

## Article

# Genome-wide Identification of the Aconitase Gene Family in Tomato (*Solanum lycopersicum*) and CRISPR-based Functional Characterization of *SlACO2* on Male-sterility

Zafer Secgin<sup>1</sup>, Selman Uluisik<sup>2</sup>, Kubilay Yıldırım<sup>3</sup>, Mohamed Farah Abdulla<sup>1</sup>, Karam Mostafa<sup>1,4</sup> and Musa Kavvas<sup>1,\*</sup>

<sup>1</sup> Faculty of Agriculture, Department of Agricultural Biotechnology, Ondokuz Mayıs University, Samsun, Turkey

<sup>2</sup> Burdur Mehmet Akif Ersoy University, Burdur Food Agriculture and Livestock Vocational School, 15030 Burdur, Turkey

<sup>3</sup> Faculty of Science, Department of Molecular Biology and Genetics, Ondokuz Mayıs University, Samsun, Turkey

<sup>4</sup> The Central Laboratory for Date Palm Research and Development, Agricultural Research Center (ARC), Giza 12619, Egypt

\* Correspondence: Associate Prof. Dr. Musa Kavvas, Ondokuz Mayıs University, Faculty of Agriculture, Department of Agricultural Biotechnology, Samsun, Turkey; Phone: +903623121919; Fax: +903624576034; E-mail: musa.kavvas@omu.edu.tr

**Abstract:** Tomato (*Solanum lycopersicum*) is one of the most cultivated vegetables in the world due to its consumption in a large variety of raw, cooked, or processed foods. Tomato breeding and productivity highly depend on the use of hybrid seeds and their higher yield, environmental adaptation, and disease tolerance. However, the emasculation procedure during hybridization raises tomato seed production costs and labor expenses. Using male sterility is an effective way to reduce the cost of hybrid seeds and ensure cultivar purity. Recent developments in CRISPR genome editing technology enabled tomato breeders to investigate the male sterility genes and to develop male-sterile tomato lines. In the current study, the tomato Aconitase (*SlACO*) gene family was investigated via in-silico tools and functionally characterized with CRISPR/Cas9-mediated gene disruption. Genome-wide blast and HMM search represented two *SlACO* genes located on different tomato chromosomes. Both genes were estimated to have a segmental duplication in the tomato genome due to their identical motif and domain structure. One of these genes, *SlACO2*, showed a high expression profile in all generative cells of tomato. Therefore, the *SlACO2* gene was targeted with two different gRNA/Cas9 construct to identify their functional role in tomato. The gene was mutated in a total of 6 genome-edited tomato lines, 2 of which were homozygous. Surprisingly, pollen viability was found to be extremely low in mutant plants compared to their wild-type (WT) counterparts. Likewise, the number of seeds per fruit also sharply decreased more than fivefold in mutant lines (10-12 seeds) compared to that in WT (67 seeds). The pollen shape, anther structures, and flower colors/shapes were not significantly varied between the mutant and WT tomatoes. The mutated lines were also subjected to salt and mannitol-mediated drought stress to test the effect of *SlACO2* on abiotic stress tolerance. The results of the study indicated that mutant tomatoes have higher tolerance with significantly lower MDA content under stress conditions. This is the first CRISPR-mediated characterization of *ACO* genes on pollen viability, seed formation, and abiotic stress tolerance in tomato.

**Keywords:** aconitate hydratase; CRISPR/Cas9; male sterility; seedless tomato

## 1. Introduction

Tomato (*Solanum lycopersicum* L.) is one of the most produced fruits or vegetable crops that is produced five million ha all over the world. Tomato contains various health-promoting compounds essential to a balanced human diet. It is the natural source of

vitamins C and E as well as large amounts of carotenoids and polyphenolics, such as lycopene, sucrose, hexoses, citrate, and malate. Tomato is an indispensable component of meals in many cultures and one of the most important ingredients in the processed food industry [1]. In addition to their commercial and dietary importance, the tomato has become a model species for plant science due to its available genome sequence, ease of transformation, and short life cycle. It has been used to understand the molecular mechanism of fruit development and identify genetic responses to biotic and abiotic stress factors [2, 3]. Since the discovery of heterosis (hybrid vigor) and the development of hybrid seed technology, most tomato varieties are grown from F1 hybrid seeds to have better offspring in terms of uniformity, yield, and stress tolerance. The common way to produce hybrid tomato seeds worldwide is the manual emasculation of male gametophytes to generate the female parent. However, this technology is a labor-intensive practice that significantly increases the cost of hybrid tomato seed production [4]. The failure of plants to produce functional anthers, pollen, or male gametes, which is also known as male sterility (MS), can be induced by genic or cytoplasmic systems, chemical hybridizing agents, or transgenic technology in plants. Incorporating male sterility into tomato breeding is a critical task to reduce the cost of hybrid seed production and ensure high varietal purity [5].

Over the years, researchers have developed several methods for integrating male sterility into the tomato genome. Several tomato genes controlling male sterility have been studied; many mutants, TILLING populations, recombinant inbred lines, and genetically modified tomato lines have been developed and tested in the last decades for this purpose [6, 7]. Despite all this intensive research, superior male-sterile tomato lines have not been used on a large scale in tomato hybrid seed production. However, recent developments in genome editing technology have opened a new era to study gene function and develop new plant cultivars suitable for any condition. The latest version of the gene editing technology, CRISPR/Cas9, provides a novel and convenient method for producing male-sterile lines in soybean, maize, rice, wheat, and even tomato [8, 9].

Aconitase, also known as citrate (isocitrate) hydratase or aconitate hydratase (EC 4.2.1.3), is a metabolic enzyme containing Fe-S cluster that catalyzes the reversible isomerization of citrate to isocitrate via cis-aconitate in the tricarboxylic acid cycle (TCA). In humans, the *ACO1* gene is a cytosolic aconitase called iron-responsive element-binding protein 1 (IRP1) that contains an AcnA\_IRP domain. IRP1 can recognize the RNA stem-loop structure of iron-responsive elements, localizes to the 5' untranslated regions of ferritin mRNA, and inhibits ferritin translation [10]. Especially in higher plants, all aconitase sequences detected to date resemble cytosolic aconitase isoforms more than mitochondrial isoforms in animals [11]. Although the evolutionary patterns of the ACO family in land plants are poorly understood, the function of ACOs has been well-studied in model plants, such as *Arabidopsis* [12]. In *Arabidopsis*, inhibition of *AtACO* activity caused disturbances in the TCA cycle or cytosolic citrate metabolism [13]. Moreover, inhibition of aconitase by nitric oxide increased the expression of alternative oxidase, which is a regulator of mitochondrial stress responses [7]. Another study showed that the *OsACO1* gene in rice was induced by heat stress [14]. Transcriptome sequencing was performed in cytoplasmic male sterile lines of *Brassica napus*. Six genes involved in the TCA cycle, mainly encoded aconitase, were shown to be significantly down-regulated [15]. Although aconitase genes are expressed at low levels during most developmental stages, a dramatic increase in its expression is observed during seed, pollen maturation, and germination [16]. Northern blot analysis was used to examine the expression of three *AtACO* genes in different plant organs, including the seed, root, stem, leaf, and flower. [17]. In the same study, the germination rate of seeds obtained by crossing ACO mutant *Arabidopsis* with non-mutant lines was reduced by 23-25%. The double mutant, *Ataco1/Ataco3*, caused abortion in seeds, implying that the two genes have a crucial role in embryo formation and seed development. Based on these studies, the relationship of the ACO genes with citric acid contents, oxidative stress, and iron homeostasis was determined in plants. However, to the best of our knowledge, there is no study characterizing an ACO gene related to pollen development and male sterility in plants. Therefore, the current study aimed to identify *SlACO* genes

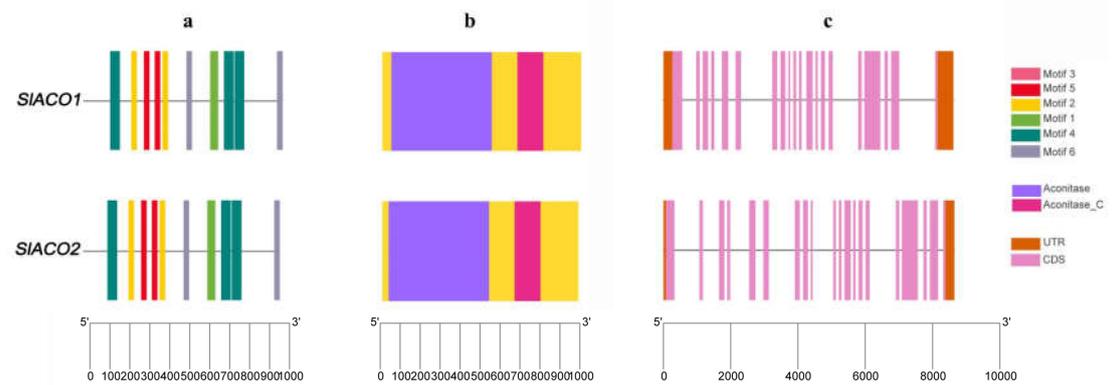
in tomato genome using bioinformatics tools. Then, CRISPR/Cas9-mediated loss-of-function assay was carried out in a mutant tomato line to investigate the potential regulatory role of *SIACO2* (cytoplasmic) genes involved in tomato male sterility and seed formation.

## 2. Results

Putative *aconitase* (*SIACO*) genes were searched *Solanum lycopersicum* genome using various in silico tools, and two genes were identified (*SIACO1*- *Solyc07g052350.3.1* and *SIACO2*- *Solyc12g005860.2.1*). The detailed information on each gene is listed in Table 1. During this analysis, two more genes (*Solyc03g005730.4.1* and *Solyc09g090900.4.1*) containing the aconitase domain but encoding the enzyme 3-isopropyl malate dehydratase were identified in the tomato genome. Since these genes do not have aconitase and swivel domains together, they were not considered true aconitase genes and were not included in subsequent studies. Both *SIACO* proteins contain an aconitase and swivel domain, and their protein lengths range from 982 to 996 amino acid residues. Their molecular weights (Mol. Wt) were found to be as 107.15kDa and 108.2 kDa, respectively. Sequence analysis of *SIACO1* and *SIACO2* showed that the isoelectric point (pI) of the *SIACO* proteins was 6.52 and 7.29, respectively, and the large mean hydrophobicity (GRAVY) was -0.178 and -0.223. The results of the instability index show that *SIACOs* are stable in the test tube. (instability index > 40). The predicted subcellular localization of *SIACO* proteins indicated that these proteins were highly active in plastid and mitochondria (Figure S1). Examining the exon-intron structure of each member of *SIACO* enabled us to discover more about the structural variation among *SIACO* genes. As illustrated in Figure 1c, both *SIACO1* and *SIACO2* genes have 20 exons. In addition, the intron length of *SIACO* genes was observed to be variable in tomato plants. It is generally agreed that protein motifs play a vital role in the interaction of different modules in transcriptional complexes and the transmission of signals and that they are also intimately related to gene classification [18]. In this way, the phylogenetic relationship and classification of *SIACO* genes were supported by motif analysis. Six conserved motifs of *SIACO* proteins were analyzed using the MEME suite (Figure 1a). Chromosomal locations of the *SIACO* genes were visualized using TBtools software, and it was found that *SIACO1* and *SIACO2* were located on chromosomes 7 and 12, respectively. Gene duplication events determined that there are segmental duplications between the *SIACO1*/*SIACO2* (Figure S2). The synonymous ratio (Ks), non-synonymous ratio (Ka), and Ka/Ks of these genes were calculated, and the results are given in Figure S2.

**Table 1.** Physiochemical properties of *SIACO* proteins.

Gene Name	Chromosome	Length (bp)	CDS (pb)	Protein Length (A.A)	Accession number	pI	Molecular Weight (Kda)	Instability index (II)	Stability	GRAV Y	Solubility/Score
<i>SIACO1</i>	ch07	8621	2988	996	XP_004243472.1	7.29	108.20	32.74	stable	-0.178	Soluble/0.552
<i>SIACO2</i>	ch12	8650	2946	982	XP_004251517.2	6.52	107.15	35.57	stable	-0.223	Soluble/0.568



**Figure 1.** Motif structure, conserved domains, and exon-intron architecture of SIACO genes. a) Structure of 6 conserved motifs analyzed through MEME suite tool online. Legends indicate the color of each motif. b) Conserved domain structure of SIACO genes in tomato. Two conserved domains were found and presented in purple and dark pink colors. c) Exon-intron architecture of SIACO genes. The light pink boxes indicate exons, the solid line indicates introns, and the brown boxes indicate the non-coding untranslated regions.

To understand the expansion of aconitase hydratase genes in different plant species, the full-length amino acid sequences of 28 ACO genes from six model plants were used to construct the phylogenetic tree using the maximum likelihood method (Figure S3). A total of 26 aconitase genes, 3 in *Arabidopsis thaliana* (AtACO1, AtACO2, and AtACO3), 7 in *Zea mays*, 6 in *Medicago truncatula*, 4 in *Populus trichocarpa*, 4 in *Sorghum bicolor*, and 2 in *Brachypodium distachyon* genome, were identified to be used in phylogenetic analysis. The result showed that the 28 ACO genes were divided into ten subgroups, here named groups I to X. SIACO genes are clustered in III subgroups. In this phylogenetic classification, subgroups I, II, and X are represented by one member each, and subgroups III, IV, V, and IX are represented by two members, but most are in subgroup VII. To figure out conserved amino acids in the swivel domain of SIACOs and AtACOs, a ClustalW alignment was performed. In the analysis performed to compare the two organisms in terms of the swivel domain consisting of 131 amino acids, only 28 amino acid differences were found; that is, 78.6% of them were completely conserved (Figure S4).

### 2.1. *In-silico functional and expressional characterization of tomato SIACO genes and determination of the SIACO2 expression level by RT-qPCR*

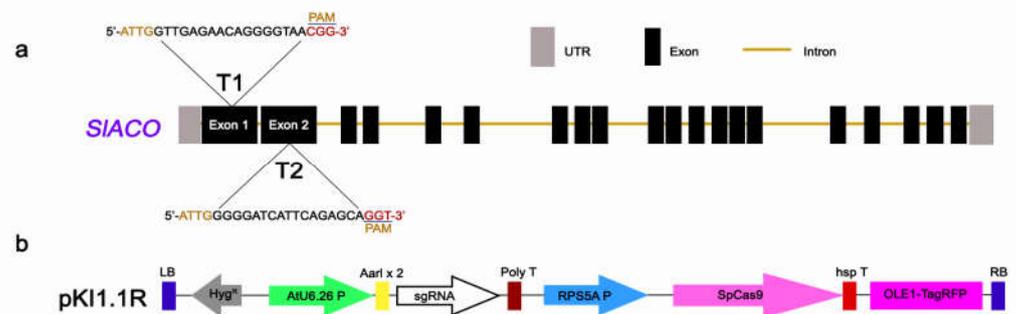
To study the tissue-specific expression patterns of members of the SIACO gene family, we used pollen cell development-related RNA-seq data to examine functional gene expression analysis in tomato [19]. As shown in Figure S5a, SIACO2 was highly expressed in microspores, mature pollen cells, and generative and sperm cells examined in the study. Among these four tissues that were used for the analysis, it was discovered that the expression of the SIACO1 gene was only down-regulated in the generative cell, while it was up-regulated in all of the other tissues. According to the results obtained from Tomato eFP browsers, it has been determined that both ACO genes have high activity in all tissues such as flower, vegetative meristem, seedling root, seedling shoot, developing fruit, mature fruit, leaves, stem (Figure S5b).

The highly correlated 25 genes related to pollen development were extracted based on the publicly available transcriptome data (GSE109672 and GSE117185). In these studies, pollen development at tetrad, post-meiotic and mature stages were investigated in different tissues, including microspores, generative cells, pollen cells, callus, fully expanded leaves, pistil, and mature pollen grain (Figure S6). SIACO genes, four genes related to male sterility; SIPHD\_MS1 [20], SIMS10 [21], SIAMS [5], SISTR1 [22], and these 25 genes were hierarchically clustered by Manhattan distance in three anatomical parts

including microspores, generative cell and pollen cell (Figure S6). Three sub-clusters were observed, and the SIACO2 gene was found to cluster separately from SIACO1. Later, the STRING program was used to understand the functional interaction between those 25 genes and two genes related to male sterility with SIACOs (Figure S6). Four gene clusters were created with 0.625 local clustering coefficient score based on the MCL inflation parameter (Table S1). Ten predicted functional partners were revealed to create the functional interaction network of SIACOs with pollen development and male sterility (Table S2). There are 31 nodes and 99 edges. Protein-protein interaction enrichment p-value is less than  $1.0 \times 10^{-16}$ . All of the findings that were obtained indicate that it is reasonable to choose SIACO2 as the target in plants to be created with the CRISPR/Cas9 for the purpose of performing functional characterization of the process of pollen formation.

## 2.2. Generation of SIACO2 mutants using the CRISPR/Cas9 gene-editing system

To investigate the male-sterility function of SIACO2, mutants were generated using the CRISPR/Cas9 gene editing technology. Two target sites, Target 1 (Exon 1) and Target 2 (Exon 2), were selected for SIACO2 knock-out studies (Figure 2a). Both gRNAs were separately cloned into the plant expression vector pKI1.1R and expressed under control Arabidopsis constitutive U6-26 promoter (Figure 2b). The resulting constructs were introduced into *Agrobacterium tumefaciens* strain GV3101, and the tomato cultivar 'Crocker' was used for transformation separately. A total of 50 cotyledon explants were used for each gRNA. Callus formation started after 10-12 days of transformation, and putative genome-edited plants were formed after 30 days. A total of 22 independent transgenic plants, 18 of which were from gRNA1, were obtained by *Agrobacterium*-mediated transformation. Using PCR and gene-specific primers for the hptII gene, we could validate the presence of transgenes (Figure S7).

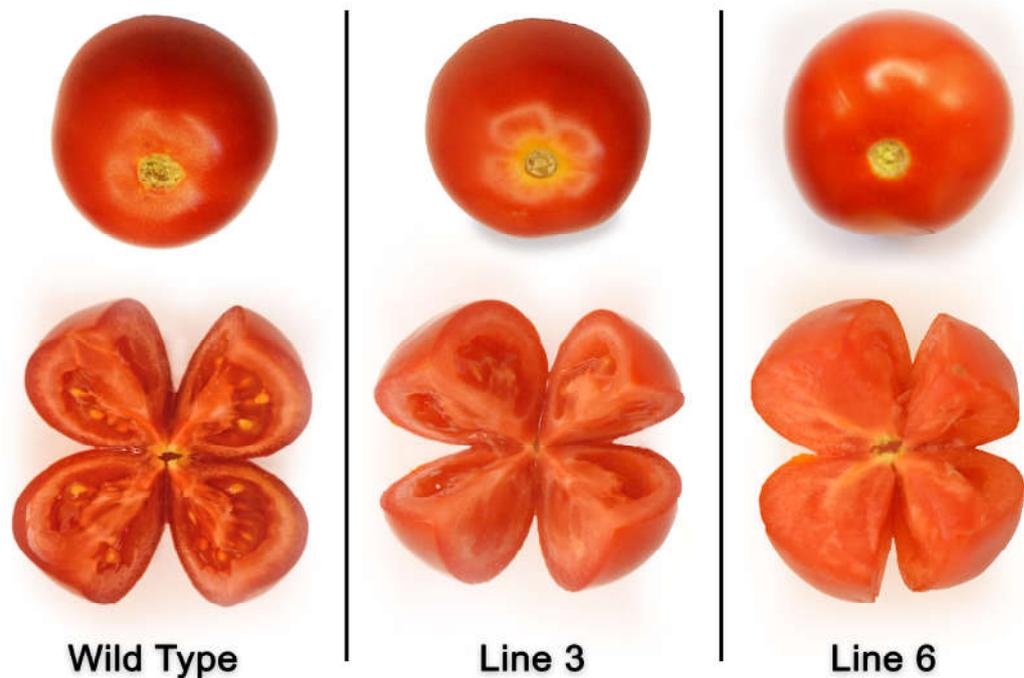


**Figure 2.** CRISPR/Cas9-mediated genome editing of SIACO2 gene a) Schematic illustration of the two target sites in SIACO2 genomic sequence. Target 1 and Target 2 sequences are shown in capital letters, and the protospacer adjacent motif (PAM) sequence is marked in red. b) Schematic diagram of Cas9/sgRNAs expression vector (pKI1.1R) used for *Agrobacterium*-mediated transformation of tomato. .

Sanger sequencing was used to check for mutations in the PCR product that was produced from four lines of gRNA2 and six lines of gRNA1. As a result of Blast and Synthego ICE analyses, it was determined that gRNA2 did not cause any mutations, and Line 1, line 2, line 4, and line 5 plants had the same mutation pattern in gRNA1. Therefore, considering these lines may be clones, only the results of line 1, line 3, and line 6 were given as a result of Sanger sequence analysis (Figure 3). As seen in Figure 3, although there is a homozygous mutation in line 3 and line 6, a heterozygous mutation is observed in line 1. In this context, 13 base deletions occurred in line 3 and 2 base deletions in line 6.



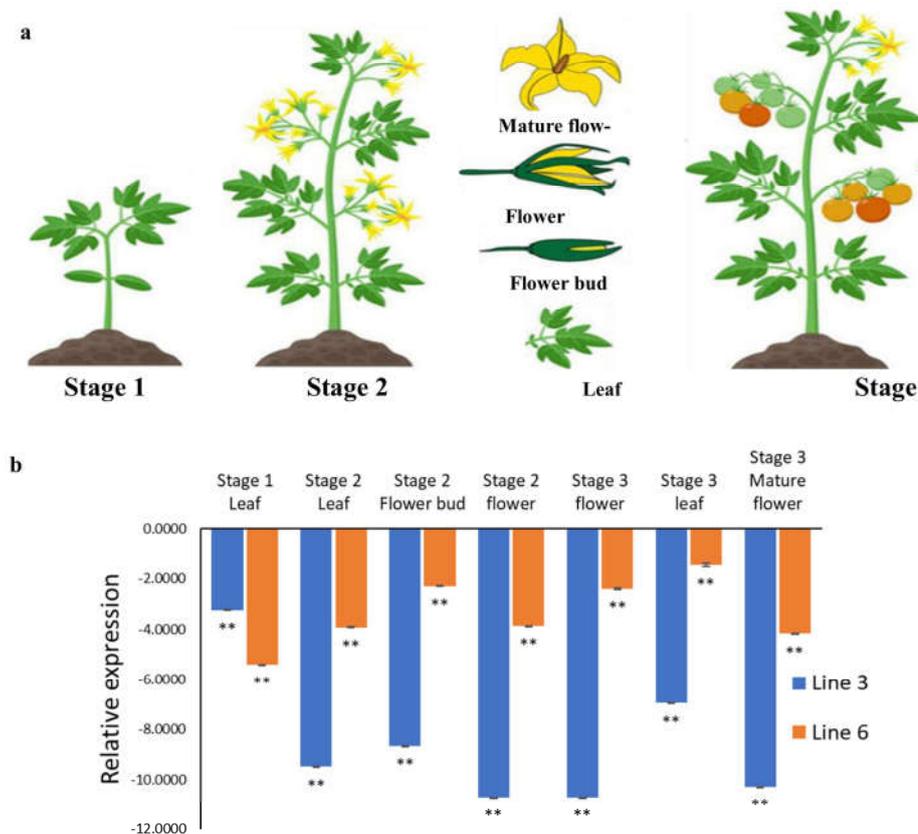
showed that the number of seeds/fruit of WT plants was 67; it was observed as 10.2 in the Line 6 plants indicating that 85% of seed formation was inhibited in mutant lines compared to WT plants (Figure S9). Although there is a fairly big difference in the total quantity of seeds produced by wild-type and mutant plants, the seeds produced by either kind do not differ in terms of their size or shape (Figure S9).



**Figure 5.** Harvested fruits and seeds formed in ripe fruits of WT and SIACO2 mutant plants.

#### 2.4. Expression analysis of SIACO2

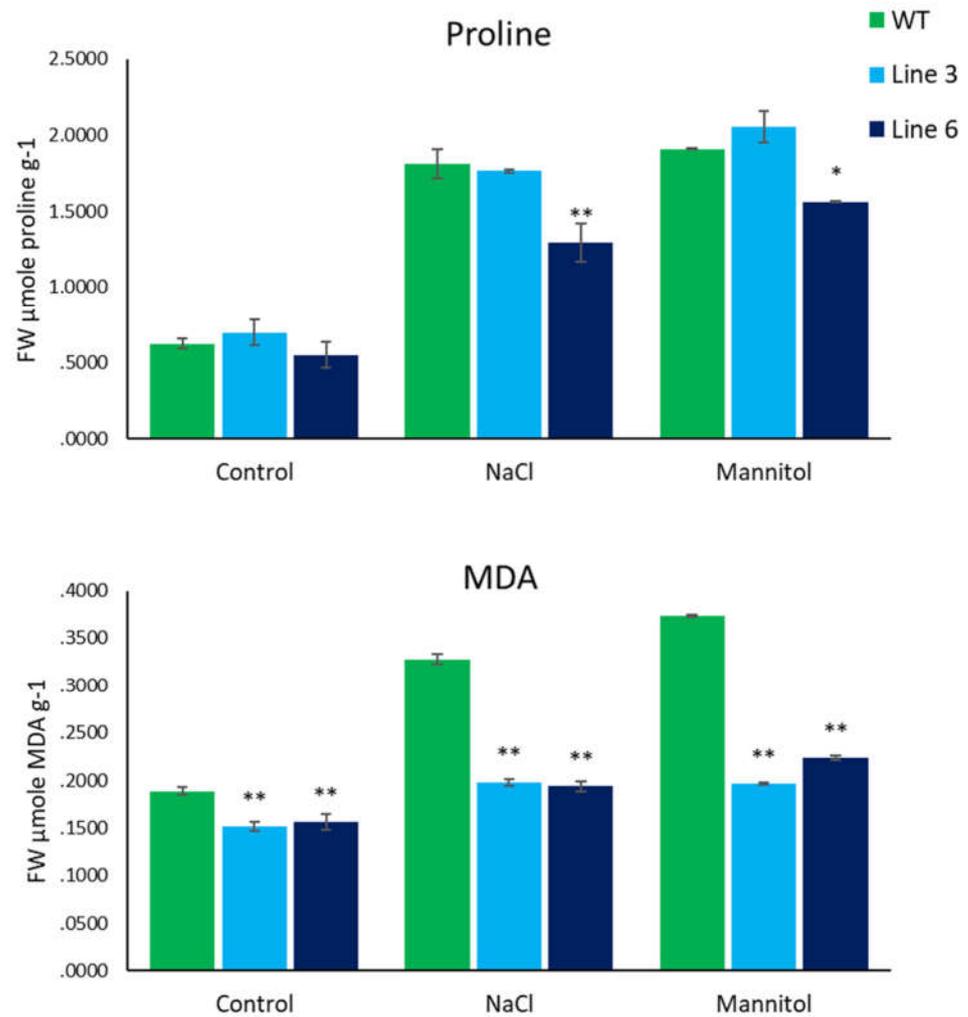
At different developmental stages, expression levels of the SIACO2 gene in various tissues, including flower, bud, and leaf, were evaluated using qRT-PCR in both WT and mutant plants to confirm the effectiveness of the CRPSR/Cas9 system (Figure 6a). It was found that the expression of the silenced SIACO2 gene in all tissues at different developmental stages analyzed was significantly decreased in genome-edited plants compared to WT plants ( $p < 0.01$ ). This decrease was more in line 3 plants. The highest decrease was observed in the 2nd and 3rd stage flower tissues when WT was compared with the plant (Figure 6b). According to the WT plant, the tissue in which the target gene works most is the 3rd stage leaf tissue isolated from Line 6.



**Figure 6.** Relative gene expression between WT and genome-edited lines at different stages. a) Explants and their stages used for qRT-PCR analysis. b) Relative expression analysis of SIACO2 knock-out and WT plants by qRT-PCR at different developmental stages of leaf and flower tissues (Actin gene was used as an internal reference gene). Data provided represent fold expression Log<sub>2</sub> (WT/genome-edited lines). Values represent the average  $\pm$  SD of three biological replicates of each reaction. \*\* indicate significant differences at  $P < 0.01$ .

### 2.5. Evaluation of the effect of salt and drought stress on genome-edited lines

To simulate salinity stress, we germinated seedlings of WT and SIACO mutant lines on MS media supplemented with 200 mM NaCl under in vitro conditions. In the same manner, 200 mM mannitol was added independently to simulate drought stress. Leaves of 10-week-old seedlings were collected for MDA and proline analysis. As shown in Figure 7, there was no significant difference in the level of proline content on tomato seedlings under control conditions. Line 6 showed a significantly lower proline content when treated with NaCl as well as mannitol treatment compared to WT. On the contrary, no significant difference was observed in the proline content comparing line 3 to WT in both treatments. A significantly lower level of MDA content was noticed under control treatment on both line 3 and line 6 compared to WT. Similarly, line 3 and line 6 mutants exhibited a significantly lower MDA content compared to that of WT when treated with NaCl and mannitol.



**Figure 7.** Comparative analysis of proline and MDA content in WT and SIACO2 mutant lines under control, 200mM NaCl, and 200mM mannitol treatments from 10 weeks old leaves. Results represent mean values; error bars indicate the  $\pm$  standard error ( $n = 3$  replicates). Single asterisks indicate statistically significant differences ( $P < 0.05$ ), and double asterisks represent a very significant difference ( $P < 0.01$ ) according to ANOVA (IBM, SPSS 26).

### 3. Discussion

Compared to cytoplasmic male sterile lines, it is much easier and more useful to produce hybrid seeds with the use of genome editing techniques [23]. Thus, the development of germplasm sources of male sterile lines depends on developing novel crop production strategies. The male sterile phenotype is usually controlled by a recessive mutation of the nuclear gene; therefore, it is possible to generate male sterile mutants using the CRISPR/Cas9 system. The SIMS10 gene, which encodes the basic helix-loop-helix transcription factor (bHLH), is involved in meiosis and programmed cell death in the tapetum during microsporogenesis. Therefore, modification of the SIMS10 via CRISPR/Cas9 leads to male sterility in tomato plants [21]. In a recent study using AMS, another gene encoding a basic helix-loop-helix (bHLH) TF was knocked out by CRISPR/Cas9 and over-expressed in tomato. SIAMS downregulation or upregulation led to abnormal pollen development, which in turn decreased pollen viability, and subsequently generated male-sterile lines [5]. In our laboratory, we have generated completely male-sterile lines by mutating the SIPHD\_MS1 gene (Solyc04g008420), encoding a PHD-type transcription factor [20]. In

another study, male-sterile lines were obtained by silencing the strictosidine synthase-like (SISTR1) gene, a homolog of the Ms45 gene of the maize plant in tomato [22]. Taken together, these results demonstrate the significance of a biotechnology-based genetic male-sterility system, including the development of novel male-sterile tomato lines.

In this study, *in vitro*, and *in silico* research methods were utilized to examine the aconitase gene found in the tomato genome. In keeping with the literature, such as the information that there are three ACO genes in both *Arabidopsis* [13] and *Citrus* [24], two ACO genes were found in the tomato genome. Phylogenetic analysis showed that two SIACO proteins, SIACO1 (Solyc07g052350) and SIACO2 (Solyc12g005860), fall within the same clade, which have the aconitase swivel superfamily domain in addition to aconitase superfamily domain (Figure S3). Moreover, it was reported that these two genes participate in the functioning and development of male gametophytes by having a role in energy-related pathways (Glycolysis and Krebs cycle) as a prerequisite for pollen maturation [25]. Additionally, the expression pattern of the SIACOs is presented in diverse tissues of tomato, including microspores, mature pollen, and generative cell (Figure S4), in which SIACO2 had the strongest expression in all tissues. Collectively, this evidence implies that SIACO2 is a possible target for editing to create male sterility in tomato.

Gene editing technologies are becoming powerful tools for crop improvement and breeding. Among them, the CRISPR/Cas9 system has been extensively used for genome editing in various organisms, especially plants, due to its simplicity, high efficiency, and multiplex capabilities. CRISPR/Cas9 technology was used in this study to knock out SIACO2 to confirm the results of *in silico* tools. The TCA cycle, including the aconitase enzyme, has significant application potential in improving crop yields [4]. There have been numerous studies showing that TCA cycle-related genes involved in seed germination [26], root development [27], flower development [28], and fruit ripening [29]. Aconitase is an iron-sulfur (Fe-S) enzyme involved in the assembly of a very complex cluster. Aco1/Aco3 double mutant of *Arabidopsis* caused the abortion of seeds, indicating the two genes had a crucial role in embryo formation and proper seed development [17]. Here, we generated six SIACO2 mutant tomato lines and confirmed the mutagenesis using molecular and physiological analysis. While two of them were found to be homozygous, the rest were heterozygous. Even though the flowers and seeds generated by homozygous mutant lines seemed identical to those produced by wild-type lines, homozygous mutant lines produced an extremely low number of seeds. A finding that scanning electron microscope analysis of mutant flower anthers reveals they contain very little pollen contributed to the credibility of this evidence. Additionally, the results of our study do not appear to be consistent with the information found in the previous research by Carrari, Nunes-Nesi [30]) in which the tomato plants created by silencing of the ACO1 gene produced darker leaves, indicating a greater quantity of carotenoids. On the contrary, we found that the leaves of wild-type plants and mutant lines were indistinguishable from one another.

To further characterize the remarkable phenotypic difference between WT and SIACO mutant lines, physiological studies were applied by treating tomato seeds with abiotic stressors in the form of salt and drought. The level of MDA content is used as a direct marker of membrane damage and lipid peroxidation to measure cell damage in plants under environmental stress [31]. In our study, mutant lines showed a significant reduction in MDA content compared to WT, indicating reduced oxidative damage caused by salinity and drought stress. This result is in accordance with previous reports for the level of MDA content as inversely proportional to the degree of tolerance to abiotic stress [32-35]. An increase in proline content in cells can function as an osmotic adjustment mediator, an eliminator of free radicals, and a potential redox buffer under environmental stress [36, 37]. In this study, no significant difference in proline content was observed in Line 3 as opposed to an increase in the proline level in Line 6 compared with WT. In line 6, the level of free proline content was significantly reduced. This underlined the fact that SIACO2 mutant line 6 might augment drought and salinity stress-induced membrane

damage. A similar result was previously reported in tomato plants with reduced proline accumulation [34, 35].

This suggests that the loss of function of SLACO might be related to lower ROS production, which can be linked to enhanced resistance to oxidative damage and increased tolerance of tomato seedlings to salt and drought stress.

#### 4. Materials and Methods

##### 4.1. Genome-wide identification and analysis of SLACO genes

The AtACO1 sequence (AT4G35830) was obtained from the NCBI (<https://www.ncbi.nlm.nih.gov/gene/>) and used as a query for blast search against the tomato genome in the Phytozome v13 (<https://phytozome-next.jgi.doe.gov/>) database to identify probable SLACO genes. After blast analysis, a sequence search was performed using the Pfam database (<http://pfam.xfam.org/>) to find the domain sequence of this gene with a length of 795 amino acids. Putative SLACO genes were identified by searching for proteins containing the Aconitase domain using Hidden Markov Model (HMM) search tool. The genes containing the Aconitase domain from the tomato plant were analyzed using the HMMER 3.1b2 database (<http://hmmer.org>). The length of the open reading frame (ORF), the isoelectric point of proteins (pI), and the molecular weight (MW) of identified ACO proteins were analyzed using the ProtParam tool on the ExPASy web server ([http://www.expasy.ch/tools/pi\\_tool.html](http://www.expasy.ch/tools/pi_tool.html)) and Phytozome v13 web tool (<https://phytozome.jgi.doe.gov/pz/portal.html>).

##### 4.2. Phylogenetic analysis and multiple sequence alignment of SLACO genes

Using the same methodology that was utilized for the identification of SLACOs, the full-length amino acid sequences of ACO proteins derived from *Arabidopsis thaliana*, *Populus trichocarpa*, *Medicago truncatula*, *Zea mays*, *Brachypodium distachyon*, and *Sorghum bicolor* were determined. Alignments of amino acid sequences of ACO proteins were performed with ClustalW, and the phylogenetic tree was generated with MEGA X software using the neighbor-joining (NJ) method with bootstrap analysis (1000 replicates) [38].

##### 4.3. Exon/intron structure and conserved motif analysis of SLACO genes

The DNA, protein, and CDS sequences of SLACOs were downloaded from the Phytozome v13 database. Conserved motif analysis of SLACO protein sequences was discovered with an online MEME (<https://meme-suite.org/meme/>) server and visualized with TBtools software [39]. In the process, the optimum width was set to 80-100 amino acids, and any number of motif repetitions and the maximum number of motifs was selected as 10.

##### 4.4. Chromosome Location and Determination of Gene Duplications of SLACO

Distributions of SLACO genes on chromosomes and gene duplications were performed using TBtools software (<https://github.com/CJ-Chen/TBtools>). The tomato-tomato block file was downloaded using the Plant Genome Duplication Database (<http://chibba.agtec.uga.edu/duplication/>) to determine gene duplications. Ka/Ks ratios were calculated using TBtools software by selecting the data of SLACO genes from the obtained Block file.

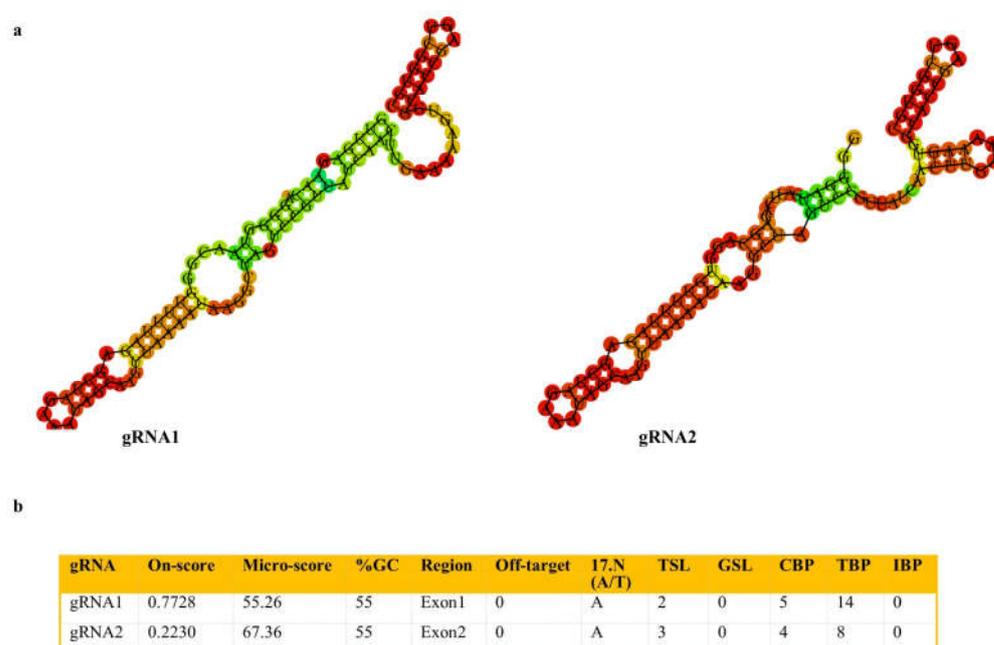
##### 4.5. In-silico Expression Analysis of Tomato SLACO genes

Expression patterns of SLACO genes were examined using transcriptome data of the developmental profile of tomato pollen cells obtained from the NCBI GEO dataset (GSE117185) [19]. Expression profiles of SLACO genes were visualized with TBtools using transcriptome data. Putative tissue-specific expression profiles of SLACO1 and SLACO2 genes were extracted based on the *Solanum lycopersicum* transcript expression database from several tissues and organs, including flowers, leaves, roots, and fruit from

different developmental stages. Expression profiles were built using the tomato plant Electronic Fluorescent Pictograph Browsers (Tomato eFP browsers) ([http://bar.utoronto.ca/eplant\\_tomato/](http://bar.utoronto.ca/eplant_tomato/)) accessed on 15 September 2022. We predicted the subcellular localization of SLACO proteins on tomatoes using an online TargetP - 2.0 server (<https://services.healthtech.dtu.dk/service.php?TargetP-2.0>) [40].

#### 4.6. Construction of Cas9/gRNA-Expressing Vectors

The gRNAs targeting SLACO2 was designed with the help of CRISPR-P V2.0 (<http://crispr.hzau.edu.cn/CRISPR2>) tools. The target site was selected based on criteria, including on-score value ((best score > 0.50) and moderate (0.20 < score < 0.50)), GC content (30-80%), number of complementary bases (TBP)  $\leq 12$  between gRNA and gRNA scaffold sequences and off-target capacity to efficiently disrupt gene function. Features of candidate gRNAs are provided (Figure 8). Each of the two synthesized gRNAs was cloned separately into the pKI1.1R (Addgene plasmid #85808) plasmid.



**Figure 8.** Some features and selection criteria of designed gRNAs targeting SLACO2. a) Minimum free energy structures of gRNAs and scaffold estimated by RNA Fold Server. b) Properties of designed gRNAs provided by Crispr-P-V2.0.

#### 4.7. Generation of transgenic CRISPR/Cas9 mutant lines in tomato

The Cas9/gRNA constructs harboring the gRNA1 and gRNA2 targeting the SLACO2 gene were transformed into *Agrobacterium tumefaciens* strain Gv3101 using electroporation separately. *Agrobacterium*-mediated transformation of tomato (*Solanum lycopersicum*) cultivar 'Crocker' were performed according to Secgin et al. (2021)[41]. All tomato lines were grown in standard glasshouse conditions of 16-h day length and 24-26°C, with a night temperature of 18°C. Plants from each line were grown in coarse potting compost (Klasmann Potgrond H) in 30L pots with standard Hoagland's solution irrigation treatment. Tomato fruits were harvested at commercial maturity and used for further analysis.

#### 4.8. Identification of transgenes and CRISPR/Cas9 mutation

Genomic DNAs from the leaves of transgenic tomato and WT plants were isolated using the Quick-DNA™ Plant/Seed Miniprep Kit (Zymo Research) to evaluate the mutation of target sites. The first PCR reaction was carried out to confirm the presence of the

transgene in T0 plants using HptII gene-specific primers (Table 2). Then, specific PCR primers were utilized to amplify the genomic regions spanning the target SIACO2 sites, which were subsequently examined to look for CRISPR/Cas9-induced mutations. Sanger sequencing was applied to find possible mutations in amplified PCR products. The obtained sequences were compared with the wild-type reference sequence of the SIACO2 using BLASTN NCBI. By using the Synthego ICE CRISPR Analysis Tool [42] and DSDecodeM [43], the mutation types and their rates were examined. In addition to the target gene, a potential off-target gene (Solyc01g014620.1) having four mismatches was also sequenced.

#### 4.9. Characterizing the morphology of mutant plants

The morphological and structural properties of T0 and WT mutant lines' flowers and pollens were studied. Triphenyl tetrazolium chloride (TTC) was used in an assay to measure pollen vitality. The slide was covered with a layer of pollen taken from various plant lines. The pollen was left exposed to TTC for two hours before being placed in the dark. Pollen was counted using a Nikon Eclipse E200 microscope equipped with 10x20 oculars and a 40/0.25 objective on samples stored in the dark. The TTC test classified undyed pollen as dead, partially active pollen as pink, and fully active pollen as red. Additionally, pollen grains and anthers of genome-edited and WT plants were visualized using scanning electron microscopy (JEOL JSM-7001F). Following this analysis, the structural makeup of flowers from various plant strains was analyzed. The physical characteristics of the flowers were compared, including their size, stigma, and stamen structure.

#### 4.10. Physiological analysis of genome-edited plants

For the functional analysis of the SIACO transgenic line against WT related to abiotic stress, T1 generation seeds of WT and transgenic lines were firstly surface sterilized using 5% sodium hypochlorite, as mentioned above. Surface sterilized seeds were then cultured on MS medium supplemented with either 200 mM NaCl to simulate salinity stress or 200 mM mannitol to simulate water deficit. Seedlings were then allowed to germinate in growth chambers set constant at  $24 \pm 2$  °C with a 16-h light and 8-h dark photoperiod at a light intensity of  $400 \mu\text{mol m}^{-2} \text{s}^{-1}$ . Treatment was carried on glass jars containing five seedlings each and three jars representing three biological replicates per treatment. Leaf samples from 10-week-old seedlings were selected randomly for physiological assay, including malondialdehyde (MDA) and proline contents, and immediately frozen in liquid nitrogen and then stored at -80 °C.

As previously described, the MDA level of transgenic lines and WT was determined using the thiobarbituric acid method [44]. In brief, 200 mg leaf samples of control, NaCl, and mannitol-treated seedlings were homogenized with 1ml 0.5% trichloroacetic acid (TCA). The total amount of ground homogenates was then centrifuged for 15 mins at 12,000g at 4 °C. A 0.6ml volume of supernatant was mixed with the same volume of 0.5% TBA containing 5% TCA in a single eppendorf tube. The mixture is then boiled for 25 mins at 96 °C. After cooling the mixture on ice to room temperature, the mixture is centrifuged for an additional 5 mins at 10,000g. Absorbance was then recorded at 532nm and 600nm using 0.5% TBA containing 5% TCA as blank.

Proline content was extracted and determined using the acid ninhydrin method as previously described [45]. 200 mg of frozen leaf samples were homogenized with 1ml 3% sulfosalicylic acid. The homogenized extract is then centrifuged for 5 mins at 14,000g at 4 °C. The supernatant (0.1ml), acid ninhydrin (0.2ml), 96% acetic acid (0.2ml), and 3% sulfosalicylic acid (0.1 ml) are mixed in a single eppendorf tube and incubated on 95 °C hot block for one hour. After the mixture is cooled down on the ice, 1 ml toluene is added and then centrifuged at 14,000g for 5min at 4 °C. The absorbance of the red supernatant was then read at 520 nm, and a preset calibration curve was used to record the proline content. All experiments were done with triplet replicates, and results were presented with standard deviations and standard errors computed using IBM SPSS 26 software. A one-way

ANOVA analysis was used to analyze the data, and a difference of  $P < 0.05$  (\*) was considered statistically significant, and  $P < 0.01$  (\*\*) was considered very significant compared with WT.

#### 4.11. Total RNA Isolation and qRT-PCR Analysis

qRT-PCR was performed to determine the gene expression of the *SlACO2* at different developmental stages of flowers and different tissues of mutant and WT tomato plants. Total RNA was isolated from various tissues of tomato plants using QIAGEN RNeasy Plant Mini Kit (QIAGEN, Valencia, California, USA) following the manufacturer's protocol. RNA concentration was quantified spectrophotometrically. According to the manufacturer's instructions, one  $\mu\text{g}$  of RNA was reverse transcribed to cDNA using the iScript<sup>TM</sup> cDNA Synthesis Kit (Biorad). All reactions were performed with a CFX96 (Biorad) following the protocol; 98°C for 30 s, followed by 40 cycles of 95°C for 15 s, and 60°C for 30 s. For each sample, three biological replicates and two technical replicates were analyzed. The actin one gene was used as an internal reference gene to obtain relative transcript abundances for SYBR green-based real-time PCR assay. The relative gene expression was calculated using the  $2^{-\Delta\Delta\text{CT}}$  method [46], and the primer sequences used for qRT-PCR are listed (Table 2). Data provided represent fold expression  $\text{Log}_2$  (WT/genome-edited lines). The one-way ANOVA was used to determine the statistical significance of changes in relative *SlACO2* expression levels. The means and standard deviation of the replications were compared by the least significant difference (LSD) test at the  $P \leq 0.01$ .

**Table 2.** List of primers used in this study.

Primer Name	Sequence 5'-3'	PCR product
sgRNA1-F	ATTGGTTGAGAACAGGGGTAACGG	
sgRNA1-R	AAACCCGTTACCCCTGTTCTCAAC	
sgRNA2-F	ATGGGGGATCATTGAGAGCAGGT	
sgRNA2-R	AAACACCTGCTCTGAATGATCCCC	
Acos Tomato Seq F	CAC TTCCGATCGCTGAGGT	
Acos Tomato Seq R	AAGCAGTGGAGCAAGTCATTT	
<i>hptII-F</i>	CGAAAAGTTCGACAGCGTC	450 bp
<i>hptII-R</i>	GGTGTCGTCCATCACAGTTTG	
<i>SlActin-F</i>	AGGCACACAGGTGTTATGGT	186 bp
<i>SlActin-R</i>	AGCAACTCGAAGCTCATTGT	
<i>SlACO2-qPCR-F</i>	TCGCTGAGGTGGAGATACGGTGT	84 bp
<i>SlACO2-qPCR-R</i>	AACAGGGGTAACGGCGGCAG	
M13 F seq primer (-20)	GTAAAACGACGGCCAGT	
M13 R seq primer (-20)	GTTTCCCAGTCACGAC	

## 5. Conclusions

In this study, we identified two aconitase genes in the tomato genome, which we think are associated with male sterility. In silico and in vitro expression analyses of these genes were performed. The *SlACO2* gene, which was determined to be highly expressed in different generative tissues, was knocked-out using the CRISPR/Cas9 system. Sequence analysis of 6 mutant lines obtained using gRNA1 was performed, and Lines 3 and 6 were determined to be homozygous mutants. In the morphological analysis, it was determined that there was no difference between mutant and WT plants. On the contrary, in physiological analysis, it was determined that the MDA content of mutant plants decreased

significantly compared to WT plants, and mutant plants were less affected by applied stresses.

**Supplementary Materials:** Figure S1: Sub-cellular location of SIACO proteins; Figure S2: Chromosomal location and duplication pattern of SIACOs; Figure S3: Phylogenetic analysis of ACO genes in different organisms; Figure S4: Alignment of aconitase swivel domains of AtACOs and SIACOs proteins; Figure S5: Expression profiles of SIACOs in various tissues and organs. a) Expression analysis of SIACOs using RNA-seq data b) Organ and tissue-specific expression of SIACOs; Figure S6: Protein-protein interaction of SIACOs; Figure S7: PCR confirmation of SIACO lines using *hptII*-specific primers; Figure S8: Scanning electron microscope images of a) pollen grains and b) anthers; Figure S9: Seed statics of WT and mutant plants. a) Seeds obtained from a) WT plants b) Line 3 c) Line 6 d) Seed numbers; Table S1: MCL Clusters of SIACOs with genes related to pollen development and male-sterility; Table S2: Predicted functional partners of SIACOs; Table S3: Changes in pollen viability in WT and SIACO2 mutant lines.

**Author Contributions:** Conceptualization, Z.S, M.K; methodology, K.M, Z.S, M.F.A; data curation, M.K., K.M, Z.S, M.F.A; writing original draft preparation, Z.S, M.F.A, S.U; writing—review and editing, Z.S, M.F.A, S.U, M.K, K.Y; funding acquisition, M.K. All authors have read and agreed to the published version of the manuscript.”

**Funding:** This research was funded by the Research Fund of Ondokuz Mayıs University, grant number PYO.ZRT.1901.17.010.

**Data Availability Statement:** Not applicable

**Acknowledgments:** We would like to thank Dr. Hiroki Tsutsui for providing the pKI1.1R plasmid.

**Conflicts of Interest:** The authors declare that they have no conflict of interest.

## References

- Li, R., et al., *CRISPR/Cas9-mediated mutagenesis of lncRNA1459 alters tomato fruit ripening*. The Plant Journal, 2018. **94**(3): p. 513-524.
- Refik, B., *Commonalities of Molecular Response in Tomato Plants against Parasitic Nematodes*. Biology Bulletin, 2021. **48**(2): p. S12-S21.
- Bozbuga, R., *Molecular analysis of nematode-responsive defence genes CRF1, WRKY45, and PR7 in Solanum lycopersicum tissues during the infection of plant-parasitic nematode species of the genus Meloidogyne*. Genome, 2022. **65**(5): p. 265-275.
- Zhang, D., et al., *Construction of a multicontrol sterility system for a maize male-sterile line and hybrid seed production based on the ZmMs7 gene encoding a PHD-finger transcription factor*. Plant Biotechnology Journal, 2018. **16**(2): p. 459-471.
- Bao, H., et al., *Gene silencing, knockout and over-expression of a transcription factor ABORTED MICROSPORES (SIAMS) strongly affects pollen viability in tomato (Solanum lycopersicum)*. BMC Genomics, 2022. **23**(1): p. 346.
- Carroll, D., *Genome Engineering With Zinc-Finger Nucleases*. Genetics, 2011. **188**(4): p. 773-782.
- Li, T., et al., *High-efficiency TALEN-based gene editing produces disease-resistant rice*. Nature Biotechnology, 2012. **30**(5): p. 390-392.
- Cong, L., et al., *Multiplex Genome Engineering Using CRISPR/Cas Systems*. Science, 2013. **339**(6121): p. 819-823.
- Wang, D., et al., *Characterization of CRISPR Mutants Targeting Genes Modulating Pectin Degradation in Ripening Tomato*. Plant Physiology, 2018. **179**(2): p. 544-557.
- Chen, G., et al., *Overexpression of iron regulatory protein 1 suppresses growth of tumor xenografts*. Carcinogenesis, 2007. **28**: p. 785-91.
- Moeder, W., et al., *Aconitase plays a role in regulating resistance to oxidative stress and cell death in Arabidopsis and Nicotiana benthamiana*. Plant Molecular Biology, 2007. **63**(2): p. 273-287.
- Heazlewood, J.L., et al., *Experimental Analysis of the Arabidopsis Mitochondrial Proteome Highlights Signaling and Regulatory Components, Provides Assessment of Targeting Prediction Programs, and Indicates Plant-Specific Mitochondrial Proteins [W]*. The Plant Cell, 2004. **16**(1): p. 241-256.
- Hooks, Mark A., et al., *Selective induction and subcellular distribution of ACONITASE 3 reveal the importance of cytosolic citrate metabolism during lipid mobilization in Arabidopsis*. Biochemical Journal, 2014. **463**(2): p. 309-317.
- Li, J., et al., *Identification and analysis of the mechanism underlying heat-inducible expression of rice aconitase 1*. Plant Science, 2015. **233**: p. 22-31.
- Du, K., et al., *Morphological Structure and Transcriptome Comparison of the Cytoplasmic Male Sterility Line in Brassica napus (SaNa-1A) Derived from Somatic Hybridization and Its Maintainer Line SaNa-1B*. Frontiers in Plant Science, 2016. **7**.
- Peyret, P., P. Perez, and M. Alric, *Structure, Genomic Organization, and Expression of the Arabidopsis thaliana Aconitase Gene: PLANT ACONITASE SHOW SIGNIFICANT HOMOLOGY WITH MAMMALIAN IRON-RESPONSIVE ELEMENT-BINDING PROTEIN (&#x2217;)*. Journal of Biological Chemistry, 1995. **270**(14): p. 8131-8137.
- Arnaud, N., et al., *The iron-responsive element (IRE)/iron-regulatory protein 1 (IRP1)-cytosolic aconitase iron-regulatory switch does not operate in plants*. Biochemical Journal, 2007. **405**(3): p. 523-531.

18. Li, M., et al., *Genome-wide characterization and analysis of bHLH transcription factors related to anthocyanin biosynthesis in spine grapes (Vitis davidii)*. Scientific Reports, 2021. **11**(1): p. 6863.
19. Liu, L., et al., *Transcriptomics analyses reveal the molecular roadmap and long non-coding RNA landscape of sperm cell lineage development*. The Plant Journal, 2018. **96**(2): p. 421-437.
20. Gökdemir, G., et al., *CRISPR/Cas9 knock-out of SIPHD\_MS1 (Solyc04g008420) gene results in complete male sterility in tomato*. Plant Growth Regulation, 2022. **98**(2): p. 329-341.
21. Jung, Y.J., et al., *Knockout of SIMS10 Gene (Solyc02g079810) Encoding bHLH Transcription Factor Using CRISPR/Cas9 System Confers Male Sterility Phenotype in Tomato*. Plants, 2020. **9**(9): p. 1189.
22. Du, M., et al., *A biotechnology-based male-sterility system for hybrid seed production in tomato*. The Plant Journal, 2020. **102**(5): p. 1090-1100.
23. Qi, X., et al., *Genome Editing Enables Next-Generation Hybrid Seed Production Technology*. Molecular Plant, 2020. **13**(9): p. 1262-1269.
24. Terol, J., et al., *The aconitate hydratase family from Citrus*. BMC Plant Biol, 2010. **10**: p. 222.
25. Paul, P., et al., *The membrane proteome of male gametophyte in Solanum lycopersicum*. Journal of Proteomics, 2016. **131**: p. 48-60.
26. Roschztardtz, H., et al., *A Nuclear Gene Encoding the Iron-Sulfur Subunit of Mitochondrial Complex II Is Regulated by B3 Domain Transcription Factors during Seed Development in Arabidopsis*. Plant Physiology, 2009. **150**(1): p. 84-95.
27. van der Merwe, M.J., et al., *Tricarboxylic Acid Cycle Activity Regulates Tomato Root Growth via Effects on Secondary Cell Wall Production*. Plant Physiology, 2010. **153**(2): p. 611-621.
28. Zhao, H., et al., *Mitochondrial citrate synthase plays important roles in anthocyanin synthesis in petunia*. Plant Science, 2021. **305**: p. 110835.
29. Araújo, W.L., et al., *Antisense Inhibition of the 2-Oxoglutarate Dehydrogenase Complex in Tomato Demonstrates Its Importance for Plant Respiration and during Leaf Senescence and Fruit Maturation*. The Plant Cell, 2012. **24**(6): p. 2328-2351.
30. Carrari, F., et al., *Reduced Expression of Aconitase Results in an Enhanced Rate of Photosynthesis and Marked Shifts in Carbon Partitioning in Illuminated Leaves of Wild Species Tomato*. Plant Physiology, 2003. **133**(3): p. 1322-1335.
31. Robinson, J.M. and J.A. Bunce, *Influence of Drought-Induced Water Stress on Soybean and Spinach Leaf Ascorbate-Dehydroascorbate Level and Redox Status*. International Journal of Plant Sciences, 2000. **161**(2): p. 271-279.
32. Liu, J., et al., *Exogenous melatonin improves seedling health index and drought tolerance in tomato*. Plant Growth Regulation, 2015. **77**(3): p. 317-326.
33. Yuan, G.-F., et al., *Effect of brassinosteroids on drought resistance and abscisic acid concentration in tomato under water stress*. Scientia Horticulturae, 2010. **126**(2): p. 103-108.
34. Li, R., et al., *Reduction of Tomato-Plant Chilling Tolerance by CRISPR-Cas9-Mediated SICBF1 Mutagenesis*. Journal of Agricultural and Food Chemistry, 2018. **66**(34): p. 9042-9051.
35. Wang, L., et al., *Reduced Drought Tolerance by CRISPR/Cas9-Mediated SIMAPK3 Mutagenesis in Tomato Plants*. Journal of Agricultural and Food Chemistry, 2017. **65**(39): p. 8674-8682.
36. Molinari, H.B.C., et al., *Osmotic adjustment in transgenic citrus rootstock Carrizo citrange (Citrus sinensis Osb. x Poncirus trifoliata L. Raf.) overproducing proline*. Plant Science, 2004. **167**(6): p. 1375-1381.
37. Claussen, W., *Proline as a measure of stress in tomato plants*. Plant Science, 2005. **168**(1): p. 241-248.
38. Kumar, S., et al., *MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms*. Mol Biol Evol, 2018. **35**(6): p. 1547-1549.
39. Chen, C., et al., *TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data*. Molecular Plant, 2020. **13**(8): p. 1194-1202.
40. Almagro Armenteros, J.J., et al., *SignalP 5.0 improves signal peptide predictions using deep neural networks*. Nature Biotechnology, 2019. **37**(4): p. 420-423.
41. Secgin, Z., M. Kavas, and K. Yildirim, *Optimization of Agrobacterium-mediated transformation and regeneration for CRISPR/Cas9 genome editing of commercial tomato cultivars*. Turkish Journal of Agriculture and Forestry, 2021. **45**(6): p. 704-716.
42. Hsiao, T., et al., *Inference of CRISPR Edits from Sanger Trace Data*. bioRxiv, 2019: p. 251082.
43. Liu, W., et al., *DSDDecode: a web-based tool for decoding of sequencing chromatograms for genotyping of targeted mutations*. Molecular plant, 2015. **8**(9): p. 1431-1433.
44. Hodges, D.M., et al., *Improving the thiobarbituric acid-reactive-substances assay for estimating lipid peroxidation in plant tissues containing anthocyanin and other interfering compounds*. Planta, 1999. **207**(4): p. 604-611.
45. Bates, L.S., R.P. Waldren, and I.D. Teare, *Rapid determination of free proline for water-stress studies*. Plant and Soil, 1973. **39**(1): p. 205-207.
46. Livak, K.J. and T.D. Schmittgen, *Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT Method*. Methods, 2001. **25**(4): p. 402-408.