THESIS

MOHAMMAD UMAR

M.Sc. Agricultural Biotechnology

Gödöllő

2023

Hungarian University of Agriculture and Life Sciences

Szent István Campus

Agricultural Biotechnology

MAKING AND CHARACTERIZATION OF DNA DEMETHYLASE 1 MUTANT IN TOMATO

Primary supervisor: Dr. György Szittya
group leader, scientific advisor

Secondary supervisor: Dr. Péter Gyula
project leader, senior scientist

Author: **Mohammad Umar**

GBX14C

Institute: Institute of Genetics and Biotechnology
Department of Plant Biotechnology

Gödöllő

2023

1. TABLE OF CONTENT

1. TABLE OF CONTENT 2
2. ABBREVIATIONS 3
3. INTRODUCTION 5
4. OBJECTIVES 9
5. LITERATURE REVIEW 10
5.1. What is epigenetics? 10
5.2. DNA methylation in plants 12
5.3. DNA demethylation in plants 15
5.4. Role of active DNA demethylation in physiological processes 15
5.5. Genome editing with the CRISPR/Cas9 system 20
5.6. Using of genome editing to improve agricultural traits in plants 24
6. MATERIALS AND METHODS 31
6.1. Designing CRISPR guide sequences 31
6.2. Modular cloning of the CRISPR construct 31
6.3. Agrobacterium transformation 34
6.4. Tomato transformation 34
6.5. Plant genomic DNA purification 35
6.6. Genotyping PCR 36
6.7. Cloning of the mutant alleles 37
7. RESULTS AND DISCUSSION 38
7.1. Designing CRISPR guide sequences 38
7.2. Modular cloning of the CRISPR construct 41
7.3. Tomato transformation 44
7.4. Genotyping of the transformed plants 45
7.5. Characterization of the sldml1 mutants 47
8. CONCLUSIONS AND FUTURE PLANS 49
9. SUMMARY 50
10. ACKNOWLEDGEMENTS 51
11. BIBLIOGRAPHY 52

2. ABBREVIATIONS

bp (base pairs)

nt (nucleotide)

rpm (rotate per minute)

DNMTs (DNA methyltransferases)

RdDM (RNA-directed DNA methylation)

DMRs (differentially methylated regions)

DRM (DOMAINS REARRANGED METHYLTRANSFERASES)

dsRNA (double-stranded RNAs)

siRNAs (small interfering RNAs)

SHH 1(Sawadee Homeodomain Homolog 1)

DCL3 (DICER-LIKE 3)

DML (*DEMETER-LIKE*)

DME (*DEMETER)*

ROS1 (*REPRESSOR OF SILENCING 1*)

Pol (POLYMERASE)

RNAi (RNA interference)

rdd (*ros1*/*dml2*/*dml3*)

EPF2 (*EPIDERMAL PATTERNING FACTOR 2*)

LMWg (low-molecular-weight glutenins)

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)

PCR (polymerase chain reaction)

TEs (transposable elements)

PAM (protospacer adjacent motif)

tracrRNA (trans-activating CRISPR RNA)

crRNA (CRISPR RNA)

pre-crRNA (premature crRNA)

sgRNA (single-guide RNA)

HR (homologous recombination)

DSBs (double-stranded breaks)

NHEJ (non-homologous end joining)

HDR (homologous direct repair)

MMEJ (microhomology-mediated end joining)

CaLCuV (cabbage leaf curl virus)

WDV (wheat dwarf virus)

ToLCV (tomato leaf curl virus)

TRV (tobacco rattle virus)

TMV (tobacco mosaic virus)

PEBV (pea early browning virus)

BNYVV (beet necrotic yellow vein virus)

BSMV (barley stripe mosaic virus)

DRs (developmental regulators)

WUS (*WUSCHEL*)

CLV (*CLAVATA*)

STM (*SHOOT MERISTEMLESS*)

MP (*MONOPTEROS*)

PDS (*PHYTOENE DESATURASE*)

QTLs (Quantitative trait locus)

PL (*PECTATE LYASE*)

ALC (ALCOBACA)

3. INTRODUCTION

In the world, tomato (*Solanum lycopersicum*) is one of the most significant vegetable crops. Tomato fruit plays an important role in the human diet and provides health-promoting compounds such as vitamins, minerals, and antioxidants. It is also one of the most popular vegetables in the world. Besides its economic and nutritional importance, tomato is also the principal model to study fleshy fruit development and ripening and to investigate the molecular bases of commercially important traits, such as fruit set, color, flavor, size, and nutritional quality (Azzi et al. 2015, Seymour et al. 2013). Fruit quality and quantity depend on the environmental conditions under which the tomato plants grow. Under natural conditions, tomato plants experience abiotic stresses, like heat, cold, drought, water, light, etc, or biotic stresses, like infection by viruses, bacteria, fungi, or wounds caused by insect bites. Also, the nutrient supply of the plant affects the development and metabolic content of the fruit. Tomato fruit set and ripening are highly light and temperature-dependent. In case of suboptimal light or temperature conditions, the fruit set is delayed or the ripening is halted. There are many different tomato variants that respond to the above-mentioned environmental conditions differently underlining the role of inherited traits in development and stress response. It has also been known for a while that besides classical genetic inheritance epigenetics also play a role in plant stress responses and development, including fruit development. Epigenetics refers to heritable changes in gene expression that occur without modification of the underlying DNA sequence. It involves histone post-translational modifications and DNA methylation which are transmitted through DNA replication (Gallusci et al. 2016). DNA methylation at cytosine residues is an important and conserved epigenetic modification in many eukaryotes, including plants, and is associated with the suppression of transposable elements (TEs), chromatin maintenance, and regulation of gene expression (Zhang et al. 2018, Zhang & Zhu 2012, Zhu 2009). The actual DNA methylation status is set by the balanced regulation of establishment, maintenance, and removal of methylated cytosines (Kumar et al. 2013, Law & Jacobsen 2010). DNA methylation status is dynamic but in some cases, it could be quite stable, resulting in the transgenerational epigenetic inheritance of DNA methylation pattern (epimutations). Epigenetic inheritance was demonstrated to be an important contributor to hybrid vigor or heterosis (increased yield and biomass of hybrid offspring relative to the parents), a phenomenon that is widely exploited in agriculture (Groszmann et al. 2014, 2011).

In plants, DNA methylation occurs in three sequence contexts: CG , CHG, and CHH (H represents A, T, or C). In *Arabidopsis thaliana*, CG and CHG methylation is maintained by

DNA METHYLTRANSFERASE 1 (MET1) and CHROMOMETHYLASE 3 (CMT3), respectively. Depending on the chromatin context, CHH methylation is maintained by CMT2 or by DOMAINS REARRANGED METHYLTRANSFERASES (DRM1 and DRM2) through the RNA-directed DNA methylation (RdDM) pathway, which is also responsible for *de novo* DNA methylation. RdDM, as the name suggests, depends on endogenous, 24-nt-long small interfering RNAs (siRNAs) that are produced by the action of DNA-DEPENDENT RNA POLYMERASE IV and V (PolIV/V), RNA-DEPENDENT RNA POLYMERASE 2 (RDR2), and DICER-LIKE 3 (DCL3). First, at dedicated sites along the genome (i.e. transposons, promoters, etc), the PolIV synthesizes a short, 30–40-nt-long RNA precursor which is immediately converted to a double-stranded RNA by RDR2. This dsRNA precursor is then cleaved by DCL3 into 24-nt siRNAs which are loaded into AGO4-type (AGO4, 6, 9 in Arabidopsis) of silencing complexes. These loaded AGOs are then bound to the AGO-binding domain-containing PolV molecules that recruit DNA methyltransferases which perform the methylation of the correspondent DNA sequence (Zhai et al. 2015).

DNA methylation can be passively lost due to the lack of maintenance or can be actively removed by DNA demethylases. In plants, active DNA demethylation is initiated by the 5-methylcytosine DNA glycosylase enzymes. These enzymes cut the methylated cytosine residue from the DNA molecule and then it is replaced with unmethylated cytosine in a multistep process catalyzed by several enzymes. There are four DNA demethylase genes in Arabidopsis: *REPRESSOR OF SILENCING 1* (*ROS1*), *DEMETER* (*DME*), *DEMETER-LIKE 2* (*DML2*), and *DML3* (Penterman et al. 2007). The roles of *ROS1*, *DML2*, and *DML3*, and later *DME*, in vegetative tissues, have already been characterized (Choi et al. 2002a, Zeng et al. 2021).

The tomato genome contains four putative *DML* genes encoding proteins with characteristic domains of functional DNA glycosylase-lyases (Mok et al. 2010). *SlDML1* (Solyc09g009080) and *SlDML2* (Solyc10g083630) are orthologous to the Arabidopsis *AtROS1* gene while *SlDML3* (Solyc11g007580) is similar to *AtDME*. *SlDML4* (Solyc03g123440) has no closely related Arabidopsis orthologue (Liu et al. 2015). All four *SlDML* genes are ubiquitously expressed in tomato plants, although *SlDML4* is expressed at a very low level in all organs and tissues. *SlDML2* mRNA abundance increases dramatically in ripening fruits, while *SlDML1*, *SlDML3*, and *SlDML4* expressions do not change during fruit ripening (Liu et al. 2015). It was recently demonstrated that deletion of *SlDML2* by the CRISPR/Cas9 system strongly inhibited the onset and progression of fruit ripening (Lang et al. 2017a). This suggests that *SlDML2* is central in mediating the promoter DNA hypomethylation necessary for the

ripening progression of tomatoes. The functions of the other *SlDMLs* have not been investigated in detail yet.

There is only one publication about the role of the *SlDMLs* other than the *SlDML2* (Liu et al. 2015). In this publication, the authors investigated a transgenic plant expressing an RNAi (RNA interference) construct that was designed to silence all the four *SlDMLs* at once. Their results indicate that in this plant, only the expression level of *SlDML1* and *SlDML2* decreased, while the expression of the other two *SlDMLs* was increased. This indicates that the silencing of these two *SlDMLs* was overridden by another mechanism that increased their expression. This is probably due to the action of the other two, actually silenced *SlDMLs*. Furthermore, the net level of *SlDML2* might be influenced by a possible feedback regulation conferred by the *SlDML2* itself. This feedback was described in Arabidopsis: the expression of *ROS1* (*SlDML2* homologue) is positively correlated with the methylation status of its own promoter, which means that the gene expression is induced when its promoter is hypermethylated but this region can be demethylated by the *ROS1* itself (Williams et al. 2015). It is possible that other DNA demethylases can also regulate the methylation status of each others promoters which could indicate the presence of an intricate regulatory network between the DNA demethylases. Since the characterization of the independent lines of the above mentioned tomato RNAi plant revealed several fruit developmental effects, it is possible that the misregulation of some other *SlDMLs* might enhance the observed phenotypes suggesting that the other *SlDMLs* also play a role in tomato fruit development and ripening.

The transcriptional regulation of tomato *SlDMLs* under conditions other than development has not been investigated yet. The role of active DNA demethylation in plant abiotic and biotic stress responses has been demonstrated (Pandey et al. 2016, Yu et al. 2013). However, it was not shown which demethylases are responsible for the observed responses. Different *SlDMLs* may have different or partially overlapping target gene sets, therefore may confer different and partially overlapping functions.

Keeping the above mentioned information in mind, we can raise several questions regarding tomato DNA demethylases:

- What is the the function of *SlDMLs* other than that of *SlDML2*? Do they regulate development or stress responses?
- Do they regulate the methylation status of each other's promoter?
- Is there a feedback regulation of DNA methylation in tomato like in Arabidopsis?

- Does their function overlap with each other or are they specialized to certain functions or expressed in different tissues?

To answer some of these questions, we decided to knock out the most abundantly expressed *SlDML1* in tomato by CRISPR/Cas9-based genome editing technique. Having the mutants in a homozygous form, we will perform different stress experiments to find out whether the *SlDML1* has a role in regulating stress responses or development. The final goal is to determine the genomic targets of SlDML1. In order to do this, we will perform genome wide DNA methylation and transcript sequencing in the mutant and wild-type plants. The analysis of these datasets will hopefully shed light on the function of *SlDML1* in tomato.

4. OBJECTIVES

DNA demethylation is well studied in *Arabidopsis thaliana* but only limited pieces of information are available from other plants, especially those of agricultural importance. In tomato, there are four DNA demethylases of which only *SlDML2* has been investigated in more detail so far because it was found to be dramatically induced during fruit ripening. Its functional role in regulation of fruit ripening was later experimentally verified using the CRISPR/Cas9-based genome editing technology. Similar expression pattern was not observed in the case of the other three *SlDMLs*. This might be the reason why researchers did not focus on the investigation of these genes. However, because some of them are ubiquitously expressed, they must have some other important functions, like regulation of developmental processes or stress responses, but it is also possible that they have some overlapping function with *SlDML2* in fruit ripening.

To investigate the function of *SlDML1* in tomato, we decided to create a null mutant of *SlDML1*, the most ubiquitously and highly expressed *DML*, using the CRISPR/Cas9-based genome editing technology. In order to achieve this, we set the following goals:

1. Design specific guide RNAs targeting the protein-coding region of *SlDML1* at two sites simultaneously to increase the chance to get a null mutant.
2. Create the binary gene construct carrying a kanamycin selection marker, the Cas9 gene cassette, and the two modules expressing the guide RNAs.
3. Transform the construct into tomato plants and regenerate transformant plants from calluses in tissue culture.
4. Verify the presence of CRISPR/Cas9-generated mutation(s) in the transformant plants.
5. Physiological characterization of the verified mutant plants.

5. LITERATURE REVIEW

5.1. What is epigenetics?

Epigenetics is a scientific field that emerged in the 1940s and 1950s with the discovery of DNA methylation and the concept of chromosomal imprinting. The term "epigenetics" was first coined by Waddington in 1942 (Waddington 1942) to describe the study of how genes interact with the environment to shape an organism's development and traits. In the past few decades, the field of epigenetics has grown rapidly, with increasing evidence that epigenetic mechanisms play a critical role in the regulation of gene expression and the development of many diseases (Bird 2007). In other words, epigenetics is the study of heritable changes in gene expression that occur without any change in the DNA sequence. It includes structural modifications, such as DNA methylation and histone modification, as well as non-coding RNAs (ncRNAs). In plants, epigenetic processes play a vital role in development, and they allow plants to adapt quickly to changing environments (Chang et al. 2020).

In agriculture, epigenetic modifications can play a critical role in hybrid vigor or heterosis, which is the phenomenon where the offspring of two different varieties exhibit enhanced growth and yield compared to either parent (Groszmann et al. 2013). In heterosis, the epigenetic changes resulting from the combination of two different genomes can lead to increased heterozygosity and diversity, thereby improving the performance of the hybrid plants. For example, inter-specific hybridization has been found to increase the yield of crops such as maize and rice (Groszmann et al. 2013, 2011). Similarly, research involving hybrid plants of *Arabidopsis* has demonstrated the existence of heterosis for various characteristics such as including vegetative biomass, size and cell number and the amount of seeds produced (Groszmann et al. 2014). Heterosis has been associated with a variety of interactive features, such as changes in gene expression, metabolic activity, and epigenetic control. As of yet, the precise molecular mechanisms behind heterosis remain an area of active research and are not fully understood. An analysis of the transcriptomes of hybrids and their parent plants revealed that many genes exhibit expression levels that deviate from the average expression levels seen in the parent plants. This deviation is referred to as a nonadditive expression (Kapazoglou et al. 2018). In recent years, researchers have used genome-wide studies to examine the epigenetic factors that contribute to heterosis (Kapazoglou et al. 2018). In *Arabidopsis*, changes in DNA methylation patterns, which impact gene expression, have been observed in hybrid plants (Shen et al. 2012). Similarly, changes in global epigenetics and transcription have been documented

in rice and maize hybrids compared to their parent plants (He et al. 2013). Epimutations, or epigenetic mutations, are also important in agriculture (Figure 1).

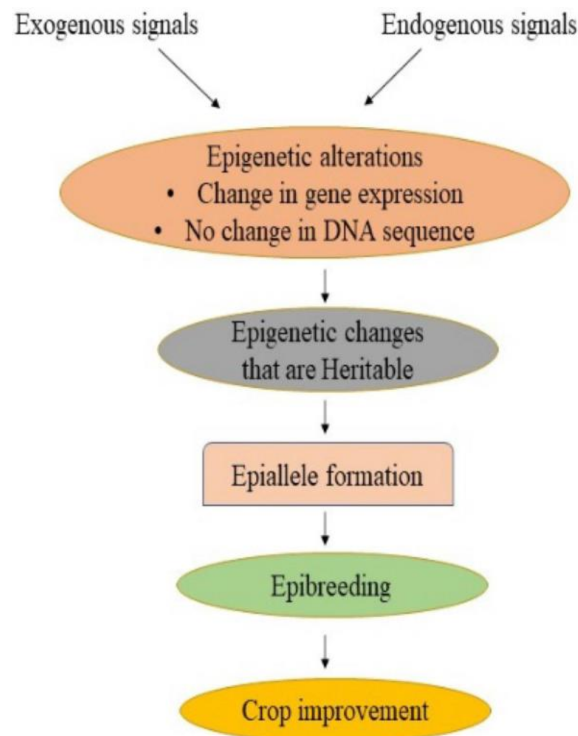


Figure 1. Epigenetics and plant breeding (Gupta & Salgotra 2022).
Epimutations are changes in gene expression or cellular phenotype caused by mechanisms other than changes in the underlying DNA sequence. These changes can happen spontaneously or be induced and can be transmitted to future generations, creating new epigenetic variants (epialleles) which can lead to further phenotypic variation.

Plants utilize the process of cytosine methylation extensively in their genomes to regulate the expression of transposable elements (TEs) and genes. Although this regulation is stringent, there can be spontaneous occurrences of loss or gain of methylation at individual or groups of cytosines (van der Graaf et al. 2015). A study by Hofmeister et al. (2017) on two separate lines of the *Arabidopsis* plant investigated the potential for spontaneous epimutations to be passed down through the generations. The results showed that, on average, approximately 43% and 46% of the newly-emerged epialleles were maintained in the following generations (Hofmeister et al. 2017). High-level environmental stress such as drought can improve epimutations rate. A study conducted with rice found that drought stress over 11 generations led to the accumulation of transgenerational epimutations, which improved the adaptability of

the offspring in the field. They also revealed that the genes in the pathways that respond to stress had increased of transgenerational epimutations, and the DNA methylation patterns of the genes that reacted to drought were altered by multi-generational drought. They observed that 30% of the alterations to methylation were stable and heritable (Zheng et al. 2017). Wang et al. conducted a study and found that 29% of the changes in DNA methylation due to drought stress were still present following the return to normal conditions (Wang et al. 2011). In another study, Kumar and Singh observed that 25% of increased methylation in the rice genotype IR-64-DTY1.1 remained even after the drought had ended (Kumar 2016). Similarly, in *Arabidopsis thaliana*, a single trait connected to seed dormancy when exposed to drought has a transgenerational epiallelic stability (Ganguly et al. 2017). Therefore, epigenetics is a key factor in the long-term ability of plants to adjust and evolve amid difficult circumstances (Zheng et al. 2017).

5.2. DNA methylation in plants

DNA methylation is a process by which a methyl group (-CH3) is added to the fifth carbon of cytosine. Methylation of DNA is an important epigenetic modification that plays a crucial role in gene expression, genomic imprinting, and other cellular processes. The process of DNA methylation is catalyzed by DNA methyltransferases (DNMTs), which transfer the methyl group from S-adenosylmethionine (SAM) to the cytosine residue of the DNA. The reverse process, i.e., removal of the methyl group, is called DNA demethylation, which is accomplished by several mechanisms (Kumar et al. 2018) (Figure 2).

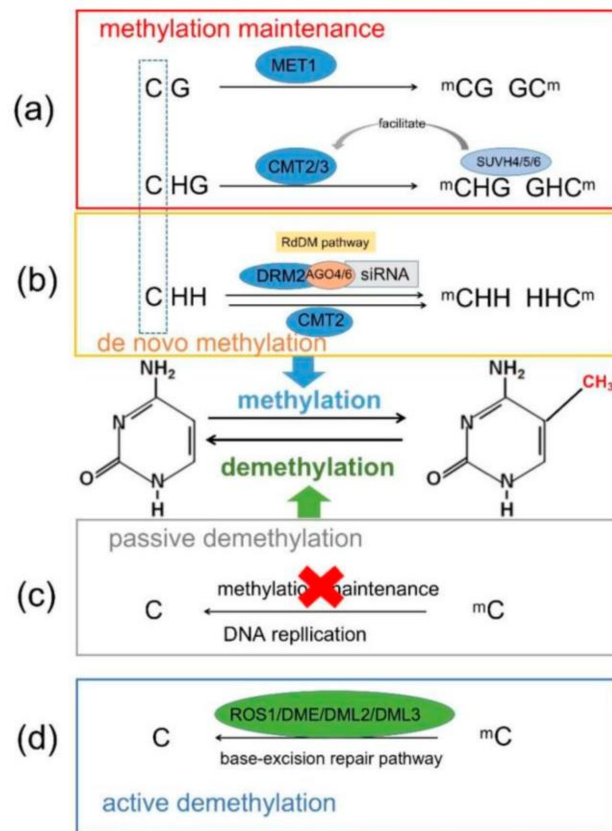


Figure 2. Dynamics of DNA methylation in plants (Fan et al. 2022).

(H=A, T or C). Two DNA methylation processes exist in plants: (a) DNA methylation maintenance and (b) *de novo* DNA methylation. (a) Methyltransferase 1 (MET1) maintains symmetric CG site methylation. CHROMOMETHYLASE (CMT2/3) maintains symmetrical CHG site methylation. The suppressor of the variegation homolog protein, SUVH4, SUVH5 and SUVH6 binds to the methylated CHG site and promotes the function of CMT3/CMT2. (b) Asymmetric *de novo* DNA methylation and CHH methylation performed by domain-rearranged methyltransferase 2 (DRM2) or CMT2, depending on the genomic region. DRM2 causes CHH methylation through the RNA-directed DNA methylation (RdDM) pathway, which depends on the 24 nt small interfering RNA (siRNA). siRNA is loaded onto the ARGONAUTE proteins (AGO), mainly AGO4 and AGO6, interacting with DRM2. DNA demethylation includes (c) passive demethylation and (d) active demethylation. (c) 5mC loses its methyl in passive demethylation during DNA replication. (d) 5mC losses are catalyzed by DNA glycosylases in active demethylation. DNA glycosylases including the repressor of silencing 1 (ROS1), Demeter (DME), Demeter-like 2 and 3 (DML2/3).

Methylation of cytosine in plants can take place in the cytosine contexts of CG, CHG, and CHH, where H represents the nucleotides A, C, or T. The RNA-directed DNA methylation

(RdDM) pathway relies on the activity of small-interfering RNAs (siRNAs), scaffold RNAs, and numerous accessory proteins to accomplish *de novo* methylation of DNA. In Arabidopsis, it is thought that the RNA polymerase IV (Pol IV) is responsible for initiating the production of 24 nt siRNA, which serves as a template for RNA-dependent RNA polymerase 2 to generate double-stranded RNAs (dsRNA) and dsRNA cleavage done by DCL3 into siRNA (Zhang et al. 2018). The Sawadee Homeodomain Homolog 1 (SHH 1) protein helps to recruit Pol IV to the targeted loci, which have dimethylated histone H3 lysine 9 (H3K9me2). The Pol IV-dependent siRNA production also requires the interaction of a chromatin remodeler, Classy 1 (CLSY1), with Pol IV (Zhang et al. 2013).

According to the recent model of 24-nt siRNA production, the double-stranded 30–40-nt-long precursor is diced by DCL3 into 24-nt siRNAs with 2-nt 3' overhangs (Zhai et al. 2015). However, it has been observed that some of the RdDM-targeted loci remain methylated in a quadruple DCL (*dcl1*/*dcl2*/*dcl3*/*dcl4*) mutant, suggesting that siRNAs may be produced by DCL-independent RdDM pathways or directly from P4 RNAs. Additionally, some intergenic loci are transcribed by Pol II, producing 24 nt siRNAs and scaffold RNAs, while some activated transposons are transcribed by Pol II and RNA-dependent RNA polymerase 6 (RDR6), generating 21 or 22 nt siRNA precursors in association with DCL2 and DCL4 (McCue et al. 2015).

Once siRNA is produced, it is loaded onto Argonaute (AGO) proteins, specifically AGO4 and/or AGO6, and directly associated with scaffold RNAs transcribed by Pol V. This complex recruits domains rearranged methylase 2 (DRM2), a DNA methyltransferase, for methylation of the target locus. AGO4 interacts with DRM2 to catalyze *de novo* methylation of cytosine in a sequence-independent manner. The DDR complex, consisting of a chromatin remodeler defective in RNA-directed DNA methylation 1 (DRD1) defective in meristem silencing 3 (DMS3), is required for the generation of scaffold RNAs and RNA-directed DNA methylation 1 (RDM1). The DDR complex also interacts with AGO4/AGO6, single-stranded methylated DNA, and DRM2. SUVH2 and SUVH9, which recognize methylcytosine, help recruit Pol V by binding to preexisting methylcytosine. The Pol V-generated scaffold RNAs lack polyadenylation at the 3' end and are long enough to be detected by PCR, differentiating them from mRNA (Kumar & Mohapatra 2021).

5.3. DNA demethylation in plants

Passive and active DNA demethylation are two distinct processes that can result in the loss or removal of DNA methylation. The former is called passive DNA demethylation, while the latter is referred to as active DNA demethylation. Passive DNA demethylation is a consequence of DNA replication where the absence of proper DNA methylation by DNA methyltransferases leads to a reduction in the level of DNA methylation (Zhang et al. 2018). This results in a diluted DNA methylation level and ultimately the loss of DNA methylation on the newly synthesized DNA strand, which is known as passive DNA demethylation (Liu & Lang 2020).

In plants, multiple enzymes work together in a sequential manner to achieve active DNA demethylation. AtROS1, also known as Repressor of Silencing 1, was the initial DNA demethylase discovered in Arabidopsis and is specific to plants (Gong et al. 2002). The ROS1 protein family, which includes AtROS1, AtDME, AtDML2, and AtDML3, are enzymes that have dual functionality as DNA glycosylase/lyases with both DNA glycosylase and apurinic/apyrimidinic (AP) lyase activities. These enzymes are capable of triggering active DNA demethylation by eliminating methylcytosine from the DNA backbone, creating a single-nucleotide gap. The gap is then filled with an unmethylated cytosine through the base-excision repair (BER) pathway (Liu & Lang 2020). The process of active DNA demethylation involves a two-step enzymatic process. Initially, the DNA glycosylase activity removes the 5mC base from the DNA backbone. Next, the lyase activity cleaves the backbone, leading to a single-nucleotide gap with either a 3′-phosphor-α, β-unsaturated aldehyde (3′-PUA) or a 3′ phosphate terminus. The 3′-PUA and 3′ phosphate termini are then processed by the AP endonuclease-like protein APE1L and 3′ DNA phosphatase ZDP, respectively, generating a 3′-OH end. This end allows the DNA polymerase and DNA ligase I (AtLIG1) to fill the gap with an unmethylated cytosine (Li et al. 2015).

5.4. Role of active DNA demethylation in physiological processes

Conceptually, DNA demethylation can be categorized into two types: passive and active. Passive demethylation encompasses the reduction of DNA methylation that occurs as a result of DNA replication when the DNA methylation process is suppressed. Active demethylation is when enzymes catalyze the process of demethylation. Conversely, active demethylation pertains to the demethylation of DNA that is catalyzed by specific enzymes

(Zhang et al. 2022). Over two decades ago, the gene that initiates active DNA demethylation was originally discovered in plants (Gong et al. 2002). Since then, additional research has uncovered that the REPRESSOR OF SILENCING 1 (ROS1)/DEMETER-LIKE (DML) family of DNA glycosylases/lyases is responsible for the initial reaction in the sequence of enzymatic processes that switch out the methylated cytosine for an unaltered cytosine base (Y. Li et al. 2018, Zhu 2009). Active demethylation is thought to play a key role in various physiological processes, including plant and growth development, fruit ripening, genomic stabilities, and stress response (Zhu 2009).

In angiosperms, the generation of the plant embryo occurs through the fertilization of the egg cell by a sperm cell. The subsequent fertilization event involving the central cell and a second sperm cell leads to the formation of the endosperm, an extraembryonic tissue that provides nourishment to the embryo. The vegetative cell in the male gametophyte serves as a companion cell, forming a tube that facilitates the transportation of the sperm cells to the female gametophyte. In *Arabidopsis*, successful reproduction and seed development require active DNA demethylation in both the vegetative and central companion cells, initiated by DEMETER (DME; Demeter-like proteins are a subgroup of plant DNA glycosylases or demethylases) (Parrilla-Doblas et al. 2019). The Arabidopsis DEMETER is primarily expressed in the central cell before fertilization and is required for maternal allele DNA demethylation in the endosperm that establishes gene imprinting (Bauer & Fischer 2011). To prevent the loss of expression of maternal alleles of imprinted genes, which can result in the termination of seeds, it is crucial to have the involvement of DME along with polycomb group genes MEA (MEDEA) and (fertilization independent seed 2) FIS2. If these imprinted genes, such as FWA (flowering wageningen), are not expressed, it can lead to abnormalities in the development of seeds. Therefore, the presence of DME is necessary to ensure proper gene expression and prevent seed abortion (Zhu 2009). The central cell of the female gametophyte is where DME-dependent DNA demethylation begins, with the activation of maternal alleles of *MEA, FWA, and FIS2* genes before fertilization. A wild-type maternal DME allele is necessary for the successful formation of seeds and their viability. In the case of a maternally inherited *dme* mutation, MEA::GFP and FWA::GFP transgenes were not expressed in the endosperm post fertilization, leading to the abortion of the embryo and endosperm (Choi et al. 2002b). DME also has a temporary presence in the male gametophyte's vegetative cell. When DME experiences mutations, it can reduce the viability and germination rates of pollen in specific ecotypes of *Arabidopsis* (Khouider et al. 2021). Furthermore, it has been demonstrated that ROS1, DML2,

and DML3 play a role in controlling the growth of certain cells that have reached their final stage of differentiation. Specifically, the density of stomatal stem cells is fold by three to four times in the *ros1* or *ros1*/*dml2*/*dml3* (*rdd*) mutants compared to the normal or wild-type version (Yamamuro et al. 2014). The appearance of this characteristic is almost identical to the one seen in mutants of *EPIDERMAL PATTERNING FACTOR 2* (*EPF2*), which is responsible for producing a cysteine-rich peptide that counteracts the growth of stomata (Hunt & Gray 2009). In the *ros1* and *rdd* mutants, the promoter region of *EPF2* is over-methylated, resulting in the suppression of *EPF2* gene expression (Parrilla-Doblas et al. 2019, Yamamuro et al. 2014).

Fruit ripening is another phase of reproductive development in which active DNA demethylation plays a critical role. During the ripening of tomatoes (*Solanum lycopersicum*), a widespread decrease in DNA methylation occurs and has been linked to the active DNA demethylation process facilitated by *SlDML2*. *SlDML2* is one of four potential DML genes in tomato that encode 5-meC DNA glycosylases (*SlDML1–4*) (Lang et al. 2017a). The gene *SlDML2*, which is homologs to *Arabidopsis ROS1*, is highly active during the process of fruit ripening and primarily affects transposable elements (TEs) located in open chromatin regions. When *SlDML2* was suppressed using RNA interference (RNAi) or altered using CRISPR/Cas technology, the ripening of the fruit was greatly inhibited because DNA methylation increased on thousands of genes that play a role in fruit ripening (Lang et al. 2017a, Liu et al. 2015). Similarly, in wheat endosperm, the *TaDME* gene, which is similar to DME, demethylates the promoter regions of seed storage proteins such as prolamins that trigger an inappropriate immune response in humans, such as gliadins and low-molecular-weight glutenins (LMWgs), causing them to be expressed at higher levels. When the transcript of *TaDME* was suppressed using RNAi, the accumulation of gliadins and LMWgs in the endosperm was reduced, offering a possible solution for creating wheat varieties that are safe for people with celiac disease (Y. Li et al. 2018, Wen et al. 2012). Furthermore, active DNA methylation is also observed in non-climatic fleshy fruits such as Strawberry (*Fragaria vesca)* and orange (*Citrus sinensis*) (Huang et al. 2019, Moglia et al. 2019). During the ripening of strawberries, DNA methylation is lost, but unlike in tomatoes, the genes responsible for DNA demethylation are not upregulated. Instead, the reduced activity of genes involved in RNA-directed DNA methylation (RdDM) is responsible for the decrease in DNA methylation during strawberry ripening (Cheng et al. 2018). This decrease in DNA methylation is associated with decreased levels of small interfering RNA. The downregulation of genes involved in RdDM during strawberry ripening suggests that the decrease in DNA methylation is caused by reduced RdDM activity. When a

DNA methylation inhibitor is applied or an important RdDM component, FvAGO4, is silenced, hypomethylation occurs and early ripening of strawberry fruit is induced (Cheng et al. 2018, Liu & Lang 2020). A widespread increase in DNA methylation occurs during sweet orange (*Citrus sinensis*) fruit ripening, potentially due to decreased DNA demethylase activity. It is suggested that DNA hypermethylation is important for the proper ripening of oranges, as evidenced by the delay in ripening caused by the application of a DNA methylation inhibitor to unripe oranges. In conclusion, both DNA methylation and demethylation play critical roles in fruit ripening, although their importance may vary depending on the fruit species (Huang et al. 2019, Li et al. 2022).

Maintenance of the integrity of the genome is of utmost importance, and it is achieved through the crucial process of suppressing transposable elements (TEs) with DNA methylation. The cells that support male and female gametophytes in Arabidopsis and rice were discovered to undergo a process of global DNA demethylation (Kim et al. 2019). Arabidopsis *ROS1*, *DML2*, and *DML3* are expressed in a wide range of plant vegetative tissues and they seem to contribute to the stability and plasticity of plant epigenome and protect the genome from excessive methylation. One of the primary genes, *ROS1*, serves as the main 5-meC DNA glycosylase in the plant's vegetative tissues, and was identified during research on mutants with abnormal expression of the repetitive *RD29A-LUC* transgene. The wild-type plants express both the transgene and the endogenous gene, but the *ros1* mutant plants display excessive methylation and transcriptional silencing of both the loci (Gong et al. 2002). This indicates that active DNA demethylation, facilitated by ROS1, helps to protect genes from incorrect silencing. In *ros1* mutants, about 5000 regions are hypermethylated, primarily located in transposable elements (TEs) and intergenic regions (Qian et al. 2012). ROS1 targets transposons that are usually located near protein-coding genes, and hypermethylation in *ros1* mutants spreads from the edges of TEs to neighboring sequences (Zhu et al. 2007). These observations suggest that ROS1-mediated DNA demethylation helps to restrict transposons and genes, preventing methylation from spreading from TEs and thus safeguarding nearby genes from being repressed (Tang et al. 2016). In vegetative cells of pollen, *DME* was the primary gene responsible for 5mC DNA glycosylase expression, but there was also some expression of *ROS1*, *DML2*, and *DML3* (Schoft et al., 2011). Altering the *DME* gene can impact the germination of pollen and the inheritance of the mutated gene in some ecotypes of Arabidopsis (Schoft et al. 2011). Furthermore, the vegetative cell, like central cells, experiences overall DNA demethylation which is dependent on DME. According to Ibarra et al. (2012), DME targets comparable

genomic regions in both central and vegetative cells, and demethylation leads to the de-repression of similar sets of genes and transposable elements. In gamete's "companion" cells, there are approximately 10,000 DME-dependent differentially methylated regions (DMRs), which is about a hundred times higher than the number of imprinted genes. This implies that establishing genome imprinting is not the primary role of demethylation in gametophytes (Ibarra et al. 2012).

In addition to its role in genome stability, ROS1 may help regulate telomere length. A study found that a mutation in the gene encoding the largest subunit of replication factor C (RFC1) acted as a suppressor for ROS 1 and resulted in longer telomeres in *ros1* mutants than in wild-type plants. Conversely, the *rfc1* mutant had shorter telomeres. These findings suggest that ROS1 may negatively regulate telomere length in Arabidopsis by modulating TERT (telomerase reverse transcriptase) or other telomere-related proteins (POT1a, BT2). In *ros1* mutants, the expression levels of these proteins were slightly higher, while they were significantly lower in *rfc1* or *rfc1/ros1* mutants than in wild-type plants (Liu et al. 2010).

Several studies have outlined the impact of DNA methylation or epigenetic regulation on how plants respond to various forms environmental biotic stress associated with symbiotic microorganisms and attack by diverse pathogens (Satgé et al. 2016). According to a recent study, that active DNA demethylation plays a role in controlling the symbiotic relationship between nitrogen-fixing soil bacteria and legumes (Satgé et al. 2016). Active DNA demethylation has also been shown to be involved in defense against bacterial infections. When the pathogen *Pseudomonas syringae* infected Arabidopsis plants, the *ros1* mutant plants (but not the *dml2* or *dml3* mutants) experienced increased bacterial multiplication and propagation (Yu et al. 2013). Furthermore, the importance of active DNA demethylation in defense against infections by pathogenic fungi has been demonstrated in multiple studies. Arabidopsis *rdd* mutants showed heightened susceptibility to the hemi-biotrophic fungal pathogen *Fusarium oxysporum* and downregulation of various plant response stress genes that contain TE sequences in their promoters, some of which were hypermethylated (Le et al. 2014).

Active DNA demethylation is also involved in abiotic stress, such as drought, salinity, temperature, and nutrient deficiencies (Zhang et al. 2022). By analyzing the DNA methylome of plants that were exposed to drought or high-salinity treatment, it was discovered that the overall DNA methylation levels were higher compared to those in wild type plants (Ganguly et al. 2017). Although the majority of DNA methylation changes remain poorly understood,

drought-induced DMRs in Arabidopsis are mostly observed in the CHH context. These changes correlate with the accumulation of small interfering RNAs (siRNAs), suggesting that RdDM increases DNA methylation in plants exposed to drought (Ganguly et al. 2017, Van Dooren et al. 2020, Zhang et al. 2022). The DNA methylome of Arabidopsis seedlings was analyzed over time to understand the effects of heat treatment (42 °C for 24 hours) followed by recovery (up to 48 hours). The findings revealed a gradual decrease in DNA methylation levels in all sequence contexts (Korotko et al. 2021). Yang et al. (2022) evaluated the impact of cold treatment in *ros1* mutants and wild-type plants. Compared to the wild type, the cold-treated *ros1* mutants showed more severe leaf deformities and higher anthocyanin levels (Yang et al. 2022). Nutritional stress has been the focus of research investigating the influence of methylation/demethylation processes. In Arabidopsis, moderate changes in total methylation were observed during zinc starvation, resulting in the hypo- and hyper-methylation of TEs, genes, and promoters mainly in CG and CHG contexts. These methylation changes are enriched in the proximity of transcriptionally responsive genes (Chen et al. 2018).

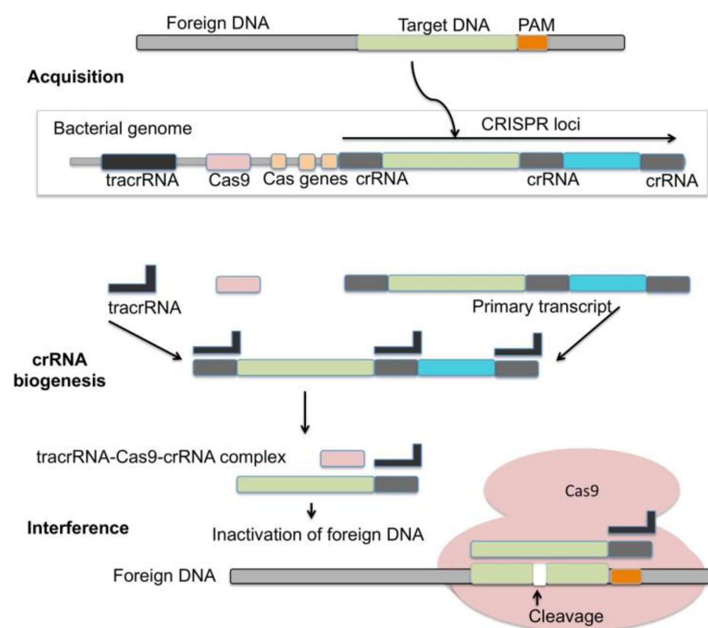
5.5. **Genome editing with the CRISPR/Cas9 system**

Genome editing is the process of directly altering an organism's DNA in order to modify its genetic code. CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is a revolutionary genome editing tool that allows scientists to make precise, targeted changes to the DNA of living organisms, including plants, animals, and humans. The CRISPR-Cas9 system works by using a specific RNA molecule, called a guide RNA, to find and bind to a specific target site in the DNA. Then, the Cas9 enzyme acts as a pair of "molecular scissors" to cut the DNA at the desired location, allowing for precise editing (Ding et al. 2016).

The Clustered Regularly Interspaced Palindromic Repeat Sequences (CRISPR/Cas) system was first reported in 1987 by Ishino *et al*. during their study on the *iap* gene in *E. coli*. They reported that the genome of *E. coli* contains a 29-nucleotide palindromic repeat sequence (Ishino et al. 1987). Numerous similar repeats were also discovered on further investigation and sequencing of other bacteria and archaea strains. These sequences were first given the name of clustered repeat elements. In 2002, the term CRISPR was first adopted by Mojica and Jansen to address this unique family of interspaced repeat sequences (Jansen et al. 2002). The in-depth study of the sequences revealed the spacer sequences to be different from the palindromic repeats. In 2005, the spacer sequences were separated from direct repeats and its sequence was associated with the phage genome or extrachromosomal DNA. CRISPR/Cas systems are highly

diverse adaptive immune systems on the basis of various Cas proteins, used by most archaea and many eubacteria to protect themselves from invading bacteriophages and mobile genetic elements. These are highly conserved sequences in which the size of CRISPR repeats varies between 23–47 bp and the size of spacers varies between 21–72 bp (Zheng et al. 2020). The genome of bacteria may contain one or more than one CRISPR locus. These loci have different and highly irregular spacer sequences even amongst closely associated species. These sequences are used to recognize the invading viral/plasmid genome. New matching viral or plasmid genomes are recognized and destroyed by adding new spacers after the infection by phage or archaea. These systems permit the cell to identify and differentiate the foreign DNA from self-DNA. CRISPR actively requires the presence of 6–20 CRISPR associated (Cas) genes present adjacent to the spacer sequence that codes for Cas proteins. This system is now used by scientists for biotechnological applications such as genome editing, and the production of phage resistant dairy culture (Hille & Charpentier 2016).

The mechanism by which the CRISPR/Cas system degrades the invader's genome



occurs through three steps i.e. adaptation or acquisition, biogenesis or expression of CRISPR RNA (crRNA) and Interference or targeting (Arora & Narula 2017) (Figure 3).

Figure 3. Mechanism of CRISPR/Cas9 action (Arora & Narula 2017).

In the acquisition phase foreign DNA gets incorporated into the CRISPR loci of bacterial genome. CRISPR loci is then transcribed into primary transcript and processed into crRNA with the help of tracrRNA during crRNA biogenesis. During interference, Cas9 endonuclease complexed with a crRNA and cleaves foreign DNA near PAM region.

The first step is the adaptation or acquisition of a spacer sequence. In adaptation, Cas proteins of the CRISPR system differentiate DNA as own or foreign by identifying the protospacer adjacent motif (PAM), is situated right beside the protospacer and is absolutely essential for adaptation and interference (Mojica et al. 2009, Zetsche et al. 2015). PAMs are 3–4 base pairs of nucleotide motifs which Cas proteins concede. On recognition of the foreign DNA, a few base pairs of nucleotides adjacent to the PAM motif are incorporated into CRISPR loci found within the pathogen. It leads to the development of remembrance of the foreign DNA within the microorganism (Hille & Charpentier 2016).

The integrated foreign sequence is now expressed in the form of RNA, and combined with tracrRNA (trans-activating CRISPR RNA) resulting in the formation of premature crRNA (pre-crRNA). After processing, the premature crRNA converts into mature crRNA which pilots the Cas endonuclease for introducing double-stranded breaks into the incoming DNA sequence (Haurwitz et al. 2010). Depending upon the specific function of crRNAs these are also known as guideRNAs. The incoming genome complementary to the completely mature crRNA and adjacent to the PAM motif would be susceptible to these double-strand breaks (Rath et al. 2015).

Interference is the last step in which crRNAs acknowledge and form a complementary base pair specific to foreign genome. This causes the destruction of the invading genome-crRNA complex by cleavage (Bhattacharjee et al. 2020). The Cas protein has two activities and acts as an endonuclease as well as helicase. The Cas nuclease in Cas9 cuts the invading genome at a specific DNA region upstream of the PAM. The PAM is a short DNA sequence present 3–4 nucleotides downstream of the region targeted for cleavage by the CRISPR system (Zheng et al. 2020). In the presence of any mutation in the PAM or discrepancy in complementarity between the spacer and invader's DNA, the severance is prohibited and is responsible for the susceptibility to infection (Hille & Charpentier 2016).

The RNA-guided Cas9 is a widely used CRISPR nuclease for DNA editing, and it functions by initiating double-stranded breaks (DSBs) at a specific genomic site with the help of a single-guide RNA (sgRNA) molecule (Anzalone et al. 2020). The CRISPR-Cas system has completely changed genome editing by providing a highly accurate and potent method for modifying DNA sequences. The production of a sgRNA by fusing tracrRNA and crRNA was one of the major advancements that opened up this technology to researchers (Bhatia et al.

2023). The sgRNA is a modified form of the naturally occurring crRNA-tracrRNA complex, which has been engineered into a continuous sequence. The sgRNA guides the Cas9 protein to bind and cleave both strands of the target sequence, leading to DSBs. Once the Cas9 cleaves the target site, the DNA damage repair process begins, with either the error-prone non-homologous end joining (NHEJ) pathway or the high-fidelity homologous direct repair (HDR) pathway repairing the DSBs (Boti et al. 2023).

The use of CRISPR/Cas9 based genome editing in plants has resulted in mutations that mostly involve short deletions of 10 base pairs or less and single nucleotide insertions, with a preference for A/T in all species. Although rare, single nucleotide substitutions were observed in soybean protoplasts, which had the highest mutation frequency. It appears that gene-specific factors may have an impact on DSB repair via microhomology mediated end joining (MMEJ), given that longer deletions were less frequent. In rice, the mutation patterns differ depending on the target (Bortesi et al. 2016), and consistent mutations have been observed across various independent cultures of soybean hairy-roots and somatic embryos. This suggests that there might be unknown regulations that determine the preferred types of mutations at a specific target. The study also revealed that the seven most effective sgRNAs resulted in only short deletions, while less efficient ones caused more insertions and substitutions (Jacobs et al. 2015). Interestingly, all the off-target mutations identified in rice had 1-base pair insertions, which suggests that the pairing of sgRNA with the target sequence could influence the type of mutation that occurs (Li et al. 2016). Higher plants generally exhibit low incidences of off-target activity, and when it occurs, only a minority of sgRNAs are typically affected. Whole-genome sequencing has confirmed that the number of unintended mutations (off-targets) is infrequent. Studies of rice and wheat have shown that off-target effects can be minimized by designing sgRNAs with mismatches outside the seed sequence. Creating more precise sgRNAs is therefore an effective strategy to avoid these types of events (Bortesi et al. 2016). In some plants like Arabidopsis (Sauer et al. 2016), barley (Lawrenson et al. 2015), soybean (Jacobs et al. 2015), and rice (Xie & Yang 2013), unexpected cutting has been observed at sites with mismatches within the seed region (Bortesi et al. 2016). This may be more likely to occur in target sequences with a high GC content (above 70%), as observed in (Li et al. 2016) research, but not in studies with lower GC content. Despite this, off-target mutations occur less frequently than on-target mutations, so it is still possible to retrieve only on-target mutations in all trials. Furthermore, it was found that off-target mutations only occurred in T1 rice plants with Cas9/sgRNA transgenes, but not in those where the CRISPR components had segregated,

suggesting that selecting appropriate T1 offspring could minimize or prevent off-target effects (Xu et al. 2015). The probability of unwanted mutations depends on the concentration of the Cas9/sgRNA RNP complex, so transient expression of the components may reduce the likelihood of off-target effects, but this may come at the expense of on-target efficiency (Bortesi et al. 2016, Tsai et al. 2015).

One way to detect mutations resulting from Cas9/sgRNA is by using a restriction enzyme site-based assay, which can be designed effectively. In the presence of a restriction enzyme site in the target sequence, the Cas9 cleavage and NHEJ DNA repair can cause specific mutations and destroy the restriction enzyme site. This makes it possible to identify mutated sequences through the digestion of genomic DNA templates or PCR amplicons with the corresponding restriction enzyme (Voytas 2013). This method has been used successfully to detect targeted mutations and evaluate the effectiveness of the editing process in various scenarios (Gao et al. 2015, Nekrasov et al. 2013). Nonetheless, it is crucial to note that this technique necessitates the presence of a restriction enzyme site in the target core sequences. Alternatively, the Surveyor nuclease and T7 Endonuclease I assays can also be utilized in a comparable manner (Guschin et al. 2010). To determine mutations induced by CRISPR/Cas9, the single-strand conformation polymorphism method can also be used. This technique works by identifying changes in the shape of single-stranded DNAs that contain nucleotide variations, leading to different migration rates on a non-denaturing PAGE gel. There is also another PAGE-based approach that can detect heteroduplex DNAs that carry targeted mutations (Zhang et al. 2017).

5.6. Using of genome editing to improve agricultural traits in plants

In order to perform genome editing within a living organism, it is necessary to introduce constructs containing the Cas9 and sgRNA expression cassettes into the cells of the plant. For many types of plants, the most efficient approach for this delivery is through Agrobacterium-mediated transformation. The predominant technique for delivering the CRISPR/Cas9 system into plants involves using T-DNA, which carries both the Cas9 and sgRNA expression cassettes, and is directly delivered into the plant genome through a type IV secretion mechanism. To initiate the process of genome editing in plants, a vector containing the sgRNA, selectable marker, and Cas9 gene is introduced into the plant material using well-established transformation techniques. The resulting T0 plants are then examined to detect any mutations that may have occurred at the target site (Liu et al. 2017) . It has been demonstrated that the

editing system can be effectively delivered to different plant species using various methods, and that the success rates of transformation vary. Barley plants (*Hordeum vulgare* cv. "Golden Promise") showed a higher number of genetically modified individuals containing the CRISPR-Cas9 reagents when they were transformed using *A. tumefaciens*, as opposed to the technique of particle bombardment (Kapusi et al. 2017). On the other side, wheat transformation has been a difficult process due to its low transformation efficiency. The success of the transformation process depends on several factors such as the type of genotype chosen, quality of the immature embryo, composition of the media, *A. tumefaciens* strain, embryo pre-treatment, and tissue handling. To overcome this challenge, researchers have optimized these factors in wheat transformation and achieved a transformation efficiency rate of up to 90% using the Agrobacterium-mediated genetic transformation method (Montecillo et al. 2020).

An alternative to Agrobacterium-mediated transformation is the Biolistic method, the delivery of CRISPR-Cas9 DNA into wheat immature embryo and callus cells using particle bombardment has resulted in significant success. This method has achieved transient expression of CRISPR-Cas9 reagents and has led to the production of transgene-free homozygous transformed plants (Zhang et al. 2016). When attempting to replace genes using CRISPR-Cas9 and delivering DNA templates alongside the CRISPR-Cas9 reagents, the biolistic transformation method is typically favored over Agrobacterium-mediated transformation. This preference is due to the biolistic method's ability to introduce numerous copies of donor DNA templates into the host, which enhances the probability of gene replacement via homologous recombination (HR) (Baysal et al. 2016). By utilizing the biolistic delivery method to introduce Cas9 genes and sgRNA genes, successful gene editing and replacement via HR was recently achieved in maize and soybean (Weeks et al. 2016).

Compared to traditional delivery methods, using viruses to deliver sgRNA offers several advantages. The first advantage is the ability to achieve high levels of sgRNAs due to the viral particle's capacity for self-replication and systemic movement throughout the plant system. The second advantage is a reduction in the time and complexity of the transformation and regeneration process, resulting in less labor-intensive operations. The third advantage is the ability to perform *in planta* transformation using this system. Finally, this method has the potential to produce transgene-free edited plants (Mahas et al. 2019). Researchers have engineered multiple RNA and DNA viruses to assist with genome editing in plants. A plant DNA virus called the cabbage leaf curl (CaLCuV) geminivirus has been found to effectively deliver sgRNAs in *Nicotiana benthamiana (N. benthamiana)*. Additionally, other viral

replicons from geminiviruses, such as yellow dwarf virus (BeYDV), wheat dwarf virus (WDV), and tomato leaf curl virus (ToLCV) have been modified to target genes in potato, wheat, rice, and tomato. Similarly, for instance, RNA viruses like tobacco rattle virus (TRV), tobacco mosaic virus (TMV), pea early browning virus (PEBV), and beet necrotic yellow vein virus (BNYVV) have been modified to deliver sgRNAs to various plants, including *N. benthamiana*, *Arabidopsis*, and *Beta macrocarpa*. In addition, barley stripe mosaic virus (BSMV), another RNA virus, has been employed as a vector for editing agents in wheat and maize. These virus-based vectors have been proven to be effective in delivering multiple sgRNAs for genome editing (Montecillo et al. 2020). Utilizing viruses to deliver genome editing reagents may result in edited plant varieties that are free of transgenes, particularly when RNA viruses are used. This is because RNA viruses don't integrate with the plant's genome as opposed to DNA viruses, preventing undesired integration into the genetic material (Gil-Humanes et al. 2017, Mahas et al. 2019). The use of viral systems to deliver genome editing agents is constrained by the size of the nucleic acid cargo they can carry. Geminivirus genomes, for example, are generally only around 3 kb in size, while the Cas9 coding sequence from *S. pyogenes* alone is over 4 kb. Consequently, viral delivery methods are most practical for transporting gRNA(s) to plants that have already established stable Cas9 expression. Despite the cargo size limitations, viral approaches are still useful for quickly screening for reverse genetic changes in plants (Liu et al. 2017).

Another delivery system is *de novo* meristem induction. Currently, tissue culture techniques limit the development of transgenic and genome-edited plants to a few plant species. However, recent efforts to enhance tissue culture involve the use of developmental regulators that influence the meristem's development and identity. By exploiting the totipotent feature of plant cells, ectopic expression of a combination of these regulators in somatic cells can induce meristem formation, as observed in Arabidopsis. For instance, when *WUSCHEL* (*WUS*), *SHOOT MERISTEMLESS* (*STM*), and a variant of *MONOPTEROS* (*MP*) were expressed in leaf cells, meristems were formed (El-Mounadi et al. 2020, Montecillo et al. 2020). A recent study by Maher et al. in 2020 demonstrated that *de novo* meristem induction could generate CRISPR-Cas9 genome-edited dicot plants, bypassing the requirement for tissue culture (Maher et al. 2020). In addition, it has been observed that *de novo* meristem induction can occur in soil-grown plants after transformation. To conduct CRISPR-Cas9-mediated genome editing, transgenic seedlings and soil-grown plants with constitutive expression of Cas9 were transformed with a suitable combination of developmental regulators (DRs) and sgRNAs that

target the *PHYTOENE DESATURASE* (*PDS*) gene. Inactivation of the *PDS* gene causes chlorophyll photobleaching, which makes it easy to identify the mutant phenotype. The utilization of DRs to produce *de novo* gene-edited meristem is a promising technique that can be applied to broaden *in planta* transformation to various plant species, ultimately leading to the efficient and speedy production of genome-edited plants (El-Mounadi et al. 2020, Montecillo et al. 2020).

Similarly, another method known as Cas9-sgRNA Ribonucleoproteins (RNPs), to create the functional RNPs in this system, the expression cassettes for Cas9 and gRNA are replaced with purified Cas9 nuclease and hybridized synthetic gRNA. These components are mixed *in vitro*, resulting in the formation of functional RNPs that can be delivered into the host through conventional transformation methods. By utilizing pre-assembled Cas9-gRNA RNPs, the potential risk of introducing foreign DNA into the host genome can be minimized, as compared to their plasmid-encoded counterparts. Furthermore, using pre-assembled RNPs can reduce the need for extensive screening of progenies and backcrossing (Cho et al. 2013). according to multiple studies, various types of plants, such as *A. thaliana,* lettuce*, Nicotiana attenuata,* rice, maize, petunia, grapevine, apple, wheat, and potato have been successfully treated with the DNA-free CRISPR-Cas9-based system (Montecillo et al. 2020). One of the benefits of this DNA-free system is that it doesn't require DNA constructs, which means there's no need for creating expression cassettes or plasmid vectors, optimizing codon usage, or identifying promoters to drive the expression of Cas9 and gRNAs. As a result, this system can be easily applied to different organisms without requiring complex adaptation procedures. While the DNA-free system offers a new way to achieve transgene-free genome editing using CRISPR-Cas9, it still faces certain challenges. Specifically, its editing efficiency is inferior to the classical DNA-based method, and the absence of a selection marker makes the selection process after transformation difficult. Moreover, the validation of successfully edited plants can only be achieved through molecular approaches or phenotype screening (Metje-Sprink et al. 2019).

Genome editing can be used to improve agriculture by creating crops with desirable traits, such as increased yield, disease resistance, and improved nutrient content. To enhance plants' resistance to viruses, two strategies have been employed. The first strategy involves the design of sgRNAs to target the virus genome, while the second strategy involves modifying plant genes that are involved in the antiviral pathway. The binding of virus genome-linked protein (VPg) to the eukaryotic translation initiation factor 4E (eIF4E) plant protein plays a critical role in the infection of plants by the potato virus Y (PVY). Plant resistance to this virus

can be achieved by altering a key site of eIF4E, which affects the virus-plant interaction (Wan et al. 2021). In cucumbers, the use of CRISPR-Cas to target the N- and C-terminal ends of eIF4E in non-transgenic homozygous plants of the T3 generation led to immunity against cucumber vein yellowing virus (CVYV) and zucchini yellow mosaic virus (ZYMV), as well as resistance to papaya ringspot virus (PRSV-W) (Chandrasekaran et al. 2016). The use of CRISPR-Cas9 technology was effective in generating mutations in both the coding and noncoding regions of geminiviruses, leading to a decrease in its pathogenicity. In the case of *N. benthamiana*, sgRNA-Cas9 constructs were utilized to target beet severe curly top virus (a geminivirus), resulting in inhibition of the virus accumulation in the leaves (Ji et al. 2015). Mutations in the noncoding regions of the geminivirus are thought to be more effective in reducing or eliminating its ability to replicate than mutations in the coding regions. Furthermore, noncoding-region mutations generate fewer viral variants compared to coding-region mutations (Ali et al. 2016, Ji et al. 2015). Similarly, to provide resistance against fungi by the application of CRISPR-Cas9 gene editing in tomatoes has revealed that knocking out the *Solyc08g075770* gene can make the plants more susceptible to *Fusarium*-mediated wilt disease (Prihatna et al. 2018). Meanwhile, in watermelons, the knockout of the *Clpsk1* gene, which produces the *PHYTOSULFOKINE* (*PSK*) precursor, has been shown to increase the plants' resistance to *Fusarium oxysporum* f. sp. *niveum* (FON) (Zhang et al. 2020). Furthermore, DOWNY MILDEW RESISTANT (DMR6) is a type of protein found in *Arabidopsis thaliana* that belongs to the Fe(II)-dependent superfamily of 2-oxoglutarate oxygenases. Its function is to regulate the levels of salicylic acid in the plant. When DMR6 is overproduced in plants, it can provide some protection against downy mildew (Zeilmaker et al. 2015). In experiments where the *DMR6* gene was knocked out in tomato plants using CRISPR-Cas9, the resulting mutation showed resistance to several plant pathogens, including *Pseudomonas syringae, Phytophthora capsici,* and *Xanthomonas spp* (Thomazella et al. 2016). The bacterial pathogen *Pseudomonas syringae* is known to cause damage to important crop plants around the world. Upon infection, this pathogen releases coronatine (COR) which triggers stomatal opening, allowing the bacteria to invade. This process requires a coreceptor called JAZ2. Researchers have used CRISPR/Cas9 gene editing to create a modified version of the *JAZ2* gene, called *SlJaz2 Δjas*, by removing the Jas domain located at the end of the gene. This modification has resulted in the production of plants that are resistant to bacterial speck disease caused by *P. syringae* (Salava et al. 2021).

The growth and yield of most horticultural crops are significantly impacted by environmental factors such as high temperatures, drought, and salinity, which are known as abiotic stresses. Although conventional plant breeding has been attempted to mitigate the effects of these stresses, it has not yet been very effective (Parmar et al. 2017). The advent of CRISPR-Cas9 genetic modification has reduced the duration of creating new types (Wan et al. 2021). *BZR1*, a gene that is involved in numerous brassinosteroid (BR)-related developmental processes, plays a part in them. By means of CRISPR, a genetic mutation was produced in *BZR1* which hindered the activation of *RESPIRATORY BURST OXIDASE HOMOLOG1* (*RBOH1*) and the creation of H_2O_2. Externally applied H_2O_2 restored the tomato *bzr1* mutant's ability to endure heat (Yin et al. 2018). In addition, gene-editing techniques can be employed to develop new strains of plants that are resistant to cold and drought conditions. For example, by editing the *CBF1* (*C-REPEAT BINDING FACTOR 1*) gene, which plays a key role in regulating a plant's ability to withstand cold temperatures, or the *MAPK3* gene, which is involved in the plant's response to drought stress and helps protecting cell membranes from peroxidative damage in tomatoes, scientists can create germplasms that are more tolerant to adverse environmental conditions (R. Li et al. 2018). Using the CRISPR/Cas9 approach, scientists developed mutations in rice that could be used to examine how STRESS/ABA-ACTIVATED PROTEIN KINASE 2 (SAPK2) contributes to stress tolerance. Similarly, in Arabidopsis plants, the same gene-editing technique was employed to generate double and triple mutants (*cbf1/3* and *cbf1/2/3* or *cbfs*) that would provide insights into the functions of C-repeat binding factors (CBFs) in response to cold stress. These mutants were obtained by introducing mutations to the *CBF1* and *CBF2* genes in a *cbf3* T-DNA insertion mutant (Bao et al. 2019). Knowledge of gene functions enables breeders to use specific genes to develop crops that are better equipped to handle environmental stresses. To create maize plants that are more drought-tolerant, Shi and colleagues employed this approach. The maize *ARGOS8* gene typically has low expression levels and is involved in regulating ethylene responses. By introducing the *GOS2* promoter sequence into the 5’-UTR of the native *ARGOS8* gene or replacing the native promoter with the *GOS2* promoter using the CRISPR/Cas9 system, the researchers were able to increase *ARGOS8* expression. These modified *ARGOS8* variants displayed improved yields under drought conditions in field experiments (Shi et al. 2017).

The main aim of breeding fruits and vegetables is to enhance their quality and extend their shelf life after they are picked. For instance, in tomatoes, the number of ovary locules is a major determinant of fruit size and is influenced by multiple quantitative trait locus (QTLs). At

Cold Spring Harbor Laboratory, scientists used CRISPR-Cas9 technology to alter the promoter region of the tomato stem cell gene *CLAVATA* (*CLV3*) using eight sgRNAs. The aim was to produce fruits that are bigger and more plentiful compared to wild-type fruits. By modifying QTLs that affect fruit size, such as locule number and fasciated number, researchers created germplasm resources that have a higher locule number (Wan et al. 2021). In fruit and vegetable production, it is important to breed crops with a longer shelf life. Scientists have used CRISPR technology to slow down fruit ripening by knocking out specific genes such as *RIPENING INHIBITOR* (*RIN*) or *DNA DEMETHYLASE 2* (*DML2*). However, this approach has also been found to negatively affect the taste and nutritional value of the fruit as well as alter its peel color (Lang et al. 2017b). In contrast, in tomatoes, the expression of *PECTATE LYASE* (*PL*) and *ALCOBACA* (*ALC*) genes were inhibited using CRISPR, which effectively extended shelf life without affecting the fruit's sensory qualities or nutritional value (Yu et al. 2017).

6. MATERIALS AND METHODS

6.1. Designing CRISPR guide sequences

The guide RNAs were designed with the CRISPOR server (http://crispor.tefor.net; Haeussler et al. 2016). For *SlDML1*, we designed two gRNAs in the protein coding part of the gene to increase the chance of getting a functional null mutant. The two guide RNAs were selected from the list of the possible sequences by their specificity score (this score is a prediction of how much an RNA guide sequence for this target may lead to off-target cleavage somewhere else in the genome. The score ranges from 0–100 with 100 being the best, meaning the search could not find a single sequence in the genome that differs from the target at up to four positions. Good guides should have a specificity score of at least 50), their possible off-targets (no off-targets within the seed region in case of 0 or 1 mismatch), and their distance from each other (they should be ~100–200 bp away from each other to ease the detection by PCR on an agarose gel in case of a large deletion between the two guides). The acceptable guides should have a 'G' at their 5' termini to be able to use with the U6 promoter, and should not contain 'UUUU' in their sequence that serves as a U6 termination signal.

6.2. Modular cloning of the CRISPR construct

To create a CRISPR-Cas9-based knock-out gene construct, we followed the protocol described by Rodríguez-Leal et al. (Rodríguez-Leal et al. 2017) utilizing modular cloning and the principle of GoldenGate assembly. For this, we used the MoClo Toolkit (Weber et al. 2011), the gRNA template pICH86966::AtU6p::sgRNA_PDS, the kanamycin resistance module pICH47732::NOSp-NPTII-OCST, and the Cas9 module pICH47742::2x35S-5'UTR-hCas9(STOP)-NOST from Addgene. The individual guide RNAs were PCR amplified with Phusion™ Hot Start II DNA Polymerase (Thermo Scientific, Waltham, MA, USA) using the template pICH86966_AtU6-sgRNA-PDS in two pieces: the U6 promoter was amplified using the common forward (gRNA_BsaI_F: 5'- TGG TCT CTG GAG ATG ATC AAA AGT CCC ACA TCG ATC -3') and reverse primer (U6pro_BsaI_R: 5'- TGG TCT CTA ATC GCT ATG TCG ACT CTA TCA -3') carrying *Bsa*I recognition sites (5′- GGTCTC(N1)/(N5) -3′) at their 5' termini resulting in a 96 bp product. The second half (the actual gRNA) was amplified with a common reverse primer matching the U6 terminator (gRNA_BsaI_R: 5'- TGG TCT CTA GCG CTA GAA AAA AAG CAC CGA CTC G -3') and a specific forward primer having the 20 bp target-specific guide sequence (underlined part of the sequence) and a *Bsa*I recognition site at the 5' termini (SlDML1_gRNA1_F: 5'- TGG TCT CTG ATT <u>GTC GCG TTA ATG</u>

TCT AGA CCG TTT TAG AGC TAG AAA TAG CAA G -3’ and SlDML1_gRNA2_F: 5’-TGG TCT CTG ATT GAT TGA CAA GGT AGT CGA TCG TTT TAG AGC TAG AAA TAG CAA G -3’). The temperature program of the PCR is described in Table 1.

Table 1. Temperature program of the PCR to amplify the guide RNA modules.

Step	Temperature	Time	Number of cycles
Initial Denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	
Primer Annealing	55 °C	10 s	30
Extension	72 °C	5 s	
Final Extension	72 °C	1 min	1
Storage	12 °C	∞	1

After purification of the PCR products from agarose gel using the GeneJET Gel Extraction Kit (Thermo Scientific, Waltham, MA, USA), the two PCR products were directly cloned into a suitable Level 1 vector (pICH47751 for gRNA1 and pICH47761 for gRNA2) in a 20 µL Golden Gate reaction using 1.2 µL *Bsa*I-HF®v2 (New England Biolabs, Ipswitch, MA, USA) and 0.4 µL T4 DNA ligase (New England Biolabs, Ipswitch, MA, USA) simultaneously in T4 ligase buffer. 100 ng was used for each plasmids. The temperature program of the Golden Gate reaction is described in Table 2.

Table 2. Temperature program of the Golden Gate assembly to create Level 1 modules.

Step	Temperature	Time	Number of cycles
Digestion and ligation	37 °C	1 h	1
Final digestion	55 °C	5 min	1
Inactivation	80 °C	5 min	1
Storage	12 °C	∞	1

2 µL of the assembly products were transformed into 100 µL *E. coli* DH5α competent cells following the standard protocol (Inoue et al. 1990) and plated on LB plates supplemented with ampicillin (50 µg/mL), IPTG (6 µg/mL) and X-gal (32 µg/mL) and grown overnight at 37

°C. Four white colonies were inoculated into 5 mL liquid LB media supplemented with ampicillin and grown overnight at 37 °C. Next day, plasmids were purified from the bacteria following the standard alkalic lysis protocol (Sambrook & Russell 2001). 1 µg of the plasmids were test-digested with *Bbs*I enzyme in a 10 µL reaction mixture following the manufacturer's instructions. The correct clones had 207 and 4352 bp fragments. The sequence of the constructs were checked by Sanger sequencing.

The Level 1 modules were used in a second assembly reaction together with pAGM4723 Level 2 vector, the Kan^R, and Cas9 modules (in pICH47732 and pICH47742, respectively), and a suitable end-linker (pICH79289) using *Bsa*I-HF®v2, *Bbs*I (New England Biolabs, Ipswitch, MA, USA), and T4 DNA ligase simultaneously in T4 ligase buffer. 100 ng was used for each plasmids. The temperature program for the Level 2 Golden Gate reaction is described in Table 3.

Table 3. Temperature program of the Golden Gate assembly to create the final product.

Step	Temperature	Time	Number of cycles
Digestion	37 °C	1 min	30
Ligation	16 °C	1 min	
Final digestion	55 °C	5 min	1
Inactivation	80 °C	5 min	1
Storage	12 °C	∞	1

2 µL of the 20 µL assembly products were transformed into *E. coli* as described above but in this case the cells were plated on LB plates supplemented with kanamycin (50 µg/mL). Four white colonies were inoculated in liquid LB media supplemented with kanamycin and grown overnight. Plasmids were purified as described above. The plasmids were digested with *Eco*RI, the correct clones should produce fragments with the following sizes: 207, 593, 1891, 4678, and 5063 bp. One positive clone was Sanger sequenced to verify the sequence. The whole cloning procedure was simulated with the SnapGene software (www.snapgene.com).

6.3. Agrobacterium transformation

The correct clones were transformed into *Agrobacterium tumefaciens* strain LBA4404 and used for transformation of the tomato variant Moneymaker following the procedure

described earlier (Chen et al. 1994). Briefly, Agrobacteria were grown in YEB medium supplemented with rifampicin (50 μg/mL) and tetracycline (12.5 μg/mL) at 30 °C overnight in a shaker, then 2 mL of this starter culture was inoculated into 50 mL YEB medium (without antibiotics) and was grown for four hours until $OD_{600} \cong 0.5$. The culture was transferred into 50 mL Falcon tubes, chilled on ice for ten minutes and then the tubes were centrifuged at 3000 rpm for ten minutes in a refrigerated centrifuge. The pelleted cells were resuspended in 1 mL ice-cold 20 mM $CaCl_2$, and 100 μL aliquots in cold 1.5 mL Eppendorf tubes were prepared for the transformations. To transform the plasmids into the cells, 1 μg plasmid DNA was mixed with 100 μL competent cells and freezed in liquid nitrogen for five minutes. Then, the cells were thawed at 37 °C for five minutes and after adding 1 mL YEB medium (without antibiotics), they were incubated for two hours at 30 °C with shaking at 200 rpm. After that, cells were collected by centrifuging the tubes at 3000 rpm for ten minutes, 1 mL of the supernatant was removed, the cells were resuspended in the remaining solution and plated onto a YEB plate supplemented with rifampicin (50 μg/mL), tetracycline (12.5 μg/mL), and kanamycin (50 μg/mL).

6.4. Tomato transformation

Tomato transformation was carried out as described earlier (Fernandez et al. 2009). Briefly, small pieces of tomato (*Solanum lycopersicum* cv. Moneymaker) cotyledons were cut out and placed on KCMS (Murashige Skoog solution containing vitamins, hormones, and the Agrobacterium inducer acetosyringone) plates overnight, at 25 °C under low light conditions. 100 mL Agrobacterium overnight cultures were also prepared in YEB solution supplemented with kanamycin (50 μg/mL), rifampicin (50 μg/mL), and tetracyline (12.5 μg/mL). The bacteria were grown until $OD_{600} \cong 1$, pelleted by centrifugation, and resuspended in KCMS solution to $OD_{600} = 0.05$. The cotyledon pieces (explants) were placed into the Agrobacterium suspension and shaken gently for 30 min at 25 °C. After this, the explants were removed, blotted on paper filters, and placed on KCMS plates abaxial side upward (30–40 explants per plate). The plates were incubated in the dark at 25 °C for two days. After this, explants were transferred to 2Z plates, a shoot regeneration medium containing the plant hormone zeatin riboside (2 μg/mL) and the antibiotic timentin (250 μg/mL) to kill Agrobacteria, and grown for two weeks under long-day conditions (16 h light at 25 °C, 8 h dark at 18 °C). The explants were transferred to fresh 2Z plates every two weeks. The healthy-looking shoots that were formed from callus tissues were cut and transferred to glass jars containing a root regeneration medium and grown for six weeks under the same conditions as described above until healthy roots were seen. The

plants were carefully removed, the roots washed with water and transferred into pots filled with soil and grown in controlled lightroom until genotyping was done and then the mutants were transferred to the greenhouse.

6.5. Plant genomic DNA purification

To isolate genomic DNA suitable for PCR, 10 mL extraction buffer was prepared (for 12 samples) by mixing 5 mL 2× Isolation Buffer (100 mM Tris-HCl pH 9.5, 600 mM NaCl, 40 mM EDTA, 1% SDS, 4% sarkosyl, 2% PVP40), 5 mL 10 M urea, and 100 µL 2 M $Na_2S_2O_5$. Next, a leaf disc was punched from a young leaf and placed in a mortar with 700 µL of the extraction buffer. The leaf was then homogenized with a pestle at room temperature for about one minute and the homogenate was poured into a 1.5 mL Eppendorf tube pre-filled with 300 µL phenol (equilibrated with 100 mM Tris-HCl pH 8.0) and 300 µL chloroform:isoamyl alcohol (24:1) and mixed thoroughly. The mixture was centrifuged at maximum speed for five minutes to separate the organic and water-soluble phases. 500 µL of the upper (water-soluble) phase was transferred to a new tube filled with 250 µL phenol and 250 µL chloroform:isoamyl alcohol and mixed thoroughly. The mixture was centrifuged at maximum speed for five minutes. 500 µL of the upper phase was transferred to a new tube and mixed with 100 µL 3 M Na-acetate pH 5.2 and 400 µL isopropanol to precipitate nucleic acids. After five minutes of incubation at room temperature, the tubes were centrifuged for five minutes. The pellet was washed twice with 1 mL 70% ethanol to remove salts and phenol traces, dried in a Speedvac, and dissolved in 50 µL TE (10 mM Tris-HCl pH 7.5, 1 mM EDTA) buffer. To remove RNA from the DNA extract, 50 µL RNase solution (49 µL TE, 1 µL 10 mg/mL RNAse A solution) was added to the 50 µL sample. After an hour of incubation, the mixture was phenol-chloroform extracted and precipitated by isopropanol as described above with the volumes adjusted (50–50 µL phenol and chloroform:isoamyl alcohol, 20 µL Na-acetate, 80 µL isopropanol). The pellet was washed twice with 70% ethanol and dried in a speedvac before being dissolved in 50 µL TE pH 7.5. Finally, 2 µL of the sample was run on an 1% agarose gel to check the integrity of the genomic DNA and the absence of RNA. After that, samples were measured with Nanodrop to check the purity and quantity of the DNA. The DNA solutions were then diluted to 50 ng/µL with TE pH 7.5 of which 2.5 µL was used for a 50 µL PCR.

6.6. **Genotyping PCR**

The high-throughput genotyping PCR was carried out using the DreamTaq Green PCR Mix (Thermo Scientific, Waltham, MA, USA) in a 10 µL of reaction mixture per sample consisting of 5 µL of Master Mix, 0.2 µL of forward primer (SlDML1_gtPCR_F: 5'- ATG AGG ACT CGA GCA GCA AC -3'), 0.2 µL of reverse primer (SlDML1_gtPCR_R: 5'- TCC CAT TGA ACT GTG CTG CT -3'), 3.6 µL of nuclease free water, and 1 µl of DNA sample (50 ng/µL). The expected PCR product was 380 bp-long in case of a wild-type allele. For the PCR, the temperature program described in Table 4 was followed.

Table 4. Temperature program of the genotyping PCR.

Step	Temperature	Time	Number of cycles
Initial Denaturation	95 °C	1 min	1
Denaturation	95 °C	30 s	40
Primer Annealing	60 °C	30 s	
Extension	72 °C	1 min	
Final Extension	72 °C	5 min	1
Storage	12 °C	∞	1

All 10 µL of the PCR were run in a 2% agarose gel along with the size marker GeneRuler 1 kb Plus (Thermo Scientific, Waltham, MA, USA). The gel was documented using the ChemiDoc™ MP Imaging System and its own ImageLab™ v5.1 software (Bio-Rad Laboratories, Hercules, CA, USA).

For the restriction digestion test, to 5 µL of the PCR product 1 µL 10× FastDigest Green buffer, 8 µL water, and 1 µL *Xba*I FastDigest enzyme were added and the mixture was incubated for 1 h at 37 °C. All the 15 µL reaction mixture was then run in a 2% agarose gel.

6.7. **Cloning of the mutant alleles**

PCRs for cloning the mutant alleles was carried out using the Phusion™ Hot Start II DNA Polymerase (Thermo Scientific, Waltham, MA, USA) in a 50 µL of reaction mixture consisting of 25 µL of Phusion Master Mix, 0.4 µL $MgCl_2$ (to compensate for the EDTA in the DNA template), 2.5 µL of forward primer (SlDML1_gtPCR_F), 2.5 µL of reverse primer

(SlDML1_gtPCR_R), 9.6 µL of nuclease free water, and 10 µl of DNA sample (50 ng/µL). The temperature program is described in Table 5.

Table 5. Temperature program of the Phusion PCR.

Step	Temperature	Time	Number of cycles
Initial Denaturation	98 °C	3 min	1
Denaturation	98 °C	10 s	40
Primer Annealing	65 °C	30 s	
Extension	72 °C	10 s	
Final Extension	72 °C	1 min	1
Storage	12 °C	∞	1

To clone the blunt-ended PCR products, we used the CloneJET PCR Cloning Kit (Thermo Scientific, Waltham, MA, USA). The ligation reaction was carried out in 20 µL consisting of 10 µL of T4 DNA ligase buffer, 1 µL pJET1.2 vector, 8 µL PCR fragment, and 1 µL T4 DNA ligase (10 U/µL). The reaction was incubated for five minutes at room temperature, then 2 µL of the ligation mixture was transformed into 100 µL DH5α ultracompetent cells following the standard protocol (Inoue et al. 1990) then plated on LB plates supplemented with 100 µg/mL ampicillin. The plates were incubated at 37 °C overnight. Next day, colony PCRs were performed in the same way as in the case of genotyping PCR. Four positive colonies were inoculated into 5 mL liquid LB medium containing ampicillin and were shaken at 37 °C overnight. Next morning, plasmids were purified from the bacterial cells following the alkaline lysis method (Sambrook & Russell 2001). After measuring the concentration with a Nanodrop spectrophotometer, the samples were Sanger sequenced at a local sequencing facility. The obtained sequences were compared to the wild-type using Clustal Omega server (Madeira et al. 2022).

7. RESULTS AND DISCUSSION

7.1. Designing CRISPR guide sequences

Our goal was to create a functional null mutant of *SlDML1* because this gene is ubiquitously expressed suggesting that it might have an important function throughout the life cycle of the plant. Also, its high similarity to *SlDML2* could reflect an overlapping function regarding fruit ripening. One way to obtain such a null mutant is that we create deletions in the coding sequence (CDS) of the gene, usually in the first coding exon, that causes a frameshift mutation. This could result either in a truncated protein due to an early stop codon or a completely different, biologically meaningless peptide. Such proteins could confer a dominant negative effect by having some residual functions (i.e. binding to some partners) but lacking others (i.e. not be able to demethylate DNA). There are other strategies to create functional null mutants using the CRISPR/Cas9 genome editing system, for example, erasing the transcriptional start site(s) that could result in no expression of the gene at all. In this case, no truncated or abnormal protein is produced. However, alternative transcription start sites can be utilized in these cases jeopardizing the effort to get a null mutant.

To increase the chance to get a mutant, we decided to design two guide RNAs in the coding sequence of the *SlDML1* gene in a way that the two gRNAs were separated by more than a hundred base pairs. Earlier, we observed that large deletions could occur between such distantly spaced gRNAs. Beside increasing the chance of getting a null mutant, such large deletions could also ease the identification of mutants by simply running the genotyping PCR products on an agarose gel without any further treatments like restriction digestion. We designed the gRNAs using the CRISPOR server (http://crispor.tefor.net; Haeussler et al. 2016) (Figure 4 and 5).

Figure 4. The input screen of the CRISPOR server.

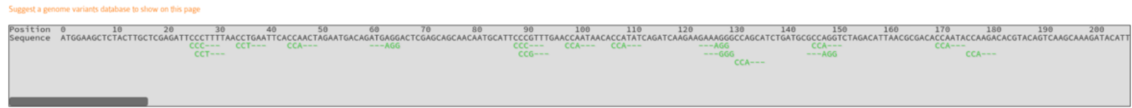
In Step 1 we need to provide an input sequence that we want to target, typically an exon that is less than 2300 bp long. In Step 2, we choose the genome version from which the sequence originates (the input sequence must perfectly match the genome). Lastly, in Step 3 we must specify the Cas enzyme and its PAM sequence to enable the server to predict possible target sites in the provided input sequence and off-target sites in the genome.

First, we specified the second exon of the *SLDML1* gene, because it contains the beginning of the coding sequence (CDS). The sequence derives from the 3.0 version of the tomato genome as can be found in the Solgenomics database, therefore we chose that one from the list. Finally, we selected the widely used Cas9 enzyme from *Streptococcus pyogenes* because the MoClo system that we wanted to use contained the precloned expression cassette of this enzyme. The output of the analysis is a table listing the possible gRNAs sorted by specificity (Figure 5). The content and the detailed explanation of the output table is given in the manual of the server (http://crispor.tefor.net/manual).

SlDML1: Solanum lycopersicum (faSolLyc3), SL3.0ch09:2410704-2412452, forward genomic strand

Your input sequence is 1749 bp long. It contains 157 possible guide sequences.
Shown below are their PAM sites and the expected cleavage position located -3bp 5' of the PAM site.
Click on a match for the PAM NGG below to show its 20 bp-long guide sequence. (Need help? Look at the CRISPOR manual)
Colors green, yellow and red indicate high, medium and low specificity of the PAM's guide sequence in the genome.

Suggest a genome variants database to show on this page



Download for: SerialCloner (free) - ApE (free) - GenomeCompiler - Benchling - SnapGene - Geneious - Vector NTI - LaserGene - Genbank - FASTA

Predicted guide sequences for PAMs

Ranked by default from highest to lowest specificity score (Hsu et al., Nat Biot 2013). Click on a column title to rank by a score.
If you use this website, please cite our paper in NAR 2018. Too much information? Look at the CRISPOR manual.

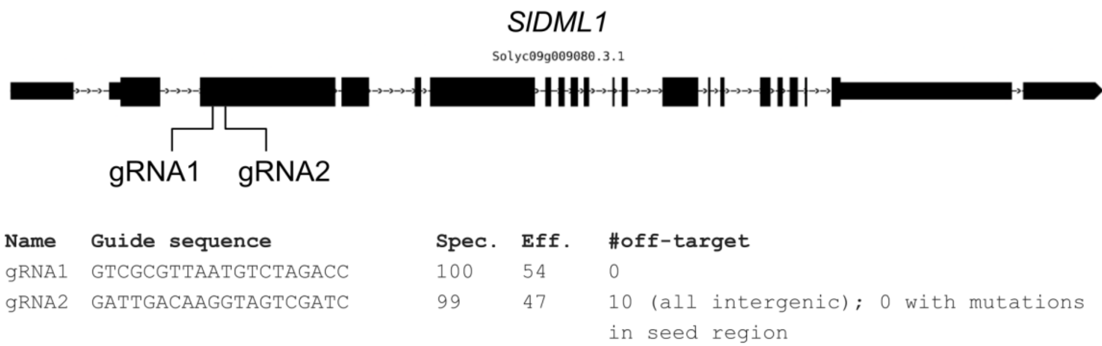
Download as Excel tables: Guides / Guides, all scores / Off-targets / Saturating mutagenesis assistant

Position/ Strand	Guide Sequence + *PAM* + Restriction Enzymes ☐ Only G- ☐ Only GG- ☐ Only A-	MIT Specificity Score	CFD Spec. score	Predicted Efficiency (Show all scores) Doench '16	Mor.-Mateos	Outcome Out-of-Frame	Lindel	Off-targets for 0-1-2-3-4 mismatches + next to PAM	Genome Browser links to matches sorted by CFD off-target score ☐ exons only ☐ SL3.0ch09 only
147 / rev	GTCGCGTTAATGTCTAGACC *TGG* Enzymes: *Hpy188III, HinP1I, XbaI, LpnPI, BstNI, StyD4I* Cloning / PCR primers	100	100	54	12	49	74	0-0-0-0-0 0-0-0-0-0 0 off-targets	
149 / fw	GGGCCAGCATCTGATGCGCC *AGG* Enzymes: *HinP1I, StyD4I, LpnPI, BstNI, MaeI* Cloning / PCR primers	100	100	54	54	49	77	0-0-0-0-1 0-0-0-0-0 1 off-targets	4:exon:mRNA:Solyc06g062870.3.1
678 / fw	CTCGACCAAGAAGATACCCC *AGG* Enzymes: *StyD4I, LpnPI, BstNI, NdeII, BseDI*	100	100	56	44	52	67	0-0-0-0-3 0-0-0-0-0	4:intergenic:mRNA:Solyc11g032090.2.1-mRNA:Solyc11g032100.2.1 4:intron:mRNA:Solyc04g007100.3.1 4:exon:mRNA:Solyc04g081820.3.1
313 / fw	GATTGACAAGGTAGTCGATC *AGG* Enzymes: *Hpy188III, TaqI, NdeII, BtsCI* Cloning / PCR primers	99	99	47	41	49	77	0-0-0-2-8 0-0-0-0-0 10 off-targets	4:intergenic:mRNA:Solyc03g044630.2.1-mRNA:Solyc03g044660.3.1 4:intergenic:mRNA:Solyc02g014140.1.1-mRNA:Solyc02g014150.3.1 4:intergenic:mRNA:Solyc12g042400.1.1-mRNA:Solyc12g042430.2.1 show all...

Figure 5. The output screen of the CRISPOR server.

The output contains some pieces of information about the input sequence, then we get a map of the possible guide sequences with their PAM and their orientation (whether they were found on the forward or reverse strand). Finally, a large table contains details of the guide sequences sorted by their specificity score. The predicted efficiency, the mutation characteristics, and the possible off-targets are also given. We have to select guides that start with a G nucleotide because we will use the U6 promoter to express the gRNA and the RNA Polymerase III that utilizes the U6 promoter needs a G at the start of the transcribed RNA. If the gRNA does not contain a G, we can add an extra G at the 5' end of the gRNA. The gRNA cannot contain a UUUU strech because it serves as a transcriptional terminator for the Pol III. Ideally, the gRNA does not have any off-targets (sites elsewhere in the genome).

When we tested the first coding exon for possible gRNAs, we could not find an acceptable pair, therefore, we tested the second coding exon. This time, we could select two good gRNAs that are separated by 163 nt (Figure 6).



Name	Guide sequence	Spec.	Eff.	#off-target
gRNA1	GTCGCGTTAATGTCTAGACC	100	54	0
gRNA2	GATTGACAAGGTAGTCGATC	99	47	10 (all intergenic); 0 with mutations in seed region

Figure 6. The location of the selected gRNA sequences within the *SlDML1* gene.
The black boxes represent the exons of the gene while the fishbone lines represent the introns. The thicker boxes mark the coding part of the gene, while the thinner parts show the untranslated regions. The location of the designed gRNAs within the third exon (second coding exon) are shown, as well as their sequence, specificity scores, efficiency scores, and the number of possible off-targets.

The first gRNA has a specificity score of 100 because it does not have any possible off-target sites, while the second one has a few possible off-target sites (all of them are in intergenic regions) only if we allow three or more mismatches in the guide sequence farther from the PAM sequence (Figure 5).

7.2. Modular cloning of the CRISPR construct

To clone the gRNA sequences, we followed the principle of the Golden Gate assembly method (Engler et al. 2008). This method utilizes Type IIS restriction enzymes that recognize asymmetric DNA sequences and cleave outside of their recognition sequence, usually within 1 to 20 nucleotides. This allows seamless stitching of DNA fragments in an oriented way since the enzyme cuts any sequences outside their recognition site, producing sticky ends depending on the orientation of the asymmetric recognition site. If the sites face toward the fragment, after cut and ligation, the original recognition sites will be lost (Figure 7).

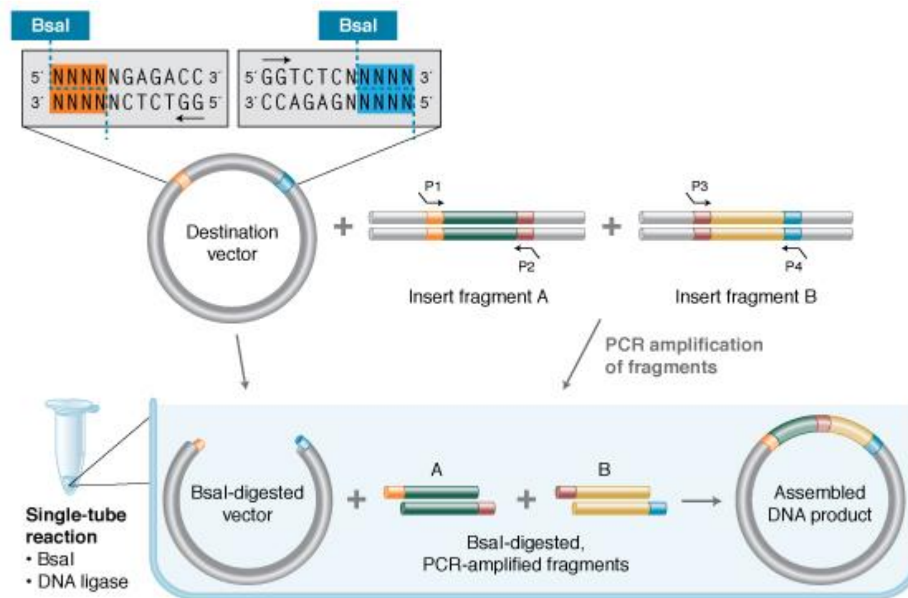


Figure 7. Principles of the Golden Gate assembly method.

The method is based on the unique features of Type IIS restriction enzymes: they have a nonpalindromic recognition sequence which provides directionality, and they cut outside their recognition site resulting in a four nucleotide sticky end of any sequence (NNNN highlighted with orange and blue colours showing that they have different sequences). Another feature of the Golden Gate method is that the restriction digestion and the ligation takes place simultaneously in one tube making this approach very efficient. Thanks to this, we can assemble large gene constructs consisting of several modules in one step.

Source: https://international.neb.com

Following this principle, first we amplified the U6 promoter (PCR A) and the rest of the gRNA scaffold (PCR B) using primers that had *Bsa*I recognition sites (5'-GGTCTC(N1)/(N5)-3') at their 5' ends and a template that has a complete gRNA expression module (pICH86966_AtU6-sgRNA-PDS) that we obtained from Addgene (Figure 8).

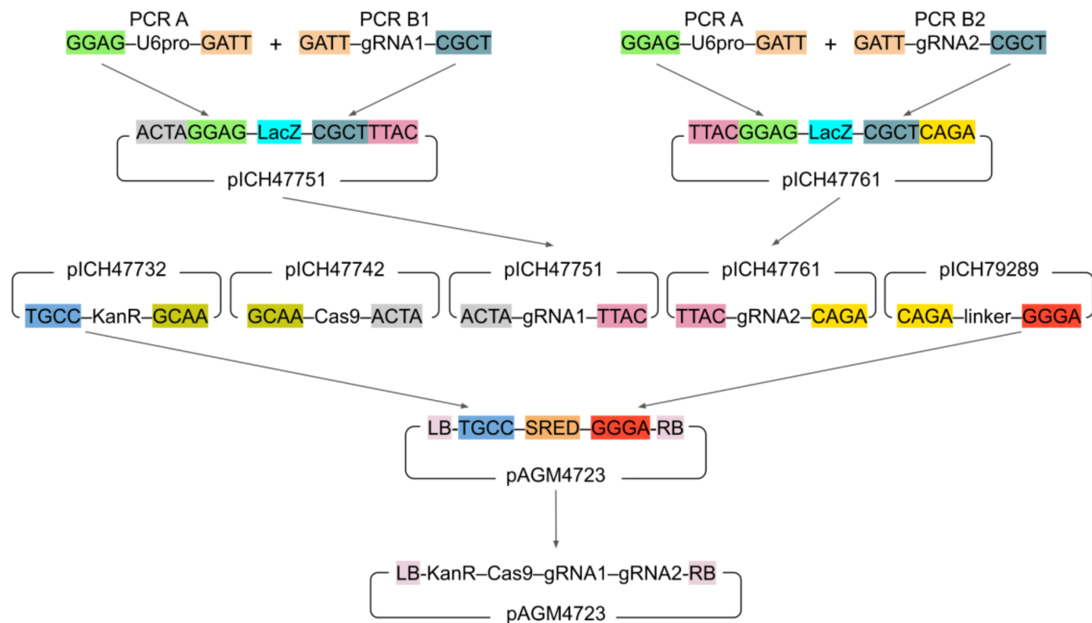


Figure 8. Modular cloning of the *SlDML1*-targeting gene editing construct.
First, we amplified the two halves of the gRNA using primers that have *Bsa*I recognition sites at their 5' ends. The first half (PCR A) is the same for any guide sequences, while the second half (PCR B1 and B2) is specific to the gRNA1 and 2 because the 5' end of the forward primers contains the 20-nt-long guide sequence and a *Bsa*I recognition site. The reverse primer is shared by PCR B1 and B2 since it matches the scaffold part of the gRNA. The sites that are cut by the *Bsa*I enzyme are highlighted with different colors to represent the different sticky end sequences. The two PCR halves were fused together and cloned into different Level 1 vectors. The LacZ enables a blue/white selection during cloning (the white colonies contain the desired insert). After cloning, the original *Bsa*I recognition sites are lost but the vectors contain *Bbs*I sites that result in different sticky ends after cutting. In this way, the different modules will have different sticky ends that makes it possible to assemble the final gene construct in a directional manner. We used a plant kanamycin selection marker module, a Cas9 expression module, the two gRNA expression modules, and an end-linker module for the final assembly. The final binary vector have a red pigment synthesis operon (SRED) that is replaced by the inserts enabling a red/white selection during cloning. The Level 1 modules carry ampicillin resistance, while the final binary vector has a kanamycin resistance gene.

These PCR products were purified from agarose gels and were used directly in a Golden Gate assembly reaction along with the vectors pICH47751 and pICH47761 (commercially available MoClo kit (Weber et al. 2011) components that contain several vectors for modular cloning utilizing the Golden Gate assembly method) for gRNA1 and 2, respectively having

compatible ends after *Bsa*I digestion (Figure 8). The reaction products were transformed into *E. coli* bacteria and grown on ampicillin and IPTG/X-gal containing agar plates. The positive clones were white compared to the negative ones that were blue due to the blue/white selection marker provided by the LacZ cassette in the vectors. These two gRNA expression modules were then used in a final Golden Gate assembly along with the plant kanamycin selection marker expression module (in pICH47732), the Cas9 expression module (in pICH47742), an end-linker (in pICH79289), and the vector (pAGM4723) to give rise to the final binary construct that can be used for tomato transformation. All the individual modules carried ampicillin resistance cassettes except for the vector, which conferred kanamycin resistance. The assembly products were transformed into *E. coli* and plated onto kanamycin containing agar plates. The positive clones were white while the negative ones were red, due to the presence of a β-carotene synthesis cassette in the vector that was replaced with the inserts in the correctly assembled product. Four white colonies were picked from the plate and plasmids were purified from them. The plasmids were digested with *Eco*RI and one correct plasmid was subjected to Sanger sequencing. According to the sequence analysis, the clone was correct, therefore, it was transformed into *Agrobacterium tumefaciens* strain LBA4404 (see Chapter 6.3 in the Materials and Methods section).

7.3. **Tomato transformation**

The correct gene construct was transformed into tomato plants as described in the Materials and Methods part (Chapter 6.4). The kanamycin resistant plantlets with strong roots were moved from the jars into soil until true leaves developed. We had more than 70 plants, some of which derived from the same callus. These plants got the same number but a letter was also assigned (D1-28A, 28B, 28C, etc). We expected that such plants would have the same mutation while truly independent lines derived from different calluses would have different mutant alleles. This is not always the case because it is impossible to tell when the mutation occurred during the development of the plant. Mosaicism is therefore frequently observed in CRISPR/Cas9 genome edited plants. It is also possible that the two chromatids of a diploid genome are edited differently, resulting in a biallelic mutant. Functionally, these plants could either be null or weak mutants depending on the nature of mutations. Most of the plant lines had wild-type phenotype, however, some lines had some interesting developmental defects (abnormal leaves or fruits, seedless fruits). Many lines did not produce flowers or fruits, therefore, we could not collect seeds from them. Some of these lines were vegetatively

propagated hoping that eventually they would produce seeds. We expected that most of the plants were heterozygous or wild-type (not edited), therefore, we genotyped them by PCR.

7.4. Genotyping of the transformed plants

The genotyping PCR primers were designed to amplify the region that includes both gRNA target sites. Ideally, a large deletion could happen between these two gRNA target sites that could easily be detected on an agarose gel without any further treatments like restriction digestion.

For genotyping the plants, genomic DNA were purified from the young leaves and PCR were performed using the DreamTaq Green PCR Master Mix (see Chapter 6.6 of the Materials and Methods section). The products were run on agarose gel hoping that large deletions would be detected. Unfortunately, no such large deletion event was observed in any of the tested lines. Smaller deletions could not be detected under such conditions because the size differences would be too small. Therefore, a restriction digestion of the PCR products with *Xba*I (that tests the mutational status of the gRNA1 site) was performed and the digestion products were resolved on an agarose gel again (Figure 9).

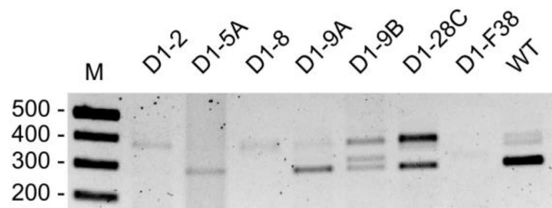


Figure 9. Genotyping PCR of a few selected lines after restriction digestion with *Xba*I. The wild-type form can be digested resulting in two smaller bands (288 and 92 bp) while the mutated ones remain intact (380 bp). Heterozygous plants are recognized by the presence of the two larger bands (380 and 288 bp). The 92 bp fragment cannot be seen because we trimmed the picture.

According to this analysis, D1-2 and D1-8 lines are homozygous mutant at the gRNA1 site, while D1-5A and D1-9A are wild-type. D1-9B and D1-28C are heterozygous. D1-9B appear to be a mosaic containing a somewhat larger deletion. D1-F38 line might be a homozygous mutant carrying a somewhat larger deletion at gRNA1 site. A large deletion

between the two gRNA sites would result in a ~217 bp product but no such event was observed. The gRNA2 site can be tested by digestion with *Taq*I but we have not performed these tests yet.

To reveal the exact nature of the detected mutations, the PCR products of the D1-2 and D1-8 lines (that appeared to be homozygous) were cloned into pJET vectors and their sequences were determined by Sanger sequencing. Aligning the sequences from the mutant lines with the wild-type one revealed the presence of smaller deletions at the gRNA1 site while there was no mutation at the gRNA2 site (Figure 10).

```
                                      PAM        gRNA1
                                      CCAGGTCTAGACATTAACGCGAC
WT          AGAAGAAAGGGCCAGCATCTGATGCGCCAGGTCTAGACATTAACGCGACACCAATACCAA
D1-2        AGAAGAAAGGGCCAGCATCTGATGCGCCAGGT--AGACATTAACGCGACACCAATACCAA
D1-8        AGAAGAAAGGGCCAGCATCTGATGCGCCAGGT-----CATTAACGCGACACCAATACCAA

WT          GACACGTACAGTCAAGCAAAGATACATTAAAGAGAGCAGAAGCAAATGATCTCCAACAGA
D1-2        GACACGTACAGTCAAGCAAAGATACATTAAAGAGAGCAGAAGCAAATGATCTCCAACAGA
D1-8        GACACGTACAGTCAAGCAAAGATACATTAAAGAGAGCAGAAGCAAATGATCTCCAACAGA
                                                                  gRNA2
                                                                 GATTGACA
WT          ACACAGAGAAGTCGGGGCTGGTCTTGAACATATCAGAACTGTCGGATAACATGATTGACA
D1-2        ACACAGAGAAGTCGGGGCTGGTCTTGAACATATCAGAACTGTCGGATAACATGATTGACA
D1-8        ACACAGAGAAGTCGGGGCTGGTCTTGAACATATCAGAACTGTCGGATAACATGATTGACA
                        PAM
            AGGTAGTCGATCAGG
WT          AGGTAGTCGATCAGGATGCTGAACAGAATAATACACCACAGCAGAAACGAAGGAAGAAGC
D1-2        AGGTAGTCGATCAGGATGCTGAACAGAATAATACACCACAGCAGAAACGAAGGAAGAAGC
D1-8        AGGTAGTCGATCAGGATGCTGAACAGAATAATACACCACAGCAGAAACGAAGGAAGAAGC
```

Figure 10. CRISPR/Cas9-induced mutations in the *sldml1-2* and *sldml1-8* mutant lines. The position of the gRNAs, their PAM sequences are shown. The *Xba*I site that cuts the wild-type allele at the gRNA1 target site and the *Taq*I site that cuts the wild-type allele at the gRNA2 site is highlighted in yellow. The deletions in the mutant lines are highlighted in red. The two mutations result in very similar truncated proteins (227 and 226 amino-acid-long) due to an immediate stop codon that occur after the site of deletion in a different frame (underlined).

It was observed earlier by others that some gRNAs did not work for an unknown reason. One explanation could be an unusual conformation of the gRNA that prevents its binding to the Cas9 enzyme. Another possibility is that the genome at the gRNA2 target site is not accessible because of closed chromatin structure. However, the observed deletions at the gRNA1 site all changed the frame of translation resulting in the occurrence of early stop codons and truncated proteins (227 and 226 amino-acid-long in the case of D1-2 and D1-8,

respectively, while the wild-type protein is 1810 amino-acid-long).

7.5. Characterization of the *sldml1* mutants

The preliminary observation of the T0 generation of *sldml1-2* and *sldml1-8* plants (which appear to be homozygous mutants) suggested that *SlDML1* might regulate fruit shape and seed formation (Figure 11).

Figure 11. Fruit shape of the *sldml1-2* and the *sldml2-8* mutants.
A, Fruit of a wild-type control plant (*Solanum lycopersicum* cv. Moneymaker). B, the *sldml1-2* line has elongated fruits, most of which were seedless. C, the *sldml2-8* line has also an altered fruit shape and all the fruits were seedless.

The *sldml1-2* had elongated, mostly seedless fruits. Fruit elongation in tomato is primarily controlled by four loci comprising *sun*, *ovate*, *sov1* and *fs8.1* (Wang et al. 2019). *SUN* controls tomato shape through redistribution of mass that is mediated by increased cell division in the longitudinal and decreased cell division in the transverse direction of the fruit. The expression of *SUN* is positively correlated with slender phenotypes in cotyledon, leaflet, and floral organs, an elongated ovary, and negatively correlated with seed weight (Wu et al. 2011). It will be interesting to test the T1 generation (there are some seeds from this line) to see if the phenotype is stable. In case of producing enough fruits, we would like to test the expression of *SUN*, *OVATE*, and *SOV1* (the molecular nature of the QTL locus fs8.1 is not known at the

moment) in the mutant and wild-type plants with RT-qPCR.

The *sldml1-8* line also had an altered fruit shape but in this line the fruits were heart-shaped with a pointy bottom. This trait is linked to auxin transport through the gene called *POINTED TIP* (*PT*) that encodes a C2H2-type zinc finger transcription factor and suppresses pointed tip formation by downregulating the transcription of *FRUITFULL 2* (*FUL2*), which alters the auxin transport (Song et al. 2022). It is possible that the lack of *SlDML1* results in an altered expression of *PT* or *FUL2* that allows pointy fruit formation. All the fruits of *sldml1-8* were seedless, therefore, the plant is propagated vegetatively.

Both mutants were mostly seedless and seedlessness is linked to auxin and gibberellic acid signaling (Bassel et al. 2008, Carrera et al. 2012, Hu et al. 2018). Also, the role of DNA demethylases in the seed development was shown in plants as described in the Literature Review section. A genome-wide DNA methylation sequencing combined with transcriptome sequencing from the developing fruits of the mutant and wild-type plants could reveal which of the components of the auxin and gibberellic acid signaling were affected in the *sldml1* mutants.

We have several other mutant lines that appeared to be heterozygous (like the D1-28C in Figure 9) and produced seeds. We are planning to determine the nature of the mutation and genotype the T1 generation looking for homozygous mutants. Ideally, the homozygous mutants would also produce elongated fruits. Unfortunately, the seedlessness phenotype of the mutants makes any further experiments difficult to perform.

Off-target effects could also be a problem to interpret the observed phenotypes. Off-targets were only predicted in the case of gRNA2, but because gRNA2 was not working for some reason, therefore, we can safely rule out the possibility that the observed phenotypes are the result of some off-target effect. To eliminate the possibility of the occurrence of off-target mutations in the future because of the constant presence of the Cas9-containing transgene, we will test the T1 generation for plants in which the Cas9 was segregated out. However, this is only possible if we have enough seeds. The investigation of independent mutant lines with different mutations would also help to establish a causative relationship between the genotype and phenotype of the *sldml1* mutants.

8. CONCLUSIONS AND FUTURE PLANS

We successfully reached our main objectives, namely, we managed to design specific guide RNAs targeting *SlDML1*, created the genome editing construct using the Golden Gate assembly method, and successfully transformed tomato plants with the construct. We also verified the presence of deletions in the *SlDML1* gene that theoretically caused frameshift in the translated protein in several independent lines some of which seemed to be homozygous already in the T0 generation. These mutants displayed aberrant fruit developmental phenotypes.

To reveal the molecular background of the observed phenotypes, we are planning to perform a genome-wide DNA methylation and transcriptome sequencing of the *sldml1-2* and *sldml1-8* mutant fruits along with the wild-type ones. We hope that by comparing the differential DNA methylation and the associated gene expression profiles of the mutant and wild-type plants we will be able to identify the genomic targets of the *SlDML1* DNA demethylase. This will be a real novelty, since currently there is no information about the role of *SlDMLs* in the regulation of fruit development. Therefore, the results probably will bring attention and make a high impact.

9. SUMMARY

The objective of this study was to investigate the function of *SlDML1* in tomato by creating a null mutant using the CRISPR/Cas9-based genome editing technology. This was done because although DNA demethylation is well studied in *Arabidopsis thaliana*, limited information is available from other plants, especially those of agricultural importance like tomato. There are four DNA demethylases in tomato, but only *SlDML2* has been investigated in more detail so far due to its significant induction during fruit ripening. However, other demethylases, including *SlDML1*, are ubiquitously expressed and may have important functions in developmental processes or stress responses.

To achieve the goal of creating a null mutant of *SlDML1*, we designed specific guide RNAs targeting the protein-coding region of *SlDML1* at two sites simultaneously to increase the chance of obtaining a null mutant. We then created a binary gene construct carrying a kanamycin selection marker, the Cas9 gene cassette, and the two modules expressing the guide RNAs. We transformed the construct into tomato plants, and transformant plants were regenerated from calluses in tissue culture. We verified the presence of CRISPR/Cas9-generated mutation(s) in the transformant plants, and we also conducted some preliminary physiological characterization of the verified mutant plants.

The results of this study showed that the CRISPR/Cas9-based genome editing technology was successful in generating a null mutant of *SlDML1* in tomato. The preliminary physiological characterization of the mutant plants revealed that they exhibited altered growth and developmental patterns compared to wild-type plants.

The significance of this study lies in its contribution to the understanding of DNA demethylation in plants, particularly in tomato. By investigating the function of *SlDML1*, a previously understudied DNA demethylase, we could hopefully provide insights into the potential role of this gene in plant development and stress responses. This information could be useful in the development of new strategies for crop improvement, including the breeding of tomato plants with enhanced stress tolerance.

10. ACKNOWLEDGEMENTS

This work was funded by the Hungarian Government Organization NRDI (National Research, Development, and Innovation Office: https://nkfih.gov.hu/about-the-office) through the grants K-134974 awarded to Dr. György Szittya, and FK-137811 awarded to Dr. Péter Gyula.

I am grateful to all who have helped me in completing this thesis, but above all, I would like to express my sincere appreciation to Almighty Allah for blessing me with the knowledge, resilience, and good health required to accomplish this task.

I would also like to thank the GBI (Institute of Genetics and Biotechnology) and SH (Stipendium Hungaricum) for providing me with access to valuable resources and facilities that have been instrumental in the successful completion of this research.

First and foremost, I would like to express my sincere gratitude to my thesis supervisor, **Dr. György Szittya,** and co-supervisor **Dr. Péter Gyula,** for their guidance, support, and encouragement throughout the research process. Their valuable feedback, insightful comments, and constructive criticism have greatly contributed to the quality of this thesis.

My heartfelt thanks go to my family and friends (**Abdul Razzak** and **Mohammad Ali**) who have supported me throughout this journey. Their unwavering encouragement, understanding, and moral support have kept me motivated and inspired throughout the writing process.

Finally, I would like to extend my gratitude to all the participants who have generously given their time and shared their experiences and perspectives, without whom this research would not have been possible.

Thank you.

11. BIBLIOGRAPHY

Ali, Z., Ali, S., Tashkandi, M., Zaidi, S.S.-A. & Mahfouz, M.M. (2016): CRISPR/Cas9-Mediated Immunity to Geminiviruses: Differential Interference and Evasion. Sci. Rep. 6(1): 26912.

Anzalone, A.V., Koblan, L.W. & Liu, D.R. (2020): Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. Nat. Biotechnol. 38(7): 824–844.

Arora, L. & Narula, A. (2017): Gene Editing and Crop Improvement Using CRISPR-Cas9 System. Front. Plant Sci. 8.

Azzi, L., Deluche, C., Gévaudant, F., Frangne, N., Delmas, F., Hernould, M. & Chevalier, C. (2015): Fruit growth-related genes in tomato. J. Exp. Bot. 66(4): 1075–1086.

Bao, A., Burritt, D.J., Chen, H., Zhou, X., Cao, D. & Tran, L.-S.P. (2019): The CRISPR/Cas9 system and its applications in crop genome editing. Crit. Rev. Biotechnol. 39(3): 321–336.

Bassel, G.W., Mullen, R.T. & Bewley, J.D. (2008): procera is a putative DELLA mutant in tomato (Solanum lycopersicum): effects on the seed and vegetative plant. J. Exp. Bot. 59(3): 585–593.

Bauer, M.J. & Fischer, R.L. (2011): Genome demethylation and imprinting in the endosperm. Curr. Opin. Plant Biol. 14(2): 162–167.

Baysal, C., Bortesi, L., Zhu, C., Farré, G., Schillberg, S. & Christou, P. (2016): CRISPR/Cas9 activity in the rice OsBEIIb gene does not induce off-target effects in the closely related paralog OsBEIIa. Mol. Breed. 36(8): 108.

Bhatia, S., Pooja & Yadav, S.K. (2023): CRISPR-Cas for genome editing: Classification, mechanism, designing and applications. Int. J. Biol. Macromol. 238124054.

Bhattacharjee, G., Mani, I., Gohil, N., Khambhati, K., Braddick, D., Panchasara, H. & Singh, V. (2020): Chapter 7 - CRISPR technology for genome editing, in: Faintuch, J., & Faintuch, S. (Eds.), Precision Medicine for Investigators, Practitioners and Providers. Academic Press, pp. 59–69.

Bird, A. (2007): Perceptions of epigenetics. Nature 447(7143): 396–398.

Bortesi, L., Zhu, C., Zischewski, J., Perez, L., Bassié, L., Nadi, R., Forni, G., Lade, S.B., Soto, E., Jin, X., Medina, V., Villorbina, G., Muñoz, P., Farré, G., Fischer, R., Twyman, R.M., Capell, T., Christou, P. & Schillberg, S. (2016): Patterns of CRISPR/Cas9 activity in plants, animals and microbes. Plant Biotechnol. J. 14(12): 2203–2216.

Boti, M.A., Athanasopoulou, K., Adamopoulos, P.G., Sideris, D.C. & Scorilas, A. (2023): Recent Advances in Genome-Engineering Strategies. Genes 14(1): 129.

Carrera, E., Ruiz-Rivero, O., Peres, L.E.P., Atares, A. & Garcia-Martinez, J.L. (2012): Characterization of the procera Tomato Mutant Shows Novel Functions of the SlDELLA Protein in the Control of Flower Morphology, Cell Division and Expansion, and the Auxin-Signaling Pathway during Fruit-Set and Development. Plant Physiol. 160(3): 1581–1596.

Chandrasekaran, J., Brumin, M., Wolf, D., Leibman, D., Klap, C., Pearlsman, M., Sherman, A., Arazi, T. & Gal-On, A. (2016): Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. Mol. Plant Pathol. 17(7): 1140–1153.

Chang, Y.-N., Zhu, C., Jiang, J., Zhang, H., Zhu, J.-K. & Duan, C.-G. (2020): Epigenetic regulation in plant abiotic stress responses. J. Integr. Plant Biol. 62(5): 563–580.

Chen, H., Nelson, R.S. & Sherwood, J.L. (1994): Enhanced recovery of transformants of Agrobacterium tumefaciens after freeze-thaw transformation and drug selection. BioTechniques 16(4): 664–668, 670.

Chen, X., Schönberger, B., Menz, J. & Ludewig, U. (2018): Plasticity of DNA methylation and gene expression under zinc deficiency in Arabidopsis roots. Plant Cell Physiol. 59(9): 1790–1802.

Cheng, J., Niu, Q., Zhang, B., Chen, K., Yang, R., Zhu, J.-K., Zhang, Y. & Lang, Z. (2018): Downregulation of RdDM during strawberry fruit ripening. Genome Biol. 19(1): 212.

Cho, S.W., Lee, Jihyun, Carroll, D., Kim, J.-S. & Lee, Junho (2013): Heritable Gene Knockout in Caenorhabditis elegans by Direct Injection of Cas9–sgRNA Ribonucleoproteins. Genetics 195(3): 1177–1180.

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E. & Fischer, R.L. (2002a): DEMETER, a DNA Glycosylase Domain Protein, Is Required for

Endosperm Gene Imprinting and Seed Viability in Arabidopsis. Cell 110(1): 33–42.
Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E. & Fischer, R.L. (2002b): DEMETER, a DNA Glycosylase Domain Protein, Is Required for Endosperm Gene Imprinting and Seed Viability in Arabidopsis. Cell 110(1): 33–42.
Ding, Y., Li, H., Chen, L.-L. & Xie, K. (2016): Recent Advances in Genome Editing Using CRISPR/Cas9. Front. Plant Sci. 7.
El-Mounadi, K., Morales-Floriano, M.L. & Garcia-Ruiz, H. (2020): Principles, Applications, and Biosafety of Plant Genome Editing Using CRISPR-Cas9. Front. Plant Sci. 11.
Engler, C., Kandzia, R. & Marillonnet, S. (2008): A One Pot, One Step, Precision Cloning Method with High Throughput Capability. PLOS ONE 3(11): e3647.
Fernandez, A.I., Viron, N., Alhagdow, M., Karimi, M., Jones, M., Amsellem, Z., Sicard, A., Czerednik, A., Angenent, G., Grierson, D., May, S., Seymour, G., Eshed, Y., Lemaire-Chamley, M., Rothan, C. & Hilson, P. (2009): Flexible Tools for Gene Expression and Silencing in Tomato. Plant Physiol. 151(4): 1729–1740.
Gallusci, P., Hodgman, C., Teyssier, E. & Seymour, G.B. (2016): DNA Methylation and Chromatin Regulation during Fleshy Fruit Development and Ripening. Front. Plant Sci. 7.
Ganguly, D.R., Crisp, P.A., Eichten, S.R. & Pogson, B.J. (2017): The Arabidopsis DNA Methylome Is Stable under Transgenerational Drought Stress. Plant Physiol. 175(4): 1893–1912.
Gao, J., Wang, G., Ma, S., Xie, X., Wu, X., Zhang, X., Wu, Y., Zhao, P. & Xia, Q. (2015): CRISPR/Cas9-mediated targeted mutagenesis in Nicotiana tabacum. Plant Mol. Biol. 87(1): 99–110.
Gil-Humanes, J., Wang, Y., Liang, Z., Shan, Q., Ozuna, C.V., Sánchez-León, S., Baltes, N.J., Starker, C., Barro, F., Gao, C. & Voytas, D.F. (2017): High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. Plant J. 89(6): 1251–1262.
Gong, Z., Morales-Ruiz, T., Ariza, R.R., Roldán-Arjona, T., David, L. & Zhu, J.-K. (2002): ROS1, a Repressor of Transcriptional Gene Silencing in Arabidopsis, Encodes a DNA Glycosylase/Lyase. Cell 111(6): 803–814.
Groszmann, M., Gonzalez-Bayon, R., Greaves, I.K., Wang, L., Huen, A.K., Peacock, W.J. & Dennis, E.S. (2014): Intraspecific Arabidopsis Hybrids Show Different Patterns of Heterosis Despite the Close Relatedness of the Parental Genomes. Plant Physiol. 166(1): 265–280.
Groszmann, M., Greaves, I.K., Albertyn, Z.I., Scofield, G.N., Peacock, W.J. & Dennis, E.S. (2011): Changes in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to hybrid vigor. Proc. Natl. Acad. Sci. U. S. A. 108(6): 2617–2622.
Groszmann, M., Greaves, I.K., Fujimoto, R., James Peacock, W. & Dennis, E.S. (2013): The role of epigenetics in hybrid vigour. Trends Genet. 29(12): 684–690.
Gupta, C. & Salgotra, R.K. (2022): Epigenetics and its role in effecting agronomical traits. Front. Plant Sci. 13.
Guschin, D.Y., Waite, A.J., Katibah, G.E., Miller, J.C., Holmes, M.C. & Rebar, E.J. (2010): A Rapid and General Assay for Monitoring Endogenous Gene Modification, in: Mackay, J.P., & Segal, D.J. (Eds.), Engineered Zinc Finger Proteins: Methods and Protocols, Methods in Molecular Biology. Humana Press, Totowa, NJ, pp. 247–256.
Haeussler, M., Schönig, K., Eckert, H., Eschstruth, A., Mianné, J., Renaud, J.-B., Schneider-Maunoury, S., Shkumatava, A., Teboul, L., Kent, J., Joly, J.-S. & Concordet, J.-P. (2016): Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol. 17(1): 148.
Haurwitz, R.E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J.A. (2010): Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease. Science 329(5997): 1355–1358.
He, G., Chen, B., Wang, X., Li, X., Li, J., He, H., Yang, M., Lu, L., Qi, Y., Wang, X. & Wang Deng, X. (2013): Conservation and divergence of transcriptomic and epigenomic variation in maize hybrids. Genome Biol. 14(6): .
Hille, F. & Charpentier, E. (2016): CRISPR-Cas: biology, mechanisms and relevance. Philos. Trans. R. Soc. B Biol. Sci. 371(1707): 20150496.
Hofmeister, B.T., Lee, K., Rohr, N.A., Hall, D.W. & Schmitz, R.J. (2017): Stable inheritance of DNA methylation allows creation of epigenotype maps and the study of epiallele inheritance patterns in the absence of genetic variation. Genome Biol. 18(1): 155.

Hu, J., Israeli, A., Ori, N. & Sun, T. (2018): The Interaction between DELLA and ARF/IAA Mediates Crosstalk between Gibberellin and Auxin Signaling to Control Fruit Initiation in Tomato. Plant Cell 30(8): 1710–1728.

Huang, H., Liu, R., Niu, Q., Tang, K., Zhang, B., Zhang, H., Chen, K., Zhu, J.-K. & Lang, Z. (2019): Global increase in DNA methylation during orange fruit development and ripening. Proc. Natl. Acad. Sci. 116(4): 1430–1436.

Hunt, L. & Gray, J.E. (2009): The Signaling Peptide EPF2 Controls Asymmetric Cell Divisions during Stomatal Development. Curr. Biol. 19(10): 864–869.

Ibarra, C.A., Feng, X., Schoft, V.K., Hsieh, T.-F., Uzawa, R., Rodrigues, J.A., Zemach, A., Chumak, N., Machlicova, A., Nishimura, T., Rojas, D., Fischer, R.L., Tamaru, H. & Zilberman, D. (2012): Active DNA Demethylation in Plant Companion Cells Reinforces Transposon Methylation in Gametes. Science 337(6100): 1360–1364.

Inoue, H., Nojima, H. & Okayama, H. (1990): High efficiency transformation of Escherichia coli with plasmids. Gene 96(1): 23–28.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. (1987): Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J. Bacteriol. 169(12): 5429–5433.

Jacobs, T.B., LaFayette, P.R., Schmitz, R.J. & Parrott, W.A. (2015): Targeted genome modifications in soybean with CRISPR/Cas9. BMC Biotechnol. 15(1): 16.

Jansen, Ruud., Embden, Jan.D.A. van, Gaastra, Wim. & Schouls, Leo.M. (2002): Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. 43(6): 1565–1575.

Ji, X., Zhang, H., Zhang, Y., Wang, Y. & Gao, C. (2015): Establishing a CRISPR–Cas-like immune system conferring DNA virus resistance in plants. Nat. Plants 1(10): 1–4.

Kapazoglou, A., Ganopoulos, I., Tani, E. & Tsaftaris, A. (2018): Epigenetics, Epigenomics and Crop Improvement, in: Advances in Botanical Research. Elsevier, pp. 287–324.

Kapusi, E., Corcuera-Gómez, M., Melnik, S. & Stoger, E. (2017): Heritable Genomic Fragment Deletions and Small Indels in the Putative ENGase Gene Induced by CRISPR/Cas9 in Barley. Front. Plant Sci. 8.

Khouider, S., Borges, F., LeBlanc, C., Ungru, A., Schnittger, A., Martienssen, R., Colot, V. & Bouyer, D. (2021): Male fertility in Arabidopsis requires active DNA demethylation of genes that control pollen tube function. Nat. Commun. 12(1): 410.

Kim, M.Y., Ono, A., Scholten, S., Kinoshita, T., Zilberman, D., Okamoto, T. & Fischer, R.L. (2019): DNA demethylation by ROS1a in rice vegetative cells promotes methylation in sperm. Proc. Natl. Acad. Sci. 116(19): 9652–9657.

Korotko, U., Chwiałkowska, K., Sańko-Sawczenko, I. & Kwasniewski, M. (2021): DNA Demethylation in Response to Heat Stress in Arabidopsis thaliana. Int. J. Mol. Sci. 22(4): 1555.

Kumar, S. (2016): Epigenetic Regulation of Abiotic Stress Tolerance in Plants. Adv. Plants Agric. Res. 5(3): .

Kumar, S., Chinnusamy, V. & Mohapatra, T. (2018): Epigenetics of Modified DNA Bases: 5-Methylcytosine and Beyond. Front. Genet. 9.

Kumar, S., Kumari, R., Sharma, Vishakha & Sharma, Vinay (2013): Roles, and establishment, maintenance and erasing of the epigenetic cytosine methylation marks in plants. J. Genet. 92(3): 629–666.

Kumar, S. & Mohapatra, T. (2021): Dynamics of DNA Methylation and Its Functions in Plant Growth and Development. Front. Plant Sci. 12.

Lang, Z., Wang, Y., Tang, K., Tang, D., Datsenka, T., Cheng, J., Zhang, Y., Handa, A.K. & Zhu, J.-K. (2017a): Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. Proc. Natl. Acad. Sci. 114(22): E4511–E4519.

Lang, Z., Wang, Y., Tang, K., Tang, D., Datsenka, T., Cheng, J., Zhang, Y., Handa, A.K. & Zhu, J.-K. (2017b): Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. Proc. Natl. Acad. Sci. 114(22): E4511–E4519.

Law, J.A. & Jacobsen, S.E. (2010): Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. 11(3): 204–220.
Lawrenson, T., Shorinola, O., Stacey, N., Li, C., Østergaard, L., Patron, N., Uauy, C. & Harwood, W. (2015): Induction of targeted, heritable mutations in barley and Brassica oleracea using RNA-guided Cas9 nuclease. Genome Biol. 16(1): 258.
Le, T.-N., Schumann, U., Smith, N.A., Tiwari, S., Au, P.C.K., Zhu, Q.-H., Taylor, J.M., Kazan, K., Llewellyn, D.J., Zhang, R., Dennis, E.S. & Wang, M.-B. (2014): DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in Arabidopsis. Genome Biol. 15(9): 458.
Li, M., Li, X., Zhou, Z., Wu, P., Fang, M., Pan, X., Lin, Q., Luo, W., Wu, G. & Li, H. (2016): Reassessment of the Four Yield-related Genes Gn1a, DEP1, GS3, and IPA1 in Rice Using a CRISPR/Cas9 System. Front. Plant Sci. 7.
Li, R., Zhang, L., Wang, L., Chen, L., Zhao, R., Sheng, J. & Shen, L. (2018): Reduction of Tomato-Plant Chilling Tolerance by CRISPR–Cas9-Mediated SlCBF1 Mutagenesis. J. Agric. Food Chem. 66(34): 9042–9051.
Li, X., Wang, X., Zhang, Y., Zhang, A. & You, C.-X. (2022): Regulation of fleshy fruit ripening: from transcription factors to epigenetic modifications. Hortic. Res. 9uhac013.
Li, Y., Córdoba-Cañero, D., Qian, W., Zhu, X., Tang, K., Zhang, H., Ariza, R.R., Roldán-Arjona, T. & Zhu, J.-K. (2015): An AP Endonuclease Functions in Active DNA Demethylation and Gene Imprinting in Arabidopsis. PLOS Genet. 11(1): e1004905.
Li, Y., Kumar, S. & Qian, W. (2018): Active DNA demethylation: mechanism and role in plant development. Plant Cell Rep. 37(1): 77–85.
Liu, Q., Wang, J., Miki, D., Xia, R., Yu, W., He, J., Zheng, Z., Zhu, J.-K. & Gong, Z. (2010): DNA Replication Factor C1 Mediates Genomic Stability and Transcriptional Gene Silencing in Arabidopsis. Plant Cell 22(7): 2336–2352.
Liu, R., How-Kit, A., Stammitti, L., Teyssier, E., Rolin, D., Mortain-Bertrand, A., Halle, S., Liu, M., Kong, J., Wu, C., Degraeve-Guibault, C., Chapman, N.H., Maucourt, M., Hodgman, T.C., Tost, J., Bouzayen, M., Hong, Y., Seymour, G.B., Giovannoni, J.J. & Gallusci, P. (2015): A DEMETER-like DNA demethylase governs tomato fruit ripening. Proc. Natl. Acad. Sci. 112(34): 10804–10809.
Liu, R. & Lang, Z. (2020): The mechanism and function of active DNA demethylation in plants. J. Integr. Plant Biol. 62(1): 148–159.
Liu, X., Xie, C., Si, H. & Yang, J. (2017): CRISPR/Cas9-mediated genome editing in plants. Methods, CRISPR-Cas systems for Genome engineering and investigation 121–12294–102.
Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A. & Lopez, R. (2022): Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res. gkac240.
Mahas, A., Ali, Z., Tashkandi, M. & Mahfouz, M.M. (2019): Virus-Mediated Genome Editing in Plants Using the CRISPR/Cas9 System, in: Qi, Y. (Ed.), Plant Genome Editing with CRISPR Systems: Methods and Protocols, Methods in Molecular Biology. Springer, New York, NY, pp. 311–326.
Maher, M.F., Nasti, R.A., Vollbrecht, M., Starker, C.G., Clark, M.D. & Voytas, D.F. (2020): Plant gene editing through de novo induction of meristems. Nat. Biotechnol. 38(1): 84–89.
McCue, A.D., Panda, K., Nuthikattu, S., Choudury, S.G., Thomas, E.N. & Slotkin, R.K. (2015): ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. EMBO J. 34(1): 20–35.
Metje-Sprink, J., Menz, J., Modrzejewski, D. & Sprink, T. (2019): DNA-Free Genome Editing: Past, Present and Future. Front. Plant Sci. 9.
Moglia, A., Gianoglio, S., Acquadro, A., Valentino, D., Milani, A.M., Lanteri, S. & Comino, C. (2019): Identification of DNA methyltransferases and demethylases in Solanum melongena L., and their transcription dynamics during fruit development and after salt and drought stresses. PLOS ONE 14(10): e0223581.
Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. (2009): Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology 155(3): 733–740.

Mok, Y.G., Uzawa, R., Lee, J., Weiner, G.M., Eichman, B.F., Fischer, R.L. & Huh, J.H. (2010): Domain structure of the DEMETER 5-methylcytosine DNA glycosylase. Proc. Natl. Acad. Sci. 107(45): 19225–19230.

Montecillo, J.A.V., Chu, L.L. & Bae, H. (2020): CRISPR-Cas9 System for Plant Genome Editing: Current Approaches and Emerging Developments. Agronomy 10(7): 1033.

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D.G. & Kamoun, S. (2013): Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nat. Biotechnol. 31(8): 691–693.

Pandey, G., Sharma, N., Sahu, P.P. & Prasad, M. (2016): Chromatin-Based Epigenetic Regulation of Plant Abiotic Stress Response. Curr. Genomics 17(6): 490–498.

Parmar, N., Singh, K.H., Sharma, D., Singh, L., Kumar, P., Nanjundan, J., Khan, Y.J., Chauhan, D.K. & Thakur, A.K. (2017): Genetic engineering strategies for biotic and abiotic stress tolerance and quality enhancement in horticultural crops: a comprehensive review. 3 Biotech 7(4): 239.

Parrilla-Doblas, J.T., Roldán-Arjona, T., Ariza, R.R. & Córdoba-Cañero, D. (2019): Active DNA Demethylation in Plants. Int. J. Mol. Sci. 20(19): 4683.

Penterman, J., Uzawa, R. & Fischer, R.L. (2007): Genetic Interactions between DNA Demethylation and Methylation in Arabidopsis. Plant Physiol. 145(4): 1549–1557.

Prihatna, C., Barbetti, M.J. & Barker, S.J. (2018): A Novel Tomato Fusarium Wilt Tolerance Gene. Front. Microbiol. 9.

Qian, W., Miki, D., Zhang, H., Liu, Y., Zhang, X., Tang, K., Kan, Y., La, H., Li, X., Li, S., Zhu, X., Shi, X., Zhang, K., Pontes, O., Chen, X., Liu, R., Gong, Z. & Zhu, J.-K. (2012): A Histone Acetyltransferase Regulates Active DNA Demethylation in Arabidopsis. Science 336(6087): 1445–1448.

Rath, D., Amlinger, L., Rath, A. & Lundgren, M. (2015): The CRISPR-Cas immune system: Biology, mechanisms and applications. Biochimie, Special Issue: Regulatory RNAs 117119–128.

Rodríguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E. & Lippman, Z.B. (2017): Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. Cell 171(2): 470-480.e8.

Salava, H., Thula, S., Mohan, V., Kumar, R. & Maghuly, F. (2021): Application of Genome Editing in Tomato Breeding: Mechanisms, Advances, and Prospects. Int. J. Mol. Sci. 22(2): 682.

Sambrook, J.F. & Russell, D. (2001): Molecular Cloning: A Laboratory Manual (3-Volume Set), 3rd ed. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press.

Satgé, C., Moreau, S., Sallet, E., Lefort, G., Auriac, M.-C., Remblière, C., Cottret, L., Gallardo, K., Noirot, C., Jardinaud, M.-F. & Gamas, P. (2016): Reprogramming of DNA methylation is critical for nodule development in Medicago truncatula. Nat. Plants 2(11): 1–10.

Sauer, N.J., Narváez-Vásquez, J., Mozoruk, J., Miller, R.B., Warburg, Z.J., Woodward, M.J., Mihiret, Y.A., Lincoln, T.A., Segami, R.E., Sanders, S.L., Walker, K.A., Beetham, P.R., Schöpke, C.R. & Gocal, G.F.W. (2016): Oligonucleotide-Mediated Genome Editing Provides Precision and Function to Engineered Nucleases and Antibiotics in Plants. Plant Physiol. 170(4): 1917–1928.

Schoft, V.K., Chumak, N., Choi, Y., Hannon, M., Garcia-Aguilar, M., Machlicova, A., Slusarz, L., Mosiolek, M., Park, J.-S., Park, G.T., Fischer, R.L. & Tamaru, H. (2011): Function of the DEMETER DNA glycosylase in the Arabidopsis thaliana male gametophyte. Proc. Natl. Acad. Sci. 108(19): 8042–8047.

Seymour, G.B., Østergaard, L., Chapman, N.H., Knapp, S. & Martin, C. (2013): Fruit Development and Ripening. Annu. Rev. Plant Biol. 64(1): 219–241.

Shen, H., He, H., Li, J., Chen, W., Wang, X., Guo, L., Peng, Z., He, G., Zhong, S., Qi, Y., Terzaghi, W. & Deng, X.W. (2012): Genome-Wide Analysis of DNA Methylation and Gene Expression Changes in Two Arabidopsis Ecotypes and Their Reciprocal Hybrids. Plant Cell 24(3): 875–892.

Shi, J., Gao, H., Wang, H., Lafitte, H.R., Archibald, R.L., Yang, M., Hakimi, S.M., Mo, H. & Habben, J.E. (2017): ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. Plant Biotechnol. J. 15(2): 207–216.

Song, J., Shang, L., Li, C., Wang, W., Wang, X., Zhang, C., Ai, G., Ye, J., Yang, C., Li, H., Hong, Z., Larkin, R.M., Ye, Z. & Zhang, J. (2022): Variation in the fruit development gene POINTED

TIP regulates protuberance of tomato fruit tip. Nat. Commun. 13(1): 5940.
Tang, K., Lang, Z., Zhang, H. & Zhu, J.-K. (2016): The DNA demethylase ROS1 targets genomic regions with distinct chromatin modifications. Nat. Plants 2(11): 1–10.
Thomazella, D.P. de T., Brail, Q., Dahlbeck, D. & Staskawicz, B. (2016): CRISPR-Cas9 mediated mutagenesis of a DMR6 ortholog in tomato confers broad-spectrum disease resistance.
Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P., Aryee, M.J. & Joung, J.K. (2015): GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat. Biotechnol. 33(2): 187–197.
van der Graaf, A., Wardenaar, R., Neumann, D.A., Taudt, A., Shaw, R.G., Jansen, R.C., Schmitz, R.J., Colomé-Tatché, M. & Johannes, F. (2015): Rate, spectrum, and evolutionary dynamics of spontaneous epimutations. Proc. Natl. Acad. Sci. 112(21): 6676–6681.
Van Dooren, T.J.M., Silveira, A.B., Gilbault, E., Jiménez-Gómez, J.M., Martin, A., Bach, L., Tisné, S., Quadrana, L., Loudet, O. & Colot, V. (2020): Mild drought in the vegetative stage induces phenotypic, gene expression, and DNA methylation plasticity in Arabidopsis but no transgenerational effects. J. Exp. Bot. 71(12): 3588–3602.
Voytas, D.F. (2013): Plant Genome Engineering with Sequence-Specific Nucleases. Annu. Rev. Plant Biol. 64(1): 327–350.
Waddington, C.H. (1942): Canalization of Development and the Inheritance of Acquired Characters. Nature 150(3811): 563–565.
Wan, L., Wang, Z., Tang, M., Hong, D., Sun, Y., Ren, J., Zhang, N. & Zeng, H. (2021): CRISPR-Cas9 Gene Editing for Fruit and Vegetable Crops: Strategies and Prospects. Horticulturae 7(7): 193.
Wang, W.-S., Pan, Y.-J., Zhao, X.-Q., Dwivedi, D., Zhu, L.-H., Ali, J., Fu, B.-Y. & Li, Z.-K. (2011): Drought-induced site-specific DNA methylation and its association with drought tolerance in rice (Oryza sativa L.). J. Exp. Bot. 62(6): 1951–1960.
Wang, Y., Clevenger, J.P., Illa-Berenguer, E., Meulia, T., van der Knaap, E. & Sun, L. (2019): A Comparison of sun, ovate, fs8.1 and Auxin Application on Tomato Fruit Shape and Gene Expression. Plant Cell Physiol. 60(5): 1067–1081.
Weber, E., Engler, C., Gruetzner, R., Werner, S. & Marillonnet, S. (2011): A Modular Cloning System for Standardized Assembly of Multigene Constructs. PLOS ONE 6(2): e16765.
Weeks, D.P., Spalding, M.H. & Yang, B. (2016): Use of designer nucleases for targeted gene and genome editing in plants. Plant Biotechnol. J. 14(2): 483–495.
Wen, S., Wen, N., Pang, J., Langen, G., Brew-Appiah, R.A.T., Mejias, J.H., Osorio, C., Yang, M., Gemini, R., Moehs, C.P., Zemetra, R.S., Kogel, K.-H., Liu, B., Wang, X., von Wettstein, D. & Rustgi, S. (2012): Structural genes of wheat and barley 5-methylcytosine DNA glycosylases and their potential applications for human health. Proc. Natl. Acad. Sci. 109(50): 20543–20548.
Williams, B.P., Pignatta, D., Henikoff, S. & Gehring, M. (2015): Methylation-Sensitive Expression of a DNA Demethylase Gene Serves As an Epigenetic Rheostat. PLOS Genet. 11(3): e1005142.
Wu, S., Xiao, H., Cabrera, A., Meulia, T. & van der Knaap, E. (2011): SUN Regulates Vegetative and Reproductive Organ Shape by Changing Cell Division Patterns. Plant Physiol. 157(3): 1175–1186.
Xie, K. & Yang, Y. (2013): RNA-Guided Genome Editing in Plants Using a CRISPR–Cas System. Mol. Plant 6(6): 1975–1983.
Xu, R.-F., Li, H., Qin, R.-Y., Li, J., Qiu, C.-H., Yang, Y.-C., Ma, H., Li, L., Wei, P.-C. & Yang, J.-B. (2015): Generation of inheritable and “transgene clean” targeted genome-modified rice in later generations using the CRISPR/Cas9 system. Sci. Rep. 5(1): 11491.
Yamamuro, C., Miki, D., Zheng, Z., Ma, J., Wang, J., Yang, Z., Dong, J. & Zhu, J.-K. (2014): Overproduction of stomatal lineage cells in Arabidopsis mutants defective in active DNA demethylation. Nat. Commun. 5(1): 4062.
Yang, L., Lang, C., Wu, Y., Meng, D., Yang, T., Li, D., Jin, T. & Zhou, X. (2022): ROS1-mediated decrease in DNA methylation and increase in expression of defense genes and stress response genes in Arabidopsis thaliana due to abiotic stresses. BMC Plant Biol. 22(1): 104.
Yin, Y., Qin, K., Song, X., Zhang, Q., Zhou, Y., Xia, X. & Yu, J. (2018): BZR1 Transcription Factor Regulates Heat Stress Tolerance Through FERONIA Receptor-Like Kinase-Mediated

Reactive Oxygen Species Signaling in Tomato. Plant Cell Physiol. 59(11): 2239–2254.
Yu, A., Lepère, G., Jay, F., Wang, J., Bapaume, L., Wang, Y., Abraham, A.-L., Penterman, J., Fischer, R.L., Voinnet, O. & Navarro, L. (2013): Dynamics and biological relevance of DNA demethylation in Arabidopsis antibacterial defense. Proc. Natl. Acad. Sci. 110(6): 2389–2394.
Yu, Q., Wang, B., Li, N., Tang, Y., Yang, S., Yang, T., Xu, J., Guo, C., Yan, P., Wang, Q. & Asmutola, P. (2017): CRISPR/Cas9-induced Targeted Mutagenesis and Gene Replacement to Generate Long-shelf Life Tomato Lines. Sci. Rep. 7(1): 11874.
Zeilmaker, T., Ludwig, N.R., Elberse, J., Seidl, M.F., Berke, L., Van Doorn, A., Schuurink, R.C., Snel, B. & Van den Ackerveken, G. (2015): DOWNY MILDEW RESISTANT 6 and DMR6-LIKE OXYGENASE 1 are partially redundant but distinct suppressors of immunity in Arabidopsis. Plant J. 81(2): 210–222.
Zeng, W., Huang, H., Lin, X., Zhu, C., Kosami, K., Huang, C., Zhang, H., Duan, C.-G., Zhu, J.-K. & Miki, D. (2021): Roles of DEMETER in regulating DNA methylation in vegetative tissues and pathogen resistance. J. Integr. Plant Biol. 63(4): 691–706.
Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., Koonin, E.V. & Zhang, F. (2015): Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell 163(3): 759–771.
Zhai, J., Bischof, S., Wang, H., Feng, S., Lee, T., Teng, C., Chen, X., Park, S.Y., Liu, L., Gallego-Bartolome, J., Liu, W., Henderson, I.R., Meyers, B.C., Ausin, I. & Jacobsen, S.E. (2015): A One Precursor One siRNA Model for Pol IV-Dependent siRNA Biogenesis. Cell 163(2): 445–455.
Zhang, H., Gong, Z. & Zhu, J.-K. (2022): Active DNA demethylation in plants: 20 years of discovery and beyond. J. Integr. Plant Biol. 64(12): 2217–2239.
Zhang, H., Lang, Z. & Zhu, J.-K. (2018): Dynamics and function of DNA methylation in plants. Nat. Rev. Mol. Cell Biol. 19(8): 489–506.
Zhang, H., Ma, Z.-Y., Zeng, L., Tanaka, K., Zhang, C.-J., Ma, J., Bai, G., Wang, P., Zhang, S.-W., Liu, Z.-W., Cai, T., Tang, K., Liu, R., Shi, X., He, X.-J. & Zhu, J.-K. (2013): DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of Pol IV. Proc. Natl. Acad. Sci. 110(20): 8290–8295.
Zhang, H. & Zhu, J.-K. (2012): Active DNA Demethylation in Plants and Animals. Cold Spring Harb. Symp. Quant. Biol. 77161–173.
Zhang, M., Liu, Q., Yang, X., Xu, Jinhua, Liu, G., Yao, X., Ren, R., Xu, Jian & Lou, L. (2020): CRISPR/Cas9-mediated mutagenesis of Clpsk1 in watermelon to confer resistance to Fusarium oxysporum f.sp. niveum. Plant Cell Rep. 39(5): 589–595.
Zhang, Y., Liang, Z., Zong, Y., Wang, Y., Liu, J., Chen, K., Qiu, J.-L. & Gao, C. (2016): Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. Nat. Commun. 7(1): 12617.
Zhang, Y., Ma, X., Xie, X. & Liu, Y.-G. (2017): Chapter Eight - CRISPR/Cas9-Based Genome Editing in Plants, in: Weeks, D.P., & Yang, B. (Eds.), Progress in Molecular Biology and Translational Science, Gene Editing in Plants. Academic Press, pp. 133–150.
Zheng, X., Chen, L., Xia, H., Wei, H., Lou, Q., Li, M., Li, T. & Luo, L. (2017): Transgenerational epimutations induced by multi-generation drought imposition mediate rice plant's adaptation to drought condition. Sci. Rep. 7(1): 39843.
Zheng, Y., Li, J., Wang, B., Han, J., Hao, Y., Wang, S., Ma, X., Yang, S., Ma, L., Yi, L. & Peng, W. (2020): Endogenous Type I CRISPR-Cas: From Foreign DNA Defense to Prokaryotic Engineering. Front. Bioeng. Biotechnol. 8.
Zhu, J., Kapoor, A., Sridhar, V.V., Agius, F. & Zhu, J.-K. (2007): The DNA Glycosylase/Lyase ROS1 Functions in Pruning DNA Methylation Patterns in Arabidopsis. Curr. Biol. 17(1): 54–59.
Zhu, J.-K. (2009): Active DNA Demethylation Mediated by DNA Glycosylases. Annu. Rev. Genet. 43143–166.

DECLARATION

on authenticity and public assess of mater's thesis

Student's name:	Mohammad Umar
Student's Neptun ID:	GBX14C
Title of the document:	Making and characterization of *dna demethylase 1* mutant in tomato
Year of publication:	2023
Department:	Department of Plant Biotechnology

I declare that the submitted master's thesis is my own, original individual creation. Any parts taken from an another author's work are clearly marked, and listed in the table of contents.

If the statements above are not true, I acknowledge that the Final examination board excludes me from participation in the final exam, and I am only allowed to take final exam if I submit another master's thesis.

Viewing and printing my submitted work in a PDF format is permitted. However, the modification of my submitted work shall not be permitted.

I acknowledge that the rules on Intellectual Property Management of Hungarian University of Agriculture and Life Sciences shall apply to my work as an intellectual property.

I acknowledge that the electric version of my work is uploaded to the repository sytem of the Hungarian University of Agriculture and Life Sciences.

Gödöllő, 2023-04-27

Student's signature

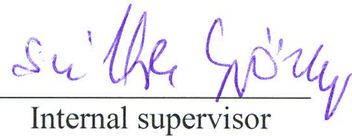
STATEMENT ON CONSULTATION PRACTICES

As a supervisor of **Mohammad Umar** (GBX14C), I here declare that the master's thesis has been reviewed by me, the student was informed about the requirements of literary sources management and its legal and ethical rules.

I recommend/don't recommend the master's thesis to be defended in a final exam.

The document contains state secrets or professional secrets: yes no

Gödöllő, 2023-04-27



Internal supervisor