G08680, is in the KOIN subgroup. With this research project, the pictured genes can be mutated and "knocked-out" to determine the redundancy of functionality for each protein. To help eliminate genes that provide the same function as the assigned and target homologs, gRNAs with more than one hit can be used, mutating multiple PLKs at once.

Expression of AT3G08680 is mostly in the carpel and cauline leaves, meaning that the mutation of this gene can impact the leaf and carpel cell structures. In Figure 7, the darker region of the pictured root

shows high expression of the gene in the cell division zone. There may be changes in the shape of leaves or carpel in the *A. thaliana* plant with the mutated version of the gene.

The first pair of guide RNAs selected, AT3G08680_6_3055 and AT3G08680_2_3055, both target AT3G08680::AT3G08680, and are located downstream, as shown in Figure 8. This pair targets on-target sites of the gene since they both have 1 hit. The second pair of guide RNAs are AT3G08680_2_1833 and AT3G08680_0_3055, which are longer and each target different sequences. In this case, AT3G08680_2_1833 targets AT3G08680::AT3G08680, and AT3G08680_0_3055 targets AT3G17840::RLK902, as shown in Figure 9. AT3G08680_0_3055 has 4 hits, which means this pair of guide RNAs can target off-target sites, genes that are homologs.

Conclusion and Future Directions

Once the guide RNAs have been designed, the nuclease can supposedly cleave the genomic sequence in reference to a PAM site with the sequence NGG and take out the portion of nucleotides between the guide RNAs. This creates a large indel, which then creates a frameshift of nucleotides for its following codons. The new set of codons are likely to code for different amino acids, producing a new set of proteins than intended and causing a mutation. This mutated version of the gene replaces the wild-type copy and potentially cause changes in the phenotype of *A. thaliana*, helping determine its functionality. A larger change in phenotype, specifically repression of cell division, would signal that the gene and KOIN are functionally redundant.

The primers chosen for the coding sequence of gene AT3G08680 is the longest out of the ones generated by the NCBI Primer BLAST. However, it is unable to cover the entire span of both guide RNA pairs, which means that the deletion may be incomplete, and the mutant will not be able to impact the function of the assigned gene.

To verify that gene AT3G08680 was “knocked out”, observations of the leaf and carpel cell structures can be made, since those are the areas with most expression. Being a part of the KOIN subgroup, the assigned gene may be involved in cell division repression. A notable observation would be a greater number of plant cells in the leaves and carpel because of the lack of cell repression due to the mutant. It is likely there would be less expression in plant roots.

If a plant had a non-functional target protein for the gene, targeted mutagenesis on sites that are on-target may not cause a mutation, and no change in phenotype would be expressed. Study on cell division in plants may be advanced because it would demonstrate that there are multiple genes involved in cell division, and further research may be needed.

Figures and Tables

Figure 1. Polar Localization of IRK (Campos, et al, 2020)

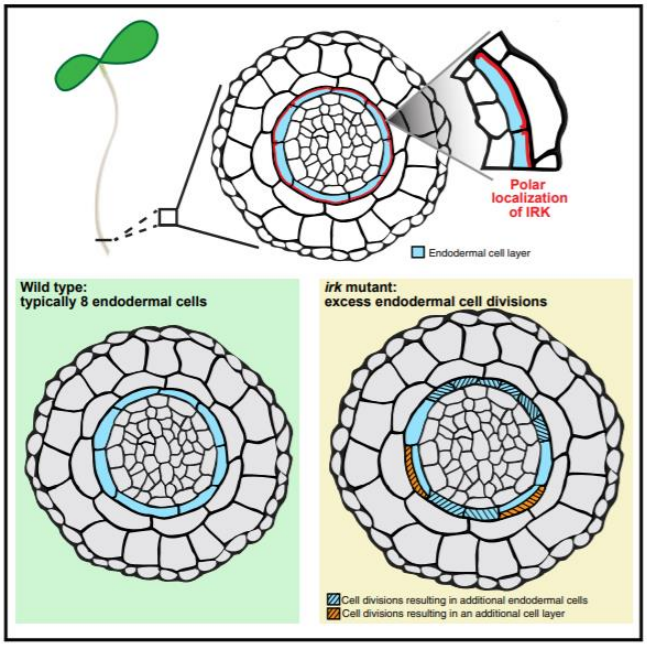


Figure 2. Receptor Kinase Subfamilies (Van Norman, 2021)

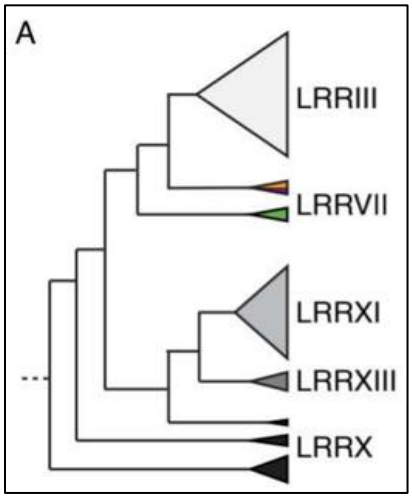
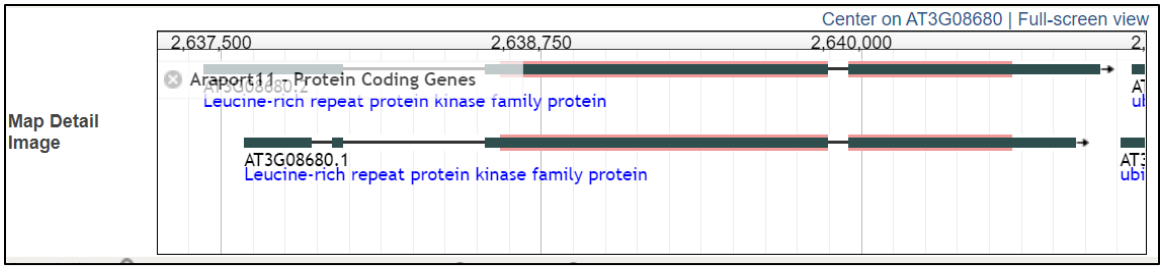
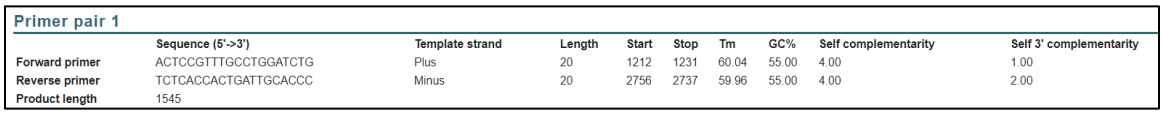


Figure 3. AT3G08680 Coding Sequence Regions (Arabidopsis.org)



Source: https://www.arabidopsis.org/servlets/TairObject?id=36238&type=locus

Figure 4. Primer Pair for CDS (NCBI Primer BLAST)



Primer pair 1

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	ACTCCGTTTGCCTGGATCTG	Plus	20	1212	1231	60.04	55.00	4.00	1.00
Reverse primer	TCTCACCACTGATTGCACCC	Minus	20	2756	2737	59.96	55.00	4.00	2.00
Product length	1545								

Source: https://www.ncbi.nlm.nih.gov/tools/primer-blast/primertool.cgi?ctg_time=1614899807&job_key=KyH0ZvwQ8bjWhvSD-ePQsYP4wYOu69qerw

Figure 5. Alignments of CDS, Primer, and gRNAs

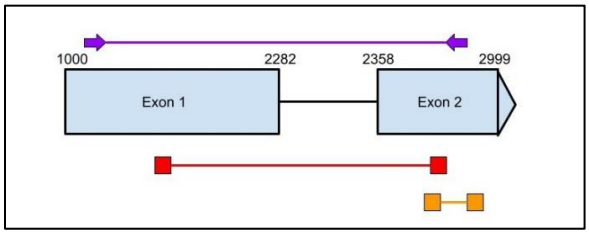


Figure 6. LRR Subfamily 3 (Van Norman, 2021)

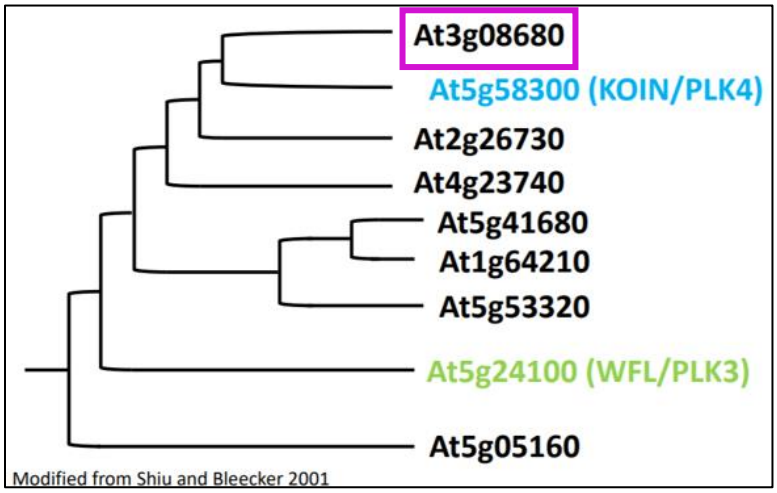
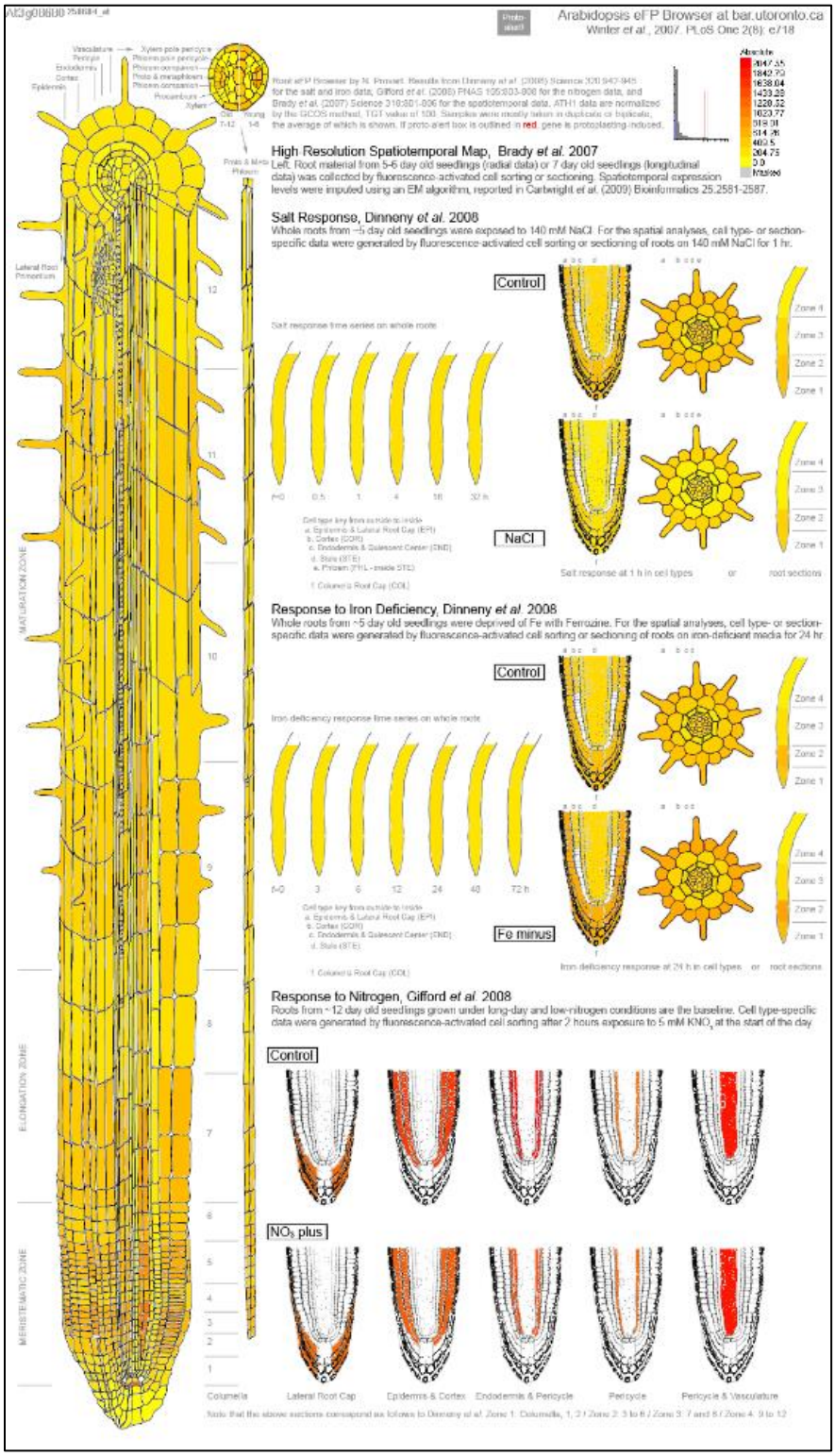
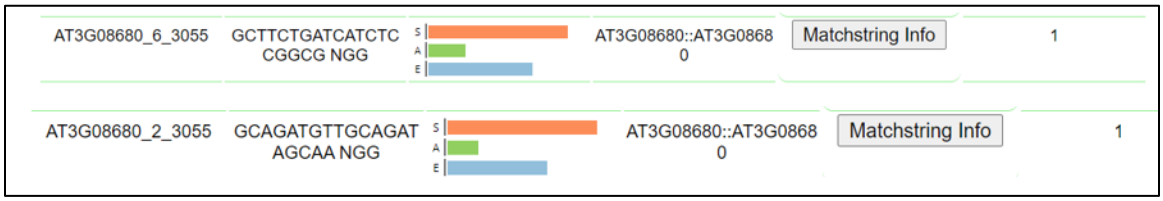


Figure 7. Gene AT3G08680 Expression in Roots (Arabidopsis.org)



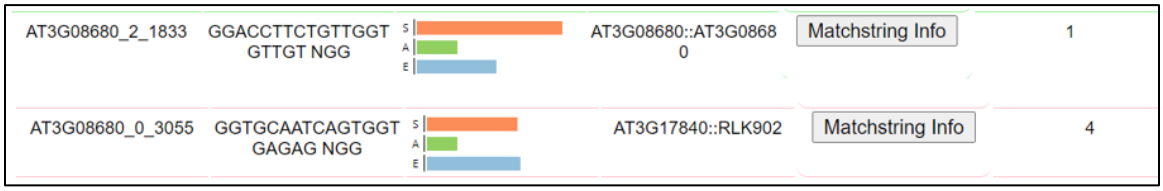
Source: https://www.arabidopsis.org/servlets/TairObject?id=36238&type=locus

Figure 8. Short gRNA Pair (E-CRISP)



Source: http://www.e-crisp.org/E-CRISP/workdir/Thu_Feb_25_21:25:54_20211614288354/index.html

Figure 9. Long gRNA Pair (E-CRISP)



Source: http://www.e-crisp.org/E-CRISP/workdir/Thu_Feb_25_21:25:54_20211614288354/index.html

Literature Cited

Arabidopsis Genome Initiative. (2000 Dec 14) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*. doi: 10.1038/35048692.

Campos, R., Goff, J., Rodriguez-Furlan, C., & Van Norman, J. M. (2020 January 27) The Arabidopsis Receptor Kinase IRK Is Polarized and Represses Specific Cell Divisions in Roots. Developmental Cell, 52 (2), 183-195.

Koornneef, M., & Meinke, D. (2010). The development of Arabidopsis as a model plant. *The Plant Journal*, 61, 909-921. doi: 10.1111/j.1365-313X.2009.04086.x

Van Norman, J. (2021). *Introductory Presentation* [PowerPoint slides]. College of Natural and Agricultural Sciences, University of California, Riverside. https://ilearn.ucr.edu/bbcswebdav/pid-4223925-dt-content-rid-44556217_1/courses/BIOL_020_004_21W/BIOL20_RootIntro_WFL%2CKOIN_11Feb2021small.pdf

Wallner, E. (2020) The value of asymmetry: how polarity proteins determine plant growth and morphology. Journal of Experimental Botany, 71 (19), 5733-5739. Doi: 10.1093/jxb/eraa329

Yang Z. (2008) Cell polarity signaling in Arabidopsis. *Annu Rev Cell Dev Biol*., 24, 551-575. doi: 10.1146/annurev.cellbio.23.090506.123233.