# Characterization of the *SlHSP70-1* Functions by Gene-to-gene Co-expression Analysis to Uncover Relationships of *SlHSP70-1*, *SlIAA9* and *SlDELLA* in the Tomato Correlation Networks

**July 2019**

**VU TUAN NAM**

# Characterization of the *SlHSP70-1* Functions by Gene-to-gene Co-expression Analysis to Uncover Relationships of *SlHSP70-1*, *SlIAA9* and *SlDELLA* in the Tomato Correlation Networks

**A Dissertation Submitted to**

**the Graduate School of Life and Environmental Sciences,**

**the University of Tsukuba**

**in Partial Fulfillment of the Requirements**

**for the Degree of Doctor of Philosophy in Agricultural Science**

**(Doctoral Program in Biosphere Resource Science and Technology)**

**VU TUAN NAM**

## Table of Contents

## Abbreviations

| | |
|---|---|
| **HSP70** | Heat shock protein 70 |
| **OE** | Over Expression |
| **Az** | Azygous |
| **WT** | Wild type |
| **MS medium** | Murashige and Skoog medium |
| **UBQ** | Ubiquitin |
| **AA** | Amino acid |
| **TILLING** | Targeting Induced Local Lesions IN Genome |
| **IAA** | Indole-3-acetic acid |
| **NBRP** | National BioResource Project |
| **EMS** | Ethyl Methane Sulfonate |
| **CDS** | Coding DNA Sequence |
| **mW** | Molecular weight |
| **pI** | Isoelectric point |
| **(kDa)** | Kilo Dalton |
| | |
| | |
| | |
| | |
| | |

## List of Tables

## List of Figures

# Chapter I. General introduction

## 1.1. General introduction

### *1.1.1. Role of tomato as a vegetable and a model plant*

Tomato (*Solanum lycopersicum*) is a valuable vegetable plant of the *Solanaceae* family; it is a good source of nutrients because it is rich in vitamins and antioxidants. In addition, its specific flavor makes it a valuable material for the food processing industry (Meissner et al., 1997; Giovannucci, 1999; Chalabi et al., 2004; Etminan et al., 2004; Jian et al., 2005; Consortium, 2012). From a scientific viewpoint, *S. lycopersicum* is a model organism for studying plant molecular biology as it has a small genome, short life cycle, small body size, and prolific seed production ability. The tomato genome is sequenced as 12 chromosomes with a size of approximately 950 Mb coding for agricultural traits, DNA markers, a large collection of mutants and expressed sequence tags (ESTs) (Tanksley et al., 1992; Menda et al., 2004; Yamamoto et al., 2005). Most recent studies used Micro-Tom, a tomato cultivar that contains most characteristics of a typical model of tomato (Meissner et al., 1997). Tomatoes represent fleshy fruit plants with a complete set of characteristics (from flowering to ripening) that cannot be studied in other model systems such as *Arabidopsis* or rice (Bertin, 2005; Mintz-Oron et al., 2008). Owing to the large number of species of this family (1000–2000 species), a more detailed understanding of the principle and dynamics of molecular plant physiology can be achieved to help create new valuable traits for vegetable and fruit plants (Aoki et al., 2013).

### *1.1.2. Improving tomato growth, development and fruit production*

Tomatoes are one of the most valuable vegetables that are cultivated and consumed widely in the world. Several cultivation methods have been applied to enhance the growth, development, and fruit production of tomatoes. Based on the physiological characterizations of tomatoes, optimizing farming conditions for tomatoes might be the first alternative. A temperature range of 22–25°C was considered as being optimal temperature for plant growth and development, as well as fruit set. However, temperatures below 18°C or above 30°C were considered suboptimal as they had adverse effects on both the plant growth rate and productivity (Mulholland et al., 2003; Bertin, 2005). Greenhouse technology has been

developed to optimize most conditions required for the stable growth of tomatoes, including temperature, lighting duration and density, nutrients, and $CO_2$ application (Vanthoor et al., 2011).

Agronomic methods such as adding fertilizers and optimizing the growth environment for tomatoes have reached their limits. However, several choices as the application of exogenous hormones have been used effectively to improve fruit formation and development of tomatoes (Delph et al., 1997; Gorguet et al., 2005). Phytohormones such as auxin, gibberellin, and cytokinin have often been used for improving the fruit set. Auxins have been reported to have a strong effect on cell division, while gibberellin contributes to cell expansion. A conventional method of combinational application of auxin and gibberellin can produce fruits of the same shape and size as those of the pollinated fruits. This shows that the cell division and cell expansion of fruits are controlled by interactions between these hormones (de Jong et al., 2009a). Cytokinin was less commonly used as compared to the use of auxin and gibberellin. However, it has been reported that cytokinin plays an important role in the growth and development of tomatoes (Matsuo et al., 2012). The application of exogenous hormones during fruit formation may result in parthenocarpy—a phenomenon of the development of ovary to a fruit without pollination and/or fertilization resulting in a seedless fruit (Spena and Rotino, 2001; de Jong et al., 2009a). This approach of crop production is not only less dependent on environmental conditions but also a favorable trait of tomato fruit, which was a focus of study in the recent years (Martinelli et al., 2009; Pandolfini, 2009). Individual or combinational applications of auxin and gibberellin had positive effects on parthenocarpic fruit development. In addition, the application of exogenous cytokinins produced similar effects (Matsuo et al., 2012; Ding et al., 2013).

Advancement in genetic technology has discovered genes that promotes the growth, development, and formation of fruits by involving in cell division and expansion or hormone signaling. Any increase or decrease in the expression of CELL CYCLE SWITCH 52 (CCS52A) causing a reduction in the size of the pericarp cells that results in the reduction of fruit size (Mathieu-Rivet et al., 2010a; Mathieu-Rivet et al., 2010b). In addition, a fruit weight 2.2 gene *(fw2.2*) has been reported to be involved in cell cycle signals. Low levels of mRNA in *fw2.2* resulted in larger fruit sizes, suggesting that this gene can participate in complexes inhibiting cell division (Cong and Tanksley, 2006). The analysis of hormone-related genes such as auxin and gibberellin revealed their role in the growth and development

of tomatoes. *IAA9*—a member of the *IAA* gene family —is considered to inhibit the transcription of auxin-responsive genes. The downregulation of the *IAA9* expression level was shown to stimulate cell division, which resulted in an increase in the size of tomatoes; this was attributed to the interaction of *IAA9* with auxin response factors (ARF) genes (Wang et al., 2005; de Jong et al., 2009b). In tomatoes, two genes of the *ARF* family, *SlARF7* and *SlARF8*, play a role in the initiation and the development of parthenocarpic fruit (Goetz et al., 2007; de Jong et al., 2009b). In addition, many genes associated with the synthesis of gibberellins (GAs) were studied. Two genes—*GA1* and *GA4*—associated with gibberellin biosynthesis were investigated for their role in the growth of tomatoes; *GA1* efficiently stimulated fruit initiation, whereas *GA4* promoted seed germination (Serrani et al., 2007; Nakaune et al., 2012). Further, it was shown that GA biosynthesis is closely related to the regulation of GA 20-oxidase (*GA20ox*) genes (Serrani et al., 2007). The downregulation of *GA20ox1* resulted in short plants, reduced pollen survival, and deformed leaves. Other *GA20ox* family members such as *GA20ox2* and *GA20ox3* were shown to play a major role in the ovaries and fruits. However, overlapping functions of *GA20ox1*, *GA20ox2*, or *GA20ox3* were reported, so that required silencing of all three *GA20ox* genes to evaluate their effect on fruit initiation and development (Xiao et al., 2006). Other studies showed that the mutation of *gib1*, *gib2*, and *gib3* genes caused a deficiency in GA resulting in the inhibition of fruit formation. This defect was overcome when the exogenous GA was applied, indicating that *GIB1, GIB2* and *GIB3* participated in the process of endogenous GA synthesis (Bensen and Zeevaart, 1990).

### *1.1.3. Role of SlIAA9 and SlDELLA genes in tomato growth, development and fruit set*

*SlIAA9* and *SlDELLA* are two genes that play a key role in the plant growth, development, and fruit production of tomato (Wang et al., 2005; de Jong et al., 2009a). *SlIAA9* is a transcriptional factor that controls the auxin-related genes involved in the plant growth, leaf structure, and formation and development of fruits. A decrease in the *SlIAA9* expression makes the auxin response genes more sensitive to auxin signals, which leads to different phenotypes resulting in the development of fruits (Wang et al., 2005). The *SlIAA9* expression conserves domains interacting with the specific regions of ARF genes, which control the activity of ARF dependent genes. It was shown that, in tomatoes, *SlARF7* and *SlIAA9* helped regulate the fruit set; while the homolog of *SlARF7* gene in *Arabidopsis* was *AtARF8* (Goetz et al., 2007; de Jong et al., 2009b). Downregulation of *SlIAA9* or *SIARF7*

led to the development of parthenocarpic fruits in tomatoes, indicating that both *SIIAA9* and *SlARF7* were negative transcription factors for the tomato fruit set. *SlARF8* sometimes works as a partner of *SlIAA9*, regulated by the *SlTIR1* (*Transport Inhibitor Response 1*) gene. In the absence of *SlIAA9*, the initiation of fruit set may be activated by the stimulation of *SlARF8* to auxin response genes. *TIR1* is an auxin receptor that plays important functions in fruit set initiation through positive regulatory mechanisms (El-Sharkawy et al., 2016). The over expression of *SlTIR1* was found to degrade the Aux/IAA protein *SlIAA9*, exhibit altered leaf morphology, and parthenocarpy, in a study on tomatoes. The simplification of leaf complexity, increase in leaf length and width, petiole diameter, and the promotion of parthenocarpy were caused by the overexpression of *SlTIR1*, which was attributed to the downregulation of *SlIAA9* (Ren et al., 2011).

Although IAA9 is a transcription factor that plays a role in controlling the activity of ancillary signalling genes, DELLA was considered a growth repressor that binds to transcription factors to control plant growth related to the gibberellin signaling response with their DNA binding domain (Harberd, 2003). *DELLAs* gene are members of the *GRAS* regulatory protein family (Bolle, 2004) that consist of five known-*DELLA* genes in *Arabidopsis*; however, it has only one single gene in tomatoes (Marti et al., 2007). In *Arabidopsis*, *AtDELLA* proteins repressed the growth of stigma, styles, and gynoecium (Fuentes et al., 2012). The downregulation of *AtDELLA* induced parthenocarpy, petiole elongation, and hypocotyl development, and it increased the light sensitivity of *Arabidopsis* plant (Djakovic-Petrovic et al., 2007). The *pro* mutant of *SlDELLA* in tomato induced internode elongation and promoted parthenocarpic fruit development (Martí et al., 2007; (Livne et al., 2015). *SlDELLA* could interact with *SlARF7* in specific regions to regulate fruit development. *SlDELLA* and *SlARF7*/*SlIAA9* built a complex mediated crosstalk between GA and auxin signaling during tomato fruit initiation (Hu et al., 2018).

### *1.1.4. Gene co-expression analysis for tomato growth and development*

The genome sequences of tomatoes and some other model organisms were fully decoded with a complete list of genes and their corresponding IDs. It was reported that the 125-megabase (Mb) genome of *Arabidopsis* contains 25,498 genes, whereas the 950 Mb genome of tomatoes contains 34,727 genes distributed in 12 chromosomes (Cao et al., 2011; Consortium, 2012). There was great challenge to elucidate the genes functions because many genes were still not fully uncovered their role during the life of plant. Rapid advancement in

biotechnology with the introduction of microarray and RNA sequencing has provided valuable information from thousands of single or multiple experimental conditions. The use of gene expression data for biological studies has become popular (Horvath et al., 2003; Wang et al., 2009; Schadt et al., 2010). In addition to evaluating the expression of a large number of genes, the analysis of the microarray data has helped discover the functions of these genes. Microarray data obtained from different studies were collected and made available in a large public database. This database was useful for gene co-expression analysis to predict and identify gene functions (Saito et al., 2008). There are two approaches for co-expression network analysis (Aoki et al., 2007). The first method uses a pre-selected gene as the core part of the network, which is comprehended in some biological processes or pathways, and other genes that appear in the network and are related to the specific processes or pathways involving the core gene. In contrast, another approach is to construct a network of all those genes using all the available data of genes. This analysis does not target a specific goal to analyze a wide range of genomes and important connections that will be extracted from that overall network (Ma et al., 2007). Co-expression analysis can be used to identify potential candidate genes associated with plant growth and development (Piya et al., 2014; Harris Amrine et al., 2015; Li et al., 2015; Lin et al., 2017). For example, gene co-expression network analysis was utilized to discover genes in the enriched co-expression module(s) in the pathway of the flavonoid biosynthetic and modules of metabolites during fruit ripening (Kurabayashi et al., 2010; DiLeo et al., 2011). Further, the creation of the Tomato Functional Genomics Database (TFGD) provided not only a fundamental database for functional genomics research but also a huge resource of microarray, metabolite, and sRNA data sets for co-expression analysis based on computational applications (Fei et al., 2011).

### 1.2. Hypothesis of this study

Gene co-expression networks could be utilized to analyze a large DNA microarray or RNA sequencing dataset for several purposes such as discovering new candidate genes with specific function in a biological process, functional annotation, and identification of regulating elements (van Dam et al., 2017). Based on the correlation between genes, gene co-expression analysis can be used to determine the role of a large number of genes under the same biological conditions. Depending on different objectives, many internet-based-packages were built to analyze the gene co-expression (Zhang and Horvath, 2005; Langfelder and Horvath, 2008; Liu et al., 2010). In tomatoes, the microarray data of *IAA9*

and *DELLA* was established and published; the roles of auxin and gibberellin in the growth and development of phytohormone signaling were also explained, respectively (Wang et al., 2005; de Jong et al., 2009a). However, *IAA9* and *DELLA* genes were only concentrated in the process of fruit initiation as they improve production and yield. Other genes controlling tomato growth and development in connection with *IAA9* and *DELLA* were not identified. Thus, in this study, MRNet package was applied to analyze the co-expression network of *SlIAA9* and *SlDELLA* from a public microarray dataset. To find out the downstream of *SlIAA9* and *SlDELLA*, networks genes around *SlIAA9* and *SlDELLA* were built using publicly available microarray data to extract genes directly connected to nodes *SlIAA9* and *SlDELLA*, respectively. A potential candidate gene can be extracted from the genes directly connected to *SlIAA9* and *SlDELLA*. The gene *Solyc06g076020.2.1* was selected as the candidate gene as it appeared in both networks of *SlIAA9* and *SlDELLA.*

To obtain insights involved in tomato growth, development, and fruit set of candidate genes, in this study, we attempted to clarify the function of the targeted gene to tomato through popular genetic approaches. In one approach, we attempted to upregulate the target gene by transferring one copy of this gene into the tomato Micro-Tom wild type through an *Agrobacterium* system. On another approach, we applied reverse genetics to find the function of the target gene by phenotyping its mutant screened from the mutation library of National BioResource Project (NBRP) Tomato by the TILLING (Targeting Induced Local Lesions in Genomes) technique. The overexpression (OE) and mutant lines would then be cultivated under a no stress condition for phenotyping. The combination of phenotypic analysis from the OE line and mutant lines of the targeted gene would give us insight into the *HSP70* functions in tomato growth and development.

### 1.3. The objective of this study

The objectives of this study were investigating the function of the candidate gene in tomato growth, development, and fruit set.

(1) We constructed a network of *SlIAA9* and *SlDELLA* to extract the candidate gene *SlHSP70*.

(2) We characterized the family gene descriptions based on their genomic sequences, coding DNA sequences (CDS), and amino acids sequences.

(3) We then attempted to determine the effects of the targeted gene in the case where the gene was overexpressed by transforming one copy of this gene into tomato host plant.

(4) Finally, we characterized the phenotypes of the *HSP70* mutant from the Mutation Library of the National BioResource Project (NBRP) Tomato by the applied TILLING technique.

## Chapter 2. Gene to gene co-expression to discover genes directed connection to *SlIAA9* and *SlDELLA*

### 2.1. Introduction

#### *2.1.1. Role of SlIAA9 and SlDELLA in tomato growth and development including fruit set*

Tomato was a highly economical vegetable which has cultivated and consumed widely used in the world. In tomato cultivation, several plant hormones such as auxin and gibberellin were applied to tomato to increase yield and production. One of the most striking effects of the exogenous hormone applying was the stimulation of the formation and development of parthenocarpy, a desired trait of tomato that ovary can develop fruit, come over the pollination and fertilization. Relating to phytohormone signalling, *Aux/INDOLE - 3 - ACETIC ACID 9* (*SlIAA9*) and *SlDELLA* were important genes for plant growth and development through cell division and cell expansion. Tomato *Aux/IAA9* acted as a negative regulator of the auxin response involved in controlling of fruit set which repressing the transcription of the auxin responsive and fruit developmental genes through the interaction with *Auxin response factor 7* (*ARF7*) (de Jong et al., 2009b). Downregulation in the transcript level of *SlIAA9* could simplify leaf shape and elongate shoot parts, leading change in plant height. The downregulation also induced parthenocarpy (Wang et al., 2005; Okabe et al., 2011; Mazzucato et al., 2015). Meanwhile, *SlDELLA* was a negative regulator of gibberellin signaling through the combination with gibberellin receptor GID1 (Yoshida et al., 2014). DELLA has been suggested to function as a transcriptional activator and the fundamental components of the gibberellin - GID1-DELLA signalling pathway. *Procera,* a *Sldella* mutant, showed morphological changes in plant elongation, branching architect and reproductive organ development and promoted parthenocarpy (Marti et al., 2007; Bassel et al., 2008; Carrera et al., 2012). The parthenocarpy development in both *Sliaa9* and *Sldella* mutants has elucidated by their downstream activities that stimulates phytohormone-related fruit development through their direct or indirect crosstalk (Hu et al., 2018). Changes of plant architecture in these mutants were investigated. However, these mechanisms remained unknown.

### *2.1.2. Introduction of the co-expression analysis in gene functions discovery and investigation of tomato*

Tomato was considered a great model tree for fleshy climacteric fruit studies. The whole tomato genome had sequenced thank to the advances in biology sequencing technology but the functions of many genes were still unknown. However, the functional status of a gene can now be identified due to the development of high-throughput technologies such as microarrays and RNA sequencing (RNA-seq). The Sol Genomics Network (SGN) was a huge information network that gathers a genetic database of plants belonging to the *Solanaceae* family (Fernandez-Pozo et al., 2015). Tomato Functional Genomics Database (TFGD) provided not only fundamental database for functional genomics research but also a huge resource of microarray, metabolite and sRNA data sets for co-expression analysis based on computational application (Fei et al., 2011). The Gene Expression Omnibus (GEO) database of the National Centre for Biotechnology Information (NCBI) which obtained available GeneChip data from individual experiments provided a valuable resource of publicly materials for building new hypothesis and knowledge of gene functions based on gene co-expression network analysis (Toro-Domínguez et al., 2018). Co-expression gene networks could support predicting and improving the understanding of function of potential candidate genes in a biological process or pathway (Aoki et al., 2007). In the gene co-expression analysis, the differences in genes expression tendency including unknown functions gene among analytical samples might suggest the role of unknown genes in specific biological pathways. For instance, it is possible to suggest the role of candidate genes as a gene regulator if this gene appears in a transcription pathway. In co-expression analysis, gene expression views can help clearly present the tendency of differential gene expression between samples. Consequently, co-expression networks with expression views can be used to associate genes of unknown function with biological processes, to discern gene transcriptional regulatory mechanisms to prioritize candidate regulatory genes. The increase in the number of reports on co-expression gene networks indicated the power of this tool in predicting and annotating gene function (D'Haeseleer et al., 2000; Aoki et al., 2007; Usadel et al., 2009; Morenorisueno et al., 2010; Li et al., 2015; Serin et al., 2016).

### *2.1.3. Objective of this study*

In this study, we applied the gene-to-gene co-expression network analysis based on the publicly available microarray data. Potential candidate genes which were directly connected with *SlIAA9* and *SlDELLA* were investigated. The relationship distance level

between potential targeted gene with *SlIAA9* and *SlDELLA* may suggest the potential functions of candidate gene for the tomato growth and developmen p t in relation with *SlIAA9* and *SlDELLA* genes.

## 2.2. Materials and methods

### *2.2.1. Construction of gene co-expression networks*

Affymetrix 307 GeneChips data obtained from GEO, ArrayExpress, and TFGD was used (Fukushima et al., 2012). To make networks, a Comprehensive R Archive Network (CRAN ver. 3.5.1) was used. The method of normalization and probe sets removal was conducted as in Fukushima et al, (2012). For the ID conversion of Affymetrix microarray's probeset-ID and ITAG ID, we used information in Sol Genomics Network. (ftp://ftp.solgenomics.net/genomes/Solanum_lycopersicum/micro-arrays_mapping/A-AFFY-87_AffyGeneChipTomatoGenome.compositeelements-_ITAG2.3-GeneID_mapping.txt). From this conversion, IDs with one-to-one correspondence were extracted, and a total of 5228 genes were used for network construction. The mrnet function package was used for the construction and the threshold was set to 0.05 (Meyer et al., 2008). MRNet generated a network using a feature selection method called minimum Redundancy Maximum Relevance (mRMR).

### *2.2.2. Drawing network diagram, GO analysis and distance measurement*

Cytoscape ver. 3.7.0 was used for network diagram. BiNGO ver. 3.0.3, an application of Cytoscape, was used for GO analysis of the neighboring genes (Maere et al., 2005). Pesca ver. 3.0, a Cytoscape application, was used to measure the distances between each gene in the network. Product name of each gene was acquired using Panther (http://pantherdb.org/) (Thomas et al., 2003; Muruganujan et al., 2012).

## 2.3. Results

### *2.3.1. Gene-to-gene correlation networks using publicly available microarray data*

*SlIAA9* and *SlDELLA* are two key genes for the auxin and gibberellin signalling. These genes play important roles in the growth and development of tomato (Wang et al., 2005; de Jong et al., 2009a). A co-expression network has been set up from 5228 genes, in which, we focused on the sub network where two key genes SlIAA9 and SlDELLA played as core gene. To find out the up- or downstream of *SlIAA9* and *SlDELLA*, networks around *SlIAA9* and *SlDELLA* were cut out from each network constructed by public available microarray data (Figure 2.1, Figure 2.2). The gene-to-gene co-expression network comprised 26 genes were

directly connected to *SlIAA9* as a main hub (Figure 2.1), while the network of *SlDELLA* had direct connection with seven different genes (Figure 2. 2).

### *2.3.2. GO enrichment analysis of neighbouring genes of SlIAA9 and SlDELLA*

We investigated neighbouring genes of *SlIAA9* and *SlDELLA*, respectively. Genes directly connected with *SlIAA9* and *SlDELLA* in the co-expression networks were listed in Table 2.1 and Table 2.2, respectively. Among 26 neighbouring genes in the *SlIAA9* network, nine genes have characterized by their functions. The most prominent one is the gene coding *SlAGAMOUS-LIKE 11* (*AGL11*) transcription factor, also known as a member of *MADS* box transcription factors that plays an important role in the process of plant growth and development, particularly for timing of flowering and fruit development (Becker and Theißen, 2003; Smaczniak et al., 2012; Puranik et al., 2014). In tomato, overexpressing *SlAGL11* has revealed abnormal stamens with poorly viable pollen (Daminato et al., 2014). The gene coding for *Ubiquitin-conjugating enzyme E2 8* (*UBC8*) was reported participating to protein modification process in the pathway of protein ubiquitination (Kraft et al., 2005), while other two genes involved in sugar metabolism in tomato (Wong et al., 1990; Cai et al., 2018).

On the other hand, analysis of neighbouring genes in the *SlDELLA* network exhibited seven genes that were directly connected to *SlDELLA* node. The gene *SAMDC* coding for *S*-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50) is vital for biosynthesis of polyamines in the *S*-adenosylmethionine biosynthesis pathway (Majumdar et al., 2017). The gene coding constitutive photomorphogenesis 9 (COP9) signalosome complex subunit 4, a component of the COP9 signalosome complex, is involved in various cellular and developmental processes relating to phytohormone auxin responses (Wang et al., 2003). The gene *Solyc06g076020.2.1* that was annotated as heat shock protein 70 *SlHSP70-1* was appeared the direct connection with *SlDELLA*. In general, genes of *HSP70* family often express in the response to stress growth conditions such as heat or drought stresses (Zhang et al., 2015b). The directed connection between *SlHSP70-1* and *SlDELLA* gave a new hypothesis that *SlHSP70-1* might be involved in tomato plant growth and development in a relationship with *SlDELLA*.

Next, we investigated distance between the targeted *SlHSP70-1* and *SlIAA9* and between that and *SlDELLA*, respectively. *SlDELLA* was directly connected to the *SlHSP70-*

*1* gene, while distance from *SlIAA9* to *SlHSP70-1* were three hops (Table 2.3). Also, the distance from *SlIAA9* to *SlDELLA* were three hops. As average of gene-to-gene distance in the network was 2.615 in this network, distance between *SlHSP70-1* and *SlDELLA* was greater than the connectivity of *SlHSP70-1* and *SlIAA9* and *SlIAA9* and *SlHSP70-1*. Besides, the overlaying the *SlLAA9* and *SlDELLA* networks could extract the *SlHSP70-1* that would have potential because this was the only gene to show direct connection with *SlDELLA* within other SlHSP70 on the microarray chip (Table 2.3). We thus focused on the *SlHSP70-1* gene for further analysis.

The results of gene ontology (GO) enrichment analysis showed the over-presented GO terms of linked to *SlIAA9* neighbouring genes were involved in various functions (Table 2.4). In the *SlDELLA* neighbouring genes, GO terms such as protein binding and proteasome complex were over-represented (Table 2.5).

**2.4. Discussion**

Gene co-expression networks can be utilized to analyse a big dataset of DNA microarray or RNA sequencing for abundant purposes such as discovering new candidate genes that have specific functions in a biological process, functional annotation and identify the regulating elements (van Dam et al., 2017). Therefore, gene co-expression networks are often used to clarify individual objectives for various plant species (Wisecaver et al., 2017; Tai et al., 2018; Rao et al., 2019). Depending on different objective, many internet-based packages were built to analyse the gene co-expression data (Langfelder and Horvath, 2008; Liu et al., 2010).

In this study, minet package was applied to construct the co-expression network of *SlIAA9* and *SlDELLA* from public microarray dataset. MRNet utilized a method of maximum relevance/minimum redundancy feature selection to decide candidate genes based on its highly relevant to selection criterion (Meyer et al., 2007). GO enrichment analysis showed that genes connecting to *SlIAA9* and to *SlDELLA* have GO terms related to plant growth and development functions in both networks. The greater number of correlated genes with *SlIAA9* in the *SlIAA9* network than that of the *SlDELLA* indicated that *SlIAA9* may participate in more biological processes at the transcript levels than *SlDELLA*. As the complex *SlIAA9* network contained many uncharacterized genes, it made difficulty to find out which candidate gene(s) should be chosen for further analysis (Figure 1). On the other hand, association between the targeted *SlHSP70-1* and *SlDELLA* with directed connection

gave us an idea that *SlHSP70-1* might have tight correlation relationships with *SlDELLA* that act as a hub in the *SlDELLA* network (Figure 2). The integrated analysis of the networks of *SlIAA9* and *SlDELLA* could highlight the *SlHSP70-1* as a potential target for further analysis. Because the targeted the *SlHPS70* gene was directly connected with *SlDELLA*. As we mentioned above, the *SlIAA9* network showed complex connections. The integrated network approach described here has a possibility to find candidate genes that may act as key genes with phytohormones in tomato (He and Maslov, 2016; Serin et al., 2016; Obayashi et al., 2017).

**Table 2. 1. List of genes in the nearest neighbouring gene group of *SlIAA9***

| Mapped ID | Gene Name / Gene Symbol | PANTHER Family/Subfamily | PANTHER Protein Class |
|---|---|---|---|
| **Solyc12g056840** | Uncharacterized protein | Phosphopantothenate--cysteine ligase (pthr12290: sf2) | Ligase (PC00142) |
| **Solyc03g019730** | Uncharacterized protein | Sumo-activating enzyme subunit 1 (pthr10953: sf162) | Ligase (PC00142)<br>transfer/carrier protein (PC00219) |
| **Solyc03g098730** | Uncharacterized protein | Cysteine protease inhibitor wscp-related (pthr33107: sf11) | protease inhibitor (PC00191) |
| **Solyc10g084920** | PRA1 family protein | Pra1 family protein (pthr12859: sf0) | Amino acid transporter (PC00046) |
| **Solyc11g011910** | Transmembrane 9 superfamily member | Transmembrane 9 superfamily member-related (pthr10766: sf14) | Transporter (PC00227) |
| **Solyc09g075000** | Uncharacterized protein | Wd repeat-containing protein 89 (pthr22889: sf0) | |
| **Solyc11g017300** | Uncharacterized protein | Cop9 signalosome complex subunit 5 (pthr10410: sf6) | Metalloprotease (PC00153) |
| **Solyc09g083150** | Uncharacterized protein | Nad(p)h-quinone oxidoreductase subunit n | Chloroplastic (PTHR35515:SF1) |
| **Solyc04g077970** | Uncharacterized protein | Adenine phosphoribosyltransferase 1 | Chloroplastic (PTHR11776:SF8) |
| **Solyc02g062680** | Anaphase-promoting complex subunit 10 | Anaphase-promoting complex subunit 10 (pthr12936: sf0) | Enzyme modulator (PC00095)<br>Ligase (PC00142) |
| **Solyc05g018410** | Uncharacterized protein | 3-hydroxyisobutyryl-coa hydrolase-like protein 3 | Mitochondrial-related (PTHR43176:SF5)<br>Acetyltransferase (PC00038)<br>Acyltransferase (PC00042)<br>Dehydrogenase (PC00092)<br>Epimerase/Racemase (PC00096)<br>Hydratase (PC00120)<br>Ligase (PC00142) |
| **Solyc03g097750** | Uncharacterized protein | Translocon-associated protein subunit beta (pthr12861: sf3) | |

| | | | |
|---|---|---|---|
| **Solyc04g081770** | Uncharacterized protein | Gdsl esterase/lipase exl1-related (pthr45642: sf25) | |
| **Solyc01g005470** | Uncharacterized protein | Plac8-like protein 1 (pthr15907: sf21) | |
| **Solyc04g078850** | Protein DCL | Ortholog | DCL Protein (DUF3223) (PTHR33415:SF2) |
| **Solyc02g085500** | Uncharacterized protein | Transcription repressor ofp10-related (pthr33057: sf68) | |
| **Solyc02g030300** | Uncharacterized protein | Subfamily not named (pthr27002: sf359) | |
| **Solyc05g049950** | Small nuclear ribonucleoprotein-associated protein | Small nuclear ribonucleoprotein-associated protein b' (pthr10701: sf0) | mRNA splicing factor (PC00148) |
| **Solyc04g076850** | Auxin-responsive protein | Subfamily not named (pthr31734: sf18) | |
| **Solyc08g028690** | Uncharacterized protein | Subfamily not named (pthr43180: sf1) | |
| **Solyc05g009390** | Uncharacterized protein | Alpha/beta-hydrolases superfamily protein (pthr11614: sf94) | Phospholipase (PC00186)<br>serine protease (PC00203) |
| **Solyc02g081160** | Pyrophosphate--fructose 6-phosphate 1-phosphotransferase subunit beta | Pyrophosphate--fructose 6-phosphate 1-phosphotransferase subunit beta 1-related (pthr43650: sf1) | carbohydrate kinase (PC00065) |
| **Solyc10g083570** | Fructose-bisphosphate aldolase | Subfamily not named (pthr11627: sf41) | |
| **Solyc11g028020** | TAGL11 transcription factor | Agamous-like mads-box protein agl11 (pthr11945: sf170) | MADS box transcription factor (PC00250) |
| **Solyc12g056100** | UBC8 | Ubiquitin-conjugating enzyme e2 29-related (pthr24068: sf78) | |
| **Solyc06g064840** | Uncharacterized protein | Agamous-like mads-box protein agl11 (pthr11945: sf170) | MADS box transcription factor (PC00250) |
| **Solyc07g041970** | Uncharacterized protein | Subtilisin-like protease sbt1.8 (pthr10795: sf335) | protease inhibitor (PC00191)<br>serine protease (PC00203) |

**Table 2. 2. Nearest neighbour gene group of *SlDELLA* gene**

| Mapped ID | Gene Name / Gene Symbol | PANTHER Family/Subfamily |
|---|---|---|
| **Solyc03g006820** | Uncharacterized protein | Fi16820p1-related (PTHR10869:SF123) |
| **Solyc09g014280** | Uncharacterized protein | Subfamily not named (PTHR31896:SF5) |
| **Solyc05g010420** | S-adenosylmethionine decarboxylase proenzyme | S-Adenosylmethionine Decarboxylase Proenzyme 3 (PTHR11570:SF15) |
| **Solyc11g011260** | DELLA protein GAI | DELLA protein RGL1-related (PTHR31636:SF47) |
| **Solyc03g111330** | Uncharacterized protein | Subfamily not named (PTHR47525:SF1) |
| **Solyc06g076020 (*SlHSP70-1*)** | Uncharacterized protein | Subfamily not named (PTHR19375:SF255) |
| **Solyc04g080160** | Uncharacterized protein | COP9 signalosome complex subunit 4 (PTHR10855:SF2) |

**Table 2. 3. The distance from *SlHSP70*-1 to *SlIAA9* and *SlDELLA* in their co-expression network**

| Gene ID | Gene name | Shortest distance to *SlIAA9* (Solyc04g076850.2) | Shortest distance to *SlDELLA* (Solyc11g011260.1) |
|---|---|---|---|
| **Solyc10g086410.2** | *SlHSP70-2* | 3 | 2 |
| **Solyc06g076020.2** | *SlHSP70-1* | 3 | 1 |
| **Solyc07g043560.2** | *SlHSP70-17* | 3 | 3 |
| **Solyc01g106260.2** | *SlHSP70-9* | 2 | 3 |
| **Solyc01g106210.2** | *SlHSP70-8* | 2 | 2 |
| **Solyc09g010630.2** | *SlHSP70-4* | 3 | 3 |
| **Solyc11g066100.1** | *SlHSP70-22* | 3 | 3 |
| **Solyc08g082820.2** | *SlHSP70-18* | 3 | 3 |
| **Solyc09g075950.1** | *SlHSP70-19* | 3 | 3 |

**Table 2. 4. GO enrichment analysis of *SlIAA9* gene on the network. (p < 0.05 adjusted by FDR)**

| Adjusted PValue_iaa9_1st2nd | description_iaa91st2nd | Name (GO_ID) | N_iaa91st2nd | nn_iaa91st2nd | X_iaa91st2nd | xx_iaa91st2nd |
|---|---|---|---|---|---|---|
| 2.44E-05 | protein binding | 5515 | 19542 | 3768 | 983 | 259 |
| 2.80E-03 | copper ion binding | 5507 | 19542 | 75 | 983 | 15 |
| 4.69E-03 | endopeptidase complex | 1905369 | 19542 | 13 | 983 | 6 |
| 4.69E-03 | proteasome complex | 502 | 19542 | 13 | 983 | 6 |
| 4.69E-03 | proteasome core complex | 5839 | 19542 | 13 | 983 | 6 |
| 4.69E-03 | peptidase complex | 1905368 | 19542 | 13 | 983 | 6 |
| 5.92E-03 | protein peptidyl-prolyl isomerization | 413 | 19542 | 9 | 983 | 5 |
| 5.92E-03 | peptidyl-prolyl cis-trans isomerase activity | 3755 | 19542 | 9 | 983 | 5 |
| 1.01E-02 | cis-trans isomerase activity | 16859 | 19542 | 10 | 983 | 5 |
| 1.49E-02 | oxidoreductase activity, acting on the CH-NH2 group of donors, disulfide as acceptor | 16642 | 19542 | 3 | 983 | 3 |
| 1.49E-02 | glycine dehydrogenase (decarboxylating) activity | 4375 | 19542 | 3 | 983 | 3 |
| 1.49E-02 | cytoplasm | 5737 | 19542 | 699 | 983 | 58 |
| 2.64E-02 | peptidyl-proline modification | 18208 | 19542 | 19 | 983 | 6 |
| 3.18E-02 | poly-pyrimidine tract binding | 8187 | 19542 | 36 | 983 | 8 |
| 3.18E-02 | poly(U) RNA binding | 8266 | 19542 | 36 | 983 | 8 |
| 3.30E-02 | proteasome core complex, alpha-subunit complex | 19773 | 19542 | 8 | 983 | 4 |
| 4.46E-02 | hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds | 16810 | 19542 | 68 | 983 | 11 |
| 4.61E-02 | NADH dehydrogenase (ubiquinone) activity | 8137 | 19542 | 22 | 983 | 6 |

**Table 2.5. GO enrichment analysis of *SlDELLA* gene on the network. (p < 0.05 adjusted by FDR)**

| Adjusted PValue_della1st2nd | description_della1st2nd | Name (GO_ID) | N_della1st2nd | nn_della1st2nd | X_della1st2nd | xx_della1st2nd |
|---|---|---|---|---|---|---|
| 3.04E-05 | endopeptidase complex | 1905369 | 19542 | 13 | 216 | 5 |
| 3.04E-05 | proteasome complex | 502 | 19542 | 13 | 216 | 5 |
| 3.04E-05 | proteasome core complex | 5839 | 19542 | 13 | 216 | 5 |
| 3.04E-05 | peptidase complex | 1905368 | 19542 | 13 | 216 | 5 |
| s4.05E-03 | unfolded protein binding | 51082 | 19542 | 55 | 216 | 6 |
| 9.10E-03 | protein binding | 5515 | 19542 | 3768 | 216 | 65 |
| 2.92E-02 | protein N-acetylglucosaminyltransferase activity | 16262 | 19542 | 3 | 216 | 2 |
| 2.92E-02 | soluble NSF attachment protein activity | 5483 | 19542 | 3 | 216 | 2 |

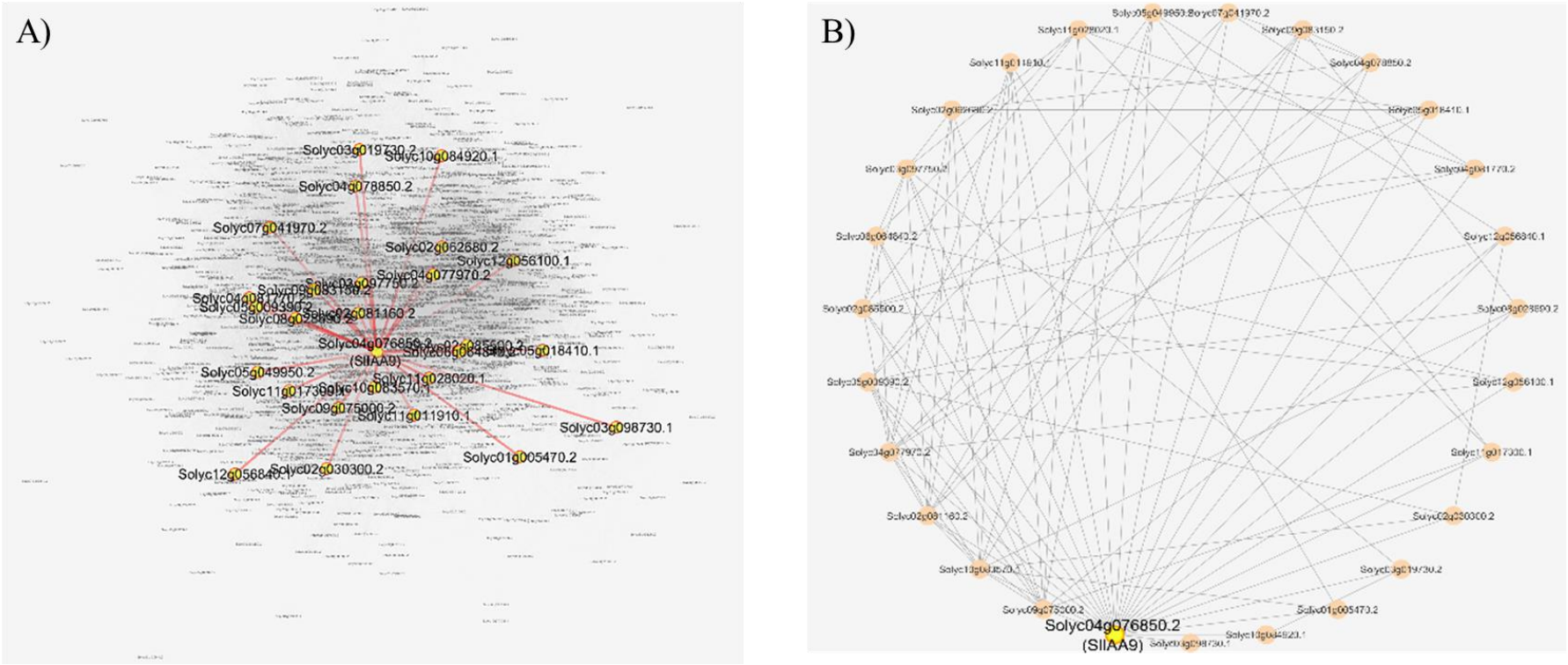



**Figure 2. 1. Co-expression network of *SlIAA9* neighbouring genes**

A) The first and second neighbouring genes of *SlIAA9* (*Solyc04g076850*.2.1) are drawn. B) The nearest neighbour genes of *SlIAA9* were extracted and drawn. *SlIAA9* (*Solyc04g076850*.2.1) is located at the bottom of network with yellow color

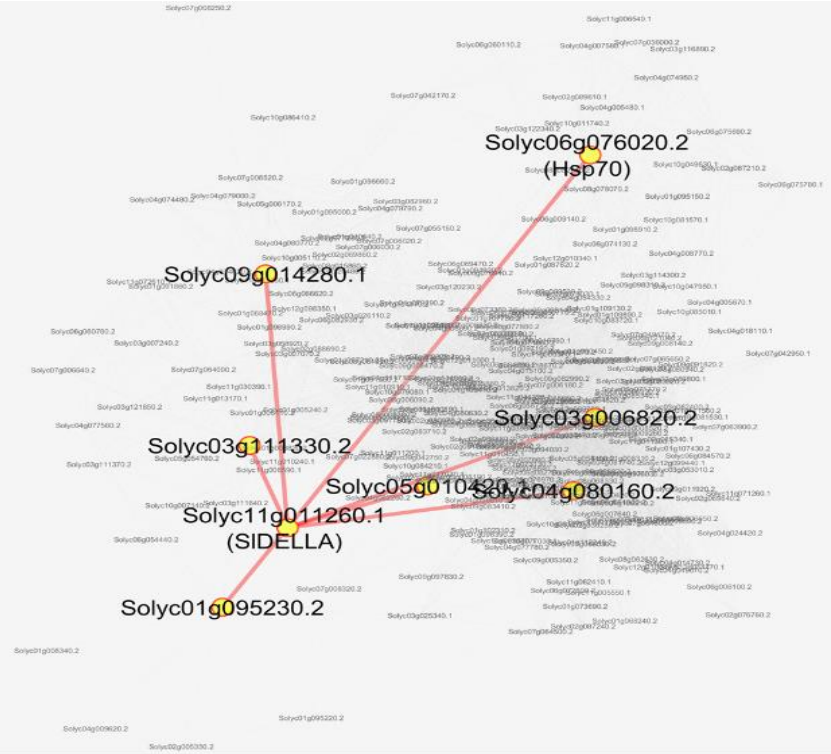



**Figure 2. 2. Co-expression network of *SlDELLA* gene neighbouring genes**

Co-expression network of *SlDELLA* gene and its neighbouring genes. The conditions of neighbouring genes and network drawing are the same as those of *SlIAA9*.

# Chapter 3. Genome-wide identification, structure characterization, phylogeny and expression patterns of tomato *SlHSP70* genes

## 3.1. Introduction

### *3.1.1. Introduction of heat shock protein HSP70*

Heat shock protein (HSP) are a family of proteins that respond to heat stress conditions and different types of cell stress in most organisms (De Maio, 1999). In plants, HSPs help resist adverse conditions caused by abiotic factors as temperature, drought, and salinity (Zou et al., 2012; Zhang et al., 2014; Augustine et al., 2015) and biotic factors such as virus or bacteria (Maimbo et al., 2007) in the growth, development and adaptation to environmental conditions. HSPs were divided into different families based on their molecular weight, in which five family genes were classified as *HSP70, HSP90, HSP60, HSP100,* and small weights heat shock protein (*sHSP*) (Wang et al., 2004). Of these, the *HSP70* family was the largest group that regulated plant growth processes and functions during heat stress (Sung et al., 2001a). HSP70 contributes to the formation of a new protein and helps stabilize the structure of the protein against the effect of stress and unfavorable conditions (Xu, 2018). HSP70 consists of two distinct functional domains: the ATPase domain (~44 kDa) and the peptide binding domain (~30 kDa), which is similar in several organisms (Sung et al., 2001a; Sharma and Masison, 2009; Smock et al., 2010). HSP70 was also identified in most organelles, showing the importance of this gene to the survival and development of the organism (Sung et al., 2001b).

### *3.1.2. HSP70 genes in plants*

Plants are organisms that cannot move themselves throughout their lifecycle, and therefore, they are greatly influenced by habitat conditions, especially those detrimental to survival and growth. In order to adapt to environmental conditions, plants developed a defense system that allows them to survive and grow under adverse conditions. HSP70 shares high similarity in its gene structure and protein characteristics among different species. Besides, the redundancy effect has made the number of member genes in each species relatively abundant. There were 22 *HSP70* genes identified in *Saccharomyces cerevisiae* while only 3 genes were identified in *E. coli* (Walsh et al., 2004). In *Arabidopsis*, there were at least 18 members in the *HSP70* superfamily, of which there were 14 genes in the *HSP70*

family, and four other genes belonging to the *HSP110* family (Lin et al., 2001). Spinach had at least 12 genes identified in the *HSP70* family while rice had 32 members of this gene family that provided the functional relevance of the *HSP70* component of the *HSP70*/J proteins bi-chaperone machine of rice (Guy and Li, 1998; Sarkar et al., 2013). However, no investigation of the HSP70 gene family in tomato plants has been conducted so far. Although the tomato genome has been completely decoded into 12 chromosomes with more than 950 Mb, the knowledge of this gene family in tomatoes is very limited and separated.

### *3.1.3. The objectives of this study*

In this chapter, we identified *HSP70* homologs in tomato based on the genomic and protein's sequences from the public database. Comparing with references on the database using free computational tools, we collected the potential genes of tomatoes *HSP70* family, characterized the gene and protein structure, and analyzed the relationship between these genes in the family. The objective of this work was establishing a list of tomato *HSP70* genes and enhancing the basic understanding of this gene family in tomato.

## 3.2. Materials and methods

### *3.2.1. Identification and annotation of SlHSP70 genes in tomato*

For the molecular description and phylogeny analysis of the *SlHSP70* gene, the genomic and amino acid sequences of the target gene were downloaded from Phytozome database ver.12.1 (available on https://phytozome.jgi.doe.gov/pz/portal.html) using a search tool with the keyword “HSP70” for the targeted gene and “tomato” for species. The amino acid of each gene was uploaded to InterPro database (https://www.ebi.ac.uk/interpro/) for protein sequence analysis, classification, and functional annotation on Pfam database (https://pfam.xfam.org/). Only proteins containing the domain of *SlHSP70* family were collected for *in silico* identification and structure characterization. The general gene annotation features, including gene identifier, locus identifier, genomic sequence, CDS sequence, amino acid sequence, and chromosomal distribution of each gene were obtained from the Phytozome database. Each gene ID was then searched against the Plant Genome and Systems Biology (PGSB) (http://pgsb.helmholtz-muenchen.de/plant/tomato/searchjsp/index.jsp) to determine the number of exons, and the position of the gene in the chromosome.

*SlHSP70* gene features were visualized with intron, exon and UTR compositions using a Gene Structure Display Server 2.0 (GSDS) (Guo et al., 2014), http://gsds.cbi.pku.edu.cn/).

### *3.2.2. Chromosomal localization and prediction of the duplication event of SlHSP70*

Gene duplication events occur when a part or whole CDS of a gene is duplicated during gene evolution. The duplication of the *SlHSP70* genes was analyzed by identifying their CDSs via alignment using the ClustalX (v. 2.1) (Larkin et al., 2007) with greater than 50% identity at the nucleotide level. The segmental duplication events were examined for the duplicated *SlHSP70* genes to determine if they distributed on different chromosomes. If duplicated *SlHSP70* genes localized on the same chromosome, they set a tandem duplication event (Feng et al., 2015).

In order to predict the duplication of the *SlHSP70* genes during gene evolution under the pressure of natural selection, we calculated the ratio of Ka/Ks using the DnaSP (v. 5.0) software (Librado and Rozas, 2009); Ks and Ka represent the values of synonymous substitutions per synonymous site and non-synonymous substitutions per non-synonymous site, respectively. The ratio Ka/Ks indicates the conservation level of the gene. If Ka/Ks > 1, it indicates that the examined protein tends to divert after time, and the opposite case is suggested for a conserved protein (Hurst, 2002). The approximate time of the duplication events expressed in "million years ago' "Mya) is estimated using $T = Ks/2\lambda$, with the mean value of clock-like rates ($\lambda$) of synonymous substitution. In tomato, $\lambda = 1.5\times10^{-8}$ substitutions per synonymous site per year (Liu et al., 2014).

### *3.2.3. Analysis of the protein features of the SlHSP70*

Information on the general characteristics of the SlHSP70 proteins, including their molecular weight (mW, kDa) and theoretical isoelectric point (pI), was explored using the ExPASy-Compute pI/Mw tool (https://web.expasy.org/compute_pi/) (Gasteiger et al., 2003). The subcellular localization of SlHSP70 proteins were predicted using the TargetP online tool (http://www.cbs.dtu.dk/services/TargetP/) (Emanuelsson et al., 2007).

### *3.2.4. Phylogeny tree construction*

Multiple sequence alignments were performed for amino acids using ClustalW (https://www.genome.jp/tools-bin/clustalw) (Thompson et al., 2003). The phylogeny analysis was inferred using the neighbor joining method (Nei and Saitou, 1987), and the phylogeny tree was constructed by MEGA7 software, with 1000 replicates for bootstrap testing (Kumar et al., 2016).

### *3.2.5. Expression analysis of the SlHSP70 genes in different tissues of tomato*

The *SlHSP70* transcript patterns in eight major organs of two tomato species *S. lycopersicum* var. M82 and *S. pennellii* were determined using the expression data obtained from the tomato eFP browse database (http://bar.utoronto.ca/efp_tomato/cgi-bin/efpWeb.cgi). The tomatoes were germinated and grown at 22°C under a mixture light of cool-white and far-red fluorescent light. The seedlings were collected three days after sowing on the plate. The shoot and the roots tissue were collected from the seedling 10 days after sowing on the plates. Vegetative meristems were collected when the third leaf reached 1 mm (30–37 days post germination). The stem between the 4th and 5th leaves and inflorescent meristem were collected at 50 days after germination for *S. lycopersicum* var. M82 and 56 days after germination for *S. pennellii*. Young green fruits and mature fruits were collected from plants in the greenhouse (Koenig et al., 2013).

## 3.3. Results

### *3.3.1. Molecular characterization of the SlHSP70 genes on the tomato genome*

Based on the genetic sequence of the tomato, obtained from the Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html), 25 genes were collected for tomato *HSP70* (*SlHSP70*), including 21 genes containing Interpro domain IPR013126 for *HSP70* and 4 genes contained Interpro domain IPR012725 for *Dna*K chaperone, an overlapping homologous subfamily of *HSP70*, which is often expressed in bacteria (Genevaux et al., 2007). Most of the 25 genes were uncharacterized (Table 3.1). To define the structural features of the *SlHSP70* genes in tomatoes, the exon/intron organization of the *SlHSP70* genes was investigated by aligning their CDSs and related genomic sequences using the gene structure display server (GSDS) tool (Hu et al., 2015). There were 22

*SlHSP70* genes were divided to several groups that shared similar genomic structures. There was, for instance, a group of genes comprising seven genes with two exons in the gene structure, while another group including three genes with six exons. Nine *SlHSP70* genes contained more than six exons, of which, two genes had seven exons, three gene had eight exons, and four genes had nine exons. In particular, one gene consisted of 13 exons while there were three intronless genes. The genomic features of the *SlHSP70* family gene are shown in Figure 3.2B and listed in Table 3.2. Multiple exon/intron gene organizations suggested a diverse structure for the *SlHSP70* gene observed in tomatoes. It has been well-established that introns are known to be essential entities to achieve an eukaryotic gene structure with numerous functions such as in exon shuffling, alteration of the gene expression, and regulation of the evolutionary rate of genes (Fedorova and Fedorov, 2003). The distinction in gene composition has shown the potential functions of intron in tomatoes in the development of the *SlHSP70* gene family.

We next obtained detailed information on the typical characteristics of all identified *SlHSP70* genes, which included protein length, molecular weight (mW), theoretical isoelectric point (pI), and subcellular localization, by subjecting their full-length protein sequence to the Expasy web-based tool for an analysis. Our data indicated that the *SlHSP70* genes possess a protein length between 80–890 amino-acid residues, as long as they brought the specific domain of the HSP70 protein. The molecular weight of *SlHSP70* were recorded from 8.68–93.88 kDa, and their pI values varied from acidic (pI = 4.32) to basic (pI = 9.3). The subcellular localization prediction showed that 25 SlHSP70 protein encoded for 2 chloroplast, 2 mitochondria, 1 plasma membrane, 4 endoplasmic reticulum, 1 nucleus, 11 cytoplasm, and 4 other proteins may localize in one or more organelles above. Some proteins showed similar descriptions of length, molecular weight (mW), and theoretical isoelectric point (pI), while their coding genes were similar in the gene structure. All gene features and protein characterizations are listed in Table 3. 3.

### *3.3.2. Chromosomal localization and prediction of the duplication events of SlHSP70 genes*

The chromosomal distributions of the identified *SlHSP70* genes were assessed based on the currently available information of each gene in the chromosome. Twenty-five genes were localized in 11 different chromosomes; in particular, no gene localized in chromosome 5, while chromosome 1 shared the highest number of *SlHSP70* genes with

five members, and it occupied 20 percent of all genes. Each chromosome 3, 6, 11, and 12 carried 3 *SlHSP70* genes, while chromosomes 7 and 9 shared 2 genes. Chromosomes of 2, 4, 8, and 10 contained one *SlHSP70* gene each. The overall distribution of the mapped *SlHSP70* genes on the annotated genome is shown in Figure 3.1. According to the chromosomal distribution of each *SlHSP70* gene, we named 25 mapped genes with their SolycID of each gene (Table 3. 1).

Further, we assessed the occurrences of duplication among the *SlHSP70* genes. According to our pre-set criteria for this analysis, 12 duplicated events with only 10 genes were detected (Table 3.4 and Figure 3.1). One duplicated event (*SlHSP70-1* and *SlHSP70-*14) was a tandem duplication event caused by localization within the chromosome while all other duplication events were determined as segmental type. Our assessment indicated that segmental duplication occurrences may play a major role in the development of the tomato gene family *SlHSP70*.

The most frequently duplication was identified with the *SlHSP70-1* gene. The gene *SlHSP70-1* was duplicated with 4 other genes, of which, only one tandem duplication event of a pair gene *SlHSP70-1/SlHSP70-14* was predicted in the family. The gene *SlHSP70-2* was duplicated with two genes while *SlHSP70-3* was duplicated with three genes. The approximate time of the duplication events expressed that the pairs of *SlHSP70-2/HSP70-3* and *SlHSP70-1/SlHSP70-21* were segregated last, which indicate they were latest separated during gene evolution. This claim is supported by the highest similarity when comparing the CDS sequences between the two genes (>85% identity). Meanwhile, the pair of *HSP70-3/SlHSP70-14* was earliest dissociations leads to the least sharing in their CDS sequences with only 50.9% between them. The abundant duplication of the *SlHSP70* genes suggested the important role of these genes in the growth, development, and stress resistance of tomato.

In terms of chromosome localization, genes on chromosome 6 tend to repeat with most genes. All three genes on chromosome 6 were replicated with others, whereas chromosome 3 replicates 2 genes, while chromosomes 4, 8, 9, 10, 11, and 12 share 1 gene with repetition. In particular, chromosome 1 carries the most genes (5 genes); however, it cannot predict gene duplication and the repeating genes are identified, similar to chromosomes 2 and 7.

### *3.3.3. Phylogenetic analysis-based on amino acid sequences*

To characterize the phylogenetic relationship of the identified *SlHSP70* genes, the full-length protein sequences of 25 *SlHSP70* were aligned, and an unrooted tree was

created using the neighbor joining method (Figure 3.2A). The results indicated that a closed relationship in the group of genes *SlHSP70-1*, *SlHSP70-2, SlHSP70-3*, *SlHSP70-4,* and *SlHSP70-21* when they were placed together in one cluster of the phylogeny tree. The protein of the targeted gene (*SlHSP70*-1) showed similarity with the other three genes *SlHSP70-2, SlHSP70-4,* and *SlHSP70-21*. The structure of the four genes has the same number of introns, exons, and similar length (Figure 3.2B). This suggests that they might share biological functions though the functions of these genes are still uncharacterized. The genes mentioned above and some other genes *SlHSP70-13, SlHSP70-22,* and *SlHSP70-23* together form a phylogenetic cluster, including the family of genes that shares high structural similarity with little exon. The next cluster includes genes *SlHSP70-6, SlHSP70-11, SlHSP70-15,* and *SlHSP70-19.* This is group of multiple exons that are very similar in terms of the genomic structure. Another cluster of multiple exons consists of genes *SlHSP70-7, SlHSP70-8, SlHSP70-9,* and *SlHSP70-20.* Two these clusters are separated by the last phylogenetic multi-exon genes cluster that consists of genes *SlHSP70-10, SlHSP70-24, SlHSP70-25,* and *SlHSP70-17* localized far away in the genome structure. Four phylogenetic clusters were found based on genomics and amino acid sequences in the family gene *HSP70*. With elevated genetic similarity and protein characterization, it can be suggested that this gene shares biological functions during the physiological development of the tomato.

### *3.3.4. Transcript patterns of the SlHSP70 genes in major organs of tomato plants during growth and development*

Organ/tissue-specific transcription patterns can suggest the role of a gene in a specific organ or tissue (Pontes et al. 2013). In this study, the transcriptome data publicly available in the Tomato eFP browser (Koenig et al., 2013) was used to explore the expression profiles of 25 *SlHSP70* genes in various tomato organs under normal growth conditions. Among 25 genes of *SlHSP70* family, gene *SlHSP70-15* had no available expression information from the transcriptome database (Figure 3.3); 24 remaining *SlHSP70* genes were differentially expressed in the examined organs. Five genes (*SlHSP70-5, SlHSP70-12, SlHSP70-14, SlHSP70-15* and *SlHSP70-23*) were expressed at very low levels in all major organs. On the contrary, genes *SlHSP70-7, SlHSP70-8, SlHSP70-11, SlHSP70-18, SlHSP70-21, SlHSP70-25, SlHSP70-2, and SlHSP70-4* were highly expressed in most tissues that were checked; in particular, gene *SlHSP*70-*4 was* expressed at the highest level compared to other genes in the family for all tissues, and it was followed by the couple of *SlHSP70-7* and *SlHSP70-21* genes in the same tissues. A

list of genes *SlHSP70-8, SlHSP70-9, SlHSP70-11, SlHSP70-24,* and *SlHSP70-25* showed the high expression level in most tissues, except at the leaf, while some other genes only expressed in a specific tissues; for example, genes *SlHSP70-6, SlHSP70-13, SlHSP70-17, SlHSP70-20,* and *SlHSP70-22* only expressed in fruit, genes *SlHSP70-1* and *SlHSP70-3* expressed in stem and fruit. Almost all genes expressed in the vegetative tissues were also expressed in the fruit of tomatoes (Figure 3. 3). The differentiation of the gene functions in plant growth and development suggested the distinct concentrations of a gene in distinct tissues. Meanwhile, a number of different genes tended to be expressed together in the same tissues, suggesting that they may be co-expressed or associated together for performing their functions. The co-expression or overlapping of genes functions may be caused by high homology or duplication, as described in the previous section.

There were also substantial variations between the genes in the same analyzed tissue in a group duplication gene. For example, while the *SlHSP70-4* gene was consistently highest expressed in most tissues, the three genes *SlHSP70-3*, *SlHSP70-14,* and *SlHSP70-1* exhibited very low levels. The gene *SlHSP70-3* only increases the level of gene expression in developing and mature fruits, which could imply that *SlHSP70-3* gene carries the ethylene response factor 21 (ER21) sequences that play an important function for tomato fruit ripening (Zegzouti et al., 1999). The genes *SlHSP70-2* and *SlHSP70-21* also expressed in most analyzed tissues with lower levels than *SlHSP70.4* gene, suggested that expression level distinction for each gene in a duplication group showed the function of each gene in each portion of the tomato plant (Figure 3.4).

### 3.4. Discussion

The *HSP70* family gene has been characterized previously in several plants such as *Arabidopsis*, rice, populus, and soybeans (Lin et al., 2001; Jung et al., 2013; Zhang et al., 2015a; Zhang et al., 2015b). In tomatoes, some individual genes belonging to this gene family were also mentioned in the responses to the virus (Gorovits et al., 2013; Gorovits and Czosnek, 2017). However, the HSP70 gene family was not examined in tomatoes. Therefore, this study performed a genome-wide analysis of the *HSP70* gene family in tomato using a combination of the phylogeny, chromosomal locations, gene structures, and expression profiles.

The *HSP70* gene family in tomato consisted of 25 genes and was larger than that in *Arabidopsis* (18 genes), moss (21 genes), ad potato (20 genes), but it had a smaller number of HSP70 gene families compared to some other plant species such as rice (32 genes *OsHSP70)* or some species of cotton family (25 to 31 genes) ((Lin et al., 2001; Jung et al., 2013; Tang et al., 2016; Xiao et al., 2017; Liu et al., 2018). *HSP70* is a multi-member gene family distributed in different organelles (Cho and Choi, 2009). In our study, the analysis of the subcellular localization of *SlHSP70* proteins showed that multiple *SlHSP70* members were identified in various cellular compartments in tomato. The *SlHSP70* proteins might be localized in all organelles as nucleus, cytoplasm, and mitochondrial membranes, in which they were distributed foremost in the nucleus and the cytoplasm, and similar localization were observed for HSP70 genes in soybean, rice, and *Arabidopsis*.

We found that the nearest gene cluster in the whole family shares similar exon/intron structures and intron numbers based on the phylogenetic and gene structures of the HSP70 gene in tomatoes. For instance, the gene *SlHSP70*-1 has the same amount of exon to almost all of its duplicate versions. These results were determined to confirm the features identified in the phylogenetic assessment (Figure 3.2). We have however identified that some duplication pairs show modifications in their structure for intron/exon, such as the pair of *SlHSP70-1/SlHSP70-14* in this study. Any changes in the structure of intron/exon can enhance or repress the importance of the gene during structural evolution among the tomato *SlHSP70* genes (Warf and Berglund, 2010).

The duplications may appear in a part of the gene, in the whole gene, in a chromosomal segments, or entire genomes for the evolution of plant genome structures (Semon and Wolfe, 2007). In this study, we evaluated the gene duplication events (tandem and segmental duplications) to understand the expansion mechanism of the tomato *SlHSP70* gene family. A tandem duplication event can be determined when two or more genes are present on the same chromosome, whereas a segmental duplication event when gene duplications occur on separated chromosomes (Schlueter et al., 2007). Segment and tandem reproduction occurrences are essential for the stabilization and enhancement of the gene functions of the *SlHSP70* gene family. In this study, the tomato HSP70 family gene recorded gene segmental duplication.

Our heat map data showed that most *SlHSP70* genes were expressed in different tissues and organs of tomato, which indicates that they can help to develop and grow the

tomato. The expression patterns for duplication pairs were similar in the same plant tissues, which revealed that pairs with high sequence similarity had similar expression patterns. However, several pairs of duplications demonstrated that distinct patterns of expression could suggest that the gene function alternated during gene evolution.

**Table 3. 1. List of *HSP70* genes identified in tomato genome and their general characteristics**

| Gene name[a] | Gene ID[c] | Transcript Name | Chromosome | Start | Stop | Strand | Description |
|---|---|---|---|---|---|---|---|
| *SlHSP70*-1 | Solyc06g076020.2 | Solyc06g076020.2.1 | Chr06 | 43585486 | 43582389 | reverse | heat shock protein (AHRD V1 ***-B2D2G5_CAPSN); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-2 | Solyc10g086410.2 | Solyc10g086410.2.1 | Chr10 | 64561395 | 64564764 | forward | Heat shock protein 70-3 (AHRD V1 ***-Q67BD0_TOBAC); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-3 | Solyc04g011440.2 | Solyc04g011440.2.1 | Chr04 | 3894918 | 3898067 | forward | heat shock protein (AHRD V1 ***-B2D2G5_CAPSN); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-4 | Solyc09g010630.2 | Solyc09g010630.2.1 | Chr09 | 3965253 | 3968837 | forward | heat shock protein (AHRD V1 ***-B2D2G5_CAPSN); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-5 | Solyc01g060400.1 | Solyc01g060400.1.1 | Chr01 | 63689969 | 63689603 | reverse | Heat shock protein (AHRD V1 **--Q8IB24_PLAF7); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-6 | Solyc01g099660.2 | Solyc01g099660.2.1 | Chr01 | 81599813 | 81602924 | forward | Heat shock protein (AHRD V1 ***-Q84KP8_CYAME); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-7 | Solyc01g103450.2 | Solyc01g103450.2.1 | Chr01 | 83821528 | 83826037 | forward | Chaperone DnaK (AHRD V1 ***-Q1SKX2_MEDTR); contains Interpro domain(s) IPR012725 Chaperone DnaK |
| *SlHSP70*-8 | Solyc01g106210.2 | Solyc01g106210.2.1 | Chr01 | 85918285 | 85922197 | forward | Chaperone DnaK (AHRD V1 ***-A2Q199_MEDTR); contains Interpro domain(s) IPR012725 Chaperone DnaK |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *SlHSP70*-9 | Solyc01g106260.2 | Solyc01g106260.2.1 | Chr01 | 85976768 | 85981140 | forward | Chaperone DnaK (AHRD V1 ***-A2Q199_MEDTR); contains Interpro domain(s) IPR012725 Chaperone DnaK |
| *SlHSP70*-10 | Solyc02g080470.2 | Solyc02g080470.2.1 | Chr02 | 39251685 | 39263151 | forward | Heat shock protein 4 (AHRD V1 ***-B6U237_MAIZE); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-11 | Solyc03g082920.2 | Solyc03g082920.2.1 | Chr03 | 46345339 | 46341372 | reverse | Heat shock protein (AHRD V1 ***-Q84KP8_CYAME); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-12 | Solyc03g117620.2 | Solyc03g117620.2.1 | Chr03 | 60776507 | 60775354 | reverse | Heat shock protein (AHRD V1 *-*-Q801X9_CARAU); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-13 | Solyc03g117630.1 | Solyc03g117630.1.1 | Chr03 | 60779574 | 60777610 | reverse | Heat shock protein (AHRD V1 ***-B2D2G5_CAPSN); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-14 | Solyc06g005440.1 | Solyc06g005440.1.1 | Chr06 | 441236 | 441592 | forward | heat shock protein (AHRD V1 ***-B2D2G5_CAPSN); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-15 | Solyc06g052050.2 | Solyc06g052050.2.1 | Chr06 | 32202991 | 32206019 | forward | Heat shock protein (AHRD V1 ***-Q84KP8_CYAME); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-16 | Solyc07g005820.2 | Solyc07g005820.2.1 | Chr07 | 659235 | 655717 | reverse | heat shock protein (AHRD V1 ***-B2D2G5_CAPSN); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-17 | Solyc07g043560.2 | Solyc07g043560.2.1 | Chr07 | 54789496 | 54781149 | reverse | Heat shock protein 4 (AHRD V1 ***-B6U237_MAIZE); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-18 | Solyc08g082820.2 | Solyc08g082820.2.1 | Chr08 | 62655311 | 62659585 | forward | Heat shock protein (AHRD V1 ***-Q84KP8_CYAME); contains Interpro domain(s) IPR013126 Heat shock protein 70 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *SlHSP70*-19 | Solyc09g075950.1 | Solyc09g075950.1.1 | Chr09 | 63079410 | 63081140 | forward | Heat shock protein 1 (AHRD V1 ***-B6SXY0_MAIZE); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-20 | Solyc11g020040.1 | Solyc11g020040.1.1 | Chr11 | 10015582 | 10019521 | forward | Chaperone DnaK (AHRD V1 ***-Q1SKX2_MEDTR); contains Interpro domain(s) IPR012725 Chaperone DnaK |
| *SlHSP70*-21 | Solyc11g066060.1 | Solyc11g066060.1.1 | Chr11 | 48824058 | 48826931 | forward | heat shock protein (AHRD V1 ***-B2D2G5_CAPSN); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-22 | Solyc11g066100.1 | Solyc11g066100.1.1 | Chr11 | 48858939 | 48856641 | reverse | heat shock protein (AHRD V1 ***-B2D2G5_CAPSN); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-23 | Solyc12g042560.1 | Solyc12g042560.1.1 | Chr12 | 43346086 | 43345106 | forward | heat shock protein (AHRD V1 *-*-B2D2G5_CAPSN); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-24 | Solyc12g043110.1 | Solyc12g043110.1.1 | Chr12 | 44216059 | 44210946 | forward | Heat shock protein 4 (AHRD V1 ***-B6U237_MAIZE); contains Interpro domain(s) IPR013126 Heat shock protein 70 |
| *SlHSP70*-25 | Solyc12g043120.1 | Solyc12g043120.1.1 | Chr12 | 44230445 | 44226370 | forward | Heat shock protein 4 (AHRD V1 ***-B6U237_MAIZE); contains Interpro domain(s) IPR013126 Heat shock protein 70 |

[a] The genes were named in this study;

**Table 3. 2. General features of the identified *SlHSP70* genes**

| Gene name[a] | Gene ID | CDS length (bp) | gDNA length (bp) | GC content (%) | number of exons |
|---|---|---|---|---|---|
| *SlHSP70*-1 | Solyc06g076020.2 | 1947 | 3098 | 41.64 | 2 |
| *SlHSP70*-2 | Solyc10g086410.2 | 1935 | 3370 | 39.55 | 2 |
| *SlHSP70*-3 | Solyc04g011440.2 | 1956 | 3150 | 39.3 | 2 |
| *SlHSP70*-4 | Solyc09g010630.2 | 1950 | 3585 | 40.39 | 2 |
| *SLHSP70*-5 | Solyc01g060400.1 | 243 | 367 | 35.42 | 2 |
| *SLHSP70*-6 | Solyc01g099660.2 | 2010 | 3112 | 37.08 | 7 |
| *SLHSP70*-7 | Solyc01g103450.2 | 2112 | 4510 | 40.66 | 8 |
| *SLHSP70*-8 | Solyc01g106210.2 | 2046 | 3913 | 39.64 | 6 |
| *SLHSP70*-9 | Solyc01g106260.2 | 2013 | 4373 | 39.36 | 6 |
| *SLHSP70*-10 | Solyc02g080470.2 | 2262 | 11467 | 35.86 | 9 |
| *SLHSP70*-11 | Solyc03g082920.2 | 2004 | 3968 | 38.08 | 7 |
| *SLHSP70*-12 | Solyc03g117620.2 | 561 | 1154 | 37.09 | 2 |
| *SLHSP70*-13 | Solyc03g117630.1 | 1965 | 1965 | 44.12 | 1 |
| *SLHSP70*-14 | Solyc06g005440.1 | 357 | 357 | 43.42 | 1 |
| *SLHSP70*-15 | Solyc06g052050.2 | 1860 | 3029 | 39.32 | 9 |
| *SlHSP70*-16 | Solyc07g005820.2 | 1965 | 3519 | 33.96 | 2 |
| *SlHSP70*-17 | Solyc07g043560.2 | 2673 | 8348 | 37.04 | 13 |
| *SlHSP70*-18 | Solyc08g082820.2 | 2001 | 4275 | 39.65 | 8 |
| *SlHSP70*-19 | Solyc09g075950.1 | 1731 | 1731 | 44.89 | 1 |
| *SlHSP70*-20 | Solyc11g020040.1 | 2079 | 3940 | 39.62 | 8 |
| *SlHSP70*-21 | Solyc11g066060.1 | 2097 | 2874 | 40.6 | 3 |
| *SlHSP70*-22 | Solyc11g066100.1 | 1965 | 2299 | 41.89 | 2 |
| *SlHSP70*-23 | Solyc12g042560.1 | 633 | 981 | 37.92 | 4 |
| *SlHSP70*-24 | Solyc12g043110.1 | 2559 | 5114 | 36.53 | 9 |
| *SlHSP70*-25 | Solyc12g043120.1 | 2541 | 4076 | 40.46 | 9 |

[a]The genes were named in this study;

**Table 3. 3. Protein feature of the identified *SlHSP70* amino acids**

| Gene name | Gene ID | Protein ID | Protein length (aa) | mW (kDa) | pI | Localization |
|---|---|---|---|---|---|---|
| *SlHSP70-1* | Solyc06g076020.2 | Solyc06g076020.2.1 | 648 | 71 | 5.04 | Cytoplasmic |
| *SlHSP70-2* | Solyc10g086410.2 | Solyc10g086410.2.1 | 644 | 70.78 | 5.07 | Cytoplasmic |
| *SlHSP70-3* | Solyc04g011440.2 | Solyc04g011440.2.1 | 651 | 71.39 | 5.13 | Cytoplasmic |
| *SlHSP70-4* | Solyc09g010630.2 | Solyc09g010630.2.1 | 649 | 71.22 | 5.13 | Cytoplasmic |
| *SlHSP70-5* | Solyc01g060400.1 | Solyc01g060400.1.1 | 80 | 8.68 | 9.3 | Mitochondrial/ Chloroplast/ cytoplasm |
| *SlHSP70-6* | Solyc01g099660.2 | Solyc01g099660.2.1 | 669 | 74.64 | 5.36 | ER |
| *SlHSP70-7* | Solyc01g103450.2 | Solyc01g103450.2.1 | 703 | 74.9 | 5.2 | Chloroplast |
| *SlHSP70-8* | Solyc01g106210.2 | Solyc01g106210.2.1 | 681 | 72.97 | 5.75 | Mitochondrial |
| *SlHSP70-9* | Solyc01g106260.2 | Solyc01g106260.2.1 | 670 | 71.88 | 5.95 | Mitochondrial |
| *SlHSP70-10* | Solyc02g080470.2 | Solyc02g080470.2.1 | 753 | 84.1 | 6.02 | Nuclear |
| *SlHSP70-11* | Solyc03g082920.2 | Solyc03g082920.2.1 | 667 | 73.46 | 5.07 | ER |
| *SlHSP70-12* | Solyc03g117620.2 | Solyc03g117620.2.1 | 186 | 21.27 | 9.73 | Mitochondrial/ nuclear |
| *SlHSP70-13* | Solyc03g117630.1 | Solyc03g117630.1.1 | 654 | 71.85 | 5.21 | Cytoplasmic |

| | | | | | | |
|---|---|---|---|---|---|---|
| *SlHSP70-14* | Solyc06g005440.1 | Solyc06g005440.1.1 | 118 | 13.18 | 4.32 | Plasma membrane |
| *SlHSP70-15* | Solyc06g052050.2 | Solyc06g052050.2.1 | 619 | 67.51 | 5.04 | ER |
| *SlHSP70-16* | Solyc07g005820.2 | Solyc07g005820.2.1 | 654 | 71.95 | 5.15 | Cytoplasmic |
| *SlHSP70-17* | Solyc07g043560.2 | Solyc07g043560.2.1 | 890 | 98.79 | 5.91 | ER/ nuclear |
| *SlHSP70-18* | Solyc08g082820.2 | Solyc08g082820.2.1 | 666 | 73.2 | 5.1 | ER |
| *SlHSP70-19* | Solyc09g075950.1 | Solyc09g075950.1.1 | 576 | 62.72 | 5.56 | Cytoplasmic |
| *SlHSP70-20* | Solyc11g020040.1 | Solyc11g020040.1.1 | 692 | 74.49 | 5.36 | Chloroplast |
| *SlHSP70-21* | Solyc11g066060.1 | Solyc11g066060.1.1 | 698 | 77.14 | 5.51 | Cytoplasmic |
| *SlHSP70-22* | Solyc11g066100.1 | Solyc11g066100.1.1 | 654 | 71.46 | 5.1 | Cytoplasmic |
| *SlHSP70-23* | Solyc12g042560.1 | Solyc12g042560.1.1 | 210 | 23.56 | 5.82 | Cytoplasmic/ nuclear |
| *SlHSP70-24* | Solyc12g043110.1 | Solyc12g043110.1.1 | 852 | 93.88 | 5.23 | Cytoplasmic |
| *SlHSP70-25* | Solyc12g043120.1 | Solyc12g043120.1.1 | 846 | 92.97 | 5.22 | Cytoplasmic |

**Table 3. 4. Prediction of duplication events among the identified *SlHSP70* genes in tomato**

| Gene Pairs | Chromosome Localization | Duplication Event | Identity Level (%) | Ka* | Ks** | Ka/Ks | T (Mya)*** |
|---|---|---|---|---|---|---|---|
| *SlHSP70-1/SlHSP70-21* | Chr06/Chr11 | segment | 86.18 | 0.02 | 0.75 | 0.026667 | 25 |
| *SlHSP70-1/SlHSP70-2* | Chr06/Chr10 | segment | 80.26 | 0.04 | 1.72 | 0.023256 | 57.33333 |
| *SlHSP70-1/SlHSP70-4* | Chr06/Chr09 | segment | 80.95 | 0.05 | 1.3 | 0.038462 | 43.33333 |
| *SlHSP70-3/SlHSP70-4* | Chr04/Chr09 | segment | 80 | 0.05 | 1.49 | 0.033557 | 49.66667 |
| *SlHSP70-3/SlHSP70-1* | Chr04/Chr06 | segment | 78.27 | 0.05 | 1.88 | 0.026596 | 62.66667 |
| *SlHSP70-3/SlHSP70-14* | Chr04/Chr06 | segment | 50.9 | 0.28 | 2.59 | 0.108108 | 86.33333 |
| *SlHSP70-4/SlHSP70-2* | Chr09/Chr10 | segment | 85.32 | 0.02 | 0.74 | 0.027027 | 24.66667 |
| *SlHSP70-11/SlHSP70-18* | Chr03/Chr08 | segment | 77.01 | 0.06 | 2.02 | 0.029703 | 67.33333 |
| *SlHSP70-11/SlHSP70-14* | Chr03/Chr06 | segment | 75.86 | 0.06 | 0.7 | 0.085714 | 23.33333 |
| *SlHSP70-13/SlHSP70-23* | Chr03/Chr12 | segment | 69.35 | 0.11 | 1.25 | 0.088 | 41.66667 |
| *SlHSP70-14/SlHSP70-4* | Chr06/Chr09 | segment | 53.78 | 0.31 | 1.37 | 0.226277 | 45.66667 |
| *SlHSP70-14/SlHSP70-1* | Chr06/Chr06 | tandem | 55.46 | 0.3 | 1.62 | 0.185185 | 54 |

*Ka, value indicating nonsynonymous substitutions per nonsynonymous site;

**Ks, value indicating synonymous substitutions per synonymous site;

***T, approximate time of the duplication event; Mya, million years ago.

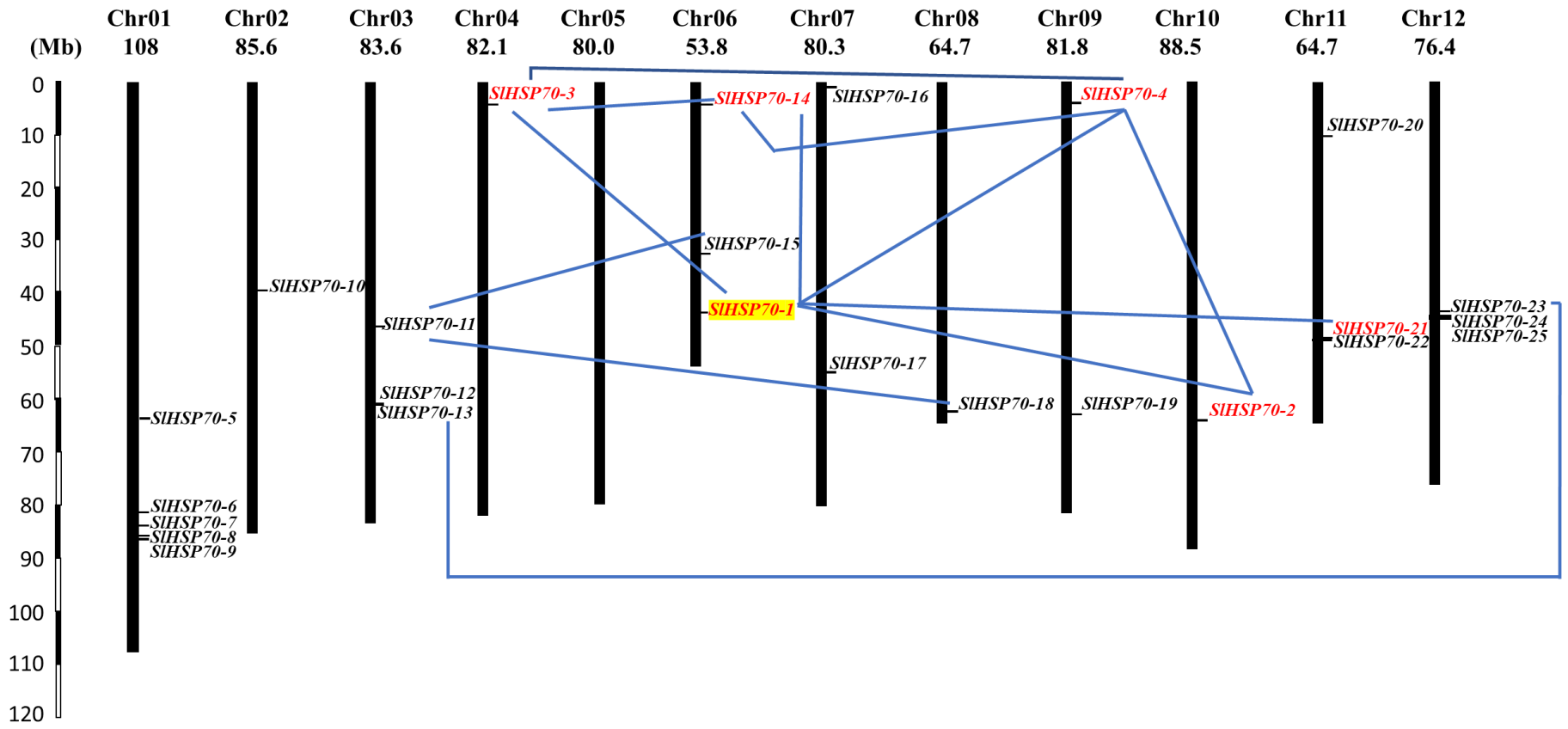



**Figure 3. 1. Chromosomal localization of 25 identified *SlHSP70* genes in the tomato genome**

Blue colour indicates the *SlHSP70* genes form duplicated pairs with the criterion of higher than 50% nucleotide identity. Targeted gene in red colour and yellow highlight, duplication version of targeted gene in red colour only.

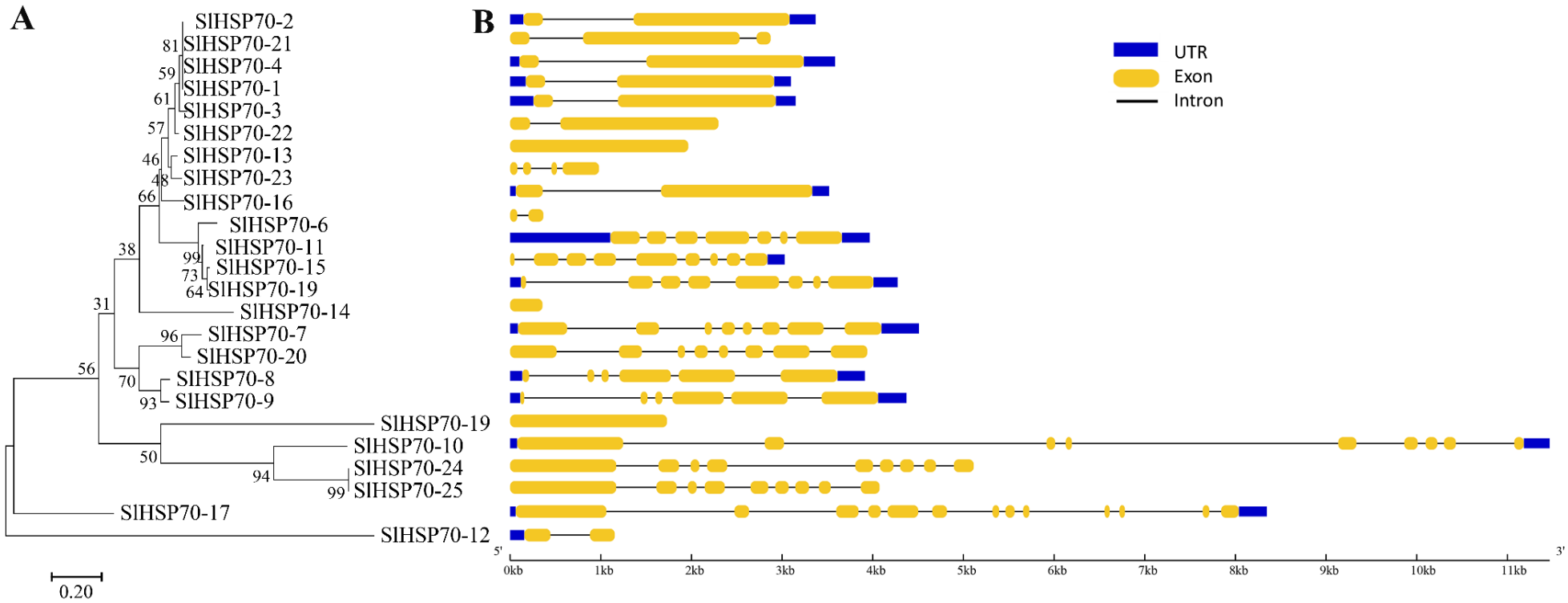



**Figure 3. 2. Phylogenetic and structural analysis of identified *SlHSP70* member's family in tomato**

A) The phylogeny analysis was inferred using the Neighbour-Joining method. Phylogeny tree was constructed by MEGA7 software, with 1000 replicates for bootstrap test, using amino acid sequences. B) Genomic structure description of *SlHSP70*. The structures were visualized by Gene Structure Display Server (GSDS) ver. 2.0 based on genomic sequence. The position of each gene was rearranged following the position of amino acid in phylogeny tree

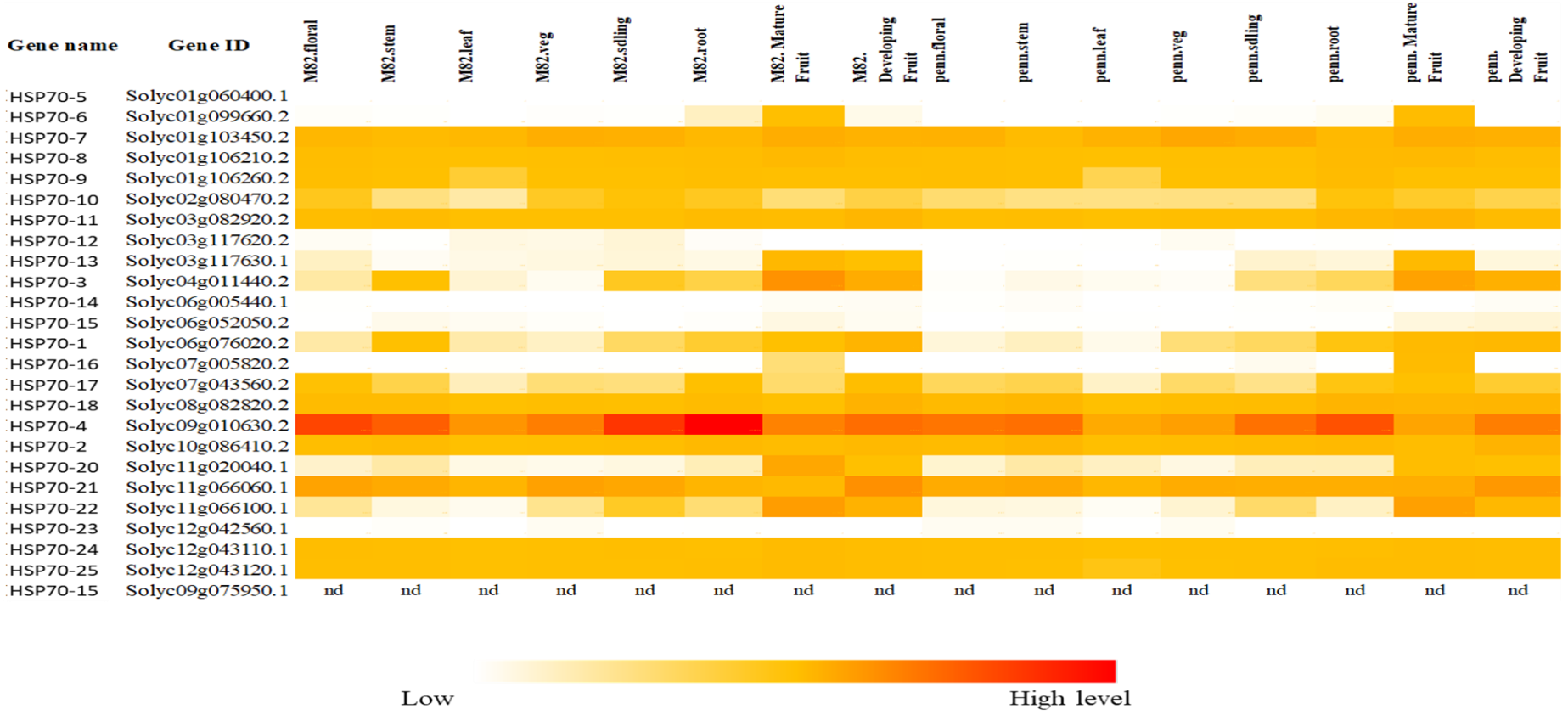



**Figure 3. 3. Gene expression level of *SlHSP70* genes in each tissues of tomato *S. lycopersicum* M82 and *S. pennelli***

The seedlings at 3 days after sowing on plate. Shoot and Roots tissue from seedling at 10 days after sowing on plates. Vegetative meristems were collected when the 3rd leaf reached 1 mm. Stem between 4th and 5th leaves and inflorescent meristem were collected at 50 days after germination for *S. lycopersicum* var. M82 and 56 days after germination for *S. pennellii*. Young green fruits and mature fruits were collected from plants in the greenhouse.

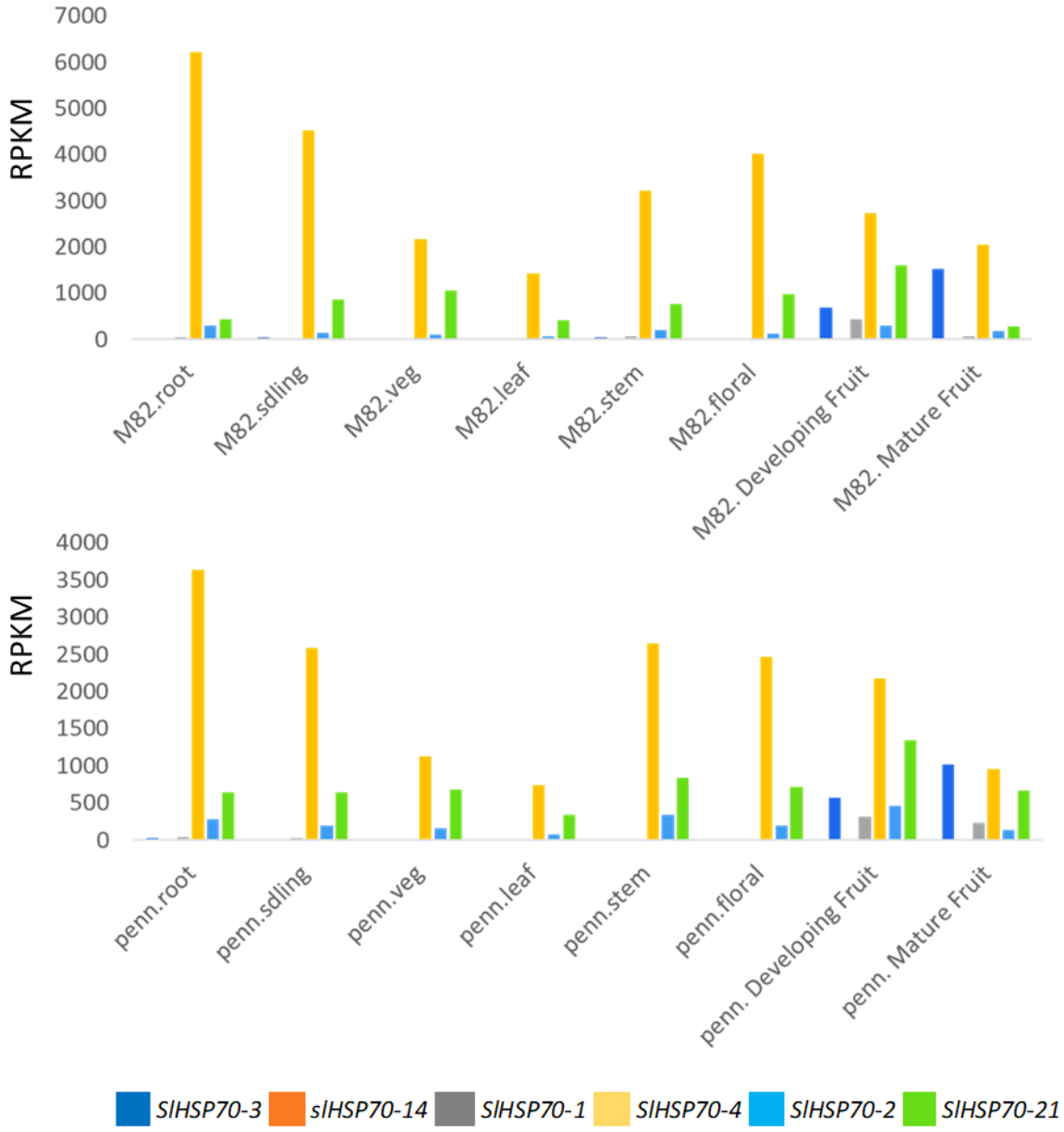



**Figure 3.4. Comparison of expression pattern of *SlHSP70*-1 and its duplications in *HSP70* family gene by tissue of tomato *S. lycopersicum* var. M82 (A) and *S. pennellii* (B).**

RPKM: Reads Per Kilobase of transcript, per Million mapped reads is a normalized unit of transcript expression. The information of each tissues was same as Figure 3.3.

# Chapter 4: Over expression of *SlHSP70-1* gene and phenotypic characterization

## 4.1. Introduction

### *4.1.1. Roles of tomato as a model plant*

Tomato (*Solanum lycopersicum*) was an important commercial vegetable that belongs to the *Solanaceae* family (Consortium, 2012). The tomato genome consists of 12 chromosomes with 950 Mb genome size and it is completely sequenced by the Tomato Genome Consortium (2012), which provides a huge data for genomic researches and reference genome for more than 3000 species (Consortium, 2012). One cultivar of tomato, namely Micro-Tom, is considered a great model because of its relatively short life cycle, small size, prolific seed production, and a small genome size with the availability of genetic and genomic resources (Meissner et al., 1997). In particular, the tomato represents typical climacteric fleshy fruits that have all characteristics from flowering to ripening that cannot be obtained from a plant such as *Arabidopsis* and rice (Shimamoto and Kyozuka, 2002; Koornneef and Meinke, 2010). Understanding the tomato's gene functions can help spread the understanding of the principle and dynamics of molecular plant physiology to create new valuable agronomic traits for vegetable and fruit plants (Aoki et al., 2013).

### *4.1.2. Function of SlIAA9 and SlDELLA in tomato growth, development, and fruit set*

*SlIAA9* and *SlDELLA* are two well-known genes that are involved in plant growth and development, including fruit set and enlargement through cell division and cell expansion. *SlIAA9* and *SlDELLA* were mentioned for inducing parthenocarpy fruits, which are desirable characteristics for producers and customers to improve manufacturing and quality of tomato fruits (Wang et al., 2005; Marti et al., 2007; Hu et al., 2018). *IAA9* is a member of *Aux*/*IAA* gene family consisting of 26 genes in tomato plants, which plays as transcriptional factors repressing auxin response (ARF) genes (Audran-Delalande et al., 2012; Wu et al., 2012) . Tomato *Aux/IAA9* acts as a negative regulator of the auxin response involved in controlling fruit set, which represses the transcription of the auxin responsive and fruit developmental genes, through the interaction with *Auxin response factors 7* (*ARF*) (de Jong et al., 2009b). Downregulation in the *SlIAA9* gene level can induce parthenocarpy, simplify leaf shape, and elongate shoot that lead to changes in plant height (Wang et al., 2005; Zhang et al., 2007;

Okabe et al., 2011; Mazzucato et al., 2015). Meanwhile, *SlDELLA* is also a negative regulator of gibberellin signaling through combination with the gibberellin receptor GID1 (Yoshida et al., 2014). DELLA has been suggested to function as a transcriptional activator and the fundamental components of the GA-DELLA signaling pathway. *Procera,* a *della* mutant, showed morphological changes in plant elongation, branching architect, and in reproductive organs development, especially to promote parthenocarpy (Marti et al., 2007; Bassel et al., 2008; Carrera et al., 2012; Lombardi-Crestana et al., 2012; Hu et al., 2018). The parthenocarpy development in both *Sliaa9* and *Sldella* mutants was clarified by their downstream activities that stimulated phytohormone-related fruit development through their direct crosstalk (Hu et al., 2018). However, changes have been surveyed in the plant architecture of these mutants, and these mechanisms stay unknown.

### *4.1.3. The objective of this study*

In the previous chapter, we constructed the gene-to-gene analysis of a co-expression network, using publicly available microarray data, to extract genes directly connected to nodes *SlIAA9* and *SlDELLA,* respectively. As the results, we discovered the directed connection of *SlHSP70-1* to *SlDELLA* and its appearance in the co-expression network of *SlIAA9*. The appearance of *SlHSP70-1* in the directed connection to *SlDELLA* obtained a new hypothesis for the function of *SlHSP70-1* in the tomato plant growth and development in a relationship with *SlDELLA.*

To gain insights into the growth, development, and fruit set of tomatoes, we applied an approach that changes the ordinary targeted gene condition by improving its expression, to explore phenotype changes caused by the target gene. In this chapter, we discussed the overexpression of the *SlHSP70-1* coding gene. Phenotyping using the lines were conducted to evaluate gene functions of the targeted gene by comparing the growth and development of a transgenic plant with that of the wild type as control grown under a no-stress condition. The upregulation of *SlHSP70-1* promoted internode elongation, leading to higher plants than wild type. The expression of *SlHSP70-1* was regulated by *SlIAA9*. This research suggested that the roles of *SlHSP70* in tomato growth and development processes in relationship with *SlIAA9* regulation.

## 4.2. Materials and Methods

### *4.2.1. Plasmid construction and Agrobacterium introduction*

The binary vector pDEST_35S_3fstop_BCKH/BCKK carried the full-length cDNA of gene *SlHSP70-1* (*Solyc06g076020.2.1*) and a kanamycin resistance gene (*NPT*II) for selection were driven by the 35S promoter (Figure 4.1). The plasmid was inserted by an electroporation technique into the LBA4404 strain of *Agrobacterium tumefaciens* (Agrobacterium). Bacterial cells were cultured in liquid Luria Broth (LB) at 28$^{o}$C in 1 hour then collecting cells by centrifugation at 3,000 rpm for 5 min and spread in LB agar 1% (w/v) added 50 mg/l Kanamycin. The colony in petri then was genotype validated by polymerase chain reactions (PCR) with specific primer for *NPTII.* Colony with positive signal for *NPT*II was re-cultured in the liquid LB medium added 100 mg/l Kanamycin at 28$^{o}$C in shaking incubator until $OD_{600}$ = 0.6-0.8 for co-cultivation with tomato explants or making glycerol stock for store in -80$^{o}$C until using.

One day before inoculation, growing *Agrobacterium* from colony or glycerol stock of *Agrobacterium* contained a binary vector in LB medium added 100 mg / L kanamycin for 24 hours at 28°C until $OD_{600}$ reach 0.6-0.8. Centrifuge the bacterial culture at 3,000 rpm at room temperature for 5 min, then discard the supernatant and resuspend the pellet at in infection medium (pH 5.8) consisting of 1.2 g sucrose, 10 μM mercaptoethanol and 100 μM acetosyringone.

### *4.2.2. Transformation*

The *Agrobacterium* strain LBA4404 carried *SlHSP70-1* gene were transformed into tomato through cotyledon explants. Cotyledons of 7-day-old seedlings were cut into two halves then dipped and soaked in Infection medium in 10 minutes without shaking. After that, taking explants out and remove the excess bacteria by absorption on sterilized paper towel in 5 minutes. The explants then were place in co-cultivation petri plate then in dark for 2 days at 24°C. After 2 days of co-cultivation, calli were transferred to petri plate of MS basal medium with 30 g/L, sucrose, 1.5 mg/L zeatin, 100 mg/L kanamycin, 375 mg/L, Augmentin and 3 g/L Gelrite at pH 5.8 at 24°C (16 h L/8 h D) to induce callus. Renew the medium every 10 days. After 3 weeks, transferred 5-7 callus to petri plate of Shoot regeneration medium containing MS basal medium adding 30 g/L sucrose, 100 mg/L kanamycin, 375 mg/L Augmentin, 1 mg/L zeatin and 3 g/L Gelrite at pH 5.8. 2cm long shoots were transferred to rooting medium containing half-strength MS basal medium, 15

g/L sucrose, 50 mg/L kanamycin, 375 mg/L Augmentin and 3 g/L Gelrite at pH 5.8 at 24ºC (16 h L/8 h D). Long shoots with lateral root were transferred to rock wo3ol supplied 1/500 Hyponex nutrient to grow to new transgenic plants. The process of plant transformation was followed by describing of (Shikata and Ezura, 2016).

### *4.2.3. Plant growth conditions*

Seeds of tomato (*Solanum lycopersicum* cv. Micro-Tom) were sterilized 0.5% (v/v) sodium hypochlorite solution in 10 minutes. After three times washing in sterilized deionized water (10 minutes each rinsing), seeds were germinated in sterilized deionized water in 2 days, before sowing in magenta box containing 70 ml MS basal medium with 30 g/L sucrose and 3 g/L Gelrite (pH 5.7). The boxes were placed in plant growth chamber under fluorescent light with 120 μmol/$m^2$/s irradiance and 16/8 (light/dark) condition at 24ºC for 1 week.

In the growth chamber, transgenic plants and wildlife were cultivated in soil under conditions of 120 μmol/$m^2$/s light density, 16/8 (light/dark) duration and 24ºC. Water and 1/500 Hyponex nutrient had supplied every two days.

### *4.2.4. Genotyping*

Genotyping transgenic plants were conducted by PCR on genomic DNA from leaves of transgenic and wild type controls using specific primers of selective gene (*NPT*II). Genomic DNA was isolated from the leaves by modified CTAB method (Doyle, 1991). Primers were used for PCR amplification of 700bp products of the selective gene were showed as in Table 4.1. PCR conditions were as follows: pre-incubation at 95ºC for 3 min, 35 cycles of denaturation at 95ºC for 30 seconds, annealing at 58ºC for 30 seconds, and extending at 72ºC for 30 seconds; and a final extension at 72ºC for 5 min. PCR products were electrophoresis analysed on 1.5% agarose gel and visualized under UV light of 254 nm wavelength.

### *4.2.5. RNA isolation and Real-time quantitative PCR analysis*

Total RNA from frozen tomato plant tissues were extracted using TRIzol Reagent (Ambion, US) and was treated with RNase-free with TURBO DNA free Kit (Invitrogen). Samples were collected from young leaf and stem of 17-days-old plants and ovaries at -2, 0, 2 and 4 day-after flowering (DAF), respectively. Young leaf and stem of *Sliaa9* mutant plants were collected at the same stage with those from transgenic plants and wild type plants for RNA extraction by the same method. The quality of RNA was assayed by electrophoresis

of 2.0 µl total RNA on 1.2% agarose gel. The cDNA was performed from 1µl total DNA free RNA using ReverTra Ace® qPCR RT Master Mix (Toyobo, Japan).

For real time PCR, 2.0 µl of 10-fold diluted cDNA was added in to 10 µl PCR reaction containing 10 µM each primer and 5.0 µl of Fast SYBR™ Green Master Mix (Thermo Fisher Scientific). To check the expression level of *SlHSP70* in leaves, stems, and fruits of tomato, the quantitative amplifications were conducted with specific primers for *SlHSP70*, while the ubiquitin gene was used as reference gene. To investigate the relationship of *SlIAA9*, *SlDELLA* and *SlHSP70-1*, the expression levels of these genes were evaluated on transgenic plant of *SlHSP70, Sliaa9* mutant and wild type plants then compared together by real time PCR using specific primers for each gene. The primers were designed by computer program OLIGO 7 Primer Analysis Software (Molecular Biology Insights, Inc. (DBA Oligo, Inc.) for Windows system and synthesized by Eurofinsgenomics corporation (https://www.eurofinsgenomics.jp) as in Table 4.1. The amplification was conducted as preheat at 95$^{o}$C for 3 min, 40 cycles of denaturation at 95$^{o}$C for 10 seconds, synthesis at 60$^{o}$C for 20 seconds, and final extension at 72$^{o}$C for 3 min on StepOnePlus™ Real-Time PCR System (Applied Biosystems™, US). The melting temperature of the product was determined at the end of PCR. The expression level of targeted gene was analyzed by comparative Ct method (ΔΔCt) using reference gene (Livak and Schmittgen, 2001). Relative expression of target gene was compared with WT control at same tissue.

### *4.2.6. Phenotyping*

The length of whole plant, internode and leaf of transgenic *SlHSP70-1* tomato and wild type at 60 days after sowing (DAS) were calculated. The number of leaves, internodes were counted from the ground to the first inflorescence of each plant. Transgenic tomato fruit size was measured and compared to controls. To measure fruit size at 0 days, 2 days 4 days, 12 days and 30 days after flowering (DAF), flowers were labelled and hand pollinated at the days of anthesis, then carefully eliminated all the petals, sepals and stamens to guard pistil and ovary. The morphology of the fruit is observed under optical microscope Stemi 2000-CS (Zeiss, Germany) with 50 folds of magnification and taken photo by camera AxioCam ERc5s (Zeiss). The size of the fruits was analyzed and processed by AxioVision software (Zeiss). To evaluate the effect of *SlHSP70-1* gene to parthenocarpy fruit set, the stamens of transgenic plants were removed at 2 days before anthesis to prevent self-pollination.

Emasculated flowers were tagged and followed the fruit development. Parthenocarpy fruit set rate was calculated by the number of parthenocarpy fruit per total emasculated flowers.

#### *4.2.7. Statistical analysis*

The significance of differences in expression level of *SlHSP70-1* gene between overexpressing (OE) and WT were tested by Student's *t-test* using Graphpad Prism 5.04 software (GraphPad Software, San Diego California, US) for Windows system. The difference between genotypes was estimated at P-value level by means value of all phenotypes. ($P \leq 0.05, 0.01, 0.001$) carried out on raw data. For each genotype, the data were shown as a mean value of each category ± Standard Deviation (SD).

## 4.3. Results

### *4.3.1. SlHSP70-1 gene transformation into tomato plants*

920 explants from tomato cotyledon were inoculated with the *Agrobacterium* solution. 424 calluses were created and transferred to shoot medium to induce 244 shoots. 117 shoots with roots were collected from rooting medium, in which 38 shoots appeared as lateral roots were transferred to the rock wool to grow in the Hyponex medium to create 38 individual transformant lines. PCR was performed with a specific primer for selective gene *NPTII* (sequences of primers were showed in Table 4.1) on the template of the genomic DNA to collect 26 lines carrying the transgene of the target and 12 lines without the insertion of the targeted gene (Figure 4.2). Flow cytometry analysis on the 26 lines of the transgenic tomato, using wild type as control obtained 13 transgenic lines with diploid genomics (2n) and 13 tetraploid genomics (4n) lines (Figure 4.3). Only transgenic plants carrying the 2n genome were selected for screening for the phenotypic evaluation. The diploid transgenic lines (T0) with the genome were named S1 to S13. Thirteen transgenic plants were achieved from 920 tomato cotyledon explants resulted approximately 1.41% at transformation efficiency.

### *4.3.2. Copy number analysis and homozygous screening transgenic plants for phenotyping*

Before phenotypic evaluation, we selected T0 transgenic lines with the strongest expression level of the targeted *SlHSP70-1* gene by quantitative real-time PCR. The RNA of T0 plants is collected from ovaries at the day of flowering. The reason we choose ovaries tissue for *SlHSP70*-1 gene expression checking because of its relation to *SlIAA9* and *SlDELLA* from vertical bioinformatics analysis based on the microarray data on ovaries.

With two replications, we decided to select the lines S10 and S13 with the relative expression of the *SlHSP70* gene that outperformed the rest of the transgenic lines (Figure 4.4). The transgenic lines were grown, and T1 seeds were collected to continue growing the next generation (Figure 4.5). 48 T1 generation transgenic plants were genotyped by PCR reactions with NPTII-selective primers. Only 33 samples for PCR-positive signals, expressed by PCR products, are displayed on agarose gel under ultraviolet light 254 nm. In order to verify the genotyping results, we collected those plants, planted, and collected seeds to verify the genetic separation of T2 generation. In T2 genetically modified plants, the copy number and homozygous genotype were determined. The seeds from Four individual plants of T1 generation were selected and collected for producing the next generation. This result confirms that the S13 line carries only single copies of the transgene (Table 4.2). Homozygous plants have no segregation of selective genes selected for phenotypic evaluation. The similar result was obtained from line S10.

### *4.3.3. Over-expression of the SlHSP70-1 promoted tomato internode elongation, but not made effects for leaf phenotypes*

To characterize the physiological functions of the targeted *SlHSP70-1* gene, a full-length cDNA of the gene was introduced into tomato cv. Micro-Tom. Thirteen transformants ($T_0$) regenerated from kanamycin-resistant callus were contained the targeted *SlHSP70-1* inserted gene. $T_0$ transgenic lines introduced mRNA of the targeted *SlHSP70-1* were screened to choose homozygous mutants (Figure 4.4). Homozygous mutant plants at $T_3$ generation of two transgenic lines (S10 and S13) were used for phenotyping during vegetative and reproductive stages.

The overexpressing *SlHSP70*-OE showed longer main shoot when compared to the controls (WT plants) (Figure 4.8, Figure 4.8). After two months of cultivation, average of plant height of the control plants was 137.4 ± 10.95 mm (mean ±SD), while average of the *SlHSP70*-OE (S13) plant height was 183.46 ± 46.26 mm (Figure 4.7 A), suggesting that the *SlHSP70*-OE plant shoots was significantly higher (133.6% (p=0.02)) of than WT (Figure 4.7 B). The transgenic line S10 was also higher (120% (p=0.01)) than WT (. Figure 4.8B).

For further analysis, the length of internodes (from cotyledon to the first inflorescence) in six independent plants of *SlHSP70*-OE were measured and then compared to WT at each same position. Length of nine internode positions of each *SlHSP70*-OE (line S13) as well as WT were compared. Internode of *SlHSP70*-OE tended to be longer than that of WT, in which,

the most remarkable difference in internode length between the *SlHSP70*-OE and WT was the fifth- and sixth internode length (Figure 4.9 A, B, C). Comparison of the sixth internode length of transgenic plants of line S13 was 147.2% (p=0.0023) and 143.2% (p=0.0065) greater than that of WT plants at the same position (Figure 4.9 C). Similarly, the transgenic line S10 showed longer internode than WT (Figure 4.10 A, B), especially the eighth internode with 207 % length of that in WT plant (p=0.0004) (Figure 4.10 D). The expression level of *SlHSP70*-OE in the stem of the transgenic mutants showed higher than that of WT (Figure 4.9 B, Figure 4.10 C). As high expression of *SlHSP70-1* gene leads to longer internode in transgenic plants, it suggested that this gene might promote the internode elongation by involvement of cell division or elongation in tomato.

Next, we compared the morphological phenotypes of *SlHSP70-1*-OE and *Sliaa9* mutant plants. The shoot length of *Sliaa9* mutants was significantly longer than that of WT with 135% greater in length (p=0.0014) under the same condition, while shoot length of *SlHSP70-1*-OE and *Sliaa9* mutants showed similar trend (p=0.0661) (Figure 4.11). It could generate a hypothesis that *SlHSP70-1* might also be involved in phytohormone signals in controlling the growth and development of tomato plants in a relationship with *SlIAA9* (Goda et al., 2004)

To investigate effects of the *SlHSP70-1* on leaf size and structure, true leaves from the first to the ninth leaf node position of the mutants at 30 DAF were compared with those of WT. The mutants and the WT have also shown similarities in leaf structures as well as leaf complexity at each leaf position (Figure 4.12 A, Figure 4.13A). There were no significant differences in leaf size between the *SlHSP70*-OE and WT plants (Figure 4.12 C, Figure 4.13 C) though the relative expression level of the *SlHSP70-1* gene in transgenic was double when compared with that of WT (Figure 4.12 B, Figure 4.13 B). It suggested that the *SlHSP70-1* expression level did not have clear influence on tomato leaf morphology.

### *4.3.4. The SlHSP70-1 showed no effects toward fruit set, formation and development*

In order to investigate effects of the *SlHSP70-1* on tomato fruit set, the first flowering period, the number of flowers and fruits were observed to calculate fruit set rate in the *SlHSP70-OE* and WT. The time for the first blooming of transgenic plants was shorter than that of WT (Figure 4.14 A). However, there was no significance of fruit set rate between SlHSP70-OE and WT (Fig. 4.14 D) though the number of flowers and fruits in WT plants were greater than that of transgenic plants (Figure 4.14 B, C).

For evaluating effects of the targeted gene on fruit formation, size of the transgenic fruits at zero, two and four DAF were measured and then compared with that of WT (Figure 4.15A). At each time period of fruit set, similar size was observed between the mutant and WT fruits (Figure 4.15 C, D). At four day-after anthesis, the level of *SlHSP70-1* was significantly higher in transgenic fruits than that in WT (Figure 4.15B), while fruit size did not show significant difference between the transgenic and WT fruits. It suggested that high expression of the *SlHSP70-1* gene was less contributed on fruit formation. Fruit size of *SlHSP70*-OE plants and WT was compared at 12 DAF and 30 DAF periods. There were no clear differences in fruit size of the *SlHSP70*-OE lines and WT. Taken together, the results suggested that the high expression of the targeted *SlHSP70-1* was not clearly contribute tomato fruit development (Figure 4.15E, F).

To evaluate the effect of *SlHSP70-1* gene to parthenocarpy fruit set, the stamens of transgenic plants were removed at 2 days before anthesis to prevent self-pollination (Figure 4.16A). The results showed that at 4 days after emasculation, the flowers without stamens stared to drop from plant (Figure 4.16B) lead to only peduncle with tagged label remain after flower dropping ((Figure 4.16B). Very few flowers did not drop after emasculation but their ovaries could not develop to become fruit (Figure 4.16D). In opposite, flowers of transgenic plants were self-pollinated (Figure 4.16E) could stay on plant (Figure 4.16F) after pollination and the ovaries of these flower could develop to become fully fruit with seed inside the pericarp (Figure 4.16G). This observation indicated that the over expression of *SlHSP70-1* cannot promoted the development of ovary into fruit without pollination. In another words, over expression of *SlHSP70-1* cannot induce the parthenocarpy fruit set in tomato.

### *4.3.5. Relationship analysis of SlIAA9, SlDELLA and SlHSP70-1*

To assess relationships between *SlIAA9*, *SlDELLA* and *SlHSP70-1* for plant growth and development, we quantified the expression level of *SlIAA9* and *SlDELLA* genes in leaves, stems and fruit tissues of the transgenic plants of *SlHSP70-1* and WT. The results showed that the gene *SlIAA9* was highly expressed in the leaves of the transgenic plants compared to WT plants ($p=0.016$), while mRNA content of *SlIAA9* showed no significant changes in stem tissue ($p= 0.063$; Figure 4.17A). Meanwhile, the *SlDELLA* gene was also highly expressed in the leaves of the transgenic plants ($p=0.0286$), but not in the stem of the transgenic plants ($p=0.0168$) (Figure 4.17B).

To evaluate the role of *SlIAA9* in the expression of *SlHSP70-1*, we quantified the expression level of *SlHSP70-1* gene in the leaves and stems of the *Sliaa9* mutant plants, and subsequently compared with that of WT. The *SlHSP70-1* gene in the *SlHSP70-1*-OE and that in *Sliaa9* were highly expressed in both tissues than in WT (Figure 4.17C). This result implied hidden relationships of *SlIAA9* and *SlHSP70-1* genes in leaves and stems of tomato plants. The result may also suggest that the expression level of the *SlHSP70-1* gene was likely to be suppressed in the presence of *SlIAA9* gene. This gene expression seemed to be up-regulated when the *SlIAA9* level was down-regulated (Figure 4.17C).

**4.4. Discussion**

The targeted gene *SlHSP70-1* was belonging to the *HSP70* family in tomato was expressed in most of organs with various expression levels (Duck et al., 1989) (Table 4.3 and Figure 4.18). The abundant exhibition of the targeted *SlHSP70-1* in both vegetative and reproductive tissues suggested that the gene family is likely to be play roles in the tomato growth, development and fruit ripening (Duck et al., 1989). There were 18 genes encoding members of the *AtHSP70* family in which 14 genes are classified in the *DnaK* subfamily, while 4 genes in the *Hsp110*/*SSE* subfamily (Sung et al., 2001b). Of these, two orthologous genes of *SlHSP70-1* in Arabidopsis were elucidated to uncover their physiological roles for plant growth and development, senescence, responding to immunity and heat shock tolerance (Sung et al., 2001b; Noël et al., 2007; Clement et al., 2011; Li et al., 2016). The two genes, *AtHSP70-1* (*AT5G02500.1*) and *AtHSP70-2* (*AT5G02490*.1), exhibited 80% identity in genomic sequence and 92% in amino acid sequence. Moreover, both genes and the targeted gene showed high similarity in amino acid sequence (80%) in genomic level the target gene *SlHSP70-1* (*Solyc06g076020.2.1*) was identity at 80% with gene *AtHSP70-1* (*AT5G02500.1*). It suggests that the *SlHSP70-1* may contribute other physiological events though further investigation should be required.

The internode elongation length of the *SlHSP70-1-OE* and *Sliaa9* mutants may involve cell elongation and/or division in tomato stem. Cell division and elongation are closely related to phytohormone signals, particularly auxin and gibberellin, in which auxin signal can promote the activity of auxin responsive genes directly or via intermediates through gibberellin biosynthesis (Ross et al., 2000; van den Heuvel et al., 2001; Campanoni and Nick, 2005). The functions of *SlIAA9* and *SlDELLA* were revealed not only in process of fruit formation and development but also in the organ's growth of tomato (Wang et al., 2005;

Bassel et al., 2008). The *SlIAA9* was highly expressed in most organs of tomato plants (Koenig et al., 2013). The point and antisense mutants showed remarkable simplification in leaf structure and internodes elongation in tomato (Wang et al., 2005; Zhang et al., 2007). *SlDELLA* also contributes morphological changes in leaf morphology and internode elongation (Bassel et al., 2008; Nir et al., 2017). In this study, the expression levels of *SlIAA9* and *SlDELLA* in *SlHSP70-OE* leaf was significantly higher than those in WT (Figure 4.16A, Figure 4.16B). The low expression level of *SlIAA9* in leaf was observed in WT tomato (Koenig et al., 2013). *SlDELLA* gene was expressed differently in distinct organs, in which, the expression level in leaf was extremely lower when compared with flowers and stems. The result implied that the high expression of the *SlIAA9* and *SlDELLA* in *SlHSP70-OE* leaf seems to maintain a stable morphology of leaves in tomato.

High expression of the *SlHSP70-1* gene in the *SlHSP70-1-OE* and *Sliaa9* mutants in stem was associated with internode elongation in tomato (Figure 8C). The result suggested that the expression of the targeted *SlHSP70-1* was probably contribute healthy stem elongation in tomato. Further investigations are necessary to understand how the *SlHSP70-1* gene make effects toward the specific phenotypes of *Sliaa9* and *Sldella* mutants.

**Table 4. 1. Primer list for PCR and real time PCR**

| **Primer name** | **Primer sequence** |
|---|---|
| NPTII-Fw: | 5’- ATGATTGAACAAGATGGATTGCAC -3’ |
| NPTII-Rv: | 5’- TCAGAAGAACTCGTCAAGAAGGCG -3’ |
| HSP70-Fw: | 5’- CAAAATGCCTAACCCGAAGCC -3’ |
| HSP70-Rv: | 5’- CGTAGTCCCCAAA-TCAATTCCGA -3’ |
| UBQ-Fw: | 5'- CACCAAGCCAAAGAAGATCA -3' |
| UBQ-Rv: | 5'- TCAGCATTAGGGCACTCCTT -3' |
| DELLA-Fw: | 5’- CTGATATGGCTGGTTGGGTACA -3’ |
| DELLA-Rv: | 5’- AGAAGAAGAAGAACCACAACCAG -3’ |
| IAA9-Fw: | 5’- TTGGTGTTGGGGAGGAGGAG -3’ |
| IAA9-Rv: | 5’- AGGGCAAGTCTCCTCACCTC -3’ |

**Table 4. 2. Segregation of selective gene *NPTII* in T2 transgenic tomato**

| **Transgenic plants** | ***NPTII* segretation ratio (positive:negative)** | | **Total plants** |
|---|---|---|---|
| | **Expected ratio** | **Real ratio** | |
| **Rep 1** | **19 : 6** | **20 : 5** | **25** |
| **Rep 2** | **18 : 5** | **14 : 9** | **23** |
| **Rep 3** | **19 : 6** | **18 : 7** | **25** |
| **$X^2$value** | 0.621752189 | | |
| **p-value** | 0.366402 | | |

**Table 4. 3. Gene expression level of targeted *SlHSP70-1* in each tissues of *tomato S. lycopersicum* M82 and *S. pennellii***

| Species | Tissue | Expression Level | Standard Deviation |
|---|---|---|---|
| *Solanum lycopersicum* var. M82 | Floral | 12.83 | 0 |
| *Solanum lycopersicum* var. M82 | Stem | 62.5 | 0 |
| *Solanum lycopersicum* var. M82 | Leaf | 12.22 | 0 |
| *Solanum lycopersicum* var. M82 | Vegetative meristem | 8.45 | 0 |
| *Solanum lycopersicum* var. M82 | Seedling | 22.44 | 0 |
| *Solanum lycopersicum* var. M82 | Root | 29.62 | 0 |
| *Solanum lycopersicum* var. M82 | Mature Fruit | 50.18 | 0 |
| *Solanum lycopersicum* var. M82 | Developing Fruit | 431.67 | 0 |
| *Solanum pennellii* | Floral | 5.47 | 0 |
| *Solanum pennellii* | Stem | 9.12 | 0 |
| *Solanum pennellii* | Leaf | 3.1 | 0 |
| *Solanum pennellii* | Vegetative meristem | 19.57 | 0 |
| *Solanum pennellii* | Seedling | 22.88 | 0 |
| *Solanum pennellii* | Root | 34.35 | 0 |
| *Solanum pennellii* | Mature Fruit | 232.5 | 0 |
| *Solanum pennellii* | Developing Fruit | 313.29 | 0 |

The seedlings at 3 days after sowing on plate. Shoot and roots tissue from seedling at 10 days after sowing on plates. Vegetative meristems were collected when the 3rd leaf reached 1 mm. Stem between 4th and 5th leaves and inflorescent meristem were collected at 50 days after germination for *S. lycopersicum* var. M82 and 56 days after germination for *S. pennellii*. Young green fruits and mature fruits were collected from plants in the greenhouse

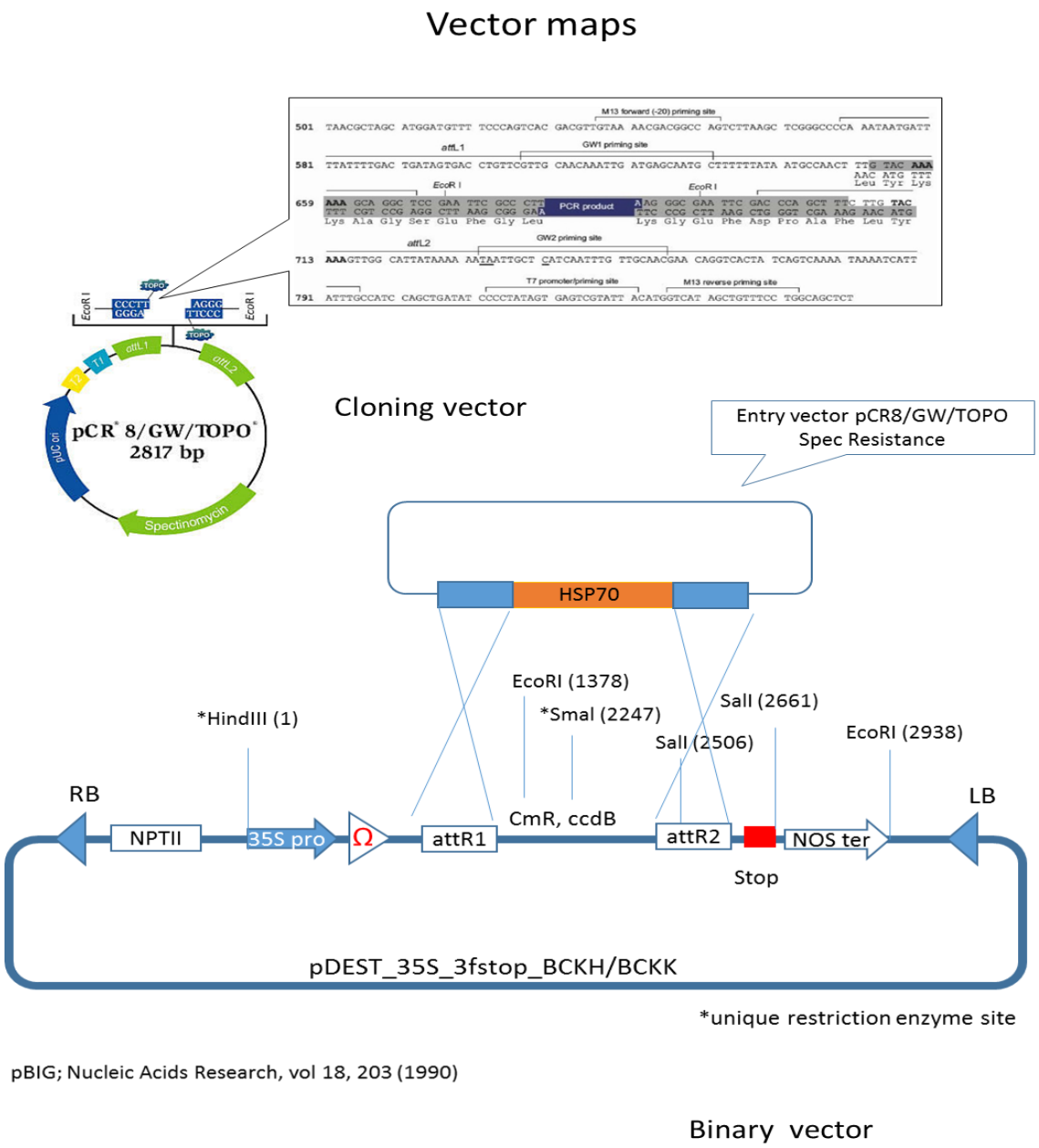



**Figure 4. 1. Vector map of *SlHSP70-1* transformation into tomato**

rsnaaa

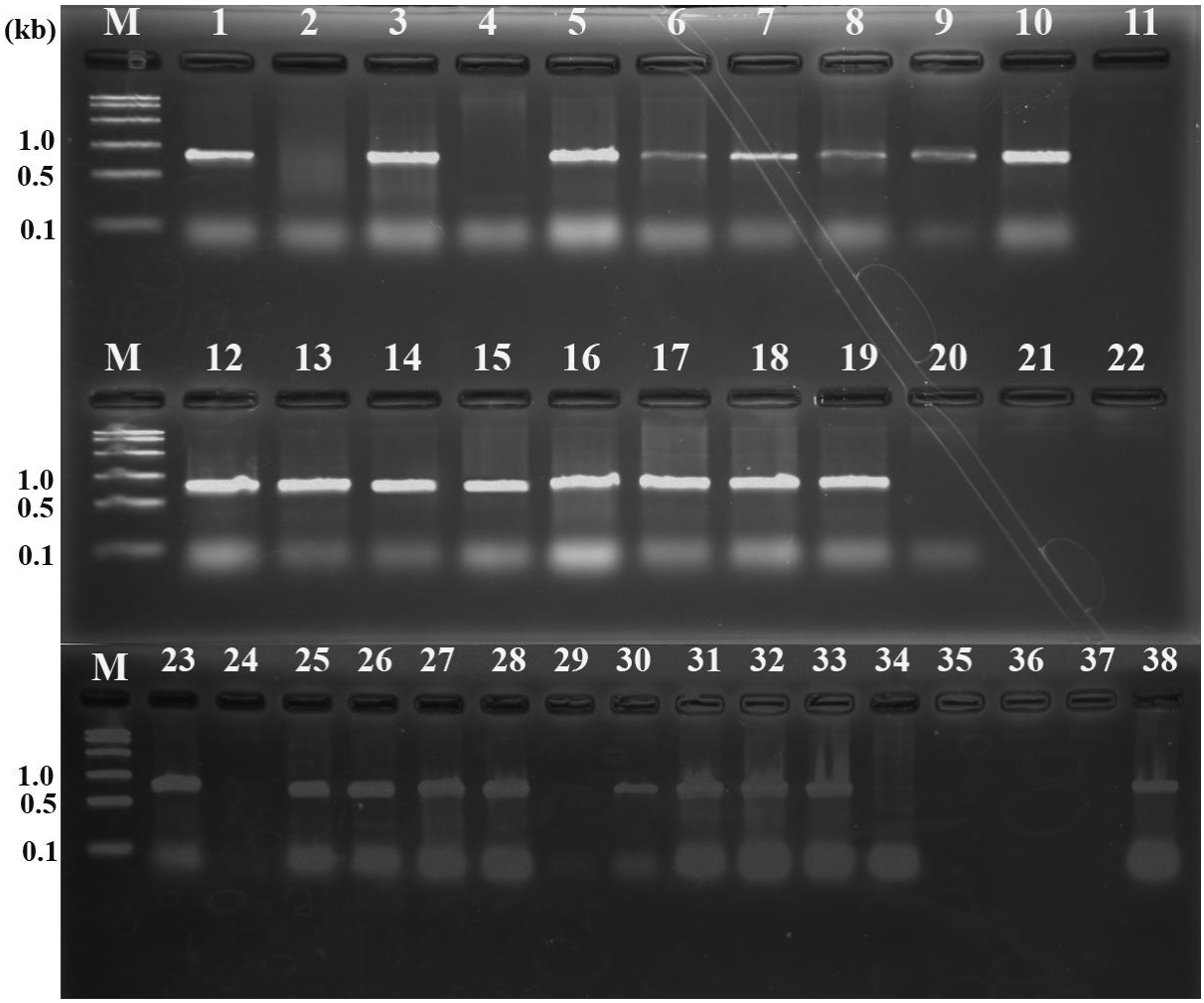



**Figure 4. 2. Screening positive line of tomato transgenic plants T0 generation**

M: DNA marker Fast2. Lane 1-38 represented to 38 lines of plants obtained after transformation. Well 1, 3, 5-10, 12-19, 23, 25-28, 30-33 and 38 showed positive signal of PCR products indicated that those plant brought transgene of *SlHSP70-1*. The rest of wells indicated that plants did not bring transgene *SlHSP70*-1.

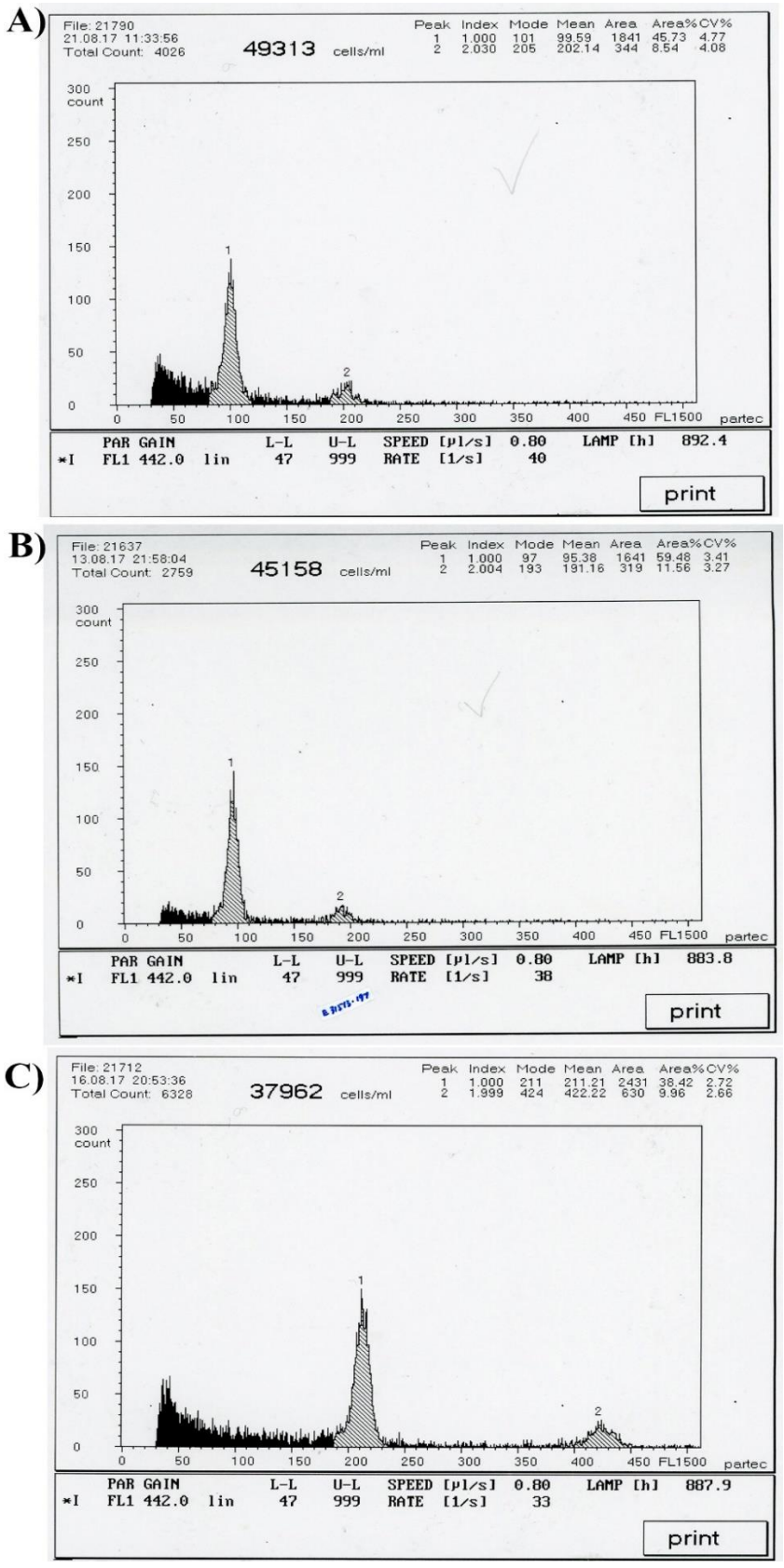



**Figure 4. 3. Flow-cytometer test for polyploid analysis of transgenic tomatoes**

A) wild type (2n); B) diploid (2n) transgenic tomato; C) tetraploid transgenic tomato.

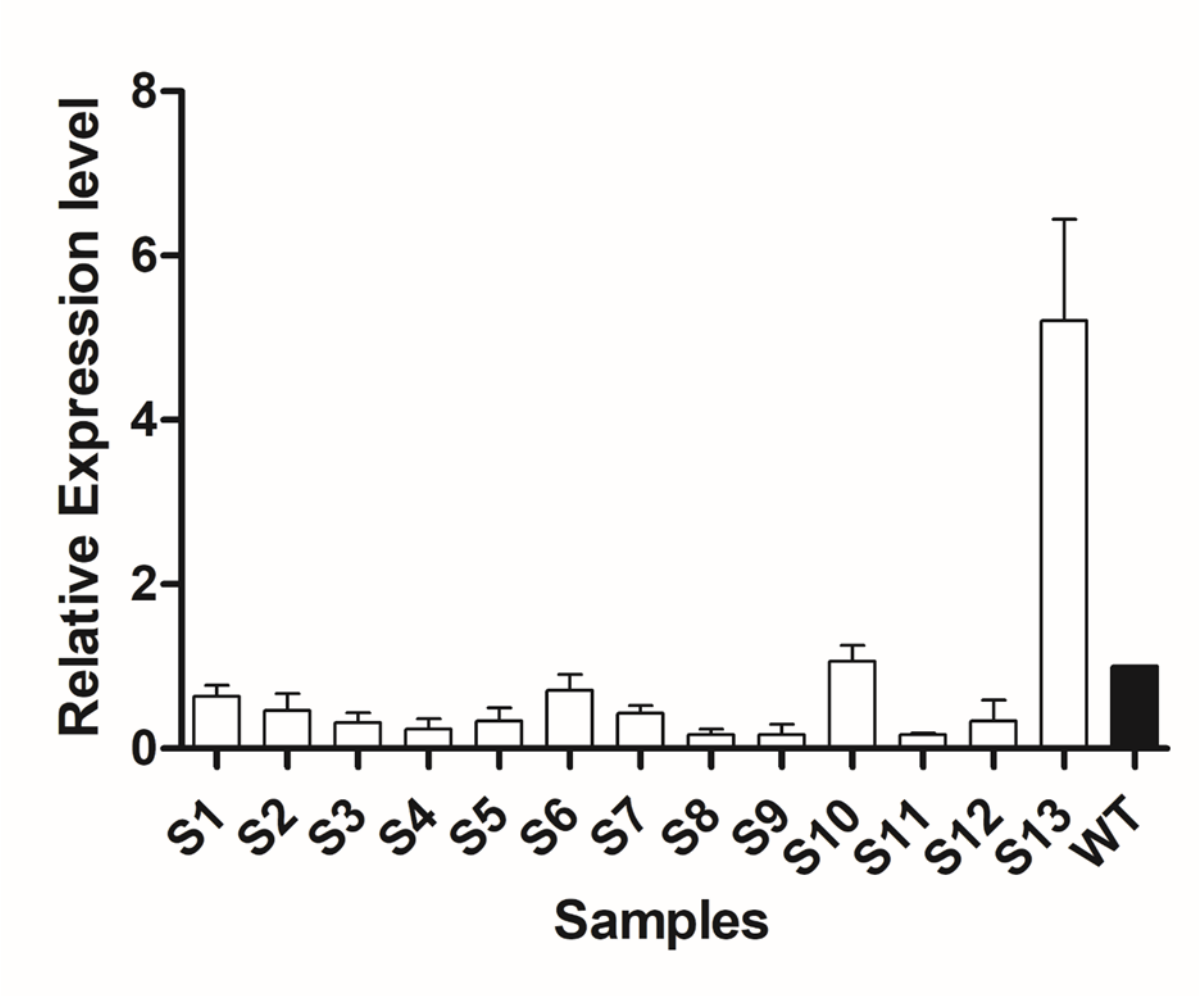



**Figure 4. 4. Relative expression level of *SlHSP70-1* gene in ovaries of T0 generation of transgenic plants.**

S1-13 were 13 lines of transformants. WT: Wild type. Bar indicated the mean value ± SD from 2 biological replications

.

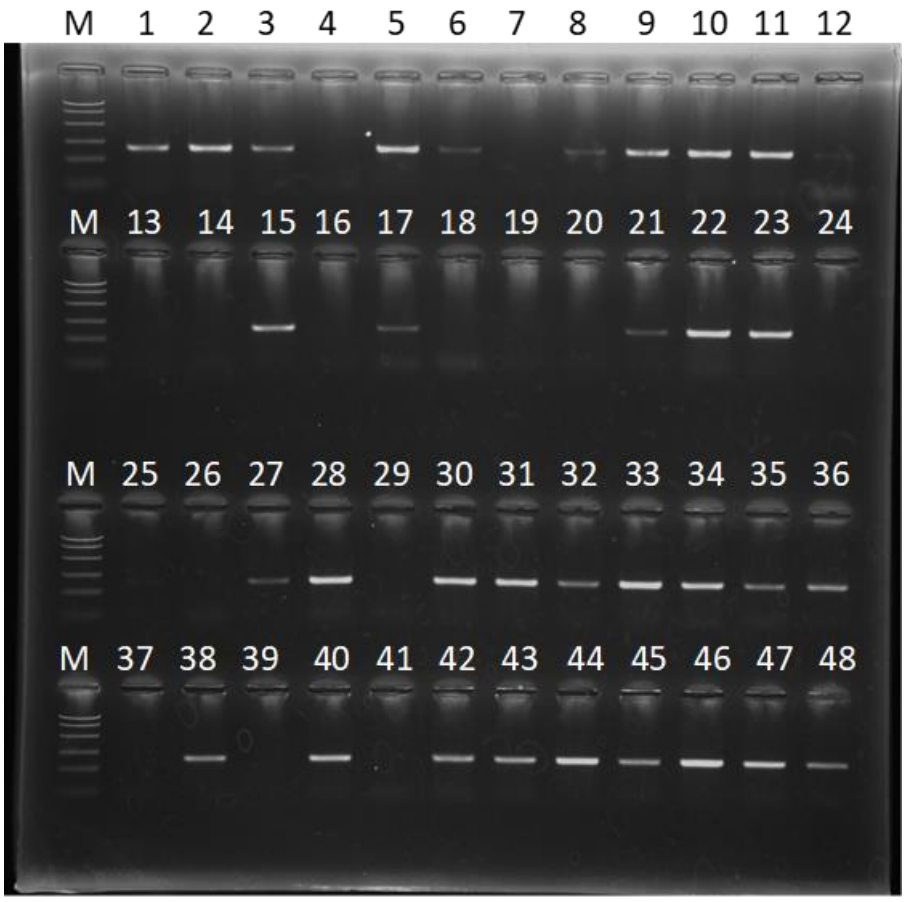



Figure 4. 5. Screening of T1 generation of transgenic tomato line S13

**Figure 4. 6. Screening T2 generation of transgenic tomato line S13**

A, B, C, D: four individual plants of SlHSP70-1 transgenic, T1 generation line S13. M: DNA marker ladder Fast2. Well 1-25 represented for individual plants of *SlHSP70-1* transgenic plants T2 generation which were progeny of A, B, C, D plants of *SlHSP70-1* transgenic T1 generation. The agarose eletrophheresis visualization showed that line A, B, D were heterozygous transgenic lines, line C was homozygous line of *SlHSP70-1* transgene.

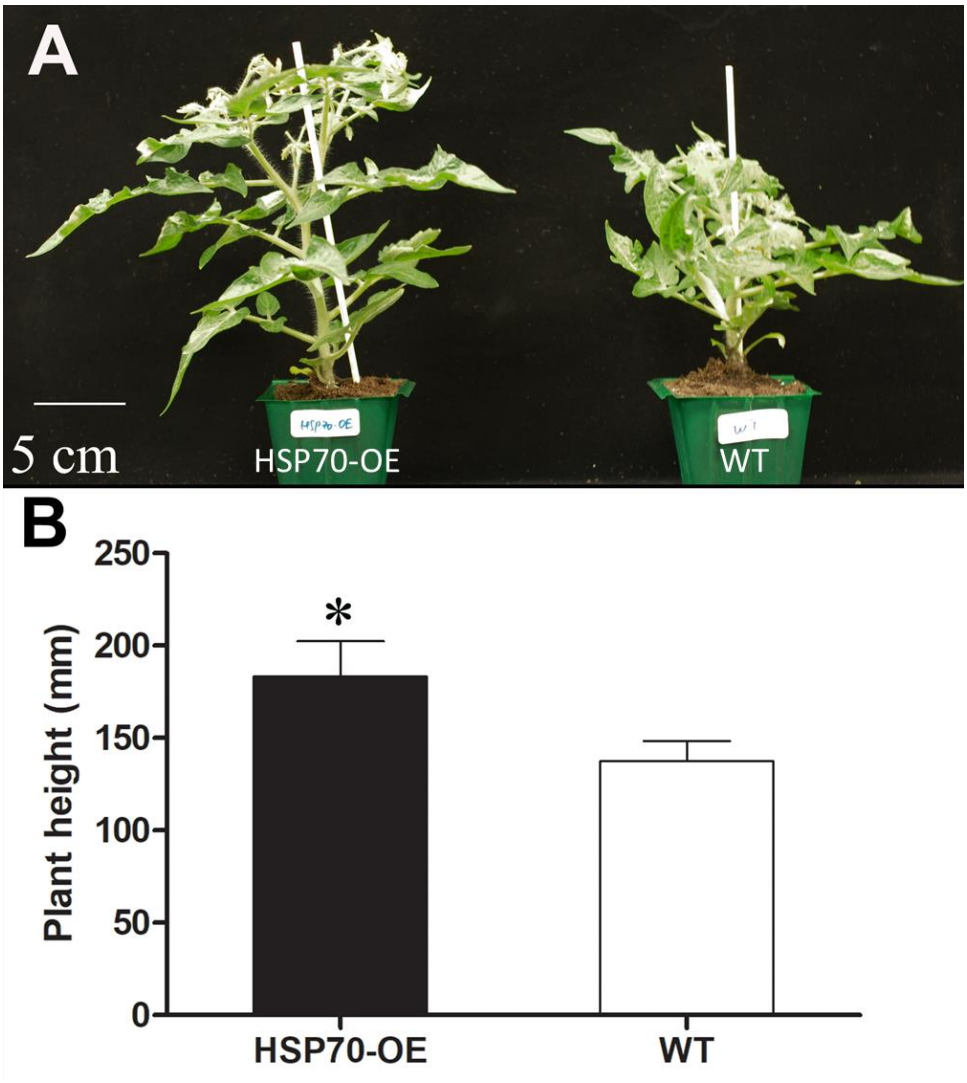



**Figure 4. 7. The *SlHSP70*-OE plant S13**

A) The height of tomato plant at 60 days after sowing (DAS). WT: wild type; *SlHSP70*-OE: Over-expression of *SlHSP70* plant. B) Mean values of plant height were compared between *SlHSP70*-OE and WT. Asterisks representative for significant difference with *P<0.05, **P<0.01 according to Student's *t-test* carried out on raw data. Bar indicated mean values of 6 biological replicates ± SD.

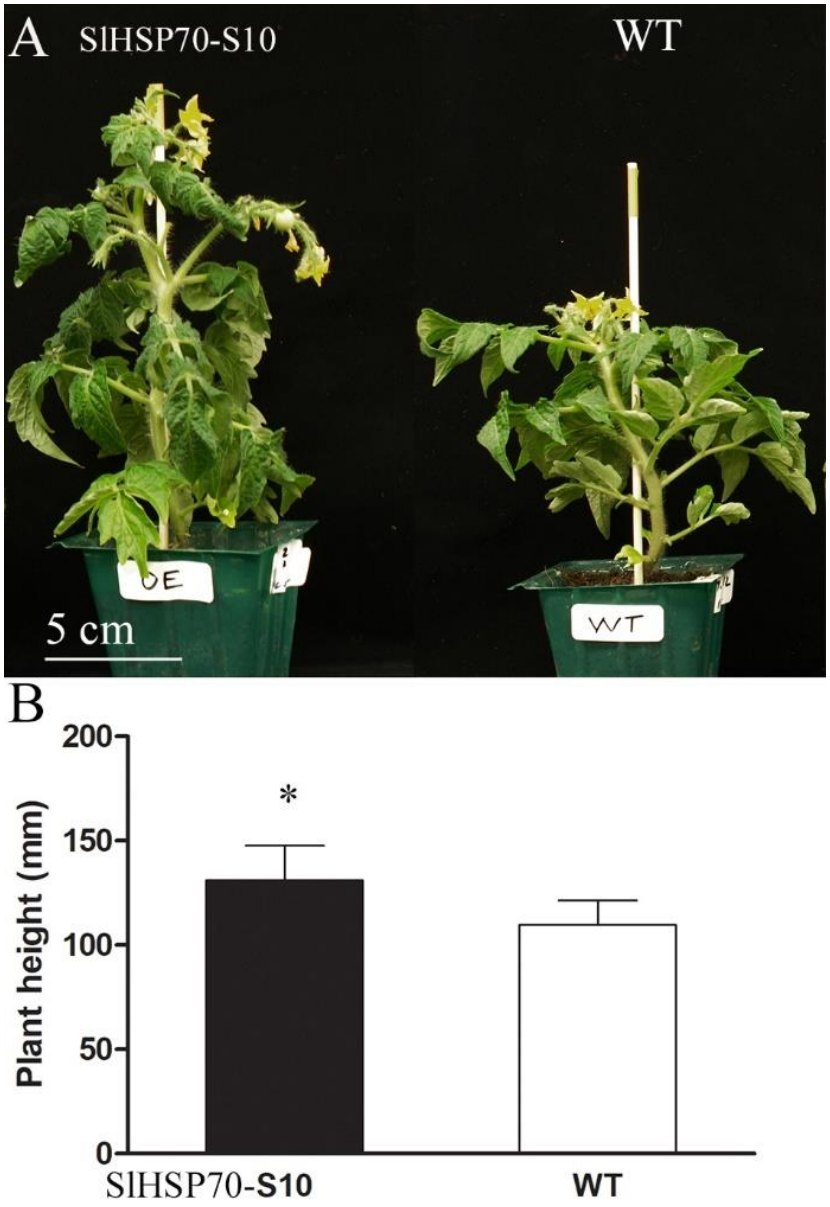



**Figure 4. 8. The *SlHSP70*-OE plant S10**

A) The height of tomato plant at 60 days after sowing (DAS). WT: wild type; *SlHSP70*-OE: Over-expression of *SlHSP70* plant. B) Mean values of plant height were compared between *SlHSP70*-OE and WT. Asterisks representative for significant difference with *P<0.05, **P<0.01 according to Student's *t-test* carried out on raw data. Bar indicated mean values of 6 biological replicates ± SD.

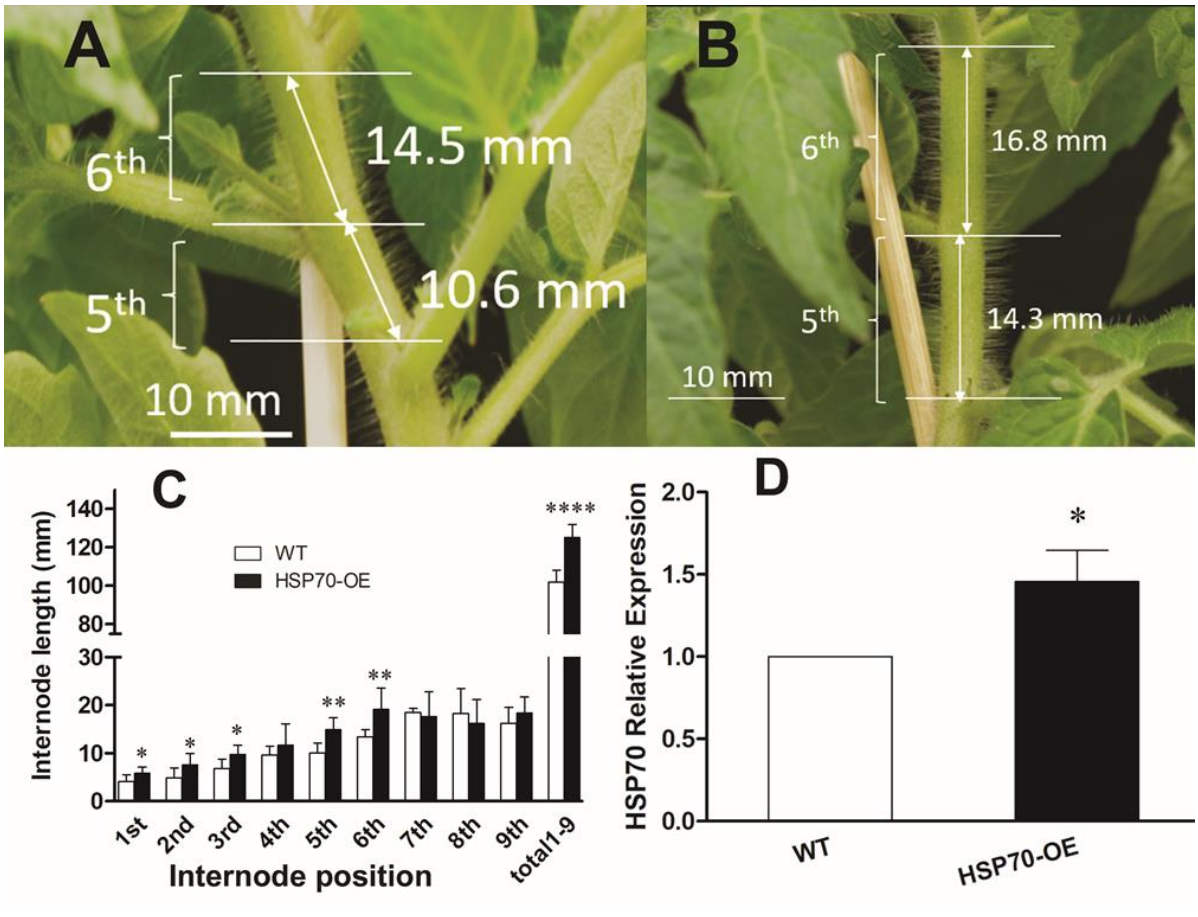



**Figure 4. 9. Internode length's comparison of *SlHSP70*-OE S13 and WT**

A) Morphological phenotypes of 5th and 6th internode of WT at 60 days after sowing. B) Morphological phenotypes of 5th and 6th internode of *SlHSP70*-OE plant at 60 days after sowing. C) Comparison the length of internode of overexpression *SlHSP70*-OE tomato and WT by internode position, respectively. Internode position was set from cotyledon to top of plant. Bar indicated the mean value of internode length of 6 independent plants for each genotype ± SD. D) Comparison of *SlHSP70-1* expression level from internode of *SlHSP70*-OE tomato and control (WT) by internode position. Bar indicated the mean value of 3 biological replications for each genotype ± SD. (*), (**), (****) representative for significant difference at P-value less than 0.05, 0.01 and 0.001 by Student's *t-test*, respectively.

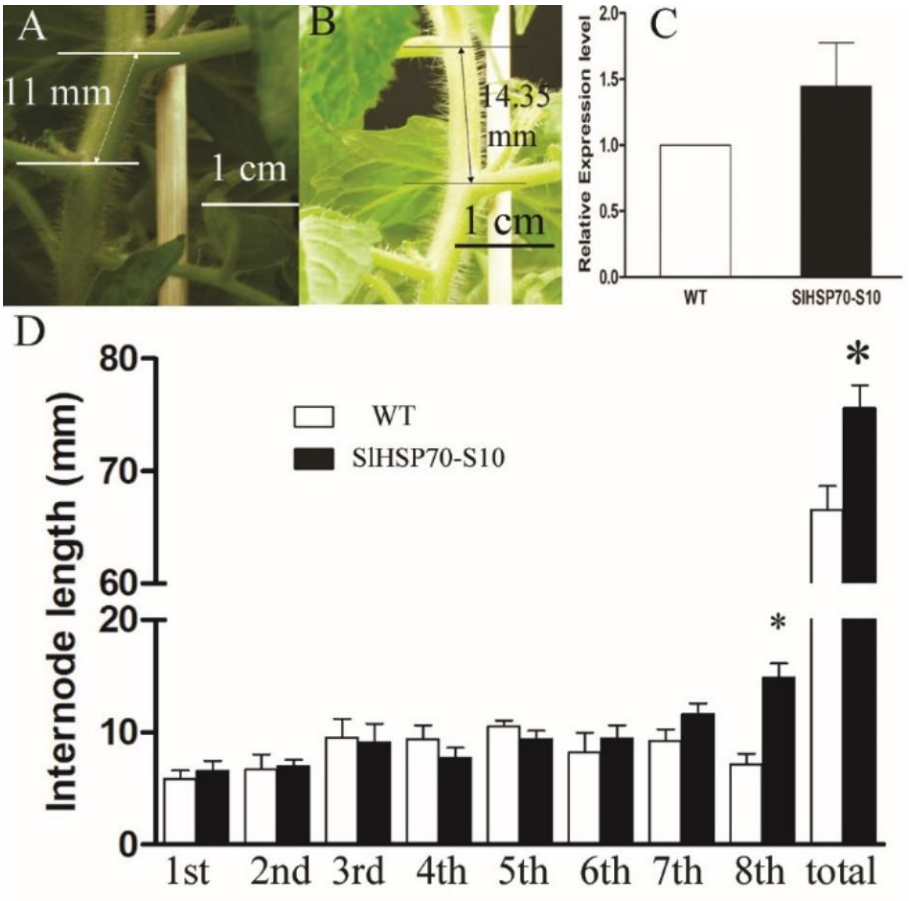



**Figure 4. 10. Internode length's comparison of *SlHSP70*-OE S10 and WT**

A) Morphological phenotypes of 8th internode of WT at 60 days after sowing. B) Morphological phenotypes of 8th internode of *SlHSP70*-OE plant at 60 days after sowing. C) Comparison the length of internode of overexpression *SlHSP70*-OE tomato and WT by internode position, respectively. Internode position was set from cotyledon to top of plant. Bar indicated the mean value of internode length of 6 independent plants for each genotype ± SD. D) Comparison of *SlHSP70-1* expression level from internode of *SlHSP70*-OE tomato and control (WT) by internode position. Bar indicated the mean value of 3 biological replications for each genotype ± SD. (*), (**), (****) representative for significant difference at P-value less than 0.05, 0.01 and 0.001 by Student's *t-test*, respectively.

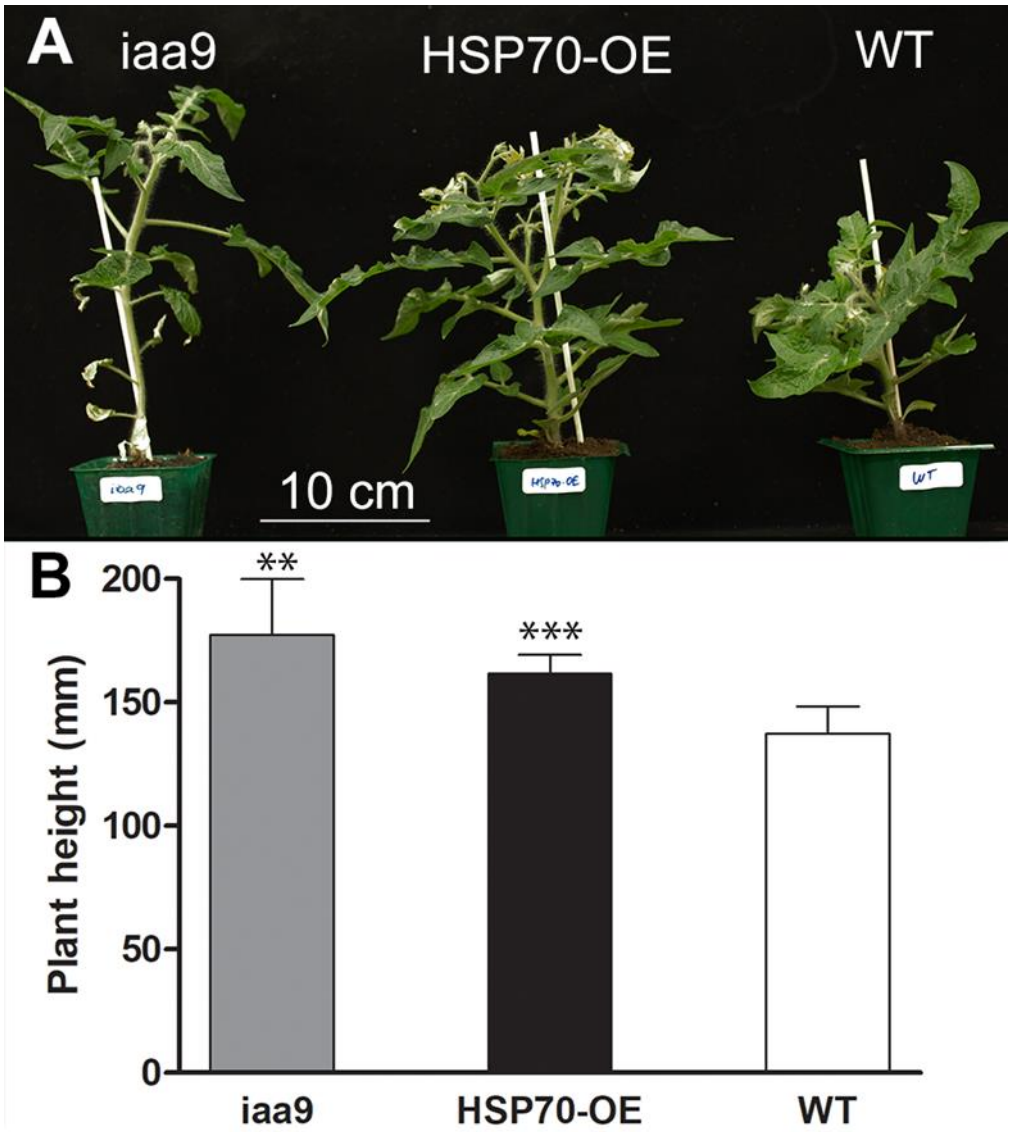



**Figure 4. 11. Plant height comparison of *SlHSP70*-OE, *Sliaa9* mutant and WT**

A) The length of tomato plant at 60 days after sowing (DAS). WT: wild type; iaa9: *Sliaa9* mutant line; *HSP70*-OE: Over-expression of *SlHSP70-1* plant. B) Mean values of plant heights were compared between *HSP70*-OE, *Sliaa9* mutant, and WT. Bar indicated mean values of 6 biological replicates ± SD Asterisks representative for significant difference with *P<0.05, **P<0.01, ***P<0.001 according to Student's *t-test* carried out on raw data.

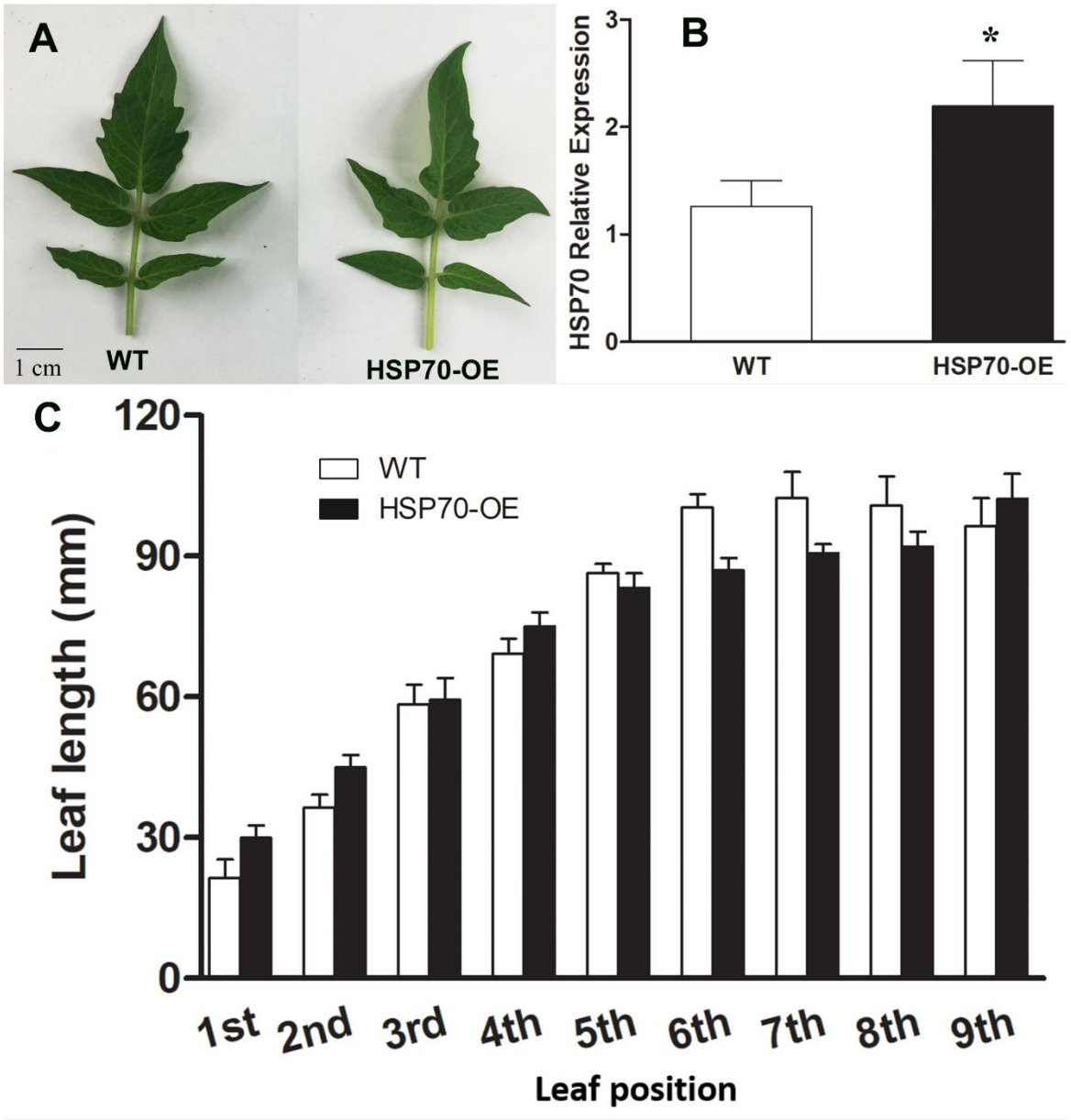



**Figure 4. 12. Comparison of leaves description between *SlHSP70*-OE S13 tomato and WT**

A) Morphological phenotypes of 6th leaves of *SlHSP70*-OE and WT plants at 30 DAS. Bar scale at 1 cm. B) Expression analysis of *SlHSP70-1* gene from young leaf of tomato at 17DAS. Bar indicated the mean value of internode length of 3 biological replications ± SD. C) Comparison of leaf length of tomato *SlHSP70*-OE and WT according to each leaf position. Bar indicated the mean value ± SD from 6 independent plants for each genotype. (*) showed the significance at $P<0.05$, by Student's *t-test*.

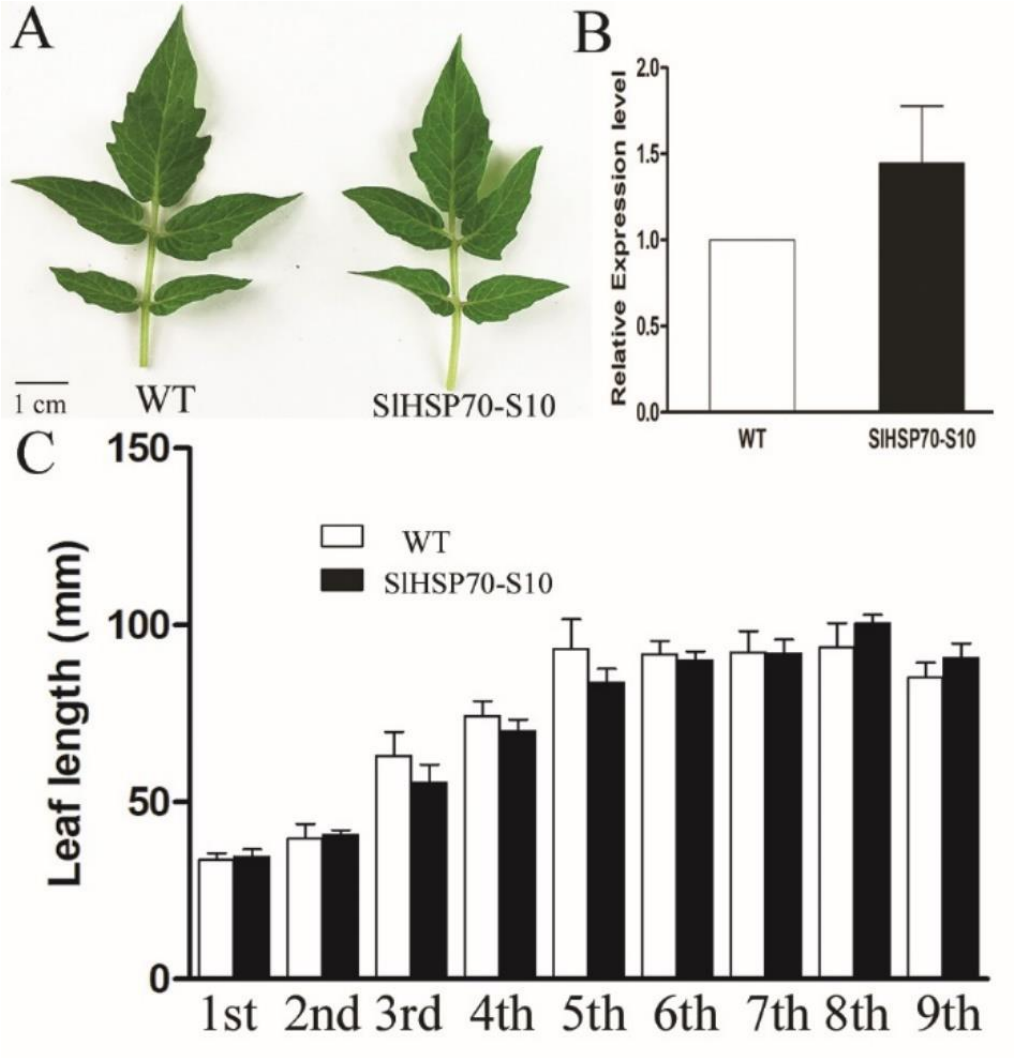



**Figure 4. 13. Comparison of leaves description between *SlHSP70*-OE S10 tomato and WT**

A) Morphological phenotypes of 6th leaves of *SlHSP70*-OE and WT plants at 30 DAS. Bar scale at 1 cm. B) Expression analysis of *SlHSP70-1* gene from young leaf of tomato at 17DAS. Bar indicated the mean value of internode length of 3 biological replications ± SD. C) Comparison of leaf length of tomato *SlHSP70*-OE and WT according to each leaf position. Bar indicated the mean value ± SD from 6 independent plants for each genotype. (*) showed the significance at $P<0.05$, by Student's *t-test*.

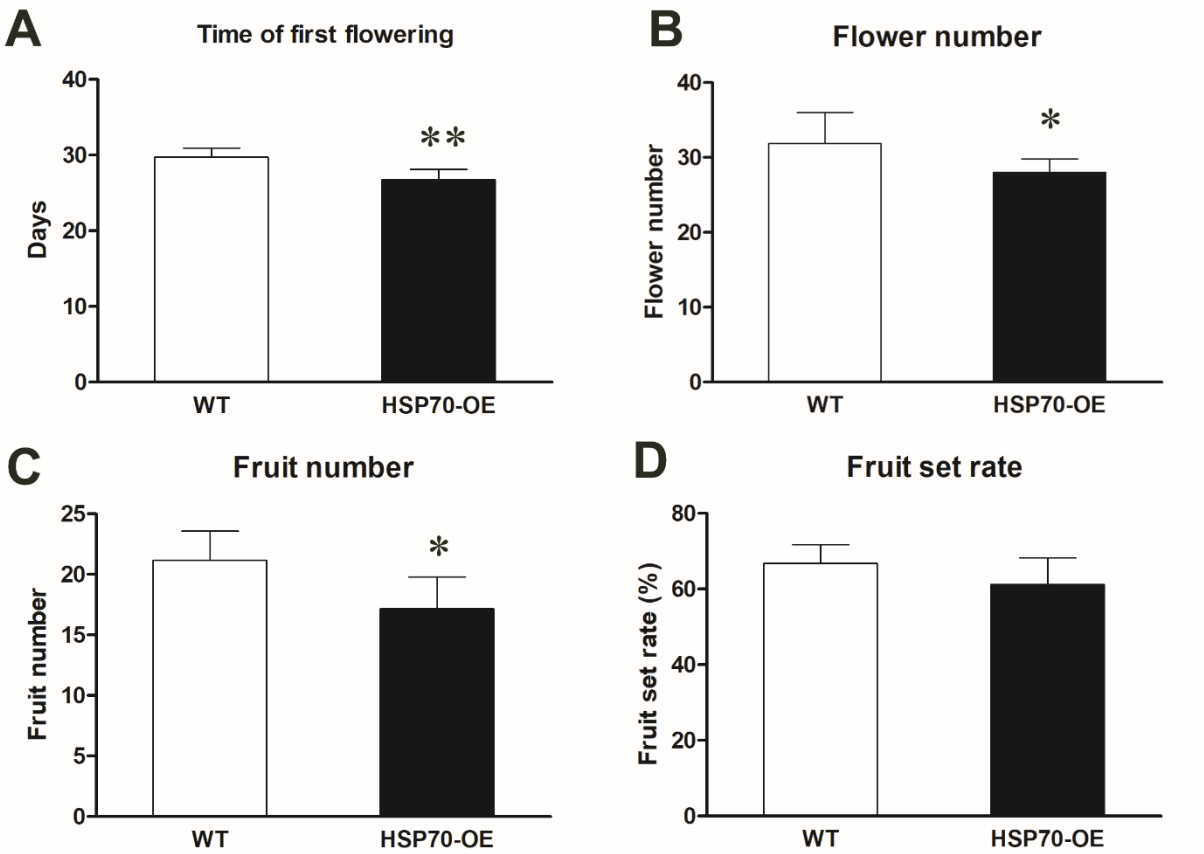



**Figure 4.14. Reproductive development of *SlHSP70*-OE tomato comparing to WT**

Comparison of A) time of first flowering; B) number of flowers; C) number of fruits; and (D) fruit set rate. Bar indicated the mean value ± SD from 6 independent plants for each genotype. (*), (**) showed the significance at $P<0.05$, $P<0.001$ by Student's *t-test.*

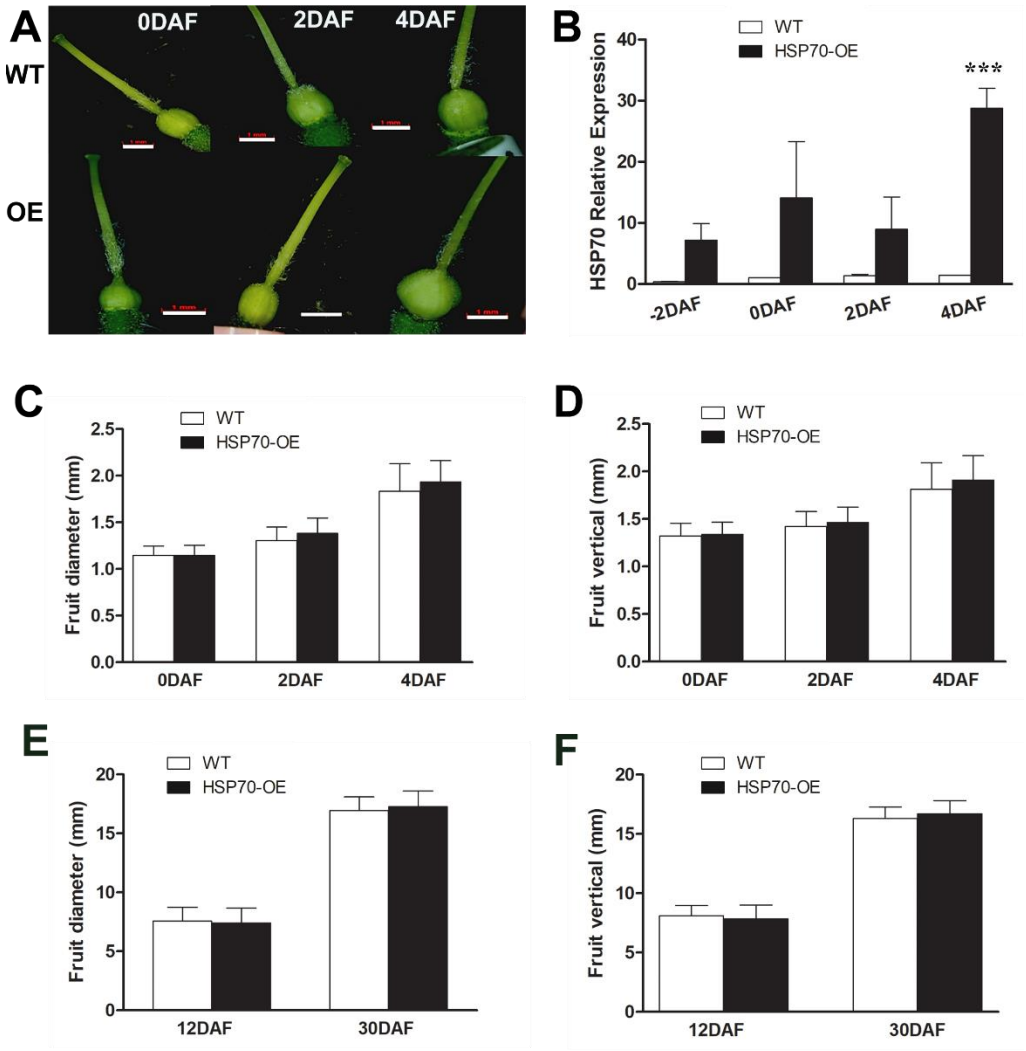



**Figure 4.15. Fruit development of SlHSP70-OE tomato**

A) Ovaries of tomato WT and *SlHSP70*-OE at 0, 2, 4 DAF, respectively. Ovaries were observed under microscope with 50x magnification. Photos were taken with bar scale at 1 mm. B) *SlHSP70* expression analysis from tomato ovaries at -2, 0, 2, 4 DAF, respectively. Bar indicated the mean value of HSP70 gene expression from 3 biological replications ± SD. The expression level of *SlHSP70* at 0DAF of WT was set as standard to represent its level at different time point. C, D) Comparison diameter and vertical size of *SlHSP70-OE* and WT fruit at 0, 2, 4 DAF, respectively. Bar indicated the mean of 36 fruits size values from 6 independent plants for each genotype ± SD. E, F) Comparison diameter and vertical size of *SlHSP70-OE* and WT fruit at 12 DAF and 30 DAF, respectively. Bar indicated the mean of 36 fruits size values from 6 independent plants for each genotype ± SD. (***) showed the significance at P<0.001, by Student *t-test*.

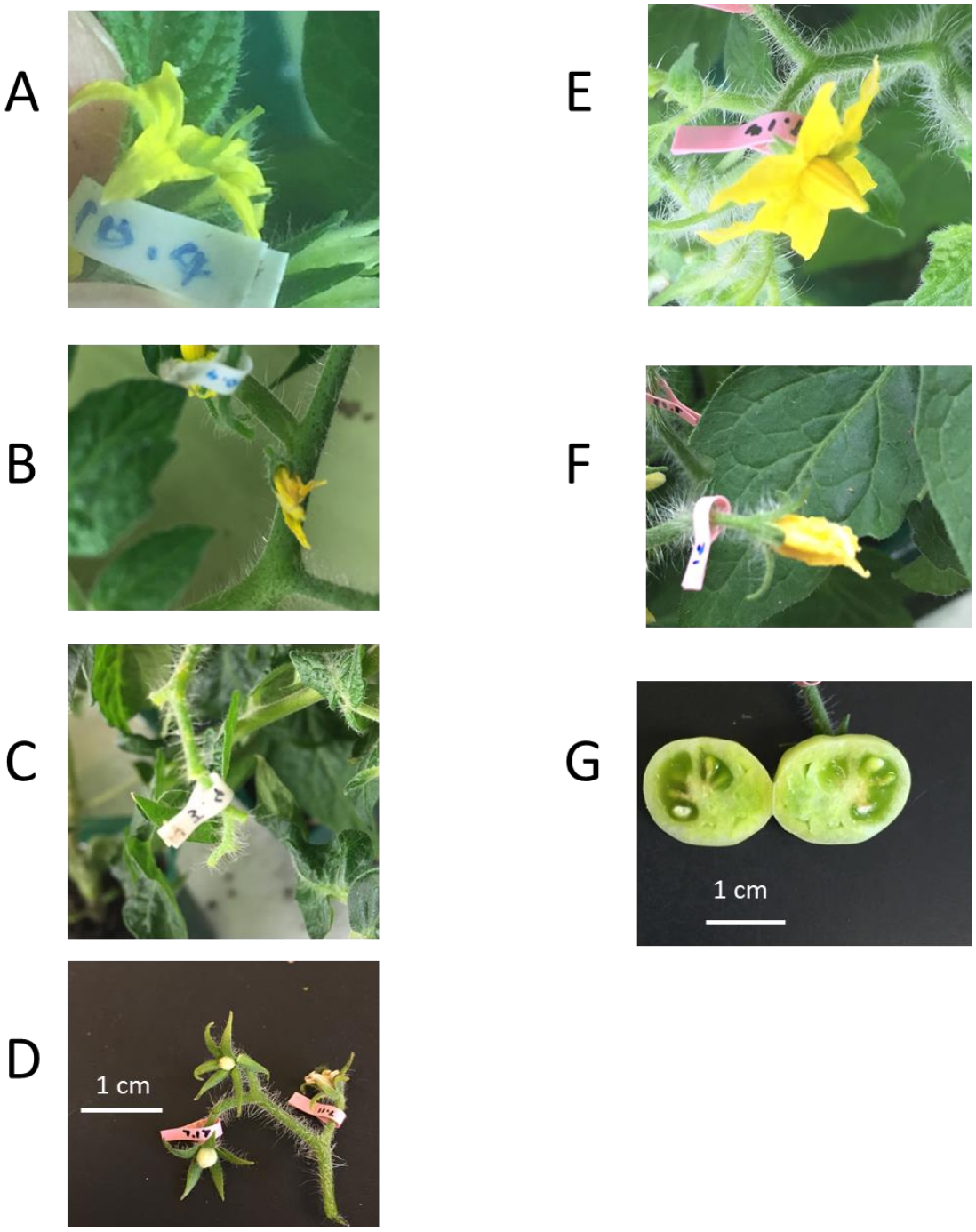



**Figure 4.16. Parthenocarpy fruit set capacity of SlHSP70-1-OE transgenic plants**

A) emasculated flowers at 1 day before anthesis; B) emasculated flowers dropped after 4 days of emasculation; C) peduncle after flower dropping; D) ovaries cannot develop after emasculation; E) self-pollinated flower at the day of anthesis; F) self-pollinated flower at 4 days of anthesis, G) Tomato fruit at 15 days after anthesis

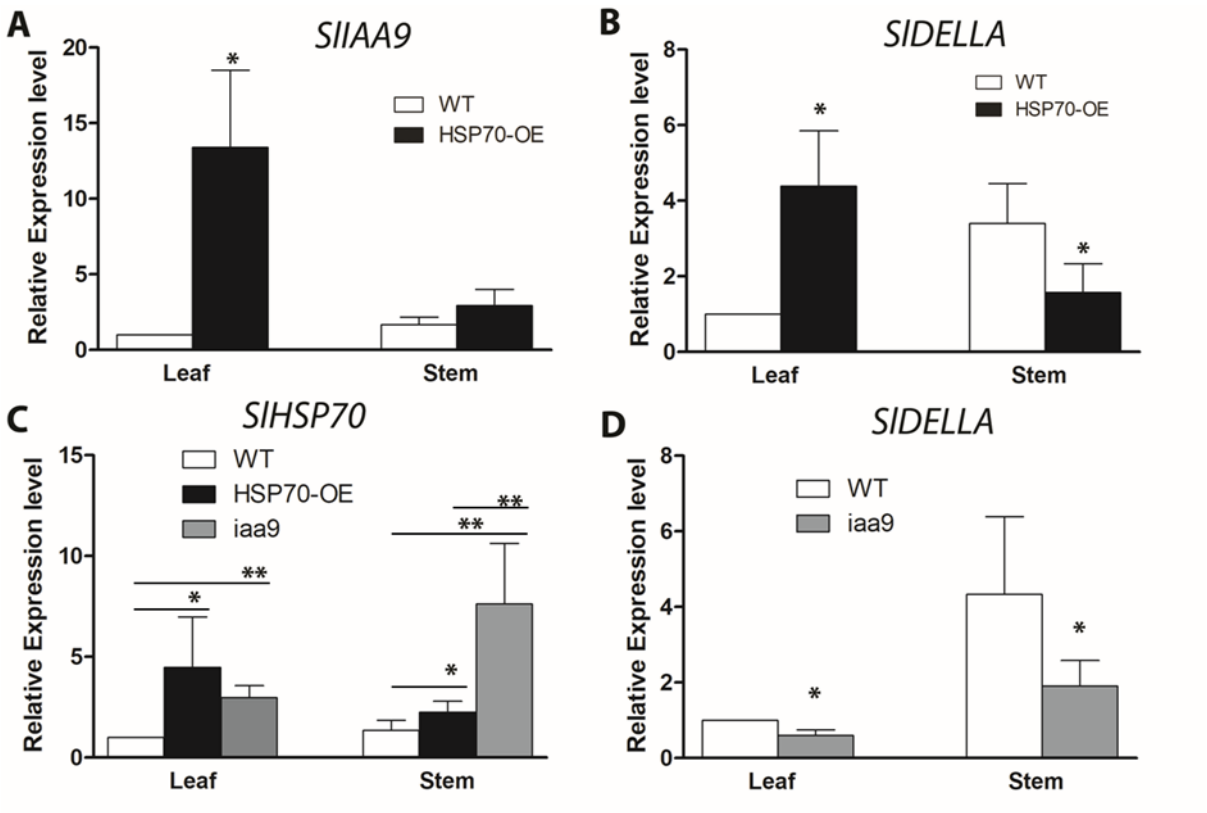



**Figure 4.17. Relationship analysis of *SlIAA9, SlDELLA* and *SlHSP70-1***

Expression analysis of A) *SlIAA9* and B) *SlDELLA* in leaves and stems of *SlHSP70*-OE tomato comparing to WT. C) Expression analysis of *SlHSP70* in leaves and stems of S*lHSP70*-OE, *Sliaa9* mutant tomato comparing to WT. D) Comparison of *SlDELLA* expression level in leaves and stems of *Sliaa9* mutant tomato comparing to WT. Bar indicated the mean value ± SD from 3 biological replications ± SD. (*), (**) showed the significance at P<0.05, P<0.001 by Student's *t-test.*

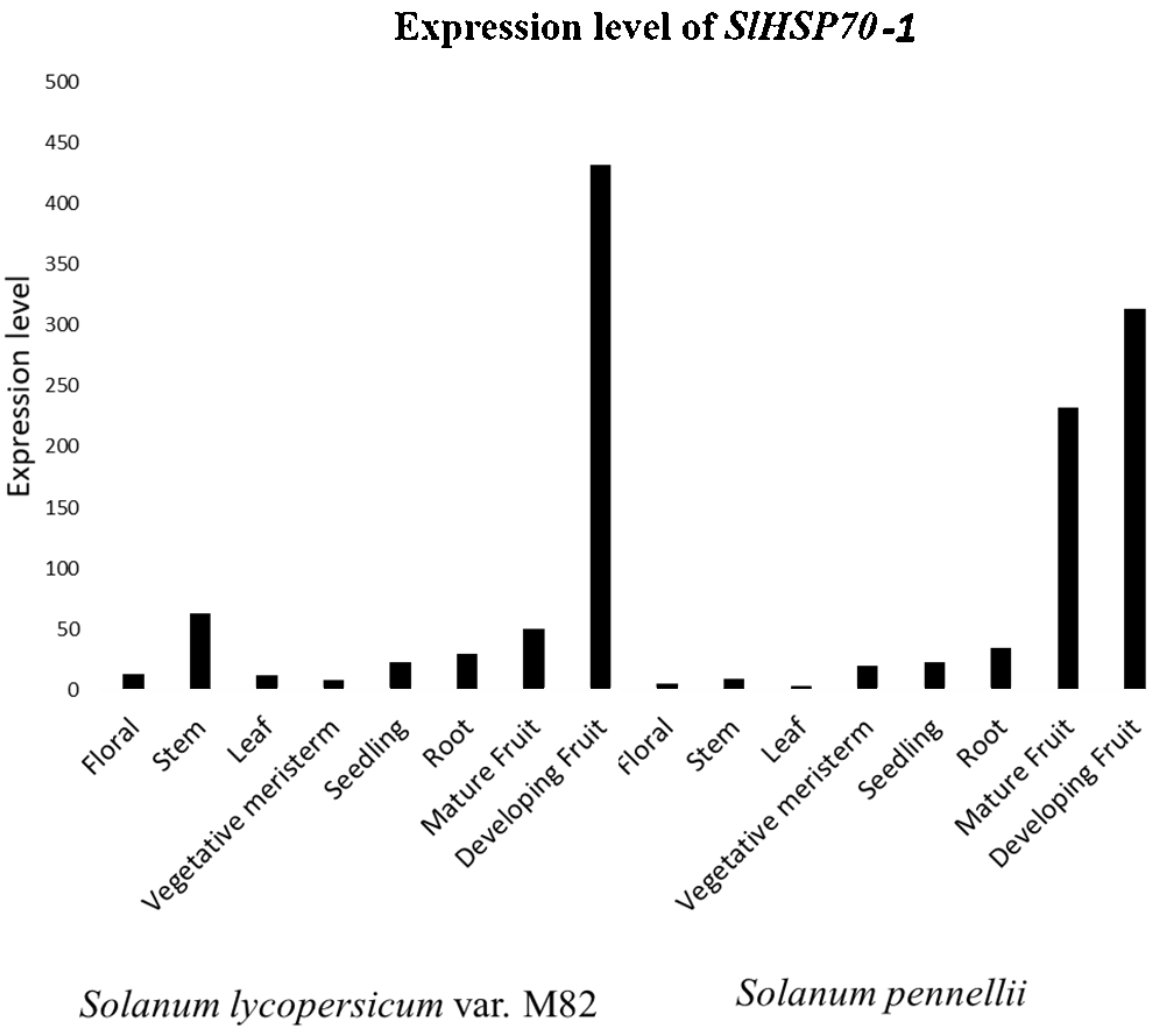



**Figure 4. 18. Gene expression level of targeted *SlHSP70* in each tissues of tomato *S. lycopersicum M82* and *S. pennellii***

The seedlings at 3 days after sowing on plate. Shoot and Roots tissue from seedling at 10 days after sowing on plates. Vegetative meristems were collected when the 3rd leaf reached 1 mm. Stem between 4th and 5th leaves and inflorescent meristem were collected at 50 days after germination for *S. lycopersicum* var. M82 and 56 days after germination for *S. pennellii*. Young green fruits and mature fruits were collected from plants in the greenhouse

# Chapter 5. Screening EMS mutations of tomato *SlHSP70-1* by TILLING and phenotypic characterization

## 5.1. Introduction

### *5.1.1. Tomato is a model system for reverse genetics analysis*

Tomato is one of highest valuable vegetable which is cultivated and consumed worldwide (Willcox et al., 2003). As an important food, the demand of tomato is increasing, which increases the need to improve tomato yield and production. The application of agricultural techniques for optimal development and nutrient improvement in agriculture has been employed; however, these approaches have been reached their limit, and productivity growth has become a challenge (Dumas et al., 2003; Erba et al., 2013). Some exogenous phytohormones have been applied as an effective solution. In tomato cultivation, several common phytohormones such as auxin, gibberellin, and cytokinin, are applied, among which auxin and gibberellin are the most commonly used (de Jong et al., 2009a; Matsuo et al., 2012; Hu et al., 2018). These hormones have been shown to stimulate a number of hormonal response genes related to tomato growth and development, in which *SlIAA9* and *SlDELLA* are the more remarkable genes (Wang et al., 2005; Goetz et al., 2007; Marti et al., 2007; de Jong et al., 2009b). The limitation of exogenous phytohormone applications come from its instability over a long time. New genetic approaches have the potential to select sustainable traits, and a large number of genes related to fruit set and production have been discovered (Ariizumi et al., 2013). The advancement of molecular biotechnology, especially the next generation sequencing (NGS) technology, have completely decoded the entire genome of tomato with of 950 Mb size on 12 individual chromosomes (Consortium, 2012). Most tomato genes have identifiers with specific ID; However, there were very limited understandings of their biological functionality and the generalization of genes is not fully described. Valuable resources for the research of gene function are found in gene mutations. A popular mutation could be induced by several agents including irradiation agents and chemicals (Oladosu et al., 2016). In tomato, more than 3,000 phenotypes were generated in the M82 tomato cultivar and collected via large EMS and fast neutron mutant collection methods (Menda et al., 2004). An EMS-induced mutation library of the tomato cultivar Micro-Tom was created, and it became valuable resource for tomato genetic studies

(Watanabe et al., 2007). Reverse genetics utilized EMS mutations as an excellent material for the investigation of gene functions by analyzing the changes in phenotype owing to the disorder in nucleotide sequence of the potential candidate gene.

### *5.1.2. Application of TILLING in reverse genetic analysis of tomato*

TILLING (Targeting Induced Local Lesions IN Genomes) is an effective and sensitive technique to identify point mutations caused by chemical mutagens responsible for functional analysis of the targeted genes (McCallum et al., 2000; Colbert et al., 2001). The knowledge of the gene nucleotide sequences was a unique requirement for TILLING screening. The main advantage is the ability to accumulate an allelic series of mutants with a variety of features, ranging from wild type to near function loss (Slade et al., 2005). TILLING showed its powerful strength when it applied to screen mutation from the EMS library, which contains a series transition mutation from A to T, G to C, and their reflection, respectively. The original principles of TILLING from the digested enzyme that recognized and cut any mismatches in the double strand of DNA were popular in EMS library. The Li-COR system was a high-resolution electrophoresis can improve the sensitivity, and it saves time and simplifies the screening process, which is an effective system for reversed genomics (Kurowska et al., 2011).

### *5.1.3. The objective of this study*

The previous chapter mentioned that the tomato *HSP70-1* (*SlHSP70-1)* has a function related to tomato growth and development. In this chapter, we applied TILLING for screening the *Slhsp70-1* mutant from the EMS mutation library. The phenotyping of the *Slhsp70-1* mutant under normal condition was conducted to obtain the phenotype changing. Any changes during tomato growth and development may address the biological function of the *SlHSP70-1* gene.

## 5.2. Materials and methods

### *5.2.1. Screening Slhsp70-1 mutant allele by TILLING technology*

To obtain *Slhsp70-1* mutant alleles, the TILLING technology was applied for screening the EMS mutant population. To attempt screening most mutations in the coding region of the *SlHSP70-1* gene, three pairs of primers were designed spanning the entire gene sequence with primer sequences as listed in Table 5.1. The primers labelled DY-681 or DY-781 are equivalent to IRDye 700 or IRDye 800 (https://www.biomers.net/), respectively.

The PCR amplification was conducted as follows: preincubation at 95°C for 3 min, 35 cycles of denaturation at 95°C for 30 s, primer binding at 58°C for 30 s, extending at 72°C for 30 s, and a final extension at 72°C for 5 min. PCR products were analyzed by electrophoresis using 1.5% agarose gel and visualization under UV light of 254 nm wavelength. The PCR products were double diluted with water then denatured at 95°C in 7 min and then cooled down to 12°C with ramp 0.1°C/s. The PCR product was then digested by the enzyme ENDO1 at 45°C for 20 min. The digested product was then purified through Sephdex G50 column and then loaded on 6% acrylamide gel for electrophoresis on a Li-COR system to visualize the fragments of the PCR that were digested and determine mutation. The process was absolutely followed as described previously (Okabe et al., 2011).

### *5.2.2. Plant Materials and Growth Condition*

After screening *Slhsp70*-1 mutations from the library, the seeds of tomato (*Solanum lycopersicum* cv. Micro-Tom) for each *Slhsp70-1* mutant were collected from the National BioResources (NBRP) Project at the University of Tsukuba (Saito et al., 2011)) with NBRP accession numbers as stated in Table 5.2. The plant genotypes were checked by TILLING. Heterozygous plants were growing and self-pollinated to produce the next generation for the next genotyping until homozygous mutant allele were obtained. From the M3 mutagenized population, homozygous mutants were used for phenotyping.

For phenotyping, the seeds of all homozygous mutant lines were germinated in sterile distilled water for two days at room temperature, and they were sowed into soil in a plastic pot (5 × 5 × 5 cm; length × width × height). The seeds of the tomato wild type (WT) and azygous (Az) line was germinated and sowed under same conditions as the control. Azygous lines were lines of tomato obtained from the segregations of heterozygous mutant lines of the targeted gene, and they did not contain any mutation in the targeted *SlHSP70*-1 gene. Hyponex 10-6-5 (Hyponex Tech, Japan) was supplied with 1/500 dilution every two days. All plants were grown under condition of 120 μmol/$m^2$/s light density, 16/8 (light/dark) duration, and 24°C.

### *5.2.3. Genomic DNA Extraction*

Genomic DNA was extracted from 30-day-old plants using the modified CTAB method (Doyle, 1991). A piece of fresh leaf sample was ground by a 5-mm-diameter bead using TissueLyser II (Qiagen, Germany). A spectrophotometer measured the DNA concentration (NanoDrop 2000c, Thermo Scientific).

### *5.2.4. Sequencing validation*

Sequencing analysis was carried out to verify the *Slhsp70-1* mutant mutation site according to the exome sequence result. The fragment of the gene amplified by PCR were used for TILLING. The PCR product after TILLING was analyzed via electrophoresis on 1.5% agarose gel and then visualized under 254 nm wavelength UV light. The products of PCR were purified by the Wizard ® SV Gel and PCR Clean-Up System (Promega, Madison, USA) following the procedure listed by the Manufacturer. The purified PCR product was then used for Sanger sequencing with primers for sequencing as listed in Table 5.3. The PCR mix reaction comprised 50 ng of template and 1 µl of primer (10 pmol), after which distilled water was added to reach a total volume of 21 µl. The Applied Biosystems/ABI 3730XL DNA Analyzer with 96-capillary (Thermo Fisher Scientific) was utilized for the sequencing analysis. The sequencing data were viewed using SnapGene Viewer software. The alignment of the nucleotide and amino acid sequences was analyzed using a parallel editor of GENETYX Ver. 11 software.

### *5.2.5. Phenotyping*

In order to investigate the effect of *SlHSP70*-1 mutant on the growth and development of tomatoes, we recorded the germination rate of the seed of each mutant line, measured the length of whole plant of *Slhsp70*-1 mutants, Az line, and WT line at 30 and 60 days after sowing (DAS), and then, we compared them together in pairs. The structure and size were evaluated and compared for 6 individual plants of each genotype (mutant, Az, and WT, respectively).

For fruit set and development analyzing, we recorded the time for first flowering (by DAS and the number of leaves at first flowering), counted the number of flowers, and the number of fruits, to calculate the rate of the fruit set. In order to analyze fruit growth, we evaluated the size of the fruit and compared it between the values for the mutant, Az, and WT lines. Data were collected from six individual plants of each genotype (mutant, Az, and WT, respectively)

### *5.2.6. Statistical analysis*

The significance of differences in the expression level of *Slhp70*-1 gene between mutants, Az, and WT were tested using Student's *t*-test using Graphpad Prism 5.04 software (GraphPad Software, San Diego California USA). The mean differences between genotypes

were estimated for all phenotypes according to the P value level (P <= 0.05, 0.01, and 0.001) carried out on raw data. Data were graphically represented as a mean value ± standard deviation (SD) of each category for each genotype.

## 5.3. Results

### *5.3.1. Isolation of the Slhsp70-1 mutants from the mutant population*

From 9216 samples of the library that were treated by EMS and double screening by TILLING, we collected 12 mutations (Figure 5.1 and Figure 5.2). The mutants and their name in EMS library were listed in Table 5.2. The position of each mutation in the gene is shown in Figure 5.3. For validation point mutation for each collected EMS line, we conducted the sequencing of *SlHSP70-1* gene by Sanger method. We recognized type of each mutation and presented them in Table 5. 4. The mutation frequency was 12/9216-0.13%, indicated that the expectation for mutation generated is 1.3 mutations per each kilobase of nucleotides on gene *SlHSP70-1*.

Two mutations in the gene were located in the first exon, seven mutations in the second, one mutation in the 5-untranslated region, two mutations in 3'-UTR, and no mutation in the intron region of the targeted gene. In terms of the effect of mutations, there were 8 missenses mutations obtained after screening, which replaced 1 nucleotide by another nucleotide, resulting in one amino acid that would be replaced by another amino acid in the primary structure of the protein. Four other mutations were identified as silent mutants that have no effect on the primary structure of the protein. No nonsenses mutation, which change a normal codon into stop codon that causes a shortened amino acid sequence, can be obtained from screening (Table 5.4). EMS is the causative agent that replaces nucleotide A by T, and replaced a nucleotide C by G, respectively. However, in this screening, only one point-mutations obtained changed from A-T, while there were six mutations from G-A, three C-T variables, and one A-C mutation. Table 5.4 and Figure 5.4 showed the distribution of mutation and the replacement of amino acids, respectively.

### *5.3.2. Slhsp70-1 mutants were not affected to tomato growth and development*

For phenotyping the mutants of *SlHSP70-1*, we requested the plant material of all missense's mutants from National BioResource Project to access the seed of the mutant. Only 4 mutant seeds were available in the library, named from *Slhsp70*-1-1 to *Slhpc70*-1-4.

These mutants were grown in a glasshouse and allowed to self-pollinate to produce the next homozygous mutant lines. TILLING was performed twice to screen and identify the homozygous lines of each mutant.

At first, we assessed the germination ability of the mutant seeds and compared them with the controls. The results showed that the germination rate of WT was about 80%, while the germination rate of the mutant lines varied from 42.75% in line *Slhsp70*-1-2 to 75.25% of the line *Slhsp70*-1-1 (Figure 5.5). There was a significant difference between WT and *Slhsp70*-1-1 ($p = 0.0003$), *Slhsp70*-1-2 ($p = 0.0017$), *Slhsp70*-1-3 ($p = 0.001$), and *Slhsp70*-1-4 ($p = 0.0014$), respectively; however, it was not significantly different between each mutant line and the Az line. This result indicated that the difference in the germination rate between the WT and EMS lines may not come from the effect of the *SlHSP70*-1 mutant. The germination rate of the mutant, Az and WT lines are shown in Figure 5.5.

Next, we evaluated the length of the plant for each mutant line and compared it with the control. The plant height of the mutant plant, Az and WT was measured at 30 and 60 DAS, and it is shown in Figures 5.6 A, B, C and D. At 30 DAS, the average height of the WT tree is about 90 mm while the mutant lines have a height ranging from 77.5 mm to 105 mm (Figure 5.6E). While at 60 DAS, the average height of the WT plant is about 157 mm and the mutant plants are high, from 127 mm to 167 mm (Figure 5.6F). The findings of the comparison showed that the height of the WT plants and the mutant plants differed significantly. However, this difference is not observed between the pairs of mutations and the Az line, respectively. The comparison results suggested that the height difference between WT and mutant plants/Az might be caused by the effect of EMS treatment.

To assess the effect of *Slhsp70*-1 mutants on leaf structure, we observed leaf morphology and compared WT and Az lines. There is no clear difference between the structure, shape, and leaf complexity between the mutant plants compared with the Az and WT plants. Leaf size is similar to the control, indicating that the *Slhsp70*-1 mutation does not affect the morphology and structure of tomato leaves Figure 5.7.

### *5.3.3. Slhsp70-1 mutants were not affected to tomato fruit set and enlargement*

The first flowering timing was recorded and compared between *Slhsp70*-1 mutant plants, WT, and Az lines. The results show that, the WT plants need an average of about 30 DAS to start flowering, while the mutant plants need 32–34 DAS and the Az lines need 31–34 DAS to reach the same state (Figure 5.8A). The first flowering timing of WT, Az and

mutant plants did not differ significantly. Based on the growth status of the tomato plant according to the number of leaves, at the time of flowering, WT plants have about 8 real leaves, while mutants and Az lines appear from 8-9 leaves (Figure 5.8B). The leaves number showed no significant difference in the process of development and formation of flowers based on the growth rate of tomato plants according to the number of leaves. The integrated results above indicated that the *Slhsp70*-1 mutant did not have a serious effect on the flowering time of tomato.

The number of flowers is counted within 1 month from the start point time of flowering. Results showed that for about 1 month, WT plants gave about 22 flowers, while mutant lines gave about 19-23 flowers and Az lines produced from 17 to 22 flowers (Figure 5.9). Line *Slhsp70*-1-2 and *Slhsp70*-1-3 produced less flowers than the two other mutant lines as well as WT. The number of flowers created in each mutant line does not differ significantly comparing to Az line under the same conditions. This proved that gene mutations do not affect the development ability of tomato flowers.

We labelled selected flowers and monitored the fruit formation process. The fruit creation rate was calculated by the number of fruits on the total number of marked flowers. 30 flowers were marked on each tree on 6 individual trees, and we calculated the fruit set rate after two weeks of flowering. According to this result, WT fruiting rate was about 66%, while mutant plants had a fruiting rate of about 45% (*Slhsp70*-1-2) to 60% (*Slhsp70*-1-1), and the Az lines resulted in a 51–65% fruit set rate (Figure 5. 10). There is a decrease in fruiting rate for the mutant line *Slhsp70*-1-2 ($p = 0.002$) and *Slhsp70*-1-3 ($p = 0.02$) compared to the WT. However, this difference is not shown between the mutant line and Az. In all mutations, the difference in the fruit production rate was not determined relative to WT. Besides, this difference is not reflected in the mutant flow and Az. We therefore assume that the difference in the pace of fruit manufacturing is not sufficient to check the difference between mutations of *Slhsp70*-1.

The diameter and the vertical sizes of the tomato fruit were measured and recorded from ripe fruits for comparison. Measuring the diameter of the fruits showed that the average diameter of the tomato WT was about 16.5 mm, while the diameter of tomato fruit in mutant lines ranged from 12.5–15.8 mm. The Az lines had fruits with a diameter of 12.5–17.2 mm. The *Slhsp70*-1-2 line produces results smaller than the rest, as well as smaller than WT ($p < 0.0001$). Meanwhile, the remaining mutant lines obtain the same diameter size as the WT

line. There is no significant difference in the fruit diameter of the mutant lines and the Az line, respectively (Figure 5.11 A).

For vertical size comparison, the average vertical size of the WT fruit was about 17.3 mm, while all off the remaining lines showed the vertical size of fruit from 14.8–16 mm. No significance was observed in the vertical size for all compared fruits. This result also suggested that the mutation of *Slhsp70*-1 does not seem to be effective for fruit development (Figure 5.11 B).

## 5.4. Discussion

### *5.4.1. Tomato missense mutant of Slhsp70-1 may be not effective to phenotype changes*

This study was conducted with materials from EMS mutation library. EMS was a monofunctional ethylating agent that has been found to be mutagenic in a wide variety of genetic test systems, a triggering agent variable by replacing nucleotide A by T or C by G (Krieg, 1963; Greene et al., 2003). However, the transition of G/C to C/G or G/C to T/A by 7-ethylguanine hydrolysis or A/T to G/C transition were also recorded in the treatment of EMS at low concentrations (Krieg, 1963). In *Arabidopsis*, the frequencies of EMS-induced stop codon were about 5%, while the missense mutations rate was about 65%, respectively (Suzuki et al., 2008). EMS mutagenesis can assist identify loss and gain-of-function mutants to comprehend the function of a protein-based amino acid (Østergaard and Yanofsky, 2004).

EMS is a mutagenesis agent that may induce many random point mutations appearring in the genome of the EMS-treated object. Point mutations can occur at different regions throughout the genome even in one gene. The benefits of this mutation technique are that a broad random, uncontrollable set of mutations could be created on the host object. For instant, more than 700 mutations could be obtained from each *Arabidopsis* line treated with EMS, and about 50,000 $M_1$ lines were created by EMS treatment (Jander et al., 2003). However, depending on the position of the point mutation, it may affect the amino acid sequences, as well as effect changes in gene function. While silent mutation could not alter the amino acid sequence, it means that there is no change in the structure and function of the protein. Missense was considered the most effective mutation for gene functional research because this sort of mutation permitted the translate amino acid sequence to expect important changes in the protein's structure and function owing to the appearance of the stop codon while missense mutations were also commonly used. In this study, 4 points mutations were

screened as missense type. There was no difference in the mutant and control phenotypes, suggesting that there was no impact on the growth and development of the mutant as well as the fruit set of the missense of the tomato *HSP70-1* gene.

### *5.4.2. Gene duplication may recover the lost function of a gene*

Phylogeny analysis indicated that there was a closed relationship between a group of genes *SlHSP70-1*, *SlHSP70-2, SlHSP70-3, SlHSP70-4,* and *SlHSP70-21* when they were placed together in one cluster of the phylogeny tree. The protein of the targeted gene showed similarity with the other three genes *SlHSP70-2, SlHSP70-4,* and *SlHSP70-21*. The four genes structures had the same number of introns, exons, and similar length (Figure 2B). It suggested that they might share their biological functions, the most frequently duplication was identified with *SlHSP70-1* gene. The gene *SlHSP70-1* duplicated with 4 other genes, including all three other heat shock cognate genes of *SlHSP70-2, SlHSP70-3,* and *SlHSP70-4* due to a high similarity when comparing the CDS sequences between two genes of each (greater 85% identity). The abundant duplication of *SlHSP70* genes may suggest the overlapping response of these for biological function. The duplication of gene function is really a test of the big reverse genetics using EMS mutant. RNA interfere was have the potential enough to solve this problem. On the other hand, the analysis of the expression pattern of the target gene between the targeted gene and its duplicated version in the family gene showed that the target gene was always expressed at very low level in most tissues, which implied this gene was essential for plant growth and development, but it was not required at high levels. Missense mutation could downregulate the gene expression but it could not silence this gene. This was probably the reason missense mutants of *Slhsp70* could not show a clear difference in their phenotype.

**Table 5. 1. List of Primers for amplifying the *HSP70-1* gene in TILLING**

| Primer | Primer sequence |
|---|---|
| **Fr1 F2** | 5'-ATCACGTCATTGATCATCAGA-3' |
| **Fr1 R1** | 5'-TAGGTCAAACACGGCTAC-3' |
| **Fr2 F2** | 5'-GTCACGTCCAAAAGCAGCAAGTC-3' |
| **Fr2 R2** | 5'-AGACGTGCAGTGTTTATATGATT-3' |
| **Fr3 F1** | 5'-GTTTTGGCAGAATAATTGCTA-3' |
| **Fr3 R2** | 5'-CTTCTTCAACGGCAAGG-3' |

**Table 5. 2. List of *Slhsp70-1* mutations in TOMATOMA mutant collection database**

| Mutant lines used in this study | Mutant line in TOMATOMA database |
|---|---|
| ***Slhsp70-1-1*** | TOMJPG7598 |
| ***Slhsp70-1-2*** | TOMJPG2209 |
| ***Slhsp70-1-3*** | TOMJPG8552 |
| ***Slhsp70-1-4*** | TOMJPG946 |
| ***Slhsp70-1-5*** | TOMJPG2212 |
| ***Slhsp70-1-6*** | TOMJPG2422 |
| ***Slhsp70-1-7*** | TOMJPG5982 |
| ***Slhsp70-1-8*** | TOMJPG8740 |
| ***Slhsp70-1-9*** | TOMJPG2603 |
| ***Slhsp70-1-10*** | TOMJPG80 |
| ***Slhsp70-1-11*** | TOMJPG2038 |
| ***Slhsp70-1-12*** | TOMJPG2201 |

**Table 5. 3. List of Primers for Sanger sequencing validation of *HSP70*-1 mutations**

| *Name of primer* | *Primer sequence* |
|---|---|
| ***285F*** | 5'-CAAGGTTATTCCTGGCCCTGGTG-3' |
| ***809F*** | 5'-GGACAGCTTGTGAGAGGGCTAAG-3' |
| ***1015R*** | 5'-GATCCACCGACAAGAACAACATC-3' |
| ***1581R*** | 5'-CTTCTGCCTTGTACTTTTCAGCT-3' |

**Table 5. 4. Description of mutation position, amino acid substitution**

| Mutant allele | Chromosome[a] | Position[b] (bp) | Within[c] | Gene[d] | Strand | Type of mutation | Nucleotide change[e] | Amino acid substitution[f] |
|---|---|---|---|---|---|---|---|---|
| ***Slhsp70-1*-1** | SL2.50ch6 | 47192147 | Exon 1 | Solyc06g076020.2.1 | Reverse | Missense | G282A | G37R |
| ***Slhsp70-1*-2** | SL2.50ch6 | 47193449 | Exon 2 | Solyc06g076020.2.1 | Reverse | Missense | G1584A | G207S |
| ***Slhsp70-1*-3** | SL2.50ch6 | 47193769 | Exon 2 | Solyc06g076020.2.1 | Reverse | Missense | G1904A | M313I |
| ***Slhsp70-1*-4** | SL2.50ch6 | 47194116 | Exon 2 | Solyc06g076020.2.1 | Reverse | Missense | A2251T | Q285H |
| ***Slhsp70-1*-5** | SL2.50ch6 | 47191947 | 5'-UTR | Solyc06g076020.2.1 | Reverse | Silent | C82T | No change |
| ***Slhsp70-1*-6** | SL2.50ch6 | 47192201 | Exon 1 | Solyc06g076020.2.1 | Reverse | Missense | G336A | G55S |
| ***Slhsp70-1*-7** | SL2.50ch6 | 47193765 | Exon 2 | Solyc06g076020.2.1 | Reverse | Missense | A1900C | N312T |
| ***Slhsp70-1*-8** | SL2.50ch6 | 47193595 | Exon 2 | Solyc06g076020.2.1 | Reverse | Silent | C1730T | No change |
| ***Slhsp70-1*-9** | SL2.50ch6 | 47193672 | Exon 2 | Solyc06g076020.2.1 | Reverse | Missense | C1807T | S281F |
| ***Slhsp70-1*-10** | SL2.50ch6 | 47194881 | 3'UTR | Solyc06g076020.2.1 | Reverse | Silent | A3016C | No change |
| ***Slhsp70-1*-11** | SL2.50ch6 | 47194235 | Exon 2 | Solyc06g076020.2.1 | Reverse | Missense | G2370A | G469S |
| ***Slhsp70-1*-12** | SL2.50ch6 | 47194321 | 3'UTR | Solyc06g076020.2.1 | Reverse | Silent | G2456A | No change |

[a] The location in the chromosome in the tomato genome

[b] Position of nucleotide substitution according to tomato genome sequence database, version SL2.50 (Sol Genomics Network)

[c] Location of nucleotide substitution of the gene in column [d]

[d] Gene mutated according to Sol Genomic Network database

[e] Reference nucleotide (left) and alternative nucleotide (right) at position of mutation in gene sequence

[f] Reference amino acid (left) and its substitution according to the position in column [e]

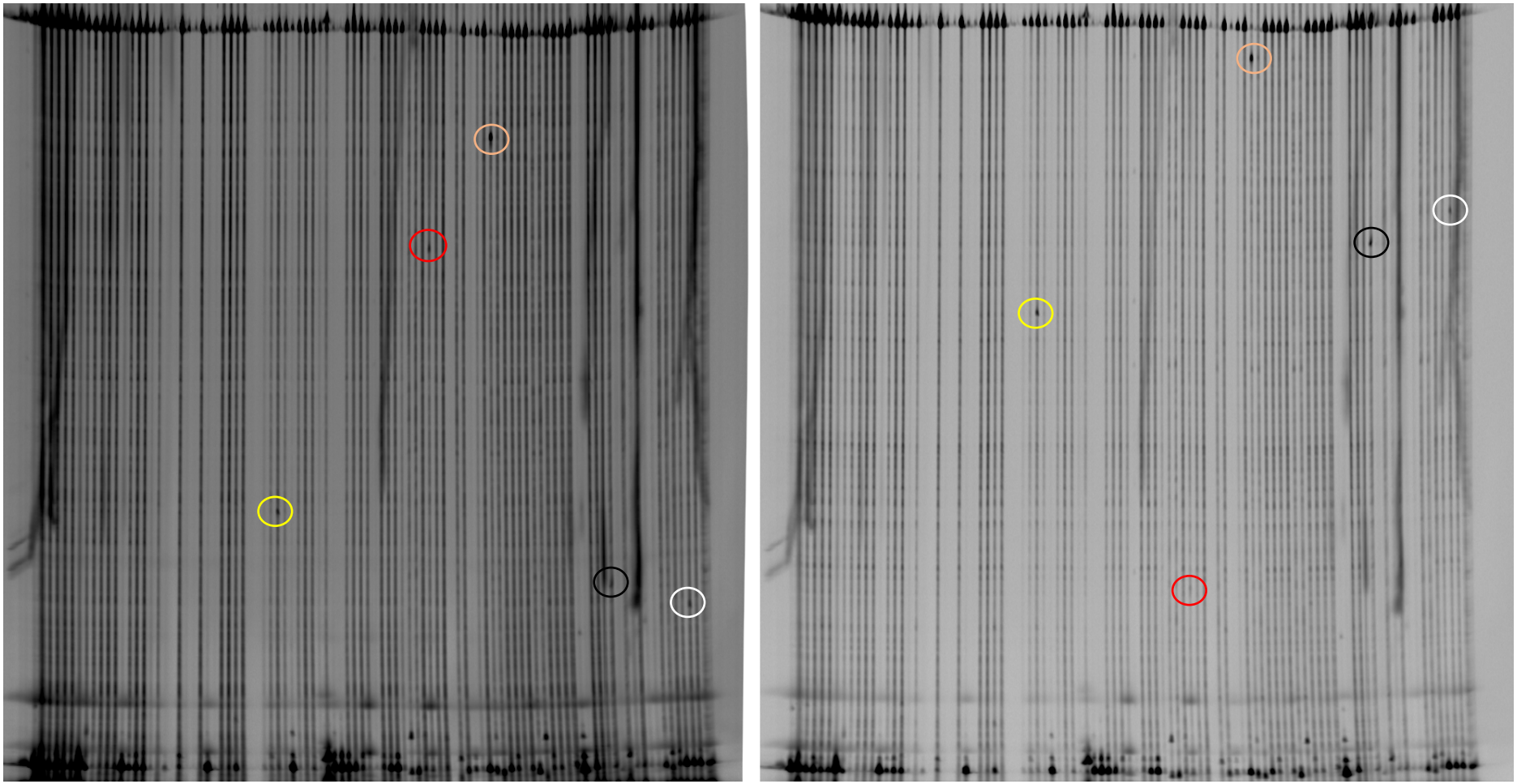


**Figure 5. 1. TILLING second screening result of *SlHSP70-1* mutations from EMS library**

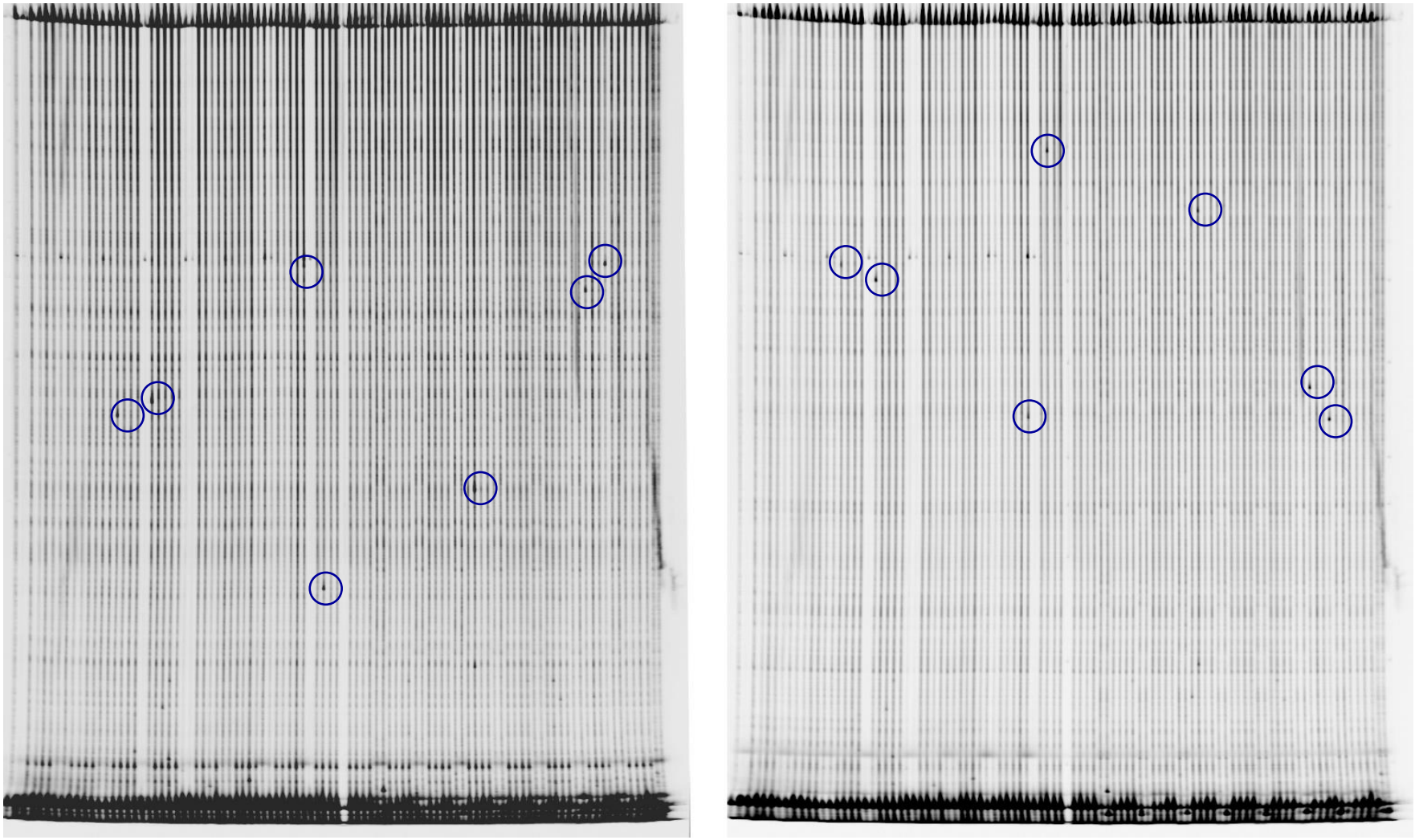


**Figure 5.2. TILLING second screening result of *SlHSP70-1* mutations from EMS library (continue)**

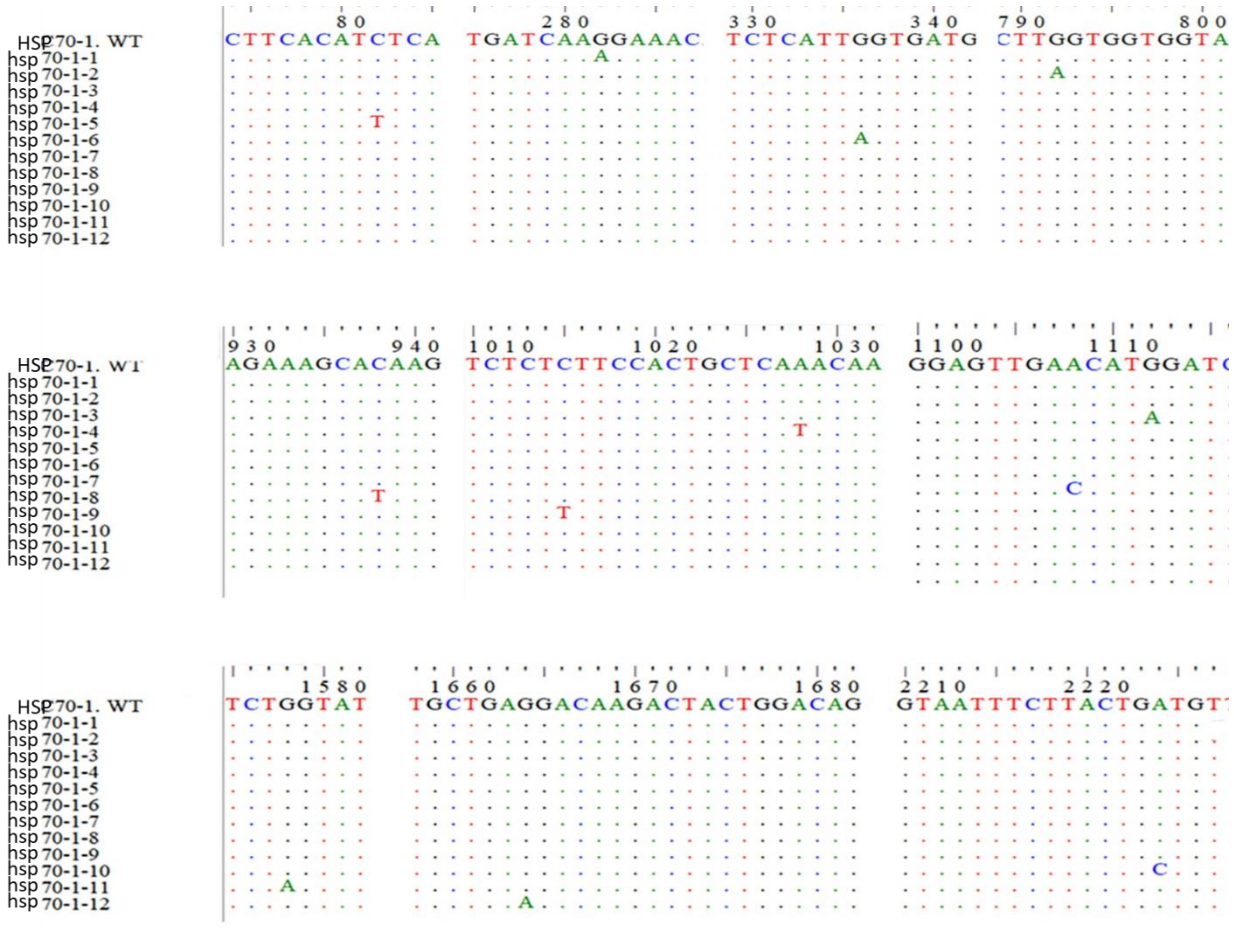



**Figure 5.3. Nucleotide alignment revealed *SlHSP70-1* point mutations**

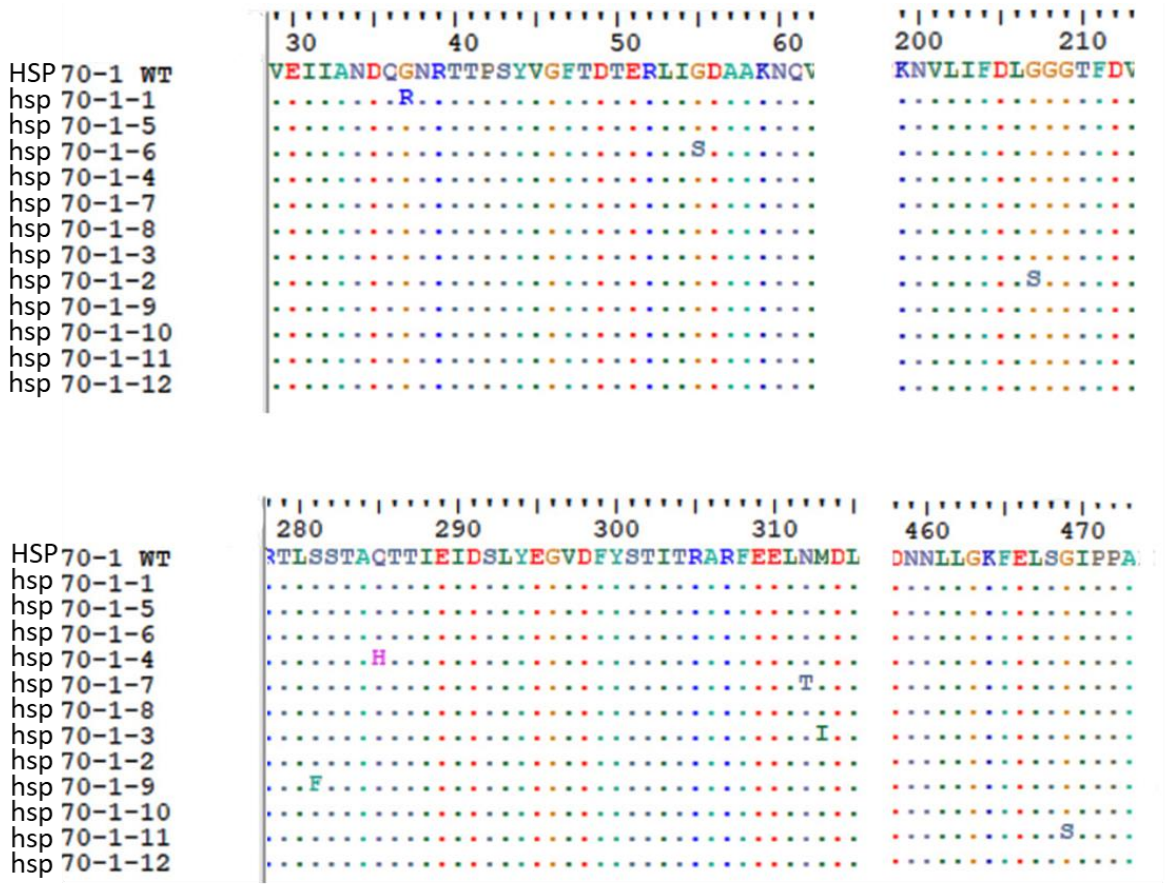



**Figure 5.4. Amino acid sequences alignment revealed impact of *SlHSP70*-1 point mutations**

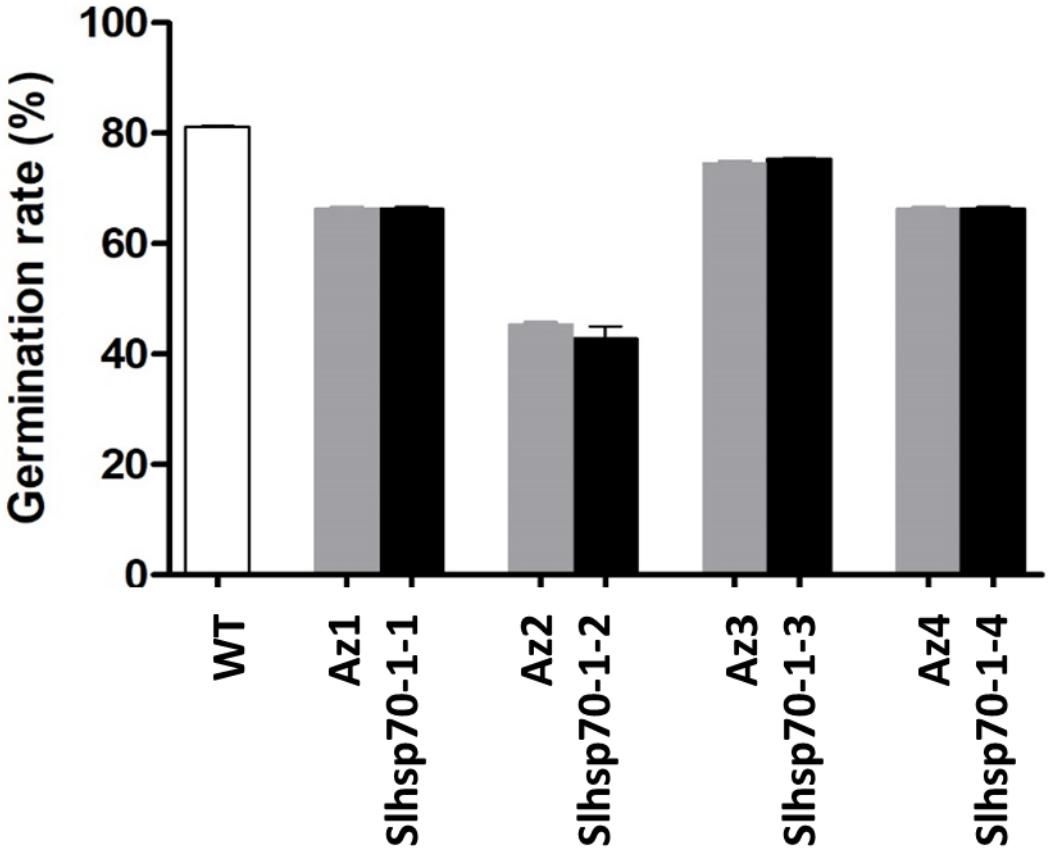
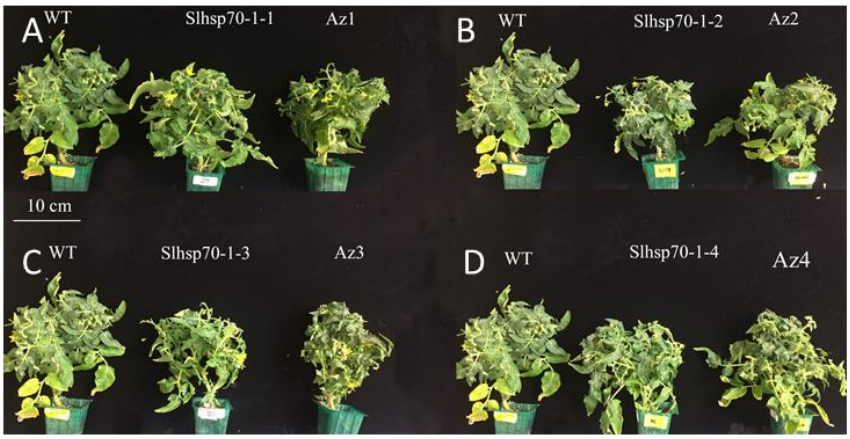
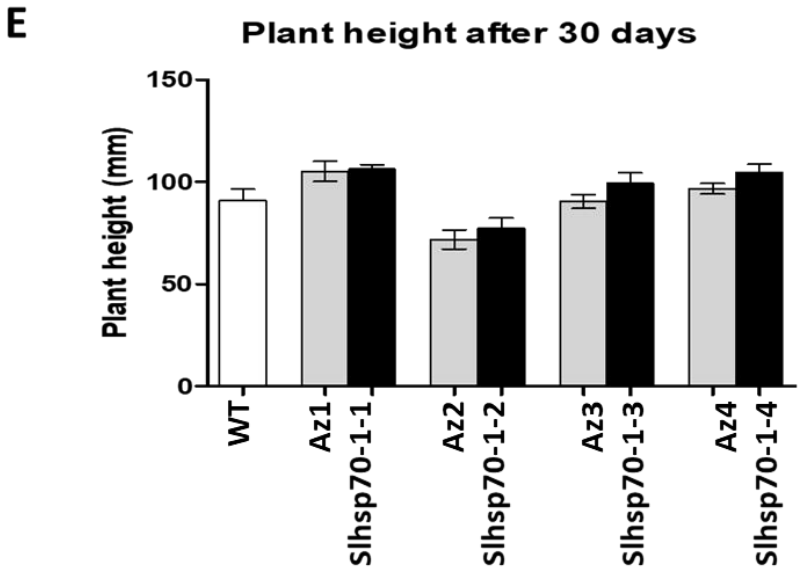



**Figure 5.5. Germination rate of *Slhsp70-1* mutant comparing to WT and Az line**



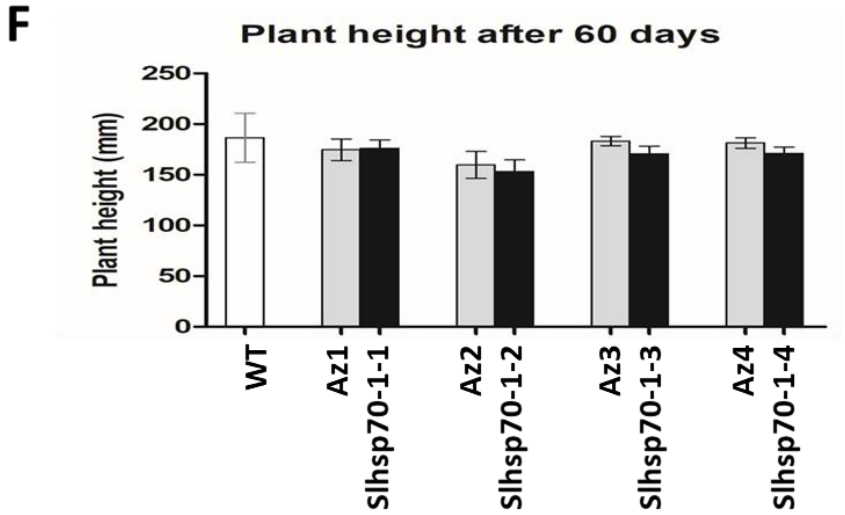




**Figure 5.6. Comparison plant height of tomato *Slhsp70*-1 mutant lines to WT and Az lines**

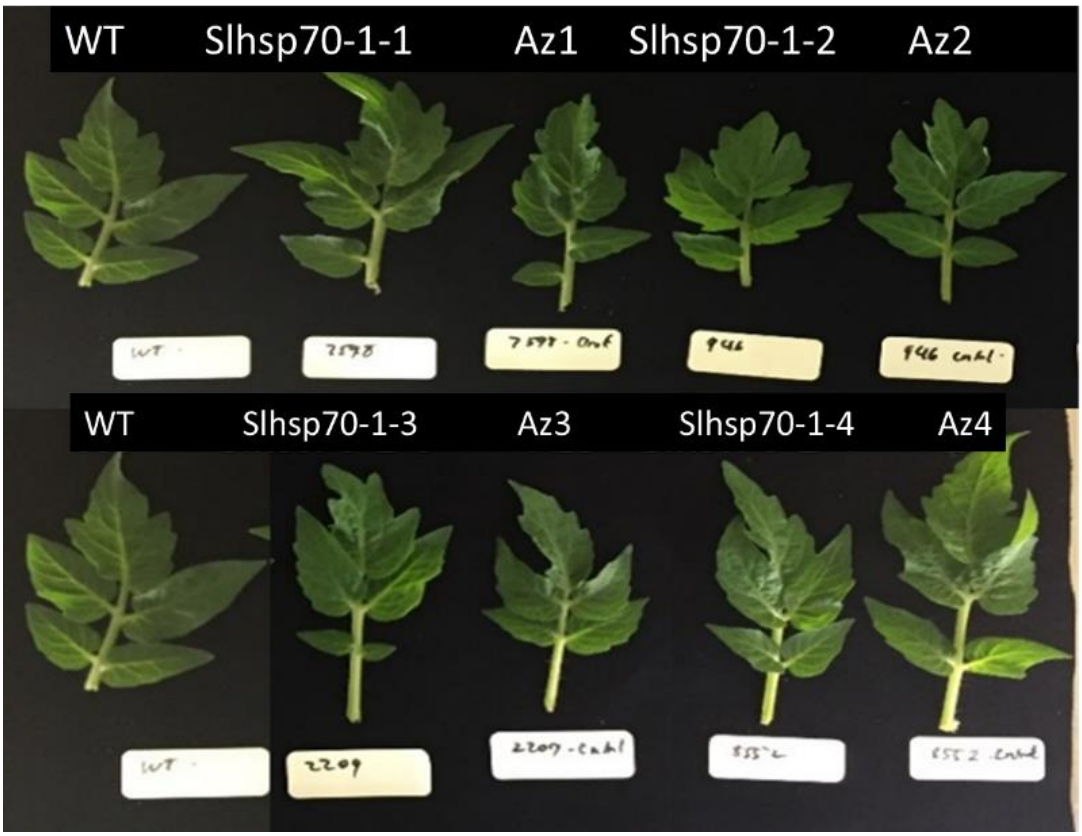



**Figure 5.7. Comparison of leaf morphology of tomato *Slhsp70*-1 mutant, WT and Az lines**

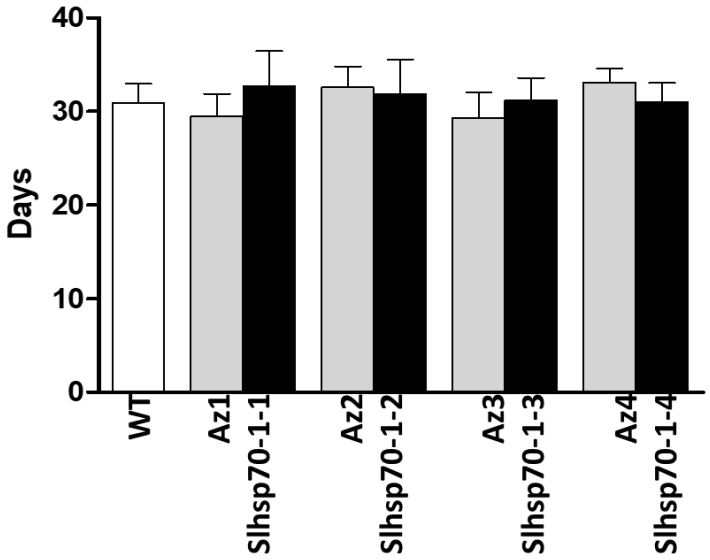



**Figure 5.8. Comparison of first flowering timing between *Slhsp70*-1 mutant, WT and Az lines**

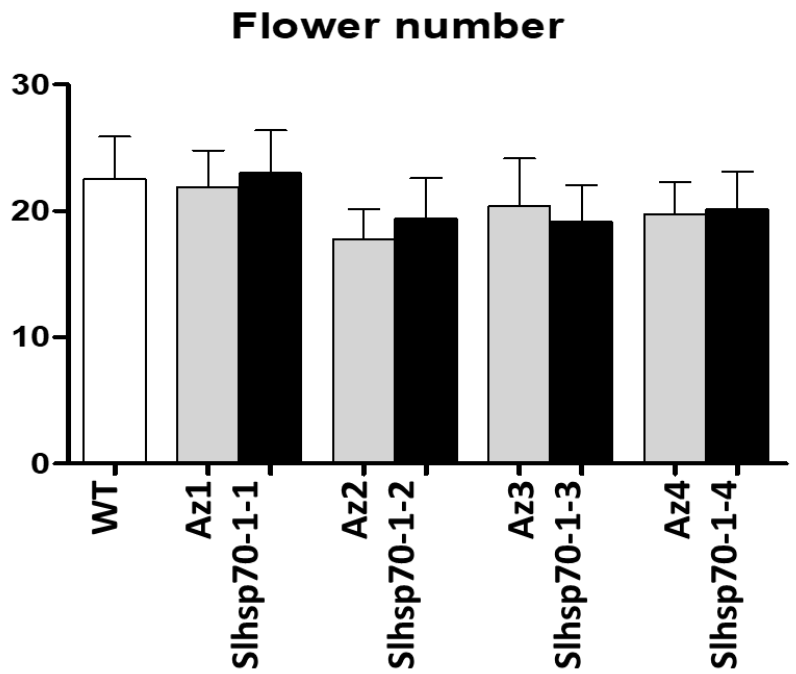



**Figure 5.9. Comparison of flowers number of *Slhsp70*-1 mutant, WT and Az lines**

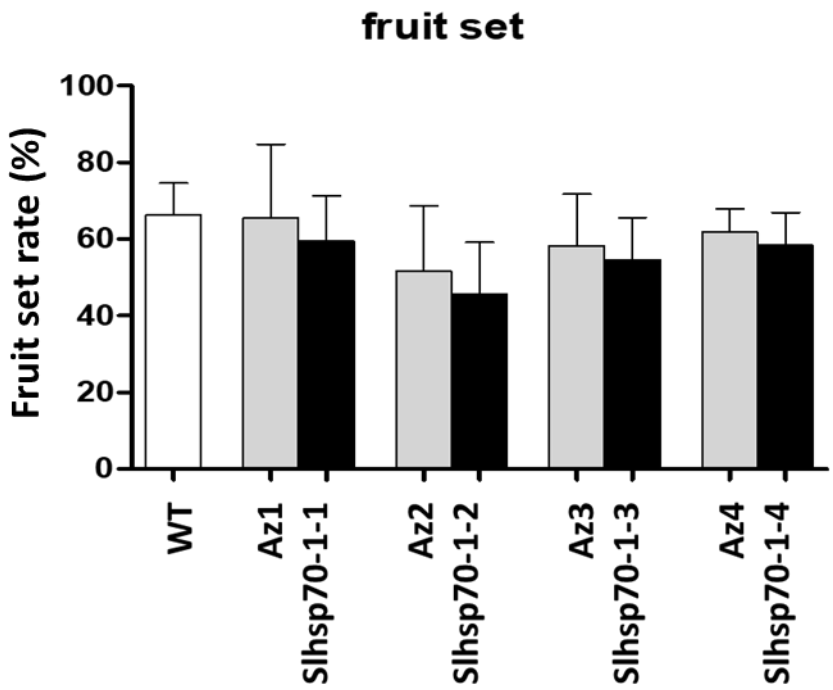



**Figure 5.10. Comparison of fruit set rate of *Slhsp70*-1 mutant, WT and Az lines**

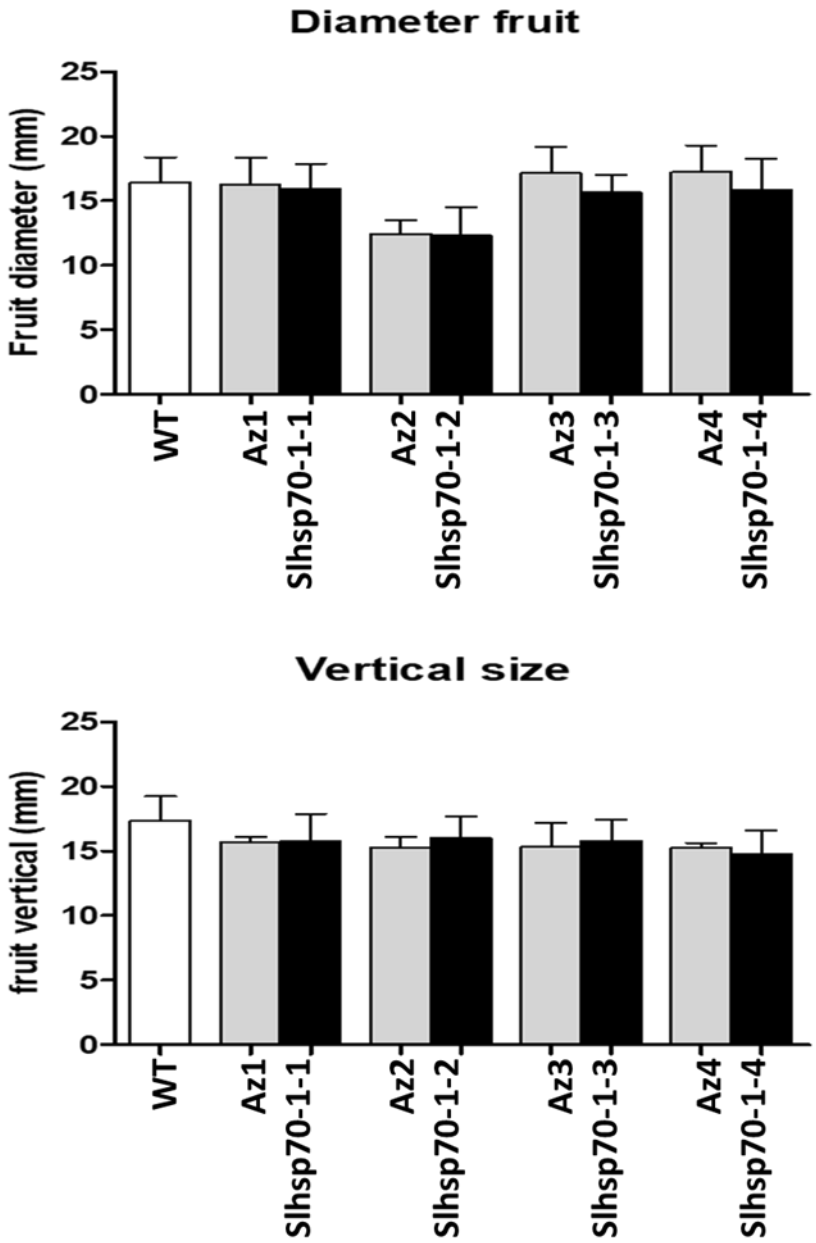



**Figure 5.11. Comparison of fruit sizes of *Slhsp70*-1 mutant, WT and Az lines**

# Chapter 6. General discussion

In this chapter, we conduct a comprehensive review of the results obtained in chapters 2 through 5. By co-expression analysis to construct the network of genes related to *SlIAA9* and *SlDELLA* during tomato growth and development, we realized that *SlHSP70-1* appeared in both *SlIAA9* and *SlDELLA* networks, hypothesized that this gene also has any functions in the growth and development of tomatoes, in relation with *SlIAA9* and *SlDELLA*. The research approach by co-expression network analysis showed that it was an effective method of finding potential candidate genes, based on microarray analysis and RNA sequencing data. This approach is commonly applied to identify functional module in bamboo growth and development (Ma et al., 2018), reveal hybrid-specific modules and find candidate genes related to plant height development in maize (Wang et al., 2018) or identify key regulators of flower and fruit development in wild strawberry (Shahan et al., 2018). Particularly, in Arabidopsis, gene co-expression analysis was applied popularly to annotate gene module and gene cluster, predict motifs in the promoter regions of co-expressed genes, and potential candidate gene relating to plant growth and development, as well as stress resistance of biotic and abiotic conditions (Mao et al., 2009; Mutwil et al., 2010; Zheng et al., 2011; Prasch and Sonnewald, 2013; Rasmussen et al., 2013; Amrine et al., 2015; Jiajie et al., 2016; van Veen et al., 2016)

In tomatoes, microarray and RNA sequencing data are collected in database National Center for Biotechnology Information Gene Omnibus (GEO-NCBI) or Tomato Functional Genomics Database (TFGD). However, the application of co-expression analysis is often not common, so far (Ozaki et al., 2010; Fukushima et al., 2012; Tzfadia et al., 2016; Mandal et al., 2018). Thus, this study was not a very effective approach to discovering and evaluating the function of a potential gene but also suggested for an effective research approach for tomato functional genomics.

*SlHSP70-1* selected from co-expression network analysis is one of 25 *SlHSP70* genes present in tomato. Since *SlHSP70* is one of a group of proteins that protect and increase expression when cells are exposed to stress conditions such as heat, it is possible that *SlHSP70-1* also has a function under high temperature conditions. Studies on the *SlHSP70-1* orthologs of Arabidopsis thaliana and rice have been reported. In rice, *OsHSP70-1* is one

of 32-member-HSP70 super family, including 24 Hsp70 family's and 8 Hsp110 family members (Sarkar et al., 2013). *OsHSP70-1* expressed in both normal growth conditions, as well as in heat-treated shock conditions, although in terms of thermal shock, its expression level is determined to be lower than other members of the gene family (Jung et al., 2013). *OsHSP70-1* was also detected to be responded in some other stress conditions, such as drought, salt, heat and light due to the presence of cis-element (nTTCnnGAAnnTT-Cn) in its promoter region (2kb before its initial ATG sequence) (Morris et al., 2008; Jung et al., 2013).

In Arabidopsis, the over-expression of *AtHSP70*-1 was enhanced for the thermo-tolerance of plants under heat stress condition (Jungkunz et al., 2011). Under normal conditions, *AtHSP7*0*-1* is detected at normal levels in leaf, root, stem, and flower, but it is detected at higher levels in leaf than other organs (Wu et al., 1994; Sung et al., 2001b). *AtHSP70-1* and its isomer *AtHSP70-2* expressed at a similar level under normal conditions, as well as both genes were induced after a heat treatment, however *AtHSP70*-1 reacted more slowly than *AtHSP70-2*, even if it was unchanged or repressed if the heat shock condition only happened for a short time (Sung et al., 2001b; Su and Li, 2008). Both genes restore levels of normal expression quickly after heat treatment (Llamas et al., 2017). Thus, the possibility that SlHSP70 in tomato will also exhibit stress tolerance when grown under high temperature stress conditions. *AtHsp70-1* mutant showed variegation and delayed growth, indicated that it can be assumed that while other *AtHSP70* members may be more prioritized in protein folding refolding for resisting stress conditions, *AtHsp70-1* played as a housekeeping gene of which functions involving in Arabidopsis growth and development (Su and Li, 2008).

In this study, we tried to enhance the expression level of *SlHSP70-1* gene in tomato plants by adding one copy of the gene to the tomato (wild type) plant and growing it under normal physiological conditions for plant's development of tomato. Transgenic tomato plants were higher than wild type plants, indicated that *SlHSP70-1* was involved in plant growth and development through elongation and cell division. This gene function is similar to its isomer in Arabidopsis mentioned above. In order to explain the role of this gene in the growth and development of tomato plants, especially the extension of internode, we hypothesized that there was something special in the gene or protein structure that was responsible for cell elongating and dividing. The Web SMART (Simple Modular

Architecture Research Tool) reported that two kinds of protein domain should be built from amino acid sequence of *SlHSP70-1* (http://smart.embl-heidelberg.de/). Besides the HSP70 domain, which characterizes commonly for the gene family, this amino acid sequence also set up another domain named MreB_MbI. MreB is similar role to actin protein in eukaryote organisms, played an important role in the process of elongated bacterial cell walls, as well as plays an important role in building the skeleton for bacteria by interacting with several proteins that have been shown to be involved in length growth (Doi et al., 1988). In *Bacillus subtilis*, MreB and its isoforms, Mbl and MreBH, controlled sidewall elongation during vegetative development (Soufo and Graumann, 2003). Thanks to MreB structure, it is possible to suggest the role of *SlHSP70-1* protein in internode elongation through involvement in cell formation and elongation in tomato.

Through co-expression network analysis, this study showed that *SlHSP70-1* played important roles in the growth and development of tomato plants in a relation with *SlIAA9* and *SlDELLA*. As we know, *SlIAA9* and *SlDELLA* are two negative regulators for plant growth and development, which have been demonstrated in in model plants such as tomato and Arabidopsis. Their loss functions cause phenotypic differences in plant growth and development, as well as fruit set and enlargement. For assessing the relationship between *SlHSP70-1*, *SlIAA9* and *SlDELLA*, we found that under normal conditions, the *SlHSP70-1* gene expressed very low while the expression level of *SlIAA9* and *SlDELLA* genes were normal. Meanwhile, in transgenic plants, the expression level of the *SlHSP70-1* gene is high in tissues that were examined such as internode, leaves and fruits. This expression pattern was similar to the trend of *SlHSP70-1* gene expression in mutant plants *Sliaa9*. This suggested a hypothesis that under normal conditions, the *SlHSP70* was inhibited by *SlIAA9*. This gene can only be escaped from *SlIAA9* control and then expressed highly when the *SlIAA9* gene decreases its expression level, and this leads to a change in the phenotype of the tomato plant, particularly the extension of the internode. The expression level of *SlDELLA* gene reduced in *Sliaa9* mutant plants compared to WT plants, showing *SlIAA9* and *SlDELLA* genes were co-expression. Additionally, the expression of *SlDELLA* gene in the internode of *SlHSP70-1* transgenic plants also tends to decrease compared to wild type as shown in Figure 4.16B suggested that *SlDELLA*, as well as *SlIAA9*, acts as a negative controller of *SlHSP70-1* during the growth and development of tomato plants. This might also explain why *SlHSP70-1* gene was at low expression level in comparison to other genes in the *SlHSP70* family gene of tomatoes.

HSP70 is a highly conservative gene family, members of them have also been considered to be high redundancy together (Sung et al., 2001). The results presented in chapter 3 showed that the *SlHSP70-1* gene was overlapping with at least 4 other members of the gene family, which presented in high levels of similarity in nucleotide and amino acid sequences. This similarity not only explained their conservative characters but also seems to be a guard for stabilising of gene function during evolution. That was probably the reason why the single mutation on the target gene does not show a difference in phenotype compared to the control as in chapter 5.

## Summary

Tomato is one of the highest value vegetables that is cultivated and consumed widely in the world. Therefore, many cultivation methods have been applied to improve the growth, development, and fruit set of the tomato. While agronomic methods such as applying nutrition and optimization of the growing environment for tomatoes have reached their limit, some exogenous hormones have been applied as an effective solution. In tomato cultivation, several common phytohormones such as auxin, gibberellin, and cytokinin are applied, in which two hormones auxin and gibberellin are most commonly used. These hormones have been found to boost the growth and development of a number of hormone response genes, in which *SlIAA9* and *SlDELLA* are the more remarkable genes. *SlIAA9* and *SlDELLA* are two key genes that play significant roles in many distinct physiological procedures requiring many distinct genes to participate in each of these procedures. The mutations of *SlIAA9* or *SlDELLA* genes cause noticeable changes related to growth as well as the formation and development of tomato fruits, such as abnormal tree height, simplification of leaf structure and especially, the induction of seedless fruits formation also called parthenocarpy. The mechanism of these procedures has nevertheless not been fully clarified and the concerned genes remain unknown. However, microarray and RNA sequencing experiments for *SlIAA9* and *SlDELLA* mutations were performed. The development of high-performance techniques, including microarrays and RNA sequencing (RNA-seq) and their corresponding techniques of data analysis, can now be used to determine the gene's functional state in co-expression or interaction. On the other hand, database libraries have been created to become bioinformatics research resources. Bioinformatics tools are powerful tools to narrow down large databases to focus on potential candidate genes. In this study, we applied bioinformatics tools to find network-related genes that control the growth and development of tomato plants in relation to *SlIAA9* and *SlDELLA*.

In chapters 2 and 4, we performed a gene-to-gene analysis of the co-expression network using publicly available microarray data to extract genes directly connected to nodes *SlIAA9* and *SlDELLA*, respectively. We chose *HSP70,* which was connected with *SlIAA9* and *SlDELLA* in the co-expression network. To validate how the extent of the *SlHSP70-1* effects on tomato growth and fruit set and development, lines overexpressing the target gene were generated. We found that the overexpression of the targeted *SlHSP70-1*

showed internode elongation, while the overexpressing lines did not show abnormal leaf shape, fruit set, and size when compared with the control. It suggests that the targeted *SlHSP70* is likely to function as an upregulation of shoot growth like *SlIAA9* and *SlDELLA*, while it is not contributed to parthenocarpy as well as fruit set. It also displayed that only one *SlHSP70* in a total of 25 genes was able to affect shoot elongation.

In chapter 3, we analyzed the characteristics of *SlHSP70* family gene base on their public information of genomic and amino acid sequence. As results, we established the phylogeny tree of all *HSP70* gene in tomato, recognized the large duplication between genes in the family, in which the targeted *SlHSP70-1* gene showed the highest duplication rate with five other genes in the family. This result suggested that many genes may have an overlapping function during the growth and development of tomato. However, the expression patterns of the targeted gene and its duplication versions were different in several tissues, which indicates gene functions are lost or gained during gene evolution.

In chapter 5, we apply reverse genetics using TILLING and analyze the function of gene *SlHSP70*-1 during the growth, development, and fruit set of tomato. 12 point-mutations were collected from the EMS library with a mutation rate of about 1.3 mutations each kilobase of nucleotide and 4 mutations were phenotypically characterized. No breakthrough of gene functions could be revealed. This reinforces the hypothesis that the duplication of the genome gene is the factor that ensures the tomato plant grows normally when the function of the main gene has been affected by the mutations.

**Acknowledgments**

In this part, let me express my deepest gratitude to all the people who supported me during my doctoral course at the University of Tsukuba. First of all, I would like to express my sincere gratitude to my respected academic supervisor, Prof. Miyako KUSANO, Professor, Faculty of Life and Environmental Sciences, University of Tsukuba. Thank you very much for giving me the opportunity to study and work in a creatively scientific environment, allowing me to join such an innovative research group, also for your warm guidance, beneficial idea and advice, unlimited support as well as enlightening discussion.

Furthermore, I would like to express great thanks to my closest teacher, Dr. Ning WANG, Assistant Professor, Faculty of Life and Environmental Sciences, University of Tsukuba. Please allow me to express my sincere thanks for your guidance since the time I started to plan my research, for continuous support and unlimited help during research, for every beneficial idea and advise, as well as great discussion during my doctoral course.

I would like to express my deep gratitude to advisory committee members, Professor Miyako KUSANO, Professor Hiroshi SHIBA, Professor Kenji MIURA and Assistant Professor Ning WANG for all your valued suggestions for my thesis. I have learned many important useful things in writing and presenting research results with your advices.

I am grateful to Mr. Ken KAMIYA for excellent collaboration. I would like to show my sincere gratitude to Mr. Naoya SUGI, my tutor and my first Japanese friend, for all his support and friendship.

Besides that, I also wish to express my sincere thanks to all the teachers in our research group, my lab mates and my friends. Thank you very much for all attention, suggestions, discussion and assistance during my doctoral course.

I sincerely thank to the Ministry of Education, Culture, Sports and Technology (MEXT), Japan for financial supporting for my doctoral scholarship.

Last but not least, I would like to express my endless and sincere thanks to my family; my parents, brother, and my grandmother for your unlimited and endless support, love, and pray. You are my steadfast supporters during this PhD. course even throughout my life.

## References


Amrine, K.C.H., Blanco-Ulate, B., and Cantu, D. (2015). Discovery of Core Biotic Stress Responsive Genes in Arabidopsis by Weighted Gene Co-Expression Network Analysis. *PLOS ONE,* 10, e0118731.

Aoki, K., Ogata, Y., and Shibata, D. (2007). Approaches for extracting practical information from gene co-expression networks in plant biology. *Plant Cell Physiol.,* 48, 381-390.

Aoki, K., Ogata, Y., Igarashi, K., Yano, K., Nagasaki, H., Kaminuma, E., and Toyoda, A. (2013). Functional genomics of tomato in a post-genome-sequencing phase. *Breeding Sci.,* 63, 14-20.

Ariizumi, T., Shinozaki, Y., and Ezura, H. (2013). Genes that influence yield in tomato. *Breeding Sci.,* 63, 3-13.

Audran-Delalande, C., Bassa, C., Mila, I., Regad, F., Zouine, M., and Bouzayen, M. (2012). Genome-Wide identification, functional analysis and expression profiling of the Aux/IAA gene family in tomato. *Plant Cell Physiol.,* 53, 659-672.

Augustine, S.M., Cherian, A.V., Syamaladevi, D.P., and Subramonian, N. (2015). *Erianthus arundinaceus* HSP70 (*EaHSP70*) acts as a key regulator in the formation of anisotropic interdigitation in Sugarcane (*Saccharum* spp. hybrid) in response to drought stress. *Plant Cell Physiol.,* 56, 2368-2380.

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

Becker, A., and Theißen, G. (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol. Phylogenet. Evol.,* 29, 464-489.

Bensen, R.J., and Zeevaart, J.A.D. (1990). Comparison of Ent-kaurene synthetase A and B activities in cell-free extracts from young tomato fruits of wild-type andgib-1, gib-2, andgib-3 tomato plants. *J. Plant Growth Regul.,* 9, 237.

Bertin, N. (2005). Analysis of the tomato fruit growth response to temperature and plant fruit load in relation to cell division, cell expansion and DNA endoreduplication. *Ann. Bot.,* 95, 439-447.

Bolle, C. (2004). The role of GRAS proteins in plant signal transduction and development. *Planta,* 218, 683-692.

Cai, B., Li, Q., Liu, F., Bi, H., and Ai, X. (2018). Decreasing fructose-1,6-bisphosphate aldolase activity reduces plant growth and tolerance to chilling stress in tomato seedlings. *Physiol Plant*, 163, 247-258.

Campanoni, P., and Nick, P. (2005). Auxin-dependent cell division and cell elongation. 1-Naphthaleneacetic acid and 2,4-dichlorophenoxyacetic acid activate different pathways. *Plant Physiol.*, 137, 939-948.

Cao, J., Schneeberger, K., Ossowski, S., Günther, T., Bender, S., Fitz, J., Koenig, D., Lanz, C., Stegle, O., Lippert, C., Wang, X., Ott, F., Müller, J., Alonso-Blanco, C., Borgwardt, K., Schmid, K.J., and Weigel, D. (2011). Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nat. Genet.*, 43, 956.

Carrera, E., Ruiz-Rivero, O., Peres, L.E.P., Atares, A., and Garcia-Martinez, J.L. (2012). Characterization of the *procera* tomato mutant shows novel functions of the *SlDELLA* protein in the control of flower morphology, cell division and expansion, and the auxin-signaling pathway during fruit-set and development. *Plant Physiol.*, 160, 1581.

Chalabi, N., Le Corre, L., Maurizis, J.C., Bignon, Y.J., and Bernard-Gallon, D.J. (2004). The effects of lycopene on the proliferation of human breast cells and *BRCA1* and *BRCA2* gene expression. *Eur. J. Cancer*, 40, 1768-1775.

Clement, M., Leonhardt, N., Droillard, M.J., Reiter, I., Montillet, J.L., Genty, B., Lauriere, C., Nussaume, L., and Noel, L.D. (2011). The cytosolic/nuclear HSC70 and HSP90 molecular chaperones are important for stomatal closure and modulate abscisic acid-dependent physiological responses in Arabidopsis. *Plant Physiol.*, 156, 1481-1492.

Colbert, T., Till, B.J., Tompa, R., Reynolds, S., Steine, M.N., Yeung, A.T., McCallum, C.M., Comai, L., and Henikoff, S. (2001). High-throughput screening for induced point mutations. *Plant Physiol.*, 126, 480.

Cong, B., and Tanksley, S.D. (2006). FW2.2 and cell cycle control in developing tomato fruit: a possible example of gene co-option in the evolution of a novel organ. *Plant Mol. Biol.*, 62, 867-880.

Consortium, T.T.G. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, 485, 635-641.

Daminato, M., Masiero, S., Resentini, F., Lovisetto, A., and Casadoro, G. (2014). Characterization of *TM8*, a MADS-box gene expressed in tomato flowers. *BMC Plant. Biol.*, 14, 319.

de Jong, M., Mariani, C., and Vriezen, W.H. (2009a). The role of auxin and gibberellin in tomato fruit set. *J. Exp. Bot.*, 60, 1523-1532.

de Jong, M., Wolters-Arts, M., Feron, R., Mariani, C., and Vriezen, W.H. (2009b). The *Solanum lycopersicum* auxin response factor 7 (SlARF7) regulates auxin signaling during tomato fruit set and development. *Plant J.*, 57, 160-170.

De Maio, A. (1999). Heat shock proteins: facts, thoughts, and dreams. *Shock*, 11, 1-12.

Delph, L.F., Johannsson, M.H., and Stephenson, A.G. (1997). How environmental factors affect pollen performance: ecological and evolutionary perspectives. *Ecology*, 78, 1632-1639.

DiLeo, M.V., Strahan, G.D., den Bakker, M., and Hoekenga, O.A. (2011). Weighted correlation network analysis (WGCNA) applied to the tomato fruit metabolome. *PLOS ONE*, 6, e26683.

Ding, J., Chen, B., Xia, X., Mao, W., Shi, K., Zhou, Y., and Yu, J. (2013). Cytokinin-induced parthenocarpic fruit development in tomato is partly dependent on enhanced gibberellin and auxin biosynthesis. *PLOS ONE*, 8, e70080.

Djakovic-Petrovic, T., de Wit, M., Voesenek, L.A., and Pierik, R. (2007). DELLA protein function in growth responses to canopy signals. *Plant J.*, 51, 117-126.

Doi, M., Wachi, M., Ishino, F., Tomioka, S., Ito, M., Sakagami, Y., Suzuki, A., and Matsuhashi, M. (1988). Determinations of the DNA sequence of the mreB gene and of the gene products of the mre region that function in formation of the rod shape of Escherichia coli cells. *Journal of Bacteriology*, 170, 4619-4624.

Doyle, J. (1991). DNA Protocols for Plants. In Molecular techniques in taxonomy, G.M. Hewitt, A.W.B. Johnston, and J.P.W. Young, eds (Springer, Berlin, Heidelberg), pp. 283-293.

Duck, N., McCormick, S., and Winter, J. (1989). Heat shock protein hsp70 cognate gene expression in vegetative and reproductive organs of *Lycopersicon esculentum*. *Proceedings of the National Academy of Sciences*, 86, 3674.

Dumas, Y., Dadomo, M., Di Lucca, G., and Grolier, P. (2003). Effects of environmental factors and agricultural techniques on antioxidantcontent of tomatoes. *J. Sci. Food Agric.*, 83, 369-382.

El-Sharkawy, I., Sherif, S., El Kayal, W., Jones, B., Li, Z., Sullivan, A.J., and Jayasankar, S. (2016). Overexpression of Plum auxin receptor *PslTIR1* in tomato alters plant growth, fruit development and fruit shelf-life characteristics. *BMC Plant. Biol.*, 16, 56.

Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Prot.*, 2, 953.

Erba, D., Casiraghi, M.C., Ribas-Agustí, A., Cáceres, R., Marfà, O., and Castellari, M. (2013). Nutritional value of tomatoes (*Solanum lycopersicum* L.) grown in greenhouse by different agronomic techniques. *J. Food Compos. Anal.*, 31, 245-251.

Etminan, M., Takkouche, B., and Caamaño-Isorna, F. (2004). The role of tomato products and lycopene in the prevention of prostate cancer: A meta-analysis of observational studies. *Cancer Epidem. Biomar.*, 13, 340-345.

Fedorova, L., and Fedorov, A. (2003). Introns in gene evolution. *Genetica*, 118, 123-131.

Fei, Z., Joung, J.G., Tang, X., Zheng, Y., Huang, M., Lee, J.M., McQuinn, R., Tieman, D.M., Alba, R., Klee, H.J., and Giovannoni, J.J. (2011). Tomato Functional Genomics Database: a comprehensive resource and analysis package for tomato functional genomics. *Nucleic Acids Res.*, 39, D1156-1163.

Feng, Z.-J., He, G.-H., Zheng, W.-J., Lu, P.-P., Chen, M., Gong, Y.-M., Ma, Y.-Z., and Xu, Z.-S. (2015). Foxtail millet NF-Y families: Genome-wide survey and evolution analyses identified two functional genes important in abiotic stresses. *Front. Plant Sci.*, 6, 1142.

Fernandez-Pozo, N., Menda, N., Edwards, J.D., Saha, S., Tecle, I.Y., Strickler, S.R., Bombarely, A., Fisher-York, T., Pujar, A., Foerster, H., Yan, A., and Mueller, L.A. (2015). The Sol Genomics Network (SGN)--from genotype to phenotype to breeding. *Nucleic Acids Res.*, 43, D1036-1041.

Fuentes, S., Ljung, K., Sorefan, K., Alvey, E., Harberd, N.P., and Østergaard, L. (2012). Fruit growth in *Arabidopsis* occurs via DELLA-dependent and DELLA-independent gibberellin responses. *Plant Cell*, 24, 3982.

Fukushima, A., Nishizawa, T., Hayakumo, M., Hikosaka, S., Saito, K., Goto, E., and Kusano, M. (2012). Exploring tomato gene functions based on coexpression modules using graph clustering and differential coexpression approaches. *Plant Physiol.*, 158, 1487.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.*, 31, 3784-3788.

Genevaux, P., Georgopoulos, C., and Kelley, W.L. (2007). The Hsp70 chaperone machines of *Escherichia coli*: a paradigm for the repartition of chaperone functions. *Mol. Microbiol.*, 66, 840-857.

Giovannucci, E. (1999). Tomatoes, tomato-based products, lycopene, and cancer: Review of the epidemiologic literature. *JNCI-J. Natl. Cancer I.,* 91, 317-331.

Goda, H., Sawa, S., Asami, T., Fujioka, S., Shimada, Y., and Yoshida, S. (2004). Comprehensive comparison of auxin-regulated and brassinosteroid-regulated genes in Arabidopsis. *Plant Physiol.,* 134, 1555.

Goetz, M., Hooper, L.C., Johnson, S.D., Rodrigues, J.C.M., Vivian-Smith, A., and Koltunow, A.M. (2007). Expression of aberrant forms of *Auxin Response Factor8* stimulates parthenocarpy in arabidopsis and tomato. *Plant Physiol.,* 145, 351.

Gorguet, B., van Heusden, A.W., and Lindhout, P. (2005). Parthenocarpic fruit development in tomato. *Plant Biol. (Stuttg),* 7, 131-139.

Gorovits, R., and Czosnek, H. (2017). The involvement of heat shock proteins in the establishment of *Tomato Yellow Leaf Curl Virus* infection. *Front. Plant Sci.,* 8.

Gorovits, R., Moshe, A., Ghanim, M., and Czosnek, H. (2013). Recruitment of the host plant heat shock protein 70 by *Tomato Yellow Leaf Curl Virus* coat protein is required for virus infection. *PLOS ONE,* 8, e70280.

Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., Comai, L., and Henikoff, S. (2003). Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. *Genetics,* 164, 731-740.

Guo, A.-Y., Hu, B., Gao, G., Zhang, H., Luo, J., and Jin, J. (2014). GSDS 2.0: an upgraded gene feature visualization server. *Bioinform.,* 31, 1296-1297.

Guy, C.L., and Li, Q.B. (1998). The organization and evolution of the spinach stress 70 molecular chaperone gene family. *Plant Cell,* 10, 539-556.

Harberd, N.P. (2003). Relieving DELLA Restraint. *Science,* 299, 1853.

Harris Amrine, K., Blanco-Ulate, B., and Cantu, D. (2015). Discovery of core biotic stress responsive genes in Arabidopsis by weighted gene co-expression network analysis. *PLOS ONE,* 10, e0118731.

He, F., and Maslov, S. (2016). Pan- and core- network analysis of co-expression genes in a model plant. *Sci. Rep.,* 6, 38956.

Horvath, D.P., Schaffer, R., West, M., and Wisman, E. (2003). Arabidopsis microarrays identify conserved and differentially expressed genes involved in shoot growth and development from distantly related plant species. *Plant J.,* 34, 125-134.

Hu, B., Jin, J., Guo, A.Y., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinform.,* 31, 1296-1297.

Hu, J., Israeli, A., Ori, N., and Sun, T.-p. (2018). The Interaction between DELLA and ARF/IAA mediates crosstalk between gibberellin and auxin signaling to control fruit initiation in tomato. *Plant Cell,* 30, 1710.

Hurst, L.D. (2002). The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet.,* 18, 486.

Jander, G., Baerson, S.R., Hudak, J.A., Gonzalez, K.A., Gruys, K.J., and Last, R.L. (2003). Ethylmethanesulfonate saturation mutagenesis in Arabidopsis to determine frequency of herbicide resistance. *Plant Physiol.,* 131, 139-146.

Jiajie, P., Tao, W., Jianping, H., Yadong, W., and Jin, C. (2016). Constructing Networks of Organelle Functional Modules in Arabidopsis. *Current Genomics,* 17, 427-438.

Jian, L., Du, C.J., Lee, A.H., and Binns, C.W. (2005). Do dietary lycopene and other carotenoids protect against prostate cancer? *Int. J. Cancer,* 113, 1010-1014.

Jung, K.H., Gho, H.J., Nguyen, M.X., Kim, S.R., and An, G. (2013). Genome-wide expression analysis of HSP70 family genes in rice and identification of a cytosolic HSP70 gene highly induced under heat stress. *Funct. Integr. Genomics,* 13, 391-402.

Jungkunz, I., Link, K., Vogel, F., Voll, L.M., Sonnewald, S., and Sonnewald, U. (2011). AtHsp70-15-deficient Arabidopsis plants are characterized by reduced growth, a constitutive cytosolic protein response and enhanced resistance to TuMV. *Plant J.,* 66, 983-995.

Koenig, D., Jiménez-Gómez, J.M., Kimura, S., Fulop, D., Chitwood, D.H., Headland, L.R., Kumar, R., Covington, M.F., Devisetty, U.K., Tat, A.V., Tohge, T., Bolger, A., Schneeberger, K., Ossowski, S., Lanz, C., Xiong, G., Taylor-Teeples, M., Brady, S.M., Pauly, M., Weigel, D., Usadel, B., Fernie, A.R., Peng, J., Sinha, N.R., and Maloof, J.N. (2013). Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. *Proc. Natl. Acad. Sci. U. S. A.,* 110, E2655.

Koornneef, M., and Meinke, D. (2010). The development of Arabidopsis as a model plant. *Plant J.,* 61, 909-921.

Kraft, E., Stone, S.L., Ma, L., Su, N., Gao, Y., Lau, O.-S., Deng, X.-W., and Callis, J. (2005). Genome analysis and functional characterization of the E2 and ring-type E3 ligase ubiquitination enzymes of Arabidopsis. *Plant Physiol.,* 139, 1597.

Krieg, D.R. (1963). Ethyl methanesulfonate-induced reversion of bacteriophage T4rII mutants. *Genetics*, 48, 561-580.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.*, 33, 1870-1874.

Kurabayashi, A., Suda, K., Yamamoto, N., Iijima, Y., Ogata, Y., Aoki, K., Shibata, D., Ozaki, S., Suzuki, T., Tsugane, T., Konishi, C., Inai, S., Fujii, T., Yamazaki, M., and Bunsupa, S. (2010). Coexpression analysis of tomato genes and experimental verification of coordinated expression of genes found in a functionally enriched coexpression module. *DNA Res.*, 17, 105-116.

Kurowska, M., Daszkowska-Golec, A., Gruszka, D., Marzec, M., Szurman, M., Szarejko, I., and Maluszynski, M. (2011). TILLING: a shortcut in functional genomics. *J. Appl. Genet.*, 52, 371-390.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.*, 9, 559.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., and Higgins, D.G. (2007). Clustal W and Clustal X version 2.0. *Bioinform.*, 23, 2947-2948.

Li, H., Wang, L., and Yang, Z.M. (2015). Co-expression analysis reveals a group of genes potentially involved in regulation of plant response to iron-deficiency. *Gene*, 554, 16-24.

Li, L., Xing, Y., Chang, D., Fang, S., Cui, B., Li, Q., Wang, X., Guo, S., Yang, X., Men, S., and Shen, Y. (2016). CaM/BAG5/Hsc70 signaling complex dynamically regulates leaf senescence. *Sci. Rep.*, 6, 31889.

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinform.*, 25, 1451-1452.

Lin, B.L., Wang, J.S., Liu, H.C., Chen, R.W., Meyer, Y., Barakat, A., and Delseny, M. (2001). Genomic analysis of the Hsp70 superfamily in Arabidopsis thaliana. *Cell Stress Chaperones*, 6, 201-208.

Lin, H., Yu, J., Pearce, S.P., Zhang, D., and Wilson, Z.A. (2017). RiceAntherNet: a gene co-expression network for identifying anther and pollen development genes. *Plant J.*, 92, 1076-1091.

Liu, B.H., Yu, H., Tu, K., Li, C., Li, Y.X., and Li, Y.Y. (2010). DCGL: an R package for identifying differentially coexpressed genes and links from gene expression microarray data. *Bioinform.,* 26, 2637-2638.

Liu, D., Sun, W., Yuan, Y., Zhang, N., Hayward, A., Liu, Y., and Wang, Y. (2014). Phylogenetic analyses provide the first insights into the evolution of OVATE family proteins in land plants. *Ann. Bot.,* 113, 1219-1233.

Liu, J., Pang, X., Cheng, Y., Yin, Y., Zhang, Q., Su, W., Hu, B., Guo, Q., Ha, S., Zhang, J., and Wan, H. (2018). The *HSP70* gene family in *Solanum tuberosum*: Genome-wide identification, phylogeny, and expression patterns. *Sci. Rep.,* 8, 16628-16628.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCt method. *Methods,* 25, 402-408.

Livne, S., Lor, V.S., Nir, I., Eliaz, N., Aharoni, A., Olszewski, N.E., Eshed, Y., and Weiss, D. (2015). Uncovering DELLA-independent gibberellin responses by characterizing new tomato *procera* mutants. *Plant Cell,* 27, 1579-1594.

Llamas, E., Pulido, P., and Rodriguez-Concepcion, M. (2017). Interference with plastome gene expression and Clp protease activity in Arabidopsis triggers a chloroplast unfolded protein response to restore protein homeostasis. *PLOS Genetics,* 13, e1007022.

Lombardi-Crestana, S., da Silva Azevedo, M., e Silva, G.F.F., Pino, L.E., Appezzato-da-Glória, B., Figueira, A., Nogueira, F.T.S., and Peres, L.E.P. (2012). The tomato (*Solanum Lycopersicum* cv. Micro-Tom) natural genetic variation Rg1 and the DELLA mutant *procera* control the competence necessary to form adventitious roots and shoots. *J. Exp. Bot.,* 63, 5689-5703.

Ma, S., Gong, Q., and Bohnert, H.J. (2007). An Arabidopsis gene network based on the graphical Gaussian model. *Genome Res.,* 17, 1614-1625.

Ma, X., Zhao, H., Xu, W., You, Q., Yan, H., Gao, Z., and Su, Z. (2018). Co-expression Gene Network Analysis and Functional Module Identification in Bamboo Growth and Development. *Frontiers in Genetics,* 9.

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinform.,* 21, 3448-3449.

Maimbo, M., Ohnishi, K., Hikichi, Y., Yoshioka, H., and Kiba, A. (2007). Induction of a small heat shock protein and its functional roles in Nicotiana plants in the defense response against *Ralstonia solanacearum*. *Plant Physiol.*, 145, 1588-1599.

Majumdar, R., Shao, L., Turlapati, S.A., and Minocha, S.C. (2017). Polyamines in the life of Arabidopsis: profiling the expression of S-adenosylmethionine decarboxylase (SAMDC) gene family during its life cycle. *BMC Plant. Biol.*, 17, 264.

Mandal, S., Ji, W., and McKnight, T.D. (2018). Candidate Gene Networks for Acylsugar Metabolism and Plant Defense in Wild Tomato Solanum pennellii. *bioRxiv*, 294306.

Mao, L., Van Hemert, J.L., Dash, S., and Dickerson, J.A. (2009). Arabidopsis gene co-expression network and its functional modules. *BMC Bioinform.*, 10, 346.

Marti, C., Orzaez, D., Ellul, P., Moreno, V., Carbonell, J., and Granell, A. (2007). Silencing of DELLA induces facultative parthenocarpy in tomato fruits. *Plant J.*, 52, 865-876.

Martinelli, F., Uratsu, S.L., Reagan, R.L., Chen, Y., Tricoli, D., Fiehn, O., Rocke, D.M., Gasser, C.S., and Dandekar, A.M. (2009). Gene regulation in parthenocarpic tomato fruit. *J. Exp. Bot.*, 60, 3873-3890.

Mathieu-Rivet, E., Gevaudant, F., Cheniclet, C., Hernould, M., and Chevalier, C. (2010a). The anaphase promoting complex activator CCS52A, a key factor for fruit growth and endoreduplication in tomato. *Plant Signal. Behav.*, 5, 985-987.

Mathieu-Rivet, E., Gevaudant, F., Sicard, A., Salar, S., Do, P.T., Mouras, A., Fernie, A.R., Gibon, Y., Rothan, C., Chevalier, C., and Hernould, M. (2010b). Functional analysis of the anaphase promoting complex activator CCS52A highlights the crucial role of endo-reduplication for fruit growth in tomato. *Plant J.*, 62, 727-741.

Matsuo, S., Kikuchi, K., Fukuda, M., Honda, I., and Imanishi, S. (2012). Roles and regulation of cytokinins in tomato fruit development. *J. Exp. Bot.*, 63, 5569-5579.

Mazzucato, A., Cellini, F., Bouzayen, M., Zouine, M., Mila, I., Minoia, S., Petrozza, A., Picarella, M.E., Ruiu, F., and Carriero, F. (2015). A TILLING allele of the tomato *Aux/IAA9* gene offers new insights into fruit set mechanisms and perspectives for breeding seedless tomatoes. *Mol. Breeding*, 35, 22.

McCallum, C.M., Comai, L., Greene, E.A., and Henikoff, S. (2000). Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol.*, 123, 439-442.

Meissner, R., Jacobson, Y., Melamed, S., Levyatuv, S., Shalev, G., Ashri, A., Elkind, Y., and Levy, A. (1997). A new model system for tomato genetics. *Plant J.*, 12, 1465-1472.

Menda, N., Semel, Y., Peled, D., Eshed, Y., and Zamir, D. (2004). In silico screening of a saturated mutation library of tomato. *Plant J.*, 38, 861-872.

Meyer, P.E., Lafitte, F., and Bontempi, G. (2008). Minet: A R/Bioconductor package for inferring large transcriptional networks using mutual information. *BMC Bioinform.*, 9, 461.

Meyer, P.E., Kontos, K., Lafitte, F., and Bontempi, G. (2007). Information-theoretic inference of large transcriptional regulatory networks. *EURASIP J. Bioinform. Syst. Biol.*, 79879.

Morris, R.T., O'Connor, T.R., and Wyrick, J.J. (2008). Osiris: an integrated promoter database for Oryza sativa L. *Bioinform.*, 24, 2915-2917.

Mulholland, B.J., Edmondson, R.N., Fussell, M., Basham, J., and Ho, L.C. (2003). Effects of high temperature on tomato summer fruit quality. *J. Hortic. Sci. Biotech.*, 78, 365-374.

Muruganujan, A., Mi, H., and Thomas, P.D. (2012). PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.*, 41, D377-D386.

Mutwil, M., Usadel, B., Schütte, M., Loraine, A., Ebenhöh, O., and Persson, S. (2010). Assembly of an Interactive Correlation Network for the Arabidopsis Genome Using a Novel Heuristic Clustering Algorithm. *Plant Physiol.*, 152, 29-43.

Nakaune, M., Hanada, A., Yin, Y.G., Matsukura, C., Yamaguchi, S., and Ezura, H. (2012). Molecular and physiological dissection of enhanced seed germination using short-term low-concentration salt seed priming in tomato. *Plant Physiol. Biochem.*, 52, 28-37.

Nei, M., and Saitou, N. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, 4, 406-425.

Nir, I., Shohat, H., Panizel, I., Olszewski, N., Aharoni, A., and Weiss, D. (2017). The Tomato DELLA protein PROCERA acts in guard cells to promote stomatal closure. *Plant Cell*, 29, 3186.

Noël, L.D., Cagna, G., Stuttmann, J., Wirthmüller, L., Betsuyaku, S., Witte, C.-P., Bhat, R., Pochon, N., Colby, T., and Parker, J.E. (2007). Interaction between SGT1 and

cytosolic/nuclear HSC70 chaperones regulates *Arabidopsis* immune responses. *Plant Cell,* 19, 4061-4076.

Obayashi, T., Aoki, Y., Tadaka, S., Kagaya, Y., and Kinoshita, K. (2017). ATTED-II in 2018: A Plant Coexpression Database Based on Investigation of the Statistical Property of the Mutual Rank Index. *Plant Cell Physiol.,* 59, e3-e3.

Okabe, Y., Asamizu, E., Saito, T., Matsukura, C., Ariizumi, T., Bres, C., Rothan, C., Mizoguchi, T., and Ezura, H. (2011). Tomato TILLING technology: development of a reverse genetics tool for the efficient isolation of mutants from Micro-Tom mutant libraries. *Plant Cell Physiol.,* 52, 1994-2005.

Oladosu, Y., Rafii, M.Y., Abdullah, N., Hussin, G., Ramli, A., Rahim, H.A., Miah, G., and Usman, M. (2016). Principle and application of plant mutagenesis in crop improvement: a review. *Biotechnol. Biotechnol. Equip.,* 30, 1-16.

Østergaard, L., and Yanofsky, M.F. (2004). Establishing gene function by mutagenesis in *Arabidopsis thaliana*. *Plant J.,* 39, 682-696.

Ozaki, S., Ogata, Y., Suda, K., Kurabayashi, A., Suzuki, T., Yamamoto, N., Iijima, Y., Tsugane, T., Fujii, T., Konishi, C., Inai, S., Bunsupa, S., Yamazaki, M., Shibata, D., and Aoki, K. (2010). Coexpression Analysis of Tomato Genes and Experimental Verification of Coordinated Expression of Genes Found in a Functionally Enriched Coexpression Module. *DNA Res.,* 17, 105-116.

Pandolfini, T. (2009). Seedless fruit production by hormonal regulation of fruit set. *Nutrients,* 1, 168-177.

Piya, S., Shrestha, S.K., Binder, B., Stewart, C.N., and Hewezi, T. (2014). Protein-protein interaction and gene co-expression maps of ARFs and Aux/IAAs in Arabidopsis 5.

Prasch, C.M., and Sonnewald, U. (2013). Simultaneous Application of Heat, Drought, and Virus to Arabidopsis Plants Reveals Significant Shifts in Signaling Networks. *Plant Physiol.,* 162, 1849-1866.

Puranik, S., Acajjaoui, S., Conn, S., Costa, L., Conn, V., Vial, A., Marcellin, R., Melzer, R., Brown, E., Hart, D., Theißen, G., Silva, C.S., Parcy, F., Dumas, R., Nanao, M., and Zubieta, C. (2014). Structural basis for the oligomerization of the MADS domain transcription factor SEPALLATA3 in *Arabidopsis*. *Plant Cell,* 26, 3603-3615.

Rao, X., Chen, X., Shen, H., Ma, Q., Li, G., Tang, Y., Pena, M., York, W., Frazier, T.P., Lenaghan, S., Xiao, X., Chen, F., and Dixon, R.A. (2019). Gene regulatory networks

for lignin biosynthesis in switchgrass (*Panicum virgatum*). *Plant Biotechnol. J.*, 17, 580-593.

Rasmussen, S., Barah, P., Suarez-Rodriguez, M.C., Bressendorff, S., Friis, P., Costantino, P., Bones, A.M., Nielsen, H.B., and Mundy, J. (2013). Transcriptome Responses to Combinations of Stresses in Arabidopsis. *Plant Physiol.*, 161, 1783-1794.

Ren, Z., Li, Z., Miao, Q., Yang, Y., Deng, W., and Hao, Y. (2011). The auxin receptor homologue in *Solanum lycopersicum* stimulates tomato fruit set and leaf morphogenesis. *J. Exp. Bot.*, 62, 2815-2826.

Ross, J.J., O'Neill, D.P., Smith, J.J., Kerckhoffs, L.H.J., and Elliott, R.C. (2000). Evidence that auxin promotes gibberellin A1 biosynthesis in pea. *Plant J.*, 21, 547-552.

Saito, K., Hirai, M.Y., and Yonekura-Sakakibara, K. (2008). Decoding genes with coexpression networks and metabolomics – 'majority report by precogs'. *Trends Plant Sci.*, 13, 36-43.

Saito, T., Ariizumi, T., Okabe, Y., Asamizu, E., Hiwasa-Tanase, K., Fukuda, N., Mizoguchi, T., Yamazaki, Y., Aoki, K., and Ezura, H. (2011). TOMATOMA: a novel tomato mutant database distributing Micro-Tom mutant collections. *Plant Cell Physiol.*, 52, 283-296.

Sarkar, N.K., Kundnani, P., and Grover, A. (2013). Functional analysis of Hsp70 superfamily proteins of rice (*Oryza sativa*). *Cell Stress Chaperones*, 18, 427-437.

Schadt, E.E., Linderman, M.D., Sorenson, J., Lee, L., and Nolan, G.P. (2010). Computational solutions to large-scale data management and analysis. *Nat. Rev. Genet.*, 11, 647.

Schlueter, J.A., Lin, J.-Y., Schlueter, S.D., Vasylenko-Sanders, I.F., Deshpande, S., Yi, J., O'Bleness, M., Roe, B.A., Nelson, R.T., Scheffler, B.E., Jackson, S.A., and Shoemaker, R.C. (2007). Gene duplication and paleopolyploidy in soybean and the implications for whole genome sequencing. *BMC Genomics*, 8, 330.

Semon, M., and Wolfe, K.H. (2007). Consequences of genome duplication. *Curr. Opin. Genet. Dev.*, 17, 505-512.

Serin, E.A.R., Nijveen, H., Hilhorst, H.W.M., and Ligterink, W. (2016). Learning from Co-expression Networks: Possibilities and Challenges. *Front. Plant Sci.*, 7, 444-444.

Serrani, J.C., Sanjuán, R., Ruiz-Rivero, O., Fos, M., and García-Martínez, J.L. (2007). Gibberellin regulation of fruit set and growth in tomato. *Plant Physiol.*, 145, 246.

Shahan, R., Zawora, C., Wight, H., Sittmann, J., Wang, W., Mount, S.M., and Liu, Z. (2018). Consensus Coexpression Network Analysis Identifies Key Regulators of Flower and Fruit Development in Wild Strawberry. *Plant Physiol.*, 178, 202-216.

Sharma, D., and Masison, D.C. (2009). Hsp70 structure, function, regulation and influence on yeast prions. *Protein Pept. Lett.*, 16, 571-581.

Shikata, M., and Ezura, H. (2016). Micro-Tom tomato as an alternative plant model system: Mutant collection and efficient transformation. In Plant signal transduction: Methods and protocols, methods in molecular biology, J.R. Botella and M.A. Botella, eds ( New York: Springer Science+Business Media), pp. 47-55.

Shimamoto, K., and Kyozuka, J. (2002). Rice as a model for comparative genomics of plants. *Annu. Rev. Plant Biol.*, 53, 399-419.

Smaczniak, C., Immink, R.G.H., Muiño, J.M., Blanvillain, R., Busscher, M., Busscher-Lange, J., Dinh, Q.D.P., Liu, S., Westphal, A.H., Boeren, S., Parcy, F., Xu, L., Carles, C.C., Angenent, G.C., and Kaufmann, K. (2012). Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development. *Proc. Natl. Acad. Sci. U. S. A.*, 109, 1560-1565.

Smock, R.G., Rivoire, O., Russ, W.P., Swain, J.F., Leibler, S., Ranganathan, R., and Gierasch, L.M. (2010). An interdomain sector mediating allostery in Hsp70 molecular chaperones. *Mol. Syst. Biol.*, 6, 414.

Soufo, H.J., and Graumann, P.L. (2003). Actin-like proteins MreB and Mbl from Bacillus subtilis are required for bipolar positioning of replication origins. *Current biology : CB*, 13, 1916-1920.

Spena, A., and Rotino, G.L. (2001). Parthenocarpy. In Current Trends in the Embryology of Angiosperms, S.S. Bhojwani and W.-Y. Soh, eds (Dordrecht: Springer Netherlands), pp. 435-450.

Su, P.-H., and Li, H.-m. (2008). Arabidopsis Stromal 70-kD Heat Shock Proteins Are Essential for Plant Development and Important for Thermotolerance of Germinating Seeds. *Plant Physiol.*, 146, 1231-1241.

Sung, D.-Y., Kaplan, F., and Guy, C.L. (2001a). Plant Hsp70 molecular chaperones: Protein structure, gene family, expression and function. *Physiol. Plantarum*, 113, 443-451.

Sung, D.Y., Vierling, E., and Guy, C.L. (2001b). Comprehensive expression profile analysis of the Arabidopsis Hsp70 gene family. *Plant Physiol.*, 126, 789-800.

Suzuki, T., Eiguchi, M., Kumamaru, T., Satoh, H., Matsusaka, H., Moriguchi, K., Nagato, Y., and Kurata, N. (2008). MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. *Mol. Genet. Genomics,* 279, 213-223.

Tai, Y., Liu, C., Yu, S., Yang, H., Sun, J., Guo, C., Huang, B., Liu, Z., Yuan, Y., Xia, E., Wei, C., and Wan, X. (2018). Gene co-expression network analysis reveals coordinated regulation of three characteristic secondary biosynthetic pathways in tea plant (*Camellia sinensis*). *BMC Genomics,* 19, 616.

Tang, T., Yu, A., Li, P., Yang, H., Liu, G., and Liu, L. (2016). Sequence analysis of the Hsp70 family in moss and evaluation of their functions in abiotic stress responses. *Sci. Rep.,* 6, 33650.

Tanksley, S.D., Ganal, M.W., Prince, J.P., de-Vicente, M.C., Bonierbale, M.W., Broun, P., Fulton, T.M., Giovannoni, J.J., Grandillo, S., Martin, G.B., Messeguer, R., Miller, J.C., Miller, L., Paterson, A.H., Pineda, O., Roder, M.S., Wing, R.A., Wu, W., and Young, N.D. (1992). High Density Molecular Linkage Maps of the Tomato and Potato Genomes. *Genetics,* 132, 1141-1160.

Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., and Narechania, A. (2003). PANTHER: A library of protein families and subfamilies indexed by function. *Genome Res.,* 13, 2129-2141.

Thompson, J.D., Gibson, T.J., and Higgins, D.G. (2003). Multiple Sequence Alignment Using ClustalW and ClustalX. *Curr. Protoc. Bioinformatics,* 00, 2.3.1-2.3.22.

Toro-Domínguez, D., Martorell-Marugán, J., López-Domínguez, R., García-Moreno, A., González-Rumayor, V., Alarcón-Riquelme, M.E., and Carmona-Sáez, P. (2018). ImaGEO: integrative gene expression meta-analysis from GEO database. *Bioinform.,* 35, 880-882.

Tzfadia, O., Diels, T., De Meyer, S., Vandepoele, K., Aharoni, A., and Van de Peer, Y. (2016). CoExpNetViz: Comparative Co-Expression Networks Construction and Visualization Tool. *Front. Plant Sci.,* 6.

van Dam, S., Võsa, U., van der Graaf, A., Franke, L., and de Magalhães, J.P. (2017). Gene co-expression analysis for functional classification and gene-disease predictions. *Brief. Bioinform.,* 19, 575-592.

van den Heuvel, K.J.P.T., Barendse, G.W.M., and Wullems, G.J. (2001). Effect of gibberellic acid on cell division and cell elongation in anthers of the gibberellin deficient gib-1 mutant of tomato. *Plant Biol.,* 3, 124-131.

van Veen, H., Vashisht, D., Akman, M., Girke, T., Mustroph, A., Reinen, E., Hartman, S., Kooiker, M., van Tienderen, P., Schranz, M.E., Bailey-Serres, J., Voesenek, L.A.C.J., and Sasidharan, R. (2016). Transcriptomes of Eight Arabidopsis thaliana Accessions Reveal Core Conserved, Genotype- and Organ-Specific Responses to Flooding Stress. *Plant Physiol.,* 172, 668-689.

Vanthoor, B.H.E., de Visser, P.H.B., Stanghellini, C., and van Henten, E.J. (2011). A methodology for model-based greenhouse design: Part 2, description and validation of a tomato yield model. *Biosyst. Eng.,* 110, 378-395.

Walsh, P., Bursac, D., Law, Y.C., Cyr, D., and Lithgow, T. (2004). The J-protein family: modulating protein assembly, disassembly and translocation. *EMBO Rep.,* 5, 567-571.

Wang, H., Gu, L., Zhang, X., Liu, M., Jiang, H., Cai, R., Zhao, Y., and Cheng, B. (2018). Global transcriptome and weighted gene co-expression network analyses reveal hybrid-specific modules and candidate genes related to plant height development in maize. *Plant Mol. Biol.,* 98, 187-203.

Wang, H., Jones, B., Li, Z., Frasse, P., Delalande, C., Regad, F., Chaabouni, S., Latche, A., Pech, J.C., and Bouzayen, M. (2005). The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. *Plant Cell,* 17, 2676-2692.

Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.,* 9, 244-252.

Wang, X., Feng, S., Nakayama, N., Crosby, W.L., Irish, V., Deng, X.W., and Wei, N. (2003). The COP9 signalosome interacts with SCF $^{UFO}$ and participates in Arabidopsis flower development. *Plant Cell,* 15, 1071-1082.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.,* 10, 57-63.

Warf, M.B., and Berglund, J.A. (2010). Role of RNA structure in regulating pre-mRNA splicing. *Trends Biochem. Sci.,* 35, 169-178.

Watanabe, S., Mizoguchi, T., Aoki, K., Kubo, Y., Mori, H., Imanishi, S., Yamazaki, Y., Shibata, D., and Ezura, H. (2007). Ethylmethanesulfonate (EMS) mutagenesis of

*Solanum lycopersicum* cv. Micro-Tom for large-scale mutant screens. *Plant Biotechnol. J.*, 24, 33-38.

Willcox, J.K., Catignani, G.L., and Lazarus, S. (2003). Tomatoes and cardiovascular health. *Crit. Rev. Food Sci. Nutr.*, 43, 1-18.

Wisecaver, J.H., Borowsky, A.T., Tzin, V., Jander, G., Kliebenstein, D.J., and Rokas, A. (2017). A global coexpression network approach for connecting genes to specialized metabolic pathways in plants. *Plant Cell*, 29, 944-959.

Wong, J.H., Kiss, F., Wu, M.-X., and Buchanan, B.B. (1990). Pyrophosphate Fructose-6-P 1-Phosphotransferase from Tomato Fruit. *Plant Physiol.*, 94, 499.

Wu, J., Peng, Z., Liu, S., He, Y., Cheng, L., Kong, F., Wang, J., and Lu, G. (2012). Genome-wide analysis of Aux/IAA gene family in Solanaceae species using tomato as a model. *Mol. Genet. Genomics*, 287, 295-211.

Wu, S.-H., Wang, C., Chen, J., and Lin, B.-L. (1994). Isolation of a cDNA encoding a 70 kDa heat-shock cognate protein expressed in vegetative tissues ofArabidopsis thaliana. *Plant Mol. Biol.*, 25, 577-583.

Xiao, J., Li, H., Zhang, J., Chen, R., Zhang, Y., Ouyang, B., Wang, T., and Ye, Z. (2006). Dissection of GA 20-oxidase members affecting tomato morphology by RNAi-mediated silencing. *Plant Growth Regul.*, 50, 179-189.

Xiao, S., Xu, J., Zhao, J., Liu, J., Wu, Q., Yu, J., and Yu, D. (2017). Genomic, evolutionary and expression profile analysis of Hsp70 superfamily in A and D genome of cotton (*Gossypium* spp.) under the challenge of *Verticillium dahliae*. *J. Plant Biol.*, 60, 11-25.

Xu, H. (2018). Cochaperones enable Hsp70 to use ATP energy to stabilize native proteins out of the folding equilibrium. *Sci. Rep.*, 8, 13213.

Yamamoto, N., Tsugane, T., Watanabe, M., Yano, K., Maeda, F., Kuwata, C., Torki, M., Ban, Y., Nishimura, S., and Shibata, D. (2005). Expressed sequence tags from the laboratory-grown miniature tomato (Lycopersicon esculentum) cultivar Micro-Tom and mining for single nucleotide polymorphisms and insertions/deletions in tomato cultivars. *Gene*, 356, 127-134.

Yoshida, H., Hirano, K., Sato, T., Mitsuda, N., Nomoto, M., Maeo, K., Koketsu, E., Mitani, R., Kawamura, M., Ishiguro, S., Tada, Y., Ohme-Takagi, M., Matsuoka, M., and Ueguchi-Tanaka, M. (2014). DELLA protein functions as a transcriptional activator

through the DNA binding of the indeterminate domain family proteins. *Proc Natl Acad Sci U S A*, 111, 7861-7866.

Zegzouti, H., Jones, B., Frasse, P., Marty, C., Maitre, B., Latch, A., Pech, J.C., and Bouzayen, M. (1999). Ethylene-regulated gene expression in tomato fruit: characterization of novel ethylene-responsive and ripening-related genes isolated by differential display. *Plant J.*, 18, 589-600.

Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.*, 4, Article17.

Zhang, J., Chen, R., Xiao, J., Qian, C., Wang, T., Li, H., Ouyang, B., and Ye, Z. (2007). A single-base deletion mutation in SlIAA9 gene causes tomato (*Solanum lycopersicum*) entire mutant. *J. Plant Res.*, 120, 671-678.

Zhang, J., Liu, B., Li, J., Zhang, L., Wang, Y., Zheng, H., Lu, M., and Chen, J. (2015a). Hsf and Hsp gene families in Populus: genome-wide identification, organization and correlated expression during development and in stress responses. *BMC Genomics*, 16, 181.

Zhang, L., Zhang, Q., Gao, Y., Pan, H., Shi, S., and Wang, Y. (2014). Overexpression of heat shock protein gene *PfHSP21.4* in *Arabidopsis thaliana* enhances heat tolerance. *Acta Physiol. Plant*, 36, 1555-1564.

Zhang, L., Zhao, H.-k., Dong, Q.-l., Zhang, Y.-y., Wang, Y.-m., Li, H.-y., Xing, G.-j., Li, Q.-y., and Dong, Y.-s. (2015b). Genome-wide analysis and expression profiling under heat and drought treatments of HSP70 gene family in soybean (*Glycine max* L.). *Front. Plant Sci.*, 6, 773.

Zheng, X., Liu, T., Yang, Z., and Wang, J. (2011). Large cliques in Arabidopsis gene coexpression network and motif discovery. *J. Plant Physiol.*, 168, 611-618.

Zou, J., Liu, C., Liu, A., Zou, D., and Chen, X. (2012). Overexpression of *OsHsp17*.0 and *OsHsp23.7* enhances drought and salt tolerance in rice. *J. Plant Physiol.*, 169, 628-635.