Speed cloning of fungal resistance genes in tomato for network analysis using transcriptomic approaches

By

Fabia Fakhar Zaman

Department of Plant Genomics and Biotechnology

PARC Institute of Advanced Studies in Agriculture

National Agricultural Research Center, Islamabad

Quaid-I-Azam University Islamabad, Pakistan

2021

Speed cloning of fungal resistance genes in tomato for network analysis using transcriptomic approaches

A Thesis

Submitted to Quaid-i-Azam University, Islamabad in the Partial fulfillment of the requirements for the degree of

MASTER OF PHILOSOPHY

In

Plant Genomics and Biotechnology

By

Fabia Fakhar Zaman

Department of Plant Genomics and Biotechnology

PARC Institute of Advanced Studies in Agriculture

National Agricultural Research Center, Islamabad

Quaid-I-Azam University Islamabad, Pakistan

PLAGIARISM REPORT

It is certified that **Miss Fabia Fakhar Zaman (**02361913006) has submitted her M. Phil thesis having Title "**Speed cloning of fungal resistance genes in tomato for network analysis using transcriptomic approaches"** that has been checked on Turnitin for similarity index (plagiarism).

Overall plagiarism =18% that lies in the limit provided by HEC (19%)

_____

Dr. Muhammad Ramzan khan
Principal Scientific Officer,
NIGAB, NARC, Islamabad

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Word	Abbreviation
EB	Early Blight
LB	Late Blight
BB	Blister Blight
RNA	Ribonucleic Acid
R genes	Resistance genes
NLR	Nucleotide-Binding Domain and Leucine-Rich Repeat
NBS	Nucleotide Binding sites
TIR	Toll-like/interleukin
CC	Coiled coil
EDS1	Enhanced disease susceptibility 1
NDR1	Non-race-specific disease resistance 1
MAS	Marker-Assisted Selection
GS	Genomic Selection
GWAS	Genome-Wide Association Studies
PDA	Potato dextrose agar
Ml	Microliter
Ml	Milliliter
cDNA	Complementary deoxyribonucleic acid
NARC	National Agriculture Research Center
NCBI	National Center for Biotechnology information
PCR	Polymerase China Reaction
qRT-PCR	Quantitative reverse transcriptase PCR
dNTPs	Deoxyribonucleotide triphosphate
GEO	Gene Expression Omnibus GEO
SRA	Sequence Read Archive

KEGG	Kyoto Encyclopedia of Genes and Genomes
GO	Gene Ontology
DNA	Deoxyribonucleic Acid
DMSO	Dimethyl sulfoxide
Mm	Millimeter

ACKNOWLEDGEMENTS

I am grateful to **ALLAH Almighty** the most merciful and beneficent, who enable me to complete this dissertation. Indeed, without his help and will, nothing is accomplished. The whole facilities, abilities, opportunities, strength and powers are blessed by **ALLAH** almighty for the level of study and research in such a vibrant environment. Countless Darood on **HOLY PROPHET HAZRAT MUHAMMAD (P.B.U.H)** w**ho** is the torch bearer of knowledge for humanity and gave the lessons of seeking knowledge from cradle to the grave.

I would like to express my sincere gratitude and special regard to my honorable supervisor, **Muhammad Ramzan Khan,** Principle Scientific Officer, National institute of Genomics and Advance Biotechnology, NARC, Islamabad for his patience, motivation, continuous support and efforts for completion of my thesis research work. I would like to say special thanks to **Dr. Nazia Rehman** Senior Scientific Officer, NIGAB, NARC, Islamabad. I am ever so grateful for her comments, advice, time and support. I would like to acknowledge and thankful to my teacher and Director National institute of Genomics and Advance Biotechnology, NARC, Islamabad **Dr. Shaukat Ali**. I convey heartiest thanks to **Mam Safeena (SO, NIGAB)** and **Dr. Muhammad Kashif Naeem (SO, NIGAB)** for their support and help**.**

I would like to thank to my colleagues of session 2019-2021, for joys, happiness, sadness, fun and all academic activities during the whole academic session. I would like to thank all my teachers and Lab fellows especially **Alia Errum and Nageen Zahra, Fatima Rasool.** All were very kind and cooperative in the research matters. I would like to pay special thanks to my colleague **Maryam Mubeen, Hina Abbas, Sehrish Bashir, Shelah Shoukat and Marya Rubab,** they all encouraged and support me allot in all the matter either related to education, research and social attraction. At the end, I am extremely grateful to my **Parents, Sisters** and **Brother** for providing me the moral and emotional support. Without my family support and love, I would not have finished this research work.

Fabia Fakhar Zaman

Abstract

Early Blight and Late Blight caused by *Alternaria solani* and oomycete *Phytophthora infestans* are major factors of destruction of tomato plants resulting 49 to 91 % yield loss. Transcriptome analysis of Early Blight and Late Blight on tomato would help to identify the NLR genes that conferred resistance mechanism. In this study some R genes are unveiled that contain NB: LRR domains and involved in regulation of genes to confer resistance mechanism in tomato plants. NLR genes were identified from the downloaded raw RNA-seq data analysis. NRC4, R1B16, RPP13, RPM1 gave High FPKM values. These genes expression was checked on three cultivated tomato accessions of *Solanum Lycopersicum* (38046, 19890, Roma) and one wild specie accession of *Solanum Chilense* (19906). NRC4, R1B16 showed upregulation and R1B12 showed down regulation in infected plants of LB and RPP13, RPM1 showed upregulation in infected plants of EB. Protein -Protein interaction represented no significant interaction between these NLR genes as many genes were not annotated. Genes Co-occurrence analysis predicted conservation of these NLR genes is higher in Solanaceae family. RPP13 showed broad spectrum resistance against diseases. In this study RPP13 gene was cloned from *Solanum lycopersicum* accession 19890. The cloned RPP13 gene showed 70% homology with existing genes sequences. Further research is required to unveil more NLR genes that are involved in resistance mechanism against fungus diseases.

Introduction

1.1 *Solanum Lycopersicum, an* Edible Berry

Tomato (*Solanum lycopersicum*) is a diploid plant species (2n = 24) that is the native to South America. This plant is the most important crop species in the world. The cultivated tomato (*Solanum lycopersicum L.)* is considered as the world's most highly consumed vegetable. Tomato is a pivotal component in a majority of food items e.g., cooked and processed food. *Solanum lycopersicum* allied to the Solanaceae family and grown for local use or for the exportation worldwide. Cultivated tomatoes originate from Peru, Ecuador and Galapagos Islands that are related to wild tomatoes. Mexico is the Centre of tomato domestication and diversification. Tropical and temperate climates are best for tomato cultivation in an open field or under greenhouse.

Climate Tomatoes need a warm climate for growth and cannot be able to condone frost. The life cycle in tomato cultivation compasses one spring and summer with its optimum temperature of day and night 26 ºC and 12 ºC respectively. For vegetative growth, Tomato plant may requisite temperature above 18ºC, however it can survive at lower temperatures (12ºC). Flower fertilization rate, plant development and fruit ripening rate may be affected by temperature above 31 ºC. 55-60% air humidity is important for production of pollen and pollination (Victoria *et al*., 2011).

The commercial valued tomato fruit can vary in color, size and shape (Vaughan *et al*., 1997). The tomato fruit may contain a large water contents, low proteins and fats carbohydrates contents and vitamins and minerals, carotenes, such as lycopene (which is responsible for red color) and beta-Carotene (orange color to fruit). The modern cultivars, fruits may quell up to 3% sugar of fresh fruit weight and tomatine, an alkaloid involved in fungicidal properties. Maturity of fruits leads to decreases the Tomatine concentration, helps to determine the taxonomy of the species and is useful in tomato crop breeding (OECD 2008; Spooner *et al*., 1993).

In Pakistan tomato is an important food consumed on a daily basis. It is considered as a complementary ingredient to make many products like ketchup and different sauces

and also used as a fresh vegetable. It may contain vitamin C, vitamin A, calcium, and iron. Tomatoes contain an antioxidant Lycopene that avoid growth of various cancerous cells (Adenuga *et al*., 2013).

From 2009 to 2015 the total global production of tomato ranged from 154 million tons to 174 million tons. The major tomato-producing countries are China and India (FAOSTAT 2017). There are a variety of geographical zones for tomato cultivation for example in open fields or greenhouses, and the fruit can be harvested by manual or mechanical means. This crop plant can be considered as perennial or semi-perennial, but commercially it is an annual crop due to some conditions like revitalization pruning, irrigation and frost protection (Geisenberg *et al*., 1986).

In Pakistan tomatoes are produced by small farms. It generates better opportunities for rural laborers for employment and raises better profits to producers. Compared to international level the crop yield of Pakistan is low. Pakistan is the 37th largest tomato producer in the world with a production rate of 0.566 million tons. According to FAO, among the top ten countries China is a leading producer with 31.47 % global tomato production and Asia has production rate 53% (FAO 2015). In Pakistan from 1995 to 2016 area under tomato cultivation has been doubled from 29.4 thousand ha to 61.9 million ha with an annual growth rate of 4.9% and production rate has also increased from 294.1 thousand tons to 587.1 thousand tons with an average annual growth of 4.4%.

In 2017 tomato production rate has declined as 569 thousand tones (MNFS&R, 2017). During 2016-2017 the tomato yield was 95,106 hg/ha in Pakistan. (FAOSTAT 2017). Semi-arid region and decrease in arable land due to urbanization has aggravated the situation of food security. Lack of improved disease resistance seeds, pesticides, organic manure, mineral fertilizers, water, and land are major constituents of lower production of tomatoes. A number of biotic and abiotic factors affect the tomato crop's final yield. Pests and diseases on tomato crop negatively affect the production of assimilates and filling of the fruits in soilless conditions that may decreasing plant photosynthetic capacity (Stacey *et al*., 1983; Bertin *et al*., 1993)

Different pathogens causing diseases in tomato plants constitute a major cause of yield reduction. Tomato plants are attacked by Viruses, bacteria, fungi and nematodes (agrios 2005). Fungal pathogens cause major destruction of tomato plants, as pathogens can attack on all stages of growth (Stone *et al*., 2000). The tomato plant diseases, blights caused by *Phytophthora infestans* and *Alternaria solani,* leaf spots by *Colletotrichum* spp., stem canker, fruit rot and wilt by *Fusarium oxysporum* f.sp. *lycopersici* are serious disease problems all over the world (Adebayo, PhD thesis 2005).

In Pakistan many factors are responsible for low production of tomato but early blight (caused by *Alternaria solani)* and late blight (caused by *Phytophthora infestans*) are major factors of destruction of tomato plants. In Pakistan, according to an estimate, 49 to 91 % yield loss is due to early blight and late blight (Azam and Shah, 2003; Akhtar *et al*., 2012 and 2016).

1.2 Early Blight

Early blight (EB), is a major devastating cause of tomato (*Solanum lycopersicum*) production worldwide. Five different *Alternaria* species, namely *Alternaria solani*, *Alternaria linariae* (syn. *Alternaria tomatophila*), *Alternaria grandis Alternaria alternata* and *Alternaria tenuissima* that may cause Early blight disease in tomatoes. *Alternaria solani* reproduces asexually, its sexual stage is unknown. warm and humid conditions. Temperature 8-32 °C humid and warm conditions are suitable for conidia germination for infection (Chaerani *et al*., 2006).

1.2.1 Symptoms of early blight

EB is reported as a three-phased disease that may infect different aerial parts of tomato to cause collar rot, leaf blight, stem end rot at any growth stage. Plant physiological maturity is associated with Early blight. However older, senectitude leaves are susceptible for disease than the younger ones. Disease appears at the lower older side first then lesions move to the upper side. Dark-colored bullseye spots are developed during the leaf blight phase which is shown as a small, dark, coalescing and concentric ring pattern necrotic in the center. The tomato leaves and fruits turn blighted and fall

when the lesions enlarge and become more plentiful resulting in gradual defoliation and sun scalding of tomato fruits that lead to depletion of quality and yield. Calyx and fruit tissues susceptibility to EB result in fruit rot, which contributes up to 79% yield losses in tomatoes (Akhtar *et al*., 2019). EB infection on tomato may produce black fruit, sunken, leathery and purple lesions on the stem-end. The fruit lesions are expanded to a crucial size and expended into the flesh of the fruit deeply. Infected fruits start dropping prematurely (Chaerani *et al*., 2006).

1.2.2 EB control measures

Three measures are used to control the EB infection to plants, cultural practices, fungicide treatment and resistant varieties. Many cultural practices are used to maintain the healthy field and crop e.g., disinfection, removal of infected vines and fruits, plant debris and volunteer weeds from the field region, and soil-directed irrigation system to lower the leaf wetness, but these are not sufficient for EB reduction. Several types of fungicides are applied at the interval of 7,10 days to control the EB infection (Adhikari *et al*. 2017). Often fungicides do not work under high disease pressure, new fungicide-resistant isolates can be produced in the result of frequent fungicides use under high selection pressure (Chaerani *et al*., 2006). Resistant varieties, and suitable control measures are used to overcome reduction caused by EB (Adhikari *et al*. 2017).

1.3 Late blight

The oomycete *Phytophthora infestans* cause Late blight (LB), occurs in the growing regions of tomato and potato in North America and worldwide depending on weather conditions. Like EB, Late Blight has been described as a major disease of tomato (Fry *et al*., 1997). The oomycete *Phytophthora infestan*s causing late blight is also described as the main cause of tomato yield losses in the greenhouses and the fields (Nowakowska *et al*., 2014). In the absence of plant protection *P. infestans* has the ability to destroy the entire tomato crop within 7 to 10 days. LB affects the economy by reducing yield, lowering fruit quality, reducing storability, and costs of fungicide application (Foolad *et al*., 2014). Commercial cultivars of tomato grown in Pakistan are prone to attack by LB that could be controlled by cultural practices and fungicides. However earlier

breeding efforts have met partial success to control the disease by producing resistant cultivars. (Akhtar *et al*., 2012)

Asexual and sexual life cycles originate *Phytophthora infestans* that act as pathogen. Humid environment, cool condition, rainy or foggy conditions are best for LB infection on plants and involved in large proportion of total monetary losses sustained by growers each growing season. High disease rate may cause extensive defoliation leading to a reduction of economic fruit yield. The optimal temperature for disease occurrence is 12 °C to 21°C on host tissue in a process taking 8-48 hours (Foolad *et al*., 2008).

1.3.1 Disease symptoms

All above-ground parts of the plant e.g., leaf and stem acquire necrosis, fruit rot and eventual plant death; it can infect tomato seed (Rubin *et al*., 2001; Rubin and Cohen, 2004). Late blight disease symptoms are shown as dark brown or black water-soaked lesions bordered with pale yellowish-green appear on healthy tissue and purple margins at leaflet margins. With the severity of disease, lesions spread more on the leaves. Moist weather is useful for the white fuzzy sporangia production on the lower part of leaves. Extensive defoliation occurs with LB progress and eventually the plant begins to die. On stem top or at a node dark brown lesion are appeared that may move down the stem. Brown, greasy lesions on the stem end and sides of green fruit, make them unmarketable (Foolad *et al*., 2008).

1.3.2 LB Control measures

According to an estimate $5 billion used for tomato and potato production annually worldwide, which also includes the cost of disease control and crop losses (Judelson *et al*., 2005). Normally cultural practices and chemical applications are used to control late blight. In 1970s, a combination of cultural practices and applications of fungicide were used to effectively control the disease.

1.3.2.1 Cultural Practices

A prime element of grower's strategy is cultural practices in LB disease management. The purpose of cultural control of LB are to reduce the infection rate, disease development conditions and disease spread. Crop revolving and elimination of volunteer tomato and sources of LB are specific cultural practices

1.3.2.2 Chemical Applications

Protectant and systemic fungicides are both included in chemical control of LB. Before the disease emergence Protectant fungicides are used and before or upon disease development systemic fungicides are given to plants (Gisi *et al*,.1996). For example, Metalaxyl fungicides are systemic fungicides that are widely applied to control LB. In fungi Metalaxyl fungicides have ability to impede ribosomal RNA (rRNA) polymerases. However, in 1980 unfortunately it showed resistance against Metalaxyl (Gisi *et al*,.1996). However, use of fungicides are expensive measure and have dangerous effect on environment and humans, and have needed proper time for its application on plants.

1.3.2.3 Biological control measure

Biologically compounds extracted from plants may be used as fungicides to control phytopathogenic fungi and bacteria. The extracted plants may contain a rich source of bioactive chemicals for example phenols, flavonoids, quinons, tannins, alkaloids, saponins and sterols (Isman 2000 and Burt 2004). These extracts may act as biodegradable to non-toxic products of fungal and bacterial pathogens, they may act as a new class of safer disease control agents.

1.3.3 Genetic Measures of Disease Control Resistance to fungal disease

Race-specific and race nonspecific resistance are two categories to genetically control *P. Infestans.* These races may be defined as disease reciprocity of pathogen and different host genotypes. A single gene can control vertical resistance, Race Specific resistance, single-gene resistance or qualitative resistance. Flax and flax rust resistance

interaction was first described by Flor (1955), which shows dependence on factors present in the host and the pathogen (Flor *et al*., 1955). Between host plants and disease organisms such interlinkage can be defined as "gene-for-gene interaction". The product of is called the R gene product or Avr gene product. Horizontal resistance has been reported for many tomato diseases such as early blight, powdery mildew caused by fungus, bacterial canker and wilt in. For late blight and powdery mildew vertical and horizontal resistances mechanism has been reported (Foolad 2007).

1.4 Resistance Genes (R genes)

Resistance genes (R genes) are the important gene family to resist and recognize the pathogen. R genes recognize avirulent (Avr) pathogen proteins and play a role in initiating the defense mechanisms culminating in a hypersensitive response (HR). Several *R* genes have been cloned from tomato, potato and pepper which contains NB-LRR domains, and are being used in current breeding technologies. The tomato genome assembly has identified the large size of the NB-LRR gene family, and potential *R* gene repository (Andolfo *et al*.,2013). The first R gene annotation was reported by the Tomato Genome Consortium by existing gene annotation and protein predictions (The Tomato Genome Consortium, 2012).

Cultivated varieties are more vulnerable to biotic factors and generally lack disease resistance due to limited genetic diversity. immune receptor proteins consist of intracellular nucleotide binding/leucine-rich repeat (NLR) are encoded by Most resistance (R) genes. R gene diversity in crops is being reduce by Domestication and intensive breeding, which make them more vulnerable to disease eruption. R gene integration into crops is arduous and corresponding with co-integration of undesirable genes. R genes clones could be engineered as a clump into crops to avert the linkage-drag and emergence of virulent pathogens. the existing R gene cloning methods needed segregation or mutant progenies but due to poor agronomic traits it is difficult to produce for many wild relatives. Positional cloning or mutational genomics are two approaches to clone R genes and they requisite screening of thousands of recombinant or mutant lines and existing single R gene of interest having a susceptible genetic background. However wild crop tomato plants frequently possess multiple R genes

which have pre-domestication traits that prohibit the use of R genes cloning methods (Arora *et al*., 2019).

1.5 Nucleotide-Binding Domain and Leucine-Rich Repeat (NLR)

Plants rely on an innate immune system as they have no specialized cells to confer adaptive immunity. For the activation of immune responses to invade the pathogens, plants rely on nucleotide binding domains and leucine rich repeats (NLR). NLR are a rapidly evolving and diverse protein family in plant species (Jacob *et al*., 2013). NLR are intracellular multi domain receptors that contain highly conserved NBS (nucleotide Binding sites) domain having similarity with the AAA-ATPase family and variable LRR (leucine rich repeats) domain (Cannon *et al*., 2002) . NLR are modular proteins that are categorized into two groups based on N-terminal domain, which is either N-terminus of the Toll-like/interleukin 1 domain (TIR-type NLR) or a coiled coil (CC) domain (non TIR-type NLR. Toll-like/interleukin 1 domain require EDS1 (enhanced disease susceptibility 1) that is involved in the activation of downstream signaling pathway and coiled coil (CC) domain require NDR1 (non-race-specific disease resistance 1) (McHale *et al*., 2006).

The Solanaceae i**s** considered as the most species-rich plant family in which major agricultural crops lie such as potato, tomato, and pepper (Sarkinen *et al*., 2013). Many NLR-type disease resistance genes from wild relatives are identified to improve the extensive breeding efforts for resistance within the Solanaceae family. From different Solanaceae species 20 NLR-type disease resistance genes have been identified, which confer resistance to many pathogens and pests, including the oomycete *Phytophthora infestans,* tomato spotted wilt virus (TSWV), and potato cyst and root-knot nematodes (Van Ooijen *et* al., 2007); Vleeshouwers *et al*., 2011). Traditional breeding process is being used to place NLR-type resistance genes in Solanaceae through agriculture (Cisgenesis, or transgenesis) (Vleeshouwers *et al*. (2011); Jo KR, *et al*. (2014).

1.6 Advanced Biotechnology tools

Due to the food demand of a growing population (25%-70%) more agricultural production is required by 2050. Modern agriculture may help to fulfill the estimated food production value from 7.7 billion to 9.8 billion by 2050 (Hunter *et al*., 2017). crop losses by different environmental factors may make it more challenging to attain the target. Crop Improvement is very important for increasing crop production to meet global food demands. Plant breeders and scientists are trying to develop improved crops for higher yield, with better nutrient profile, resistant to pest and pathogens and climate-resilient. Conventional breeding is a prolonged and laborious process in which large populations of crops are grown and examined over several generations, with chances of co-integration of lethal genes (linkage drag). multiple times of backcrossing and selection to re-establish the elite background limit the co-integration of lethal genes (Lidder *et al*., 2012)

Marker-Assisted Selection (MAS) and Genomic Selection (GS) are techniques that assist the breeding program by facilitating the selection efficiency (Collard *et al*,.2007; Desta *et al*., 2014). Random recombination events and undirected mutagenesis, extra improvement of existing elite varieties is an extensive and exhausting practice and domestication of elite varieties are the factors that make traditional breeding techniques limited (Shi *et al*., 2017). Genetic engineering is being used to improve the crop with direct modification of genetic makeup of any organism via biotechnology (Christou *et al*., 2013). The genetically modified Crop (GM) has an important role in sustainable agriculture and food security but despite its potential applications, their usage raised bio-safety concerns.

Genome-Wide Association Studies (GWAS) help to identify the correlated traits in a genetically varied population by using recombination events which are already present in natural populations. The limiting factor of GWAS is the dependence on a reference genome. It limits the sequence retrieval that have considerably deviated from the reference, for example, R genes. However due to this limitation, sub-sequences (k-mers) based trait associations have been used (Rahman *et al*., 2018).

1.7 Association Genetics and Sequence Enrichment'' (AgRenSeq) or speed cloning

To avoid conventional breeding and genetic engineering limitations, the scientists are trying to develop new technology which assists to reduce the generation time and speed up crop development. With the aim of crop improvement, “Association Genetics and Sequence Enrichment'' (AgRenSeq) or speed cloning, genome editing and speed breeding are techniques innovated by scientists to attain their goal.

An international union of researchers has developed a new revolutionized technique to speedily identify disease-resistance genes from wild plants for transfer into domestic crops and fulfill the global food demand. Scientists working at John Innes Centre in the UK in collaboration with the scientists in Australia and US have developed Speed cloning or AgRenSeq “Association Genetics with R gene enrichment Sequencing” technique. According to Dr Brande Wollf “AgRenSeq technique is quite useful to increase yields and reduce pesticide applications”. Speed cloning accelerates the struggle against pathogens which threaten major food crops globally such as wheat, maize, rice, soybean and potato.

According to Habans Bariana, this technique reinforced the rapid identification and characterization of new disease-resistant genes. Scientists have cloned R genes by selecting a diversified panel of wild diploid wheat accessions to identify their resistance against different pathogens. According to Brian Steffenson, “Other than wheat, AgRenSeq may be an important step forward to assist and speed up the development of more resilient crop plant varieties”. To minimize the crop losses, we can breed new varieties with broad-based, multi-gene resistance.

Breeders will be able to use the AgRenSeq technique to breed crops having resistance against pests and diseases thus bringing down the production cost and the price of food products may benefit the consumer. Speed cloning is a great innovation of breeders and scientists. However, this technique can not only be used to speedily clone the R genes but also can be used to clone genes for other traits improvement including yield, nutrition profile and climate tolerant crops to make them accessible to the world’s grower (Nazia Rehman *et al*. 2020)

1.8 Objectives

1- Identification of NLR-derived R genes for early and late blight resistance in tomato using transcriptome approach

2- Validation of key differentially expressed NLR genes using quantitative real time PCR

3- Network analysis of NLR genes

4- Cloning of NLR-derived R genes for early blight and late blight resistance in tomato

Materials and Methods

2.1 Plant materials

Seeds of tomato accessions and progenitors acquired from Bio-resource Conservation Institute (BCI) NARC, Islamabad Pakistan. These tomato seeds included tomato varieties (*Solanum lycopersicum*) and wild species of *S. chilense*, *S. pimpinilifolium* and *S. pennellii* from Pakistan. These seeds were sown in trays, after 3-4 weeks plants were shifted to soil for proper growth.

2.2 Inoculum preparation

For the inoculum preparation PDA media (39g of PDA powder mixed in 1L of water) was prepared. Infected leaves of early blight and late blight were cut from the plant and washed with 70% ethanol. Cut the infected part of leaves and placed it on PDA media and let it for growth on room temperature 25ºC. After one-week circular rings of fungus appeared on media. Fungus culture colonies were identified using microscope. Black circular rings of *Alternaria solani* were formed on media and white circular rings of *Phytophthora infestans* were formed on media.

2.3 Infection conditions and symptoms screening

Leaves of two-month-old plants in field were inoculated with spores of fungus inoculum on dorsal surface of leaves and kept them growing for 3-5 days for disease symptoms appearance. After 5 days a pale-yellow base with black circular rings on leaves appeared. On the visual appearance of disease symptoms on inoculated plants, from 50 tomato plant accessions four plant accessions of *Solanum lycopersicum* (38046, 38037, Roma, 38039) and *S. chilense* (19906) were used for RNA extraction and expression. Healthy and diseased leaves tissue samples were plugged from plants and stored at -80 ºC for RNA extraction.

2.4 Extraction of RNA

Total RNA from the leaf tissues of five healthy and diseased tomatoes plant accessions are extracted by using RNA kit (Invitrogen). By using liquid nitrogen, the leaf tissues were grounded mechanically with autoclaved mortar and pestle. Samples were taken in the Eppendrof tube containing 1ml lysis buffer and 10µl mercaptoethanol. The homogenization took place and then centrifuged the homogenized mixture for 3 minutes at 14,000 ×g. After centrifugation transferred the supernatant into clean tube, an equal volume of 70% ethanol was added into it. Vortex the mixture thoroughly and then 700 ml of mixture was transferred into the spin cartridge (with the collection tube). Centrifuged the mixture at 12,000 ×g for 30 seconds at room temperature, discarded the flow-through and repeated this step until the entire sample had been processed. 700µl wash buffer I was added to spin cartridge and centrifuged at 12,000 ×g for 30 seconds, discarded the flow through with collection tube. 500µl wash buffer II was added in spin-to-spin cartridge and centrifuged at 12,000 ×g for 30 seconds, discarded the flow through and repeated the step. The spin cartridge was centrifuged 12,000 ×g for 1-2 minutes to dry the membrane with bound RNA. Discarded the flow through and inserted the spin cartridge into recovery tube. 50µl RNase-free water was added to the center of spin cartridge, incubated at room temperature for 1 minute and then centrifuged at 12,000 ×g for 2 minutes. Eluted RNA was collected in tube and stored at -80℃.

2.5 Gel Electrophoresis

Gel electrophoresis was used to observe RNAs of different tomato accessions. Total RNA was subjected to 1% agarose gel prepared in a flask using 70ml of TAE buffer and 0.7 grams of agarose. The mixture was then heated in a microwave oven for a few seconds until it was dissolved and cooled down completely. 4µl ethidium bromide was added on the gel for RNA visualization. The mixture then poured into gel tray solidify after several minutes. After this 4µl of RNA samples were loaded into the wells with 3µl (Thermo Scientific 6 X DNA Loading dye), 2.5ul of 1000bp ladder (Thermo Scientific Gene Ruler 1000bp DNA ladder) was loaded as a standard in parallel. The tray containing 1x TAE buffer was set on 200 volts and

300 A for 10 minutes. The presences of RNA bands were confirmed after 10 minutes under iBright (Invitrogen imaging system).

2.6 Quantification of RNA

BioSpec-nano Spectrophotometer for LIFE SCIENCE instrument was used to analyze the quality, quantity and integrity of RNA. RNA quantification wavelength was 260/280nm. Three readings of RNA were found in distinct samples of plants at 280/260 nm wavelength of a spectrophotometer.

2.7 Synthesis of cDNA

The synthesis of cDNA was accomplished by using Thermo Scientific Revert Aid reverse transcriptase. To make 20µl sample variable quantity of RNA and water was mixed in 1µl oligo-dT and vortex gentle to mix reagents carefully. For 5 minutes the mixture was incubated at 65 ºC then transfer to ice for 1 minute. To make 4µl reaction, 5x reaction RT buffer, 2 µl dNTPs and 1µl reverse transcriptase enzyme were added and fused, incubated at 42ºC for 60 minutes. Hence, the reaction was abort at 72 ºC for 10 minutes.

2.8 Transcriptome analysis

Raw RNA seq data for early blight was downloaded from GEO database under the accession number **GSE75923 (**Sarkar, D *et al*., 2017)**.** For late blight raw RNA seq data was downloaded from GenBank Sequence Read Archive (SRA) under accession number **SRP041501 (**Zuluaga, A. P *et al*., 2016) and the reference genome was downloaded from Sol genomics database (http://solgenomics .net). Reference genome was indexed using Hisat2-build tool. Raw reads were mapped with reference genome using **HiSat2 tool (version 2.1.0)**, which generated sam files. Sam files were converted into Bam files using **Samtools (version 1.10). GFF** files contain information about exons location on genome which defines the overall transcript structure, with optional *CDS* features to specify the coding segments. This information is required for quantification of reads. GFF file of reference genome was downloaded from Sol genomics (http://solgenomics .net). and converted into GTF file using **gffread** tool as

required for next step. **Stringtie (version V2.1.5)** tool was used to quantify the Bam files and GTF files. Resulting GTF files, one for each sample was given as input into Ballgown package in R, which generated expression profile of whole transcriptome. The expression values were in FPKM. NLR genes with high expression values for early blight and late blight were retrieved from the expressed data analysis.

2.9 KEGG and GO analysis

Gene ontology and KEGG enrichment analysis was performed to identify the significantly enriched biological, molecular processes and molecular pathways in tomato by using online tool **g: Profiler** version e104_eg51_p15_3922dba (https://biit.cs.ut.ee/gprofiler/gost).

2.10 Primer designing

Primers of both early blight and late blight were designed for the expression of genes RPP13, RPM1, NRC4, R1B12, and R1B16 by NCBI primer blast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) with setting the primer parameters as PCR product size 70 to 130 and Tm ranging 57ºC to 60ºC and default setting of Primer pair specificity checking parameters.

Table 2.1: list of primers for expression analysis

Primers	**Sequences (5' to 3')**
RPM1-F	CGATGCTGCAAATGATCGCA
RPM1-R	CTCCACGAGCCAACCGATTA
RPP13-F	TCGTCTCAAGGCTTGTGCTT

RPP13-R	GCGAGCGATGTCCATGATTT
NRC4-F	CTTCAGGCAACCCCTGTTCT
NRC4-R	AACCCGACCACAGCATCATT
R1B16-F	TGGAGTACTGAGGCATGGGA
R1B16-R	GGGCAGTTGACATAATCCGC
R1B12-F	TGATGCTCTACCAACACGGG
R1B12-R	GAAGGAGGCCCTATAAGCCG

2.11 qRT-PCR expression analysis

The expression of NLR genes of *Solanum lycopersicum* in leaf tissues were checked using comparative ΔCT method in real-time PCR (Applied Biosystems) with StepOnePlus software. Gene specific primers of early blight (RPP13, RPM1) and late blight (NARC4, R1B16, and R1B12) were used. The kit used for RT-PCR was Thermo Scientific Maxima SYBR GREEN/ROX qPCR. Elongation factor (EF) was used as endogenous control and the target genes were (RPP13, RPM1, NARC4, R1B16, R1B12. The total reaction volume was kept 10µl in which 5ul SYBER Green was used. The profile of RT-PCR was set as initial 94ºC temperature for denaturation stage for 10 minutes, followed by 40 cycles at 95ºC in second stage for 30 seconds, then 60ºC for 30 seconds 72ºC for 30 seconds, the third stage is set at 95ºC for 15 seconds and 60ºC for 30 seconds. Finally, melt curve study was complete at 52ºC to 95 ºC. Standard curve Amplification plot, ΔCT values data and melt curve plot were recorded. These all data

were more analyzed with Step One Plus software system. T. Test was applied to check the significance of results.

2.12 Isolation and purification of fungal resistant gene

Primers for the gene isolation were designed manually by using cds sequences of fungal resistant gene RPP13 (early blight resistant gene). The isolation of genes was accomplished by using Q5 High-Fidelity DNA Polymerase. The 25µl, 1X amplification reaction was executed by addition of following reagents 5µl of Q5 Reaction buffer, 0.5µl of dNTPs, 0.25µl of Q5 DNA Polymerase, 1.25µl of forward and reverse primers, 2.5µl of cDNA sample and 13.75µl water. The PCR tubes were transferred in thermal cycler (Applied Biosystem® Veriti® 96 well) and the profile was set as: denaturation stage for 3 min at 95ºC, the second stage was set for 35 cycles at 94 ºC for 30 seconds, 56 ºC for 40 seconds, and 68 º C for 2 minutes and 30 seconds, the third stage at 68 ºC for 10 minutes. 1% agarose gel (stained with ethidium bromide) was used for PCR product loading and band visualization using ultraviolet light. The desired amplicons were purified using Gene JET PCR purification kit cat. No. KO70 (Thermo scientific). The amplicons were dissolved in 600 µl of binding solution and incubated for 10 minutes at 56 ºC. Solution was shifted in GeneJET purification columns, for 1 minute centrifuged the solution and then discarded the flow through. 700 µl Wash Buffer was used for washing and then added the elution buffer (10mM Tris-HCL, pH 8.5) in the middle of tube and centrifuged for 1 minute at room temperature to elute the desire DNA. Gene isolation primer sequences 5'-3'.

RPP13-F ATGGCTGATGCCTTTGTGTC

RPP13-R TCAAGAATAAGACTTCATAT

2.13 RPP13 gene cloning in vector

The RPP13 gene amplified PCR product was ligated in 2.8kb ampicillin-resistant cloning vector pTZ57-R/T (Thermo scientific #K1214). using PCR cloning kit (InsTAclone^{TM} PCR cloning kit) The ligation was executed by adding following reagents: 1 µl T4 DNA ligase, 6 µl 5X ligation buffer, 3 µl TZ57-R/T and nuclease free

water up to a volume of 30 µl in 1.5ml Eppendrof tube, then incubated the reaction samples for 16 hours at 4 º C to get more ligated products.

2.14 Electro-competent cells preparation

E. coli DH5α electro-competent cells were prepared by following method; the DH5α strain was streaked on LB media (Loria Broth media contains 10g NaCl, 10g tryptone, 14g agar and 5g yeast extract to make final volume of 1000 ml) and then incubated overnight at 37 º C. *E. coli* DH5α single colony was picked using a tip and placed it into 5ml of liquid LB media. The culture media was incubated overnight on a shaker at 37 º C and 350 rpm. The culture was diluted with LB media to get 1:100 dilutions for example 10 ml culture is added in 1000 ml LB media and then it is incubated for 3-4 hours to acquire the culture density 0.4-0.5 at OD_{600}. For 30 minutes the culture was chilled on ice and then it was placed into 50ml centrifuged tube, spun the culture for 10 minutes with 8000 rpm at 4 º C. Deionized water was used for washing and step was repeated for three times. The harvested cells were re-suspended in cryoprotectant (7% dimethyl sulfoxide DMSO) after third washing. 50 µl of cell aliquots are made in pre chilled Eppendrof tubes for electroporation and kept at -80 º C freezer.

2.15 Electro transformation in DH5α and Blue-white assay

Frozen aliquots of electrocompetant cells were used with the quantity of 50µl. These were thawed and 7 µl of ligated RPP13 gene was added in aliquot cells and mixed it mildly by pipetting. Cells and DNA mixture was allocated in 1mm electroporation cuvette. The cuvette was placed in gene pulser apparatus and pulse of 1800 Volts was given. 1ml of liquid LB was added and incubated at 37 º C in shaker for 1 hour. Transformed E. coli cell (100-200 µl) were spread on LB plates containing 50mg Ampicillin, 100 µl X-gal and 50 µl IPTG. Plates were incubated at 37 º C overnight. The recombinant clones were represented as white colony while the non-recombinant cells were shown as blue colonies on media.

2.16 Plasmid extraction and re-combinant clone's confirmation

The recombinant vector pTZ57R/T with RPP13 gene was analyzed by PCR using gene specific primers. Clones were isolated by thermo fisher scientific (GeneJET Plasmid Miniprep Kit). The recombinant colonies (white colonies) were elected and inoculated in 10ml of LB contain 50µg/ml ampicillin and then incubated at 37 º C overnight on 200rpm. The harvested cells grown overnight were centrifuged at 8000 rpm for 2 minutes at 4 º C in Eppendrof tubes. lysis buffer 200 µl was added and inverted the tubes for 4-6 times. When lysate become clear 300 µl neutralized buffer was added and inverted the tube for 10 times immediately and centrifuged at 13,000 rpm for 5 minutes. The supernant was shifted to column and centrifuged for 30 seconds again and discard the flow-through. 500 µl wash buffer was added and centrifuged at 14,000 rpm for 30-60s, the flow-through was discarded and additional centrifugation was done to remove all the residues. 60 µl elution buffers was added in the center of column, incubated it for 2 minutes and centrifuged it to elute the pDNA. The plasmid DNA was confirmed by agarose gel electrophoresis.

Gene specific RPP13-F/R primers was used to confirm the desired recombination by PCR amplification process. M13, universal sequencing primers were used to sequence the cloned amplicon of RPP13 in pTZ57R/T vector.

2.17 RPP13 gene in-silico analysis

RPP13 gene was sequenced from Humanizing genomics Macrogen using Sanger sequencing method. Forward and Reverse gene sequencing fragments were assembled using Bio edit sequence alignment editor (version 7.2). DNA sequence of *solanum lycopersicum* disease resistance protein -RPP13 was downloaded from NCBI (XM_010321022.3). Clustal W was used to align the sequence in Bio edit.

2.18 Phylogenetic analysis

Phylogenetic tree was constructed by MEGA-X using the Neighbor-Joining method with 1000 bootstrap replications. Protein sequences from eight species, *Actinidia chinensis, Arabidopsis thaliana, Capsicum annuum, Manihot esculenta, Nicotiana*

attenuata, Populus trichocarpa, Solanum lycopersicum, Solanum tuberosum, and *Vitis vinifera*, for phylogenetic analysis were downloaded from EnsemblPlants database (https://plants.ensembl.org/). Protein sequence alignment was conducted through CLUSTALW. Alignment was then used to build the phylogenetic tree.

Results

3.1 Plant infection and disease screening

Tomato plant leaves were inoculated with spores of *A. solani* using grafting inoculation method, and chlorosis around the inoculums appeared within 48 hours post-inoculation followed by the development of typical blight symptoms at 96 hours (Fig: 3.1B). yellow based with black circular rings appeared on the leaves.

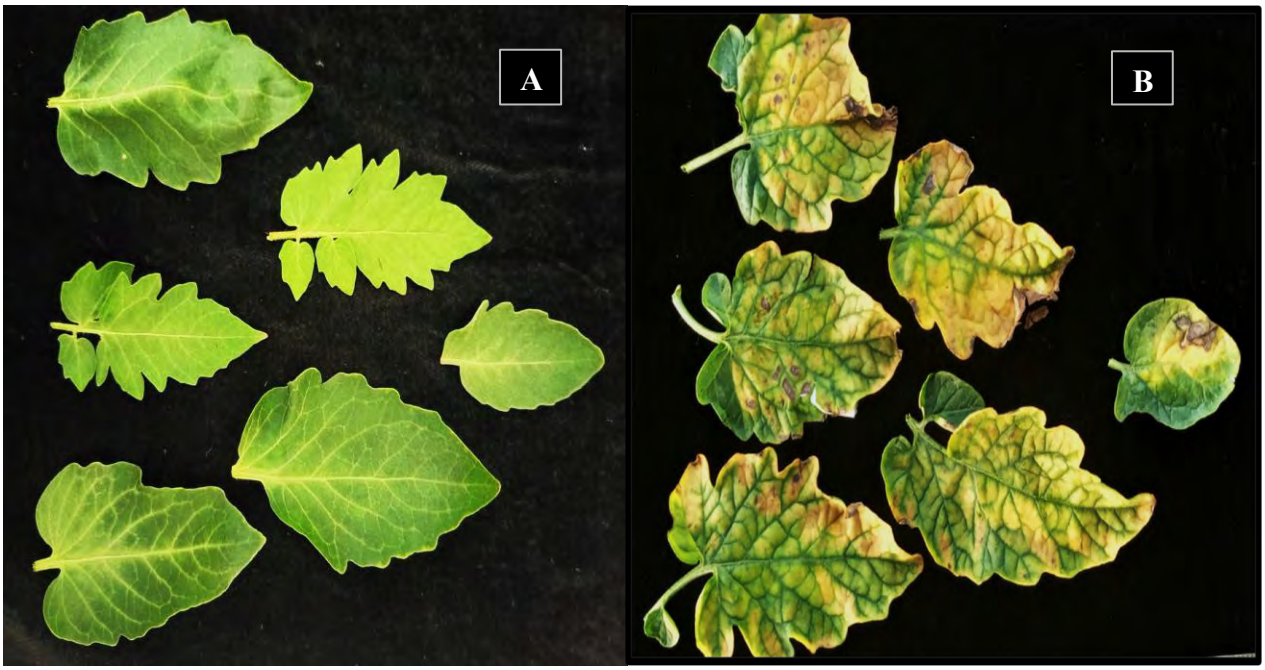


Figure 3.1: Early Blight disease plants leaves. (A) Show healthy tomato leaves, (B) show diseased tomato leaves

Healthy and diseased samples were collected at 96 hours after infection and RNA was isolated from healthy and disease leaves of four accessions *Solanum Lycopersicum* (38046, 38039, 19890 and Roma) and *Solanum Chilense* (19906). RNA bands were confirmed by gel electrophoresis. (Fig:3.2). RNA concentration was checked using BioSpec-nano Spectrophotometer for LIFE SCIENCE instrument and cDNA was prepared for further analysis.

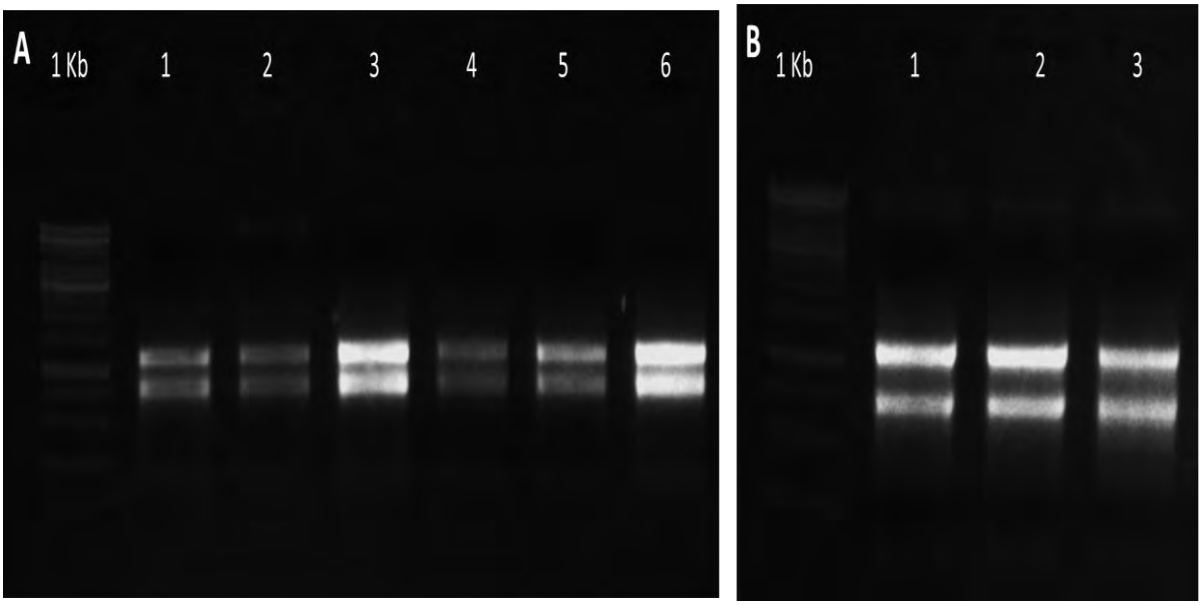


Figure 3.2: RNA extraction of Healthy and Diseased plant leaves.

(A) lane 1-6 showing RNA extracted from *Solanum lycopersicum* accessions (38039 S, 38046 H, Roma H, 19890) *Solanum Chilense* accession (19906 H) respectively and (B) lane 1-3 showing 38046 D, 19906 D, Roma D with 1Kb ladder

3.2 Transcriptomic data analysis identified differentially expressed genes

Raw RNA-Seq data for early blight and late blight was downloaded from GEO database under the accession number **GSE75923** and GenBank Sequence Read Archive (SRA) under accession number **SRP041501** respectively (Sarkar *et al*., 2017; Zuluaga *et al*., 2016). Using different bioinformatic tools expression vales in FPKM were generated. NLR genes sequences and ids of *Solanum lycopersicum* were downloaded from Tomato Genomics Resources Database and retrieved the expression values of these NLR genes from the transcriptomic analysis expression vales. A heat map of expression values of retrieved 16 NLR genes for early blight (Fig 3.3) and 75 NLR genes for late blight (Fig.3.4) was generated which showed upregulation and down regulation of NLR genes. From 267 NLR genes of *Solanum lycopersicum,* 15 NLR genes were retrieved in which 4 genes (Solyc01g086810.2, Solyc04g009110.1, Solyc04g009150.1,

Solyc04g009250.1, Solyc04g008170.1) showed upregulation against early blight. In case of late blight 75 NLR genes gave differential expression values in which 21 NLR genes showed upregulation while the rest of NLR genes showed either downregulation or no regulation after disease treatments. Approximately 6-7 NLR genes showed high upregulation values, 5 upregulated and 1 down regulated NLR genes were marked for further analysis.

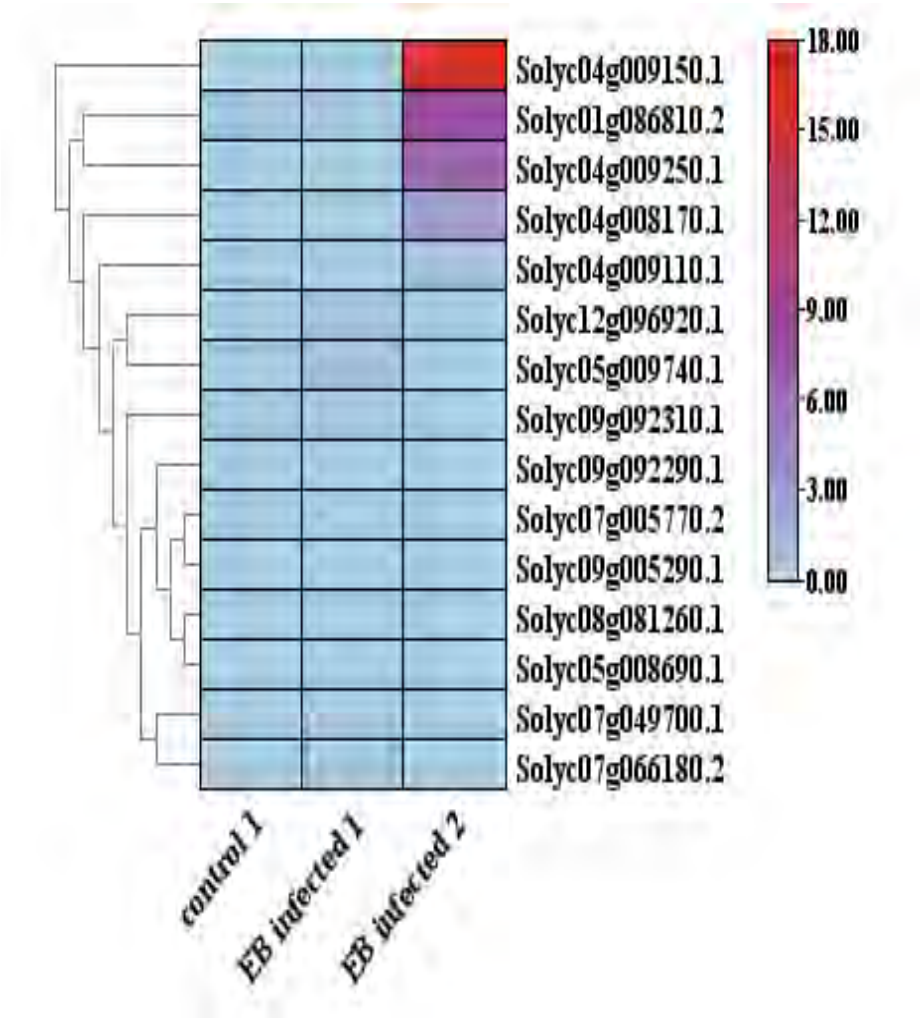


Figure 3.3: Heatmap showing RNA-seq data for NLR genes against early blight disease. Color patterns indicate the upregulation and downregulation of NLR genes under control and disease conditions.

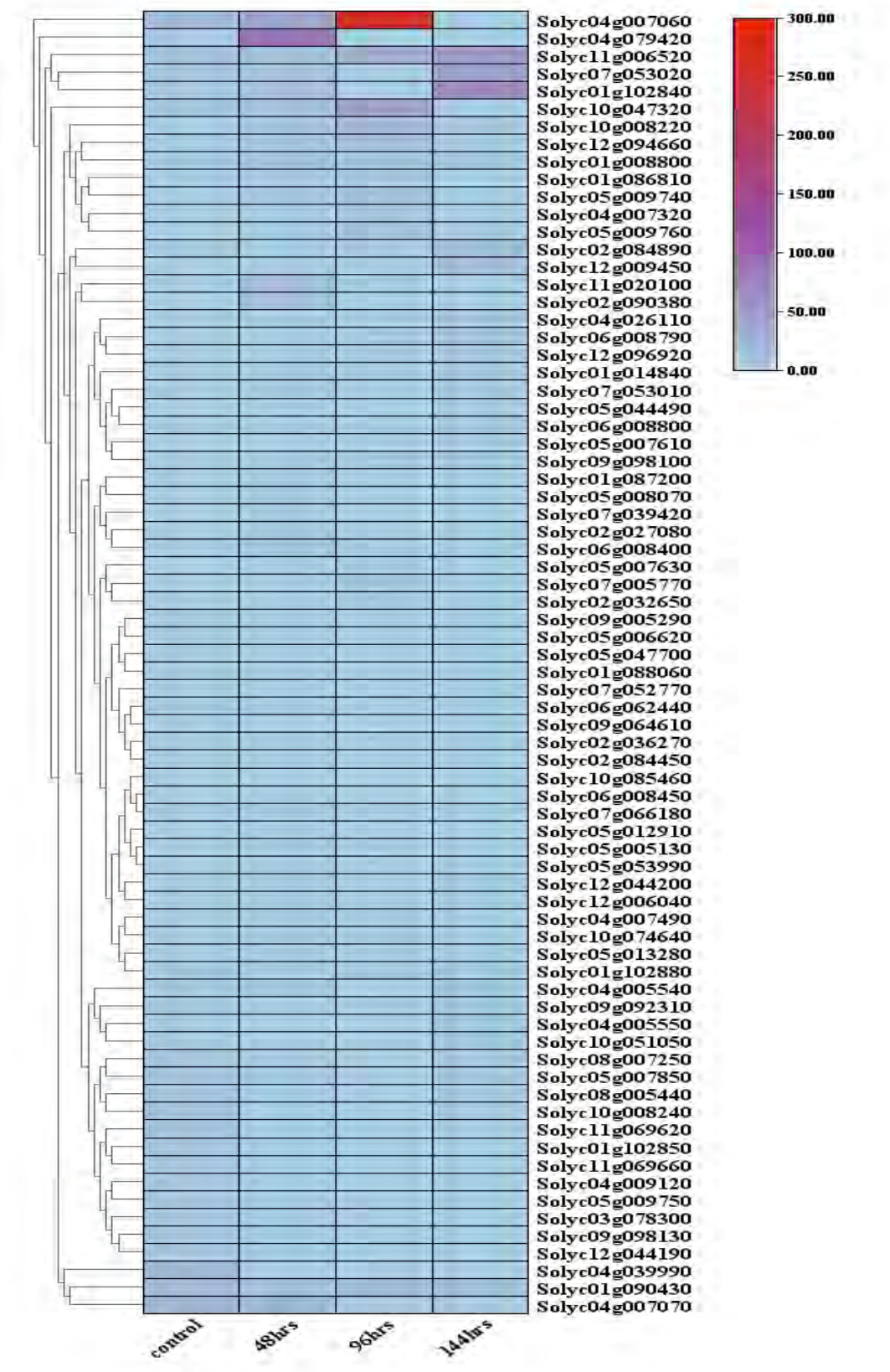


Fig3.4: Heatmap of NLR genes against late blight

3.3 g: Profiler results represented GO Enrichment analysis and KEGG analysis of NLR genes

g: Profiler was performed for the bioinformatics analysis and the enrichment information of candidate genes encoding proteins. From GO, KEGG databases, g: Profiler enriched a large number of terms. Only 1 pathway was enriched by g: Profiler. The analysis reveals that pathway is involved in plant pathogen interaction.

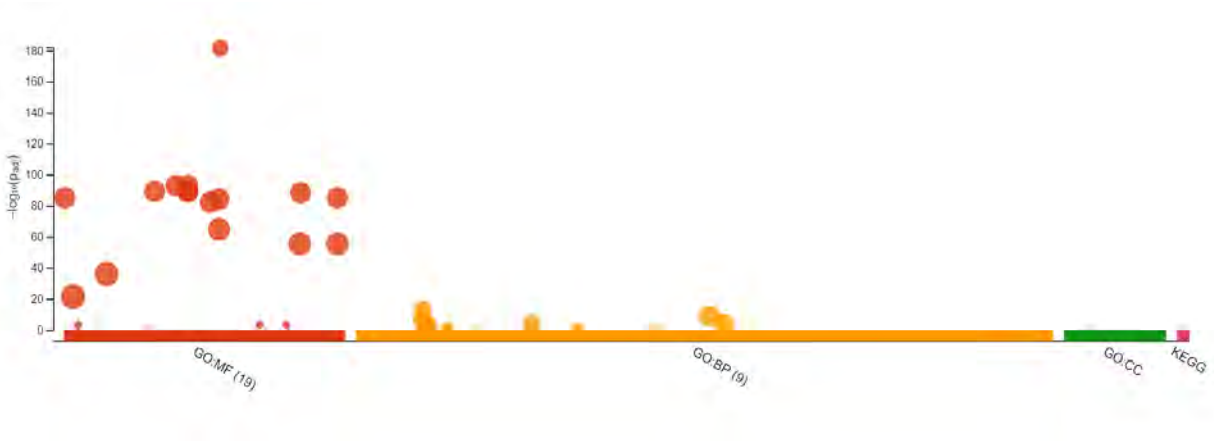


Figure 3.5: The significantly changed terms enriched by GO, KEGG databases

Biologic process (BP), molecular function (MF) and cellular component (CC) induced by candidate gene proteins are enriched term by GO enrichment analysis. 19 BP, 9 MF and 1 CC terms were enriched by g: Profiler in BP, MF and CC, respectively. Classification analysis revealed that in GO: BP genes were mainly concentrated in defense response, stimulus response and signal transduction. ADP binding, adenyl ribonucleotide binding, purine nucleotide binding, anion binding, carbohydrate derivative binding, nucleoside phosphate binding, small molecule and NAD+ nucleosidase activity were the main pathways of GO:MF. Extrinsic component of the plasma membrane pathway was enriched in GO: CC.

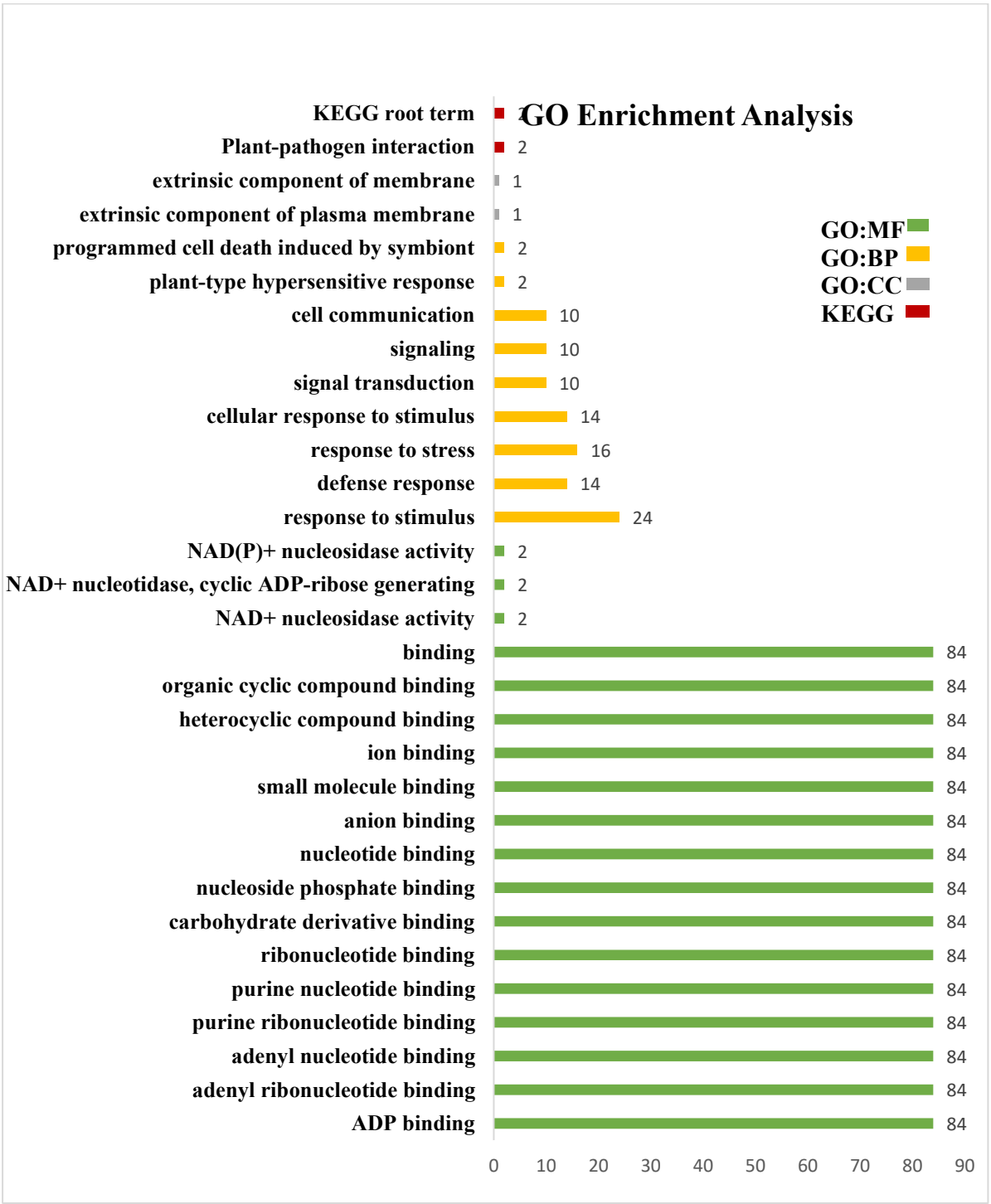


Figure 3.6: GO enrichment analysis by g: Profiler

Table.3.1: No. of NLR genes involved in GO enrichment analysis

GO enrichment	No. of NLR genes
GO: BP	24
GO:MF	84
GO: CC	1
KEGG	2

3.4 Upregulated genes of early and late blight represented high expression in diseased/infected plant tissues

The expression level of NLR resistance genes RPP13, RPM1, NRC4, R1B16 and R1B12 in *Solanum lycopersicum* accessions (Roma, 38046, 38037, 19890) and *Solanum Chilense* (19906) were carried out using Real-Time PCR analysis. The tissues selected for the expression of resistance genes were healthy and diseased plant leaves tissue. Our results suggested high expression values in Roma diseased plants. NLR genes were more expressed in diseased plants as compared to healthy plants but Roma showed high expression values in diseased plant. Comparative analysis described that there are some variations in expression of different NLR genes but the Roma healthy and diseased plant showed the same expression pattern against NLR resistance genes. *Solanum Chilense* show up and down genes regulations but *S. Lycopersicum* exhibited higher expression values

The graph bars are defining the expression difference and co-relation among the healthy and diseased tissues. Fig 3.7.1 represented the R1B12 gene gave low expression in diseased plant tissues compared to healthy plant tissues. Fig 3.7.2 late blight resistant

gene NRC4 exhibits comparative higher expression in 38046 and Roma diseased plant. Fig 3.7.3 RPM1 showed some variations but expression in infected tissues is higher as compared healthy plant. However, Roma diseased tissues expressed high expression values Fig, 3.7.4 RPP13 show upregulation in infected plant. Fig 3.7.5 R1B16 exhibited high expression in infected showing upregulation of NLR gene.

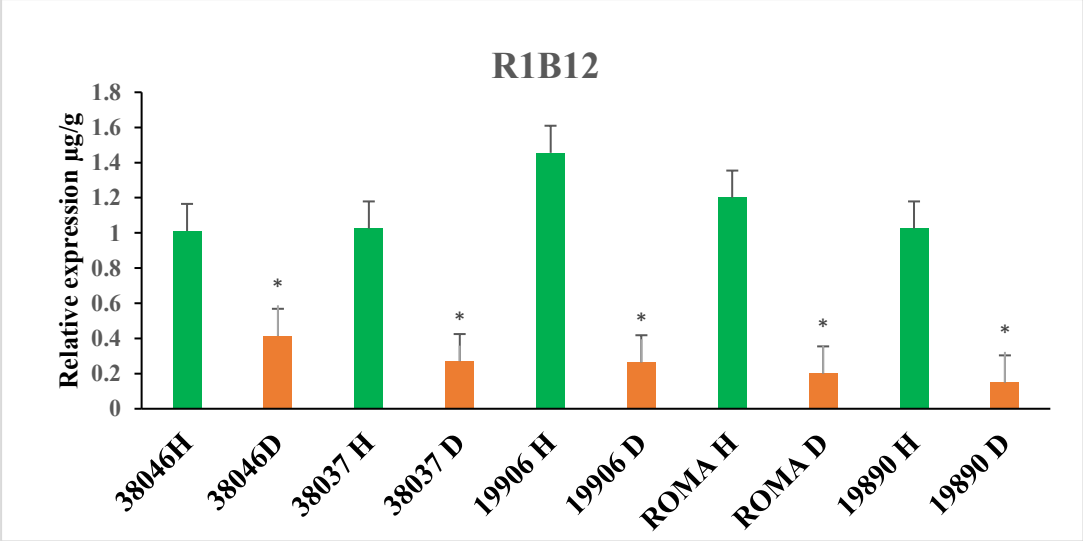


Figure 3.7.1: Expression analysis of NLR genes R1B12. Expression analysis of R1B12 show down regulation of gene in infected plants.

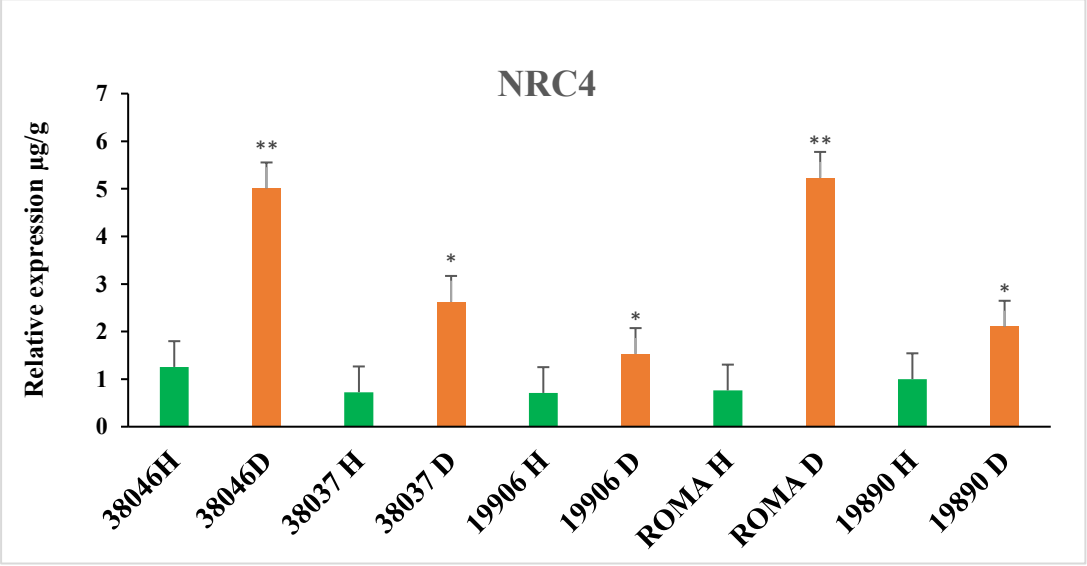


Figure 3.7.2: Expression analysis of NLR genes NRC4. Expression analysis reveal differential upregulation of NRC4 gene in infected plants

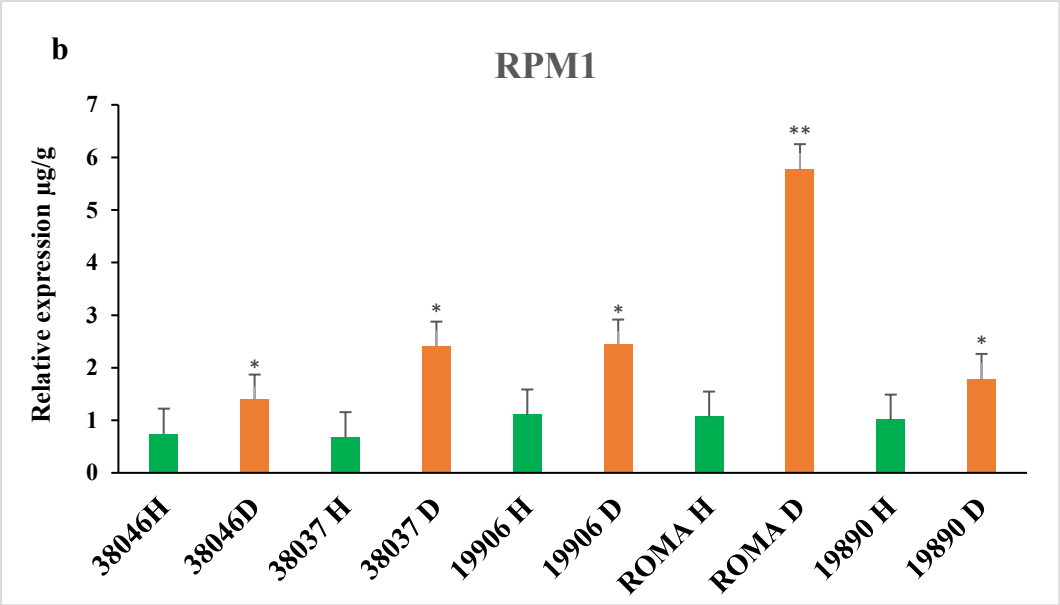


Figure 3.7.3: Expression analysis of NLR genes RPM1. Expression analysis revealed upregulation of gene in infected plant tissues.

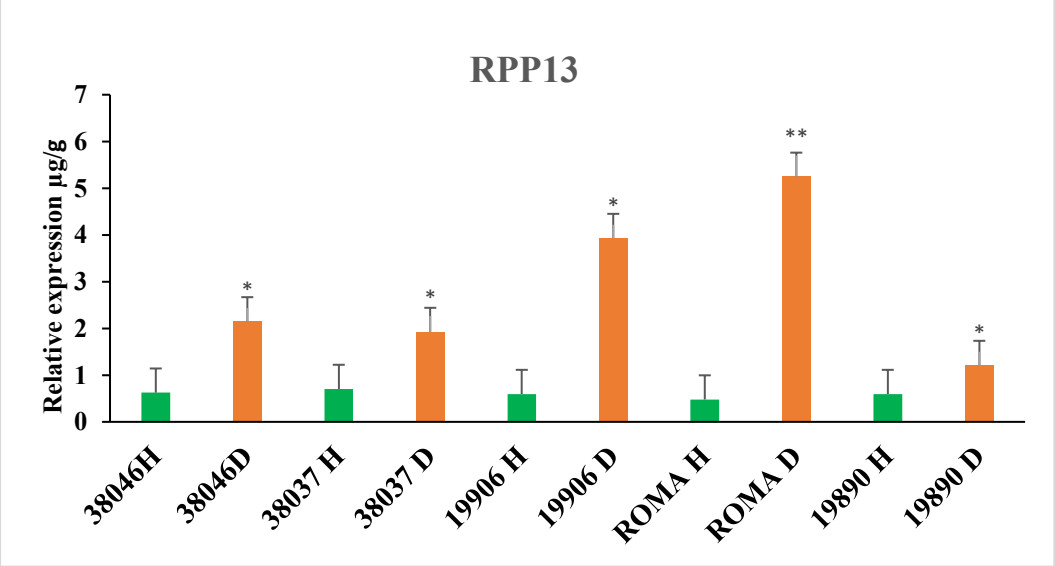


Figure 3.7.4: Expression analysis of NLR genes RPP13. Gene is highly expressed in infected plants leaves espeasially Roma D showing upregulation of gene.

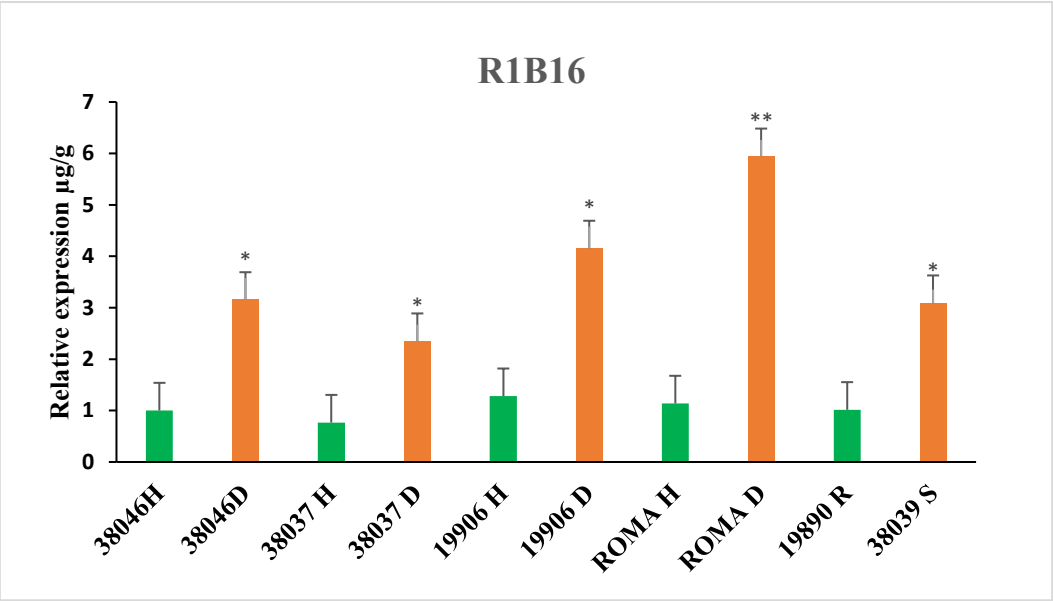


Figure 3.7: Expression analysis of NLR genes R1B16. Exression analysis exhibited upregulation of gene in infected plants.

3.5 Network analysis of NLR genes showed no significant Protein-Protein interaction

The interactions among NLR genes related to early blight and late blight disease resistance was investigated using information from STRING database version (11.5) (https://string-preview.org/cgi/network) to construct a PP1 network. Interaction network was predicted from 15 nodes proteins with enrichment *p*-value 1.63×10^{7} at medium confidence level. RPM1(Solyc01g086810.2.1) show powerful PPI network and interacted with LRR proteins Solyc09g059430.2.1, Solyc08g005440.2.1, Solyc08g005440.2.1. the other predicted NLR genes RPP13(Solyc04g009150.1.1), NRC4(Solyc04g007060.2.1), R1B16(Solyc04g007070.2.1), R1B12 (Solyc10g008240) show no interaction with each other.

Differential expressed NLR genes belongs to NB-LRR family and their PPI reflected no significant interaction with each other. However, RPM1 (Solyc01g086810.2.1) showed some interaction with 1st shell interactors and R1B12 (Solyc10g008240.2.1) showed interaction with Pto kinase; Serine/threonine protein kinase Pto (2^{nd} shell interactors) Fig 3.10A.

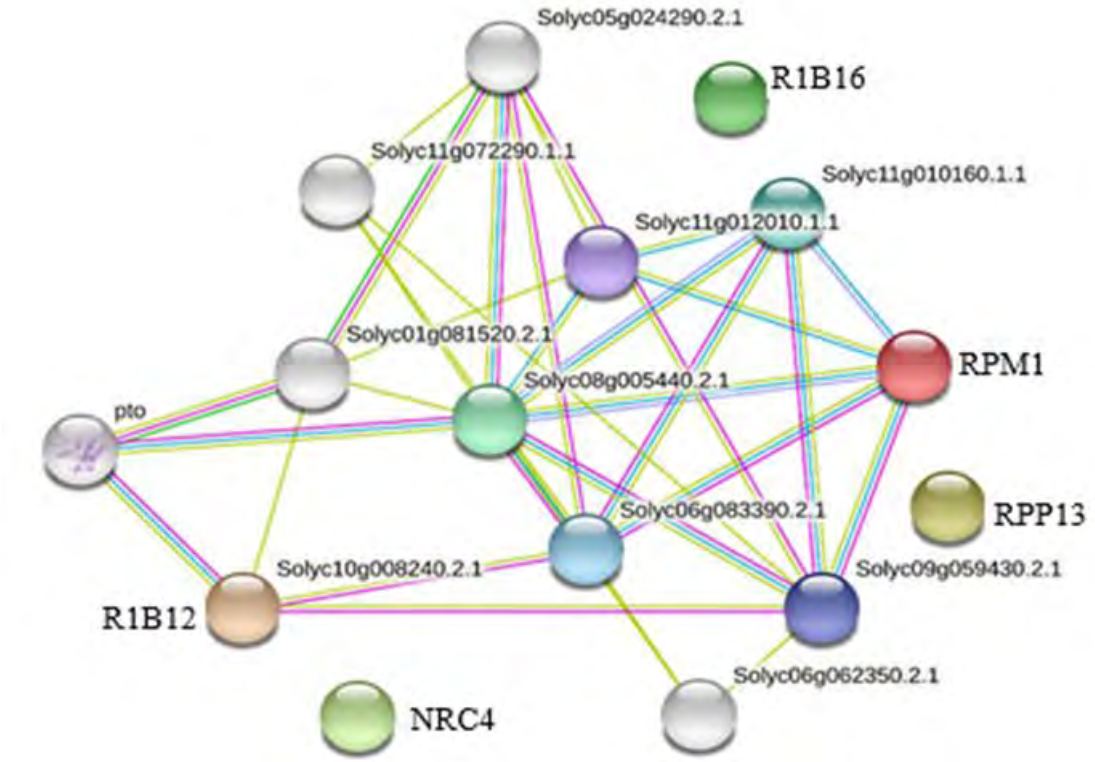


Figure 3.8: Network analysis of differential expressed NLR genes**.** Experimental evidence is exhibited by the purples. The green lines exhibit gene neighborhood, the blue lines describe gene co-occurrence database evidence, the yellow lines and black lines represent text mining evidence the co-expression evidence respectively.

3.6 Solanaceae species showed maximum Co-occurrence of NLR genes in Phylogenetic tree

The occurrence of NLR genes families across genomes with maximum similarities was investigated using information from STRING database version (11.5) (https://string-preview.org/cgi/network).

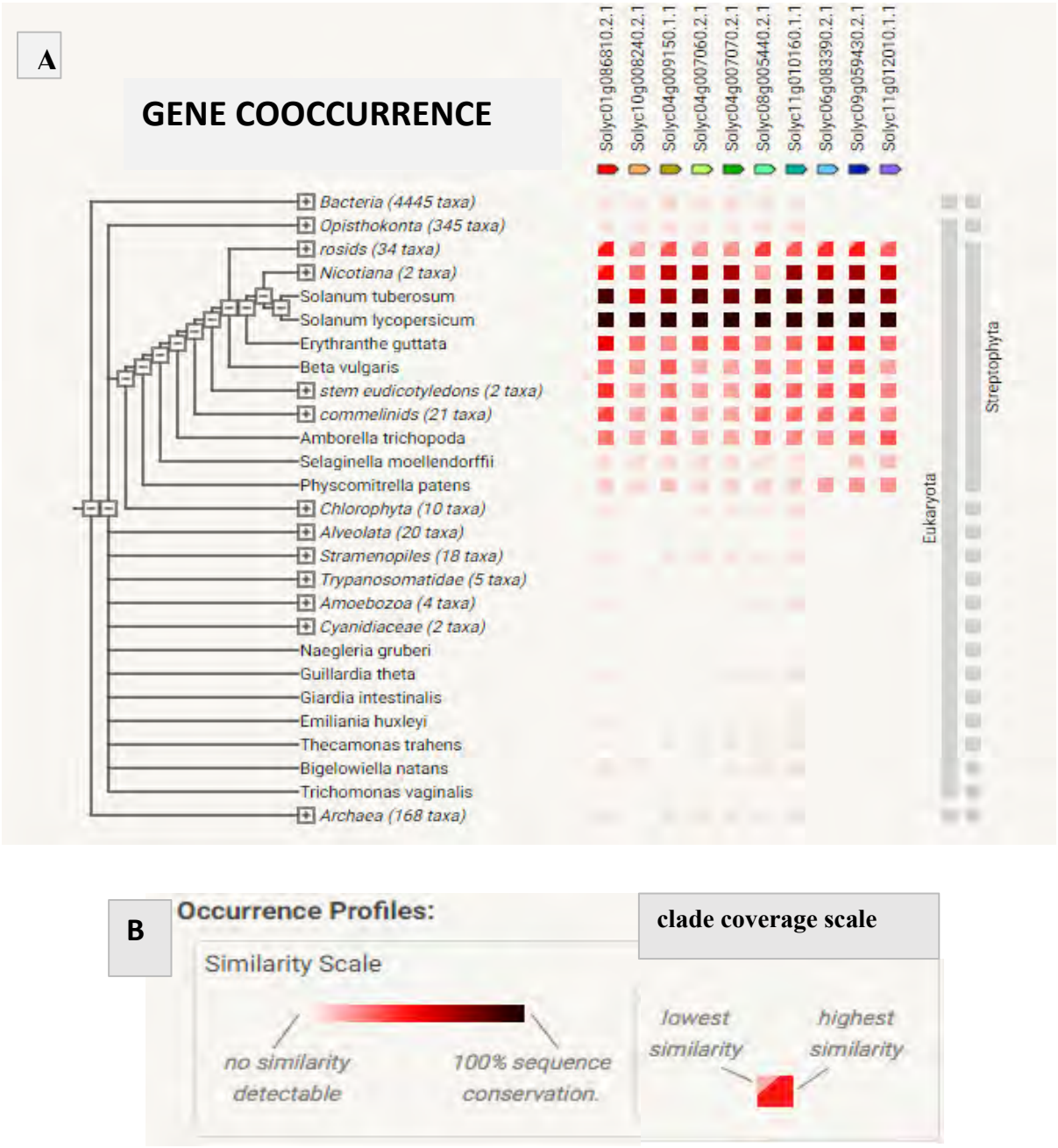


Figure 3.9. NLR Gene co-occurrence in other genomes (A), Occurrence Profiles (B)

Co-occurrence network analysis represented that Solanaceae species have maximum similarities Fig 3.11. A. The dark red color represented 100% sequence conservation while light color show low similarity or no similarity. Correlation of these color presence and absence-based profiles can predict interactions. The group of genomes

that collapsed in the phylogenetic tree depicted two distinct color showing lowest and highest similar conservation within clade Fig 3.11.B.

3.7 UP-regulated early blight RPP13 gene was isolated from *Solanum lycopersicum* accession 19890 and cloned in pTZ 57R/T vector

For amplification of 2535 bp RPP13 CDS sequence, cDNA of *Solanum lycopersicum* accession 19890 was used as a template with specific forward gene primer 5' -ATGGCTGATGCCTTTGTGTC-3' and reverse gene primer 3'-TCAAGAATAAGACTTCATAT-5'using standard PCR settings. 1% agarose gel was used to visualize the resulting PCR product. Appearance of single band of 2535 kb on gel identifying the specificity of primers for isolation of desire gene RPP13 (Figure 3.12). The fragment was separated on gel and purified by using gel purification kit. T4 DNA ligase enzyme was used to ligate the quantified PCR product into pTZ 57R/T vector (Figure: 3.13)

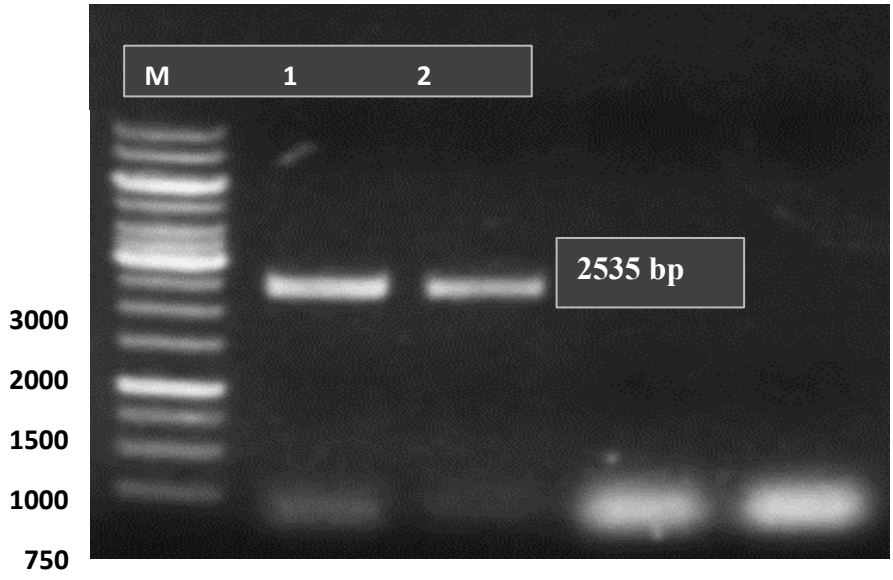


Figure: 3.10: Agarose gel electrophoresis showing amplicon of RPP13 *from Solanum lycopersicum* accession 19890 with 1Kb DNA ladder as Marker

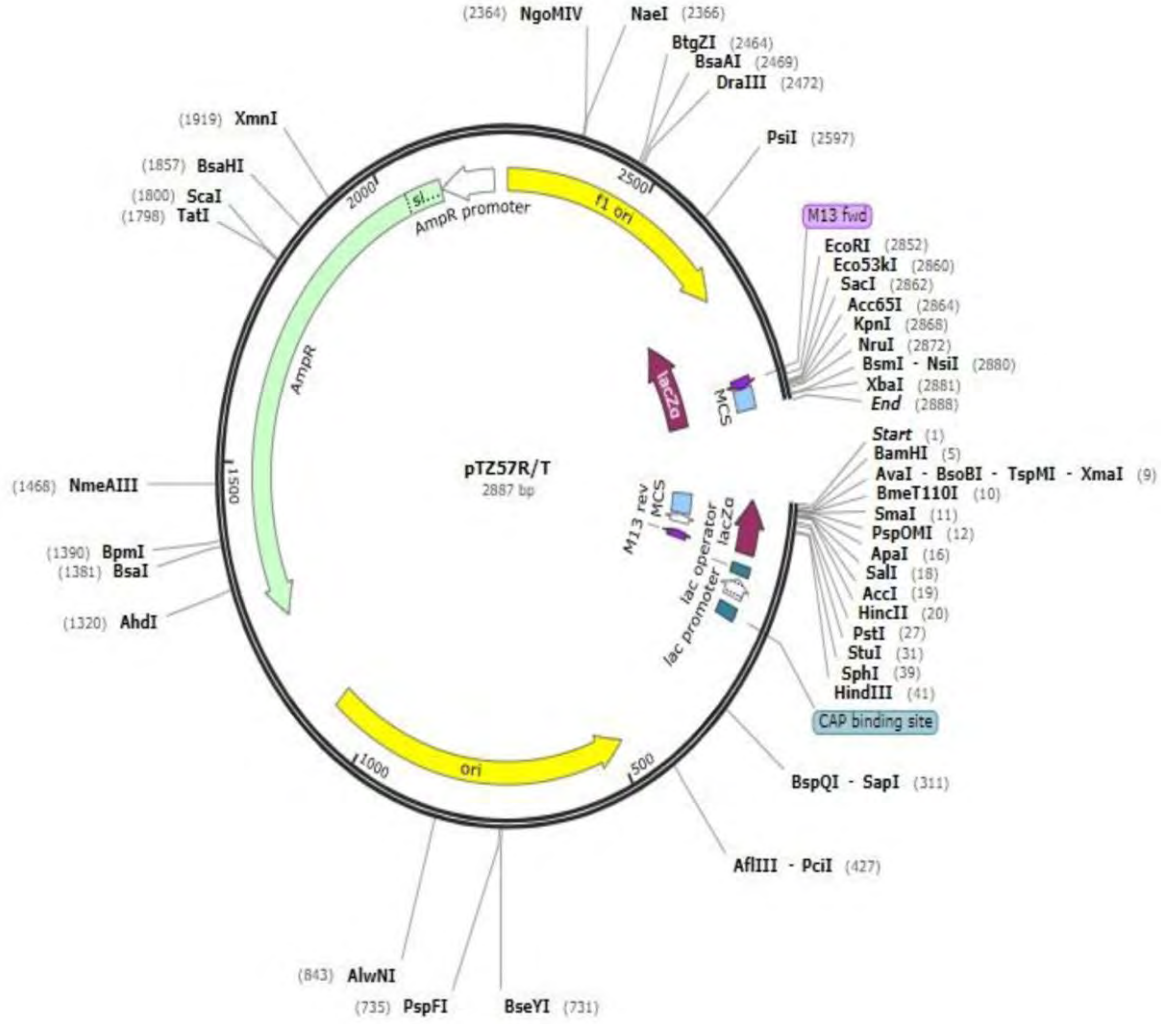


Figure 3.11: Map of T vector pTZ57R/T (Fermentas, Germany)

3.8 Cloning of RPP13 gene in E. coli

PCR amplification of the NLR gene RPP13 showed an expected amplicon size of 5.3bp (Figure 3.14). Cloning of the NLR gene RPP13 fragment using pTZ57R/T vector in *E. coli* DH5α host strain resulted in formation of recombinant clones nurture the gene. A 5.3 Kb size band of orientated gene in pTZ57R/T vector was visualized by Agarose gel electrophoresis.

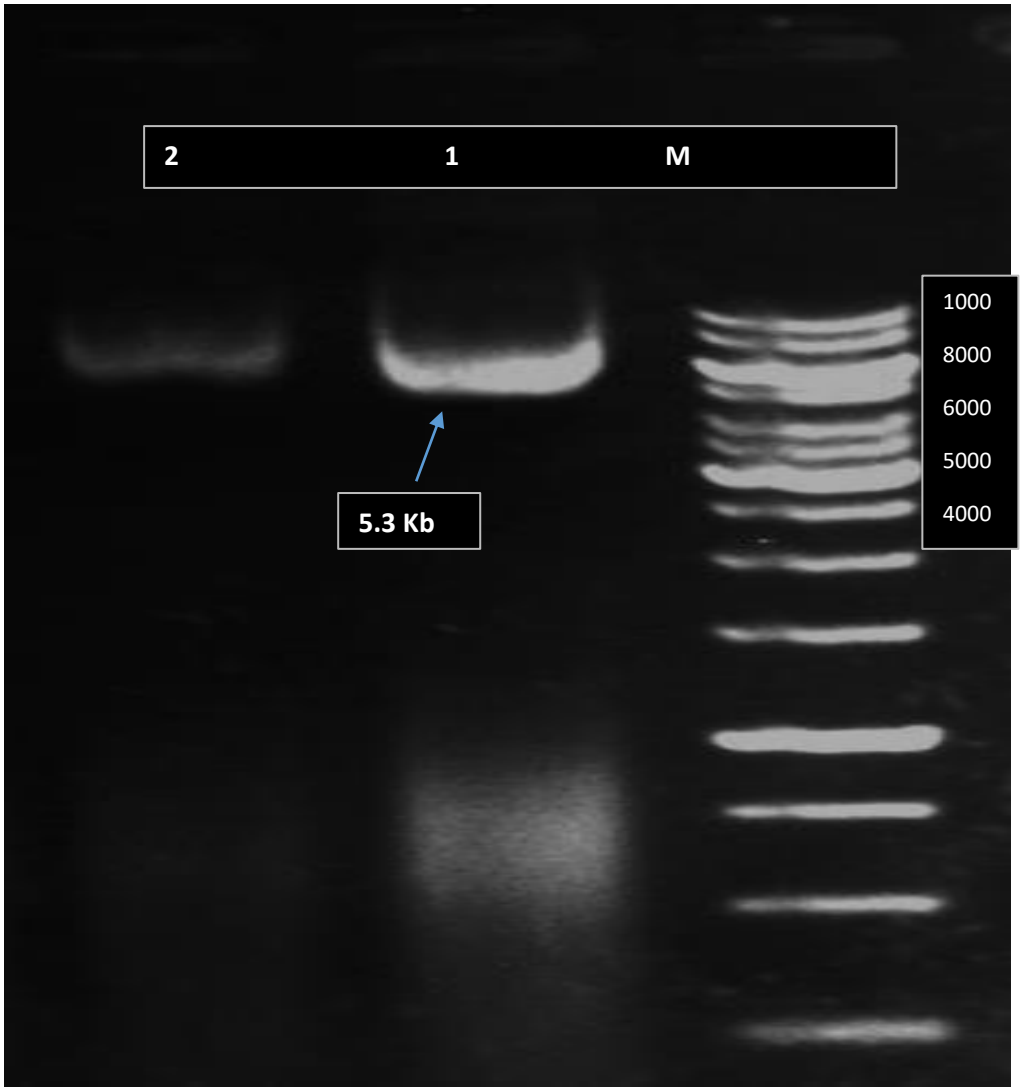


Figure 3.12: Agarose gel electrophoresis analysis for PCR amplification of the RPP13, gene orientated within the recombinant pTZ57R/T vector. Lane 1-2 show 5.3 Kb size orientated gene in pTZ57R/T vector with 1Kb DNA Ladder as Marker

3.9 In silico analysis of RPP13 gene show interaction with other genomes

RPP13 clone was sequenced using sanger sequencing method and assembled the sequence using bioinformatic tool. DNA sequence of *solanum lycopersicum* disease resistance protein -RPP13 was downloaded from NCBI and aligned it with cloned RPP13 sequence in Bio-Edit. In Figure 3.13 Asterisk sign (*) represented conserved regions of both genes. The alignment results described substitution mutation, delection and gaps are present in sequenced gene.

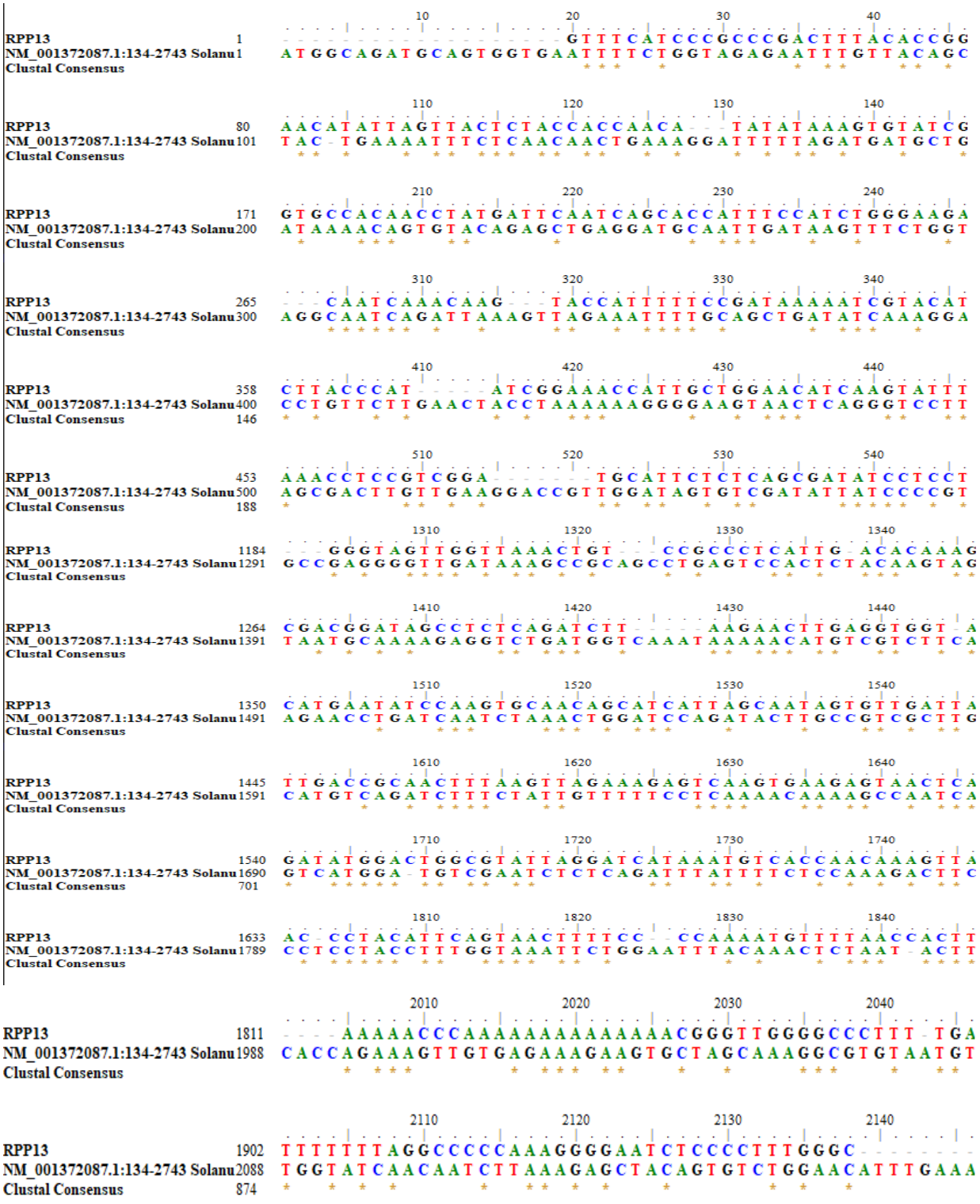


Figure 3.13: In silico sequence analysis of cloned RPP13 gene. Asterisk sign shoed conserved regions of genes which show 70 % homology of sequences.

3.10 Phylogenetic analysis of cloned RPP13 protein

Phylogenetic tree using cloned RPP13 sequence with other plant genomes was constructed by neighbor- joining method. The evolutionary relationship was described in tree. The cloned sequence was name as 19890. In (Figure 3.14) 19890 showed closed relationship with *Actinidia chinensis* and *Vitis vinifera* and shared common cluster with *Solanum lycopersicum*, *Solanum tuberosum* and *Arabidopsis thaliana*. *Capsicum annum, Nicotiana attenuata, Manihot esculenta, Populas trichocarpa* are present in different cluster and show divergence from 19890.

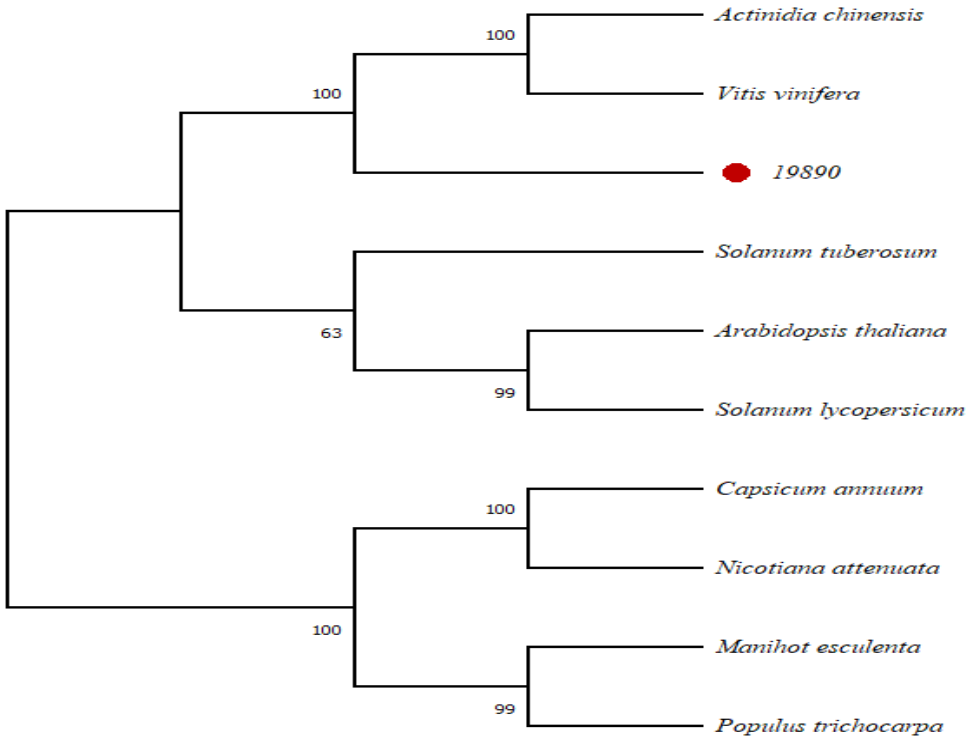


Figure 3.14: Neighbor -joining tree of cloned RPP13 genes. Red dot represented the cloned sequence protein sequence of 19890 accession. Cloned gene and S. Lycopersicum present in the same cluster.

Discussion

Many factors are responsible for low production of tomato but early blight (*Alternaria solani)* and late blight (*Phytophthora infestans*) are major factors of destruction of tomato plants all over the world. Early blight disease attack on calyx and fruit tissues contributes up to 79% yield losses in tomato (Chaerani and Voorrips 2006) and late blight involves in yield loss up to 65-70% (Amin *et al*., 2013). In Pakistan according to an estimate 49 to 91 % yield loss is due to early blight and late blight (Azam and shah 2003; Akhtar *et al*., 2012 and 2016). Control measures, chemicals, fungicides are used to overcome the destruction due to EB and LB. Conventional breeding, Marker-Assisted Selection (MAS) and Genomic Selection (GS) are techniques that assist the breeding program by facilitating the selection efficiency (Collard *et al*,.2007; Desta *et al*., 2014). Genetically modified crops and Genome-Wide Association Studies (GWAS) help in modification of genetic makeup of any organism and to identify the correlated traits in a genetically varied population (Christou *et al*., 2013; Rahman *et al*., 2018). Resistance genes (R genes) are the important gene family to resist and recognize the pathogen). Several *R* genes have been cloned from tomato, potato and pepper which contains NB-LRR domains, and are being used in current breeding technologies. The tomato genome assembly has identified the large size of the NB-LRR gene family, and potential *R* gene repository (Andolfo *et al*.,2013).

Genes were identified after transcriptomic analysis of raw RNA-seq data for early blight and late blight. The genes showing high FPKM value was marked as upregulated genes and downregulated genes with low FPKM for further analysis. Four accessions of *Solanum lycopersicum* (38046, 19890, Roma) and one accession of Solanum Chilense (19906) were used to study the resistance conferred by NRC4, R1B12, R1B16, RPP13, RPM1 genes belonging to NB-LRR class of plant resistant gene. These NLR genes were retrieved from RNA-seq analysis data. The RPP13 gene is under high selection pressure exhibiting highly diverse allele. The CC:NBS domain of RPP13 protein have been shown to be under selection for amino acid conservation and LRR is under extreme levels of diversifying selection (Bakker *et al*., 2006; Ding *et al*., 2007; Rose *et al*., 2004). The RPM1 was initially identified in *Arabidopsis thaliana* that

conferred resistance against *P. syringae* (Debener *et al*., 1991). R1B16 and R1B12 are putative late blight resistance protein homolog transcript variant (accession XM_019211150 and XM_019216095) (https://www.ncbi.nlm.nih.gov). NRC4 are helper NLR proteins and mediate the immunity of Rpi-blb2 in potato against *P. infestans* (Wu *et al*., 2017).

Expression analysis of NLR genes with real time PCR for late blight resistance genes (NRC4, R1B12, R1B16) and early blight resistance genes (RPP13, RPM1) were revealed by different gene expression found in tomato healthy and diseased plants. In Arabidopsis thaliana upregulation of RPP13 and RPM1 resulted defense against blister blight (BB) (Bittner *et al*.,2000). NRC4 act as helper NLR and involved in defense mechanism (Wu *et al*., 2017). Results of expression analysis by RT-PCR represented upregulation of NLR genes RPP13, RPM1, NRC4 and R1B16 in different varieties of tomato cultivars. Downregulation of R1B12 in different cultivar resulted in resistance. Roma plant showed upregulation of NLR genes that comprises resistance in plants. GO enrichment analysis revealed that RPP13, R1B16, RPM1, R1B12 and NRC4 (contains leucine rich repeat receptor) are involved in molecular functioning e.g., ADP binding, adenyl ribonucleotide binding, purine nucleotide binding, anion binding, carbohydrate derivative binding, nucleoside phosphate binding, small molecule and NAD+ nucleosidase activity.

Only one NLR gene RPM1 is involved in KEGG analysis and involved in plant pathogen interaction and involved in effector trigger immunity as show pathway of interaction with AvrRpm1, RIN4, Avr2, AvrXa21, AvrB, RIN4, RPS2 Effector and show hypersensitivity response (http://www.genome.jp/kegg/)(Kanehisa *et al*. 2006). Protein-Protein Interaction of NLR genes revealed no significant interaction, as many of the NLR genes are not annotated and these genes might show interaction with other genes. RPM1 showed some interaction with other genes. In *Arabidopsis thaliana* direct and indirect association of RPM1 might be act as defense against Blister Blight (BB) (Spoel *et al*., 2012). The gene co-occurrence phylogenetic tree has identified the higher conservation of NLR genes (RPP13, R1B16, RPM1, R1B12 and NRC4) in Solanaceae species (*Solanum tuberosum* and *Nicotiana*).

Pyramiding resistance genes by classical method is time consuming, it may take several years. Genetic transformation of individual cloned *R*-genes in to cultivars via different mechanisms, can be faster and it might be overcome the problem of individual genes segregating in backcross populations (Klymiuk *et al*., 20018). Speed cloning is a great innovation of breeders and scientists. However, this technique can not only be used to speedily clone the R genes but also can be used to clone genes for other traits improvement including yield, nutrition profile and climate tolerant crops to make them accessible to the world's grower (Nazia Rehman *et al*. 2020). About 30 functional resistance genes from Solanaceae (tomato and potato) have been cloned. Resistance gene enrichment sequencing (RenSeq) from some Solanaceae plants is being used for identification of putative and functional NLR genes (Jupe *et al*., 2012; Andolfo *et al*., 2013; Kim *et al*., 2014).

NLR gene RPP13 is under high selection pressure exhibiting highly diverse allele and show broad spectrum resistance against diseases (Bakker *et al*., 2006; Ding *et al*., 2007; Rose *et al*., 2004). RPP13 is unique NBS: LRR R-gene that have ability to retain its function in rar1, ndr1, eds1, pad4, npr1, and double eds1, ndr1 mutant plants (Bittner *et al*., 2001) NLR gene RPP13 was amplified using conventional PCR and a fragment of 2.5kb was isolated. The isolated gene was successfully cloned using pTZ 57R/T vector of 2.8kb in *E. coli* DH5α host strain and sequenced the clone to show the variations present in isolated gene. In silico and phylogenetic analysis revealed conserved region, mutations and gaps present in gene sequences.

Future perspective

Variations like spontaneous mutations, chemically induced mutations or physical irradiation are rare to occur in populations and might take many years to achieve the targets. Advancement in biotechnology and its integration in breeding program reduce the generation time. Speed cloning and genome editing techniques implementation may help to reduce the generation time. CRISPR/Cas system is a versatile and rapid tool that target the any part of genome of an organism. It may help to unveil more NLR genes for further studies. speed cloning and CRISPR/Cas system are safer and easier tools that may help to generate improved crop varieties.

REFERENCES

Adebayo, O. (2005). Integrated control of bacterial and Fusarium wilt disease of tomato in Ibadan, Nigeria. *A Ph. D Thesis submitted to the Department of Crop Protection and Environmental Biology, University of Ibadan*.

Adenuga, A. H., Muhammad-Lawal, A., & Rotimi, O. A. (2013). Economics and technical efficiency of dry season tomato production in selected areas in Kwara State, Nigeria. *Agris on-line Papers in Economics and Informatics*, *5*(665-2016-44983), 11-19.

Adhikari, P., Oh, Y., & Panthee, D. R. (2017). Current status of early blight resistance in tomato: an update. *International journal of molecular sciences*, *18*(10), 2019.Agrios, G. N. (2005). Introduction to plant pathology. *Elsevier Academic Press Publication.*

Akhtar, K. P., Saleem, M. Y., Iqbal, Q., Asghar, M., Hameed, A., & Sarwar, N. (2016). Evaluation of tomato genotypes for late blight resistance using low tunnel assay. *Journal of Plant Pathology*, 421-428.

Akhtar, K. P., Ullah, N., Saleem, M. Y., Iqbal, Q., Asghar, M., & Khan, A. R. (2019). Evaluation of tomato genotypes for early blight disease resistance caused by Alternaria solani in Pakistan. *Journal of Plant Pathology*, *101*(4), 1159-1170.

Akhtar, K. P., Saleem, M. Y., Asghar, M., Ali, S., Sarwar, N. I. G. H. A. T., & Elahi, M. T. (2012). Resistance of Solanum species to *Phytophthora infestans* evaluated in the detached-leaf and whole-plant assays. *Pak. J. Bot*, *44*(3), 1141-1146.

Amin, M., Mulugeta, N., & Selvaraj, T. (2013). Field evaluation of new fungicide, Victory 72 WP for management of potato and tomato late blight (*Phytophthora infestans* (Mont) de Bary) in West Shewa Highland, Oromia, Ethiopia. *Journal of Plant Pathology and Microbiology*, *4*, 192.

Andolfo, G., Sanseverino, W., Rombauts, S., Van de Peer, Y., Bradeen, J. M., Carputo, D., ... & Ercolano, M. R. (2013). Overview of tomato (*Solanum lycopersicum*) candidate pathogen recognition genes reveals important Solanum R locus dynamics. *New Phytologist*, *197*(1), 223-237.

Arora, S., Steuernagel, B., Gaurav, K., Chandramohan, S., Long, Y., Matny, O., ... & Wulff, B. B. (2019). Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nature biotechnology*, *37*(2), 139-143.

Azam, F., & Shah, S. J. (2003). Exploring the role of farmer led management practices on various tomato and cucumber diseases in Peshawar and Dragai areas of NWFP. *Final Reports on PHP funded research projects*, 1-75.

Bakker, E. G., Toomajian, C., Kreitman, M., & Bergelson, J. (2006). A genome-wide survey of R gene polymorphisms in Arabidopsis. *The Plant Cell*, *18*(8), 1803-1818.

Bertin, N., & Heuvelink, E. (1993). Dry-matter production in a tomato crop: comparison of two simulation models. *Journal of Horticultural Science*, *68*(6), 995-1011.

Bittner-Eddy, P. D., & Beynon, J. L. (2001). The Arabidopsis downy mildew resistance gene, RPP13-Nd, functions independently of NDR1 and EDS1 and does not require the accumulation of salicylic acid. *Molecular plant-microbe interactions*, *14*(3), 416-421.

Bittner-Eddy, P. D., Crute, I. R., Holub, E. B., & Beynon, J. L. (2000). RPP13 is a simple locus in Arabidopsis thaliana for alleles that specify downy mildew resistance to different avirulence determinants in Peronospora parasitica. *The Plant Journal*, *21*(2), 177-188.

Burt, S. (2004). Essential oils: their antibacterial properties and potential applications in foods—a review. *International journal of food microbiology*, *94*(3), 223-253.

Cannon, S. B., Zhu, H., Baumgarten, A. M., Spangler, R., May, G., Cook, D. R., & Young, N. D. (2002). Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. *Journal of Molecular Evolution*, *54*(4), 548-562.

Chaerani, R., & Voorrips, R. E. (2006). Tomato early blight (Alternaria solani): the pathogen, genetics, and breeding for resistance. *Journal of general plant pathology*, *72*(6), 335-347.

Christou, A., Manganaris, G. A., Papadopoulos, I., & Fotopoulos, V. (2013). Hydrogen sulfide induces systemic tolerance to salinity and non-ionic osmotic stress in strawberry plants through modification of reactive species biosynthesis and transcriptional regulation of multiple defence pathways. *Journal of experimental botany*, *64*(7), 1953-1966.

Collard, B. C., & Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1491), 557-572.

Debener, T., Lehnackers, H., Arnold, M., & Dangl, J. L. (1991). Identification and molecular mapping of a single Arabidopsis thaliana locus determining resistance to a phytopathogenic Pseudomonas syringae isolate. *The Plant Journal*, *1*(3), 289-302.

Desta, Z. A., & Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. *Trends in plant science*, *19*(9), 592-601.

Ding, J., Cheng, H., Jin, X., Araki, H., Yang, Y., & Tian, D. (2007). Contrasting patterns of evolution between allelic groups at a single locus in Arabidopsis. *Genetica*, *129*(3), 235-242.

FAOSTAT (2017). "Production – Crops – Area harvested/ Production quantity – Tomatoes – 2014", FAO Statistics online database, Food and Agriculture Organization, Rome, www.fao.org/faostat/en (accessed 22 Sept. 2017).

Flor, H. H. (1955). Host-parasite interactions in flax rust-its genetics and other implications. *Phytopathology*, *45*, 680-685.

Food and Agriculture Organization Statistics, 2017

Foolad, M. R. (2007). Genome mapping and molecular breeding of tomato. *International journal of plant genomics*, *2007*.

Foolad, M. R., Sullenberger, M. T., Ohlson, E. W., & Gugino, B. K. (2014). Response of accessions within tomato wild species, Solanum pimpinellifolium to late blight. *Plant Breeding*, *133*(3), 401-411.

Foolad, M. R., Merk, H. L., & Ashrafi, H. (2008). Genetics, genomics and breeding of late blight and early blight resistance in tomato. *Critical Reviews in Plant Sciences*, *27*(2), 75-107.

Fry, W. E., & Goodwin, S. B. (1997). Resurgence of the Irish potato famine fungus. *Bioscience*, *47*(6), 363-371.

Geisenberg, C., & Stewart, K. (1986). Field crop management. In *The tomato crop* (pp. 511-557). Springer, Dordrecht.

Gisi, U., & Cohen, Y. (1996). Resistance to phenylamide fungicides: a case study with *Phytophthora infestans* involving mating type and race structure. *Annual review of Phytopathology*, *34*(1), 549-572.

Hunter, M. C., Smith, R. G., Schipanski, M. E., Atwood, L. W., & Mortensen, D. A. (2017). Agriculture in 2050: recalibrating targets for sustainable intensification. *Bioscience*, *67*(4), 386-391.

Isman, M. B. (2000). Plant essential oils for pest and disease management. *Crop protection*, *19*(8-10), 603-608.

Jacob, F., Vernaldi, S., & Maekawa, T. (2013). Evolution and conservation of plant NLR functions. *Frontiers in immunology*, *4*, 297.

Jo, K. R., Kim, C. J., Kim, S. J., Kim, T. Y., Bergervoet, M., Jongsma, M. A., ... & Vossen, J. H. (2014). Development of late blight resistant potatoes by cisgene stacking. *BMC biotechnology*, *14*(1), 1-10.

Judelson, H. S., & Blanco, F. A. (2005). The spores of Phytophthora: weapons of the plant destroyer. *Nature Reviews Microbiology*, *3*(1), 47-58.

Jupe, F., Pritchard, L., Etherington, G. J., MacKenzie, K., Cock, P. J., Wright, F., ... & Hein, I. (2012). Identification and localisation of the NB-LRR gene family within the potato genome. *BMC genomics*, *13*(1), 1-14.

Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., ... & Hirakawa, M. (2006). From genomics to chemical genomics: new developments in KEGG. *Nucleic acids research*, *34*(suppl_1), D354-D357.

Kim, S., Park, M., Yeom, S. I., Kim, Y. M., Lee, J. M., Lee, H. A., ... & Choi, D. (2014). Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. *Nature genetics*, *46*(3), 270-278.

Klymiuk, V., Yaniv, E., Huang, L., Raats, D., Fatiukha, A., Chen, S., ... & Fahima, T. (2018). Cloning of the wheat Yr15 resistance gene sheds light on the plant tandem kinase-pseudokinase family. *Nature communications*, *9*(1), 1-12.

Lidder, P., & Sonnino, A. (2012). Biotechnologies for the management of genetic resources for food and agriculture. *Advances in genetics*, *78*, 1-167.

Matsukura, C., Aoki, K., Fukuda, N., Mizoguchi, T., Asamizu, E., Saito, T., ... & Ezura, H. (2008). Comprehensive resources for tomato functional genomics based on the miniature model tomato Micro-Tom. *Current Genomics*, *9*(7), 436-443.

McHale, L., Tan, X., Koehl, P., & Michelmore, R. W. (2006). Plant NBS-LRR proteins: adaptable guards. *Genome biology*, *7*(4), 1-11.

MNFS&R (Ministry of National Food Security and Research (2017). Fruit, Vegetable & Condiments Statistics of Pakistan: Economic Wing, MNFS&R. Government of Pakistan, Islamabad www.amis.pk/files/F&V%20Statistics%202015-16.pdf

Nowakowska, M., Nowicki, M., Kłosińska, U., Maciorowski, R., & Kozik, E. U. (2014). Appraisal of artificial screening techniques of tomato to accurately reflect field performance of the late blight resistance. *PLoS One*, *9*(10), e109328.

OECD (2008), "Consensus document on compositional considerations for new varieties of tomato: Key food and feed nutrients, toxicants and allergens", Series on the Safety of Novel Foods and Feeds, No. 17, OECD, Paris, www.oecd.org/env/ehs/biotrack/46815296.pdf

Rahman, A., Hallgrímsdóttir, I., Eisen, M., & Pachter, L. (2018). Association mapping from sequencing reads using k-mers. *Elife*, *7*, e32920.

Rose, L. E., Bittner-Eddy, P. D., Langley, C. H., Holub, E. B., Michelmore, R. W., & Beynon, J. L. (2004). The maintenance of extreme amino acid diversity at the disease resistance gene, RPP13, in Arabidopsis thaliana. *Genetics*, *166*(3), 1517-1527.

Rubin, E., & Cohen, Y. (2004). Oospores associated with tomato seed may lead to seedborne transmission of *Phytophthora infestans*. *Phytoparasitica*, *32*(3), 237.

Rubin, E., Baider, A., & Cohen, Y. (2001). *Phytophthora infestans* produces oospores in fruits and seeds of tomato. *Phytopathology*, *91*(11), 1074-1080.

Saleem, M. Y., Akhtar, K. P., Iqbal, Q., Asghar, M., & Shoaib, M. (2015). Development of high yielding and blight resistant hybrids of tomato. *Pakistan Journal of Agricultural Sciences*, *52*(2).

Sarkar, D., Maji, R. K., Dey, S., Sarkar, A., Ghosh, Z., & Kundu, P. (2017). Integrated miRNA and mRNA expression profiling reveals the response regulators of a

susceptible tomato cultivar to early blight disease. *DNA Research*, *24*(3), 235-250.

Särkinen, T., Bohs, L., Olmstead, R. G., & Knapp, S. (2013). A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. *BMC evolutionary biology*, *13*(1), 1-15.

Shi, J., Gao, H., Wang, H., Lafitte, H. R., Archibald, R. L., Yang, M., ... & Habben, J. E. (2017). ARGOS 8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant biotechnology journal*, *15*(2), 207-216.

Spoel, S. H., & Dong, X. (2012). How do plants achieve immunity? Defence without specialized immune cells. *Nature reviews immunology*, *12*(2), 89-100.

Stacey, D. L. (1983). The effect of artificial defoliation on the yield of tomato plants and its relevance to pest damage. *Journal of Horticultural Science*, *58*(1), 117-120.

Stone, J. K., Bacon, C. W., & White Jr, J. F. (2000). An overview of endophytic microbes: endophytism defined. *Microbial endophytes*, 17-44.

Tomato Genome Consortium, X. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, *485*(7400), 635.

van Ooijen, G., van den Burg, H. A., Cornelissen, B. J., & Takken, F. L. (2007). Structure and function of resistance proteins in solanaceous plants. *Annu. Rev. Phytopathol.*, *45*, 43-72.

Vaughan, J. G., & Geissler, C. A. (1997). Vegetable fruits. *The New Oxford Book of Food Plants. Oxford University Press, Oxford, New York*, 122-137.

Victoria, N. G., van der Valk, O. M. C., & Elings, A. (2011). *Mexican Protected Horticulture: Production and market of Mexican protected horticulture*

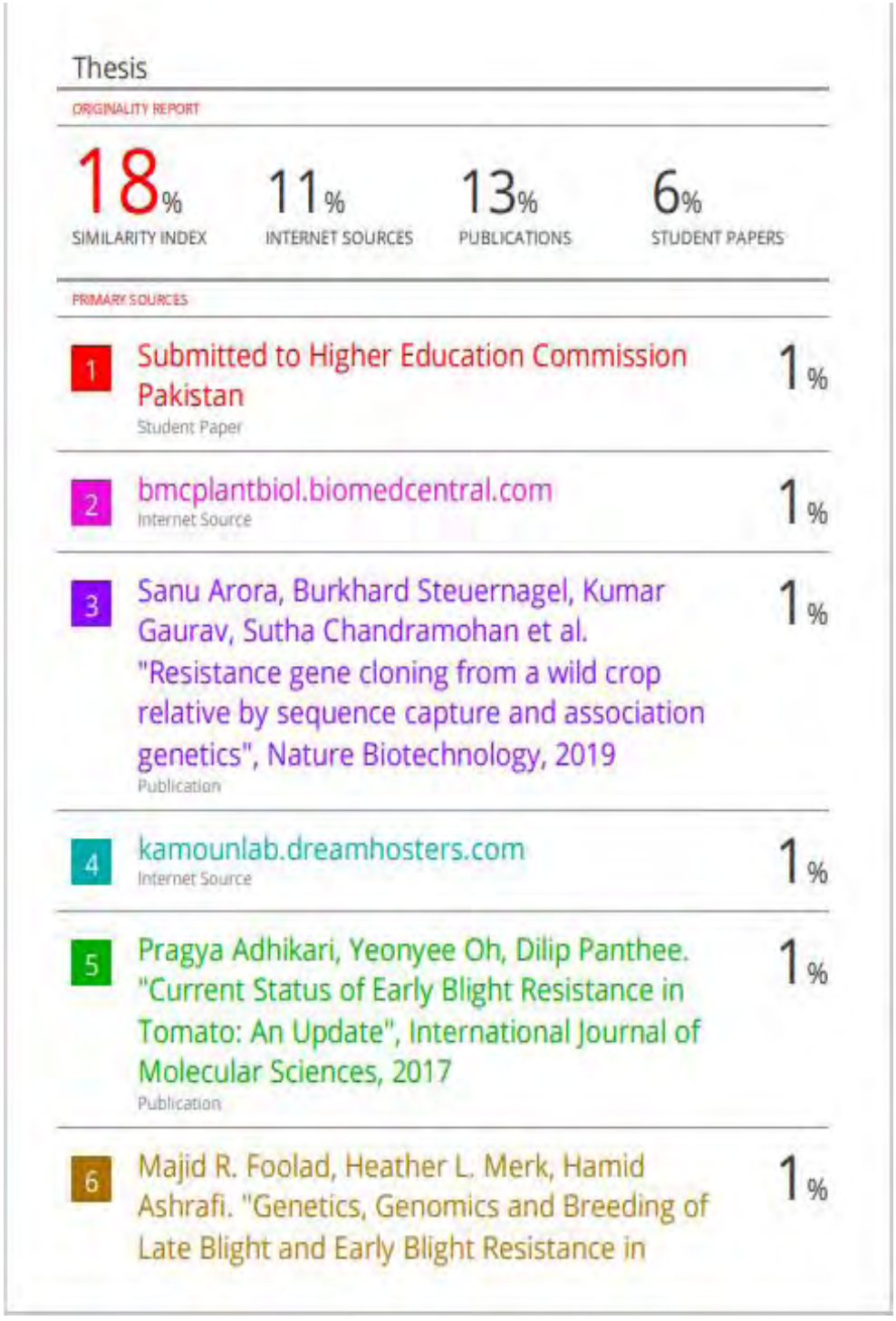
described and analysed (No. 1126). Wageningen UR Greenhouse Horticulture/LEI.

Vleeshouwers, V. G., Raffaele, S., Vossen, J. H., Champouret, N., Oliva, R., Segretin, M. E., ... & Kamoun, S. (2011). Understanding and exploiting late blight resistance in the age of effectors. *Annual review of phytopathology*, *49*, 507-531.

Wu, C. H., Abd-El-Haliem, A., Bozkurt, T. O., Belhaj, K., Terauchi, R., Vossen, J. H., & Kamoun, S. (2017). NLR network mediates immunity to diverse plant pathogens. *Proceedings of the National Academy of Sciences*, *114*(30), 8113-8118.

Zuluaga, A. P., Vega-Arreguín, J. C., Fei, Z., Matas, A. J., Patev, S., Fry, W. E., & Rose, J. K. (2016). Analysis of the tomato leaf transcriptome during successive hemibiotrophic stages of a compatible interaction with the oomycete pathogen *Phytophthora infestans*. *Molecular plant pathology*, *17*(1), 42-54.

Speed cloning of fungal resistance genes in tomato for network analysis using transcriptomic approaches



Thesis

ORIGINALITY REPORT

18%	11%	13%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Higher Education Commission Pakistan Student Paper	1%
2	bmcplantbiol.biomedcentral.com Internet Source	1%
3	Sanu Arora, Burkhard Steuernagel, Kumar Gaurav, Sutha Chandramohan et al. "Resistance gene cloning from a wild crop relative by sequence capture and association genetics", Nature Biotechnology, 2019 Publication	1%
4	kamounlab.dreamhosters.com Internet Source	1%
5	Pragya Adhikari, Yeonyee Oh, Dilip Panthee. "Current Status of Early Blight Resistance in Tomato: An Update", International Journal of Molecular Sciences, 2017 Publication	1%
6	Majid R. Foolad, Heather L. Merk, Hamid Ashrafi. "Genetics, Genomics and Breeding of Late Blight and Early Blight Resistance in	1%

Supervisors' Signature