Molecular and In Silico Characterization of Tomato LBD Transcription Factor

Lee Yun Sheng (70165)

Bachelor of Science with Honours
(Resource Biotechonology)

2022

Molecular and In Silico Characterization of Tomato LBD Transcription Factor

Lee Yun Sheng

A thesis submitted in partial fulfilment of the Requirement of The Degree Bachelor of Science with Honours
(Resource Biotechnology)

SUPERVISOR: Madam Safarina Binti Ahmad

Programme of Resource Biotechnology
Faculty of Resource Science and Technology
UNIVERSITI MALAYSIA SARAWAK
2022

UNIVERSITI MALAYSIA SARAWAK

Grade: _____

Please tick (√)

Final Year Project Report	✓
Masters	
PhD	

DECLARATION OF ORIGINAL WORK

This declaration is made on the 15th day of June 2022.

Student's Declaration:

I Lee Yun Sheng, 70165, Faculty of Resource Science and Technology (PLEASE INDICATE STUDENT'S NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled, Molecular and In Silico Characterization of Tomato LBD Transcription Factor is my original work. I have not copied from any other students' work or from any other sources except where due reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

15th June 2022
Date submitted


Lee Yun Sheng (70165)

Supervisor's Declaration:

I, Mdm Safarina binti Ahmad (SUPERVISOR'S NAME), hereby certify that the work entitled, Molecular and In Silico Characterization of Tomato LBD Transcription Factor (TITLE) was prepared by the above named student, and was submitted to the "FACULTY" as a * partial/full fulfillment for the conferment of Bachelor of Science with Honours (Resource Biotechnology) (PLEASE INDICATE THE DEGREE), and the aforementioned work, to the best of my knowledge, is the said student's work

Received for examination by: 
(Safarina Binti Ahmad)

Date: 6 July 2022

I declare this Project/Thesis is classified as (Please tick (√)):

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*** RESTRICTED** (Contains restricted information as specified by the organisation where research was done)*****

☑ **OPEN ACCESS**

Validation of Project/Thesis

I therefore duly affirmed with free consent and willingness declared that this said Project/Thesis shall be placed officially in the Centre for Academic Information Services with the abide interest and rights as follows:

- This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
- The Centre for Academic Information Services has the lawful right to make copies for the purpose of academic and research only and not for other purpose.
- The Centre for Academic Information Services has the lawful right to digitise the content to for the Local Content Database.
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis for academic exchange between Higher Learning Institute.
- No dispute or any claim shall arise from the student himself / herself neither third party on this Project/Thesis once it becomes sole property of UNIMAS.
- This Project/Thesis or any material, data and information related to it shall not be distributed, published or disclosed to any party by the student except with UNIMAS permission.

Student's signature_____ Supervisor's signature: _____

(15th June 2022) (6th July 2022)

Current Address:
454, Lorong Taman Hui Sing 6/2, Taman Hui Sing, 93350 Kuching, Sarawak.

Notes: ***** If the Project/Thesis is **CONFIDENTIAL** or **RESTRICTED**, please attach together as annexure a letter from the organisation with the period and reasons of confidentiality and restriction.

[The instrument was duly prepared by The Centre for Academic Information Services]

ACKNOWLEDGEMENT

For the successful completion of my Final Year Project 2 (FYP2) during this last semester, I would like to give my sincere appreciation to my supervisor, Madam Safarina binti Ahmad, and lecturer Associate Professor Dr Hairul Azman bin Roslan for giving me consultation and guidance regarding the project and thesis write up through both face-to-face and online sessions. Thank you very much for supervisor and lecturer's help and support throughout the process and encourage me to finish the project successfully.

Besides, I would also like to give my sincerest thank you to my family members who gave me emotional support throughout the FYP period. I would also like to thank you to my course mates for giving support to each other and learning the process together to complete task in FYP.

Molecular and In Silico Characterization of Tomato LBD Transcription Factor

Lee Yun Sheng

Resource Biotechnology Programme
Faculty of Resource Science and Technology
Universiti Malaysia Sarawak

ABSTRACT

Lateral Organ Boundaries Domain (LBD) is a family of plant specific transcription factors. LBD plays regulatory roles in various biological process including organ development, regeneration, pollen development, biotic and abiotic stress response, anthocyanin, and nitrogen metabolisms in various plant species. Lateral root growth and development is controlled by two important LBD genes, namely LBD Containing Protein 16 (LBD16) and LBD18 in *Arabidopsis thaliana*. The expression of LBD16 and LBD18 could be upregulated or downregulated by environmental stress which affect the growth of root in plant. LBD16 and LBD18 were reported as the key LBD genes in *A. thaliana* which involved in lateral root formation. However, there are no comprehensive studies were found yet to discuss the LBD16 and LBD18 genes in tomato, scientifically known as *Solanum lycopersicum*. In this study, the retrieval data for LBD genes in *A. thaliana* and *S. lycopersicum* had been carried out and phylogenetic tree was constructed by using MEGA 11 software to find the corresponding homolog group of *S. lycopersicum* that contain LBD16 and LBD18. The identified group of *SlLBD* genes that might contain LBD16 and LBD18 were analyzed for its sequence similarity with BLAST tool by comparing with the model plant *A. thaliana* sequence. Further analysis, namely collinearity and expression analysis had been carried out on the selected *SlLBD* genes by TBtools to analyze its evolutionary relationship with model plant and its gene expression pattern under various conditions on heat map. The selected *SlLBD* genes were confirmed to be the candidate genes that contain the LBD16 and LBD18. Further analysis of these genes needs to be carried out in future work by laboratory work with various treatments to test its expression level on tomato root.

Key words: Lateral Organ Boundaries Domain, lateral root development, *Solanum lycopersicum*.

ABSTRAK

Domain Sempadan Organ Lateral (LBD) ialah keluarga faktor transkripsi khusus tumbuhan. LBD memainkan peranan pengawalseliaan dalam pelbagai proses biologi termasuk pembangunan organ, penjanaan semula, pembangunan debunga, tindak balas tekanan biotik dan abiotik, antosianin dan metabolisme nitrogen dalam pelbagai spesies tumbuhan. Pertumbuhan dan perkembangan akar lateral dikawal oleh dua gen LBD yang penting, iaitu LBD Mengandungi Protein 16 (LBD16) dan LBD18 dalam Arabidopsis thaliana*. Ekspresi LBD16 dan LBD18 boleh dikawal atau dikurangkan oleh tekanan persekitaran yang menjejaskan pertumbuhan akar dalam tumbuhan. LBD16 dan LBD18 dilaporkan sebagai gen LBD utama dalam* A. thaliana *yang terlibat dalam pembentukan akar sisi. Walau bagaimanapun, tiada kajian komprehensif ditemui lagi untuk membincangkan gen LBD16 dan LBD18 dalam tomato, secara saintifik dikenali sebagai* Solanum lycopersicum*. Dalam kajian ini, data perolehan untuk gen LBD dalam* A. thaliana *dan* S. lycopersicum *telah dijalankan dan pokok filogenetik telah dibina dengan menggunakan perisian MEGA 11 untuk mencari kumpulan homolog sepadan* S. lycopersicum *yang mengandungi LBD16 dan LBD18. Kumpulan gen SlLBD yang dikenal pasti yang mungkin mengandungi LBD16 dan LBD18 telah dianalisis untuk persamaan jujukannya dengan alat BLAST dengan membandingkan dengan jujukan tumbuhan model A. thaliana. Analisis lanjut, iaitu analisis kolineariti dan ekspresi telah dijalankan ke atas gen SlLBD terpilih oleh TBtools untuk menganalisis hubungan evolusinya dengan tumbuhan model dan corak ekspresi gennya di bawah pelbagai keadaan pada peta haba. Gen SlLBD yang dipilih telah disahkan sebagai gen calon yang mengandungi LBD16 dan LBD18. Analisis lanjut gen ini perlu dijalankan dalam kerja masa hadapan dengan kerja makmal dengan pelbagai rawatan untuk menguji tahap ekspresinya pada akar tomato.*

***Kata kunci**: Domain Sempadan Organ Sisi, perkembangan akar sisi,* Solanum lycopersicum

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

		Page
LOB	Lateral Organ Boundaries	1
LBD	Lateral Organ Boundaries Domain	1
TFs	Transcription Factors	1
GWAS	Genome-wide analysis	1
SlLBD	*Solanum lycopersicum* Lateral Organ Boundaries Domain	1
GAS	Gly-Ala-Ser	4
CRISPR/Cas9	Clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9	4
ASL20	Asymmetric Leaves2-Like20	4
ARF7	Auxin Response Factor 7	4
IAA4	Indole-3-acetic acid	5
OsLBD	*Oryza sativa* Lateral Organ Boundaries Domain	6
MEME	Multiple EM for Motif Elicitation	
BLAST	Basic Local Alignment Search Tool	8
NCBI	National Center for Biotechnology Information	8
PlantTFDB	Plant Transcription Factor Database	8
TAIR	The *Arabidopsis* Information Resource	8
MSA	Multiple Sequence Alignment	8
MEME	Multiple EM for Motif Elicitation	8
PlantCARE	Plant Cis-Acting Regulatory Element	10

CHAPTER 1

INTRODUCTION

1.1 Study Background

Lateral Organ Boundaries (LOB) Domain (LBD) transcription factors are a plant-specific gene family that regulates the plant growth and development. Further study revealed that LBD genes are important in regulation of the organ development, pollen development, regeneration, disease susceptibility, photomorphogenesis, secondary growth and metabolism. There are genome-wide analyses (GWAS) studies on LBD genes and the diverse functional studies of LBD protein had been done (Zhang et al., 2020) from plant such as *Arabidopsis thaliana*, *Glycine max, and Zea mays* (Gupta & Gupta, 2021). In 2020, *SlLBD40* was functionally characterised to be the negative regulator of drought tolerance in *SlLBD40* knockout mutant line in tomato (Liu et al., 2020). In 2021, *SlLOB1* was found to be the LBD TF genes that regulates the fruit ripening in tomato (Shi et al., 2021). Within the same year, a GWAS was conducted by Gupta and Gupta (2021) and 47 *SlLBD* genes were identified by referencing *Arabidopsis thaliana* model plant sequence. Through their research, candidate genes *SlLBD3, SlLBD20, SlLBD22,* and *SlLBD47* responsible for fruit development and *SlLBD5* responsible for hormone signaling were identified.

Despite the findings above, there is lack of information in defining the role of LBD genes in lateral root formation. Lateral root is important in plant which act as an anchorage system to support them as well as for facilitating the water and nutrient uptake (De Smet et al. 2006). Root development is highly controlled by transcriptional regulators and affected by environmental factors such as water and nutrients availability in soil (Nibau et al. 2008).

Plus, Brady et al. (2007) stated that changes in concentration of auxin hormone will affect the root growth. In most research report, *A. thaliana* often used as the model plant for the root development study (Campos et al. 2016). *LBD16 and LBD18* were reported as the key LBD genes in *A. thaliana* which involved in lateral root formation. In fact, there are no comprehensive studies were found yet to discuss the LBD16 and LBD18 genes in tomato.

Thus, an in-silico analysis of tomato LBD transcription factor was conducted to identify the candidate genes containing LBD16 and LBD18 in *S. lycopersicum* which believed might have functional role in lateral root development in tomato as reported in the model plant *A. thaliana*. The objectives of the study are

1) To identify the group of *SlLBD* genes that contain LBD16 and LBD18 by conducting phylogenetic analysis of LBD family proteins for *A. thaliana* and *S. lycopersicum*.
2) To determine the gene localization patterns and shared ancestry of genes on chromosomes in *A. thaliana* and *S. lycopersicum* by collinearity and synteny analysis.
3) To compare the RNA expression levels of *SlLBD* genes in root using TOMEXPRESS and the visualization of gene expression levels towards abiotic stress using ePlant program.

CHAPTER 2

LITERATURE REVIEW

2.1 Lateral Organ Boundaries Domain (LBD) Transcription Factors of Tomato

Tomato, scientifically known as *Solanum lycopersicum*, is from the large family *Solanaceae* and breeding has been done to improve its fruit productivity, disease resistance and abiotic stresses (Kimura and Sinha, 2014). Tomato is diploid plant, which has a large genome sequence data available. It has a total pair of 12 chromosomes in its genome, comprises size of 900 megabase which encodes for 35000 genes.

Lateral Organ Boundaries Domain Transcription Factors are a plant-specific gene family which initially believed by scientist to be the gene that separates organ by the boundary forming cells (Iwakawa et al., 2002) as well as for lateral organ development such as root development (Hochholdinger, Yu and Marcon, 2018), and leaf development (Moon and Hake, 2011). As the study on LBD genes getting further, there are results proved that LBD proteins not only helps in lateral organ development, but also play important roles in photomorphogenesis (Mangeon et al., 2011), petiole development (Ge et al., 2014), biotic resistance (Hu et al., 2014), hormone response (Yu, Gutjahr, Li and Hochholdinger, 2016), and metabolism regulation (Albinsky et al., 2010). The diverse function of LBD genes can be well understood through the phylogenetic analysis because LBD genes that possess certain function mostly from the same phylogenetic clade (Zhang et al., 2020).

According to Zhang et al. (2020), in the molecular level, the LBD TFs protein are made of conserved N-terminal region and variable C-terminal region. In N-terminal, there are 3 significant structures, the zinc finger-like motif (CX2CX6CX3C) act as a DNA binding

site, leucine-zipper-like coiled-coil motif (LX6LX3LX6L) is the region for protein binding, and GAS block (Gly-Ala-Ser). These structure of N-terminal is classified as Class I LBD proteins. Meanwhile, Class II LBD protein lacks intact leucine-zipper-like domain. The C terminal is responsible for the transcriptional activation or repression of the target gene. These characteristics of LBD protein structure also reported the same by Gupta and Gupta (2021) in their tomato characterisation research.

At present, there are very few tomatoes LBD protein study done by researchers. Liu et al. (2020) through their study found the novel *SlLBD40* genes from LBD family suggesting the genes may negative regulator of drought resistance in tomato. Liu et al. (2020) used the CRISPR/Cas9 technology to generate *SlLBD40* knockout line tomato plant to study gene's function and found that the knockout lines were wilted slightly but the overexpressed *SlLBD40* genes cause severe wilting in tomato plant in drought condition. Shi et al. (2021) and Gupta & Gupta (2021) both showed the role of tomato LBD proteins in fruit ripening. Shi et al. (2021) suggested the *SlLOB1* as the candidate of tomato ripening while Gupta & Gupta (2021) suggested *SlLBD3*, *SlLBD20*, *SlLBD22*, *SlLBD47* and *SlLBD5* as the candidate genes in fruit ripening.

According to Kumar et al. (2021), in *A. thaliana,* LBD16/ASL18 (LBD16 or also known as ASL18), LBD18/ASL20 and LBD29/ASL16 are the genes that specifically expressed in the lateral root formation. The induction of these genes on lateral root formation are controlled by 2 factors, namely ARF7 (Auxin Response Factor) and ARF19. Other LBD genes such as LBD1 and LBD13, with gene Id Solyc06g005090 and Solyc06g007390 respectively, are also expressed in the roots of tomato. Ren, Liu, Gu and Dong (2016) reported that *SlARF2* (*S. lycopersicum* auxin response factor 2) is the regulatory element for lateral root formation as well as flower organ senescence. LBD18 of *A. thaliana* is candidate genes that will response to the salinity and drought stress. This is supported by the evidence

from Ma et al. (2021) stating that the transgenic *A. thaliana* with overexpressing *CaNAC46* transcription factor upregulates the expression of both *LBD18* and *IAA4* (indole-3-acetic acid) when the transgenic plant is treated under high salinity and drought stress.

2.2 Phylogenetic Analysis of LBD Family Protein Members

In the past, there are many functional and evolutionary studies on LBD family proteins and lead to many novel discoveries that is now used as a reference for the studies of new LBD protein function in other species especially tomato. However, due to the large family of LBD members, the function discovered and published on journal seems overwhelming because the genes from different phylogenetic clade have widely diversify function. Due to the confusion about the function of large member of LBD family, Zhang et al. (2020) in their review article stated that referring the functional reports of LBD proteins to the phylogenetic tree is an effective way to analyse respective LBD proteins' function. In addition to that, through this method the LBD proteins of *A. thaliana* they reviewed and found that the genes within the same clades of dendrograms (phylogenetic tree) most likely exhibit similar function in the plant's development.

Recently, this method had been used by Gupta and Gupta (2021) in their in-silico characterization of tomato to categorise the different motifs of tomato LOB domain based on different phylogeny clades. On top of that, dendrogram is used to represent the closely related *SlLBD40* genes within its subfamily which aids in the study of drought resistance in tomato (Liu et al., 2020). Besides, back in few years ago, there are many genomes wide and revolutionary analysis using phylogenetic tree for research on various plants species including *Vitis vinifera* (Cao et al., 2016), *Morus notabilis* (Luo et al., 2016), *Hordeum*

vulgare L. (Guo et al., 2016), *Camellia sinensis* (Wang et al., 2018), *Eucalyptus grandis* (Lu et al., 2018) and *Gossypium* (Yu et al., 2020).

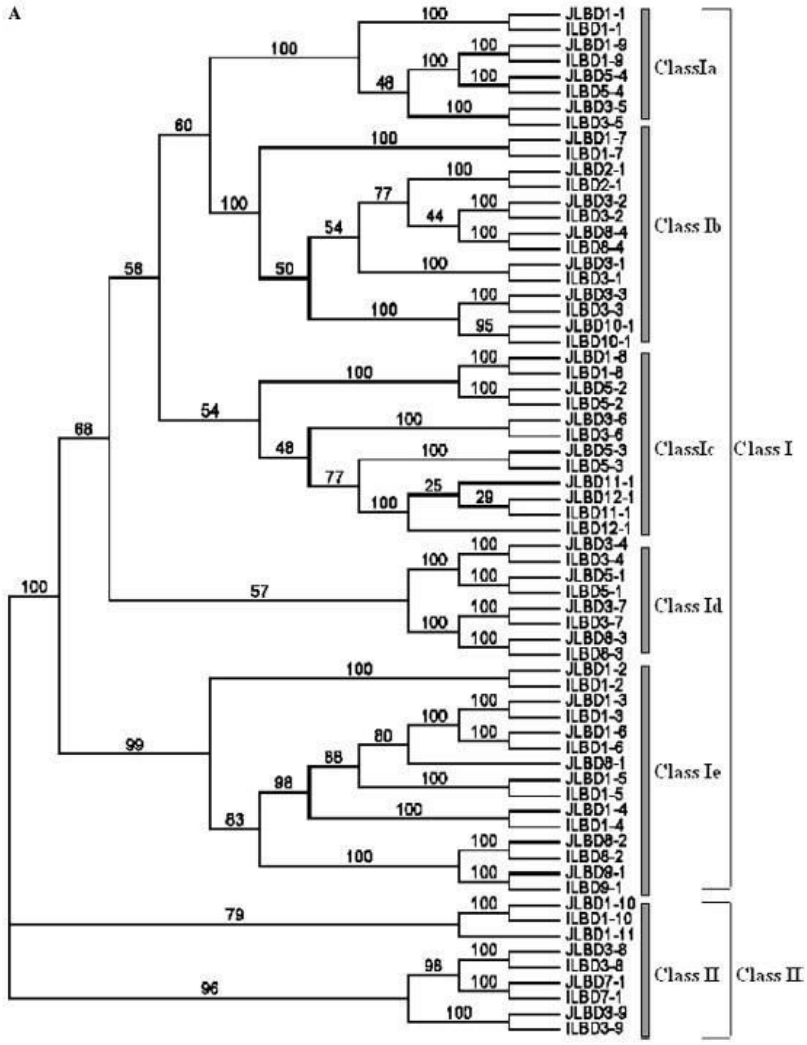

Figure 2.1 Dendrogram above is constructed to analyse the *OsLBD* genes. (Adapted from Yang et al., 2006)

2.3 The Classification and Molecular Structure of LOB Domain

As mentioned in the previous part, the large members of LBD transcription factors generally can be classified into 2 subfamilies, which are Class 1 and Class 2. There are 4 clades under Class 1, namely Class IA, IB, IC and IE, while under Class 2 there are 2 clades, namely Class IIA and Class IIB. This classification is determined by the difference in

leucine-zipper-like coiled coil motif's part that can be found within the conserved region N-terminal of LBD. The leucin-zipper-like coiled coil region is responsible for the protein dimerization (Majer & Hochholdinger, 2011). The conserved motif analysis is very useful in the determination of clades or subclades in the circle phylogenetic tree when LBD family protein from 2 or more species are used. The classification in the phylogenetic tree helps the researcher to identify the homologues of LBD genes from different species, for instance JLBD1-1 and ILBD1-1 are within the same clades with high confidence level of 100% as shown in ***Figure 2.1*** indicates that they are potential homologues with highly conserved motif sequences in their LBD genes.

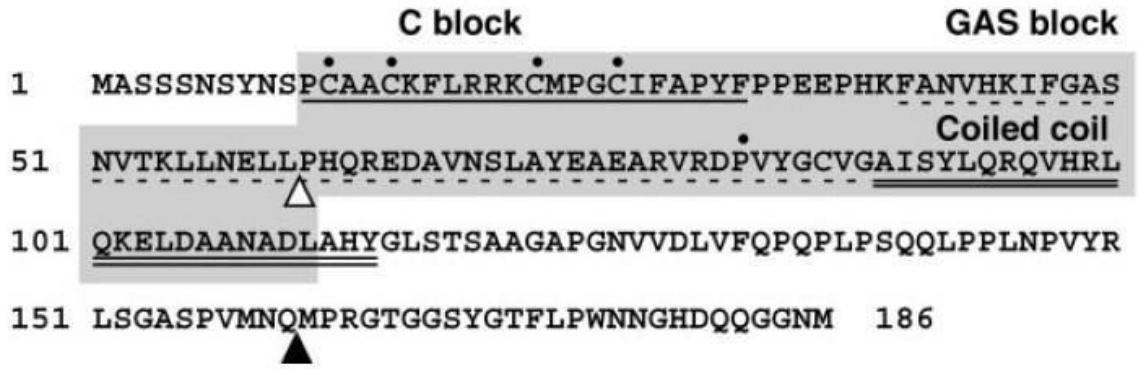

Figure 2.2 LOB domain amino acid sequence which consists of highlighted C block (a.k.a zinc-finger-like motif), GAS block and coiled coil region (a.k.a leucine-zipper-like coiled coil). C block region is annotated with single underline; GAS block is annotated with dash line while the predicted coiled coil is annotated with double line. The Cys residues which makes up the C block and the Pro residue that makes up the GAS block are annotated with black dot respectively. The C block and GAS block are conserved region while the coiled coil is the variable region that determines the class of LBD proteins as well as their functions. (Adapted from Shuai et al., 2002)

2.4 Bioinformatic Tools

According to Bayat (2002), bioinformatic is a computation analysis for biological data interpretation which is crucial in the data management for biology field study. Bioinformatics are widely applied in the functional characterization of proteins for omics

study. Homology searching is one of the important applications in bioinformatics to characterize the function of a newly identified sequence from certain species. Homologue search helps researchers to trace the common ancestry by comparing the two species gene structure through bioinformatic tools.

Basic Local Alignment Search Tool (BLAST) is built in many different databases, such as Plant Transcription Factor Database (PlantTFDB), Phytozome v13 and The *Arabidpsos* Information Resource (TAIR). There are many types of BLAST, which are Blastn, Blastp, Blastx, Blastx, tBlastn and tBlastx (Rani, 2017). BLAST is designed to search for local similarities within the gene sequence region between nucleotide or protein sequences. Statistical computation will be generated as the result which comprises significant data such as e value, bit scores, sequence similarities, and gaps between the query sequences to find the significant match.

Multiple Sequence Alignment (MSA) is an alignment algorithm used in many different bioinformatic tools to align more than 2 sequences which is important in some applications, for instance, the construction of phylogenetic tree for common ancestry analysis, and the conserved motifs finding within the query protein sequence. MSA can be found in MEGA 11 software which was used for the in-silico characterization of tomato LBD protein by phylogenetic analysis.

Plant specific database, such as TAIR for *A. thaliana*, TomExpress for tomato, Phytozome v13 and PlantTFDB databases which store bioinformatic information for wide range of plant species are useful in bioinformatic in-silico analysis. Data such as gene sequence, protein sequence, promoter sequence and protein family search are available .

CHAPTER 3

METHODOLOGY

3.1 The Identification of LBD genes in *Arabidopsis thaliana* and *Solanum lycopersicum*

The *A. thaliana* genes containing LBD16 and LBD18 were identified from TAIR database. PlantTFDB was used to retrieve *A. thaliana* and *S. lycopersicum* LBD protein family sequences.

3.2 Conserved Motif Analysis

Conserved motifs identification was carried out by using Multiple EM for Motif Elicitation (MEME) (Timothy et al., 2015) program (Bailey and Elkan, 1994) for 47 SlLBD genes. Within that program, the minimum width of amino acid was set to 8 and the number of conserved motifs to be searched by the program was set to 10. The conserved motifs prediction was utilized to classify the 47 *SlLBD* genes into Class I and Class II based on its molecular structure. The identified group of genes will be used to mapped on the phylogenetic tree constructed from *A. thaliana* and *S. lycopersicum* LBD protein sequences for identification of *SlLBD* genes that contain *A. thaliana* LBD16 and LBD18.

3.3 Phylogenetic Analysis

Phylogenetic tree was constructed by neighbour-joining (NJ) tree with 1000 bootstraps, p-distance and pairwise deletion parameters settings in Mega 11 software

(https://www.megasoftware.net) for *A. thaliana* and *S. lycopersicum* LBD family protein. Class I and Class II identified from conserved motif analysis were labelled on the phylogenetic tree where Class I were further divided into subclades Ia, Ib, Ic, Id, Ie, If and Ig based on the bootstraps value. The subclades of *A. thaliana* containing LBD16 and LBD18 were identified and the corresponding *SlLBD* genes were identified as well and were further analyse.

3.4 Sequence Similarity (Forward and Reverse BLAST)

The protein sequence of *A. thaliana* LBD16 (AT2G42430) and LBD18 (AT2G45420) retrieved from PlantTFDB were BLAST against *S. lycopersicum* using PlantTFDB BLAST tool. NCBI Blast Tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to reverse BLAST the identified *SlLBD* genes against *A. thaliana* species to ensure the accuracy of identity of *SlLBD* genes.

3.5 Cis-acting Regulatory Elements Analysis

All *SlLBD1* to *SlLBD47* genes protein sequence were BLAST in Phytozome v13 database to find each LBD genomic sequence. The promoter sequence up to 1500 bp were found from each LBD genomic sequence. The retrieved promoter sequence for each LBD gene were submitted to Plant Cis-Acting Regulatory Element (PlantCARE) database (Lescot et al., 2002). The result was obtained via email sent by PlantCARE and the raw data containing all the cis-acting regulatory element in all *SlLBD* genes were filtered. The

significant analysis information after filtering were used to construct the cis-regulatory element graph using TBtools (Chen et al., 2020).

3.6 Collinearity and Synteny Analysis

The gene sequence of *A. thaliana* and *S. lycopersium* in GFF3 and Fasta format were retrieved from Phytozome v13 database. Multiple Collinearity Scan Tool (MCScanX), subset of TBtools software was used to generate collinearity file format by input gene sequence (GFF3 file and Fasta file) of *A. thaliana* and *S. lycopersicum*. The generated collinearity file format, as well as CTL file and GFF file from MCScanX tool were later inserted into Dual Synteny Plot for MCScanX tool in TBtools software to generate the graph. The homologue gene pairs were mapped on the synteny plot.

3.7 Expression Analysis

The selected *S. lycopersicum* LBD gene Id were browsed in the TomExpress database (Zouine et al., 2017) to find the RNA expression data (Zouine et al., 2017). The retrieved data were presented in graph format to analyse the *SlLBD* genes expression pattern in tomato roots. ePlant Tomato program (https://bar.utoronto.ca/) was used to visualise the location where the selected *SlLBD* genes are expressed on the tomato plant in graphic format. The expression level of LBD genes under auxin and abiotic stress treatment were also visualised using Expression Angler tool in ePlant program.

CHAPTER 4

RESULT

4.1 LBD Genes of *A. thaliana* and *S. lycopersicum*

The *A. thaliana* genes that contain LBD16 and LBD18 were retrieved from TAIR database. The respective gene locus and description were presented as in the ***Table 4.1.1*** below.

Locus	Description
AT2G42430	LOB Domain-Containing Protein 16 (LBD16). The gene contains one auxin-responsive element, abbreviated as AuxRE. LBD16 and AuxRE in the promoter regulates the lateral root formation in *A. thaliana*.
AT2G45420	LOB Domain-Containing Protein 18 (LBD18).

Table 4.1 Locus and description for LBD16 and LBD18 of A. thaliana. (Source: www.arabidopsis.org)

4.2 Conserved Motifs Analysis and Classification of LBD Genes

Conserved motif analysis of all 47 *SlLBD* genes was presented in ***Figure 4.1*** to identify Class I and Class II of *SlLBD* protein based on the intact leucine-zipper-like coiled coil region molecular structure. From *Figure 4.1.1*, 6 out of 47 *SlLBD* genes, namely *SlLBD5, SlLBD11, SlLBD15, SlLBD20, SlLBD22* and *SlLBD23*, were identified to have lack of intact leucine-zipper-like coiled coil region (LX6LX3LX6L) in their molecular structure which indicates that they are belong to the Class II LBD gene family. Other *SlLBD genes*

have conserved C-block (zinc-fingerlike-like motif), GAS block and coiled coil region which indicates that they are within the Class I of LBD gene family.

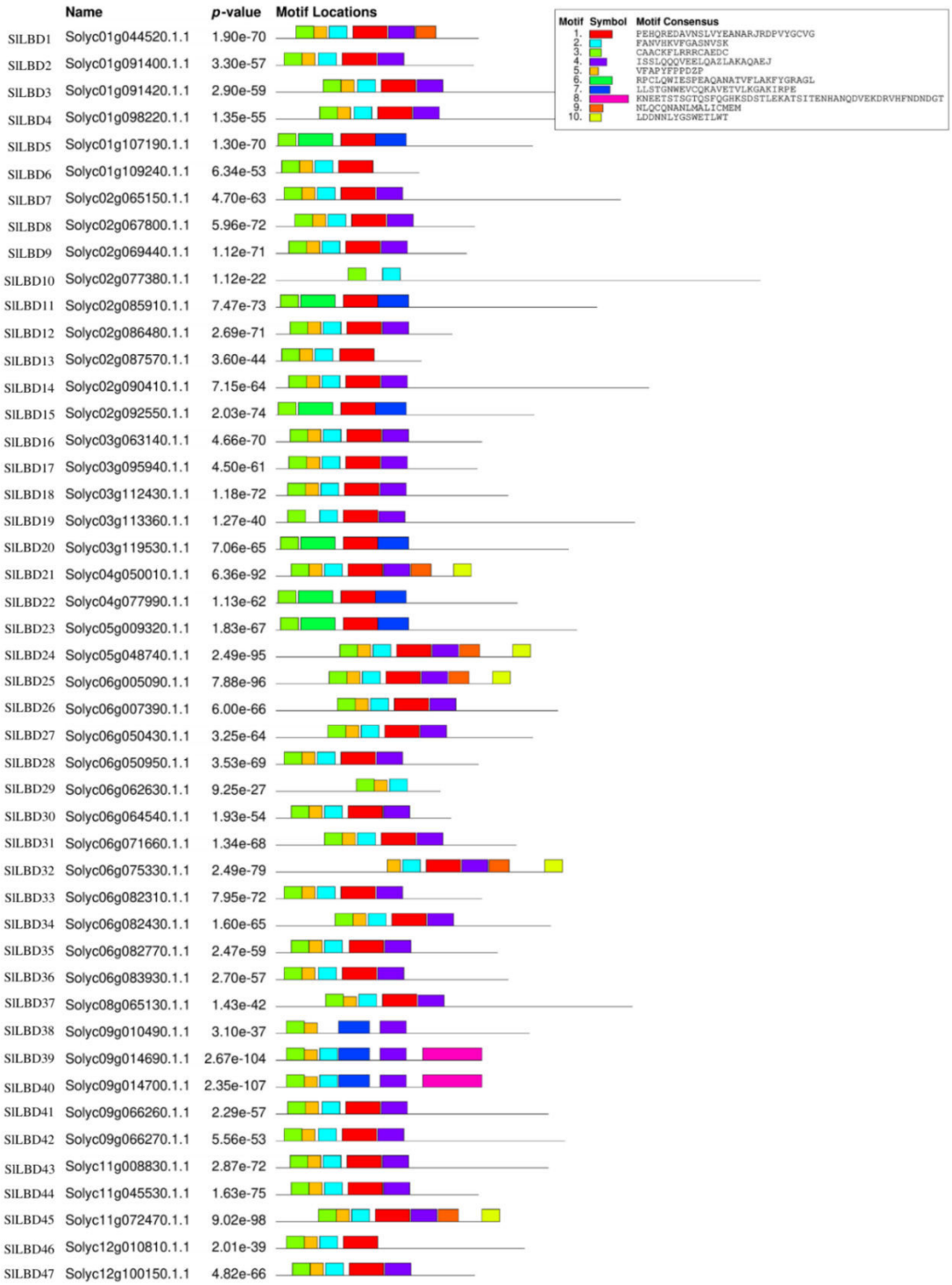

Figure 4.1 Motif analysis of *S. lycopersicum* LBD genes using MEME programme (https://meme-suite.org/meme/) to distinguish Class I and Class II genes based on the present of leucine-zipper-like coiled coil (LX6LX3LX6L) region. Each gene Id were renamed from *SlLBD1* to *SlLBD47* for easy reference.