Comparison of genome alterations and phenotypic variations between random mutagenesis and site-directed genome editing in tomato plants

by

Gergely Motolai

B.S., Kansas State University, 2020

AN ABSTRACT OF A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree.

DOCTOR OF PHILOSOPHY

Department of Horticulture and Natural Resources

College of Agriculture

KANSAS STATE UNIVERSITY

Manhattan, Kansas

2024

Abstract

The enhancement of crop performance and adaptability in response to the challenges posed by an expanding global population and climate variability is considered pivotal through plant breeding. This study investigates the comparative impacts of ethyl methanesulfonate (EMS) mutagenesis and CRISPR-Cas9 gene editing on tomato plants. EMS, a chemical mutagen known for inducing random mutations, is compared with CRISPR-Cas9, which offers precise, targeted genomic alterations. A purple tomato phenotype monitoring system was established to assess these techniques' mutational and phenotypic effects on transgenic Rubion tomato lines containing the *Del/Ros1* gene.

Whole genome sequencing revealed that EMS-treated lines exhibited a higher frequency of single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) compared to CRISPR-edited lines, which displayed fewer off-target effects. Phenotypic, nutrient, and phytochemical content analyses indicated no significant differences between EMS and CRISPR treatments, with variations attributed to genetic and environmental factors. This study underscores the complementary roles of EMS and CRISPR in plant breeding, highlighting their respective advantages in inducing genetic diversity and achieving targeted genetic modifications.

The findings have significant implications for advancing agricultural innovation and informing regulatory frameworks, ensuring the development of safe, nutritionally enhanced crops. This research contributes to the development of resilient crops, supporting sustainable agricultural practices and addressing global food security challenges.

Comparison of genome alterations and phenotypic variations between random mutagenesis and site-directed genome editing in tomato plants

by

Gergely Motolai

B.S., Kansas State University, 2020

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree.

DOCTOR OF PHILOSOPHY

Department of Horticulture and Natural Resources

College of Agriculture

KANSAS STATE UNIVERSITY

Manhattan, Kansas

2024

Approved by:

Major Professor

Sunghun Park

Abstract

The enhancement of crop performance and adaptability in response to the challenges posed by an expanding global population and climate variability is considered pivotal through plant breeding. This study investigates the comparative impacts of ethyl methanesulfonate (EMS) mutagenesis and CRISPR-Cas9 gene editing on tomato plants. EMS, a chemical mutagen known for inducing random mutations, is compared with CRISPR-Cas9, which offers precise, targeted genomic alterations. A purple tomato phenotype monitoring system was established to assess these techniques' mutational and phenotypic effects on transgenic Rubion tomato lines containing the *Del/Ros1* gene.

Whole genome sequencing revealed that EMS-treated lines exhibited a higher frequency of single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) compared to CRISPR-edited lines, which displayed fewer off-target effects. Phenotypic, nutrient, and phytochemical content analyses indicated no significant differences between EMS and CRISPR treatments, with variations attributed to genetic and environmental factors. This study underscores the complementary roles of EMS and CRISPR in plant breeding, highlighting their respective advantages in inducing genetic diversity and achieving targeted genetic modifications.

The findings have significant implications for advancing agricultural innovation and informing regulatory frameworks, ensuring the development of safe, nutritionally enhanced crops. This research contributes to the development of resilient crops, supporting sustainable agricultural practices and addressing global food security challenges

Table of Contents

List of Figures

List of Tables

List of Supplementary Figures

Acknowledgments

I want to express my deepest gratitude to everyone who has supported my studies and professional development during my time at K-State. The family-like environment in the Horticulture and Natural Resources department has been fundamental to my academic journey.

Foremost, I extend my most sincere appreciation to Dr. Sunghun Park, my Principal Investigator and mentor. Dr. Park's guidance has been indispensable, offering insightful advice on professional development. His critical approach to research outcomes has profoundly influenced my scientific perspective. Dr. Park's warm and kind demeanor, coupled with his genuine concern for his students' well-being, is truly inspiring. Without his support, I would not be where I am today. Thank you, Dr. Park.

I am fortunate to have had the support of my current and past lab members. Dr. Tej Man Tamang, Dr. Tayebeh Kakeshpour, Kim Park, Dr. Aaron Avalos-Calleros, and Nathan Stewart contributed significantly to my research success. Our lab’s collaborative and supportive atmosphere has made this journey more enjoyable and rewarding.

I am equally grateful to my committee members, each of whom has played a unique and crucial role in my academic development. Dr. Kimberly Williams has been pivotal throughout my academic journey, from undergraduate to graduate studies. Her mentorship and dedication to student learning have significantly shaped my professional growth. Her expertise in greenhouse production and nutrient management was instrumental in ensuring the success of my greenhouse experiments.

Dr. CB Rajashekar, my supervisor for Hort 350 lab instruction, has deepened my expertise in plant science. His extensive knowledge of phytochemical content and environmental stress has tremendously benefited my research. Thank you, Dr. Rajashekar, for your advice and support.

I am also deeply indebted to Dr. Harold N. Trick for his plant transformation and molecular biology expertise. His approachable nature and innovative suggestions have significantly advanced my research. Thank you, Dr. Trick.

I sincerely thank Dr. Chad Miller for his investment in my academic career. His support and encouragement have been vital to my scholarly and teaching development.

I want to thank our greenhouse manager, Jacob Hueste, for his countless help and efforts to help me maintain our greenhouses and or plants. Without his help, it would have been challenging to succeed.

I want to thank the other faculty members who had a significant impact on me. Dr. Stuart Sprague has been a vital part of my experience as an instructor for Hort 350 labs. Collaborating with him to develop lab curricula and assist students has honed my skills in educational methodologies and practical horticulture. His support and guidance have been invaluable.

I would like to express my gratitude to the teachers who have influenced me and had a great impact on my studies over the years. Without their dedication and excellent teaching, I couldn't be here. I would like to say thank you to Vargáné Tünde, Isaias Mendez, Loretta Shull, and Dr. Raymond Cloyd.

Finally, I am profoundly grateful to my family and friends for their unwavering support. My mom's encouragement and belief in my abilities have been a constant source of strength. Thank

you to everyone who has been part of this incredible journey. Your support and guidance have made all the difference.

Dedication

To my mom Anna and my sister Veronika, whose unwavering support, encouragement, and sacrifices have provided the foundation for all my endeavors. Your love and guidance have been the cornerstone of my journey. To my extended family, thank you for your support and understanding throughout this journey. Your presence in my life has constantly reminded me of the importance of family and love. This work is a testament to your endless love, support, and belief in me. I could not have accomplished this without you.

With all my love and gratitude,

Gergely Motolai

Chapter I

Comparison of Genome Alterations and Phenotypic Variations Between Random Mutagenesis and Site-Directed Genome Editing in Tomato Plants

Abstract

The enhancement of crop performance and adaptability in response to the challenges posed by an expanding global population and climate variability is considered pivotal through plant breeding. This study investigated the comparative impacts of ethyl methanesulfonate (EMS) mutagenesis and CRISPR-Cas9 gene editing on tomato plants. EMS, a chemical mutagen known for inducing random mutations, was compared with CRISPR-Cas9, which offers precise, targeted genomic alterations. A monitoring system utilizing the purple tomato phenotype was established to assess these techniques' mutational and phenotypic effects on transgenic Rubion tomato lines containing the *Del/Ros1* gene. It was revealed through whole genome sequencing that EMS-treated lines exhibited a higher frequency of single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) compared to CRISPR-edited lines, which displayed fewer off-target effects. Phenotypic, nutrient, and phytochemical content analyses indicated that no significant differences were observed between EMS and CRISPR treatments, with variations attributed to genetic and environmental factors. The complementary roles of EMS and CRISPR in plant breeding were underscored by this study, highlighting their respective advantages in inducing genetic diversity and achieving targeted genetic modifications. The findings have significant implications for advancing agricultural innovation and informing regulatory frameworks, ensuring the development of safe, nutritionally enhanced crops. This research contributes to the development of resilient crops, thereby supporting sustainable agricultural practices and addressing global food security challenges.

Introduction

Plant breeding is essential for enhancing crop performance and adaptability to meet the needs of an expanding global population. It plays a crucial role in improving crop yields, resilience, and environmental adaptability, which are critical as climate variability increases (Raina et al., 2016). By incorporating plant breeding into multi-crop systems, breeders can develop varieties that are robust against pests, diseases, and climatic stresses while sustaining high productivity levels. This strategy not only increases the efficiency of crop production but also promotes the ecological sustainability of agricultural systems through the utilization of genetic diversity. Consequently, advanced plant breeding techniques are crucial for agricultural innovation and sustaining elevated productivity under a rapidly changing climate.

Plant breeders utilize a spectrum of methodologies to cultivate desirable traits in crops. Classical breeding, which forms the foundation of this field, involves selecting and crossbreeding plants with favorable characteristics over successive generations to reinforce desired traits (Acquaah, 2012; Charlesworth, 2006; Heslot et al., 2012). In the field of mutagenesis, chemical agents such as ethyl methanesulfonate (EMS) are employed for their ability to induce high rates of mutation. EMS specifically targets guanine bases, causing GC to AT transitions that frequently result in significant genetic modifications, including missense mutations or complete gene knockouts (Henry et al., 2014; Greene et al., 2003). These mutations can be identified using techniques like TILLING, enabling breeders to select plants based on precise phenotypic traits (Till et al., 2003).

Technological advances have paved the way for transgenic methods, where foreign DNA is inserted into the plant genome to impart new characteristics such as enhanced pest resistance or

improved drought tolerance (James, 2015). Moreover, the advent of genome editing tools like CRISPR-Cas9 has revolutionized plant breeding. CRISPR-Cas9 facilitates precise, targeted DNA modifications by guiding the Cas9 enzyme to specific genomic sites, where it creates double-strand breaks. These breaks are typically repaired in a manner that alters or deletes sequences at the target location, enabling the cultivation of plants with specifically engineered traits that enhance agricultural efficiency and resilience (Gaj et al., 2016; Lim et al., 2016).

Chemical mutagens such as EMS are widely used for their ability to induce high-frequency mutations with relative ease (Sikora et al., 2011). EMS specifically alkylates guanine bases, leading to the erroneous replacement of cytosine with thymine during DNA replication, facilitating GC to AT transitions (Sega, 1984; Lawly & Martin, 1975; Jiang & Ramachandran, 2010). Such nucleotide transitions predominantly result in missense or knockout mutations. These mutations can be identified by phenotype screening or TILLING techniques (Henry et al., 2014). In Arabidopsis, EMS treatment has been shown to introduce mutations predominantly as G/C to A/T transitions, accounting for 99.5% of observable mutations (Greene et al., 2003). Beyond these transitions, EMS can also induce transversions, insertions, and deletions (indels), which can affect one or several nucleotides (Henry et al., 2014; Minoia et al., 2010). For example, in mutated rice populations, EMS treatments have resulted in a mutation distribution of 70% GC to AT, 11% AT to GC, 4% GC to TA, and 15% AT to TA (Till et al., 2007)—notably, an EMS treatment in Solanum lycopersicum cv. Red Setter led to a significant number of transversions, including various types of nucleotide substitutions (Minoia et al., 2010). Substantial indels were also observed in EMS-treated rice, including deletions that resulted in the complete loss of 31 genes, with a deletion frequency of 6.25% observed in mutated wheat (Henry et al., 2014; Chen et al., 2020).

Recent research highlights the significant potential of mutagenesis techniques in expanding genetic diversity and facilitating plant breeding. For instance, a comprehensive genome-wide survey of artificial mutations induced by ethyl methanesulfonate (EMS) and gamma rays in the tomato Micro-Tom genome revealed a diverse spectrum of mutations, including single nucleotide variations and insertion/deletion mutations. The study identified 5920 induced mutations, with EMS predominantly causing C/G to T/A transitions, while gamma rays resulted in a broader range of mutations, including deletions and various transversions. Interestingly, over 90% of these mutations were located in intergenic regions, indicating their potential to generate novel genetic variations without severely impacting essential gene functions (Shirasawa et al., 2015). Such findings underscore the utility of mutagenesis in generating a wide array of genetic variants, which can be harnessed in plant breeding programs to develop crops with improved traits and enhanced resilience to environmental stresses.

In contrast, recent advancements in genetic engineering have favored the use of targeted nucleases such as transcription activator-like effector nucleases (TALENs), zinc-finger nucleases (ZFNs), homing endonucleases or meganucleases, and the CRISPR-Cas9 system (Gaj et al., 2016). The CRISPR-Cas9 system, noted for its simplicity and versatility, has been particularly transformative in the field of targeted genome editing. It employs a two-component mechanism: the Cas9 nuclease, which induces double-strand breaks, and a single-guide RNA (sgRNA) that directs the Cas9 to specific DNA sequences (Gaj et al., 2016; Lim et al., 2016). While the resulting double-strand breaks are typically repaired through non-homologous end joining, leading to indels, the specificity of the CRISPR-Cas9 system allows for precise gene editing (Gaj et al., 2016). This technology has facilitated the development of genetically modified organisms with desired traits, such as the 'Sicilian Rouge High GABA' tomato, the first CRISPR-edited commercial plant

product (Waltz, 2022; Ezura, 2022). However, despite its specificity, off-target activity remains a significant concern (Tycko et al., 2019; Young et al., 2019; Manghwar et al., 2020).

In this study, we established a purple tomato monitoring system to elucidate the effects and potential unintended consequences of CRISPR and EMS techniques on plant genomes and phenotypes. We utilized tomato fruit color as a visible marker, which was altered using either CRISPR/Cas9-mediated gene editing or EMS-induced mutation in transgenic Rubion tomato lines containing the *Del/Ros1* gene under the control of the E8 promoter (Butelli et al., 2008). The reason for developing this purple fruit phenotype monitoring system is to mimic a breeding program focused on identifying one specific phenotype—in this case, the loss of purple fruit color, which indicates a reversion to the red fruit color. By monitoring for this specific change, we can effectively compare the off-target mutations caused by different breeding methods. This comparison is particularly valuable in lines exhibiting the desired traits, as it allows for a comprehensive analysis of the genetic stability and potential unintended consequences of these methods. This dual approach enabled us to compare the specificity and mutational outcomes of CRISPR and EMS treatments, contributing to a deeper understanding of their respective impacts on targeted and off-target genetic regions. This approach facilitates a deeper understanding of the efficacy and precision of various breeding techniques, ultimately contributing to the refinement of methods used to achieve specific phenotypic outcomes in crops.

Materials and methods

Generation of Purple Tomato Monitoring System

The purple tomato was created using genetic engineering techniques by introducing the *Delila* (*Del*) and *Rosea1* (*Ros1*) genes from snapdragon (Antirrhinum majus) into tomato plants. These genes, which encode transcription factors that induce anthocyanin biosynthesis, result in purple pigmentation. The *Del* and *Ros1* genes were incorporated into pCLD04541 binary expression vector, which harbored the *Del/Ros1* construct with fruit-specific E8 promoter to drive the expression of the genes and a kanamycin resistance gene as a selectable marker. The binary vector was introduced into A. tumefaciens strain LBA4404, which was then used to transform tomato cotyledons. The explants were co-cultivated with *Agrobacterium* for 3 days before being transferred to a selection medium containing 100mg/L kanamycin and 250 mg/L clavacin (Clavamox) to eliminate non-transformed cells and *Agrobacterium*, respectively. Regenerated shoots were subsequently selected based on kanamycin resistance. This work was done by Dr. Tayebeh Kakeshpour and Dr. Kim Park

Selection of Target Sequences and Vector Construction

Two distinct vector constructs were engineered, each harboring an identical set of gRNAs but differing in their promoter structures. In the first construct, the gRNAs of 20 nucleotides were driven by the AtU6 promoter, and those of 18 nucleotides were controlled by the AtU6 promoter. The complete assembly of gRNA scaffolds, incorporating four sets of promoters, gRNAs, and terminators, was synthesized and cloned into the pYLCRISPR/Cas9P35S-H binary vector via the Bsa I site (Figure 1-1). Conversely, the second vector featured all gRNAs linked by tRNA spacers

and regulated by a single CmYLCV viral promoter and terminator. This gRNA scaffold was synthesized and similarly integrated into the binary vector using the Bsa I site.

For the implementation of multiplex CRISPR/Cas9 gene editing, two distinct target sites within each of the *Del* and *Ros* genes were selected. Guide RNAs (gRNAs) were systematically designed using the CRISPR-P 2.0 platform and CRISPR direct software to ensure specificity and minimize potential off-target effects, which were validated using CHOPCHOP v3. The gRNAs, varying in length—18 and 20 nucleotides—were strategically chosen to enhance editing efficiency (Figure 1-2)

These vectors were introduced into a kanamycin-resistant transgenic tomato line expressing *Del* and *Ros* genes, with hygromycin serving as the selection marker in the constructs.

EMS Mutagenesis of Transgenic Purple Tomato Lines

Chemical mutagenesis was conducted using ethyl methanesulfonate (EMS) on transgenic purple tomato lines. Initially, seeds were treated in two EMS concentrations (0.7% and 1% v/v) for 12 hours at 22°C with gentle agitation. Following three subsequent treatments with 1% EMS for 16 hours each, the seeds were extensively washed and germinated under controlled conditions. The M1 generation was cultivated in BM6 media within a greenhouse setup, maintaining a temperature regime of 24/20°C (day/night) and a 16-hour light cycle. Supplemental lighting was provided using high-pressure sodium (HPS) lamps to ensure a 16-hour photoperiod, with light intensity maintained at approximately 300 μmol m^{-2} s^{-1} PAR (Photosynthetically Active Radiation) to maintain optimal growing conditions. Regular fertilization was carried out using Miracle-Gro Tomato Plant Food following manufacturer recommendations.

From the M1 generation, 814 independent lines were advanced to the fruit ripening phase for phenotypic screening. M2 seeds were extracted, treated with 2% bleach, and stored at 4°C for subsequent testing. A total of 492 lines were selected from which 5880 M2-generation plants were evaluated for fruit color phenotype in BM6 media under similar greenhouse conditions. Plants expressing the red fruit phenotype were selected for further testing, while plants with the purple fruit phenotype and other non-desirable characteristics were not selected for additional testing.

DNA Isolation

Total DNA was extracted from several Rubion lines, including transgenic, segregated nontransgenic, EMS-treated transgenic, and CRISPR-Cas9 edited lines, using the DNeasy® Plant Mini Kit. Quality and quantity assessments of the DNA samples were performed using a NanoDrop ND-1000 spectrophotometer and electrophoresis on a 1% TopVision Agarose gel. In 2022, high-quality DNA samples from the selected lines were sent to Novogene Corporation Sacramento CA for whole genome sequencing

Whole genome sequencing SNP and INDEL Identification

Whole genome sequencing was conducted at Novogene using Illumina NovaSeq X Plus and NovaSeq 6000 platforms, generating 150 bp paired-end reads. To identify Single Nucleotide Polymorphisms (SNPs) and Insertions/Deletions (INDELs), we employed the Genome Analysis Toolkit (GATK) version 4.1.0.0, chosen for its robust performance in variant detection. Our analysis pipeline mandated a minimum of 100 sequencing reads per variant across all samples to minimize the impact of sequencing errors and ensure reliable variant detection. Variant filtration parameters were set with an allele frequency minimum of 0.2 to exclude rare potential artifacts, and spatial clustering of variants was controlled by allowing no minimum distance between features. The heterozygous non-reference calls were constrained with a lower bound of zero and

an upper bound of 0.6 to balance sensitivity and specificity. Additionally, variants with more than 20% missing data were excluded to maintain the integrity and statistical power of our analyses. The genotypic constitution was stringently monitored, limiting homozygous allele frequencies to 0-1% and allowing heterozygosity in a range of 0-60%, thus including plausible heterozygous variants while excluding outliers likely resulting from technical anomalies.

Phenotypic Data Collection

Phenotypic parameters, including height, fruit yield, size, and nutrient content, were assessed alongside phytochemical analyses. The third-generation (T2) CRISPR knockout and third-generation (M3) EMS-treated tomato plants were initially germinated in 4-inch pots using PRO-MIX BX Mycorrhizae general-purpose media (Premier et al.). Upon reaching 15-20 cm in height, the plants were transferred to 1-gallon pots and cultivated under optimal greenhouse conditions with temperatures maintained at 24/20°C (day/night). Supplemental lighting was provided using high-pressure sodium (HPS) lamps to ensure a 16-hour photoperiod, with light intensity maintained at approximately 300 μmol m^{-2} s^{-1} PAR (Photosynthetically Active Radiation) to maintain optimal growing conditions. Weekly fertilization was conducted using Miracle-Gro Tomato Plant Food (ScottsMiracle-Gro) following manufacturer guidelines.

At 90 days post-germination, plant height was measured, and growth was monitored until day 120, when fruit yield and individual fruit size and weight were recorded. The harvested fruits were oven-dried at 60°C for seven days, ground into a fine powder using a coffee grinder, and shipped to the Kansas State University soil testing lab for analysis of mineral content (N, P, K, Ca, Mg, Na, S, Fe, Cu, Mn, and B).

Phytochemical Analysis

Carotenoid Compound Measurement

Lycopene and beta-carotene concentrations were determined using a method established by Nagata and Yamashita (1992). Frozen samples were homogenized in an Oster #6694 blender, and approximately one gram of the homogenate was vigorously shaken with 16 mL of an acetone-hexane solution (2:3 ratio) in a 50 mL test tube. After freezing the mixture at -20°C for one hour to facilitate phase separation, 350 µL of the supernatant was transferred into a 96-well microplate (Costar #3364, Corning, MA). Absorbance was measured at 663, 645, 505, and 453 nm using a Synergy H1 spectrophotometer Lycopene and beta-carotene concentrations were calculated with specified formulas and expressed as µg/g dry weight.

Flavonoids and Phenolics Compound Measurement

To determine the concentrations of flavonoids and phenolics, High-Performance Liquid Chromatography (HPLC) was employed. Sample preparation involved homogenizing the previously frozen biological material using an Oster #6694 blender to ensure uniformity. Approximately 0.5 grams of each homogenized sample were extracted with 10 mL of 80% methanol at room temperature under constant agitation for 24 hours. The extracts were then filtered through a 0.45 µm PTFE filter (Millipore et al., USA) to remove particulate matter.

Chromatographic separation was achieved using an Agilent 1260 Infinity HPLC system (Agilent et al., USA) equipped with a reverse-phase C18 column (250 mm x 4.6 mm, 5 µm particle size). The mobile phase consisted of a gradient mixture of water with 0.1% formic acid (solvent A) and acetonitrile with 0.1% formic acid (solvent B). The gradient program started with 10% B, which was linearly increased to 50% B over 30 minutes, held at 50% B for 5 minutes, and then

returned to 10% B over 5 minutes for column re-equilibration. The flow rate was maintained at 1 mL/min, and the injection volume was 20 µL.

Detection of the analytes was performed using a diode-array detector with data collection at multiple wavelengths appropriate for flavonoids (280 nm, 330 nm) and phenolics (280 nm). Identification of individual compounds was based on a comparison of their retention times and UV-Vis spectra with those of known standards. Quantification was carried out using calibration curves established with pure standards.

Statistical Analysis

All statistical analysis was done using GraphPad Prism 10.2.2. Outliers were identified using ROUT with Q=1% for all samples. One-way was performed using 95% confidence interval with the combination of Tukey multiple comparison test on the dataset.

Control Experiment for Genetic Screening

Approximately 700 transgenic Line 7 *Del/Ros1* expressing transgenic tomato plants were germinated and grown under the previously described greenhouse conditions. Upon reaching 10 cm in height and displaying three true leaves, one leaf tissue sample was taken per plant. Samples were pooled by combining similarly sized tissue from five individuals, promptly frozen in liquid nitrogen, and ground with a chilled mortar and pestle. DNA was extracted using the CTAB method. The resulting DNA samples were quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA), and assessed for quality on a 1% TopVision Agarose gel (Thermo et al., USA). PCR amplification of the partial segment of *Del/Ros1* was performed using Q5® High-Fidelity 2X Master Mix (New England BioLabs, Ipswich, MA, USA) to screen for the deletion of the *Ros1* gene in the EMS lines.

Results

SNP and INDEL Mutation Frequency of EMS and CRISPR Lines Compared to *Del/Ros1* Expressing Transgenic Plants

The EMS treatments resulted in varying numbers of SNPs and INDELs across different lines (Table 1-1). Specifically, EMS 13, EMS 60, EMS 61, and EMS 92 lines exhibited significant genetic variations with total mutation counts of 687, 412, 546, and 2076, respectively. Large deletions were observed in the *Del/Ros1* construct (Table 1-2) in the tested EMS lines, which was observed between the repeating CMV terminator sequences (Figure 1-3). The largest mutation load was observed in EMS 92, predominantly consisting of SNPs (1981 SNPs), indicating a high mutagenic effect of the EMS treatment. (Figure 1-4)

In contrast, CRISPR-edited lines showed more targeted mutations. The lines CRISPR_pCMY 1, CRISPR_pCMY 2, CRISPR_pCMY 3, CRISPR_pCMY 4, CRISPR_pATU 1, CRISPR_pATU 8, and CRISPR_pATU 10 displayed lower total SNP counts of 248, 240, 7, 315, 364, 342, and 345, respectively. Notably, the CRISPR_pCMY 3 line showed the least mutations, emphasizing the precision of CRISPR technology in generating specific genetic alterations with insertions and deletions with minimal off-target effects. (Table 1-3)

Flavonoids and Phenolics Content Comparison of EMS, CRISPR, and WT

The comparative analysis of chicoric acid, chlorogenic acid, kaempferol, and naringenin across CRISPR, EMS, and WT treatments revealed distinct patterns of phytochemical accumulation. Average concentrations and standard errors were determined to establish the presence of statistically significant differences between treatments (see Figures 1-5, 6, 7, and 1-8).

In the case of chlorogenic acid, CRISPR (11.71 mg/g DW) and EMS (11.70 mg/g DW) did not differ significantly from each other. However, they were both significantly different from WT (14.23 mg/g DW), which showed a substantially higher average concentration. When chlorogenic acid, levels are observed at individual lines, they behave similarly except for the wild-type plants (Figure 1-5)

For chicoric acid, the EMS treatment resulted in a significantly higher average concentration (9.69 mg/g DW) compared to both WT (7.07 mg/g DW) and CRISPR (8.00 mg/g DW). When chicoric acid levels are observed at individual lines, they all behave similarly(Figure 1-6).

The kaempferol content revealed that the EMS treatment (11.13 mg/g DW) was significantly lower compared to WT (11.31 mg/g DW) and CRISPR (11.23 mg/g DW). When kaempferol levels are observed at individual lines, they show a similar range. From this, we can interpolate that the observed differences are due to variations in greenhouse growing conditions and not due to the treatments (Figure 1-7). Lastly, naringenin levels were significantly lower in WT samples (6.60 mg/g DW) compared to CRISPR (6.92 mg/g DW) and EMS (6.91 mg/g DW), signifying that WT is less efficient in naringenin accumulation than the other treatments. When naringenin levels are broken down by individual lines, they behave similarly except for the wild-type plants.(Figure 1-8).

Carotenoids Content Comparison of EMS, CRISPR, and WT

The concentration of lycopene and beta-carotene in CRISPR, EMS, and wild-type (WT) samples was measured in milligrams per gram of dry weight (mg/g DW), and the standard errors were calculated to evaluate the consistency of the measurements.

For lycopene, CRISPR samples demonstrated a significantly higher average concentration (0.164 mg/g DW) compared to EMS samples (0.119 mg/g DW). However, the difference between CRISPR and WT (0.124 mg/g DW) was not statistically significant, indicating that genetic editing does not affect the natural variation in lycopene content. (Figure 1-9) When observed individually, all lines behave similarly, with the exception of line pATU1.

In contrast, the average beta-carotene content did not significantly differ among the three treatments. WT samples had the highest average beta-carotene concentration (0.088 mg/g DW), but this was not statistically distinct from CRISPR (0.078 mg/g DW) or EMS (0.075 mg/g DW) treatments (Figure 1-10). The natural variation observed in the individual lines is expected under greenhouse conditions.

Nutrient Comparison of EMS, CRISPR, and WT

Regarding Total Carbon (C%), EMS presented a significantly lower average (42.384%) when compared to WT (44.986%). However, the difference between EMS and CRISPR (42.845%) was not statistically significant. CRISPR and EMS shared similar C% levels, suggesting comparable carbon assimilation processes between these two treatments, with natural variation expected in a greenhouse environment. (Figure 1-11)

Total Nitrogen (N%) exhibited no significant differences among the treatments despite the highest mean value being reported for EMS (4.078%) followed by WT (4.238%) and CRISPR (3.701%)(Figure 1-12). Similarly, the analysis revealed no significant variance in the

concentrations of Phosphorus (P%), Potassium (K%), Sulfate-Sulfur (SO4-S%), Magnesium (Mg%), and Copper (Cu) among the different treatments. This suggests a homogenous distribution of these nutrients regardless of the genetic treatment applied. The observed natural variation is likely caused by the environmental conditions of the greenhouse. (see Figure 1-13, 14, 15, 16, 17)

Significant disparities were noted in the Calcium (Ca%) content, where WT samples (0.115%) were significantly lower than those from CRISPR (0.252%) and EMS (0.219%). It is noteworthy that the Rubion tomatoes are highly susceptible to blossom end rot, which can cause higher-than-expected variances in Calcium levels. When looking at lines individually, they behave similarly but higher variations are observed from line to line compared to other nutrients. (Figure 1-16) This suggests a potential differential regulation of calcium uptake or sequestration in the WT variant, likely influenced by the greenhouse environment.

In the context of trace elements, while Copper (Cu) concentrations did not differ significantly, a notable contrast was observed in Iron (Fe) content. EMS (89.405 ppm) was significantly higher compared to WT (64.26 ppm), yet no significant difference was observed when compared to CRISPR (93.821 ppm) (Figure 1-18 and 1-19). Furthermore, the comparison between WT and CRISPR yielded no significant difference, indicating a conserved iron accumulation mechanism in these two treatments. The natural variation observed in these measurements is likely due to the greenhouse environment.

Phenotypic Comparison of EMS, CRISPR, and WT

The average fruit weight per plant followed a comparable trend, with EMS-treated plants producing heavier fruits (average = 43.10 g) than CRISPR-treated plants (average = 32.43 g). WT plants had an average fruit weight (average = 42.12 g) that did not differ significantly from CRISPR or EMS treatments. (Figure 1-20) Consistency within the EMS treatment was evidenced by the lowest standard error (SE = 1.64), with natural variation expected in a greenhouse environment.

The number of fruits per plant was significantly higher in EMS-treated plants (average = 10) compared to CRISPR-treated plants (average = 7.18), but (average = 9.33) did not show a significant difference from either CRISPR or EMS treatments, suggesting a moderate response to genetic treatments concerning fruit count. (Figure 1-21) The natural variation in fruit production is likely influenced by the greenhouse environment. The standard error was lowest in EMS (SE = 0.61), indicating a consistent outcome within this treatment group.

Similarly, the sum of fresh weight per plant or yield showed significant variation between CRISPR (average = 214.48 g) and EMS (average = 430.19 g) treatments. WT plants exhibited an intermediate yield (average = 404.44 g) that was not significantly different from either CRISPR or EMS treatments despite the higher variability suggested by the standard error (SE = 83.89). The observed natural variation in yield is expected under greenhouse conditions due to the individual lines behaving fairly similarly, with only pCMY3 and pATU1 being significantly lower compared to the other lines. (Figure 1-22)

Notably, average plant height revealed a distinct divergence in phenotypic expression between treatments. CRISPR-treated plants were significantly shorter (average = 63.57 cm) than both EMS (average = 75.77 cm) and WT (average = 77.17 cm). (Figure 1-23) Meanwhile, EMS

and WT did not differ significantly from each other. The observed natural variation in plant height is likely influenced by the greenhouse environment based on the individual lines behaving similarly, with the exception of the pCMY 3 line.

Discussion

This study provides an extensive comparison between CRISPR-Cas9 mediated gene editing and ethyl methanesulfonate (EMS) mutagenesis in tomato plants, elucidating their respective impacts on genetic and phenotypic alterations. Our results highlight the precision and targeted nature of CRISPR-Cas9 compared to the broader mutational spectrum induced by EMS, demonstrating the strengths and limitations of each technique.

Genetic Mutations and Specificity

The whole genome sequencing results underscore the distinct mutational profiles of CRISPR and EMS treatments. CRISPR lines exhibited fewer mutations with higher specificity, aligning with previous findings that emphasize its precision in creating targeted genetic modifications (Gaj et al., 2016; Kamburova et al., 2017). In contrast, EMS lines displayed a higher frequency of SNPs and indels, reflecting the random nature of chemical mutagenesis (Henry et al., 2014; Greene et al., 2003). The significant number of SNPs in EMS-treated lines, particularly the EMS 92 line, highlights the extensive genetic variability that EMS can introduce, which can be both an advantage and a limitation depending on breeding objectives (Jiang & Ramachandran, 2010; Sikora et al., 2011).

Notably, our results show that while CRISPR technology is highly specific, off-target effects remain a concern. Studies have documented unintended modifications at genomic sites with high sequence similarity to target sites, which can lead to unpredictable phenotypic outcomes

(Manghwar et al., 2020; Tycko et al., 2019). Advances in CRISPR technology, such as high-fidelity Cas9 variants, are being developed to mitigate these off-target effects, further enhancing the precision of genome editing (Haeussler et al., 2020).

EMS-Induced Large Deletions

Ethyl methanesulfonate (EMS) is a potent chemical mutagen widely used in plant breeding due to its ability to induce random mutations at a high frequency. EMS is particularly effective at causing large deletions, especially in regions with repetitive sequences, resulting in substantial genetic variation and phenotypic diversity, which are critical for breeding programs (Henry et al., 2014).

EMS induces mutations by alkylating guanine bases, leading to GC to AT transitions. When repetitive sequences, such as tandem repeats or transposable elements, are present, the repair processes following EMS-induced damage can cause misalignment and unequal crossing over during homologous recombination. This often results in deletions encompassing one or more repeats (Jiang & Ramachandran, 2010).

In this study, EMS treatment on transgenic tomato lines resulted in significant deletions within the *Del/Ros1* construct due to the repetitive nature of the construct design, particularly between the regions with repetitive CMV terminator sequences (see Figure 1-4). Recombination-based DNA repair is a crucial cellular mechanism for ensuring genetic stability by fixing breaks or errors in DNA, thereby preventing mutations. Double-strand breaks (DSBs) in DNA, caused by EMS or other factors such as radiation, chemicals, or replication errors, pose a significant threat as these breaks can result in the loss of genetic information or chromosomal rearrangements (Miller et al., 2021). These deletions observed in our study are likely due to the DNA repair mechanisms responding to the breakages caused by EMS.

Homologous recombination (HR) is primarily used for repairing these breaks, relying on a homologous sequence, typically a sister chromatid available during the S and G2 phases of the cell cycle, as a template for accurate repair. The HR process begins with the recognition of the DSB, followed by the processing of the broken DNA ends to produce single-stranded DNA (ssDNA) tails. These tails then invade a homologous DNA sequence, forming a displacement loop (D-loop) with the help of proteins like RAD51. The invading strand is extended by DNA polymerase using the homologous sequence as a template. The resulting structure, often a Holliday junction, undergoes branch migration and resolution to complete the repair (Van der Auwera et al., 2008).

In contrast, non-homologous end joining (NHEJ) is used as an alternative pathway that directly ligates the broken ends without requiring a homologous template. Although NHEJ is more error-prone than HR and is often utilized when cells are not in the S or G2 phases, it provides a quick fix to DNA breaks. The choice between HR and NHEJ is tightly regulated to maintain genomic stability, with crucial roles played by proteins like BRCA1 and BRCA2 in this regulation (Gorbunova et al., 1997). Overall, recombination-based DNA repair, particularly through homologous recombination, is vital for preserving genome integrity and preventing the accumulation of mutations that can lead to serious genetic mutations.

The use of EMS as a chemical mutagen has been extensively documented to induce a high frequency of mutations, particularly large deletions in regions with repetitive sequences. A comprehensive genome-wide survey of EMS-induced mutations in the tomato genome by Shirasawa et al. (2015) demonstrated that EMS treatment resulted in a significant number of mutations, including 47 deletion mutations, among the 4690 total mutations identified. These mutations were predominantly located in repeat-rich regions, which highlights the tendency of EMS to cause deletions in these areas. Furthermore, the deletions observed in the study were

associated with regions of the genome that contain repetitive sequences, which are prone to misalignment and unequal crossing over during DNA repair processes (Shirasawa et al., 2015). This propensity for inducing large deletions in repetitive sequences underscores the utility of EMS in creating substantial genetic variation, which is invaluable for plant breeding and functional genomics studies.

Whole genome sequencing confirmed these deletions, illustrating the mutagenic potential of EMS. For instance, the EMS 92 line exhibited a high mutation load, including large deletions, highlighting the extensive genomic alterations induced by EMS (Henry et al., 2014).

While large deletions can naturally occur, EMS significantly increases their likelihood. This is beneficial in functional genomics studies, where knocking out large genomic regions can reveal gene functions and interactions. In breeding programs, these large deletions can introduce new alleles with desirable traits, enhancing genetic diversity. However, the randomness of EMS-induced mutations necessitates thorough screening to identify beneficial changes while avoiding deleterious effects (Greene et al., 2003; Chang et al., 2008; Broun & Tanksley, 1996)).

Moreover, EMS-induced large deletions can be utilized to study regulatory regions and non-coding DNA, which are often enriched with repetitive sequences. Altering these regions through EMS mutagenesis can provide insights into their functional dynamics (Sikora et al., 2011).

Phenotypic Variations

Phenotypic analysis revealed that the variations in traits such as fruit production, plant height, and lycopene content are likely due to natural variance rather than the specific effects of CRISPR or EMS treatments. When broken down into individual lines, both CRISPR and EMS treatments displayed similar behaviors, suggesting that the observed phenotypic differences are

more attributable to inherent genetic diversity and environmental factors than to the treatments themselves.

Nutrient and Phytochemical Content

Similarly, the variations in nutrient and phytochemical content across treatments can be attributed to natural variance. The testing methodology also played a role in the observed variance. Whole fruit samples were used for phytochemical content, which is less accurate than measurements using only the fruit's pericarp(Yahia, 2017). The higher calcium and iron content observed in EMS-treated plants and the higher lycopene content in CRISPR-treated plants do not significantly deviate from the expected natural variation within individual lines. This suggests that environmental conditions and intrinsic genetic factors play a more substantial role in influencing these traits than the mutagenesis treatments.

Genotypic Variation

The genotypic variation observed in this study provides critical insights into the underlying genetic mechanisms that contribute to trait development in tomato plants. EMS-induced mutations result in a broad spectrum of genetic changes, including SNPs and indels, which can lead to a wide array of phenotypic outcomes (Henry et al., 2014). This extensive genetic diversity can be harnessed to identify novel alleles associated with desirable traits, thereby enriching the genetic pool available for breeding programs (Greene et al., 2003).

CRISPR-Cas9, on the other hand, facilitates precise, targeted modifications, resulting in fewer unintended genetic changes. This specificity allows for the direct manipulation of genes known to influence important traits, streamlining the breeding process (Gaj et al., 2016). The reduced off-target effects observed in CRISPR-treated lines further underscore its potential for generating stable, predictable genetic modifications (Tycko et al., 2019).

Implications for Plant Breeding

The results of this study underscore the complementary roles of CRISPR and EMS in plant breeding. CRISPR's precision and efficiency in creating specific, targeted genetic changes make it an invaluable tool for developing crops with enhanced traits and reduced off-target effects (Gaj et al., 2016; Tycko et al., 2019). The successful use of CRISPR in creating the 'Sicilian Rouge High GABA' tomato, the first CRISPR-edited commercial plant product, highlights the potential for this technology in agriculture (Waltz, 2022; Pandita, 2023).

However, the broad mutational spectrum of EMS can uncover novel phenotypes and enhance traits that might be overlooked in targeted approaches (Greene et al., 2003; Roychowdhury et al, 2013). For example, EMS mutagenesis has been used to develop drought-tolerant wheat varieties by inducing mutations in key genes involved in stress response (Sikora et al., 2011). Therefore, integrating both methodologies could provide a balanced approach, leveraging the strengths of each to achieve comprehensive improvements in crop performance and resilience.

Impact on Regulatory Agencies and Food Safety

The implications of this research extend to regulatory agencies tasked with ensuring food safety and environmental sustainability. The adoption of CRISPR and EMS in crop development presents distinct regulatory challenges and opportunities.

CRISPR technology, with its precision and specificity, has sparked significant interest and debate among regulatory bodies. The European Union, for instance, has classified organisms developed using CRISPR as genetically modified organisms (GMOs), subjecting them to stringent regulatory frameworks (European Court of Justice, 2018). In contrast, the United States Department of Agriculture (USDA) has opted for a more lenient approach, excluding certain

CRISPR-edited crops from GMO regulations if they do not contain foreign DNA (USDA, 2020). The findings from this study, particularly the lower off-target effects, and specific trait enhancements could inform regulatory decisions by highlighting the safety and efficacy of CRISPR technology in producing high-quality, nutritionally superior crops.

EMS mutagenesis, being a more traditional method, typically faces fewer regulatory hurdles. However, the broad spectrum of mutations it induces necessitates thorough screening to ensure food safety. Regulatory agencies such as the Food and Drug Administration (FDA) and the European Food Safety Authority (EFSA) require comprehensive safety assessments, including allergenicity and toxicity evaluations, for new crop varieties (EFSA, 2011; FDA, 2015). The higher nutrient and phytochemical content observed in EMS-treated plants in this study could support the case for their safety and nutritional benefits, potentially easing regulatory approvals.

The integration of advanced genomic techniques and chemical mutagenesis in plant breeding could also facilitate the development of crops that meet specific regulatory standards for enhanced nutritional quality and environmental sustainability. For instance, biofortified crops with higher levels of essential nutrients could address micronutrient deficiencies in populations, aligning with public health goals and regulatory priorities (Saltzman et al., 2013). Moreover, crops with improved stress resistance could contribute to sustainable agricultural practices, reducing the need for chemical inputs and mitigating the environmental impact of farming (Tilman et al., 2011).

Future Directions

Future research should focus on optimizing the integration of CRISPR and EMS techniques to harness their full potential in developing crops that can meet the challenges of global food security and climate change. Combining these approaches with advanced phenotyping and genomic selection methods could accelerate the development of superior crop varieties (Hickey et

al., 2017). Additionally, exploring the use of other genome editing tools, such as TALENs and ZFNs, in conjunction with chemical mutagenesis could further expand the possibilities for plant breeding (Gaj et al., 2016).

Moreover, it is essential to address the regulatory and ethical considerations surrounding the use of genome editing technologies in agriculture. Public perception and acceptance of genetically modified crops remain significant challenges that must be addressed through transparent communication and evidence-based policy-making (Qaim, 2020). Ensuring the safety and sustainability of these technologies will be crucial for their successful adoption in global agriculture.

To further assess the safety of non-desirable mutations, lines produced by EMS with unwanted characteristics should be evaluated and compared to lines with chromothripsis induced by CRISPR-Cas9 (Leibowitz et al., 2021). This comparison aims to understand better the potential off-target mutations and their negative consequences associated with these two methods.

Conclusion

This study highlights the distinct advantages and limitations of CRISPR and EMS in plant breeding, providing valuable insights into their impacts on genetic and phenotypic variation in tomato plants. The precision of CRISPR-Cas9 and the extensive mutational capacity of EMS offer complementary tools for advancing plant breeding. Future research should focus on optimizing the integration of these techniques to harness their full potential in developing crops that can meet the challenges of global food security and climate change.

References

Bourgaud, F., Gravot, A., Milesi, S., & Gontier, E. (2001). Production of plant secondary metabolites: A historical perspective. Plant Science, 161(5), 839-851.

Briat, J. F., Dubos, C., & Gaymard, F. (2015). Iron nutrition, biomass production, and plant product quality. Trends in Plant Science, 20(1), 33-40.

Broun, P., & Tanksley, S. D. (1996). Characterization and genetic mapping of simple repeat sequences in the tomato genome. Molecular and General Genetics MGG, 250, 39-49.

Butelli, E., Titta, L., Giorgio, M. et al. Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. Nat Biotechnol 26, 1301–1308 (2008). https://doi.org/10.1038/nbt.1506

Chang, S. B., Yang, T. J., Datema, E., Van Vugt, J., Vosman, B., Kuipers, A., ... & De Jong, H. (2008). FISH mapping and molecular organization of the major repetitive sequences of tomato. Chromosome research, 16, 919-933.Acquaah, G. (2012). Principles of Plant Genetics and Breeding. Wiley-Blackwell.

Charlesworth, D. (2006). Evolution of plant breeding systems. Current Biology, 16(17), R726-R735.

Chen, T., Huang, L., Wang, M., Huang, Y., Zeng, R., Wang, X., ... & Zhang, L. (2020). Ethyl methyl sulfonate-induced mutagenesis and its effects on peanut agronomic, yield and quality traits. Agronomy, 10(5), 655.

European Court of Justice. (2018). Organisms obtained by mutagenesis are GMOs and are, in principle, subject to the obligations laid down by the GMO Directive (Press Release No 111/18). Retrieved from https://curia.europa.eu

Ezura, H. Letter to the Editor : The World ' s First CRISPR Tomato Launched to a Japanese Market : The Social-Economic Impact. 2022, 63, 731–733.

Food and Drug Administration (FDA). (2015). Voluntary Labeling Indicating Whether Foods Have or Have Not Been Derived from Genetically Engineered Plants. Retrieved from https://www.fda.gov

Gaj, T., Gersbach, C. A., & Barbas, C. F. III. (2016). "ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering". Trends in Biotechnology, 31(7), 397-405.

Gorbunova, V., & Levy, A. A. (1997). Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. Nucleic Acids Research, 25(22), 4650-4657.

Greene, E. A., Codomo, C. A., Taylor, N. E., Henikoff, J. G., Till, B. J., Reynolds, S. H., ... & Henikoff, S. (2003). Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. Genetics, 164(2), 731-740.

H, Akhunova A, Akhunov E, Dubcovsky J, Tai TH, et al (2014) Efficient Genome-Wide Detection and Cataloging of EMS-Induced Mutations Using Exome Capture and Next-Generation Sequencing. Plant Cell 26: 1382–1397

Haeussler, M. (2020). CRISPR off-targets: a question of context. Cell biology and toxicology, 36(1), 5-9.

Henry, I. M., Nagalakshmi, U., Lieberman, M. C., & Comai, L. (2014). Efficient genome-wide detection and cataloging of EMS-induced mutations using exome capture and next-generation sequencing. Plant Cell, 26(4), 1382-1397.

Heslot, N., Yang, H. P., Sorrells, M. E., & Jannink, J. L. (2012). Genomic selection in plant breeding: a comparison of models. Crop science, 52(1), 146-160.

Hickey, J. M., Chiurugwi, T., Mackay, I., Powell, W., & Hickey, J. M. (2017). Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. Nature Genetics, 49(9), 1297-1303.

James, C. (2015). Global Status of Commercialized Biotech/GM Crops: 2015. ISAAA Brief No. 51. ISAAA: Ithaca, NY.

Jankowicz-Cieslak J, Till BJ (2016) Chemical Mutagenesis of Seed and Vegetatively Propagated Plants Using EMS. Curr Protoc Plant Biol. doi: 10.1002/cppb.20040

Kamburova, V. S., Nikitina, E. V., Shermatov, S. E., Buriev, Z. T., Kumpatla, S. P., Emani, C., & Abdurakhmonov, I. Y. (2017). Genome editing in plants: an overview of tools and applications. International Journal of Agronomy, 2017.

Leibowitz, M. L., Papathanasiou, S., Doerfler, P. A., Blaine, L. J., Sun, L., Yao, Y., Zhang, C. Z., Weiss, M. J., & Pellman, D. (2021). Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. Nature genetics, 53(6), 895–905. https://doi.org/10.1038/s41588-021-00838-7

Lim, Y.; Bak, S.Y.; Sung, K.; Jeong, E.; Lee, S.H.; Kim, J.S.; Bae, S.; Kim, S.K. Structural roles of guide RNAs in the nuclease activity of Cas9 endonuclease. Nat. Commun. 2016, 7, 13350.

Manach, C., Scalbert, A., Morand, C., Rémésy, C., & Jiménez, L. (2004). Polyphenols: Food sources and bioavailability. The American Journal of Clinical Nutrition, 79(5), 727-747.

Manghwar, H., Lindsey, K., Zhang, X., & Jin, S. (2020). CRISPR/Cas system: recent advances and future prospects for genome editing. Trends in Plant Science, 25(3), 528-547.

Manghwar, H.; Li, B.; Ding, X.; Hussain, A.; Lindsey, K.; Zhang, X.; Jin, S. CRISPR/Cas Systems in Genome Editing: Methodologies and Tools for sgRNA Design, Off-Target Evaluation, and Strategies to Mitigate Off-Target Effects. Adv. Sci. 2020, 7, 1902312.

Miller, V., Beying, N., Schmidt, C., & Puchta, H. (2021). Double strand break (DSB) repair pathways in plants and their application in genome engineering. In Genome editing for precision crop breeding (pp. 27-61). Burleigh Dodds Science Publishing.

Minoia, S., Petrozza, A., D'Onofrio, O., Piron, F., Mosca, G., Sozio, G., ... & Barbieri, G. (2010). A new mutant genetic resource for tomato crop improvement by TILLING technology. BMC Research Notes, 3(1), 69.

Nidhi, S.; Anand, U.; Oleksak, P.; Tripathi, P.; Lal, J.A.; Thomas, G.; Kuca, K.; Tripathi, V. Novel CRISPR–Cas Systems: An Updated Review of the Current Achievements, Applications, and Future Research Perspectives. Int. J. Mol. Sci. 2021, 22, 3327.

Pandita, D. (2023). Unlocking CRISPR/Cas-Mediated Editing Potential for Designing Climate-Smart Crop Plants. In Climate-Resilient Agriculture, Vol 1: Crop Responses and Agroecological Perspectives (pp. 873-893). Cham: Springer International Publishing.

Pham A-T, Bilyeu K, Chen P, Boerma HR, Li Z (2014) Characterization of the fan1 locus in soybean line A5 and development of molecular assays for high-throughput genotyping of FAD3 genes. Mol Breed 33: 895–907

Qaim, M. (2020). Role of new plant breeding technologies for food security and sustainable agricultural development. Applied Economic Perspectives and Policy, 42(2), 129-150.

Raina, A., Laskar, R. A., Khursheed, S., Amin, R., Tantray, Y. R., Parveen, K., & Khan, S. (2016). Role of mutation breeding in crop improvement-past, present and future. Asian Research Journal of Agriculture, 2(2), 1-13.

Rao, A. V., & Agarwal, S. (2000). Role of antioxidant lycopene in cancer and heart disease. Journal of the American College of Nutrition, 19(5), 563-569.

Roychowdhury, R., & Tah, J. (2013). Mutagenesis—A potential approach for crop improvement. Crop improvement: new approaches and modern techniques, 149-187. Lawley, P. D., & Martin, C. N. (1975). Molecular mechanisms in alkylation mutagenesis. Induced reversion of bacteriophage T4 r II AP72 by ethyl methanesulphonate in relation to extent and mode of ethylation of purines in bacteriophage deoxyribonucleic acid. Biochemical Journal, 145(1), 85-91.

Sega, G. A. (1984). A review of the genetic effects of ethyl methanesulfonate. Mutation Research/Reviews in Genetic Toxicology, 134(2-3), 113-142.

Shirasawa, K., Hirakawa, H., Nunome, T., Tabata, S., & Isobe, S. (2016). Genome-wide survey of artificial mutations induced by ethyl methanesulfonate and gamma rays in tomato. Plant biotechnology journal, 14(1), 51-60.

Sikora, P., Chawade, A., Larsson, M., Olsson, J., & Olsson, O. (2011). Mutagenesis as a tool in plant genetics, functional genomics, and breeding. International Journal of Plant Genomics, 2011, 1-13.

Till, B. J., Colbert, T., & Tompa, R. (2007). High-throughput TILLING for Arabidopsis. Methods in Molecular Biology, 553, 127-135.

Till, B. J., Reynolds, S. H., Greene, E. A., Codomo, C. A., Enns, L. C., Johnson, J. E., ... & Henikoff, S. (2003). Large-scale discovery of induced point mutations with high-throughput TILLING. Genome research, 13(3), 524-530.

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. Proceedings of the National Academy of Sciences, 108(50), 20260-20264.

Tycko, J., Myer, V. E., & Hsu, P. D. (2019). Methods for optimizing CRISPR-Cas9 genome editing specificity. Molecular Cell, 68(3), 686-702.

Tycko, J.; Wainberg, M.; Marinov, G.K.; Ursu, O.; Hess, G.T.; Ego, B.K.; Aradhana; Li, A.; Truong, A.; Trevino, A.E.; et al. Mitigation of off-target toxicity in CRISPR-Cas9 screens for essential non-coding elements. Nat. Commun. 2019, 10, 4063.

United States Department of Agriculture (USDA). (2020). SECURE Rule Frequently Asked Questions. Retrieved from https://www.usda.gov

Van der Auwera, G., Baute, J., Bauwens, M., Peck, I., Piette, D., Pycke, M., ... & Depicker, A. (2008). Development and application of novel constructs to score C: G-to-T: A transitions and homologous recombination in Arabidopsis. Plant physiology, 146(1), 22-31.

Waltz, E. GABA-enriched tomato is first CRISPR-edited food to enter market. Nat. Biotechnol. 2022, 40, 9–11.

White, P. J., & Broadley, M. R. (2003). Calcium in plants. Annals of Botany, 92(4), 487-511.

Yahia, E. M. (2017). Tomato (Solanum lycopersicum). In Fruit and Vegetable Phytochemicals. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119158042

Young, J.; Zastrow-Hayes, G.; Deschamps, S.; Svitashev, S.; Zaremba, M.; Acharya, A.; Paulraj, S.; Peterson-Burch, B.; Schwartz, C.; Djukanovic, V.; et al. CRISPR-Cas9 Editing in Maize: Systematic Evaluation of Off-target Activity and Its Relevance in Crop Improvement. Sci. Rep. 2019.

Figures and Tables

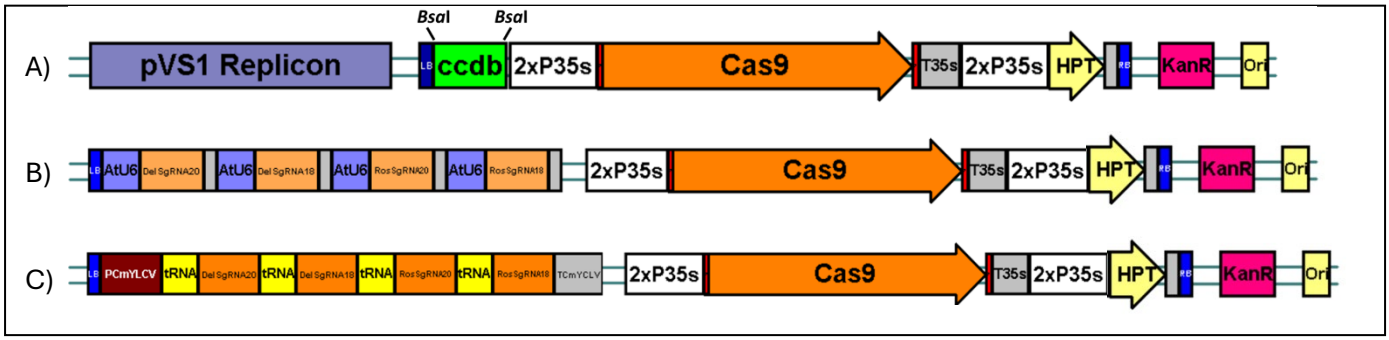


Figure 1-1) Schematic maps of the constructs. A) Backbone pYLCRISPR/Cas9P35S-H. B) Construct using the AtU6 promoter. C) Construct using the PCmYLCV promoter.

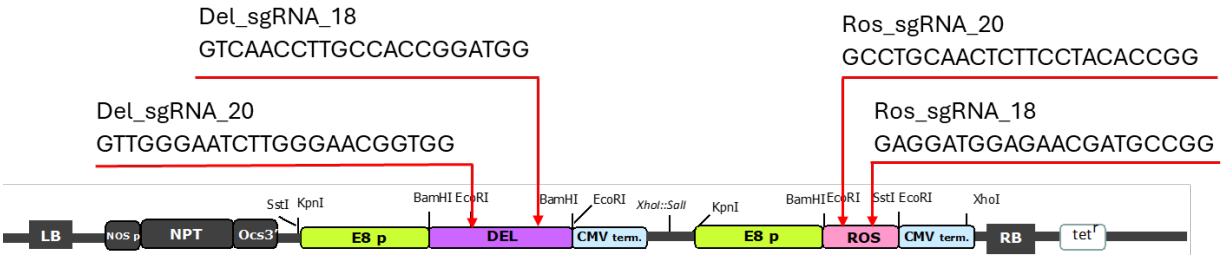


Figure 1-2) Four sgRNA sequences used in the CRISPR/Cas 9 constructs and target location within the Del/Ros1 construct.

Table 1-1) The total number of genetic variations found in analyzed sequences compared to transgenic Del/Ros1 expressing purple tomato.

SNPs and INDELs relative to the T7
(D/R expressing line)

Line	SNP	INDEL	Total	Diff (%)
EMS 13	580	107	687	1.4×10^{-3}
EMS 60	330	82	412	8.4×10^{-4}
EMS 61	460	86	546	1.1×10^{-3}
EMS 92	1981	95	2076	4.2×10^{-3}
EMS Exp 3A	267	160	427	8.6×10^{-4}
CRISPR pCMY 1	163	85	248	5.2×10^{-4}
CRISPR pCMY 2	163	77	240	4.9×10^{-4}
CRISPR pCMY 3	6	1	7	1.4×10^{-5}
CRISPR pATU 1	255	109	364	7.3×10^{-4}
CRISPR pATU 8	247	95	342	6.6×10^{-4}

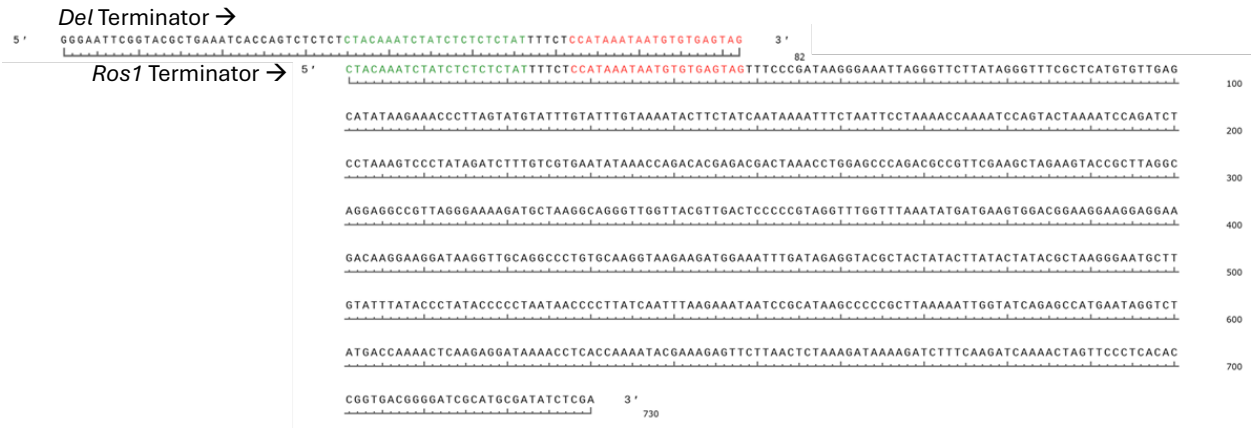


Figure1-3) CMV Terminator sequence showing from Del terminator to Ros1 terminator

Table 1-2) Location of amplified sequences found in the EMS lines at the edge of deleted locations

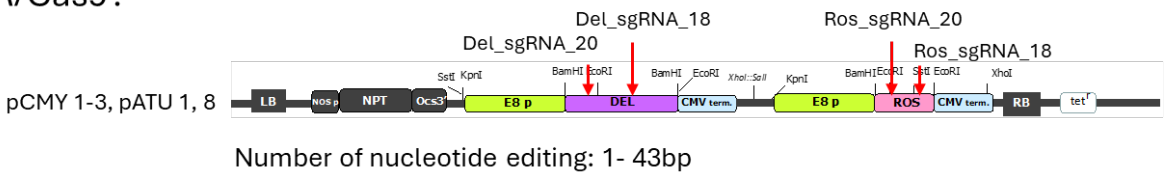
EMS Lines	5’ end	3’ end
EMS 13	5’-CCATAAANAATGTGTGAGTAG…-3’	5’-…CTACAAATCTATCTCTCTCTAT-3’
EMS 60	5’-CCATAAATAATGTGTGAGTAG…-3’	5’-…TACAAATCTATCTCTCTCTAT-3’
EMS 61	5’-CCANAAANAATGTGTGAGTAG…-3’	5’-…CTACAAATCTATCTCTCTCTAT-3’
EMS 92	5’-CCATAAATAATGTGTGAGTAG…-3’	5’-…TACAAATCTATCTCTCTCTA-3’
EMS Exp3A	5’-…ATGGTTGAACAAGATGGATTG-3’	5’-GGCGCCCGGTTCTTTTTGTCA…-3’

Table 1-3) Site-specific editing nucleotide changes and their relative locations to the PAM sequence found in the CRISPR/Cas9 edited lines.

CRISPR Line

pCMY 1	Del_sgRNA_20 5'-TTGGGAACGGTGGCAGGACTGTCCAA-3' 5'-TTGGGAA~~CGGTGG~~CAGGACTGTCCAA-3' (Deletion)
	Ros_sgRNA_20 5'-CCGGTG-TAGGAAGAGTTGCAGGC-3' 5'-CCGG~~TG~~-TAGGAAGAGTTGCAGGC-3' (Deletion) 5'-CCGGTG**T**TAGGAAGAGTTGCAGGC-3' (Insertion)
pCMY 2	Del_sgRNA_18 5'-CCATCCGGTGGCAAGGTTGACAAAGTATCAATACTAGACCATACAATAGATTACTTGAGAGGGCTTGA-3' 5'-CCATCCGGTGGCAAGGTTGACAAAGTATCAATACTAGACCATACA~~ATAGATTACTTGA~~GAGGGCTTGA-3' (Deletion)
	Ros_sgRNA_20 5'-ACACAGAGCAGGGTTGAACCGGTGTAGGAAGAGTTGCAGG-3' 5'-ACACA~~GAGCAGGGTTGAACCGGTGTAGGAAGAGTT~~GCAGG-3' (Deletion)
pCMY 3	Del_sgRNA_20 5'-CCTGCCACCG-TTCCCAAGATTCCCAACTATGTCTCC-3' 5'-CCTGCCACCG**A**TTCCCAAGATTCCCAACTATGTCTCC-3' (Insertion)
pATU 1	Del_sgRNA_20 5'-CCTGCCACCGTT-CCCAAGATTCCCAACT-3' 5'-CCTGCCACCGTT**T**CCCAAGATTCCCAAC-3' (Insertion)
	Ros_sgRNA_18 5'-CGAGGATGGAGAACGAT-GCCGG-3' 5'-CGAGGATGGAGAACGAT**T**GCCGG-3' (Insertion)
pATU 8	Del_sgRNA_18 5'-CTAGTCCCATCC-GGTGGCAAGGTTGACAA-3' 5'-CTAGTCCCATCC**G**GGTGGCAAGGTTGACAA-3' (Insertion)
	Ros_sgRNA_18 5'-CGAGGATGGAGAACGAT-GCCGG-3' 5'-CGAGGATGGAGAACGAT**T**GCCGG-3' (Insertion)

A) CRISPR/Cas9:



B) EMS:

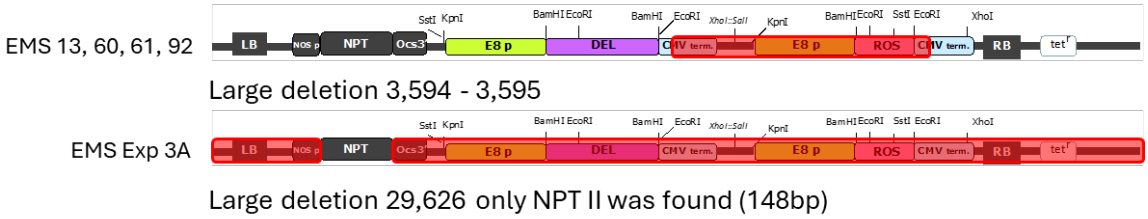


Figure 1-4) Nucleotide changes found in the Del/Ros1 construct. A) Site-specific deletions found in CRISPR/Cas9 edited lines. B) Location of large deletions EMS treated lines.

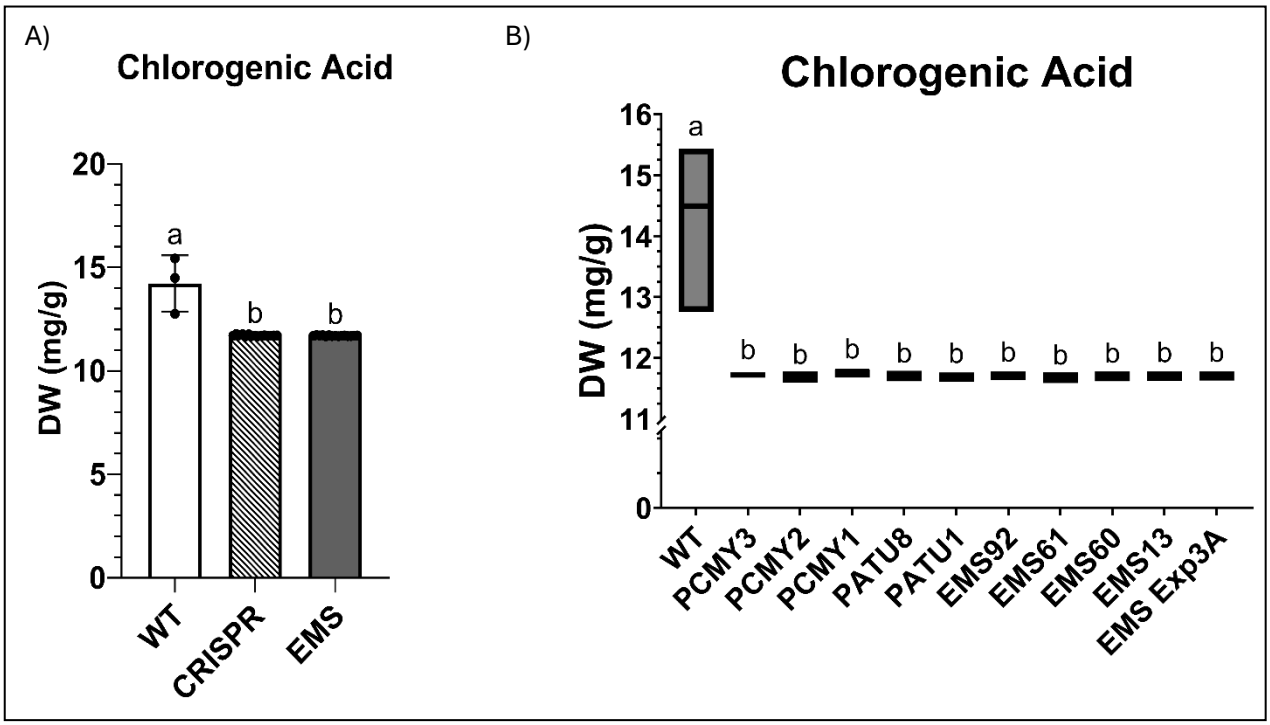


Figure 1-4) A) Combined treatment average of Chlorogenic Acid content. B) Individual line breakdown of Chlorogenic Acid content.

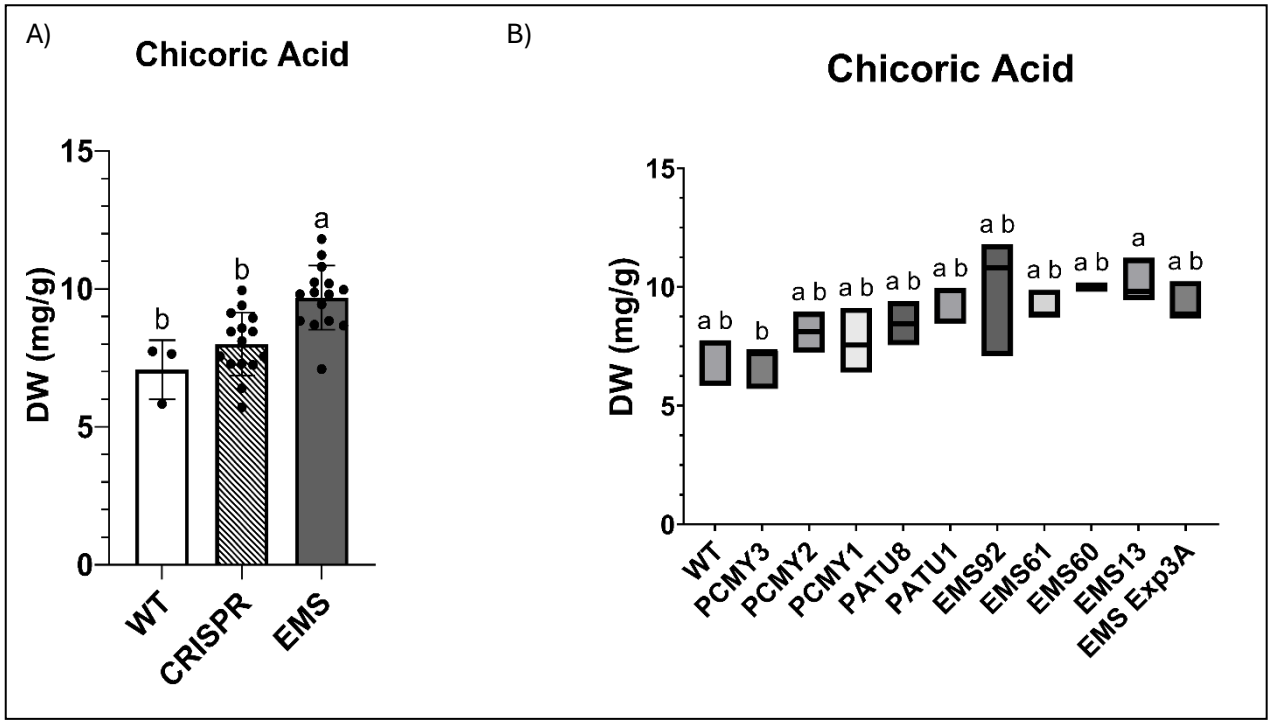


Figure 1-5) A) Combined treatment average of Chicoric Acid content. B) Individual line breakdown of Chicoric Acid content.

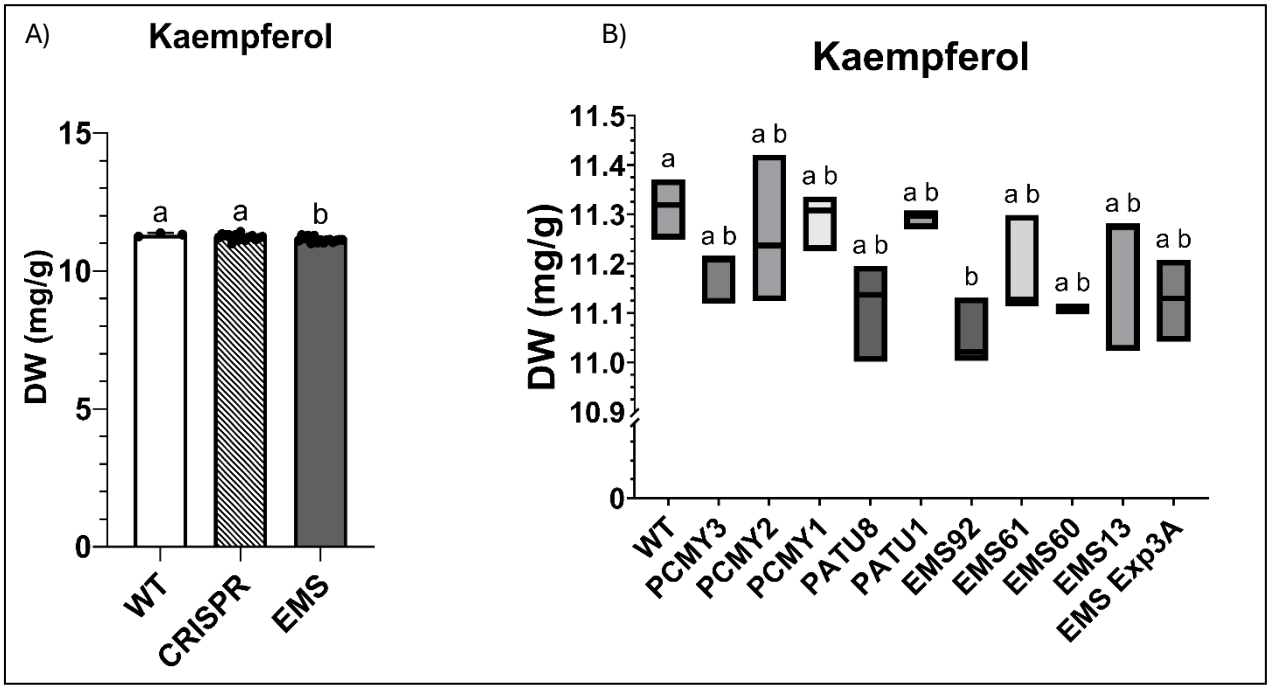


Figure 1-7) A) Combined treatment average of Kaempferol content. B) Individual line breakdown of Kaempferol content.

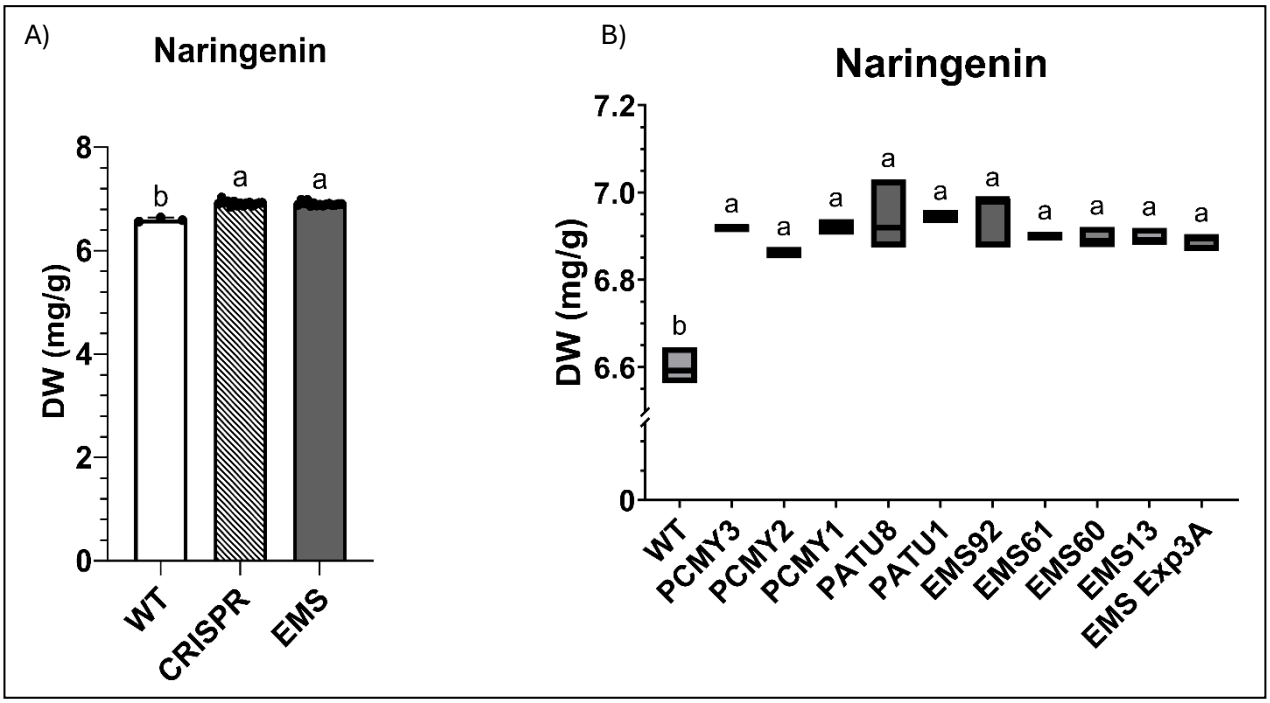


Figure 1-8) A) Combined treatment average of Naringenin content. B) Individual line breakdown of Naringenin content.

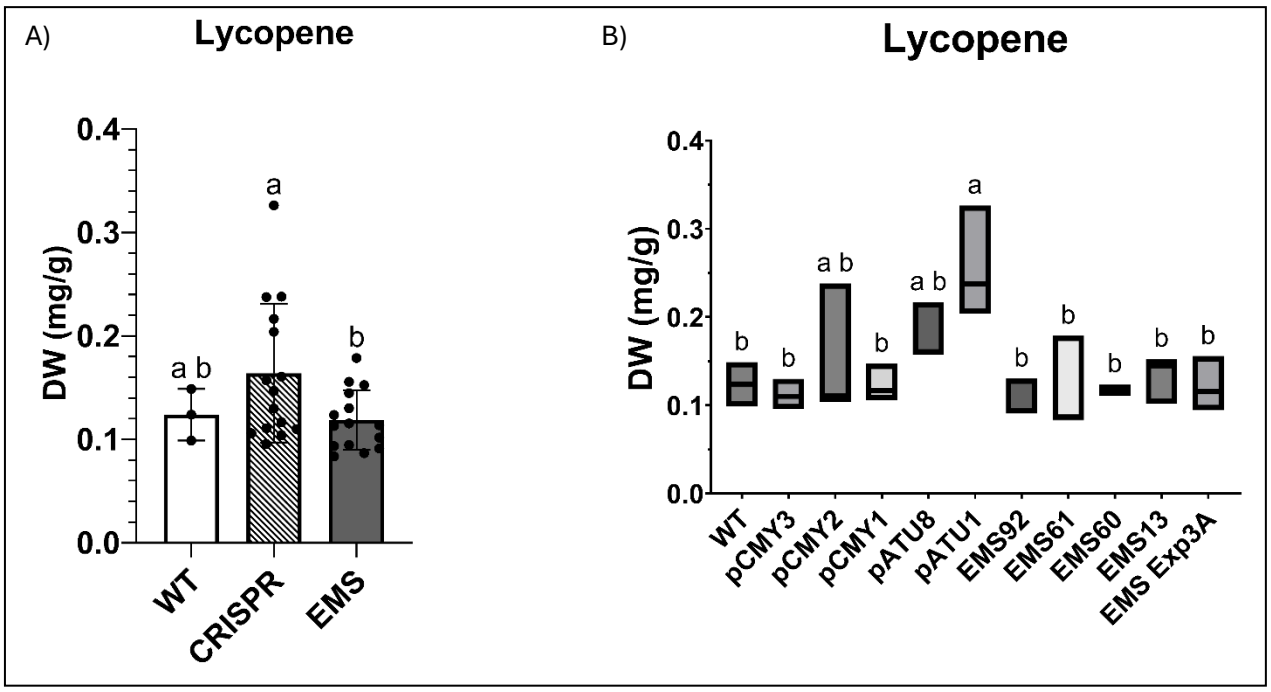


Figure 1-9) A) Combined treatment average of Lycopene content. B) Individual line breakdown of Lycopene content.

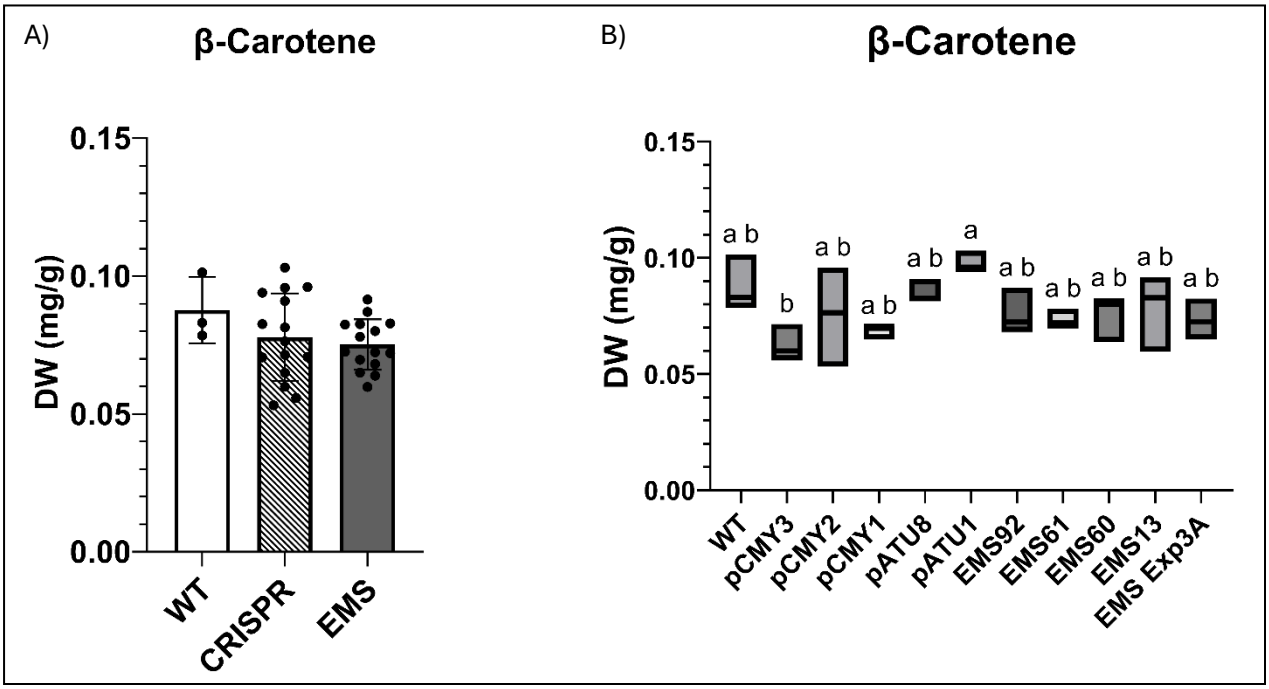


Figure 1-10) A) Combined treatment average of β-Carotene content. B) Individual line breakdown of β-Carotene content.

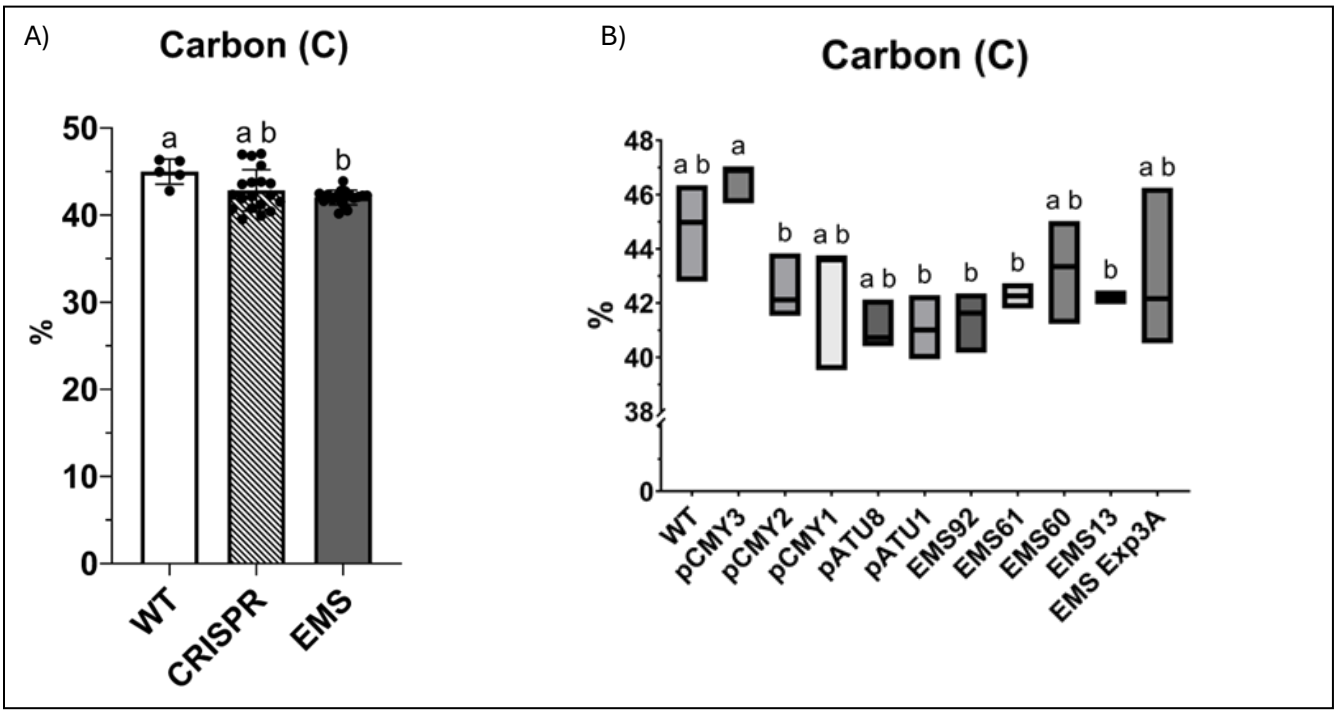


Figure 1-11) A) Combined treatment average of Carbon content. B) Individual line breakdown of Carbon content.

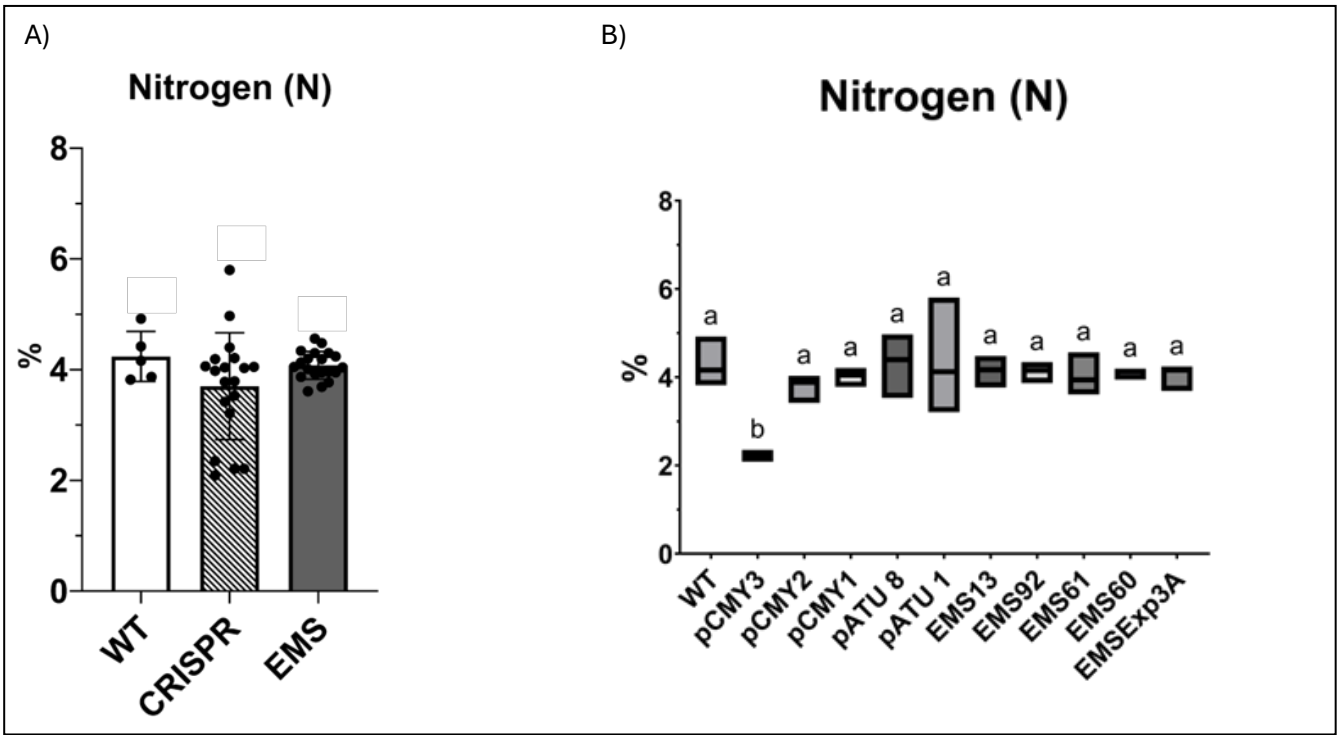


Figure 1-12) A) Combined treatment average of Nitrogen content. B) Individual line breakdown of Nitrogen content.

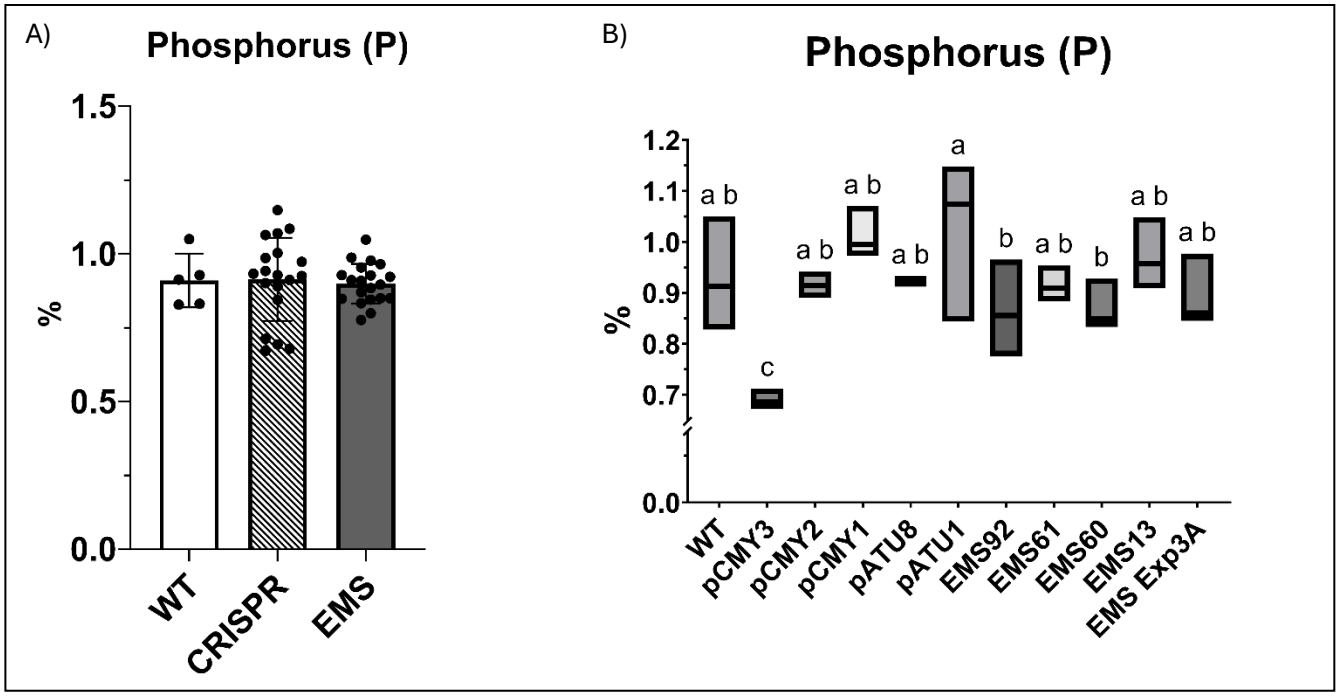


Figure 1-13) A) Combined treatment average of Phosphorus content. B) Individual line breakdown of Phosphorus content.

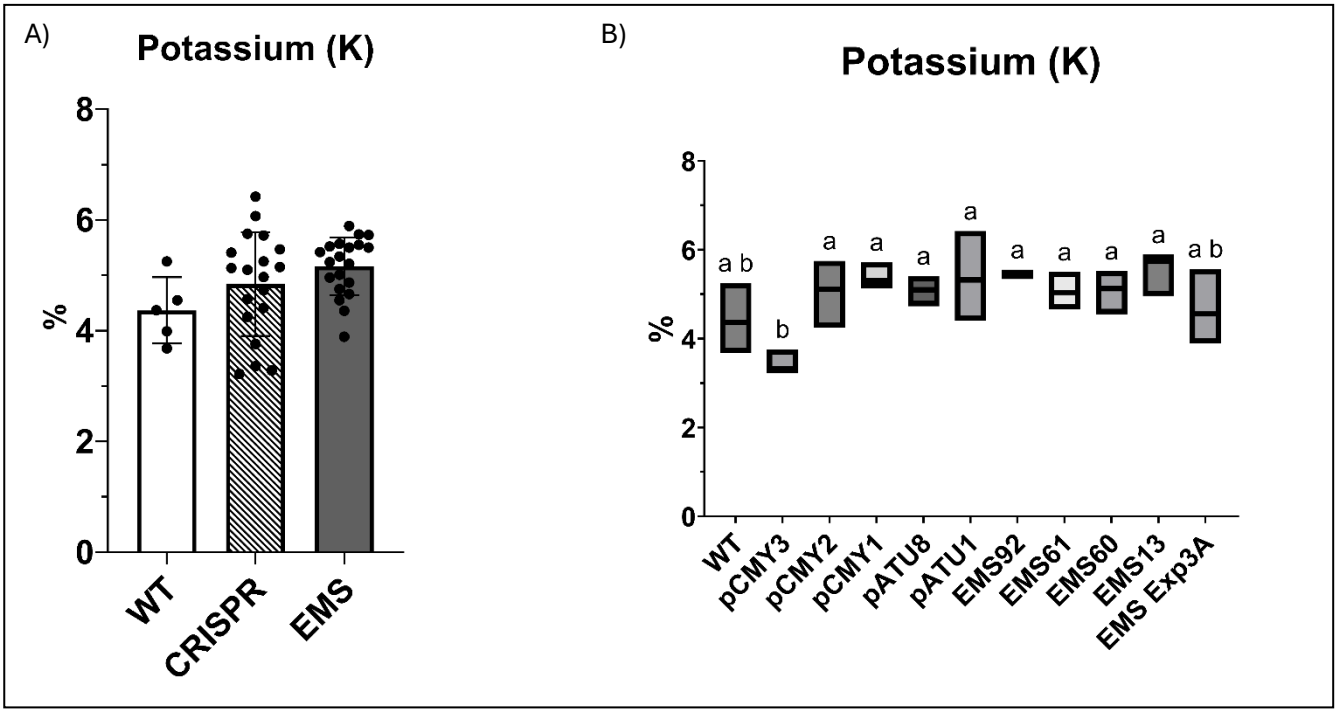


Figure 1-14) A) Combined treatment average of Potassium content. B) Individual line breakdown of Potassium content.

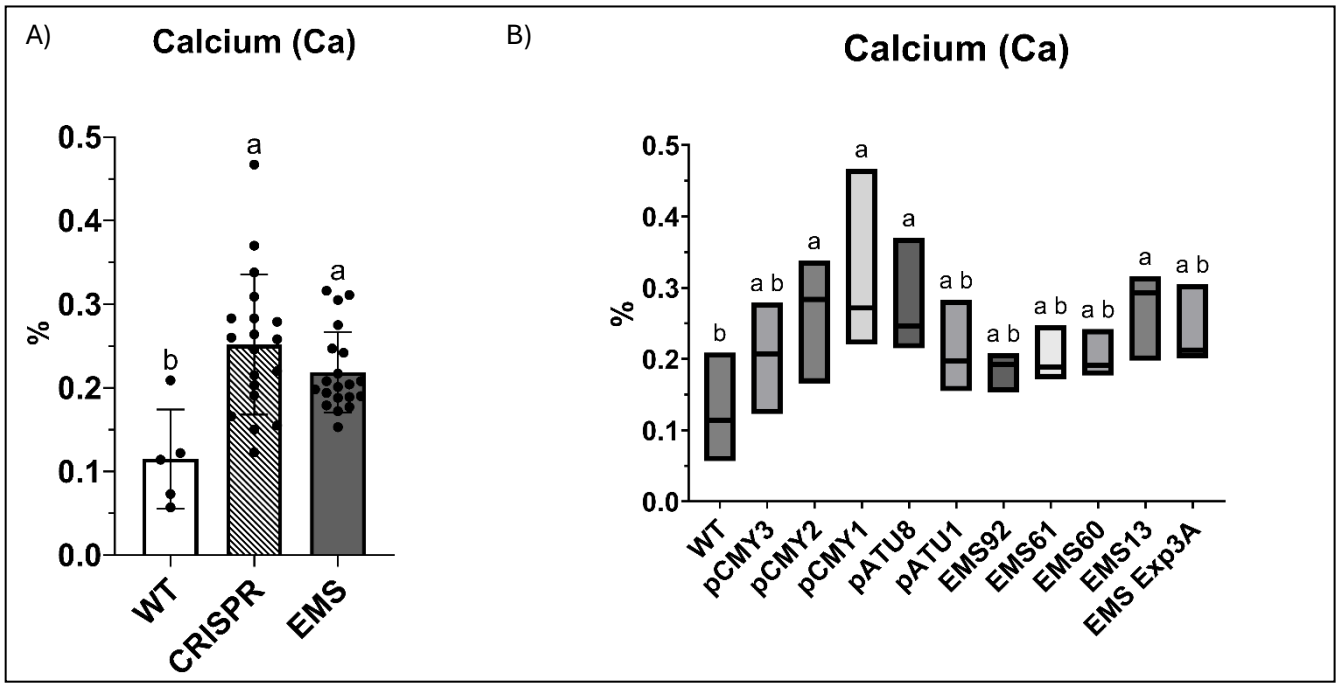


Figure 1-15) A) Combined treatment average of Calcium content. B) Individual line breakdown of Calcium content.

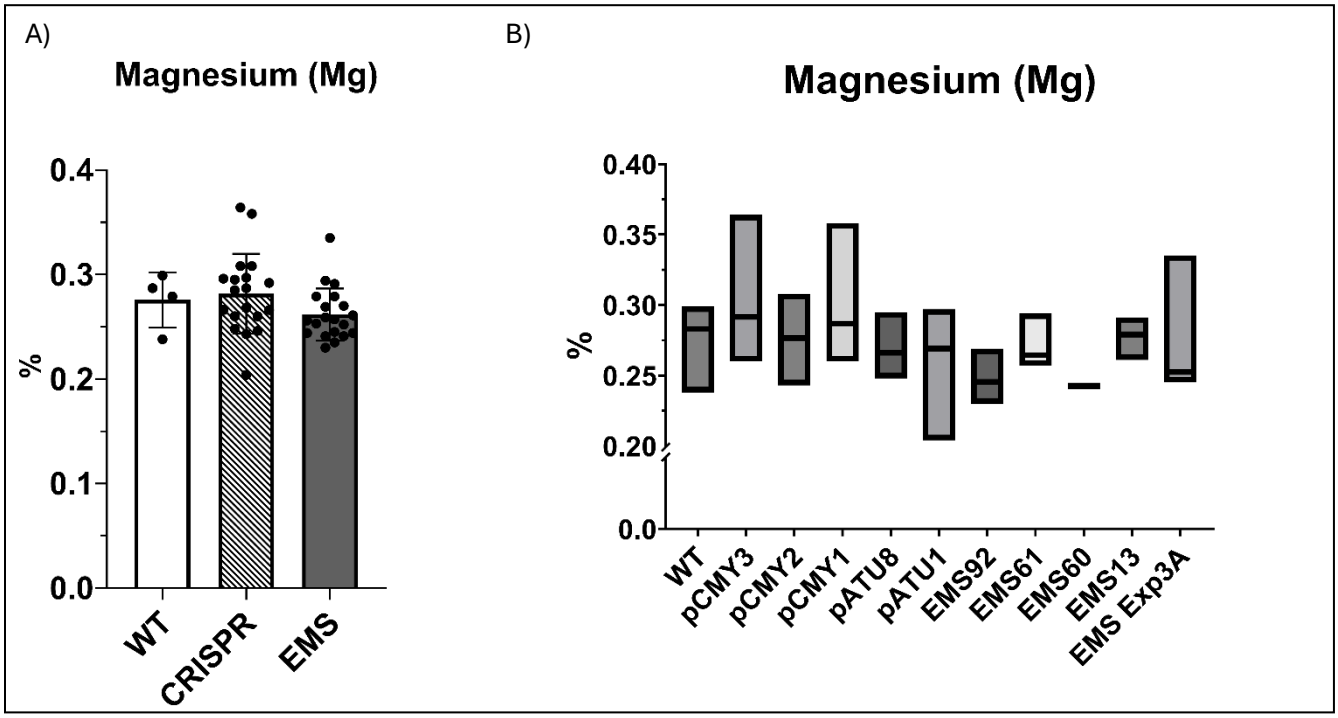


Figure 1-16) A) Combined treatment average of Magnesium content. B) Individual line breakdown of Magnesium content.

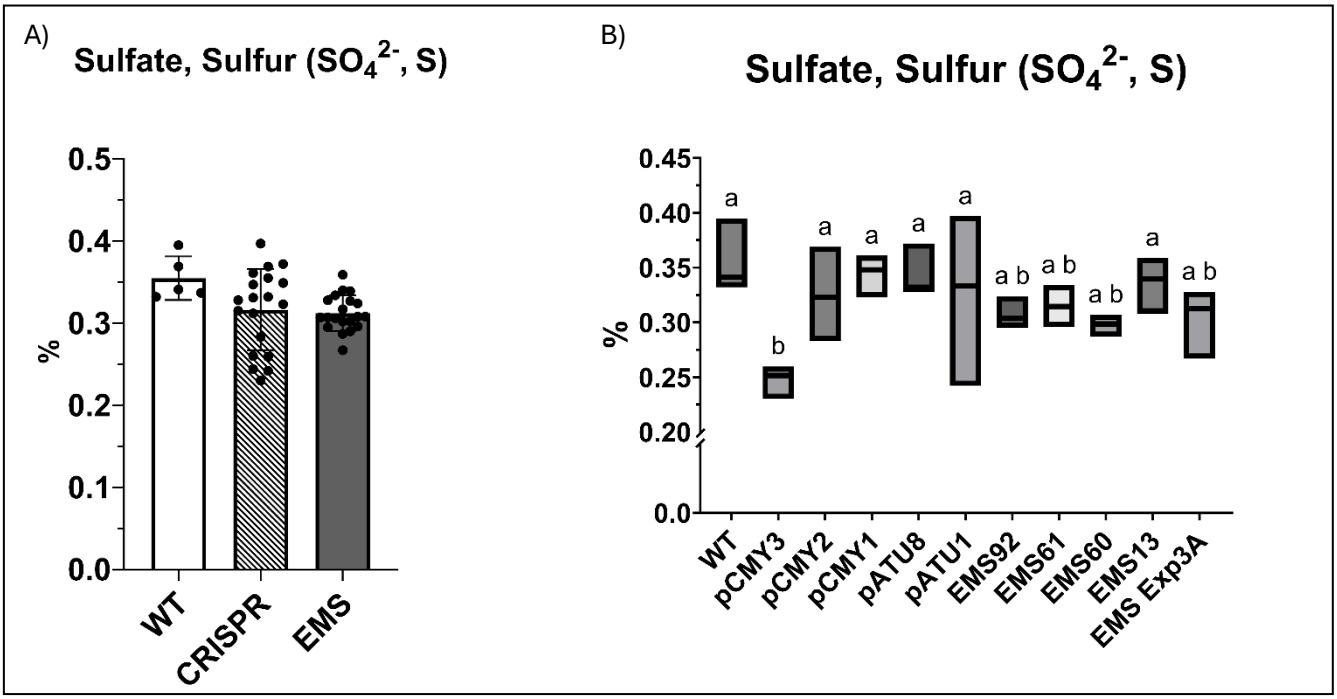


Figure 1-17) A) Combined treatment average of Sulfate, Sulfur content. B) Individual line breakdown of Sulfate, Sulfur content.

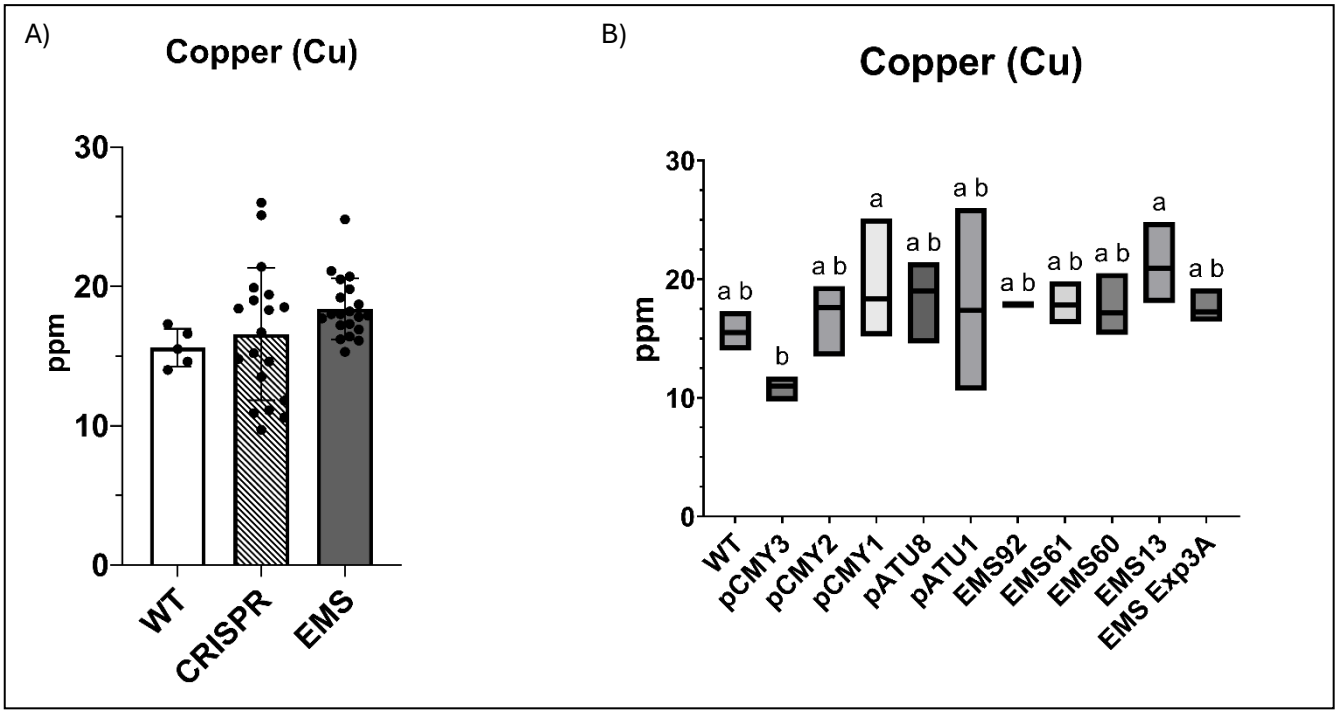


Figure 1-18) A) Combined treatment average of Copper content. B) Individual line breakdown of Copper content.

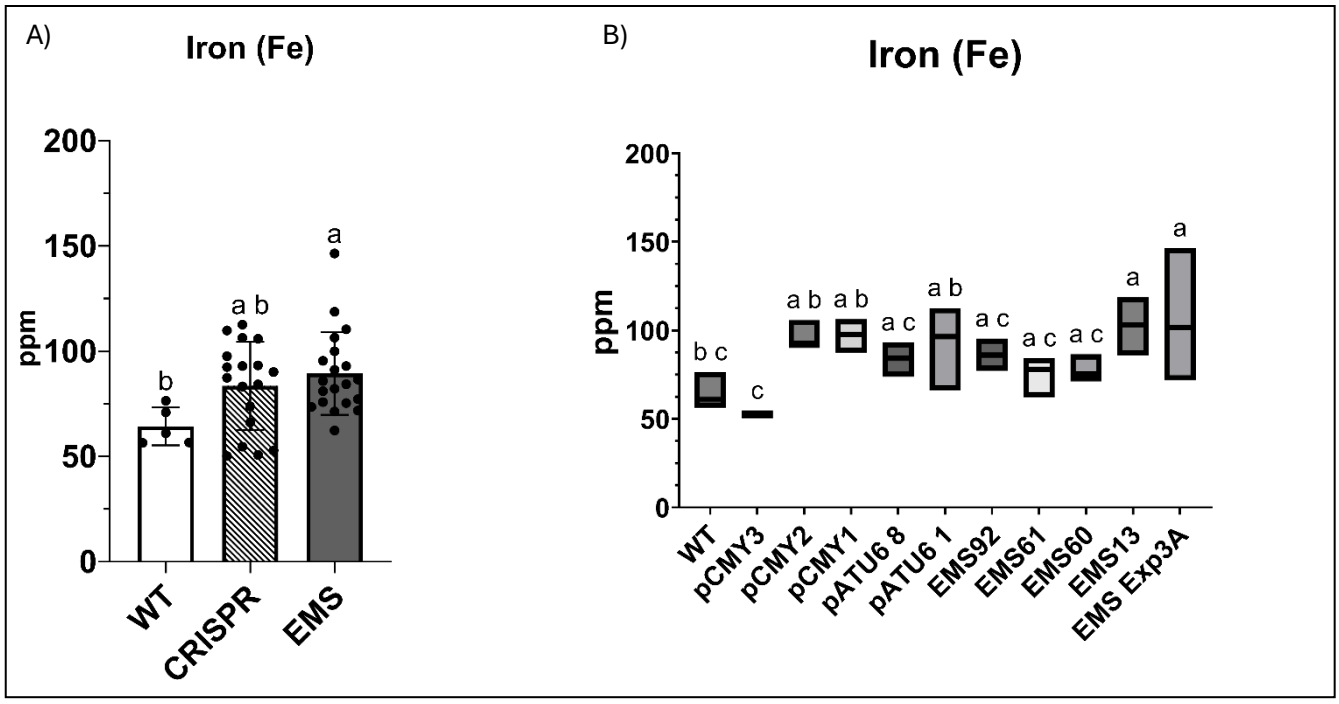


Figure 1-19) A) Combined treatment average of Iron content. B) Individual line breakdown of Iron content.

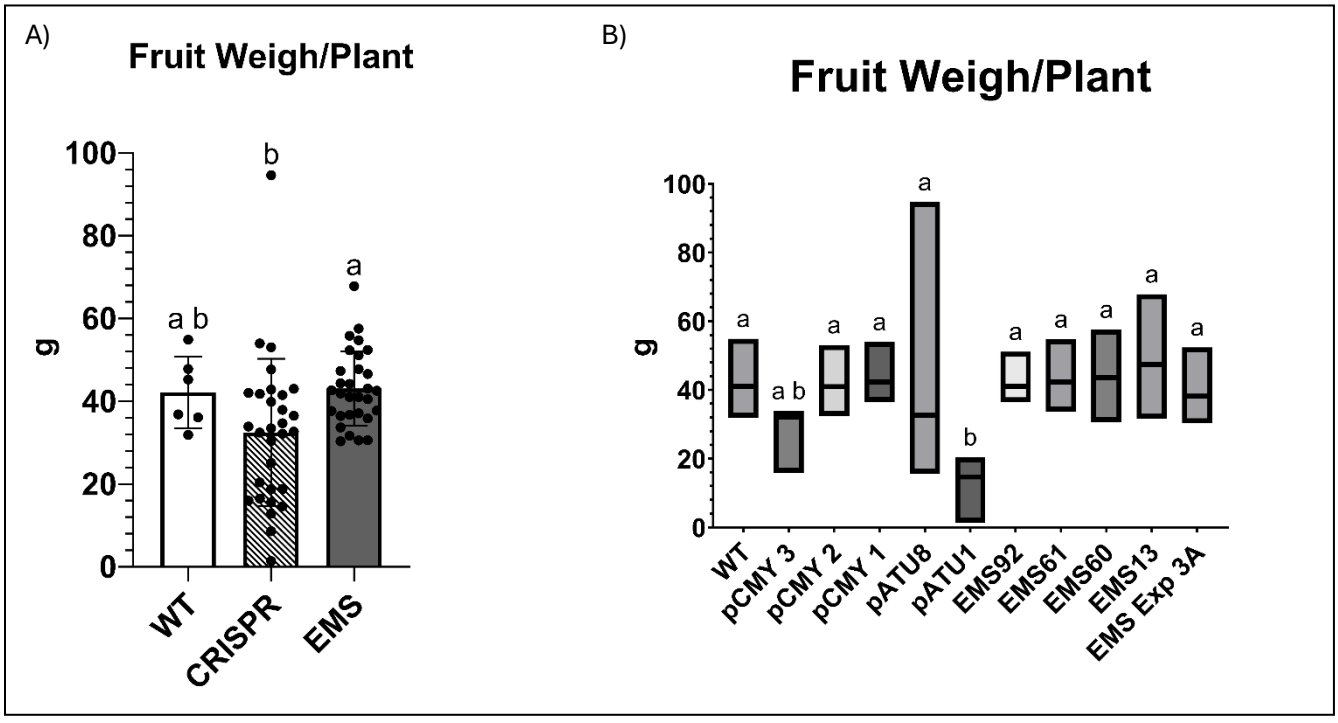


Figure 1-20) A) Combined treatment average of Fruit Weight/Plant. B) Individual line breakdown of Fruit Weight/Plant.

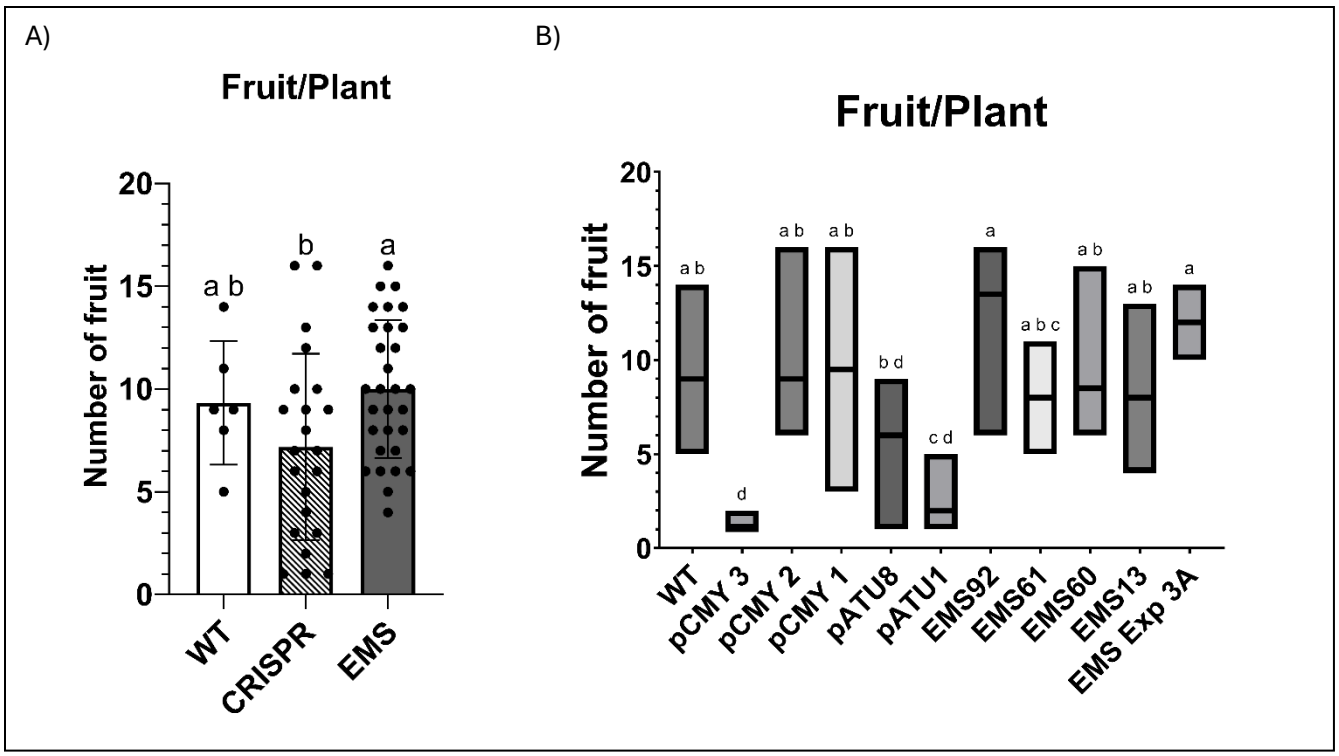


Figure 1-21) A) Combined treatment average Amount of Fruit/Plant. B) Individual line breakdown of Amount of Fruit/Plant.

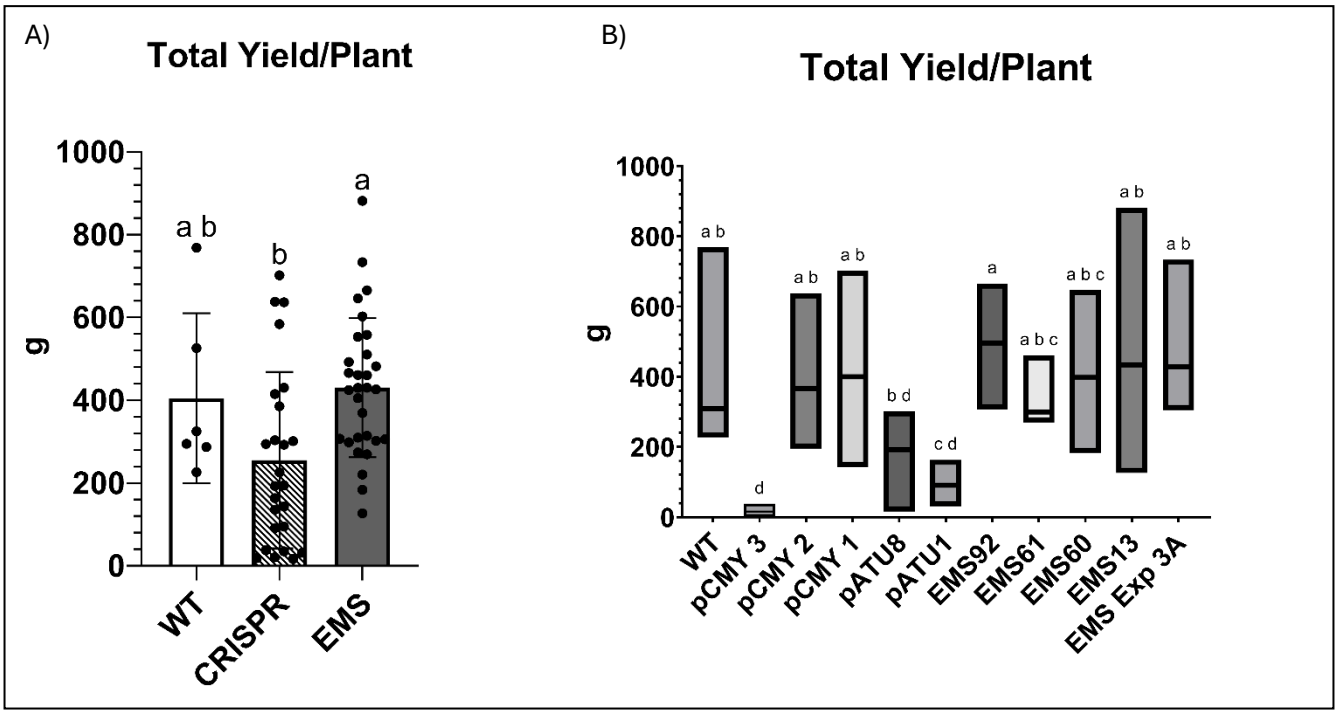


Figure 1-222) A) Combined treatment average of Yield/Plant. B) Individual line breakdown of Yield/Plant.

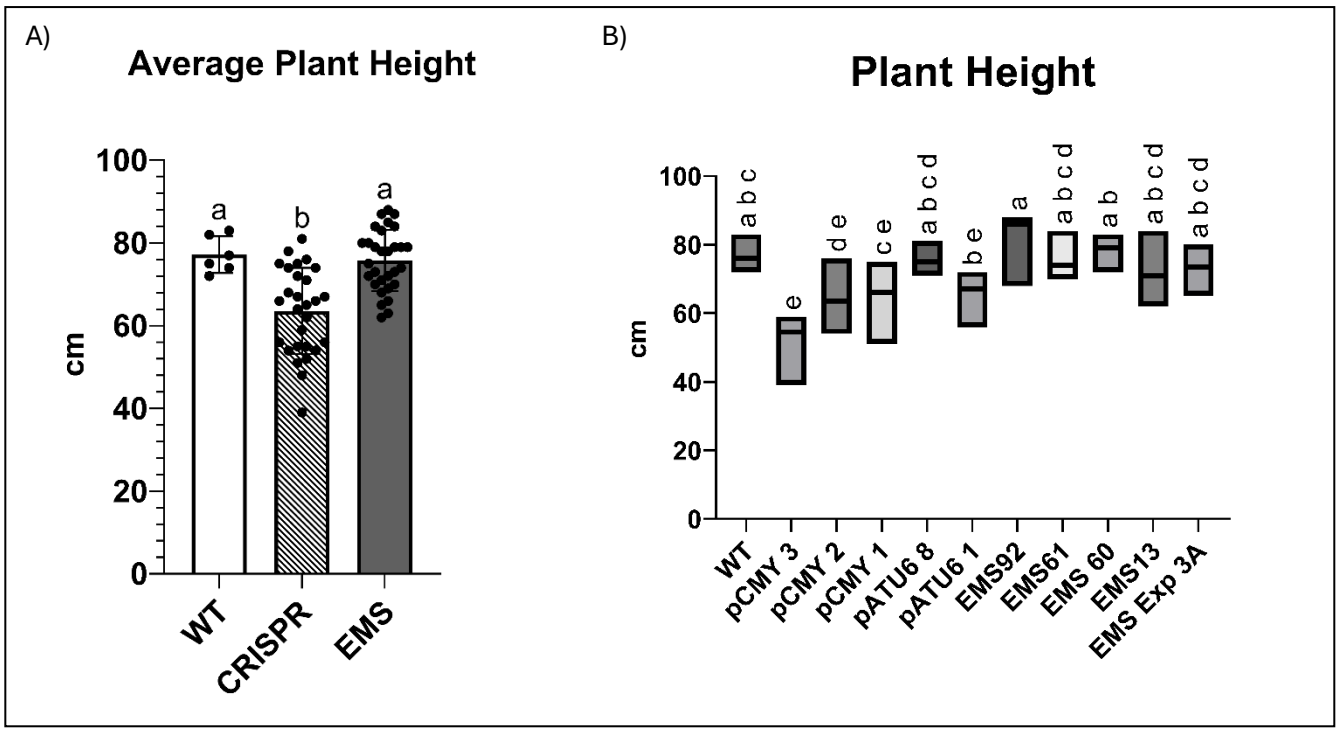


Figure1-23) A) Combined treatment average of Plant Height. B) Individual line breakdown of Plant Height.

Chapter II

Efficiency Comparison of Direct and Indirect Somatic Embryogenesis of *Agrobacterium*-mediated Watermelon Transformation

Abstract

Watermelon (*Citrullus lanatus*) is a vital crop with significant nutritional and economic value, widely cultivated and contributing substantially to agricultural economies globally. In the United States, watermelon production spans 44 states, generating over $500 million annually. Watermelon is highly valued for its high water content (92%) and rich supply of vitamins, minerals, and bioactive compounds, including vitamin C, vitamin A, potassium, and lycopene, a potent antioxidant linked to reduced risks of certain cancers and cardiovascular diseases. Genetic transformation offers promising avenues to enhance watermelon's nutritional value and stress tolerance by increasing beneficial compounds like lycopene and beta-carotene. Moreover, genetic modifications can improve disease resistance and stress tolerance by incorporating genes responsible for these traits.

This study compares the efficiency of direct and indirect somatic embryogenesis in *Agrobacterium*-mediated watermelon transformation. The results show that direct somatic embryogenesis yields a higher number of regenerated shoots within 2 to 4 weeks, but with higher incidences of chimeric expression, indicating partial transgene integration. In contrast, indirect somatic embryogenesis, though yielding fewer shoots, produced plants with stable and consistent transgene expression, suggesting a more effective method for stable transformation.

Key challenges in watermelon transformation include plant regeneration, and stable gene integration. The use of the RUBY genes as a visual marker proved effective, although slower growth rates in explants expressing the RUBY genes suggest the need for further optimization. This study's findings have broader implications for the genetic transformation of other recalcitrant crop species, providing insights that can improve transformation protocols for crops like soybean and cotton.

Future research should focus on optimizing transformation protocols by adjusting growth regulators, exploring alternative explant sources, and employing novel techniques such as CRISPR/Cas9-mediated gene editing. Long-term studies are necessary to assess the heritability and stability of transgene expression across generations. Addressing these challenges is critical to fully realizing the potential of genetic manipulation in watermelon, enhancing its nutritional value, disease resistance, and stress tolerance.

Introduction

Watermelon (*Citrullus lanatus*) is a globally significant crop, renowned for its nutritional benefits and economic value. It is widely cultivated and highly valued, contributing substantially to agricultural economies around the world. In the United States alone, watermelon production spans 44 states, generating over $500 million in commercial revenue annually (USDA Economic Research Service, 2012). This economic impact is further amplified by effective marketing and promotional efforts, which have significantly increased per capita consumption of watermelon (Kaiser, 2016).

Nutritionally, watermelon is prized for its high-water content (about 92%) and its rich supply of vitamins, minerals, and bioactive compounds. It is an excellent source of vitamin C and a good source of vitamin A, both of which are vital for various bodily functions including immune response, skin health, and vision (Wu & Morris, 2007). Watermelon also provides significant amounts of potassium, essential for heart health and muscle function, along with other minerals such as magnesium, calcium, sodium, zinc, iron, and copper (Adesanya et al., 2011). Additionally, red-fleshed varieties of watermelon are rich in lycopene, a potent antioxidant linked to reduced risks of certain cancers and cardiovascular diseases (Perkins-Veazie, Collins, Davis, & Roberts, 2006).

One of the most promising avenues for enhancing watermelon's nutritional value and stress tolerance is through genetic transformation. These carotenoids are crucial for human health due to their potent antioxidant properties, as demonstrated by Perkins-Veazie et al. (2006), who showed significant enhancement of carotenoid levels in genetically modified watermelon cultivars. Furthermore, transgenic watermelons can be engineered to contain higher concentrations of essential vitamins and minerals, thereby improving their nutritional value and addressing dietary

deficiencies (Li, Tang, Qin, Li, & Li, 2012). Another critical benefit of watermelon transformation is improved disease resistance. The introduction of genes such as Pti4 has been shown to confer resistance to the watermelon mosaic virus (WMV), reducing reliance on chemical treatments and minimizing crop losses (Li et al., 2012; Thomas, 1971). Moreover, incorporating resistance genes from other species can provide resistance against a range of fungal and bacterial pathogens, enhancing the overall resilience of watermelon crops (Martyn., 2014).

Genetic transformation also holds promise for improving stress tolerance in watermelon. Wild watermelons, which exhibit high tolerance to drought and excess light, serve as a valuable genetic resource. By incorporating genes responsible for the synthesis of protective compounds such as citrulline, which scavenges harmful radicals, these stress tolerance traits can be introduced into cultivated varieties (Akashi, Morikawa, & Yokota, 2005). Such advancements can ensure better growth and yield under adverse environmental conditions (Li et al., 2012).

Despite these potential benefits, watermelon transformation faces significant challenges, particularly in plant regeneration and stable gene integration. Watermelons are known to be recalcitrant to genetic transformation, making it difficult to regenerate mature plants from explants (Akashi et al., 2005). Post-transformation, explants often suffer from browning and necrosis, which affects the viability and success rate of producing transgenic plants (Li et al., 2012).

The selection and screening process for transformed cells is another hurdle. High concentrations of antibiotics such as kanamycin and hygromycin, necessary for effective selection, can be toxic to plant tissues, complicating the regeneration process (Akashi et al., 2005; Li et al., 2012). Furthermore, plants may develop as chimeras during the selection process, where only some cells are transformed, complicating the screening process since these plants might not uniformly exhibit the desired traits (Akashi et al., 2005; Li et al., 2012).

Ensuring stable integration and expression of introduced genes across generations is also a technical challenge. Techniques like Southern blot analysis are required to confirm stable gene integration, adding to the complexity and cost of the process (Akashi et al., 2005; Li et al., 2012). Moreover, transformed watermelon plants may exhibit increased ethylene production, which negatively affects plant growth and gene transfer efficiency. Managing this response is crucial for the success of transformation efforts (Li et al., 2012).

In summary, while genetic transformation of watermelon offers significant benefits in terms of improved nutrient content, disease resistance, and stress tolerance, considerable challenges remain. Addressing these challenges is critical to realizing the full potential of this biotechnological advancement.

Materials and methods

Explant Preparation

Mature watermelon seeds (Citrullus lanatus cv. Crimson Sweet) are thoroughly rinsed under running tap water for 5 minutes, then surface sterilized by immersion in 75% ethanol for 5 minutes, followed by a 20-minute treatment with 01% sodium hypochlorite, with the addition of TWEEN® 20 surfactant, under constant shaking. The seeds are rinsed three times with sterile distilled water to remove any sodium hypochlorite residue. The seeds are germinated on basic Murashige and Skoog (MS) medium containing standard salts, 30 g/L sucrose, and 8 g/L agar, with the pH adjusted to 5.7 before autoclaving. After seven days in darkness, cotyledon explants are excised from seven-day-old seedlings and cut into proximal and distal halves, with the proximal halves being used as for transformation.

Agrobacterium Preparation

Agrobacterium tumefaciens strain LBA 4404, containing the plasmid pKL2299 virulence plasmid with the 35S:RUBY plasmid genes, (Figure 2-1) is cultured on yeast extract peptone (YEP) solid medium supplemented with kanamycin 100 mg/L and streptomycin 100 mg/L for 48 hours. A single colony is transferred to 10 ml of YEP liquid medium with the same antibiotics and cultivated at 26°C with shaking until an OD600 of 0.6-0.8 is reached. The bacterial suspension is centrifuged and resuspended in a liquid MS medium containing 100 µM acetosyringone to an OD600 of 0.5 for inoculation of explants.

Transformation via Direct Somatic Embryogenesis

Resting

Explants of proximal cotyledons are pre-cultivated on MS medium supplemented with 2 mg/L benzyladenine (BA) for two days in darkness at 26°C.

Inoculation and Co-cultivation

The pre-cultivated explants are soaked in *Agrobacterium* tumefaciens suspension for 10 minutes in the absence of light at 26 °C. They are then removed from the suspension, and blotted dry, After inoculation, the explants are transferred to MS medium with BA 2 mg/L and incubated for 72 hours in darkness at 27°C

Shoot Development Media with Selection

After inoculation, the explants were transferred from dark to light with a 16h photoperiod and placed in shooting medium with selection to MS medium with BA 2 mg/l, kanamycin 100 mg/L, and Clavacillin 250 mg/L for selection. Subculturing is performed every two weeks until shoots start to form. The explants that survive kanamycin selection are considered putative transformants.

Shoot Elongation Media with Selection

During the selection phase, well-developed shoots are transferred to MS medium supplemented with BA 2 mg/L, kanamycin 100 mg/L, and Clavacillin 250 mg/L for elongation. These plants stay under the same 16h photoperiod at 26 °C conditions until

Root Development and Acclimation

Roots are induced on half-strength MS medium supplemented BA 2 mg/L, kanamycin 100 mg/L, and Clavacillin 250 mg/L.

Transformation via Indirect Somatic Embryogenesis

Resting

Explants of proximal cotyledons are pre-cultivated on MS medium supplemented with 2 mg/L BA, 0.2mg/L 1-Naphthaleneacetic acid (NAA), and 0.1 mg/L Myo-inositol for 48 hours in darkness at 26°C.

Inoculation and Co-cultivation

The pre-cultivated explants are soaked in *Agrobacterium* tumefaciens suspension for 10 minutes in the absence of light at 26 °C. They are then removed from the suspension and blotted dry. After inoculation, the explants are transferred to MS medium supplemented with 2 mg/L BA, 0.2mg/L NAA, and 0.1 mg/L Myo-inositol for 72 hours in darkness at 26°C.

Shoot Development Media with Selection

After inoculation, the explants were transferred to MS medium supplemented with 2 mg/L BA, 0.2 mg/L NAA, and 0.1 mg/L Myo-inositol, kanamycin 100 mg/L, and Clavacillin 250 mg/L for selection. The explants were placed back in the dark, and subculturing was performed every two weeks until shoots started to form within 4 to 6 weeks. The explants that survive kanamycin selection are considered putative transformants.

Shoot Elongation Media with Selection

Well-developed shoots are transferred to MS medium supplemented with 2 mg/L BA, 0.5 mg/L NAA, and 0.1 mg/L Myo-inositol, kanamycin 100 mg/L, and Clavacillin 250 mg/L were exposed to light for a 16h photoperiod at 26 °C conditions for the following 2 to 4 weeks

Root Development and Acclimation

Roots are induced on half-strength MS medium supplemented with NAA 0.5 mg/L, kanamycin 100 mg/L, and Clavacillin 250 mg/L. The surviving explants were exposed to light for a 16h photoperiod at 26 °C conditions for the following 2 to 4 weeks or until they rooted.

Acclimation of Rooted Plants to a Low-Humidity Environment

Plantlets with developed roots are transplanted into plastic pots with BX potting mix, covered with clear Ziplock bags for 2 days to maintain high humidity, and incubated at 26°C under a 16-hour photoperiod. Humidity is gradually reduced over the next two weeks to allow adequate

time for acclimation to a low-humidity environment found in the growth chamber. The plants are watered with nutrient solution and water alternately.

DNA Isolation and PCR Confirmation of Transgenic Plants

Leaf samples were collected from two-week-old potted plants, and leaf samples were also collected from shoots expressing the Ruby genes in tissue culture due to time limitations for DNA isolation. DNA was extracted using the DNeasy® Plant Mini Kit (Qiagen) following the manufacturer's instructions. The quality and quantity of the DNA samples were assessed using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific), and the integrity of the DNA was confirmed by electrophoresis on a 1% TopVision Agarose gel (Thermo Scientific). DNA samples were visualized under UV light to ensure high-molecular-weight DNA and to check for any degradation.

For PCR amplification, specific primers were used: RubyForward (5'-CCCTCAGACTCGGCTCTGTGA-3') and RubyReverse (5'-ggcGATCAGCTTGTcGAAGGA-3'). The PCR reaction was performed using the DreamTaq DNA Polymerase Kit (Thermo Scientific). The reaction mix had a total volume of 25 µL, containing 1 µL of template DNA (~50 ng), 2.5 µL of 10X DreamTaq Buffer, 1 µL of each primer (10 µM), 0.5 µL of dNTP mix (10 mM each), 0.25 µL of DreamTaq DNA Polymerase (5 U/µL), and nuclease-free water to the final volume.

PCR amplification was carried out in a thermal cycler (Bio-Rad) under the following conditions: an initial denaturation at 94°C for 3 minutes; 35 cycles of denaturation at 94°C for 30

seconds, annealing at 58°C for 30 seconds, and extension at 72°C for 1 minute; followed by a final extension at 72°C for 10 minutes.

The PCR products were analyzed by electrophoresis on a 1% agarose gel prepared by dissolving 1 g of agarose in 100 mL of 1X TAE buffer. The solution was heated until the agarose was completely dissolved, cooled to approximately 60°C, poured into a gel tray with a comb, and allowed to solidify. PCR products were mixed with 6X loading dye, and 5 µL of each sample was loaded into the wells along with a 1kb bp DNA ladder (Thermo Scientific) as a molecular weight marker. Electrophoresis was performed at 100 volts for 30-45 minutes, and the gel was stained with GelRed (Biotium). Bands were visualized under UV light using a gel documentation system (Bio-Rad).

A specific 539 bp region was successfully amplified in transgenic plants, as indicated by the presence of a distinct band corresponding to this size, confirming the integration of the transgene.

Results

Transformation via Direct Somatic Embryogenesis

Out of 50 explants subjected to direct somatic embryogenesis, 12 shoots were successfully regenerated within 2 to 4 weeks (Figure 2-3). However, chimeric expression of the RUBY genes was observed, indicating that not all cells within the explants were uniformly transformed. This chimeric expression suggests partial integration of the transgene, which may affect the consistency of trait expression in subsequent generations.

Transformation via Indirect Somatic Embryogenesis

In the case of indirect somatic embryogenesis, 4 shoots were regenerated from 25 explants. The selection process took longer, with shoot development occurring over 6 to 8 weeks (Figure 26 C) D)). Unlike the direct method, the entire plant expressed the RUBY genes consistently, suggesting more stable integration of the transgene. The higher levels of RUBY expression observed in these plants indicated a more efficient transformation process compared to the direct method.

Evaluation of Ruby Genes as a Visual Marker

The RUBY genes proved to be a reliable visual marker for transformation. Explants expressing the RUBY genes and kanamycin resistance gene exhibited slower growth compared to those with chimeric expression. Under light conditions, the expression of the RUBY genes (Figure 2-1) was clearly visible, although it was harder to detect in some areas of the explants. This visual distinction facilitated the identification of successfully transformed plants during the selection process.

PCR Confirmation of Transgenic Plants

PCR analysis using the RubyForward (5’-CCCTCAGACTCGGCTCTGTGA -3’) and RubyReverse (5’-GGCGATCAGCTTGTcGAAGGA-3’) primers confirmed the presence of the RUBY genes in the transgenic plants. A specific 539 bp fragment was successfully amplified from the DNA of putative transformants regenerated via indirect somatic embryogenesis, indicating the integration of the RUBY transgene (Figure 2-2 B). The PCR results were consistent across multiple samples, confirming the stable transformation of the selected explants.

However, when analyzing plants from direct somatic embryogenesis transformation, two plants did not amplify the specific 539 bp fragment. Instead, a larger ~750 bp genomic segment

was observed (Figure 2-2 A), which was also present in wild-type (WT) plants. This indicates potential difficulties in the transformation process, possibly due to challenges inherent to watermelon transformation, as highlighted in recent studies utilizing morphogenic genes for better results (Pan et al., 2021)

Summary of Findings

The transformation efficiency differed between the direct and indirect somatic embryogenesis methods. Direct somatic embryogenesis resulted in a higher number of regenerated shoots, and the regenerated plants produced roots faster compared to the indirect somatic embryogenesis-derived plants. In contrast, indirect somatic embryogenesis, though yielding fewer shoots and these plants did not produce any roots, the produced shoots had a stable and consistent transgene expression. The RUBY genes served as an effective visual marker, aiding in the selection of transformed plants despite some challenges in growth rates and visibility under certain conditions.

Discussion

Efficiency of Direct vs. Indirect Somatic Embryogenesis

Transformation via Direct and Indirect Somatic Embryogenesis differ in many ways (Table 2-1). The use of the correct method is crucial to achieve a high rate of transformation. The efficiency of genetic transformation in watermelon via direct and indirect somatic embryogenesis presents distinct outcomes and challenges. The results demonstrated that direct somatic embryogenesis yielded a higher number of regenerated shoots within a shorter period (2 to 4 weeks) compared to indirect somatic embryogenesis (6 to 8 weeks). However, plants regenerated via direct somatic embryogenesis only expressed the RUBY genes temporarily, and the regenerated plants were not transgenic. (Figure 2-3), indicating a non-stable integration of the transgene. This

non-stable integration can affect the consistency of trait expression in subsequent generations if the transformation is successful, posing a significant challenge for stable genetic transformation (Li et al., 2012; Akashi et al., 2005).

Conversely, indirect somatic embryogenesis, despite producing fewer shoots and developing at a slower pace, resulted in more stable and consistent expression of the RUBY genes across the entire plant. This suggests that indirect somatic embryogenesis may be more effective for achieving stable transformation (Figure 2-4), as evidenced by the uniform expression of the transgene (Deo et al., 2010). The longer selection process in indirect somatic embryogenesis could contribute to the enhanced stability of transgene integration, as it allows for more rigorous selection of successfully transformed cells.

Importance of Establishing a Stable Transformation Protocol

Establishing a stable transformation protocol is critical for the successful genetic manipulation of watermelon. The challenges associated with watermelon transformation, such as browning and necrosis of explants post-transformation, (Figure 2-5) necessitate optimization of the transformation process (Li et al., 2012; Akashi et al., 2005). The use of the RUBY genes as a visual marker in this study provided a reliable method for identifying transformed plants, although the slower growth rates of explants expressing the RUBY genes suggest that further optimization is needed.

Recent advancements in plant transformation techniques, such as the incorporation of antioxidants like ascorbic acid to mitigate visible browning on the plants browning and the use of less toxic selection agents, could improve the efficiency and stability of watermelon transformation protocols (Dan Y., 2008). Additionally, exploring the use of alternative visual markers and selection strategies, such as GFP or GUS reporter genes, could enhance the identification and

selection of transformed plants (Choi et al., 1994). Recent advancements have shown that the use of developmental regulator genes such as *AtGRF5* can significantly improve transformation efficiency in watermelon, increasing the rate to about 25%, which is a 40-fold increase compared to traditional methods (Pan et al., 2021). This suggests that incorporating such strategies could potentially overcome the observed difficulties and improve the reliability of transformations.

Evaluation of the RUBY Genes as a Visual Marker

The RUBY genes proved to be an effective visual marker for the transformation process, facilitating the identification of successfully transformed plants. The distinct coloration provided by the RUBY genes allowed for easy differentiation of transformed explants under light conditions, although its expression was harder to detect in some areas of the explants. This variability in visibility suggests that the use of RUBY genes as a marker could be further optimized to enhance its reliability (He et al., 2020).

Moreover, the slower growth observed in explants expressing the RUBY genes raises questions about the metabolic burden imposed by the transgene. The addition of tyrosine to the media has been shown to increase growth rates in transformed tobacco plants, suggesting a potential strategy to mitigate the growth retardation observed in this study. (Jogam et al., 2024) Investigating the effects of supplementing transformation media with supplemental tyrosine could improve the growth rates of transformed watermelon explants, thereby enhancing the overall efficiency of the transformation process.

Broader Implications and Future Directions

The findings from this study have broader implications for the genetic transformation of other recalcitrant crop species. The challenges observed in watermelon transformation are not unique and are shared by other crops, such as soybean and cotton, which also exhibit difficulties

in achieving high transformation efficiency and stable gene integration (Kumar et al., 2001). Lessons learned from optimizing watermelon transformation protocols can be applied to these crops, potentially improving the overall success rate of genetic transformation in recalcitrant species.

Future research should focus on further refining the transformation protocols to enhance both efficiency and stability. This includes optimizing the concentrations of growth regulators and selection agents, exploring alternative explant sources, and investigating the use of novel transformation techniques. (Tay et al., 2022). Additionally, long-term studies are needed to assess the heritability and stability of transgene expression across multiple generations, ensuring that the desired traits are consistently expressed in progeny plants (Ma et al., 2020).

Conclusion

The comparison of direct and indirect somatic embryogenesis for *Agrobacterium*-mediated watermelon transformation reveals trade-offs between the methods. Direct somatic embryogenesis offered a higher regeneration rate but did not result in a successful transformation. In contrast, indirect somatic embryogenesis provided more stable transgene expression despite a lower regeneration rate. The RUBY genes, used as a visual marker, were effective but impacted growth rates, indicating a need for further optimization. Future research should aim to enhance the stability and efficiency of transformation protocols to maximize the potential of genetic manipulation in watermelon for improved nutrition, disease resistance, and stress tolerance.

References

Adesanya, A. O., Olaseinde, O. O., Oguntayo, O. D., Otulana, J. O., & Adefule, A. K. (2011). Effects of Methanolic Extract of Citrullus lanatus Seed on Experimentally Induced Prostatic Hyperplasia. *European Journal of Medicinal Plants, 1*(4), 171-179.

Akashi, K., Morikawa, K., & Yokota, A. (2005). *Agrobacterium*-mediated transformation system for the drought and excess light stress-tolerant wild watermelon (Citrullus lanatus). *Plant Biotechnology, 22*(1), 13-18.

Akashi, R., Morikawa, K., & Yokota, A. (2005). Citrulline, a novel compatible solute in drought-tolerant wild watermelon leaves, is an efficient hydroxyl radical scavenger. FEBS Letters, 508(3), 438-442.

Choi, P. S., Soh, W. Y., Kim, Y. S., Yoo, O. J., & Liu, J. R. (1994). Genetic transformation and plant regeneration of watermelon using *Agrobacterium* tumefaciens. Plant cell reports, 13, 344-348.

Dan, Y. (2008). Biological functions of antioxidants in plant transformation. In Vitro Cellular & Developmental Biology-Plant, 44(3), 149-161.

Deo, P. C., Tyagi, A. P., Taylor, M., Harding, R., & Becker, D. (2010). Factors affecting somatic embryogenesis and transformation in modern plant breeding. The South Pacific Journal of Natural and Applied Sciences, 28(1), 27-40.

Jogam, P., Anumula, V., Sandhya, D., Manokari, M., Venkatapuram, A. K., Achary, V. M. M., ... & Allini, V. R. (2024). Monitoring genetic transformation with RUBY visible reporter in Nicotiana tabaccum L. Plant Cell, Tissue and Organ Culture (PCTOC), 157(1), 23.

He, Y., Zhang, T., Sun, H., Zhan, H., & Zhao, Y. (2020). A reporter for noninvasively monitoring gene expression and plant transformation. Horticulture research, 7.

Kaiser, H. M. (2016). An Economic Analysis of Market Impacts of the National Watermelon Promotion Board. National Watermelon Promotion Board.

Kumar, S., & Fladung, M. (2001). Controlling transgene integration in plants. Trends in Plant Science, 6(4), 155-159.

Li, J., Tang, Y., Qin, Y., Li, X., & Li, H. (2012). *Agrobacterium*-mediated transformation of watermelon (Citrullus lanatus). *African Journal of Biotechnology, 11*(24), 6450-6456. doi: 10.5897/AJB11.1626

Ma, X., Zhu, Q., Chen, Y., & Liu, Y. G. (2020). CRISPR/Cas9 platforms for genome editing in plants: developments and applications. Molecular Plant, 13(9), 1317-1347.

Martyn, R. D. (2012). Fusarium wilt of watermelon: a historical review.

Pan, W., Cheng, Z., Han, Z., Yang, H., Zhang, W., & Zhang, H. (2021). Efficient transformation and genome editing of watermelon assisted by genes that encode developmental regulators. bioRxiv, 2021-11.

Perkins-Veazie, P., Collins, J. K., Davis, A. R., & Roberts, W. (2006). Carotenoid content of 50 watermelon cultivars. *Journal of Agricultural and Food Chemistry, 54*(7), 2593-2597. doi: 10.1021/jf052066p

Tay Fernandez, C. G., Nestor, B. J., Danilevicz, M. F., Marsh, J. I., Petereit, J., Bayer, P. E., ... & Edwards, D. (2022). Expanding gene-editing potential in crop improvement with pangenomes. International Journal of Molecular Sciences, 23(4), 2276.

Thomas, W. (1971). The incidence and economic importance of watermelon mosaic virus. *New Zealand Journal of Agricultural Research, 14*(1), 242-247. doi: 10.1080/00288233.1971.10421320

USDA Economic Research Service. (2012). Per capita consumption of watermelon in the United States. USDA ERS.

USDA Economic Research Service. (2012). Watermelon statistics. Retrieved from USDA ERS Website

Wu, G., & Morris, S. M. (2007). Arginine metabolism: nitric oxide and beyond. *Biochemical Journal, 336*, 1-17.

Figures and Tables

Table 2-1 Step-by-step comparison of differences between direct and indirect somatic embryogenesis

Transformation Step	Media Composition	Conditions	Purpose
Explant Preparation	Basic MS medium, 30 g/L sucrose, 8 g/L agar, pH 5.7	7 days in darkness	Germination of seeds
Pre-cultivation (Direct SE)	MS medium, 2 mg/L BA	2 days in darkness, 26°C	Preparation of explants for transformation
Pre-cultivation (Indirect SE)	MS medium, 2 mg/L BA, 0.2 mg/L NAA, 0.1 mg/L Myo-inositol	48 hours in darkness, 26°C	Preparation of explants for transformation
Inoculation and Co-cultivation (Direct SE)	MS medium, 2 mg/L BA, 100 µM acetosyringone	72 hours in darkness, 27°C (Direct SE)	Co-cultivation with *Agrobacterium*
Inoculation and Co-cultivation (Indirect SE)	MS medium, 2 mg/L BA, 0.2 mg/L NAA, 0.1 mg/L Myo-inositol, 100 µM acetosyringone	72 hours in darkness, 26°C (Indirect SE)	Co-cultivation with *Agrobacterium*
Shoot Development (Direct SE)	MS medium, 2 mg/L BA, 100 mg/L kanamycin, 250 mg/L Clavacillin	16-hour photoperiod, 26°C	Selection and shoot induction
Shoot Development (Indirect SE)	MS medium, 2 mg/L BA, 0.2 mg/L NAA, 0.1 mg/L Myo-inositol, 100 mg/L kanamycin, 250 mg/L Clavacillin	16-hour photoperiod, 26°C 4-6 weeks	Selection and shoot induction
Shoot Elongation (Direct SE)	MS medium, 2 mg/L BA, 100 mg/L kanamycin, 250 mg/L Clavacillin	16-hour photoperiod, 26°C 2-4 weeks	Shoot elongation
Shoot Elongation (Indirect SE)	MS medium, 2 mg/L BA, 0.5 mg/L NAA, 0.1 mg/L Myo-inositol, 100 mg/L kanamycin, 250 mg/L Clavacillin	16-hour photoperiod, 26°C, 2-4- weeks	Shoot elongation
Root Development (Direct SE)	Half-strength MS medium, 2 mg/L BA, 100 mg/L kanamycin, 250 mg/L Clavacillin	16-hour photoperiod, 26°C 2-4 weeks	Root induction
Root Development (Indirect SE)	Half-strength MS medium, 0.5 mg/L NAA, 100 mg/L kanamycin, 250 mg/L Clavacillin	16-hour photoperiod, 26°C, 2-4 weeks	Root induction
Acclimation	BX potting mix	Gradual reduction of humidity, 26°C	Acclimation of rooted plants to low-humidity environment

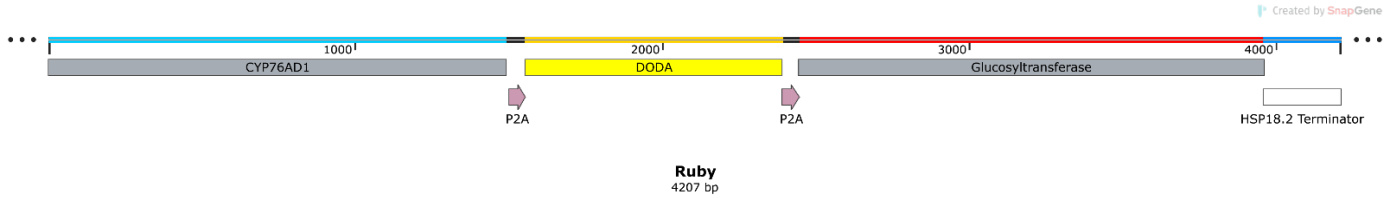


Figure 2-1) Map of the Ruby construct pathway

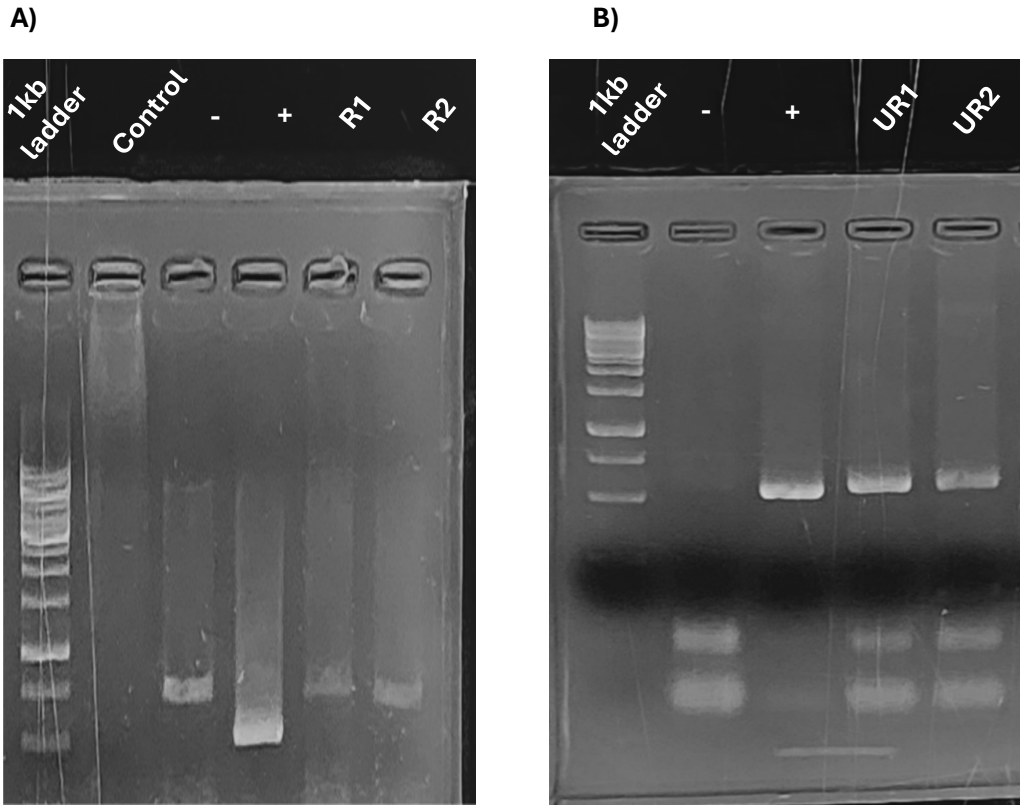


Figure 2-2) A) PCR confirmation of rooted non-expressing lines. -(WT), rooted plant1 (R1) and rooted plant2 (R2) show non-target PCR amplification. B) PCR confirmation of Ruby expressing lines + (plasmid), unrooted plant 1 (UR1), and unrooted plant 2 (UR2) show Ruby genes PCR amplification (transgenic)

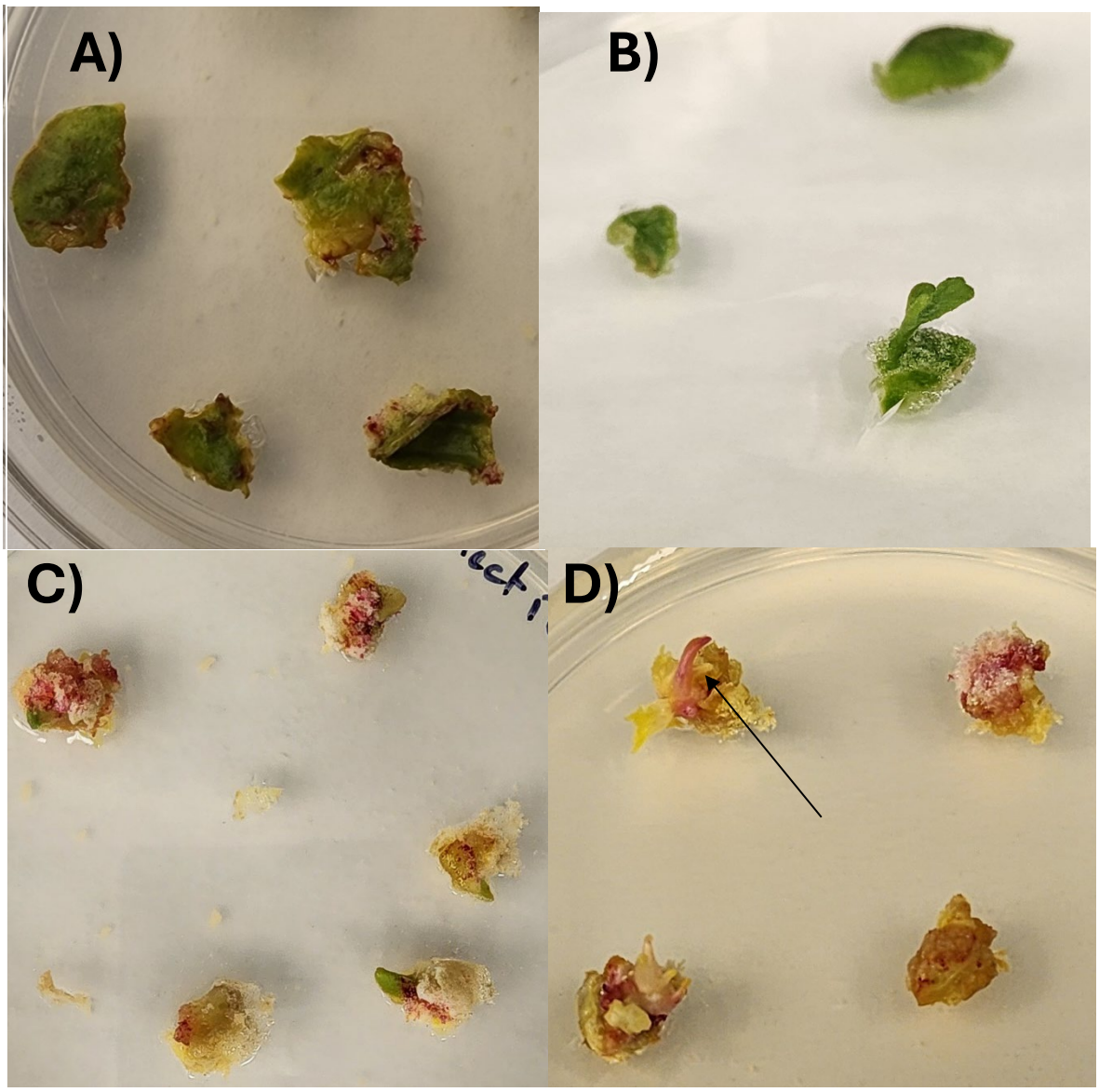


Figure 2-3) A) Direct somatic embryogenesis 2 weeks after inoculations. B) Direct somatic embryogenesis shoot development. C) Indirect somatic embryogenesis 2 weeks after inoculations with callus formation. Shoot formation in Indirect somatic embryogenesis.

A)

B)

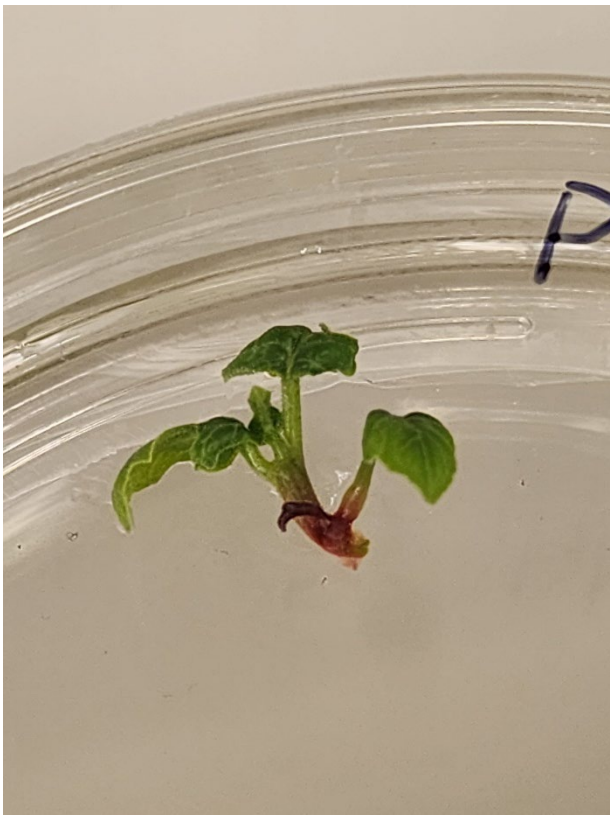
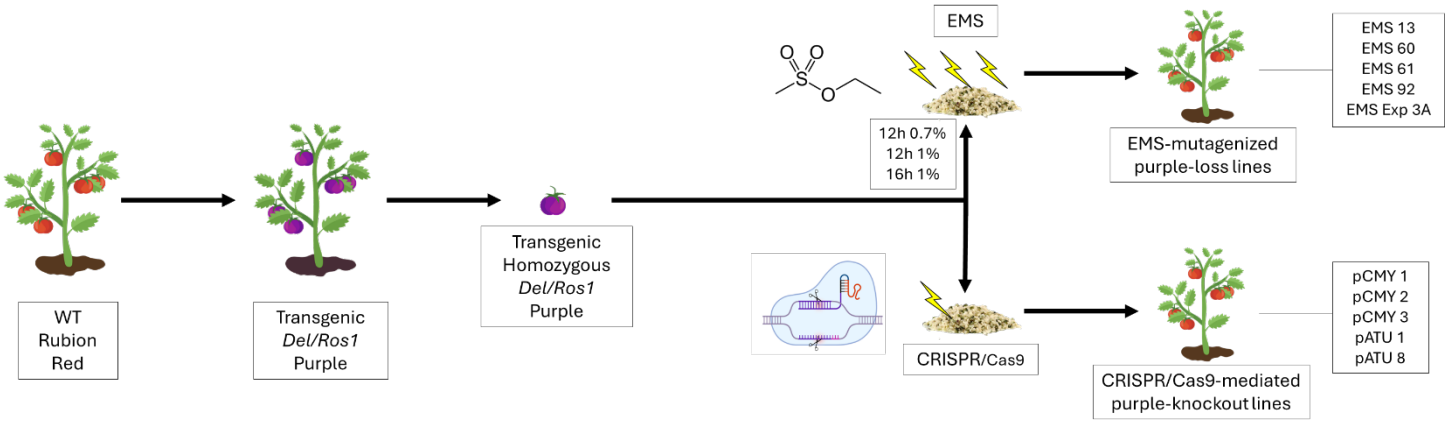


Figure 2-4) A) Direct somatic embryogenesis-derived plant. B) Indirect somatic embryogenesis derived transgenic plant.

Figure 2-5 Direct somatic embryogenesis derived plant in soilless media with browning.

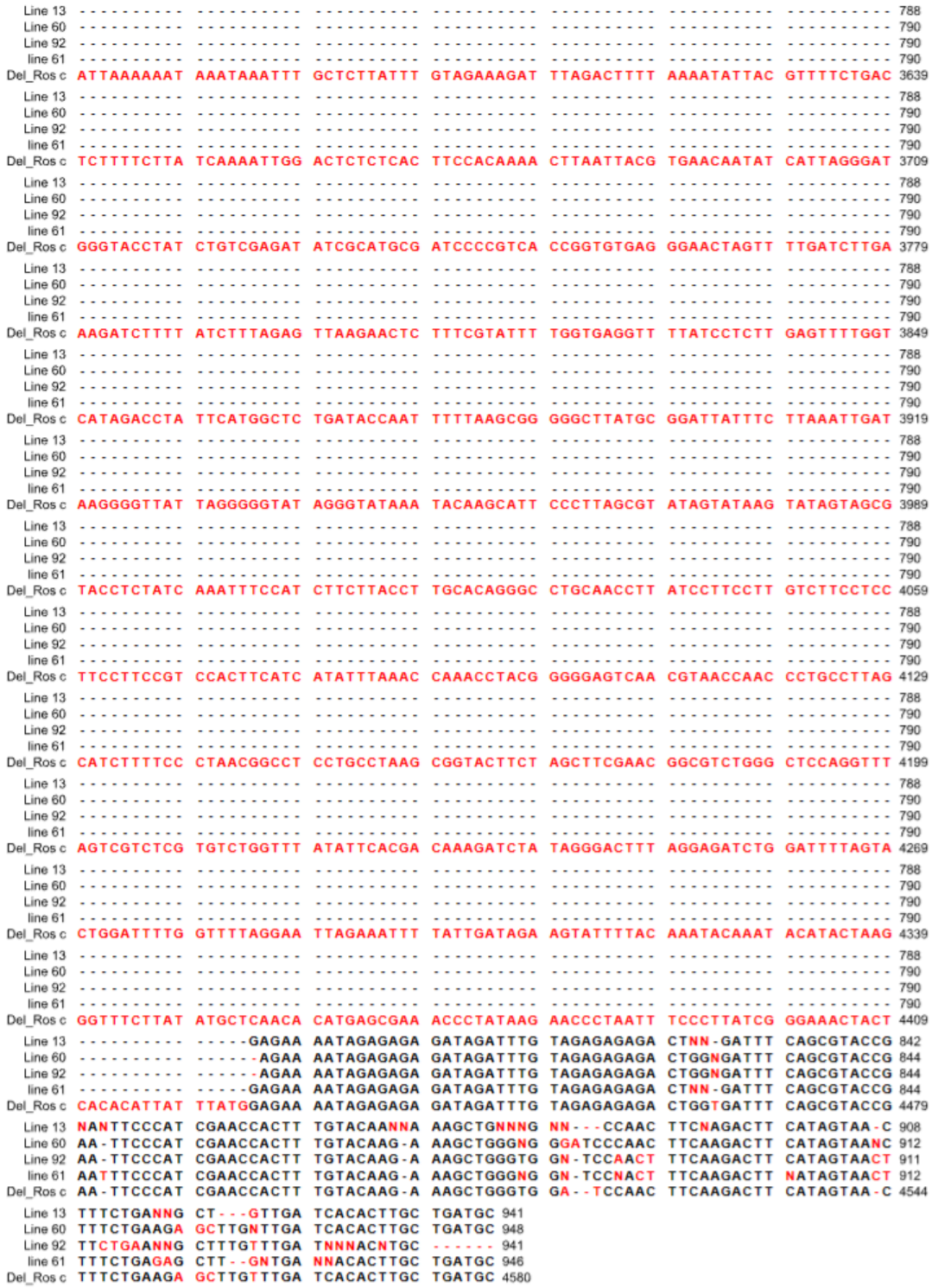
Supporting Information



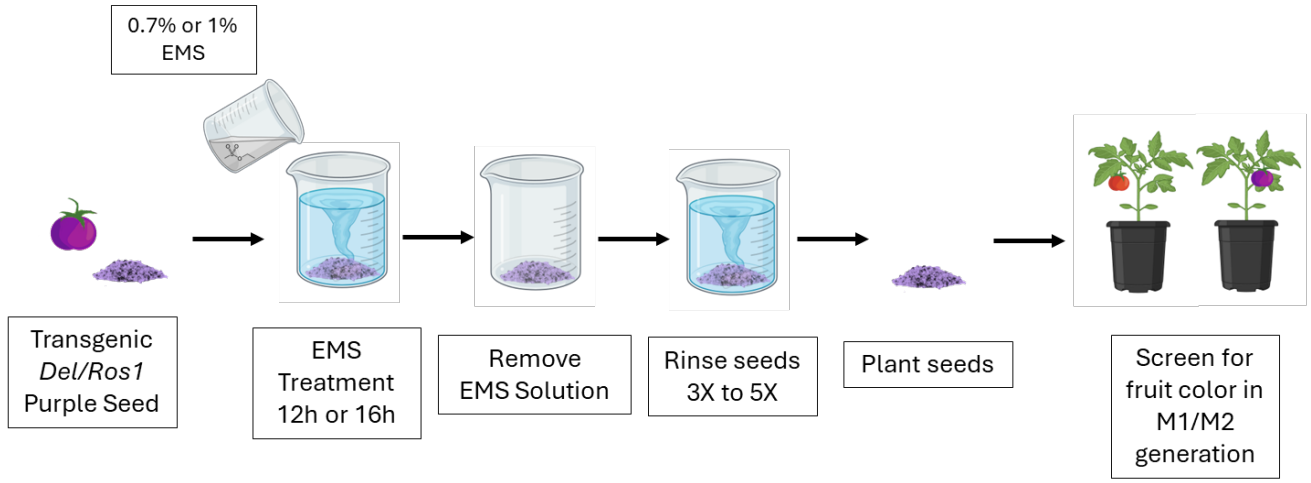
Supplementary Figure 1 Experimental design


Line 13 CAGGAAACAG CTATGACCAT GATTACGCCA AGCTCGAAAT TAACCCTCAC TAAAGGGAAC AAAAGCTGGT 70
Line 60 CAGGAAACAG CTATGACCAT GATTACGCCA AGCTCGAAAT TAACCCTCAC TAAAGGGAAC AAAAGCTGGT 70
Line 92 CAGGAAACAG CTATGACCAT GATTACGCCA AGCTCGAAAT TAACCCTCAC TAAAGGGAAC AAAAGCTGGT 70
line 61 CAGGAAACAG CTATGACCAT GATTACGCCA AGCTCGAAAT TAACCCTCAC TAAAGGGAAC AAAAGCTGGT 70
Del_Ros c CAGGAAACAG CTATGACCAT GATTACGCCA AGCTCGAAAT TAACCCTCAC TAAAGGGAAC AAAAGCTGGT 70
Line 13 ACGTACCGGG CCCCCC-TCG AGATATCGCA TGCGATCCCC GTCACCGGTG TGAGGGAACT AGTTTTGATC 139
Line 60 ACGTACCGGG CCCCCCCTCG AGATATCGCA TGCGATCCCC GTCACCGGTG TGAGGGAACT AGTTTTGATC 140
Line 92 ACGTACCGGG CCCCCCCTCG AGATATCGCA TGCGATCCCC GTCACCGGTG TGAGGGAACT AGTTTTGATC 140
line 61 ACGTACCGGG CCCCCCCTCG AGATATCGCA TGCGATCCCC GTCACCGGTG TGAGGGAACT AGTTTTGATC 140
Del_Ros c ACGTACCGGG CCCCCCCTCG AGATATCGCA TGCGATCCCC GTCACCGGTG TGAGGGAACT AGTTTTGATC 140
Line 13 TTGAAAGATC TTTTATCTTT AGAGTTAAGA ACTCTTTCGT ATTTTGGTGA GGTTTTATCC TCTTGAGTTT 209
Line 60 TTGAAAGATC TTTTATCTTT AGAGTTAAGA ACTCTTTCGT ATTTTGGTGA GGTTTTATCC TCTTGAGTTT 210
Line 92 TTGAAAGATC TTTTATCTTT AGAGTTAAGA ACTCTTTCGT ATTTTGGTGA GGTTTTATCC TCTTGAGTTT 210
line 61 TTGAAAGATC TTTTATCTTT AGAGTTAAGA ACTCTTTCGT ATTTTGGTGA GGTTTTATCC TCTTGAGTTT 210
Del_Ros c TTGAAAGATC TTTTATCTTT AGAGTTAAGA ACTCTTTCGT ATTTTGGTGA GGTTTTATCC TCTTGAGTTT 210
Line 13 TGGTCATAGA CCTATTCATG GCTCTGATAC CAATTTTTAA GCGGGG-CTT ATGCGGATTA TTTCTTAAAT 278
Line 60 TGGTCATAGA CCTATTCATG GCTCTGATAC CAATTTTTAA GCGGGGGCTT ATGCGGATTA TTTCTTAAAT 280
Line 92 TGGTCATAGA CCTATTCATG GCTCTGATAC CAATTTTTAA GCGGGGGCTT ATGCGGATTA TTTCTTAAAT 280
line 61 TGGTCATAGA CCTATTCATG GCTCTGATAC CAATTTTTAA GCGGGG-CTT ATGCGGATTA TTTCTTAAAT 279
Del_Ros c TGGTCATAGA CCTATTCATG GCTCTGATAC CAATTTTTAA GCGGGGGCTT ATGCGGATTA TTTCTTAAAT 280
Line 13 TGATAAGGGG TTATTAGGGG GTATAGGGTA TAAATACAAG CATTCCCTTA GCGTATAGTA TAAGTATAGT 348
Line 60 TGATAAGGGG TTATTAGGGG GTATAGGGTA TAAATACAAG CATTCCCTTA GCGTATAGTA TAAGTATAGT 350
Line 92 TGATAAGGGG TTATTAGGGG GTATAGGGTA TAAATACAAG CATTCCCTTA GCGTATAGTA TAAGTATAGT 350
line 61 TGATAAGGGG TTATTAGGGG GTATAGGGTA TAAATACAAG CATTCCCTTA GCGTATAGTA TAAGTATAGT 349
Del_Ros c TGATAAGGGG TTATTAGGGG GTATAGGGTA TAAATACAAG CATTCCCTTA GCGTATAGTA TAAGTATAGT 350
Line 13 AGCGTACCTC TATCAAATTT CCATCTTCTT ACCTTGCACA GGGCCTGCAA CCTTATCCTT CCTTGTCTTC 418
Line 60 AGCGTACCTC TATCAAATTT CCATCTTCTT ACCTTGCACA GGGCCTGCAA CCTTATCCTT CCTTGTCTTC 420
Line 92 AGCGTACCTC TATCAAATTT CCATCTTCTT ACCTTGCACA GGGCCTGCAA CCTTATCCTT CCTTGTCTTC 420
line 61 AGCGTACCTC TATCAAATTT CCATCTTCTT ACCTTGCACA GGGCCTGCAA CCTTATCCTT CCTTGTCTTC 419
Del_Ros c AGCGTACCTC TATCAAATTT CCATCTTCTT ACCTTGCACA GGGCCTGCAA CCTTATCCTT CCTTGTCTTC 420
Line 13 CTCCTTCCTT CCGTCCACTT CATCATATTT AAACCAAACC TACGGGGGAG TCAACGTAAC CAACCCTGCC 488
Line 60 CTCCTTCCTT CCGTCCACTT CATCATATTT AAACCAAACC TACGGGGGAG TCAACGTAAC CAACCCTGCC 490
Line 92 CTCCTTCCTT CCGTCCACTT CATCATATTT AAACCAAACC TACGGGGGAG TCAACGTAAC CAACCCTGCC 490
line 61 CTCCTTCCTT CCGTCCACTT CATCATATTT AAACCAAACC TACGGGGGAG TCAACGTAAC CAACCCTGCC 489
Del_Ros c CTCCTTCCTT CCGTCCACTT CATCATATTT AAACCAAACC TACGGGGGAG TCAACGTAAC CAACCCTGCC 490
Line 13 TTAGCATCTT TTCCCTAACG GCCTCCTGCC TAAGCGGTAC TTCTAGCTTC GAACGGCGTC TGGGCTCCAG 558
Line 60 TTAGCATCTT TTCCCTAACG GCCTCCTGCC TAAGCGGTAC TTCTAGCTTC GAACGGCGTC TGGGCTCCAG 560
Line 92 TTAGCATCTT TTCCCTAACG GCCTCCTGCC TAAGCGGTAC TTCTAGCTTC GAACGGCGTC TGGGCTCCAG 560
line 61 TTAGCATCTT TTCCCTAACG GCCTCCTGCC TAAGCGGTAC TTCTAGCTTC GAACGGCGTC TGGGCTCCAG 559
Del_Ros c TTAGCATCTT TTCCCTAACG GCCTCCTGCC TAAGCGGTAC TTCTAGCTTC GAACGGCGTC TGGGCTCCAG 560
Line 13 GTTTAGTCGT CTCGTGTCTG GTTTATATTC ACGACAAAGA TCTATAGGGA CTTTAGGAGA TCTGGATTTT 628
Line 60 GTTTAGTCGT CTCGTGTCTG GTTTATATTC ACGACAAAGA TCTATAGGGA CTTTAGGAGA TCTGGATNTT 630
Line 92 GTTTAGTCGT CTCGTGTCTG GTTTATATTC ACGACAAAGA TCTATAGGGA CTTTAGGAGA TCTGGATTTT 630
line 61 GTTTAGTCGT CTCGTGTCTG GTTTATATTC ACGACAAAGA TCTATAGGGA CTTTAGGAGA TCTGNNTTTT 629
Del_Ros c GTTTAGTCGT CTCGTGTCTG GTTTATATTC ACGACAAAGA TCTATAGGGA CTTTAGGAGA TCTGGATTTT 630
Line 13 AGTACTGGAT TTTGGTTTTA GGAATTAGAA ATTTTATTGA TAGAAGTATT TTACAAATAC AAATACATAC 698
Line 60 AGTACTGGAT TTTGGTTTTA GGAATTAGAA ATTTTATTGA TAGAAGTATT TTACAAATAC AAATACATAC 700
Line 92 AGTACTGGAT TTTGGTTTTA GGAATTAGAA ATTTTATTGA TAGAAGTATT TTACAAATAC AAATACATAC 700
line 61 AGTACTGGAT TTTGGTTTTA GGAATTAGAA ATTTTATTGA TAGAAGTATT TTACAAATAC AAATACATAC 699
Del_Ros c AGTACTGGAT TTTGGTTTTA GGAATTAGAA ATTTTATTGA TAGAAGTATT TTACAAATAC AAATACATAC 700
Line 13 TAAGGGTTTC TTATATGCTC AACACATGAG CGAAACCCTA TAAGAA-CCC TAATTTCCCT TATCGGGAAA 767
Line 60 TAAGGGTTTC TTATATGCTC AACACATGAG CGAAACCCTA TAAGAA-CCC TAATTTCCCT TATCGGGAAA 769
Line 92 TAAGGGTTTC TTATATGCTC AACACATGAG CGAAACCCTA TAAGAA-CCC TAATTTCCCT TATCGGGAAA 769
line 61 TAAGGGTTTC TTATATGCTC AACACATGAG CGAAACCCTA TAAGAAACCC TAATTTCCCT TATCGGGAAA 769
Del_Ros c TAAGGGTTTC TTATATGCTC AACACATGAG CGAAACCCTA TAAGAA-CCC TAATTTCCCT TATCGGGAAA 769
Line 13 CTACTCACAC ATTNTTTATG G---------- ---------- ---------- ---------- ---------- 788
Line 60 CTACTCACAC ATTATTTATG G---------- ---------- ---------- ---------- ---------- 790
Line 92 CTACTCACAC ATTATTTATG G---------- ---------- ---------- ---------- ---------- 790
line 61 CTACTCACAC ATTNTTTNTG G---------- ---------- ---------- ---------- ---------- 790
Del_Ros c CTACTCACAC ATTATTTATG GAGAAAATAG AGAGAGATAG ATTTGTAGAG AGAGACTGGT GATTTCAGCG 839
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c TACCGAATTC CCGGGTTAAT TTCCAATTTG TTGGGCCTCC TCGAATAGGT TTCCTAATTC CCCATCCTCC 909
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c GTTGTTTCTA GCAACTTACT CCACCACTGA ATGCAGTCTT CAACTTCATC TTGTGGCGAC GCAACATCAT 979
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c TGTAAAATTG CGTTTGCTTC TCACAATCTG GAATCTCATC AGTTGTTAAC CGGACATTTG AAAATTCATC 1049
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c GGTTTTTCCG ACTTCTCTCG GCCAAGTAAC GTGCAATCCG GTGAAGGTCC GAGCTCGGGG TCTTACGATA 1119
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c TTAGTCAGCT TAATGGTTTT TGTGTTCATA ACATTTTTCC GGCATCGTTC TCCATCCTCG CCTAAATTCT 1189

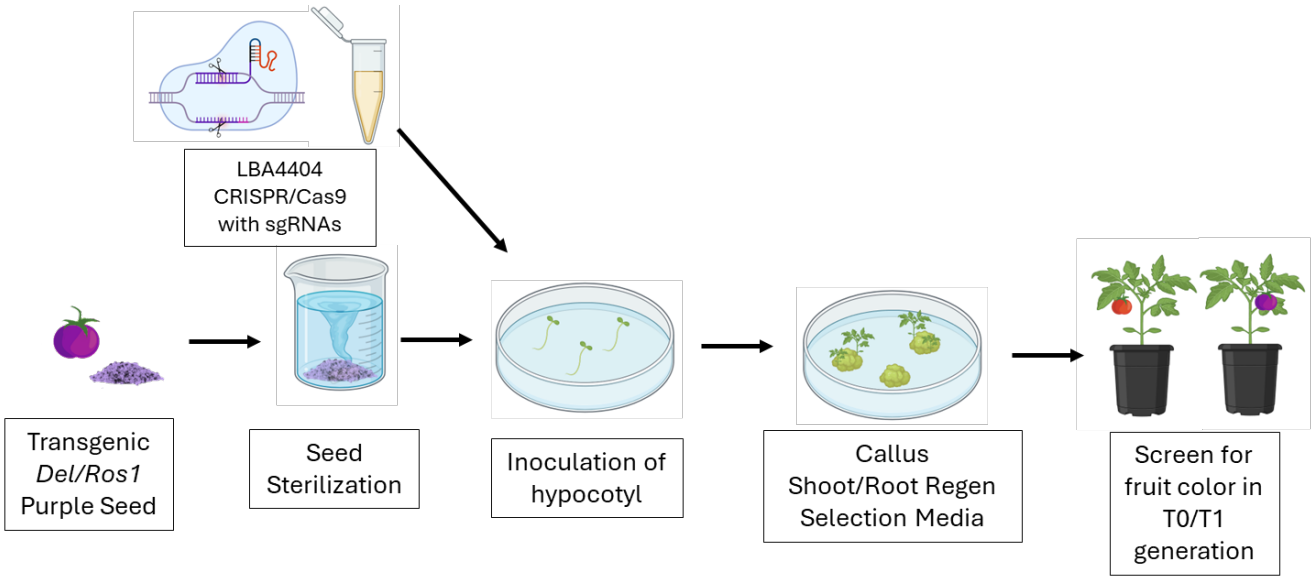

Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c TCCCCACATG AGTATTCCAA AAGTTCTTCA CGTCATTAGC TGTCCTTCCA GGAATTCTAC CAGCAATCAG 1259
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c CGACCATTTG TTACCCAACA GCTTATGAAG CCTCACAATT AGGTCCACTT CATCTCTCGA AAACCGACCT 1329
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c CTTTTGATAT TTGGCCTCAG ATAATTCAAC CACCTCAGCC TGCAACTCTT CCTACACCGG TTCAACCCTG 1399
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c CTCTGTGTGG AACTTGATGC CATTTCCCTT CACCATACTC TTCTATACAT TGCCTCAAGA GAGTGTCTTC 1469
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c TTCTTTGGTC CAAGTACCTT TTCTCACTCC ACGACAATTC TTTTCCATGG ATCCGCACTG TGAATGATTA 1539
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c GAATAATTTC TAAAAATCCC AATATGAGGA TGCCATATTT ATAATAGAAT AAAATAAAAT GTGAACAAAG 1609
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c AAAGAGATAA AGTAGTTCAC TTTTTGAAAT CTAAGAGAGA AATGGGAACA AGAAAGAGAC AAAAGTAGTT 1679
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c TCAAACAAAC TTCTCTTCTA AGTTTAGTCC CTTTTTAAAA TATGAAACCC AATACGTCTG ATTAAGAATA 1749
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c GAAAAATATC AAATTTTCAA TATAATTTAT ACTAATCGTT TTGAATTTTT CATACTGATA TAGTGTACGT 1819
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c TTCATCATAA CAACCAAAAC GTTGTTGCTT CACAACAATA ATATAGTAGT AGTTAATTTA TTATTTAGTA 1889
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c ATAAGTGGTC CTAAAAATTA GATAAATATT ACTATGATAA TATAAAAATA TTTGAGTCAG TCCTAAAAAA 1959
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c TTATTTAGTA TTCATACATG AATCAAACTG ATTAGTTAAG TGTCAACAAT TGGACAAGTG GCATGGAGGT 2029
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c TGTAAAAGAA TGACATAAGC CAACTGCTAT TTTTATCCAA AAAAAAGAAG ACAACTTGAC AACTACATTT 2099
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c CTTTTATTTT TATAAATTTA CTAATATCTT CTATGCAAAA TTATTCGGTG CCTTTCTAAA CTTTAAGGTT 2169
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c TTTATTTGAT GTACACCTAA ATTATATTTT ATTTTAATCA CTCCACTGGA CTTGTTTATT CCTTCATCAT 2239
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c ATACACCTAC TCCTATTATG ACTACAAGTT GGCAAAAGTA ATGATATGAA TTTCTACTTA AATAAATAAT 2309
Line 13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 788
Line 60 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Line 92 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
line 61 ---------- ---------- ---------- ---------- ---------- ---------- ---------- 790
Del_Ros c AGTCACCTAG ATAAATTAAT TTAACAAAAG ATAAATATCA AACCTTCTCA CCTAAAATTT TGAGCAAAAC 2379



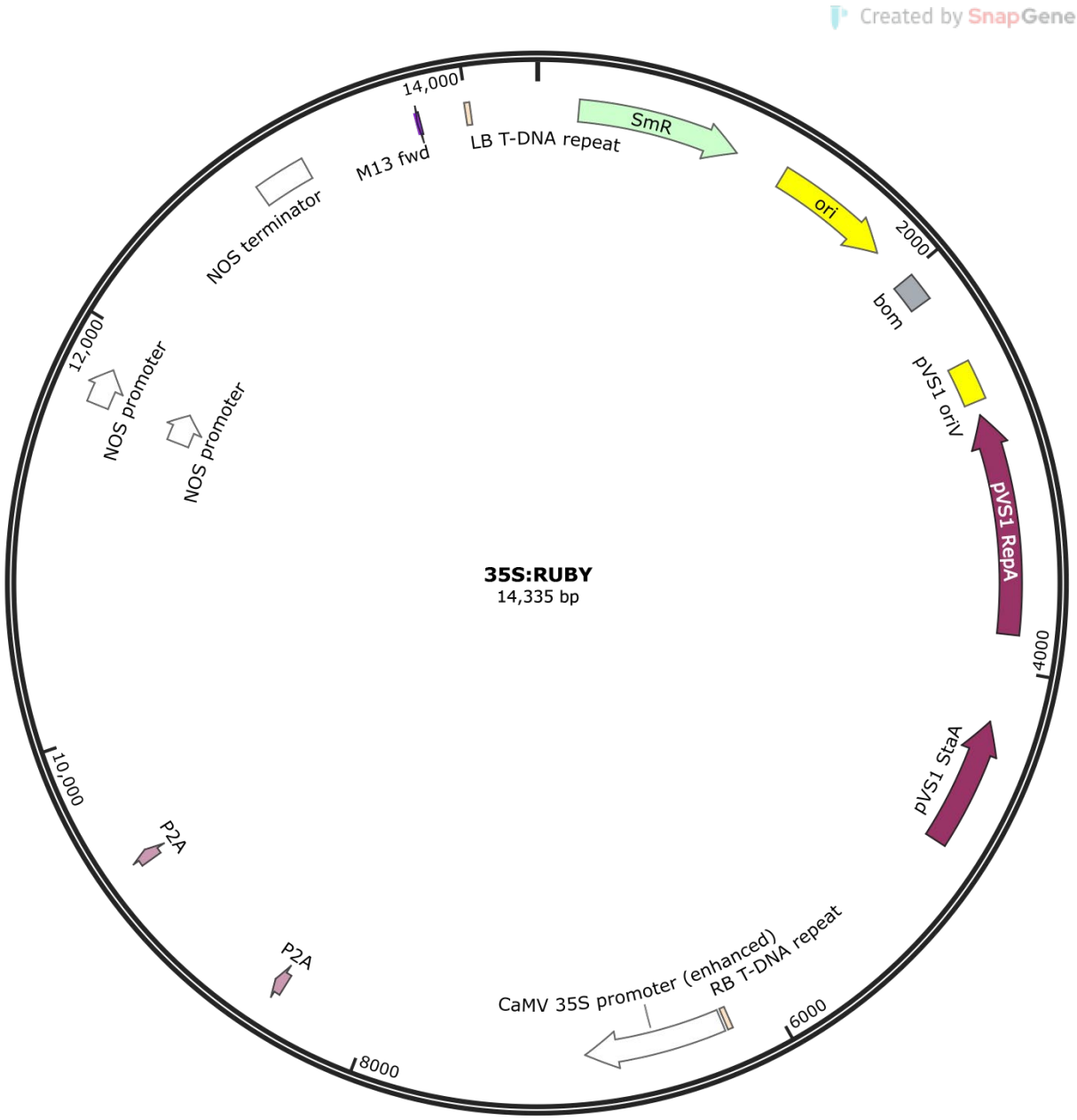
Supplementary Figure 2 Line-specific sequence analysis of EMS lines. In the Del/Ros1 Construct



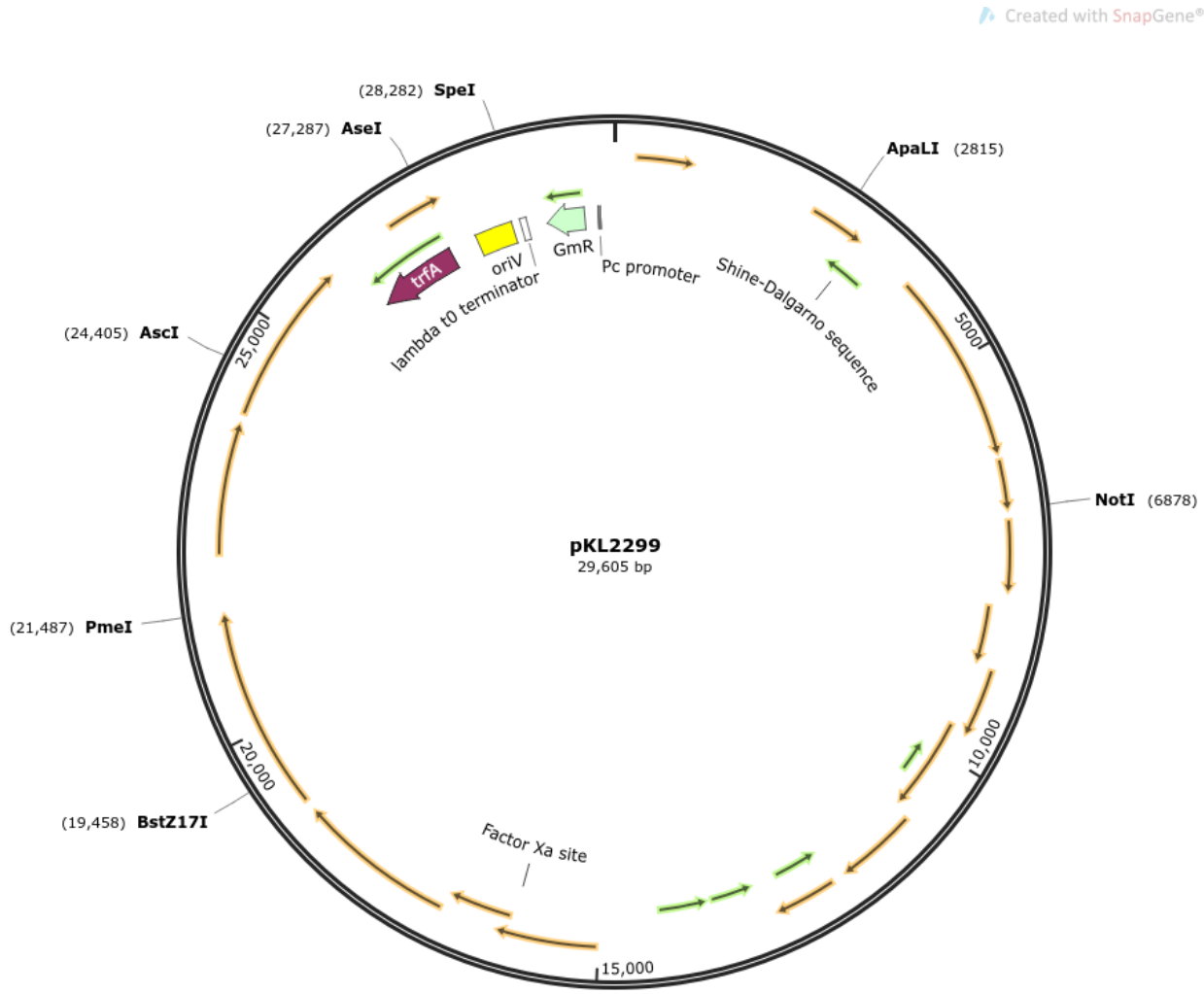
Supplementary Figure 3 EMS Treatment of Del/Ros1 expressing transgenic seeds



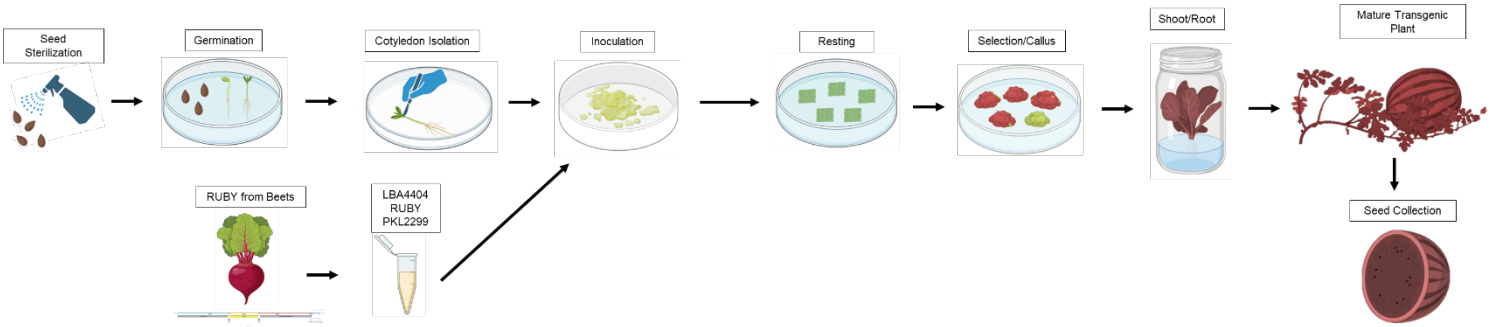
Supplementary Figure 4 CRISPR/Cas9 editing of Del/Ros1 expressing transgenic seeds



Supplementary Figure 5 35S: Ruby construct map



Supplementary Figure 6 Map of virulence genes from Agrobacterium tumefaciens pTiBo542 in PKL 2299 plasmid



Supplementary Figure 7 Process of watermelon transformation