UNIVERSITY OF CALIFORNIA SAN DIEGO

Implementation of CRISPR/Cas9 Genome Engineering Tools in Polyploid Plant Genomes

A Thesis submitted in partial satisfaction of the requirements for the degree Master of Science

In

Biology

by

Maria Yurievna Omelchenko

Committee in charge:

Professor Martin F. Yanofsky, Chair
Professor José L. Pruneda-Paz, Co-Chair
Professor Mark Estelle

2020

The Thesis of Maria Yurievna Omelchenko is approved, and is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

Co-Chair

_____

Chair

University of California San Diego

2020

DEDICATION

I would like to dedicate this thesis to my family and friends. Without your love, unconditional support and consistent encouragement throughout this process I would not have been able to get to where I am now.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to first say thank you to Dr. Martin Yanofsky for giving me the opportunity to do research and be a Master's student in the lab. I am thankful for his guidance and aid that he has provided me throughout the program. I would also like to thank the committee members, Dr. José L. Pruneda-Paz and Dr. Mark Estelle for agreeing to be a part of my committee and for taking the time to attend my defense. I'd also like to thank JJ for all of his help, support, and mentorship throughout my time in the lab. Without his wise words, encouragement and assistance, I would not have been able to make it this far. Lastly, I would like to thank the Tata Institute for Genetics and Society (TIGS) for granting me the 2019-2020 fellowship and supporting this thesis. Thank you.

ABSTRACT OF THE THESIS

Implementation of CRISPR/Cas9 Genome Engineering Tools in Polyploid Plant Genomes

by

Maria Yurievna Omelchenko

Master of Science in Biology

University of California San Diego, 2020

Professor Martin F. Yanofsky, Chair
Professor José L. Pruneda-Paz, Co-Chair

Plants serve the basic energy and food source for all living organisms on Earth and have been bred for centuries in order to maximize qualities that are desirable for humans, including organ size or yield. Fruit are the harvested product in many crops, and thus, many fruit traits have been targeted in breeding programs for decades. Current global factors such as climate change, reduced water availability or the availability of arable land, will largely impact breeding and crop yield in the 21st century, jeopardizing food security of our increasing human population. The rise of CRISPR/Cas9 genome editing technologies has provided new promising venues for crop improvement. More recently, ground-breaking Active Genetics tools have provided novel opportunities to explore new techniques for faster breeding and circumvent the constraints of

classical Mendelian inheritance. This becomes even more important when considering that many of our crops are polyploid. As an initial step towards implementing Active Genetics in polyploid species, we need to find a configuration in which CRISPR is efficient. In this study, we have investigated different CRISPR configurations for efficiently targeting polyploid plant genomes with the goal of being able to transfer this knowledge to crops. We hope that we can then accelerate plant breeding in order to elaborate new cultivars to meet future food and yield demands.

INTRODUCTION

Plants are the basic energy and food source for all living organisms on Earth. Plants have been bred for centuries in order to maximize their size and yield. However, there are a number of global factors that currently impact plant breeding and crop yield in the 21st century. Such factors include climate change, reduced water availability, carbon capture and carbon fuel. In the upcoming years, an increase in global temperature as well as a decrease in precipitation is predicted to result in a higher variability and lower yields amongst crops, which will be detrimental to the increasing global food supply (Kang et al., 2009). These aforementioned factors pose a challenge for the rising human population, which has already increased nearly four-fold during the 20th century and is predicted to increase to 9.1 billion by 2050 (Roser et al., 2013; "How to Feed the World in 2050"), as compared to 8 billion in the world today. Establishing global food security under these circumstances will continue to be challenging as hunger still persists across many nations in the current world (Beyaz and Yildiz, 2017). However, even if there is enough food to feed everyone, hunger will still continue to exist as many people do not have the money or resources to access food supply (Holt-Giménez et al., 2012). For nearly 10,000 years, informal selection by farmers for desirable traits lead to a gradual improvement of crop yields. Following Mendel's discovery of the principles of inheritance, this informal selection gave way to systematic plant breeding approaches ("Genetic Engineering of Plants: Agricultural Research Opportunities and Policy Concerns", 1984). Fortunately, genetic research and continuous breeding to produce elite crop genotypes has been an ongoing development in order to tackle these difficulties (Prohens 2011). Recently the ground-breaking CRISPR/Cas9 genome editing technology has opened the door toward rapid crop modification. Indeed, the ability to rapidly modify crop plants may prove

essential as we try to feed a growing human population, while at the same time, face rapid fluctuations in environmental conclusions due to human-caused climate change. Fruits and the seeds they contain are perhaps the most important agricultural product as they are what are commonly harvested in many crop species and have an abundance of nutritional and medicinal benefits. Additionally, fruit have also been harvested to be used as sustainable biofuel sources, such as using corn (*Zea mays*) to produce ethanol (Shapouri et al., 2002). The studies of *Arabidopsis thaliana*, the reference organism in plants (*Arabidopsis* hereafter), and its fruit were crucial for the understanding of the underlying mechanisms that controlled how a fruit opens to release its seeds (Dinneny et al., 2005; Alonso-Cantabrana et al., 2007; González-Reig et al., 2012; Roeder et al., 2006; Gu et al., 1998; Liljegren et al., 2000; Ferrándiz et al., 2000; Roeder et al., 2004; Lilijegren et al., 2004; Roeder et al., 2005; Sorefan et al., 2009; Ripoll et al., 2011; Romera-Branchat et al., 2013). These findings became crucial for the development of another important crop, *Brassica napus*, in which fruits are key for their production is canola. The oil is taken from the seeds and it is considered to have an abundance of health benefits (Lin et al., 2013). Now with the development of tools such as molecular breeding the process of plant breeding to increase crop yield and fruit abundance can be expedited. Current molecular breeding approaches have also been targeting drought and salt tolerance in crop species by understanding how plants develop under saline conditions as well as how their root growth and stomata operate in response to water scarcity (Jenks et al., 2007). Additionally, molecular breeding advances have been made by using artificial microRNAs to target gene activity in plants in order to modify and improve agronomic traits that would be beneficial towards the overall goal of improving crop yield (Liu and Chen, 2009). One of the hurdles in being able to rapidly modify crops for increased yield has been the ability to create precise genome edits. Recent advances of genome editing using CRISPR/Cas9 have opened

the door to a new revolution in crop modification. In order to bypass some of these challenges, we need to find a configuration in which CRISPR is efficient in polyploid genomes, since many of the most important crops such as wheat, potato, cotton, apple, and canola (Weeks 2017) are polyploid. For the delivery of CRISPR/Cas9 machinery, *Agrobacterium tumefaciens* is mainly used for the delivery of CRISPR/Cas9 but, there are some plant species that are reluctant to *Agrobacterium*-mediated transformation. However, new strategies have been developed mainly based on the use of gold particles (Gan 1989) and more recently artificial molecules based on carbon nanoparticles (Demirer et al., 2019). Even so, the ability to modify more complex polyploid plant genomes remains a critical challenge and efficient editing based on CRISPR for polyploid genomes is far from efficient. Currently, there is ongoing research for developing additional strategies in order to find a configuration in which CRISPR is efficient in polyploid genomes by using *Arabidopsis thaliana* (*Arabidopsis* hereafter). *Arabidopsis* has a diploid genome and the vast majority of the CRISPR work in *Arabidopsis* has been done using diploid (2n). As an initial step towards efficiently applying CRISPR to polyploid crop species we employed an *Arabidopsis* polyploid strain as a "proof of concept" for our studies. Furthermore, we can continue to improve the CRISPR strategy in polyploids by introducing a newly developed technology known as Active Genetics that we believe can revolutionize agricultural biotechnology.

More recently, the Bier Lab at UC San Diego has developed Active Genetic methodology using CRISPR in order to carry out Super-Mendelian inheritance that permits the passage of desired traits to the next generation circumventing the constrains of normal mendelian inheritance. The implementation of Active Genetics in different systems has opened new venues for engineering genomes and create genetic combinations that following classical Mendelian inheritance are virtually impossible. Therefore, and in addition to identifying efficient

configurations for efficient CRISPR/Cas9-mediated editing in polyploids, we also wanted to evaluate whether the tools developed for this end can be also exploited for Active Genetics in polyploid crop genomes. If successful, we can then accelerate plant breeding employing Active Genetics and create the desired genetic combinations needed to mitigate these environmental and population challenges and improve agricultural efforts in producing an improved crop supply. On the other hand, these strategies can also provide an excellent tool that can expedite any breeding program in order to meet future food and yield demand and accelerate basic research investigations (Figure 1).

The polyploid challenge

Polyploidy is abundant amongst many crop species such as wheat, cotton, potatoes, oats, and coffee (Figure 2). Polyploid crops contain more than two sets of homologous chromosomes, which are known to dramatically contribute to an increased mass yield compared to their diploid equivalents (Corneille et al., 2019). An additional benefit to polyploidy is heterosis, which contributes to increased biomass and gene redundancy, as it protects the plant from inheriting detrimental recessive mutations (Comai 2005; Osborn et al., 2003). Polyploidy increases the complexity of genomes, and represents a challenge for breeding programs and crop engineering (Byfield and Wendel 2014). Moreover, polyploidy increases gene copy number, which can be also translated into higher gene redundancy making the interrogation of gene functions more difficult and classical mutagenesis strategies less efficient (Mao et al., 2019).

In this context, we propose to identify an efficient CRISPR/Cas9-based approach to edit and engineer polyploid genomes as an initial step towards implementing and Active Genetics to assist in circumventing the issues associated to polyploid genomes. As proof of principle, we are employing *Arabidopsis* for our studies. In addition to the large number of advantages that

Arabidopsis offers as reference model, our lab has obtained polyploid *Arabidopsis* strains (4n, 8n) to test different CRISPR-Cas9 editing configurations and evaluate the feasibility of Active Genetics. Interestingly, *Arabidopsis* has many relatives that are crops and are cultivated as polyploid species including members of the *Brassica* species, such as canola (Stephenson et al., 2019). Additionally, *Arabidopsis* has many genetic similarities with other crop species such as tomato, wheat, barely, potato, and cotton, which makes it an ideal model organism for research in agricultural improvement (Penin et al., 2019; Peng et al., 2015; Gebhardt et al., 2003; Xu et al., 2014).

History of Gene Editing. Pre-CRISPR era.

The plant genome has been naturally modified over the course of thousands of years, beginning with the major crops still present today being domesticated approximately 5,000-10,000 years ago. As human civilizations evolved over time, agriculture had a parallel evolutionary timeline as well. The informal selection of desired plant traits has been prevalent throughout the course of human history as farmers would naturally pick their desired individuals from the natural variation present in a crop population (Kossman 2012). The development of hybrid maize due the result of hybrid vigor, or heterosis, which is when the offspring from the cross of two inbred lines are much higher yielding greatly contributed to the increase of corn yield (Barber et al., 2012). Then in the mid-20th century, the Green Revolution emerged and spread throughout the world, particularly impacting developing countries such as Pakistan, India, and the Philippines (Borlaug and Narvaez, 2001). The Green Revolution was a term initially coined by William Gaud, who was the Administrator for the United States Agency for International Development. Gaud noted that there was an agricultural revolution occurring in which the surplus of wheat is dramatically

increasing in the aforementioned countries (Gaud 1968). The resulting product of the Green Revolution was what is known as the 'golden harvest,' which resulted in a global increased yield of wheat, rice, and maize (Borlaug and Narvaez, 2001). The Green Revolution dramatically increased the grain production by 174% between 1950 and 1990, which surpassed the increasing population growth during that time period (Otero and Pechlaner 2008). The Green Revolution was largely successful due to the development of genetically enhanced crops (Khush 2001). Norman Borlaug, also known as the "Father of the Green Revolution" dramatically improved crop yields by introducing semi-dwarf grain varieties that were highly abundant in yield (Swaminathan 2009). Although it was a ground-breaking discovery for the 20th century, our current world is facing a completely different scenario that will worsen in the next decades. We will now require powerful tools to cope through the upcoming challenges that come with an increased population size as well as continuous global climate change and its effects on crop biomass. Genome engineering and gene editing appear to have the answers to the challenges that years of previous research have yet to conquer.

Prior to the development of the current tool-kit for genome editing there were some classical methodologies for manipulating genomes and creating mutations and new alleles. The development of genome editing methodologies has been on a rise in order to understand the regulation and foundation of biological systems and improve crops (Kamburova et al., 2017). There are various genome editing techniques that have superseded traditional breeding and hybridization techniques that have been implemented over time and have provided more efficient and faster methods of breeding and improving crop yield. One of them is known as mutation breeding in which chemicals or radiation are applied to generate mutant plants that contain desirable traits. Radiation breeding, a subclass of this technique, was discovered by Lewis Stadler

who used X-Rays to induce mutations in crop species including maize (Stadler 1928). Mutation breeding has contributed to new crop varieties (Kharkwal and Shu, 2009). However, one of the downfalls of these techniques is that the resulting mutations are random and can result in large detrimental deletions or duplications of chromosomal segments (Hartwell 2017). Continuous development of genome editing *in planta* resulted in another technique called marker-assisted selection (MAS) to be implemented in plant breeding. This technique was proven to be effective because it uses DNA markers that are close to the gene of interest target and can thereby track the such genomic sequence during crossing. Consequently, this method allows for better identification and quantification of the genetic variation within a population of plants or crops (Xu and Crouch 2008; Tanksley et al., 1989; Tanksley and McCouch, 1997; Gur and Zamir, 2004). Molecular markers for plant breeding were first utilized in the early 1980s using isozyme markers, however these markers very limited in terms of the polymorphisms required for efficient gene tracking (Goodman et al., 1980). A few years later, Beckmann and Soller used restriction fragment length polymorphism (RFLP) markers to improve crops (Bechmann and Soller 1986a). RFLPs were particularly useful in maize because there were many polymorphisms in maize, which allowed for successful genetic mapping. (Helentjaris et al., 1986; Burr et al., 1988; Hoisington, 1989; Coe et al., 1995; Davis et al., 1999). Though these sequences were suited for MAS experimentation, one disadvantage to using RFLPs was that they were particularly lengthy sequences (Ragot and Lee, 2007). Additionally, Beckmann and Soller discussed the theoretical concerns of MAS in back-crossing (MABC) in order to improve the plant trait qualities (Beckmann and Soller 1986a). MABC is a useful technique when there is a variety of elite plants that require a single-trait improvement such as disease-resistance or stress tolerance (Cregan et al., 1999; Cahill and Schmidt, 2004; Johnson, 2004; Niebur et al., 2004; Eathington, 2005; Crosbie et al., 2006; Ragot

and Lee, 2007; reviewed by Xu, 2003; Miklas et al., 2006; Dwivedi et al., 2007). Mostly, MABC is used for selecting transgenes and removing undesired regions of DNA that reduce crop fitness (Fehr, 1987). Lastly, MABC reduces the number of generations that it would take to achieve a desired phenotype if done using conventional back-crossing technique (Crosbie et al., 2006; Ragot et al., 1995). One major limitation to MAC, however, is the cost of production marker data points and generating high-grade phenotypic data (Ragot and Lee, 2007).

A key element for understanding how a biological process works is to interrogate all the gene functions involved and study the resulting phenotypes. This concept is universal to all fields of biology, and is of especial interest in agricultural research and crop improvement. In this context the research community has developed a battery of different strategies. One of them involves the use of RNAi (RNA interference). RNAi is a common gene knockdown strategy by which a small RNA molecule can bind to and silence targeted genes within cells. RNAi has been shown to be useful in crops as it has contributed to developing insect and disease resistance as well as affecting flowering time and improved nutritional benefits (Guo et al., 2016). One of the downfalls of this technology, however, is that the results may not be reproducible, gene silencing is only temporary, and off-target effects are common.

The onset of genetic modification tools emerged as a promising tool for understanding gene functions. New strategies such as zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) that involved creating double-stranded breaks (DSBs) in the plant genome appeared (Gaj et al., 2013). Zinc fingers are protein modules that were first introduced in 1996 as a useful genetic engineering tool when linked with the *FokI* endonuclease (Kim et al., 1996). Together, this technology came to be called zinc finger nucleases (ZFNs) as these nucleases were site-specific and could recognize their desired DNA target sequence when fused to a zinc

finger domain (Gaj et al., 2013). ZFNs have shown to be extremely advantageous in that they offer high specificity and efficiency and have fewer off target effects (Kamburova et al., 2017). TALENs are another class of nucleases that emerged as an additional powerful gene editing tool. These nucleases consist of the *FokI* endonuclease fused to the DNA-binding domains of the transcription activator-like effector (TALE) proteins (Gaj et al., 2013), which can bind to and activate specific plant promoters by recognizing certain repeat domains (Jankele and Svoboda, 2014). TALENs were successfully introduced into *Arabidopsis* as well as into several crop species such as rice, wheat, potato, and tomato, proving to be a useful tool in crop engineering (Xiong et al., 2015). Such tools expedited plant genome engineering research and provided additional tools for the ongoing improvement of crop species in order to increase global food supply (Weeks at al., 2016). Although these technologies provided new venues for engineering genomes, the very low variety of sequences that can be targeted or the complexity in the elaboration of the corresponding construct, represent a drawback that limit their abilities to needs in basic research, or more importantly, in agriculture.

Perhaps one of the one of the most universal gene editing technologies that has been developed during the 21st century is the type II clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9) (Gaj et al., 2013). Though the CRISPR system was first been identified in *Escherichia coli* in 1987 by Ishino *et al* as part of the prokaryotic adaptive immune system (Ishino *et al*., 1987), it wasn't until 2012 that the Cas9 protein was discovered (Jinek *et al*., 2012). Figure 3 describes the timeline of the gene editing strategies explained in detail above.

CRISPR/Cas9

Clustered Regularly Interspaced Short Palindromic Sequences (CRISPR) are short repeating sequences are naturally located in a specific CRISPR loci within the prokaryotic genome and are separated by spacer DNA, which are unique sequences that have been acquired over time from pathogens that have previously invaded the bacterial or archaea cells (Rath et al, 2015., Sontheimer and Barrangou, 2015; Mojica et al., 2005). Additionally, genes for Cas proteins are located in the CRISPR loci that mediate the immune response (Amitai and Sorek, 2016). During the first stage of the prokaryotic immune response, which is known as adaptation, the spacer DNA is incorporated into the genome. It is heritable and can be passed on to the bacterial or archaea progeny, which in turn receive increased adaptive immunity to certain pathogens over the generations (Sontheimer and Barrangou, 2015). During the expression stage of the immune response, the spacer DNA is then transcribed into CRISPR RNA (cRNA). During the last stage of the immune response, the interference stage, the cRNA forms a complex with the Cas protein that can identify the foreign DNA due to the ability of cRNA to complementary base pair with the region of interest, signaling the Cas protein to degrade the pathogenic genetic material (Amitai and Sorek, 2016). The CRISPR/Cas9 RNA-mediated adaptive immunity system (CRISPR type II), has since been widely used as a powerful genome editing tool. This system uses the Cas9 endonuclease, which makes double stranded breaks at a targeted region of DNA (Farasat and Salis, 2016). A guide RNA sequence, consisting of a cRNA and trans-activating crRNA (tracrRNA) is designed in order to bind to the targeted DNA region of interest, as well as to the Cas9 endonuclease, which in turn makes a cut at the target region. Additionally, the Cas9-RNA complex recognizes the PAM site, a non-target DNA sequence region located adjacent to the 3’ end of the gRNA target region. This recognition allows the Cas9 target the DNA sequence and cut it

(Sternberg et al, 2014). The bacterial repair system can use non-homologous end joining (NHEJ), which results in random insertions or within the targeted genomic region and creates a knock-out gene. The bacterial repair system can also use homology-directed repair if a specific DNA template is provided to create a knock-in within the genome (Rouet et al., 1994a; Rouet et al., 1994b). (Figure 4).

The CRISPR/Cas9 technology has been used in a variety of prokaryotic and eukaryotic organisms, including bacteria, flies, plants, mice, rabbits, and even human cells (Farasat and Salis, 2016). Using various transformation methods, this technology was first introduced in 2013 into the plant field using *Arabidopsis thaliana*, *Nicotiana benthamiana*, and rice (Feng et al., 2013, Li et al., 2013, Nekrasov et al., 2013, Shan et al., 2013, Xie and Yang, 2013). Feng et al. successfully targeted and disrupted the following genes: *BRASSINOSTEROID INSENSITIVE 1 (BRI1)*, *JASMONATE-ZIM-DOMAIN PROTEIN 1 (JAZ1)* and *GIBBERELLIC ACID INSENSITIVE (GAI)*, which led to the production of the CRISPR modified transgenic *Arabidopsis* plants.

CRISPR/Cas9 technology is a powerful tool to expedite breeding, but whereas important crop species such as cotton, potatoes, and wheat are polyploid (Weeks 2017), this tool has been mainly used for diploid plants genomes. In this context, we want to see if the efficiency of generating CRISPR/Cas9 mutations in a polyploid organism is just as efficient as it would be in a diploid organism. A limiting factor that has been evident is that the transformation efficiency of polyploid plants is much lower than that in diploid, as there are more alleles that require targeting by CRISPR/Cas9 technology, which decreases the accuracy. Additionally, with an increased number of alleles comes a possibility of having different mutations in the alleles, which results in hetero-allelic (or poly-allelic) plants. This possibility can be problematic, in particular when for

breeding purposes a specific or a unique allele is required because it becomes uncertain as to which mutated allele is contributing to the resulting phenotype. CRISPR paved the way for additional technologies that have been employed in order to develop more elaborated ways to target genomes. One of these new ground-breaking technologies is Active Genetics.

Active Genetics: A Tool for Cheating Mendel's Laws

Moreover, we want to see if the mutations induced using CRISPR could be spread in a faster way using 'super mendelian' strategies known as Active Genetics. Active Genetics is a strategy that bypasses Mendelian Laws of inheritance. This technology has been previously used by Gantz *et al* to produce autocatalytic homozygous loss-of-function mutations in *Drosophila melanogaster* by skipping the need for multiple generations of breeding (Figure 5) (Gantz and Bier, 2015). Active Genetics strategies use a so-called gene drive that consists of the expression of Cas9 and a gRNA encoding cassette that is designed to target a gene of interest, and homology arms that flank the Cas9/gRNA cassette and correspond to the left and right borders of the cas9 cutting site. Cas9 makes a DSB cut at the corresponding allele, causing the system to repair itself using HDR and the Cas9/gRNA gene drive as a template. Expression of the CRISPR construct in the inserted allele should subsequently lead to the Cas9 making a cut at the alternative allele, which should lead to the HDR repaired insertion of the construct within that region as well. The targeted chromosome should now contain both mutated alleles for a particular gene. This concept of creating a cassette containing Cas9 and gRNA to be properly inserted at the target region where the gRNA makes its cut has been deemed Mutagenic Chain Reaction, because like the Polymerase Chain Reaction (PCR), it can lead to the copy and amplification of DNA, except in this case it involving trans-acting autocatalytic mutagenesis (Figure 6). (Gantz and Bier, 2016). This

technology might be the answer for the challenges that breeding programs, crop yield and crop improvement have faced over the past 20 years. To evaluate the suitability of CRISPR technologies and thus the feasibility of Active Genetics in polyploid plant genomes (as many crops) we decided to use the *Arabidopsis* tetraploid plant as a platform for our studies. As it will be discussed in more detail below, *Arabidopsis* has many suitable features to facilitate this approach. In summary, the goal of this study is not only to find efficient methodologies for targeting polyploid plant genomes via CRISPR/Cas9 but also to start investigating the feasibility of active genetics in polyploid plant genomes.

Arabidopsis as a reference plant model

Arabidopsis is a small, flowering plant belonging to the Brassicaceae family, which consists of mustard-related crops such as canola, broccoli, and cabbage. *Arabidopsis* is considered a model organism for research because it is easy to grow, and it has a short lifespan of 5-6 weeks, which allows for faster breeding (Meyerowitz, 1989). Additionally, *Arabidopsis* has a small genome size of approximately 132 Mb, which is compacted into five chromosomes containing greater than 20,000 protein-coding genes, making this organism extremely advantageous for developmental and genetic research in plant biology. There is an abundance of information and data for *Arabidopsis*, which makes it an ideal model plant to utilize for research. *Arabidopsis* is extremely self-fertile, producing as many as tens of thousands of seeds per individual plant, which are small and can remain dormant for a while, which makes the processes of storage and dispersal easier (Woodward and Bartel 2018). Many of the genes that have been shown to play critical roles in *Arabidopsis* they also do so in crop species, such as tomatoes, strawberries, and soybean (Ku et al., 2002, Mouhu et al., 2009, Jung et al., 2012). This makes the study of *Arabidopsis* biology not

only important for basic research purposes but also for application in agriculture in order to improve crop yield. In fact, homologs of important genes required for fruit dehiscence in *Arabidopsis* (*IND, ALC, and SHP*; to be discussed in more detail in later sections) were identified in *Brassica napus*, an important crop from which canola oil is produced. When these genes were knocked out it was found that that they could contribute to indehiscence of the fruit, which ultimately led to the solving the pod shattering problem that *Brassica napus* faces annually (Ferrándiz 2000; Liljegren et al. 2000, 2004; Rajani and Sundaresan 2001; Groszmann et al. 2011). More recently, *Arabidopsis* is favored in transgenic studies using CRISPR/Cas9 because of its simple transformation methods based on flower dipping using *Agrobacterium tumefaciens* (Li et al., 2013, Clough and Bent 1998). This process allows for the insertion of the T-DNA, a specific portion of the Ti-plasmid in *A. tumefiaciens*, which gets incorporated randomly into the plant genome. The T-DNA portion can be engineered to contain the CRISPR/Cas9 machinery, which can then create the desired cut within the plant genome, generating mutant transgenic progeny in the next generation (Yajko and Hegeman 1971, Alonso et al. 2003; Gao and Zhao 2014). Because the CRISPR/Cas9 technology has been implemented successfully in *Arabidopsis*, we want to take advantage of that and take advantage of the polyploid strains in *Arabidopsis*.

Impact of *Arabidopsis* fruit research

One of the morphogenetic programs that has a direct impact upon crop production is fruit development. In *Arabidopsis*, the fruit derives from the gynoecium and is constituted by the stigma, style and basal ovary with emerging ovules. After fertilization, the fruit begins to grow and, on its surface, three unique tissue types can be distinguished: the valves, the replum and the valve margins which, later will mediate fruit dehiscence and seed dispersal (Ferrándiz et al., 1999; Ripoll et al., 2011). There is a set of genes that are required for the formation of the valve margin,

known as the valve margin identity genes that are expressed at the valve and replum boundary (Liljegren et al., 2000; Rajani and Sundaresan, 2001; Liljegren et al., 2004; Roeder et al., 2003) (Figure 7). The MADS-box transcription factors *SHATTERPROOF1* (*SHP1*) and *SHATTERPROOF* 2 (*SHP2*) are involved in valve margin formation and *Arabidopsis* fruit dehiscence. (Liljegren et al., 2000; Ripoll et al., 2011), and thus when mutated *shp1 and shp2 mutant* fruit fail to differentiate the valve margin and become indehiscent (Liljegren et al, 2000). *SHP1* and *SHP2* positively regulate the expression of two additional regulatory genes, *INDEHISCENT* (*IND*), and *ALCATRAZ* (*ALC*), which encode bHLH transcription factors also required for fruit valve margin formation (Liljegren et al, 2000; Liljegren et al, 2004; Rajani and Sundaresan, 2001). *IND* is significantly important for the differentiation of the lignified and separation layers of the valve margin as *ind* mutants lack the differentiation of the valve margin layers (Liljegren et al., 2004). *ALC* is only involved with the cellular differentiation of the separation layers of the valve margin (Rajani and Sundaresan, 2001).

Another MADS-box gene, *FRUITFUL (FUL)* is required for valve differentiation and post-fertilization fruit growth (Gu et al., 1998; Ripoll et al., 2015). The homeobox gene *REPLUMLESS* (*RPL*) is required for replum development. These genes confine the expression of the valve margin genes *SHP*, *IND*, and *ALC*, ensuring that the expression of the valve margin identity genes are only present at the valve-replum boundaries (Roeder et al., 2003; Liljegren et al., 2004). Thus, when *rpl* is mutated the replum adopt valve margin identity (Roeder et al., 2003). Similarly, *ful* mutants valves fail to differentiate due to the ectopic expression of the valve margin identity genes in the valves (Ferrándiz et al, 2000a; Liljegren et al., 2004). In line with this, in *rpl ful* double mutants, cells that would normally develop into the valve and replum develop into the valve margin cell instead (Roeder et al., 2003; Liljegren et al., 2004; Ripoll et al., 2011). Along with RPL, the

KNOX-family gene *BREVIPEDICELLUS* (*BP*) is required for replum formation. During vegetative development, *BP* is negatively controlled by the MYB transcription factor *ASYMMETRIC LEAVES1* (*AS1*) and the *LATERAL ORGAN BOUNDARY (LOB)* domain protein *ASYMMETRIC LEAVES2* (*AS2*) (Byrne et al., 2000; Byrne et al., 2002; Guo et al., 2008; Ori et al., 2000; Semiarti et al., 2001) to allow correct meristematic development and allow leaf differentiation. Interestingly, this regulation also occurs in the fruit, where both *AS1* and *AS2*, active in the valves, and in collaboration with the AP2/EREBP family transcription factor APETALA2 (AP2) repress BP expression to circumscribe it to the replum tissues (Alonso-Cantabrana et al., 2007, Ripoll et al., 2011). Thus, *ap2* or *as* mutants have abnormally larger repla and valve margins (Ripoll et al., 2011).

The final steps of fruit development are fruit maturation, senescence and fruit dehiscence (which occurs when the valves separate from the replum) followed by seed dispersal, (Ferrándiz et al., 2000b, Ripoll et al., 2015, Ripoll et al., 2019). This research in studying fruit development following classical approaches done in *Arabidopsis* contributed to the improvement of canola oil, a globally important crop.

Solving the pod-shattering mystery in canola using *Arabidopsis* fruit as reference

There are many examples in which basic research upon reference plant species were vital to discovering important methodologies and techniques to apply in crops. For example, the research performed by the Yanofsky Lab led to the elaboration of a regulatory framework model for *Arabidopsis* fruit development and in particular for valve margin formation. These investigations provided with a number of candidate genes to investigate fruit related processes, such as dehiscence (aka pod shattering), and served to help improve crop yield in species in which pre-harvest pod shattering is a concern. In *Arabidopsis*, pod-shattering consists of the silique

opening and releasing seeds with the application of very little mechanical force (Dinneny and Yanofsky, 2005). Although dehiscence is important for seed dispersal and continuous crop growth, it is also a contributing factor to crop yield loss. Decreasing a crops' susceptibility to pod-shattering is critical for increasing yield supply (Ogutcen et al., 2018).

Arabidopsis is closely related to *Brassica napus* (or canola), which is responsible for producing canola oil. This crop suffers devastating pod-shattering problems where up to 50 percent of its yield can be lost annually due to poor weather conditions (Child et al, 1998; MacLeod 1981). Canola is the second largest globally used crop and is the source of oil and protein for various foods (Raymer 2002), and it is one of many crops that experience yield loss due to premature dehiscence and pod-shattering.

As mentioned previously, the Yanofsky Lab discovered a suite of genes (*ALC*, *SHP*, and *IND*) critical for the formation of valve margin and thus fruits dehiscence (pod shattering) in *Arabidopsis* (Ferrándiz 2000; Liljegren et al. 2000, 2004; Rajani and Sundaresan 2001; Groszmann et al. 2011) and interestingly, these genes have their counterparts (orthologs) in canola. Two of the *IND* genes, *IND-A1* and *IND-C1*, were then successfully mutated in canola to engineer a modified canola crop species that had improved resistance to pod shattering (Laga et al., 2014). (Figure 8).

Due to the research done in Arabidopsis fruit (dehiscence) the Yanofsky lab was able to translate that knowledge for improving crops, such as developing pod-shattering prevention techniques that were later implemented in order to increase canola oil yield. The main issue with classical approaches is time, because this process took over twenty years from its discovery to the its initial field trials. Therefore, we need to identify strategies to not only advance basic research, but also expedite crop improvement and accelerate the transfer of this basic knowledge to

translational endeavors in agriculture to meet current and future challenges. Now, CRISPR/Cas9 technologies have opened the door to put these translational strategies on the fast track.

The study of fruit morphogenesis was critical to solving the pod-shattering problem in canola. There are also additional molecules that are involved in fruit morphogenesis that our lab has investigated and utilized in our research. In particular, we have studied microRNAs and their role in fruit development. In order to optimize our CRISPR strategy in polyploids, we were selective of the genes chosen to target as we wanted to utilize the resources that would generate robust results, and for this reason one of the genes we chose to target was *microRNA160*, a gene that will be discussed in more detail in the section below.

The importance of MicroRNAs – Key players during fruit morphogenesis

Although many non-coding RNAs have long been considered "molecular garbage", a good amount of these have been found to have important roles in regulating gene expression in critical developmental programs. MicroRNAs (miRNAs) are small, non-coding RNA molecules (approximately 20-24 nucleotides long) that are involved in post-transcriptional regulation processes. These single-stranded molecules are formed from double-stranded hairpin precursors, the pre-miRNA. A dicer enzyme then breaks apart this double-stranded structure into shorter pieces (Rhoades et al, 2006). One of the strands of each piece gets degraded, leaving the single-strand miRNA to associate and form a complex with other proteins that can bind to a complementary mRNA sequence. Once bound to the mRNA, this complex can either degrade mRNA or block its translation in order to silence gene expression (Bartel 2004).

Unlike animal miRNAs, which can be found within intron sequences, plant miRNAs are generally located in non-protein coding regions of the genome and are sometimes grouped together, indicating that multiple miRNAs are likely to be transcribed from a single primary transcript. Additionally, plant miRNAs are conserved throughout evolution and also likely to be transcribed by RNA Polymerase II, although the true transcriptional mechanisms are still poorly understood (Rhoades et al, 2006, Reinhart et al., 2002). miRNAs have been found to be involved in many plant development processes such as embryonic development, leaf development, stem cell differentiation, or responding to environmental abiotic and biotic stress (Carrington and Ambros, 2003; Liu et al., 2011; Sun 2012). Additionally, miRNAs have been found to be important for breeding and agricultural research because of their roles in various crop species such as rice, wheat, peanuts, and potatoes (Djami-Tchatchou et al., 2017). Currently, there are 26 distinct miRNAs found in *Arabidopsis* (Maher et al., 2006) and this number is likely to continue increasing.

Whereas the role of miRNAs was discovered for a number of developmental programs, their role in fruit development was not described. The Yanofsky Lab demonstrated that one of the roles of miR172 is fruit morphogenesis. In fruit, miR172 represses AP2 in the valves, and restricts its activity to the the valve margin and replum, which allows fruit valve elongation after fertilization. When *miR172* activity was blocked, fruit growth is subsequently blocked, confirming the critical role of miR172 in fruit growth and development (Ripoll et al, 2015). Interestingly, one of the miR172-encoding genes, The *MIR172C* is expressed in valves and activated by the combined actions of the MADS-box *FUL* and the ARF transcription factors ARF6 and ARF8. This regulation allows correct miR172 functions and appropriate fruit valve growth (Ripoll et al, 2015).

At this point our lab wanted to investigate if there were additional microRNAs involved in fruit development. A previous Master's student identified that miRNA160 played an important role during fruit and leaf development. Previous studies have shown that miRNA160 is involved in post-transcriptionally repressing the ARF transcription factors *ARF10, ARF16* and *ARF17* (Rhoades et al, 2006). Additionally, previous data from a former Master student in the Yanofsky lab revealed that miRNA160 works together with *AS1* and *JAG* in order to regulate fruit morphogenesis. In particular, miRNA160 represses *ARF10*, which, in turn, negatively regulates *AS1*, *JAG*, and *FIL*, essential for correct carpel and valve formation via the repression of the replum identity genes (Bailey, L.J. 2012). However, this study could not be move forward because we lacked mutants for miRNA160-encoding genes. In order to continue the project and overcome this challenge, a senior research scientist in the Yanofsky Lab generated a CRISPR/Cas9 tool-kit and was able to knockout *MIRNA160A* in *Arabidopsis* diploid (2n) plants, which produced a serrated leaf phenotype and shorter fruit with smaller valves and big repla, similar to those when *AS1*, *JAG* or *FIL* functions are compromised.

The fact that the CRISPR/Cas9 tools worked to knockout miRNA160 in diploid *Arabidopsis*, so we decided to leverage these tools in order to see if we could edit miRNA160 in a polyploid *Arabidopsis* strain. We wanted to take advantage of the conspicuous phenotypes that *miRNA160* mutants have in order to evaluate our system in a 4n *Arabidopsis* strain. Additionally, we can also see whether in a polyploid miRNA160 plays a similar role as it does in the diploid. We were able to use this gRNA in order to establish consistent transgenic lines in a 2n plant. For this reason, we wanted to use the gRNA targeting *microRNA160* (referred to as gRNA∝160) in tetraploid Arabidopsis. In this thesis we have investigated different ways for targeting polyploid genomes via CRISPR using *Arabidopsis* 4n tetraploid as a platform for our experiments.

Additionally, we used a gRNA (referred to as gRNA∝AS1) targeting the middle of the *AS1* gene in order to make a null allele and see a distinct phenotype affecting the leaf and fruit development of diploid and tetraploid *Arabidopsis.* Additionally, mutating *AS1* produces a phenotype that can be detected early on in development. We did this in order to continue to develop strategies for polyploid *Arabidopsis* and to compare the gene editing efficiency using two different guides.

For my thesis, we decided to investigate whether the CRISPR/Cas9 tools we developed for 2n *Arabidopsis* could also be employed in tetraploid (4n) *Arabidopsis*. However, getting transgenic plants using polyploids has proven to be quite challenging. Even though it is difficult, it is still possible to use CRISPR technology in order to generate mutations and transgenic plants. Ultimately, the purpose of this study is to develop an effective CRISPR/Cas9 strategy for targeting polyploid genomes with the aim to transfer these methodologies towards advancements in crop and agricultural biotechnology. We also envision that, if successful, we could employ these tools for developing Active Genetic methodology in polyploids and expedite crop breeding programs improve yield. Thus, our research provides one of the first foundations for not only devising a system for efficient editing in polyploids but also providing a platform for demonstrating the viability of Active Genetics in polyploid genomes.

MATERIALS AND METHODS

Plant materials and growth conditions

For this study we have employed diploid and tetraploid Arabidopsis wild-type 2n and 4n strains in the Columbia-0 (Col-0) accession. The original CRISPR mutant allele for *MIR160A* was previously isolated in our lab in the Col-0 background. For seed germination, seeds were sown directly on soil or on plates containing Murashige-Skoog media supplemented with 5% sucrose and the corresponding antibiotic (see below). The *DD45::Cas9* (aka p3J1) transgenic driver line was generated by a member of the Yanofsky lab.

Cloning strategies, and bacteria transformation

We cloned the gRNA expressing modules (containing either 1 or 2 gRNA expressing units) via Gibson DNA assembly into the p3J1 T-DNA binary vector (pGreen derivative) using the SpeI unique site or into pGreenII0229 using the same restriction site. For targeting *MIR160A* (AT2G39175) we generated a 2xgRNA cassette using the following primer combinations oJJRR961+oJJR1001; oJJR1002+oJJR1003; oJJR1004+oJJR967. For generating the gRNA. Cassette to target *AS1* (AT2G37630) we used the primer combinations: oJJR961+oJJR1614 and oJJR1615+oJJR967 (see sequences in Table 1 in Appendix B). These sets of primers were used for PCR reactions employing Phusion proof-reading Taq-polymerase (NEB) and the plasmid pJJJ14 as template for the reactions. The PCR amplicons were run on an agarose gel and gel isolated using the QIAGEN gel isolation kit following manufactures indications (Qiagen Inc.). For cloning in the p3J1 vector we used Gibson DNA assembly protocol. We then transform *E. coli* and

detect the positive clones by colony PCR. Positive colonies were grown in liquid LB and plasmid was isolated using Promega plasmid prep isolation kit (Promega Inc.).

PCR amplicons:

For mutant screening in our plant population, we used Choice-Taq (Thomas Scientific) and the PCR amplicons were visualized on an agarose gel. The products were done gel extracted and isolated using Qia-gen gel extraction kit following manufacture guidelines.

Sequencing:

For sequencing our plasmids and amplicons we used a local company (Retrogen Inc.) following their guidelines for their sample preparation).

Generating and isolating transgenic plants:

Our resulting pGreen-based constructs were co-transformed into *Agrobacterium tumefaciens* (AGL-0) with the helper plasmid pSOUP (Hellens et al., 2000) via electroporation. We used the floral dipping method (Clough and Bent, 1998) to transform constructs into the corresponding strain. T1 transgenic plants were sown on MS plates containing 20mg/mL Hygromycin and resistant seedlings were then transplanted onto soil after a 1-2 weeks. For our split system, Cas9 driver lines was transformed with the pGreenII0229 containing the 2xgRNA cassette targeting MIR160A and the resulting T1 seeds were screened by spraying with BASTA

(120mg/mL ammonium sulfate; Finale, AgrEvo, Montvale, NJ) two to three times a week for 2 weeks.

RESULTS

The use of the CRISPR/Cas9 system allows for targeted modifications in the plant DNA to achieve a desired genotype and phenotype. The T-DNA is delivered randomly into the plant genome by *Agrobacterium tumefaciens* via floral dipping (Lee et al., 2019; Bent 2006). The T-DNA then expresses a gRNA that is specific to a locus within the genome, which allows the Cas9 to make the targeted double-stranded breaks. In diploids the desired mutation can be in one allele, both alleles, or neither allele. In tetraploids this has not been studied yet in detail.

We know from previous studies that our lab has done that small RNAs play a significant role in fruit morphogenesis. We then decided to focus our research plan on fruit-related genes and their role in fruit morphogenesis. In particular, our experimental plan will be focused on two genes: *microRNA160* (*miRNA160*) and *AS1*. As mentioned earlier, one of the previous Master's students in our lab has shown that *miRNA160* is involved in fruit development. Additionally, we chose *miRNA160* because its phenotype is prominent not only in the fruits but also the leaves, showing a serrated leaf phenotype as compared to the wild-type col-0, which has rounder leaves. We also know that the guide RNA that targets *miRNA160* is effective and provides transgenic edits in the diploids. Previous work in our lab has been done by another Master's student to target this gene using CRISPR/Cas9, proving that the guide RNA designed to target *miRNA160* to be efficient in generating robust edits. We chose *AS1* because it has a wealth of information about its role in fruit morphogenesis and it also has an evident phenotype during leaf and fruit development.

Our overall goal was to create an effective gene targeting strategy using CRISPR/Cas9 that could be applied towards crops. We decided to first carry out these experiments in diploid *Arabidopsis* as a control because we knew that targeting and creating homozygous edits was efficient in 2n. Many of the the most important crop species are polyploid, and one of the

challenges of creating edits in polyploids is to be able to develop homozygous mutants so that the trait can then be passed to all subsequent offspring (Passricha et al., 2016). Consequently, generating targeted edits in polyploids is known to be more difficult than in diploids since there are more alleles therefore more possibilities for bi-allelic edits. Due to these challenges, there is no specific strategy for generating targeted mutations. Therefore, our goal was to propose an initial mechanism for generating edits in polyploid plants. We decided to take advantage of 4n *Arabidopsis* as our main model instead of crops because it is easier to analyze and transform so that it can be easier to develop the framework for the mechanism that can effectively target polyploid genomes. We additionally want to emphasize that any result in the 4ns is novel because there haven't been many successful edits done in polyploids as previous studies have generated bi-allelic edits. After generating a strategy to efficiently generate edits in 4n plants, we wanted to apply this knowledge to do Active Genetics and knock in strategies in the 4n as well. We ultimately hope that the use of Active Genetics in our optimized strategy will demonstrate enhanced plant characteristics which can then be carried out to crop species.

I. Targeting *MIRNA160A* in *Arabidopsis* diploid employing an *in cis* CRISPR/Cas9 configuration

One of the major goals of this study is to elaborate an efficient and reliable strategy to efficiently perform CRISPR/Cas9 genome editing in both diploid (2n) and tetraploid (4n) *Arabidopsis* plants. To further expand our knowledge on how miRNAs impact fruit development, a former Master's student in the Yanofsky Lab showed that the gRNA∝*miR160A* (a gRNA

targeting the *MIR160A* gene) is effective in combination with Cas9 expressed from the UBQ10 constitutive promoter (Tran, Q.H., 2019). However, this strategy, although successful, had some drawbacks including the appearance of mosaic individuals or the isolation of homo-allelic individuals with low percentage. Thus, we wanted to continue improving our CRISPR/Cas9 toolkit by assaying additional promoters and identifying promoters proved to be the most effective for successfully generating edits. We found that the egg-cell promoter DD45 would also be successful in diploids. Previous research has shown that this promoter has a high efficiency in creating stable homo-allelic transgenic plants (Wang et al., 2015). We therefore generated a T-DNA construct containing a Cas9 module in which the Cas9 is expressed from the DD45 promoter together and a 2 gRNA (2xgRNAα160) expressing cassette that would result in the generation of deletion alleles (Figure 9c). Note that based on our unpublished data, 1xgRNA cassettes leads to a more diverse population of alleles when comparing to a 2xgRNA strategy for targeting the same gene. Additionally, we hypothesized that the two guide RNA strategy challenges the repair machinery of the cell, resulting in a larger segment of mutated DNA, which can generate a more reliable knockout.

We used our construct and transformed wild-type col-0 plants and selected for transgenic plants on Hygromycin plates (Figure 10a). From the T1 population we identified a high number of homo-allelic mutants amongst 56 lines, demonstrating the efficiency of using the DD45 promoter along with 2xgRNAα160 to create edits in 2n plants. We were able to identify the mutants by initially observing the phenotype of the transgenic T1s, which had noticeable differences in leaf development as compared to Col-0 (Figure 9b). We then checked by PCR followed by sequencing the integrity of the *MIR160A* gene (Figure 10a), and confirmed that there were edits. Interestingly, the sequencing analysis of the samples revealed that most of the T1s

sampled were homo-allelic mutations in which a large portion of the targeted *miR160A* locus was deleted. The remaining alleles were small insertions and deletions of portions of the targeted locus flanked by the 2 gRNAs (Figure 9a). This data suggests that this all-in-one strategy is effective and capable of generating edits in *miRNA160* when using *Arabidopsis* 2n. We also realized that the efficiency of the Cas9 is likely dependent on the kind of promoter it is expressed by. Therefore, we wanted to choose the right promoter to continue developing our CRISPR/Cas9 strategies in the 4n.

II.Comparing UBQ10 2n vs 4n (CRISPR/Cas9 editing in 2n vs 4n).

At this point, we wanted to assay if the promoters that we successfully used to generate transgenic edits in Arabidopsis 2n were also capable of generating edits in tetraploid *Arabidopsis* 4n. For example, we were able to generate transgenic edits when using the constitutive UBQ10 promoter to express Cas9 in 2n plants, so we then wanted to see if UBQ10 promoter would be useful to generate transgenic edits in 4n *Arabidopsis* plants. However, we isolated more than 31 T1 lines and unfortunately, no mutant phenotype or edits were found. We then concluded that it is very likely that a constitutive promoter is not the best strategy for targeting *Arabidopsis* 4n genome. We found that DD45 was more efficient than UBQ10 in generating mutant alleles in 2n plants as we identified via sequencing more homo-allelic plants in the T1 generation (Figure 9a). Therefore, we decided to carry out the remainder of our experiments using a T-DNA containing DD45::Cas9.

III. Targeting *MIRNA160A* in *Arabidopsis* tetraploid employing an *in cis* CRISPR/Cas9 configuration

Knowing that the DD45 was more effective at generating edits, we then transformed 4n *Arabidopsis* with the DD45 promoter with the 2xgRNA cassette. We were able to identify 2 transgenic plants from this in cis method (all-in-one). However, the sequencing results of these two lines showed that these plants were wild-type for the deletion that was present in the original 2n T4 C92#8 homozygous (HOM) line (Figure 12c). This data suggests that the transgene is still present in the genome but Cas9 did not produce an edit at the targeted region. Despite the initial numbers being low, possibly due to the age of plants used for transformation, we can still conclude that the all-in-one strategy in which the gRNA and Cas9 were present in the same T-DNA to be not as not as effective of generating edits in miRNA160 in the 4n as compared to the 2n. In the future, we hope to repeat this transformation with younger plants in order to see if we could generate a greater amount of transgenic plants that also could carry the deletion.

IV. Creating Cas9 driver lines and targeting *MIRNA160A* in *Arabidopsis* diploid employing an *in trans* (split-system) CRISPR/Cas9 configuration

To further expand the versatility of our CRISPR editing system, we decided to explore the possibility of splitting our system into two elements. On one hand, we could create Arabidopsis transgenic lines harboring Cas9 constructs in which Cas9 expression is under control of different promoters (Cas9 driver lines). We wanted to ideally create homozygous Cas9 driver lines in 4n since we know that getting few transgenics in polyploids is common and expected. On the other

hand, we could generate T-DNA constructs containing our desired gRNA cassette that could be then either transformed into our Cas9 driver lines, or into wild-type and then cross these gRNA lines to the Cas9 driver lines.

Our lab has generated a large collection of Cas9 transgenic lines in which Cas9 is under the control of different promoters including DD45. These lines were selected using Hygromycin antibiotic and checked for the transgene using PCR. This method allowed us to be more versatile with the gRNAs that we wanted to utilize for making the transgenic edits. To evaluate this strategy, we first targeted MIRNA160 in diploid *Arabidopsis*. We transformed homozygous lines for *DD45::Cas9* with a T-DNA containing a 2xgRNA∝ *miR160,* and checked for the presence of the gRNA using BASTA and the subsequent edits in *miR160* using PCR. Figure 2b and 2c depict the general schematic of the split system described above. Interestingly, all the T1s identified showed the phenotype associated with the lack of miR160 activity and thus presumptive to contain mutations (i.e edits) within the *MIR160* coding region (data not shown). This contrasts with our previous observation just using the all-in-one methods described above. Because these systems were so successful in the diploids, we wanted to see which system, the all-in-one or the split system, was going to be more effective in the 4n Arabidopsis strain.

IV. Creating Cas9 driver lines in tetraploid *Arabidopsis*

Our experiments in 2n *Arabidopsis* and the preliminary data in *Arabidopsis* tetraploid, support DD45 as the promoter of choice for efficiently creating edits. Moreover, our data also shows that transforming a T-DNA harboring a gRNA expressing cassette into a Cas9 driver line (*DD45::Cas9*) leads to efficient editing and creating a high percentage of homo-allelic lines. We

therefore decided to develop a split strategy in Arabidopsis 4n (Figure 10b, 10c). We first had to generate 4n Cas9 driver lines. To do so, we transformed our T-DNA containing the DD45::Cas9 (aka p3J1) into Col-0 4n *Arabidopsis* plants. As expected, T1 plants showed wild-type phenotype consistent with the absence of a gRNA in our system. We verified the presence and integrity of the Cas9 transgene by PCR. We selected 5 lines for further analysis (T1-1, 3, 4, 5, 6).

VI. Transformation of the Cas9 4n driver line with a gRNA construct

We generated this split system in the 2n using the DD45 egg-cell promoter to drive the Cas9 and found it to be effective. The knowledge of this successful data pioneered our drive to produce a generalized strategy for the whole system. We then decided to see if we could achieve an effective strategy for targeting the 4n genome by splitting the CRISPR/Cas9 system (Figure 10b, 10c). We transformed Cas9 driver lines with DD45::Cas9 with a T-DNA containing a 2xgRNA∝*MIR160A,* similarly as we had done in the 2n. We selected individuals that contained the gRNA using BASTA and tried to see first if we could identify potential edits by observing the phenotype. We identified T1s that likely had edits because they showed the serrated leaf phenotype as compared to wild-type col-0 (Figure 11a, 11b). Additionally, sequencing results confirmed that edits were made in these T1s. More specifically, T1-1 p3J1 (e) w/ 2xgRNA has sequencing that are consistent with this line being homo-allelic. However, in order to confirm and verify this we need to analyze the next generation and sequence it to see that every individual from that line shows the same mutation. T1-2 and T1-3 p3J1 w/ 2xgRNA also show a deletion and insertion respectively that are likely to be hetero-allelic (Figure 12b). We would also have to confirm this by observing the segregation of alleles in the next generation. These results appear promising and

demonstrate that the split system is more efficient and capable of generating edits in polyploid genomes as compared to the all-in-one system.

VII. Using CRISPR/Cas9 to target *AS1*

The set of experiments described above evaluated the possibility of editing 2n and 4n *Arabidopsis* genome using our collection of CRISPR tools for targeting the MIRNA160A gene. However, we wanted to further investigate if these tools are capable to target additional genes. As mentioned in the introduction, the gene *AS1* is required for correct leaf and fruit morphogenesis and when mutated, leaves and fruit show abnormal phenotypes (see figure 14c and references). Interestingly, a former Masters student in our lab discovered that *AS1* and miR160 functions act together during fruit patterning (Bailey, L.J., 2012). In this context we decided to design a gRNA and elaborate the corresponding constructs to target the gene *AS1* via CRISPR in 2n and 4n *Arabidopsis* strains. Following the approach used for MIRNA160, we decided to first target *AS1* in 2n Arabidopsis and generate a construct in which Cas9 was under the DD45 promoter control. We decided to use a single gRNA located in the middle of the *AS1* gene (Figure 13 and 14a). To evaluate the efficiency of the *gRNA∝AS1* we transformed Col-0 diploids and identified T1 individuals with the associated *as1* mutant phenotype. We next used the primers (oJJR56 and oJJR57) to genotype by PCR and sequencing the T1s that showed the desired phenotype as compared to wild-type colp0 (Figure 13, 14b, 14c). We confirmed the presence of mutations in such plants. Our sequencing results identified several individuals with different kinds of edits (Figure 14a). We could then conclude that this gRNA can efficiently recognize the *AS1* locus and generate edits. Additionally, we observed wild-type looking AS1 lines that were resistant to

BASTA and that upon sequencing confirmed that no edits were made at the targeted locus. However, when these lines were sown into the T2 generation, we were able to observe the mutant phenotype. By sequencing we saw edits that confirmed the presence of edits (Figure 14a). We then wanted to use this additional gRNA construct in order to target the *AS1* gene in *Arabidopsis* tetraploids using the all-in-one system as well as using our split system, similarly as we did for *MIRNA160*. Unfortunately, and the current situation caused by the COVID-19 pandemic limited our ability to continue with these experiments and therefore this would be something to re-visit in the future. We were not able to finalize these assays. However, this goes beyond the scope of the current study. Eventually we will also transform with the corresponding constructs and evaluate whether the split system is as efficient to use in *AS1*. Not only that, but knowing this gRNA is efficient, we can design a construct using this gRNA to do a knock-in via CRISPR/Cas9. However, the pandemic has limited us in these efforts, and this is also something we will have to re-visit again in the future.

In summary, we ruled out that using the UBQ10 promoter to drive Cas9 in our CRISPR/Cas9 strategies in the 4n *Arabidopsis* because we could not identify any T1s with edits. We assayed additional promoters and found that DD45 was efficient not only in the 2n but in the 4n as well. We then continued to develop the rest of our experimental tools by using a T-DNA containing DD45::Cas9. We also found that the all-in-one system is effective in the 2n because we did see edits, however our current data demonstrates that it is not effective in the 4n as we did not generate any edits. We could then conclude that it is likely the split-system rather than the all-in-one that is more effective to use in the 4n. We also can conclude that the gRNA∝*160A* is also effective in both the 2n and 4n when the Cas9 is expressed from the DD45 promoter. Lastly, we

could also conclude that the gRNA∝*AS1* is also effective at generating edits and therefore can serve as another tool for further developing our CRISPR tools. We believe that we have identified tools and strategies for targeting 2n and 4n genomes that we can also utilize for doing a knock-in (KI) in Arabidopsis as well as provide a platform to do Active Genetics in polyploids.

DISCUSSION

The current world is facing an unprecedented climate crisis and population changes that represents a continuous pressure upon food production and distribution. Genome modification technologies such as zinc finger nucleases (ZFNs), transcription activator-like effector molecules (TALENs), mutation breeding and marker assisted selection (MAS) have been used in the past to assist in crop breeding programs. However, in our changing world it is necessary to find new strategies to overcome the challenges that will soon affect agriculture. CRISPR/Cas9 is an effective gene editing tool adopted from the bacterial immune system response to foreign pathogens for eukaryote genome engineering. This system uses the Cas9 protein and a guide RNA in order to induce targeted double-stranded breaks in the genome, resulting in the formation of insertion or deletion mutations induced by the host cell's NHEJ repair pathway (Farasat and Salis, 2016; Sternberg et al., 2014; Rath et al., 2015) and has been successfully used in animals, insects, and plants (Reardon 2016; Feng et al., 2013).

Many crops such as potato, wheat, apple, and canola are polyploid (Weeks 2017), and whereas CRISPR has been reported to be an effective tool for generating edits in diploid plant genomes, how this system can be used to precisely target polyploid genomes is still at its infancy. We wanted to design our experimental plan using tetraploid (4n) *Arabidopsis* as the reference model. Therefore, our goal was to evaluate different approaches and elaborate a general method for introducing targeted mutations polyploid genomes. We therefore decided to employ a tetraploid (4n) strain of the plant reference Arabidopsis as platform for our studies. We expect to provide the principal foundation for future applications towards gene editing in polyploid crop species.

The Yanofsky lab has devoted numerous efforts at studying the roles of post-transcriptional regulators including microRNAs in fruit development. Previous Master's students had shown that miR160 played a crucial role in fruit morphogenesis and that it could be effectively targeted using a CRISPR/Cas9 toolkit that was developed by a senior scientist in our lab. The toolkit employed two guide RNAs for generating a complete knockout of the *MIR160* gene. Because we know this construct was efficient, we chose to develop our CRISPR/Cas9 strategy in polyploid *Arabidopsis* by first targeting *MIR160A* in our experimental approach.

In order to further test our CRISPR/Cas9 strategy and provide additional supporting evidence of its efficiency, we also chose an additional gRNA to target the *AS1* gene, which has been previously studied in the Yanofsky lab for its role in *Arabidopsis* fruit development. We planned to also use this construct to compare its efficiency in 2n and 4n *Arabidopsis* strains. However, due to the COVID-19 pandemic, we were unable to pursue these experiments, so this is something we would have to revisit in the future. We hope that targeting two loci and determining whether we can achieve similar results in both 2n and 4n can serve as evidence that the system we had developed is efficient and can therefore be transferred for editing other polyploid plant genomes including crops.

We employed two systems that lead us to the development of an effective gene editing strategy in the 2n that we then optimized in the 4n *Arabidopsis*. We focused on comparing the all-in-one system (the T-DNA used for plant transformation containing both the Cas9 and gRNA was delivered into *Arabidopsis* 2n and/or 4n) to the split system (T-DNA containing a gRNA is transformed into a 2n or 4n Cas9 driver transgenic line respectively). Our results in 4n strongly suggest that the split-system is more efficient as we were able to detect mutant individuals and edits in the T1; whereas the all-in-one system resulted in T1s that had the T-DNA but were wild-

type looking and no edits within the gene of interest (Figure 12c). It is very likely that the difference in efficiency between all-in-one vs. split in 4n has to do to with the Cas9 levels present in the ovules. For our studies we controlled the expression of Cas9 using the *DD45* promoter, which has been previously reported to be highly active in ovules (Steffen et al., 2007; Ye et al., 1999) (tissue targeted by *Agrobacterium*). It is conceivable that the levels of Cas9 in the ovules are higher in the driver lines derived from our split approach than in those from all in one, as the Cas9 transgene is a present longer in the former than in the latter and makes enough Cas9 because the promoter driving Cas9 is expressed at higher levels earlier in ovule development. It is likely that there is a shift in levels of promoter activity as the T-DNA is being introduced into the plant genome. Our data also likely suggests that while it is possible that using the all-in-one method can generate edits in the 2n, the efficiency is null in 4n *Arabidopsis* because because there are more alleles that need to be targeted. Although our observations are consistent with these conclusions, additional experiments will be required to further support them. For example, we can harvest the seeds from the transgenic T1s and sow the next generation (T2) in order to see if we can identify any edits. Additionally, if we do see edits we can check to see if they are homo-allelic by tracking the segregation of alleles in the following generation (T3). This approach is similar to what we saw in the T2 *as1* mutant population which were likely homo-allelic for the edits but were wild-type in the T1 generation (Figure 14a). The data suggests that this could be the case because Cas9 is already present in the plant and therefore should have higher levels of expression. Additionally, this is similar to what we have seen when following the split-system approach, in which driver lines already expressing Cas9 are transformed with the gRNA. These lines are more likely to have edits than the all-in-one system in which Cas9 and the T-DNA are introduced simultaneously. Our sequencing results indicated that while using the split-system in the 4n we had identified one line

that is likely homo-allelic (Figure 12b), however additional experiments are required to further validate this data.

The experiments that we have developed in this study allowed us to evaluate certain elements of the system that show promising results for targeted gene editing in polyploid plant genomes (4n *Arabidopsis*). For example, we know that targeting *AS1* and *MIRNA160* using the *DD45* promoter is an effective strategy for generating mutant alleles in both 2n and 4n. Now that we have these fundamental elements established and we also know that it is possible to use this CRISPR system to target 4n genomes, we need to focus on increasing the of the system in the 4n. We additionally envisioned an alternative approach to the split system. In this alternative version, we would again establish Cas9 driver lines as well as gRNA containing lines that we can cross to each other. Establishing independent Cas9 and gRNA-containing lines can allow us to test the same gRNA line with different Cas9 transgenic strains just by crossing them to each other.

We know that because we were able to use CRISPR to target the 4n genome that we can then explore the possibility of doing Active Genetics in 4n. We wanted to find the right strategy to target the tetraploids and apply the knowledge of what we have gathered about gene drives in the 2n and 4n *Arabidopsis* in order to set up the foundation for developing Active Genetics in polyploid plants. The data we have generated from our initial studies shows promising evidence that Active Genetics can be done in polyploids. Due to the pandemic it was difficult to continue with many of the experiments we had set out to see through during the spring quarter. Therefore, additional work will be required to further support these claims. Nevertheless, we still believe that this primary data opens a new venue to explore Active Genetics and could represent a new revolution in plant breeding and it can expedite crop genome tailoring the challenges of our changing world.

In addition to initiating the steps for implementing CRISPR, we also decided to evaluate the feasibility of the tools we generated in our study to achieve knock-ins and implement Active Genetics in 4n *Arabidopsis*. We identified 2 gRNA sets (gRNA∝*miR160* and gRNA∝*AS1*) for generating and mobilizing mutations, or generating a knock-in gene drive for Active Genetics. In plants it is very challenging to achieve a knock-in (KI) because the prevalence of the NHEJ pathway over the homology-directed repair (HDR), in which the template is used to repair the double-stranded breaks (DSBs). Additionally, HDR has a number of limiting factors, which makes it an inefficient repair pathway to use for gene editing and therefore poses a challenge for KI strategies (Zhang et al., 2017).

There have also been alternative approaches that our lab is trying to take in order to increase the efficiency of generating knock-ins. One of the limiting factors to the method is the low transformation in the 4n. One of the promising techniques we have implemented is the use of nanotubes to deliver DNA or protein into a biological system. Previous research has shown that the use of carbon nanotubes to deliver exogenous DNA bypasses the necessity of implementing a transgene into the plant cells and can subsequently increase the transformation efficiency (Demirer et al., 2019). We hope to further explore such possibilities in the 4n as well.

We hope that the tools that we have generated in our lab can serve as an initial foundation for establishing both Active Genetics and KI strategies into *Arabidopsis* 4n. Our world is currently facing a myriad of global challenges induced by climate change and now the COVID-19 pandemic, which will no doubt impact food production and agriculture in the time to come. Therefore, we also hope that these tools can establish an initial foundation for further research in agricultural biotechnology in order to improve crop species and meet the needs of our changing world.

APPENDIX: FIGURES AND FIGURE LEGENDS

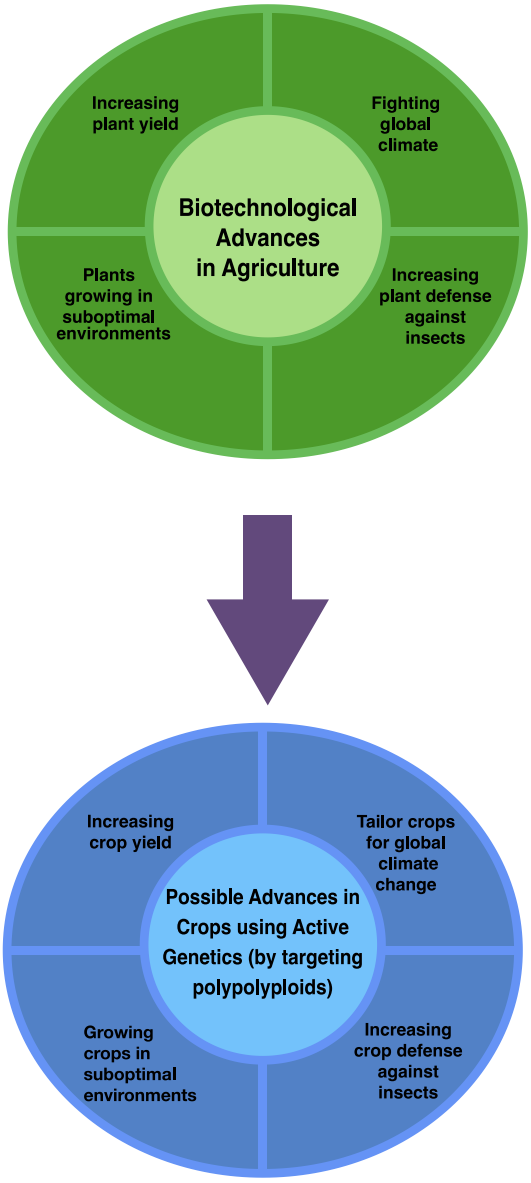

Figure 1: Biotechnological Advances in Agriculture and Application to Active Genetic Strategies

Schematic of biotechnology advances that are currently being research in plants and that can be further expedited using Active Genetics in order to be applied to crops.

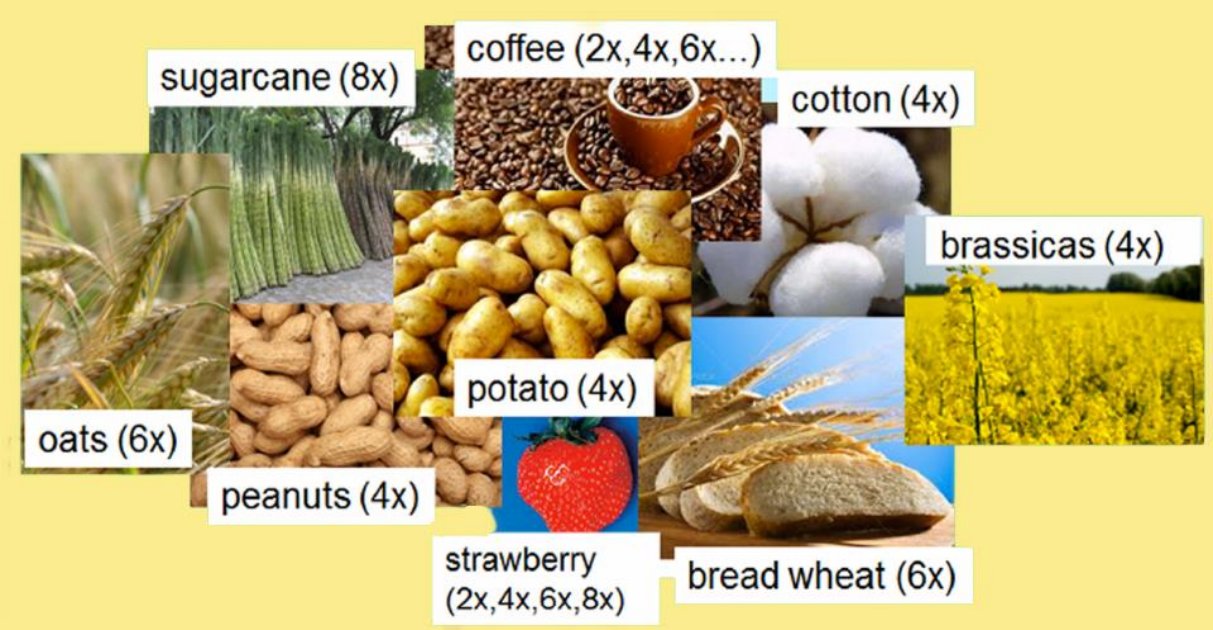

Figure 2: Examples of Crops with Polyploid Genomes
Many of the most common crop species are polyploid such as sugarcane, coffee, cotton, brassicas, oats, peanuts, strawberries, potatoes, and bread wheat.

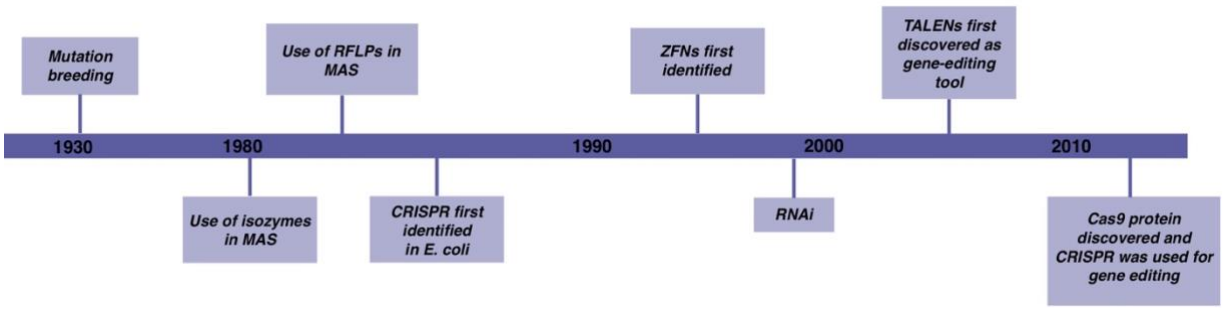

Figure 3: Timeline of Gene Editing Strategies

Timeline of various gene editing strategies that were used in editing plant genomes and breeding.

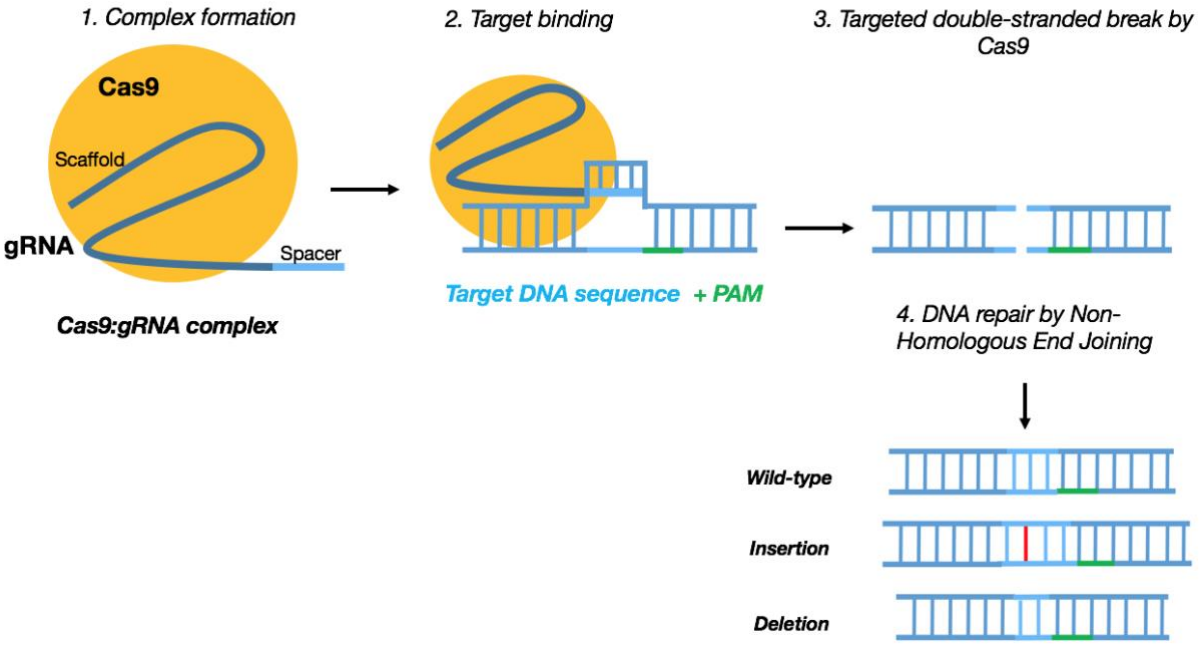

Figure 4: General Mechanism of Gene Editing Using CRISPR/Cas9

General Mechanism of Gene Editing via CRISPR/Cas9 in which 1. Cas9 and gRNA form a complex and then 2. Bind together to the target sequence via PAM recognition. 3. Cas9 produces a double-stranded break and 4. The cell repairs the DNA break via non-homologous end joining (NHEJ) and can either result in the wild-type strand or an insertion or deletion (indels).

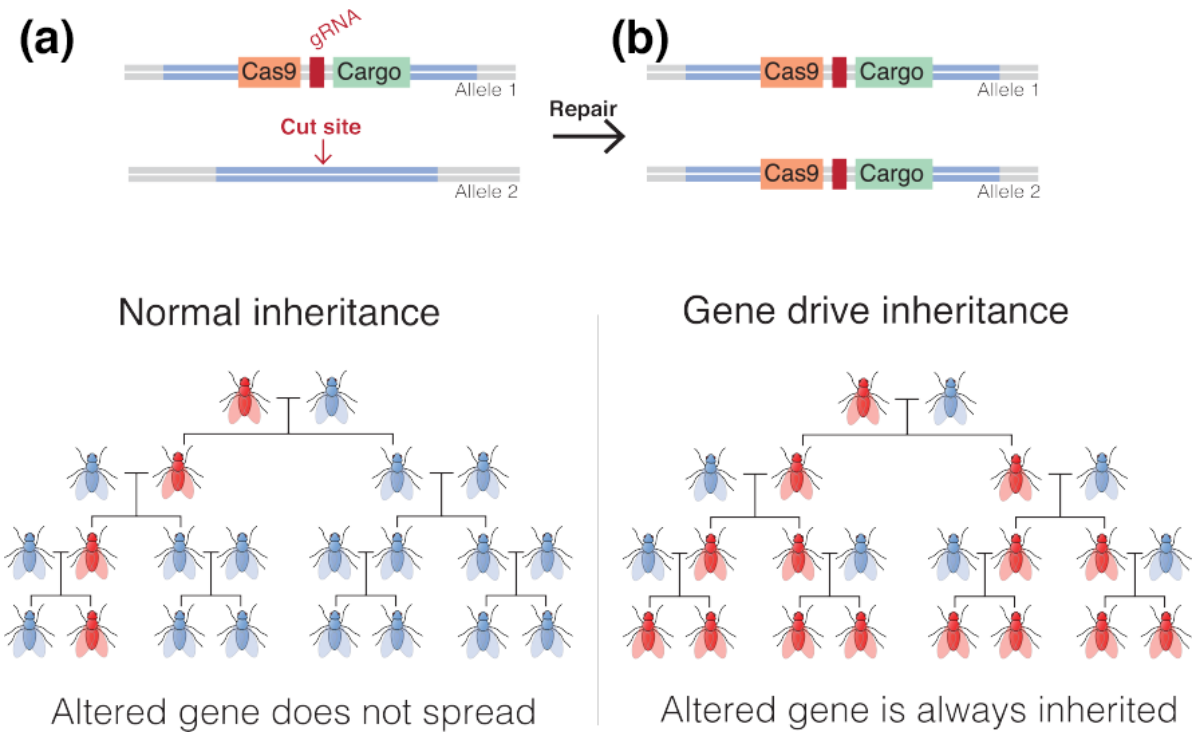

Figure 5: Active Genetic Gene Drive (Figure edited and based on Mariuswalter)
a) Cas9 and gRNA are introduced into allele 1 via transformation. Cas9:gRNA complex are active and make a cut at the target site at allele 2. Altered allele inheritance follows normal Mendelian Laws. b) If allele 2 receives the Cas9 and gRNA now both alleles are homozygous for the CRISPR/Cas9 construct. Therefore, the altered allele is likely to be inherited at a higher percentage and bypasses Mendelian laws via a strategy known as Active Genetics.

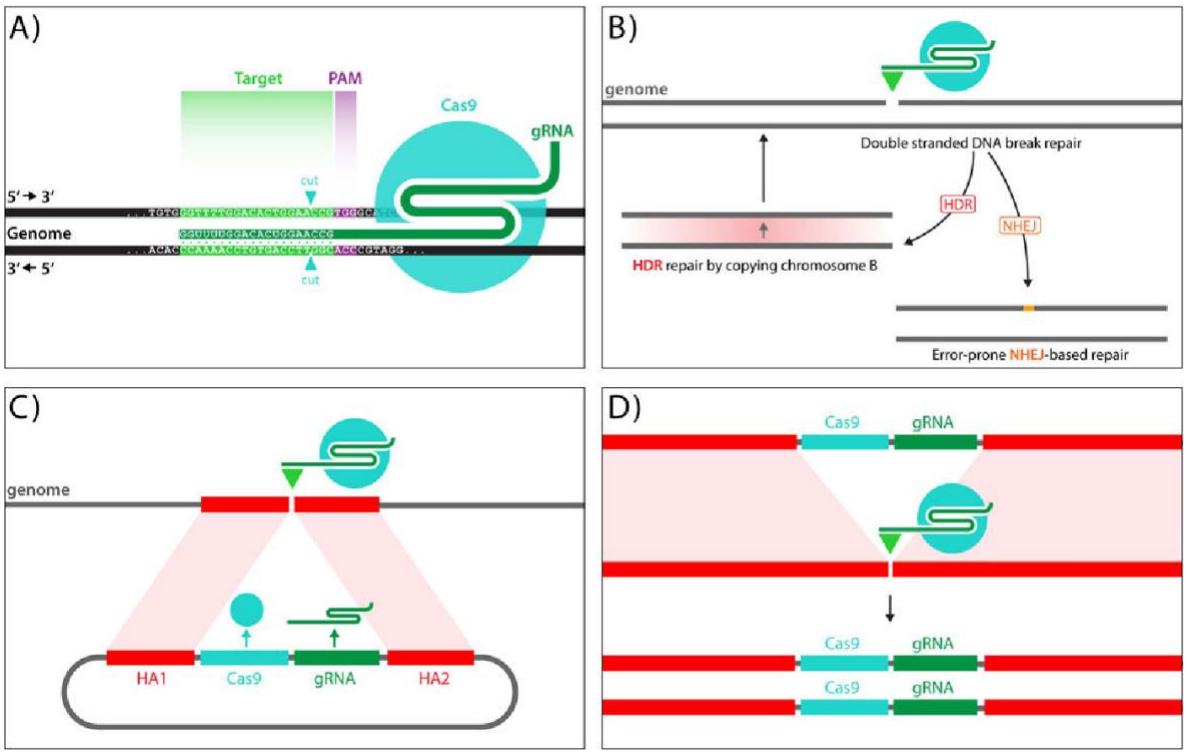

Figure 6: General Active Genetics Methodology (Figure from Gantz and Bier, 2015)
a) Cas9 and gRNA complex bound to target region. b) Cas9:gRNA complex make a double-stranded break at targeted region in which cell's repair system will either use NHEJ which usually results in indels or HDR using another template. c) Cas9:gRNA complex will make a cut at desired region and the provided template containing Cas9 and gRNA construct with two homology arms will be inserted into the target region. d) expressed Cas9 in the allele will then make a cut in second allele to create a homozygous mutation.

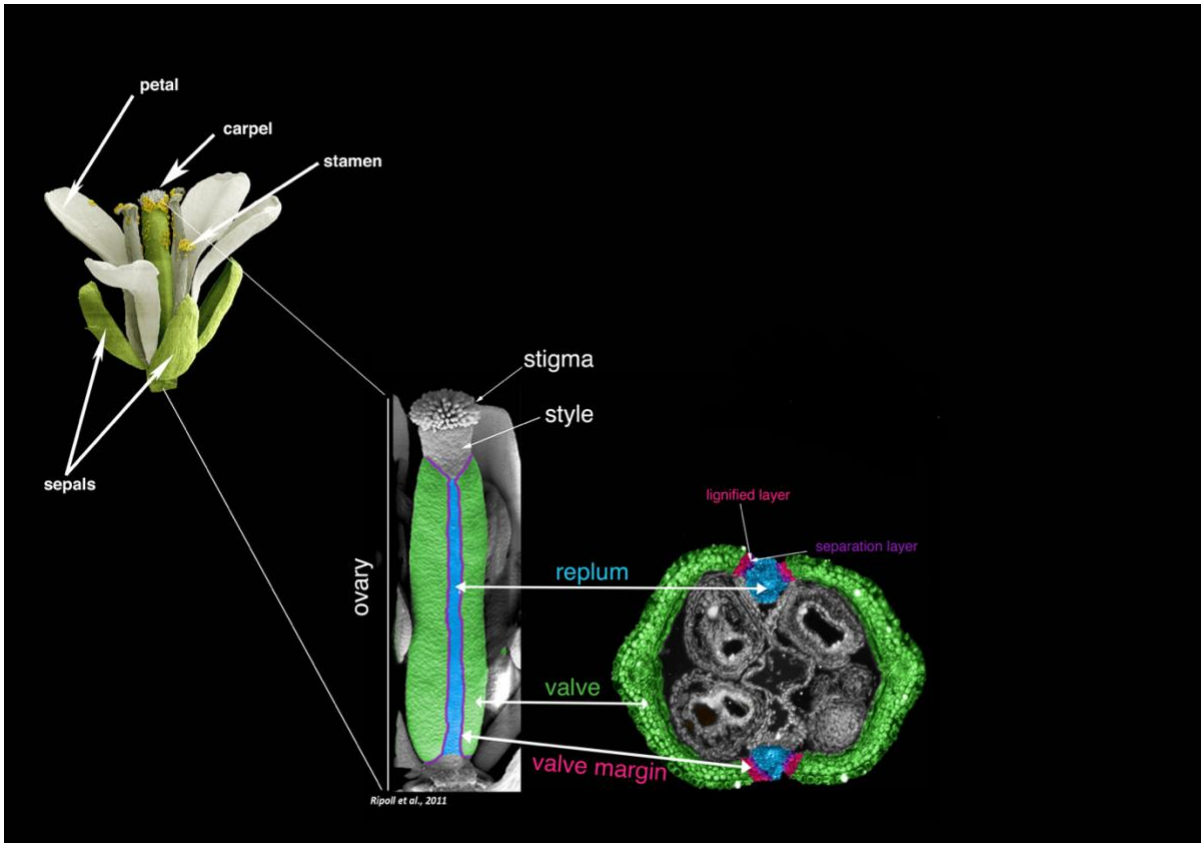

Figure 7: Anatomy of the Arabidopsis thaliana fruit and flower (Modified from Ripoll et al., 2011)

On the left is the Arabidopsis flower, with its four basic parts labeled. The schematic of the fruit is presented, with the ovary labeled with the important tissue regions: replum is labeled in blue, valve is labeled in green, valve margin is in pink and is composed of the lignified layer (pink) and the separation layer (purple).

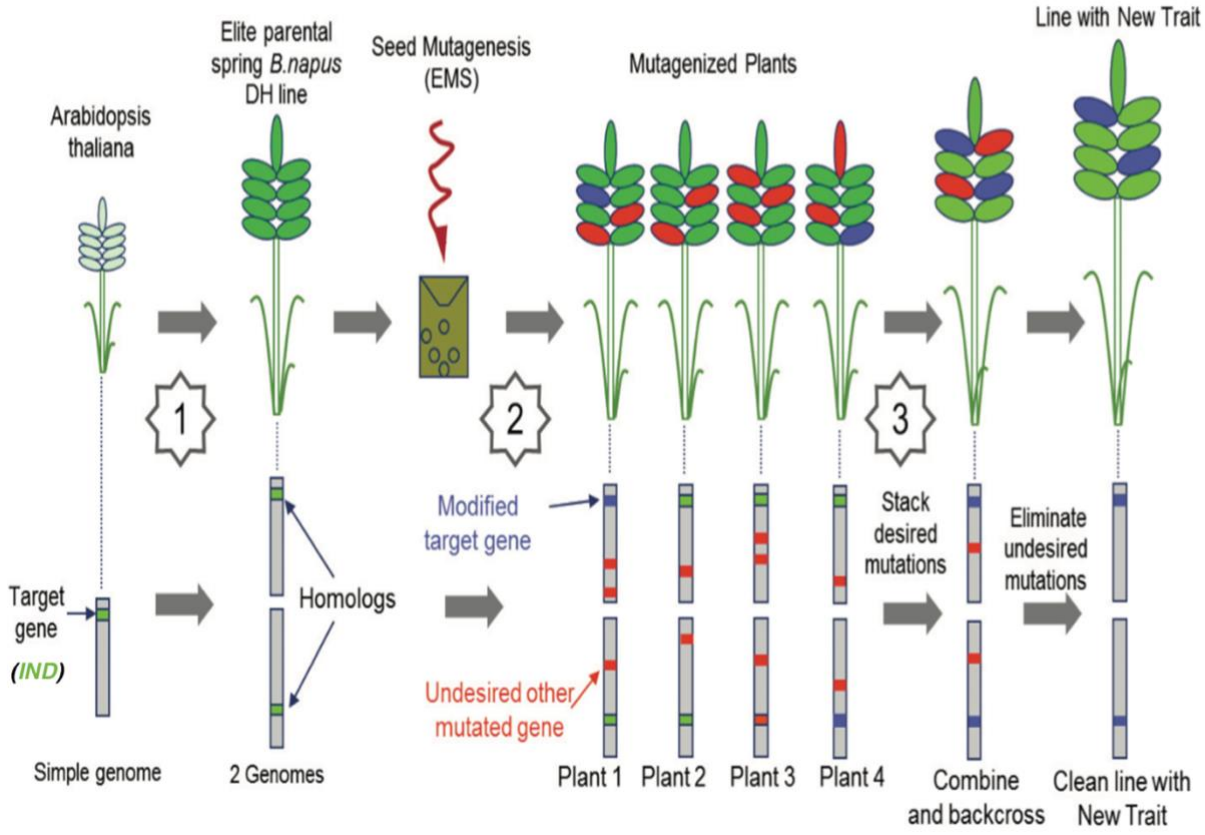

Figure 8: Identification of Target Gene in Arabidopsis and Modification of *Brassica napus* to Achieve Desired Trait (Modified from Lambert et al., 2015)

IND gene which is involved in fruit dehiscence was identified in Arabidopsis. This gene had its homologs identified in oilseed rape (*B. napus*) which underwent chemical mutagenesis and were backcrossed until the line with the new desired trait was identified.

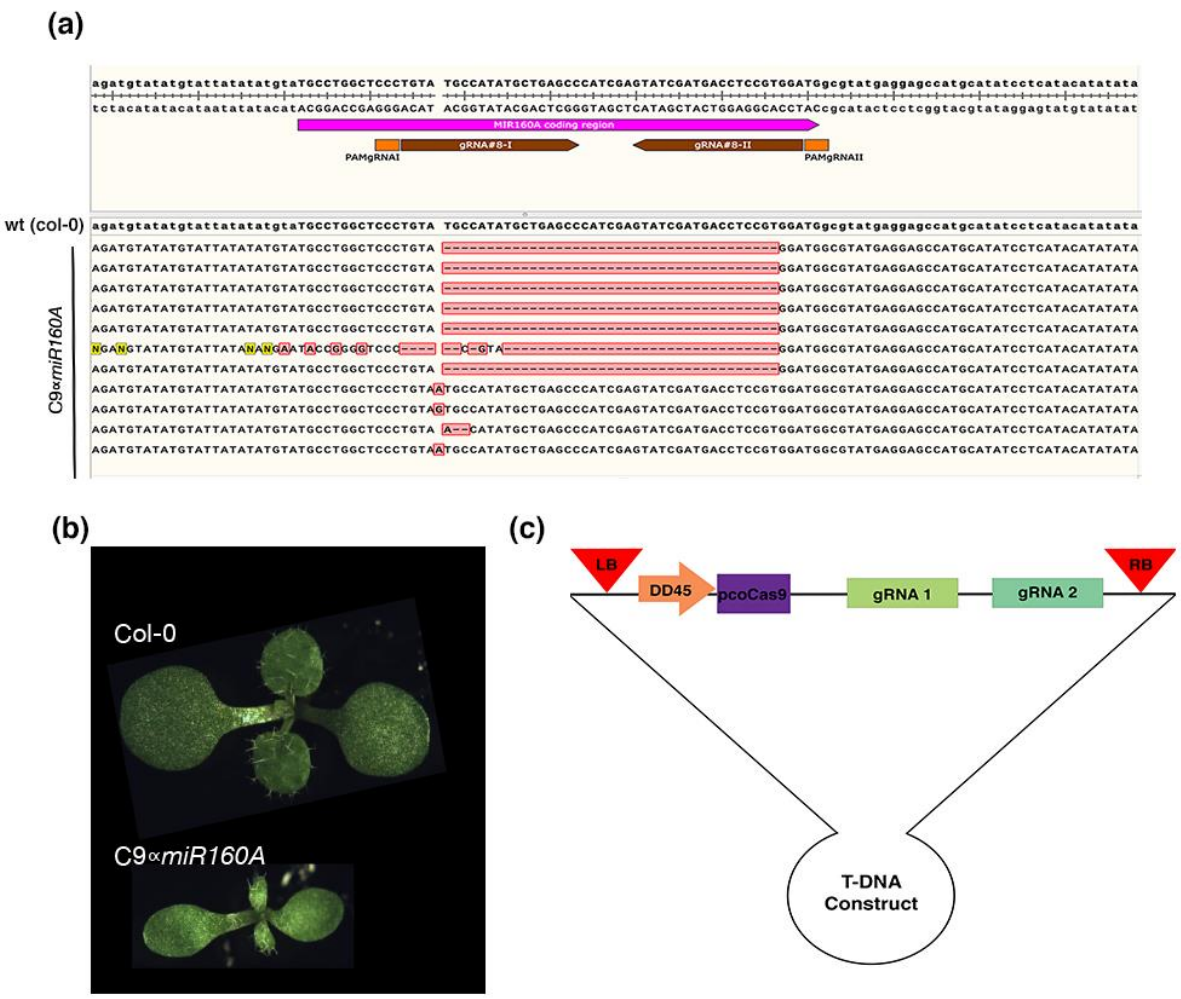

Figure 9: Targeting 2n C9∝*miR160A* in *cis* (all-in-one method).
(a) Sequencing results for *miR160A* locus between C9∝*miR160A* and Col-0. (b) Phenotype of mir160 mutant (C9∝*miR160A)* vs wild-type Col-0 plants (c) T-DNA construct for delivering the Cas9 and 2 gRNAs targeting *miR160A* driven under the DD45 egg-cell promoter.

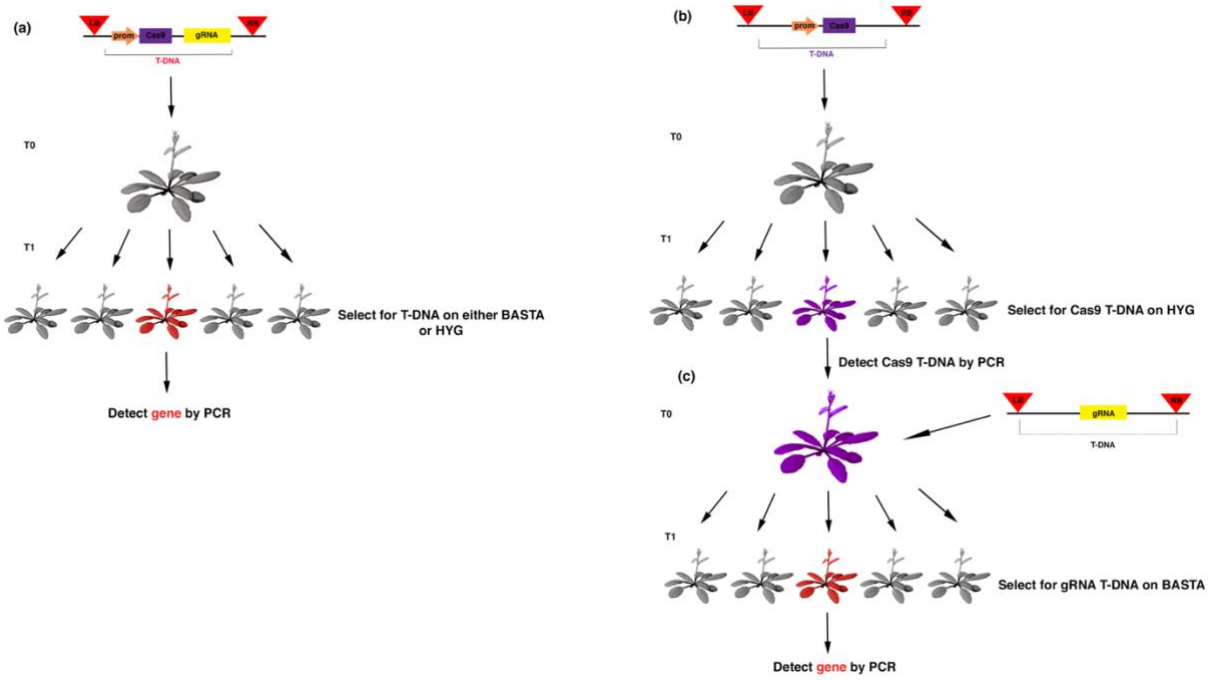

Figure 10: General workflow for the all-in-one vs the split method for delivering Cas9 and gRNA to *Arabidopsis thaliana*.

(a) All-in-one (cis) method in which the T-DNA harbors both the Cas9 and the gRNA expressing modules (b) Developing Cas9 driver lines (c) Representation of the split system in which a Cas9 driver line is transformed with a T-DNA containing the desired gRNA.

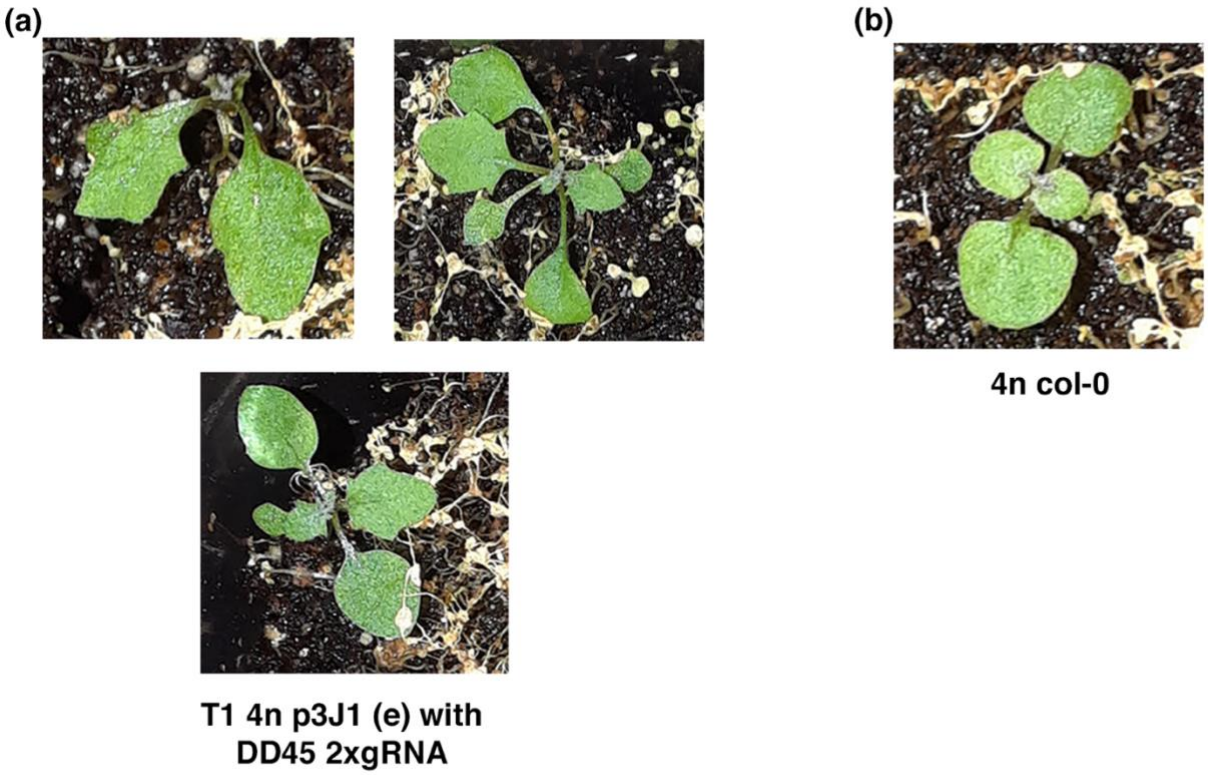

Figure 11: Comparison of phenotypes of T1 4n p3J1 (e) transformed with 2xgRNA and wild-type 4n Col-0.

(a) Serrated leaf phenotype of 4n T1 plants generated using the split system method in which 4n p3J1 (e) driver lines were transformed with 2xgRNA∝*miR160A.* (b) 4n wild-type Col-0 phenotype of non-transformed T1 plant.

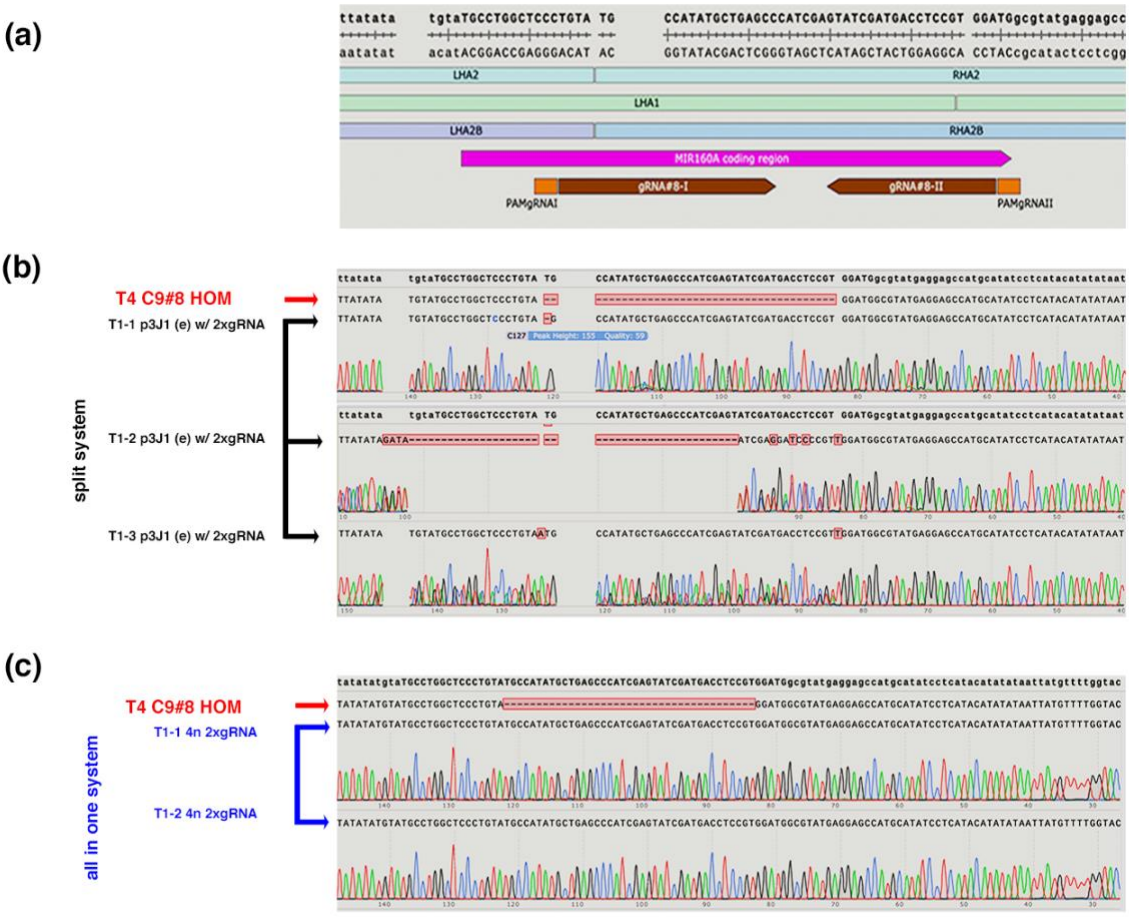

Figure 12: Sequence results for plants transformed with split system vs all-in-one
a) MIR160A locus target region flanked by two guide RNAs. b) Sequencing of original allele deletion in 2n T4 homozygous line compared to the split system. Chromatograms of the transformed plants by the split system depicted. c) Sequencing of original allele deletion in T4 homozygous line compared to the all-in-one system. Chromatograms of the transformed plants by the all-in-one system depicted.

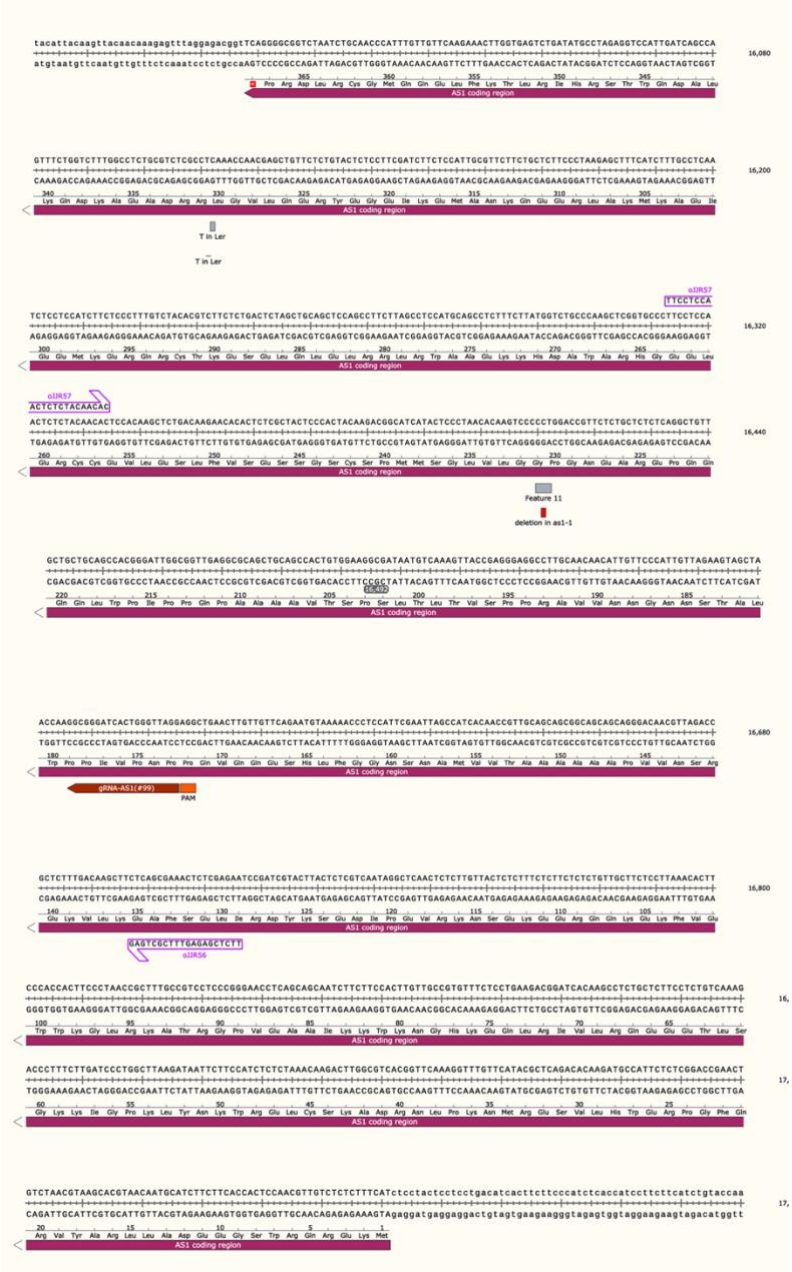

Figure 13: Sequence of *AS1* locus targeted by gRNA

Sequence results for gRNA∝*AS1* designed in the center of the *AS1* locus that was amplified using primers oJJR56 and oJJR57. Shown a deletion mutation in that was identified in a previous *as1-1* allele. In this particular line (line T2-13) we identified a larger deletion that deleted more than half of the gene.

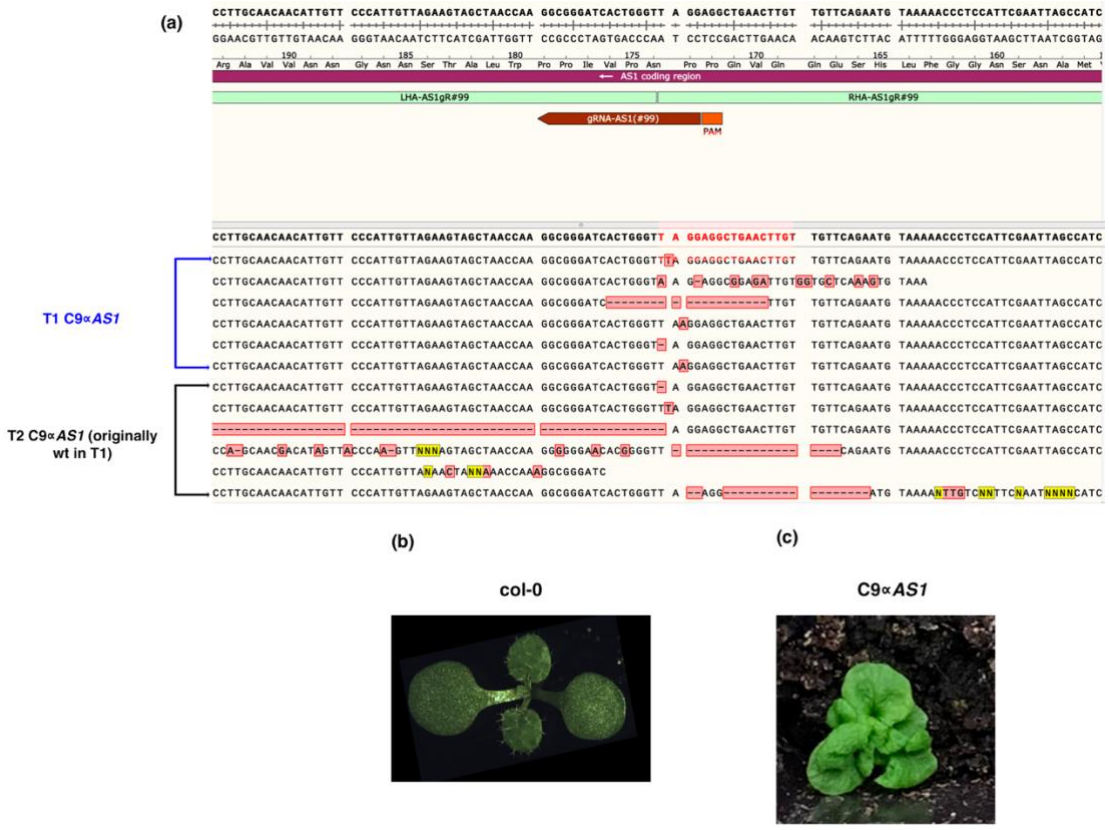

Figure 14: Identifying edits in 2n T1 and T2 C9∝*AS1*

a) Sequencing results in T1 vs T2 C9∝*AS1* in which edits were generating using gRNA∝AS1 in which the gRNA is designed in the middle of the *AS1* locus. b) 2n col-0 *Arabidopsis* wild-type phenotype c) 2n C9∝*AS1* mutant phenotype.

APPENDIX B: TABLES AND TABLE LEGENDS

Table 1: List of primers used to generate cassettes used to create Cas9 knock-out mutants in *MIR160A* gene.
The primers listed were used to generate the 2xgRNA cassette to target the *MIR160A* gene

Primer Name	**Sequence**
oJJR961	CCCGGGGGATCCACTAGTCGACTTGCCTTCCGCACAATAC
oJJR967	GCGGCCGCTCTAGAACTAGTTATTGGTTTATCTCATCGGAACTG
oJJR1001	TGTATGCCATATGCTGAGCCAATCACTACTTCGACTCTAGCTG
oJJR1002B	GGCTCAGCATATGGCATACAGTTTTAGAGCTAGAAATAGCAAG
oJJR1003B	aaacTCCACGGAGGTCATCGATACAATCTCTTAGTCGACTC
oJJR1004B	GTATCGATGACCTCCGTGGAGTTTTAGAGCTAGAAATAGCAAG

Table 2: List of primers used to generate cassettes used to create Cas9 knock-out mutants in *AS1* gene.
The primers listed were used to generate the 1xgRNA cassette to target the *AS1* gene

Primer Name	**Sequence**
oJJR961	CCCGGGGGATCCACTAGTCGACTTGCCTTCCGCACAATAC
oJJR967	GCGGCCGCTCTAGAACTAGTTATTGGTTTATCTCATCGGAACTG
oJJR1614	CCTAACCCAGTGATCCCGCCaatcactacttcgactctagctg
oJJR1615	GGCGGGATCACTGGGTTAGGgttttagagctagaaatagcaag

REFERENCES

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel,D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C., Ecker, J.R., 2003. Genome-Wide Insertional Mutagenesis of Arabidopsis thaliana. Science 301, 653- 657.

Alonso-Cantabrana, H., Ripoll, J.J., Ochando, I., Vera, A., Ferrándiz, C., Martínez- Laborda, A., 2007. Common regulatory networks in leaf and fruit patterning revealed by mutations in the Arabidopsis ASYMMETRIC LEAVES1 gene. Development 134, 2663-2671.

Amitai G, Sorek R. CRISPR-Cas adaptation: insights into the mechanism of action. *Nat Rev Microbiol*. 2016;14(2):67-76. doi:10.1038/nrmicro.2015.14

Barber WT, Zhang W, Win H, Varala, K.K., Dorwiler, J.E., Hudson, M.E.., Moose, S.P. Repeat associated small RNAs vary among parents and following hybridization in maize. *Proc Natl Acad Sci U S A*. 2012;109(26):10444-10449. doi:10.1073/pnas.1202073109

Bailey, Lindsay J.. "Expanding the genetic network controlling fruit development in Arabidopsis thaliana." (2012).

Bartel, D.P., 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116, 281-297.

Beckmann JS, Soller M (1986a) Restriction fragment length polymorphisms and genetic improvement of agricultural species. Euphytica 35:111–124

Bent A. (2006) *Arabidopsis thaliana* Floral Dip Transformation Method. In: Wang K. (eds) Agrobacterium Protocols. Methods in Molecular Biology, vol 343. Humana Press

Beyaz R, Yildiz M. The Use of Gamma Irradiation in Plant Mutation Breeding, Plant Engineering, Snježana Jurić, IntechOpen, 2017. DOI: 10.5772/intechopen.69974. Available from: https://www.intechopen.com/books/plant-engineering/the-use-of-gamma-irradiation-in-plant-mutation-breeding

Borlaug, N. E. (1969). A green revolution yields a golden harvest: Technology canbe more revolutionary than any ‘‘ism’’; agricultural, more than industrial,technology is transforming the internal economies of developing countries andmay soon be a source of critically needed foreign exchange. Columbia Journal ofWorld Business, 4: 9–19.

Burr, B., Burr, F.A., Thompson, K.H., Albertsen, M.C. & Stuber, C.W. 1988. Gene mapping with recombinant inbreds in maize. Genetics 118: 519–526.

Byrne, M.E., Barley, R., Curtis, M., Arroyo, J.M., Dunham, M., Hudson, A., Martienssen, R.A.,

2000. Asymmetric leaves1 mediates leaf patterning and stem cell function in Arabidopsis. *Nature* 408, 967-971.

Byrne, M.E., Simorowski, J., Martienssen, R.A., 2002. ASYMMETRIC LEAVES1 reveals knox gene redundancy in Arabidopsis. *Development* 129, 1957-1965.

Cahill, D.J. & Schmidt, D.H. 2004. Use of marker-assisted selection in a product development breeding program. *Proc. 4th Internat. Crop Sci. Congr.* (available at www.cropscience.org.au.).

Carrington JC, Ambros V. Role of microRNAs in plant and animal development. *Science*. 2003;301(5631):336-338. doi:10.1126/science.1085242

Child, R. D., Chauvaux, N., John, K., Ulvskov, P. & Onckelen, H. A. Ethylene biosynthesis in oilseed rape pods in relation to pod shatter. J. Exp. Bot. 49, 829±838 (1998).

Clough, S.J. and Bent, A.F., 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The plant journal, 16(6), 735-743.

Coe, E., Hancock, D., Kowalewski, S. & Polacco, M. 1995. Gene list and working maps. Maize Genet. Coop. Newsl. 69: 191–267.

Comai L. The advantages and disadvantages of being polyploid. Nat. Rev. Genet. 2005;6:836–846.

Corneillie S., De Storme N., Van Acker R., Fangel, J.U., Bruyne, M.D., Rycke, R.D., Geelen, D., Willats, W.G.T., Vanholme, B., Boerjan, W. Polyploidy Affects Plant Growth and Alters Cell Wall Composition. *Plant Physiol*. 2019;179(1):74-87. doi:10.1104/pp.18.00967

Cregan, P.B., Mudge, J., Fickus, E.W., Danesh, D., Denny, R. & Young, N.D. 1999. Two simple sequence repeat markers to select for soybean cyst nematode resistance conditioned by the *rhg1* locus. *Theor. Appl. Genet*. 99: 811–818.

Crosbie, T.M., Eathington, S.R., Johnson, G.R., Edwards, M., Reiter, R., Stark, S., Mohanty, R.G., Oyervides, M., Buehler, R.E., Walker, A.K., Dobert, R., Delannay, X., Pershing, J.C., Hall, M.A. & Lamkey, K.R. 2006. Plant breeding: past, present, and future. In K.R. Lamkey & M. Lee, eds. Plant Breeding: The Arnel R. Hallauer Internat. Symp. Ames, IA, USA., Blackwell Publishing.

Davis, G.L., McMullen, M.D., Baysdorfer, C., Musket, T., Grant, D., Staebell, M., Xu, G., Polacco, M., Koster, L., Melia-Hancock, S., Houchins, K., Chao, S. & Coe Jr., E.H. 1999. A maize map standard with sequenced core markers, grass genome reference points and 932 expressed sequence tagged sites (ESTs) in a 1736-locus map. Genetics 152: 1137–1172.

Demirer, G.S., Zhang, H., Goh, N.S., González-Grandío, E., Landry, M.P. Carbon nanotube–mediated DNA delivery without transgene integration in intact plants. *Nat Protoc* 14, 2954–2971 (2019). https://doi.org/10.1038/s41596-019-0208-9

Dinneny, J. R., and Yanofsky, M. F. (2005). Drawing lines and borders: how the dehiscent fruit of Arabidopsis is patterned. Bioessays 27, 42–49.

Dinneny, J.R., Weigel, D., Yanofsky, M.F. A genetic framework for fruit patterning in Arabidopsis thaliana. Development 132, 4687–4696 (2005).

Djami-Tchatchou AT, Sanan-Mishra N, Ntushelo K, Dubery IA (2017) Functional roles of microRNAs in agronomically important plants-potential as targets for crop improvement and protection. Front Plant Sci 8:378

Dwivedi, SL., Crouch, J.H., Mackill, D.J., Xu, Y., Blair, M.W., Ragot, M., Upadhyaya, H.D., Ortiz R (2007) The molecularization of public sector crop breeding: progress, problems and prospects. Adv Agron 95:163-318.

Eathington, S.R. 2005. Practical applications of molecular technology in the development of com- mercial maize hybrids. *Proc. 60th Ann. Corn and Sorghum Seed Res. Conf.* Washington, D.C., USA, American Seed Trade Association.

Farasat I, Salis HM (2016) A Biophysical Model of CRISPR/Cas9 Activity for Rational Design of Genome Editing and Gene Regulation. PLoS Comput Biol 12(1): e1004724. doi:10.1371/journal. pcbi.1004724

Fehr, W.R. 1987. Principles of cultivar development. New York, USA, Macmillan.

Feng, Z., Zhang, B., Ding, W., Liu, Z., Yang, D., Wei, P., Cao, F., Zhu, F., Mao, y., Zhu, J. Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res* 23, 1229–1232 (2013). https://doi.org/10.1038/cr.2013.114

Ferrándiz, C., Pelaz, S., Yanofsky, M.F., 1999. Control of carpel and fruit development in Arabidopsis. Annu Rev Biochem 68, 321-354. Fornara, F., Coupland, G., 2009. Plant Phase Transitions Make a SPLash. Cell 138, 625-627.

Ferrándiz, C., Gu, Q., Martienssen, R. and Yanofsky, M. F. (2000a). Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. Development 127, 725-34.

Ferrándiz, C., Liljegren, S.J., Yanofsky, M.F., 2000b. Negative regulation of the SHATTERPROOF genes by FRUITFULL during Arabidopsis fruit development. *Science* 289, 436-438.

Food and Agriculture Organization. (2009) How to Feed the World in 2050. Retrieved from http://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf

Gaj T, Gersbach CA and Barbas CF, III. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* 2013; 31: 397–405.

Gan C. (1989) Gene Gun Accelerates DNA-Coated Particles To Transform Intact Cells. The Scientist, Vol.3, No.18, p. 25

Gantz VM, Bier E. Genome editing. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science*. 2015;348(6233):442-444. doi:10.1126/science.aaa5945

Gantz VM, Bier E. The dawn of active genetics. *Bioessays*. 2016;38(1):50-63. doi:10.1002/bies.201500102

Gaud, William S. (8 March 1968). "The Green Revolution: Accomplishments and Apprehensions". AgBioWorld. Retrieved 8 August 2011.

Gebhardt C, Walkemeier B, Henselewski H, Barakat A, Delseny M, Stüber K. Comparative mapping between potato (Solanum tuberosum) and Arabidopsis thaliana reveals structurally conserved domains and ancient duplications in the potato genome. *Plant J*. 2003;34(4):529-541. doi:10.1046/j.1365-313x.2003.01747.x

Goodman, M.M., Stuber, C.W., Newton, K. & Weissinger, H.H. 1980. Linkage relationships of 19 enzyme loci in maize. Genetics 96: 697–710.

S. González-Reig, J. J. Ripoll, A. Vera, M. F. Yanofsky, A. Martínez-Laborda, Antag- onistic gene activities determine the formation of pattern elements along the me- diolateral axis of the Arabidopsis fruit. PLoS Genet. 8, e1003020 (2012).

Groszmann M, Paicu T, Alvarez JP, Swain SM, Smyth DR. SPATULA and ALCATRAZ, are partially redundant, functionally diverging bHLH genes required for Arabidopsis gynoecium and fruit development. *Plant J*. 2011;68(5):816-829. doi:10.1111/j.1365-313X.2011.04732.

Gu, Q., Ferrándiz, C., Yanofsky, M.F., Martienssen, R., 1998. The FRUITFULL MADS- box gene mediates cell differentiation during Arabidopsis fruit development. Development 125, 1509-1517.

Guo, M., Thomas, J., Collins, G., Timmermans, M.C., 2008. Direct repression of KNOX loci by the ASYMMETRIC LEAVES1 complex of Arabidopsis. Plant Cell 20, 48- 58.

Guo Q, Liu Q, Smith NA, Liang G, Wang MB. RNA Silencing in Plants: Mechanisms, Technologies and Applications in Horticultural Crops. *Curr Genomics*. 2016;17(6):476-489. doi:10.2174/1389202917666160520103117

Gur, A.and Zamir, D. (2004) Unused natural variation can lift yield barriers in plant breeding. PLoS Biology 2(10), e245

Hartwell L. *Genetics: From Genes to Genomes*, 6th ed. New York: McGraw-Hill Education, 2017.

Helentjaris, T., Slocum, M., Wright, S., Schasfer, A. & Nienhuis, J. 1986. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theor. Appl. Genet. 72: 761–769.

Hellens, R.P., Edwards, E.A., Leyland, N.R., Bean, S. and Mullineaux, P.M., 2000. pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant molecular biology, 42(6), 819-832.

Hoisington, D. 1989. Working linkage maps. Maize Genet. Coop. Newsl. 63: 141–15.

Holt-Giménez, E., Shattuck, A., Altieri, M., Herren, H., & Gliessman, S (2012) We Already Grow Enough Food for 10 Billion People ... and Still Can't End Hunger, Journal of Sustainable Agriculture, 36:6, 595-598, DOI: 10.1080/10440046.2012.695331

Jankeleand, R and Svoboda, P. "TAL effectors: Tools for DNATargeting," *Briefings in Functional Genomics*, vol. 13, no. 5, pp. 409–419, 2014.

Jenks MA, Hasegawa PM, Mohan Jain S (2007) Advances in Molecular Breeding toward Drought and Salt Tolerant Crops. Springer, Dordrecht, The Netherlands

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E., 2012. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science, 337(6096), 816-821.

Johnson, G.R. 2004. Marker-assisted selection. *In* J. Janick, ed. *Plant Breeding Rev.* 24: 293–310.

Jung C-H, Wong CE, Singh MB, Bhalla PL (2012) Comparative Genomic Analysis of Soybean Flowering Genes. PLoS ONE 7(6): e38250. doi:10.1371/ journal.pone.0038250

Kamburova, V.S., Nikitina, E.V., Shermatov, S.E., Buriev, Z.T., Kumpatla, S.P., Emani, C., Abdurakhmonov, I.Y., 2017. Genome editing in plants: an overview of tools and applications. Inter. J. Agronomy 2017, 7315351.

Kang Y, Khan S, Ma X. Climate change impacts on crop yield, crop water productivity and food security—a review. *Prog Nat Sci.* 2009; 19: 1665-1674

Kharkwal MC, Shu QY. The role of induced mutations in world food security. In: Shu QY, editor. Induced plant mutations in the genomics era. Rome: Food and Agriculture Organization of the United Nations; 2009. p. 3338.

Khush GS. Green revolution: the way forward. *Nat Rev Genet*. 2001;2(10):815-822. doi:10.1038/35093585

Kim., Y., Cha, J., and Chandrasegaran, S "Hybrid restriction enzymes: zinc nger fusions to Fok I cleavage domain," *Proceed- ings of the National Academy of Sciences of the United States of America*, vol. 93, no. 3, pp. 1156–1160, 1996.

Kossmann J. Grand challenges in plant biotechnology. Front Plant Sci. 2012;3:61. DOI: 10.3389/fpls.2012.00061

Ku HM, Vision T, Liu J, Tanksley SD. Comparing sequenced segments of the tomato and Arabidopsis genomes: large-scale duplication followed by selective gene loss creates a network of synteny. *Proc Natl Acad Sci U S A*. 2000;97(16):9121-9126. doi:10.1073/pnas.160271297

Laga, B., Boer, B., Lambert, B. 2014. *Brassica Plant Compromising a Mutant Indehiscent Allele.* 8809635.

Lambert B, Denolf P, Engelen S, Golds T, Haesendonckx B, Ruiter R, Robbens S, Bots M, Laga B. 2015. Omics-directed reverse genetics enables the creation of new productivity traits for the vegetable oil crop canola. Procedia Environmental Sciences 29: 77–78

Lee K, Eggenberger AL, Banakar R, McCaw, M.E., Zhu, H., Main, M., Kang, M., Gelvin, S.B., Wang, K. CRISPR/Cas9-mediated targeted T-DNA integration in rice. *Plant Mol Biol*. 2019;99(4-5):317-328. doi:10.1007/s11103-018-00819-1

Li J. F., Norville J. E., Aach J., McCormack M., Zhang D., Bush J., Church, G.M., Sheen, J. (2013). Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. Nat. Biotechnol. 31, 688–691. 10.1038/nbt.2654

Liljegren, S. J., Ditta, G. S., Eshed, Y., Savidge, B., Bowman, J. L. and Yanofsky, M. F. (2000). SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature 404, 766-70.

Liljegren, S.J., Roeder, A.H.K., Kempin, S.A., Gremski, K., Østergaard, L., Guimil, S., Reyes, D.K., Yanofsky, M.F. Control of fruit patterning in Arabidopsis by INDEHISCENT. Cell 116, 843–853 (2004).

Lin L, Allemekinders H, Dansby A, Campbell, L., Durance-Tod, S., Berger, A., Jones, P. J.H. Evidence of health benefits of canola oil. *Nutr Rev*. 2013;71(6):370-385. doi:10.1111/nure.12033

Liu Q. & Chen Y.Q. (2010) A new mechanism in plant engineering: the potential roles of microRNAs in molecular breeding for crop improvement. Biotechnology Advances 28, 301–307.

Liu, Z., Jia, L., Wang, H., He, Y., 2011. HYL1 regulates the balance between adaxial and abaxial identity for leaf flattening via miRNA-mediated pathways. *Journal of experimental botany* 62, 4367-4381.

MacLeod, J. in Oilseed Rape Book 107 ± 119 (Cambridge Agricultural, Cambridge, 1981).

Maher C, Stein L, Ware D. Evolution of Arabidopsis microRNA families through duplication events. *Genome Res*. 2006;16(4):510-519. doi:10.1101/gr.4680506

Mariuswalter, from Wikimedia Commons, *licensed under Creative Commons Attribution-Share Alike 4.0 International license;* https://creativecommons.org/licenses/by-sa/4.0/

Meyerowitz, Elliot M. "Arabidopsis, a Useful Weed." Cell, vol. 56, no. 2, 27 Jan. 1989, pp. 263–269.

Miklas, P.N., Kelly, J.D., Beebe, S.E. & Blair, M.W. 2006. Common bean breeding for resistance against biotic and abiotic stresses: from classical to MAS breeding. *Euphytica* 147: 105–131.

Mojica FJ, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol*. 2005;60(2):174-182. doi:10.1007/s00239-004-0046-3

Mouhu, K., Hytönen, T., Folta, K, Rantanen, M., Paulin, L., Auvinen, P., Elomaa, P. Identification of flowering genes in strawberry, a perennial SD plant. *BMC Plant Biol* 9, 122 (2009). https://doi.org/10.1186/1471-2229-9-122

Niebur, W.S., Rafalski, J.A., Smith, O.S. & Cooper, M. 2004. Applications of genomics technologies to enhance rate of genetic progress for yield of maize within a commercial breeding program. *In* T. Fischer, N. Turner, J. Angus, L. McIntyre, M. Robertson, A. Borrell & D. Lloyd, eds. *New directions for a diverse planet. Proc. 4th Internat. Crop Sci. Congr.,* Brisbane, Australia. (available at www.cropscience.org.au/icsc2004).

National Research Council. 1984. *Genetic Engineering of Plants: Agricultural Research Opportunities and Policy Concerns.* Washington, DC: The National Academics Press. https://doi.org/10.17226/10.

Nekrasov V., Staskawicz B., Weigel D., Jones J. D., Kamoun S. (2013). Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. Nat. Biotechnol. 31, 691–693. 10.1038/nbt.2655

Ogutcen, E.; Pandey, A.; Khan, M.K.; Marques, E.; Penmetsa, R.V.; Kahraman, A.; Von Wettberg, E.J.B. Pod Shattering: A Homologous Series of Variation Underlying Domestication and an Avenue for Crop Improvement. *Agronomy* 2018, *8*, 137.

Ori, N., Eshed, Y., Chuck, G., Bowman, J.L., Hake, S., 2000. Mechanisms that control knox gene expression in the Arabidopsis shoot. Development 127, 5523-5532. Park, M.Y., Wu, G., Gonzalez-Sulser, A., Vaucheret, H., Poethig, R.S., 2005. Nuclear processing and export of microRNAs in Arabidopsis. Proc Natl Acad Sci U S A 102, 3691-3696.

Osborn T.C., Pires J.C., Birchler J.A., Auger D.L., Chen Z.J., Lee H.-S., Comai L., Madlung A., Doerge R.W., Colot V., Martienssen, R. A. Understanding mechanisms of novel gene expression in polyploids. Trends Genet. 2003;19:141–147.

Otero, G. and G. Pechlaner. 2008. Latin American agriculture, food, and biotechnology: Temperate dietary pattern adoption and unsustainability. In: G. Otero, ed. Food for the few: Neoliberal globalism and biotechnology in Latin America. Austin: University of Texas Press, pp 31–56.

Passricha N, Saifi S, Khatodia S, Tuteja N. Assessing zygosity in progeny of transgenic plants: current methods and perspectives. *J Biol Methods*. 2016;3(3):e46. Published 2016 Jul 18. doi:10.14440/jbm.2016.114

Peng FY, Hu Z, Yang RC. Genome-Wide Comparative Analysis of Flowering-Related Genes in Arabidopsis, Wheat, and Barley. *Int J Plant Genomics*. 2015;2015:874361. doi:10.1155/2015/874361

Penin AA, Klepikova AV, Kasianov AS, Gerasimov ES, Logacheva MD. Comparative Analysis of Developmental Transcriptome Maps of Arabidopsis thaliana and Solanum lycopersicum. *Genes (Basel)*. 2019;10(1):50. Published 2019 Jan 15. doi:10.3390/genes10010050

Prohens J. Plant Breeding: A Success Story to be Continued Thanks to the Advances in Genomics. *Front Plant Sci*. 2011;2:51. Published 2011 Sep 16. doi:10.3389/fpls.2011.00051

Ragot, M., P.H. Sisco, D.A. Hoisington, and C.W. Stuber. 1995. Molecularmarker-mediated characterization of favorable exotic alleles at quantitative trait loci in maize. Crop Sci. 35:1306–1315.

Ragot, M., and M. Lee. 2007. Marker-assisted selection in maize: Current status, potential, limitations and perspectives from the private and public sectors. p. 117–150. In E.P. Guimarães, E., Ruane, J., Scherf, B., Sonnino, A., Dargie, J (ed.) Marker-assisted selection, current status and future perspectives in crops, livestock, forestry, and fish. FAO, Rome.

Rajani S, Sundaresan V. 2001. The *Arabidopsis* myc/bHLH gene ALCATRAZ enables cell separation fruit dehiscence. *Curr. Biol.* 11:1914-22

Rath D, Amlinger L, Rath A, Lundgren M. The CRISPR-Cas immune system: biology, mechanisms and applications. *Biochimie*. 2015;117:119-128. doi:10.1016/j.biochi.2015.03.025

Raymer, P.L. 2002. Canola: An emerging oilseed crop. p. 122–126. In: J. Janick and A. Whipkey (eds.), Trends in new crops and new uses. ASHS Press, Alexandria, VA.

Reardon, S. (2016). The CRISPR zoo. *Nature* 531, 160-163.

Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. MicroRNAs in plants [published correction appears in Genes Dev 2002 Sep 1;16(17):2313]. *Genes Dev*. 2002;16(13):1616-1626. doi:10.1101/gad.1004402

Renny-Byfield S, Wendel JF. Doubling down on genomes: polyploidy and crop plants. *Am J Bot*. 2014;101(10):1711-1725. doi:10.3732/ajb.1400119

Rhoades, M.W; Bartel, D.P; Bartel B. MicroRNAs and Their Regulatory Roles in Plants. Annual Review of Plant Biology 2006 57:1, 19-53.

J. J. Ripoll, A. H. Roeder, G. S. Ditta, M. F. Yanofsky, A novel role for the floral homeotic gene APETALA2 during Arabidopsis fruit development. Development 138, 5167–5176 (2011).

Ripoll, J.J., Bailey, L.J., Mai, Q.A., Wu, S.L., Hon, C.T., Chapman, E.J., Ditta, G.S., Estelle, M., Yanofsky, M.F., 2015. microRNA regulation of fruit growth. Nature Plants 1, 15036.

Ripoll, J.J., Zhu, M., Brocke,S., Hon, C.T., Yanofsky, M.F., Boudaoud, A., Roeder, A. Growth dynamics of the *Arabidopsis* fruit is mediated by cell expansion. *Proceedings of the National Academy of Sciences* Dec 2019, 116 (50) 2533325342;
DOI: 10.1073/pnas.1914096116

Roeder, A.H., Ferrándiz, C., Yanofsky, M.F. The role of the REPLUMLESS homeo-domain protein in patterning the Arabidopsis fruit. Curr. Biol. 13, 1630–1635 (2003).

Roeder, A.H., Lijegren, S., Ferrándiz, C., Yanofsky, M.F. Drawing a line in the Arabidopsis fruit. Dev. Biol. 283, 658 (2005).

Roeder, A.H., Yanofsky, M.F., "Fruit development in Aabidopsis" in The Arabidopsis Book, C. R. Somerville, E. M. Meyerowitz, Eds. (American Society of Plant Biologists, Rockville, MD, 2006).

Romera-Branchat, M., Ripoll, J.J., Yanofsky, M.F., Pelaz, S. The WOX13 homeobox gene promotes replum formation in the Arabidopsis thaliana fruit. Plant J. 73, 37–49 (2013).

Roser, M., Ritchie, H., and Ortiz-Ospina, E. (2020) - "World Population Growth". *Published online at OurWorldInData.org.* Retrieved from: 'https://ourworldindata.org/world-population-growth'

Rouet, P., Smih, F. and Jasin, M., 1994a. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Molecular and cellular biology, 14(12), 8096-8106.

Rouet, P., Smih, F. and Jasin, M., 1994b. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proceedings of the National Academy of Sciences, 91(13), 6064-6068.

Semiarti E, Ueno Y, Tsukaya H, Iwakawa H, Machida C, Machida Y. The ASYMMETRIC LEAVES2 gene of Arabidopsis thaliana regulates formation of a symmetric lamina, establishment of venation and repression of meristem-related homeobox genes in leaves. *Development*. 2001;128(10):1771-1783.

Shan Q., Wang Y., Li J., Zhang Y., Chen K., Liang Z., Zhang, K., Liu, J., Xi, J.J., Qiu, J., Gao, C (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. 31, 686–688. 10.1038/nbt.2650

Shapouri, H., McAloon, A., "The 2001 net energy balance of corn ethanol'' (U.S. Department of Agriculture, Washington, DC, 2004). Also available at www.usda. gov/oce/oepnu.

Sontheimer EJ, Barrangou R. The Bacterial Origins of the CRISPR Genome-Editing Revolution. *Hum Gene Ther*. 2015;26(7):413-424. doi:10.1089/hum.2015.091

Sorefan, K., Girin, T., Liljegren, S.J., Ljung, K., Robles, P., Galván-Ampudia, C.S., Offringa, R., Friml, J., Yanofsky, M.F, Østergaard, L. A regulated auxin minimum is required for seed dispersal in Arabidopsis. Nature 459, 583–586 (2009).

Stadler LJ. Genetic Effects of X-Rays in Maize. *Proc Natl Acad Sci U S A*. 1928;14(1):69-75. doi:10.1073/pnas.14.1.69

Steffen JG, Kang IH, Macfarlane J, Drews GN. Identification of genes expressed in the Arabidopsis female gametophyte. *Plant Journal*. 2007;51:281–92.

Stephenson, P., Stacey, N., Brüser, M, Pullen, N., Ilyas, M., O'Neill, C., Wells, R., Østergaard, L. The power of model-to-crop translation illustrated by reducing seed loss from pod shatter in oilseed rape. *Plant Reprod* 32, 331–340 (2019). https://doi.org/10.1007/s00497-019-00374-9

Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*. 2014;507(7490):62-67. doi:10.1038/nature13011

Sun, G. MicroRNAs and their diverse functions in plants. *Plant Mol Biol* 80, 17–36 (2012). https://doi.org/10.1007/s11103-011-9817-6

Swaminathan, M. Norman E. Borlaug (1914–2009). *Nature* 461, 894 (2009). https://doi.org/10.1038/461894a

Tanksley, S.D., N.D. Young, A.H. Paterson, and M.W. Bonierbale. 1989. RFLP mapping in plant breeding: New tools for an old science. Biotechnology (N. Y.) 7:257–263.

Tanksley, S.D., and S.R. McCouch. 1997. Seed banks and molecular maps: Unlocking genetic potential from the wild. Science 277:1063–1066.

Tran, Q.H. "Adapting CRISPR/CAs9 Gene Editing to Characterize Gene Functions During Arabidopsis thaliana Fruit Development" (2019).

Wang, Z., Xing, H., Dong, L., Zhang H., Han, C., Wang, X., Chen, Q. Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome Biol* 16, 144 (2015). https://doi.org/10.1186/s13059-015-0715-0

Weeks, D.P., Spalding, M.H., and Yang, B. Use of designer nucleases for targeted gene and genome editing in plants. *Plant Biotechnology Journal*, vol. 14, no. 2, pp. 483–495, 2016.
Woodward, A.W., and Bartel, B. Biology in Bloom: A Primer on the *Arabidopsis thaliana* Model System. GENETICS. *April 1, 2018 vol.208 no.413371349; https://doi.org/10.1534/genetics.118.300755*

Weeks DP. Gene Editing in Polyploid Crops: Wheat, Camelina, Canola, Potato, Cotton, Peanut, Sugar Cane, and Citrus. *Prog Mol Biol Transl Sci*. 2017;149:65-80. doi:10.1016/bs.pmbts.2017.05.002

Xie K., Minkenberg B., Yang Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc. Natl. Acad. Sci. U.S.A. 112, 3570–3575. 10.1073/pnas.1420294112

Xiong, J., J. Ding, J., Li, Y. Genome-editing technologies and their potential application in horticultural crop breeding. *Horticulture Research*, vol. 2, article 15019, 2015.

Xu, Y (2003) Developing marker-assisted selection strategies for breeding hybrid rice. Plant Breed Rev 23:73-174.

Xu, Y., and J.H. Crouch. 2008. Marker-assisted selection in plant breeding: From publications to practice. Crop Sci. 48:391–407.

Xu, Y., Iacuone, S., Li, S.F., Parish, R.W. MYB80 homologues in Arabidopsis, cotton and Brassica: regulation and functional conservation in tapetal and pollen development. *BMC Plant Biol* 14, 278 (2014). https://doi.org/10.1186/s12870-014-0278-3

Yajko DM, Hegeman GD. Tumor induction by Agrobacterium tumefaciens: specific transfer of bacterial deoxyribonucleic acid to plant tissue. *J Bacteriol*. 1971;108(3):973-979.

Yanfei Mao, Jose Ramon Botella, Yaoguang Liu, Jian-Kang Zhu, Gene editing in plants: progress and challenges, *National Science Review*, Volume 6, Issue 3, May 2019, Pages 421–437, https://doi.org/10.1093/nsr/nwz005

Yangbin Gao, Yunde Zhao. Gene editing in plants. Proceedings of the National Academy of Sciences Mar 2014, 111 (12) 4357-4358; DOI: 10.1073/pnas.1402295111

Ye GN, Stone D, Pang SZ, Creely W, Gonzalez K, Hinchee M. Arabidopsis ovule is the target for Agrobacterium in planta vacuum infiltration transformation. *Plant J*. 1999;19(3):249-257. doi:10.1046/j.1365-313x.1999.00520.x

Yoshizumi Ishino, Mart Krupovic, Patrick Forterre. History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. *Journal of Bacteriology* Mar 2018, 200 (7) e00580-17; doi 10.1128/JB.00580-17

Zhang, J., Li, X., Li, G. Chen, W., Arakaki, C., Botimer, G.D., Baylink, D., Zhang, Lu., Wen, W., Fu, Y., Xu, J., Chun, N., Yuan, w., Cheng, T., Zhang, X. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol* 18**,** 35 (2017). https://doi.org/10.1186/s13059-017-1164-8