OnLine Journal of Biological Sciences

Original Research Paper

Characterization of $Mi_{1.2}$ Whitefly (*Bemisia tabaci*) Resistance Gene

[1]Sherin Amin Mahfouze, [2]Shipra Saxena,
[1]Heba Amin Mahfouze and [2]Manchikatla Venkat Rajam

[1]*Department of Genetics and Cytology, National Research Centre, Genetic Engineering and Biotechnology Division, Dokki, 12622, Egypt*
[2]*Department of Genetics, University of Delhi South Campus, Benito Juarez Road, New-Delhi, India*

Article history
Received: 17-08-2017
Revised: 05-09-2017
Accepted: 9-11-2017

Corresponding Author:
Sherin Amin Mahfouze
National Research Centre, Genetic Engineering and Biotechnology Division, Genetics and Cytology Department, Dokki, 12622, Egypt
E-mail : sherinmahfouze@yahoo.com

Abstract: Tomato (*Solanum peruvianum*) *Mi* gene provides resistance to whitefly (*Bemisia tabaci*), potato aphids and nematode making *Mi* a useful source in integrated pest management program. The aim of this work was to isolate, clone and sequence $Mi_{1.2}$ gene from *S. peruvianum*. In addition, physico-chemical identification of amino acids deduced from $Mi_{1.2}$ gene was done. Secondary (2D) and tertiary (3D) structures of $Mi_{1,2}$ protein were also predicated. Distinct amplicons of 620, 600, 3300 and 1993 bp were successfully amplified using PCR amplification. The full-length DNA (5.4 kbp) and cDNA (4 kbp) of $Mi_{1.2}$ gene was isolated using specific primers. Fragments 620 and 600 bp cloned into *Escherichia coli* XL-1 Blue and sequenced. Sequencing results of both assembled fragments (620 and 600 bp) joined at the overlap region (1440 bp). A BLAST search confirmed that the DNA sequence from the amplified fragments was $Mi_{1.2}$ gene. It shared 98% identity and deduced amino acids shared 97% identity with $Mi_{1.2}$ gene published in GenBank. An Open reading frame (ORF) of $Mi_{1,2}$ protein encoded for 479 amino acid residues with molecular weight 54.59 KDa and isoelectric point (PI) 5.52 was calculated using Expasy's ProtParam server. 2D and 3D structures of $Mi_{1.2}$ protein was analyzed using SOPMA and Swiss-Prot software, respectively.

Keywords: Whitefly, Insect Resistance, $Mi_{1,2}$ Gene, Gene Cloning, Protein Structure

Introduction

Tomato (*Solanum lycopersicum* L.) is an important vegetable crop and it is produced worldwide under both the glasshouse and open field (Kaur *et al*., 2014; McDaniel *et al*., 2016; Alatar *et al*., 2017). Whitefly (*Bemisia tabaci*) of *Aleyrodidae* family (Order: *Hemiptera*), is among the most harmful insects of tomato and causes significant yield loss. It effects on harvest directly by phloem feeding and indirectly imputable to the plant viruses transmission via their saliva such as *Tomato Yellow Leaf Curl Virus* (TYLCV) (Momotaz *et al*., 2010; Chen *et al*., 2015). A family of *Mi* genes arising from wild tomato (*Solanum peruvianum* L.) confers resistance against several of pests such as whiteflies (*Bemisia tabaci*), potato aphids (*Macrosiphum euphorbiae*) and root-knot nematodes (*Meloidogyne* sp.) (Nombela *et al*., 2003; Pallipparambil *et al*., 2014).

The *Mi* genes have three homologues, viz., $Mi_{1.1}$, $Mi_{1.2}$ and $Mi_{1.3}$. Only $Mi_{1.2}$ provides whitefly, aphid and nematode resistance in tomato (Seah *et al*., 2004). $Mi_{1.2}$ gene produces a transcript of approximately 4 kbp that encodes a putative protein of 1,257 amino acid residues (Rossi *et al*., 1998). The protein is identified by the presence of a nucleotide binding site (NBS) and a leucine-rich repeat motif (LRR) (Milligan *et al*., 1998). Proteins of the NBS-LRR motif structure composed of the largest category of cloned plant resistance genes against viruses, fungi, bacteria, insects and nematodes (Dangl and Jones, 2001). $Mi_{1.2}$ is a potential gene in tomato integrated pest management program (Nombela *et al*., 2003; Mahfouze *et al*., 2015).

The aim of this work was to isolate, clone and sequence $Mi_{1.2}$ gene from *S. peruvianum*. In addition, physico-chemical identification of amino acids deduced from $Mi_{1.2}$ gene was done. Secondary (2D) and tertiary (3D) structures of $Mi_{1.2}$ protein were also predicated.

Materials and Methods

Plant Materials

Young leaves of tomato (*Solanum peruvianum)* plants were obtained from Indian Agricultural Research Institute (IARI), New Delhi, India. All the collected plant material were kept at -80°C for storage.

Design of Oligonucleotide Primers for $Mi_{1.2}$ Gene

A total of four primers with different degrees of specificity were designed according to the public sequence of $Mi_{1.2}$ gene (GenBank accession number AF039682.1) using SMS Sequence Manipulation Suite (http://bioinformatics.org/sms2/index.html).TomMi3 and TomMi4 primers containing chimeric regions complementary to one another. These chimeric overlapping sequences, which amplified $Mi_{1.2}$ gene were mixed and annealed at the overlaps. Primers used in this study were designed with various factors in consideration: GC content, melting temperature for primer set, formation of hairpin loops and dimerization of oligos. Various oligonucleotides used in this study are listed in Table 1. Primers were synthesized by Sigma Aldrich Chemicals, Bangalore, India.

Extraction of Genomic DNA and PCR Amplification of $Mi_{1.2}$ Gene

DNA was extracted from fresh *S. peruvianum* tomato leaves using CTAB method (Fulton *et al*., 1995). PCR reactions contained sterile distilled water 37.25, 5.0 µL 10×PCR buffer (Promega Corp.), 1.5 µL $MgCl_2$ (50 mM), 1.0 µL dNTPs mix (10 mM), the two primer combinations 3.0 µL (1.5 µL each = 150 ng), 0.25 µL *Taq* DNA polymerase, recombinant (Biotools, Spain) (5 Units/µL). 2.0 µL of DNA template (~400 ng) was added to the reaction. PCR cycles were 94°C for 5 min; 35 cycles of: 94°C for 30s, Annealing temperature (AT°C) as shown in Table 1 for 1 min, 72°C for 3 min; 72°C for 10 min. PCR reactions were performed in Bio-Rad, C1000 thermal cycler, California, USA.

Isolation of Full Length DNA $Mi_{1.2}$ Gene

PCR reactions contained sterile distilled water 36.8, 5.0 µL 10×PCR buffer B (Promega Corp.), 5 µL dNTPs (10 mM each), the two primer combinations (F.P. TomMi3and R.P TomMi4) (25 mM) (1.0 µL each), 0.2 µL Takara *Taq* DNA polymerase (5 Units/µL). 1 µL of DNA template (~500 ng) was added to the reaction. PCR was performed as follows: 94°C for 1 min; then 94°C for 15s, 60°C for 30s, 72°C for 6 min, 40 cycles; followed by 72°C for 10 min and held at 4°C.

Reverse Transcriptase (RT-PCR) and Isolation of Full Length of cDNA

Total RNA was extracted from tomato leaves using Trizol reagent according to the manufacture's recommended protocol (Sigma, India). cDNA was synthesized from total RNA (1 µg/sample) using Reverse transcriptase kit and oligo-DT primers (Thermo Scientific, UK) in 10 µL reaction at room temperature for 15 min and then heated at 65°C for 10 min to inactivate DNaseI. The DNaseI-treated total RNA was then reverse-transcribed using the RT-PCR kit (Thermo Scientific, Uk). $Mi_{1.2}$ cDNA was PCR amplified using F.P. TomMi3and R.P TomMi4 primers. The β-actin gene-specific primers (Table 1) were added to the same RT-PCR reactions as internal standards for RNA quantity. PCR reactions of $Mi_{1.2}$ amplification contained sterile distilled water 36.4, 10 µL 10× PCR buffer HF (Promega Corp.), 1.0 µL dNTPs (10 mM), the two primer combinations (F.P. TomMi3and R.P TomMi4) (0.3 µL each) (0.5 µM), 0.5 µL Phusion *Taq* DNA polymerase (Thermo Secientific, UK) (0.02 Units/µL). 1.5 µL of DNA template (~400 ng) was added to the reaction. PCR cycles were 98°C for 3 min; 35 cycles of: 98°C for 30s, 64.3°C for 30s, 72°C for 3 min; 72°C for 10 min. Similar PCR conditions were used for the β actin primers, with the exception that the annealing temperature was 55°C.

PCR products for all samples was electrophoresed on 0.8% agarose containing ethidium bromide (0.5 µg mL^{-1}) in 1x TBE buffer (89 mM Tris-HCl, 89 mM Boric acid, 2.5 mM EDTA, pH 8.3) at 75 constant volt and determined with UV transilluminator. The size of each fragment was determined with reference to a size marker of 1 kbp DNA ladder (Thermo Scientific, UK).

Cloning of the $Mi_{1.2}$ Gene

Fragments of the expected size, 620 and 600 bp for $Mi_{1.2}$ gene was excised from the agarose gels and further purified using Gene Jet Gel DNA purification kit (Thermo Scientific, UK). The quality and concentration of the purified products was confirmed by gel-electrophoresis in a 0.8% agarose gel in 1xTBE buffer and by measuring the absorbance ratio at 260 nm wavelength using a NanoDrop ND-1000 spectrophotometer. The purified PCR products were ligated into pGEM®-TEasy vector (Promega, Mannheim, Germany). Ligation reactions were prepared containing the appropriate quantities of vector and insert (1:3), 1 µL of the 2 × ligase buffer and 2.5 U T4 DNA ligase supplied with the kit. The reaction volume was made up to 10 µL with sterile dsH_2O and the reactions were incubated at 4°C overnight. Ligated plasmids were transformed into *E. coli* XL-1 Blue competent cells. Isolation of plasmid DNA from *E. coli* XL-1 Blue was done by the alkaline lysis method according to Sambrook *et al*. (1989).

Digestion of Plasmid DNA with Restriction Enzyame EcoR1

To confirm the presence of positive intact clones, restriction enzyme digestion of plasmid DNA was also carried out with *EcoR1* at 37°C overnight.

Table 1: Primers used in this study.

Primers	Single nucleotide sequence (5'-3')	GC (%)	Annealing temperature (AT)°C	Molecular size (bp)	Nucleotide position
IMOF1	AGCCATGCTTGCTTCACTTT	45.0	55	998	2395-2414
IMOR1	AGAGGACCCACAGTGGTTTG	55.0			3373-3392
F.P. TomMi 1	TGAAAG CCC CAAATT CAT CT	40.0	60	620	1201-1220
R.P. TomMi 1	CCATGC ACGAAG GTC AAAAT	45.0			1801-1820
F.P. TomMi 2	ATT TTG ACC TTC GTG CATGG	45.0	60	600	1801-1820
R.P. TomMi 2	ATGGCTTGAGGTGATGTGGt	50.0			2381-2400
F.P. TomMi3	GGATCCAATAGCTTCAACATTATT	33.3	58	3300	4-21
R.P. TomMi3	CAATAAGATACCTCTTTCCAAACAGTTGTTTCCGC	40.0			1931-1965
F.P. TomMi4	GCGGAAACAACTGTTTGGAAAGAGGTATCTTATT	40.0	58	1993	1931-1964
R.P. TomMi4	TCTAGAGGAATCTCATCACAGGA	43.4			3907-3923
β-Actin F	ACAATGAGCTCCGAGTTGCT	45.0	55	900	-
β-Actin R	TTGATCTTCATGCTGCTTGG	50.0			-

Sequencing

Partial nucleotide sequence of *$Mi_{1.2}$* gene was done by Applied Biosystems (Inst model/Name 3100/3130XL-1468-009, India using gene-specific primers. The sequence was aligned with corresponding sequences from the database using BLAST from the website http://www.ncbi.nlm.nih.gov/blast. Multiple alignments and phylogenetic tree of protein were performed using CLC Main Workbench 5 program, Denamark.

2D and 3D Structures of $Mi_{1.2}$ Protein

The primary amino acid sequence of $Mi_{1.2}$ protein was subjected to predict its secondary and tertiary structures using SPOMA (Geourjon and Deleage, 1995) and a SWISS-MODEL workspace servers (http://swissmodel.expasy.org/workspace), respectively (Arnold *et al*., 2006).

Results

Amplification of $Mi_{1.2}$ Gene

PCR was used for the amplification of *$Mi_{1.2}$* gene by using different specific primers. These primers designed from conserved region of *$Mi_{1.2}$* gene available in the GenBank (AF039682.1). The expected sizes of amplicons were 998 (Bendezu, 2004), 620, 600, 3300 and 1993 bp of FIMO/RIMO, F.P. TomMi1/R.P. TomMi1, F.P. TomMi2/R.P. TomMi2, F.P. TomMi3/R.P. TomMi3 and F.P. TomMi4/R.P. TomMi4 primers, respectively as shown in Fig. 1.

Isolation of Full Length Genomic DNA and cDNA of Target Gene

The size of *$Mi_{1.2}$* gene using the specific primers the F.P. TomMi3 and R.P. TomMi4 was 5.4 kb (Fig. 2). In the present study, the total RNA was isolated using TRIzol reagent. The RNA profile on 0.8% agarose gel indicated the intactness of different subunits of RNA. cDNA was synthesized by RT-PCR using oligodT as 5' and 3' primers. The cDNA product of around 4 kb is shown in Fig. 3.

Cloning of $Mi_{1.2}$ Gene Fragments

The fragments of the expected sizes 620 and 600 bp were excised from agarose gels and the DNA products were cleaned up by a Gene Jet Gel DNA purification Kit (Thermal Scientific, UK). The quality and concentration of the purified products was confirmed by gel-electrophoresis in a 0.8% agarose gel by measuring the absorbance. The purified PCR products were ligated into pGEM®-T-Easy vector (Promega, Mannheim, Germany). To screen positive colonies, four or five white colonies were picked from pGEM: $Mi_{1,2}$ construct, Mini-preparations were performed with all colonies. To determine the insert orientation within pGEM-T-Easy vector was performed by digestion of *EcoR1* (Fig. 4 and 5) and sequencing. Fig. 4 and 5 indicated that the transformation with pGEM: *$Mi_{1.2}$* was successful and plasmid with correct insert orientation.

Multiple Sequence Alignments and Phylogenetic Tree

Sequencing results of both assembled fragments (620 and 600 bp) joined at the overlap region (1440 bp) (Fig. 6). The sequence was submitted at the GenBank with the accession number KU886265. BLAST analysis showed that the *$Mi_{1.2}$* gene under study had the identity ranged 98-82% to the root-knot nematode resistance *$Mi_{1.2}$* genes recorded in GenBank (Table 2). On the other hand, the deduced amino acids sequence of $Mi_{1.2}$ protein gave the homology 97-61% (Table 3). The amino acids sequence of $Mi_{1.2}$ protein was aligned with six different accessions of other Mi proteins published in GenBank by CLC Main Workbench 5 program, Denamark (Fig. 7). The phylogenetic tree applied by using CLC Main Workbench 5 program, Denamark with the UPGMA method is presented in Fig. 8. A close relationship was found between our $Mi_{1.2}$ protein and other NBS-LRR proteins (Fig. 8).

Table 2: Sequences alignments of $Mi_{1.2}$ gene under study using BLAST analysis.

Accession No.		Identity
U65668.1	*Lycopersicon esculentum* putative *Mi-1* copy 2 nematode-resistance gene	98%
AF091048.1	*L.esculentum* disease resistance protein (*Mi-1*) gene, complete cds	98%
U81378.1	*L. esculentum* disease resistance gene homolog *Mi*-copy2 gene, complete cds; resistance gene pseudogene, complete sequence; disease resistance	98%
DQ863287.1	*Solanum* sp. VFNT NBS-LRR resistance protein-like protein (*Mi-1.4*) gene, complete cds	98%
NM_001247134.1	*S. lycopersicum* root-knot nematode resistance protein (*Mi-1.2*), mRNA	98%
DQ465824.1	*Capsicum annuum* root-knot nematode resistance protein gene, complete cds	98%
DQ863290.1	*S. lycopersicum* NBS-LRR resistance protein-like protein (*Mi-1C*) gene, complete cds	98%
DQ863293.1	*S. lycopersicum truncated* NBS-LRR resistance protein-like protein (*Mi-1G*) gene, complete cds	97%
U65667.1	*L. esculentum* putative *Mi-1* copy 1 nematode-resistance gene	96%
NM_001247693.1	*S. lycopersicum* plant resistance protein (*Mi-1.1*), mRNA	96%
XM_015303248.1	*S. tuberosum* putative late blight resistance protein homolog R1B-17 (LOC107058010), mRNA	92%
XM_015304640.1	*S. tuberosum putative* late blight resistance protein homolog R1A-3 (LOC102582957), transcript variant X1, mRNA	90%
FJ231739.1	*C. annuum* NBS-LRR root-knot nematode resistance protein mRNA, complete cds	85%
XM_009616696.1	*Nicotiana tomentosiformis* putative late blight resistance protein homolog R1A-3 (LOC104107800), transcript variant X1, mRNA	84%
XM_009772561.1	*N. sylvestris* putative late blight resistance protein homolog R1A-3 (LOC104221492), transcript variant X2, mRNA	82%

Table 3: Sequences alignments of $Mi_{1.2}$ protein under study using BLAST analysis.

Accession No.		Identity
NP_001234063.1	Root-knot nematode resistance protein [*Solanum lycopersicum*]	97%
ABE68835.1	Root-knot nematode resistance protein [*Capsicum annuum*]	96%
ABI96218.1	Truncated NBS-LRR resistance protein-like protein [*S. lycopersicum*]	95%
AAC32253.1	Disease resistance gene homolog Mi-copy1 [*S. lycopersicum*]	94%
XP_015078202.1	Putative late blight resistance protein homolog R1A-3 [*S. pennellii*]	92%
XP_015158734.1	Putative late blight resistance protein homolog R1B-17 [*S. tuberosum*]	88%
XP_004240523.1	Putative late blight resistance protein homolog R1A-3 [*S. lycopersicum*]	87%
XP_015160126.1	Putative late blight resistance protein homolog R1A-3 isoform X1 [*S. tuberosum*]	85%
XP_015160128.1	Putative late blight resistance protein homolog R1A-3 isoform X3 [*S.tuberosum*]	84%
XP_006340022.1	Putative late blight resistance protein homolog R1A-3 [*S. tuberosum*]	82%
ACI43068.1	NBS-LRR root-knot nematode resistance protein [*C.annuum*]	79%
XP_015162455.1	putative late blight resistance protein homolog R1B-8 [*S. tuberosum*]	78%
AJW77761.1	Pvr9-like protein 1 [*C. annuum*]	75%
XP_009608035.1	Putative late blight resistance protein homolog R1A-4 isoform X1 [*Nicotiana tomentosiformis*]	74%
XP_009763450.1	Putative late blight resistance protein homolog R1B-12 isoform X3 [*N. sylvestris*]	70%
XP_004240205.1	Putative late blight resistance protein homolog R1A-3 isoform X2 [*S. lycopersicum*]	61%

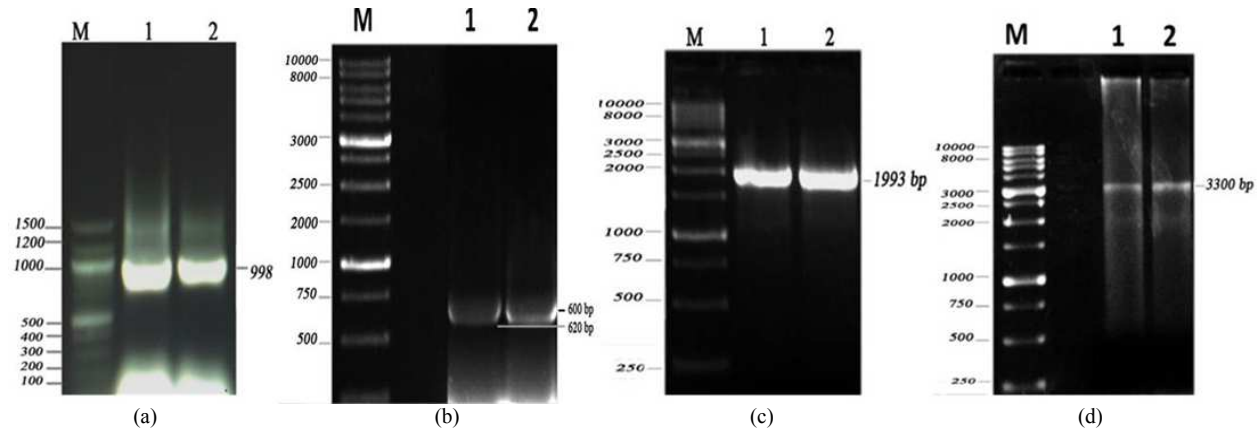


Fig. 1: PCR products of the $Mi_{1.2}$ gene from *S. peruvianum* laves using four specific primers (a) 998 bp, (b) 620 bp,(c) 1993 bp and (d) 3300 bp. Lane M: 1 kb DNA ladder , lanes 1 and 2: PCR amplicons of $Mi_{1.2}$ gene. The arrow indicates amplification of the $Mi_{1.2}$ gene.

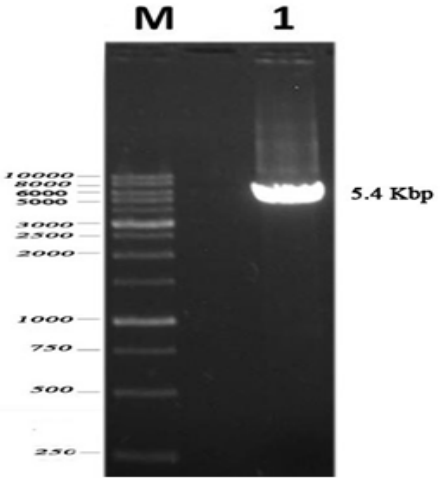


Fig. 2: PCR amplification of the $Mi_{1.2}$ gene from *S. peruvianum* leaves using F.P.TomMi3and R.P. TomMi4 primers. Lane M: 1 kb DNA size marker lane 1: PCR product of $Mi_{1.2}$ gene.

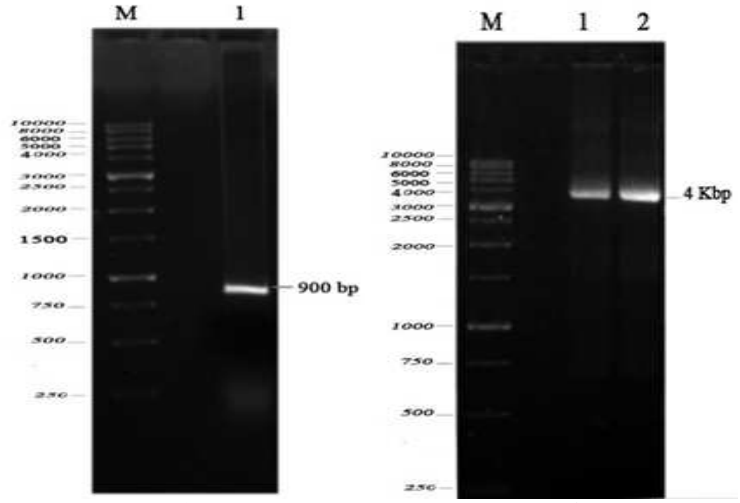


Fig. 3: Electrophoresis of cDNA on 0.8% agarose gel using F.P.TomMi3 and R.P.TomMi4 primers. Lane M: DNA size marker 1 Kbp. (a) lane 1: ampilcon size of β-actin cDNA. (b) lanes 1 and 2: size of PCR product of $Mi_{1.2}$ cDNA.

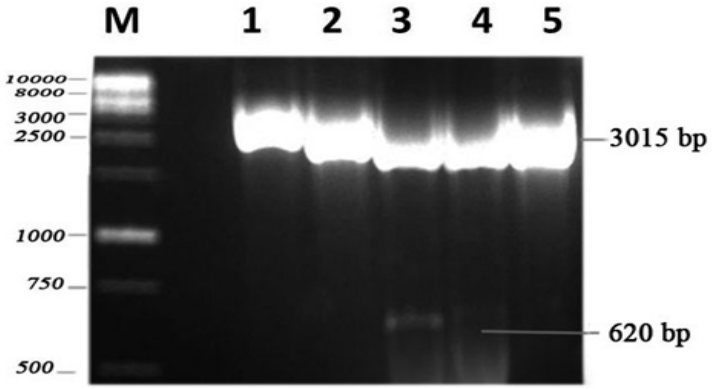


Fig. 4: *EcoR1* restriction digests of pGEMT plasmid (3015 bp) contained the insert (620 bp). Lanes 1,2 and 5 undigested clones. Lanes 3 and 4 featured the correct insert orientation digested clones.

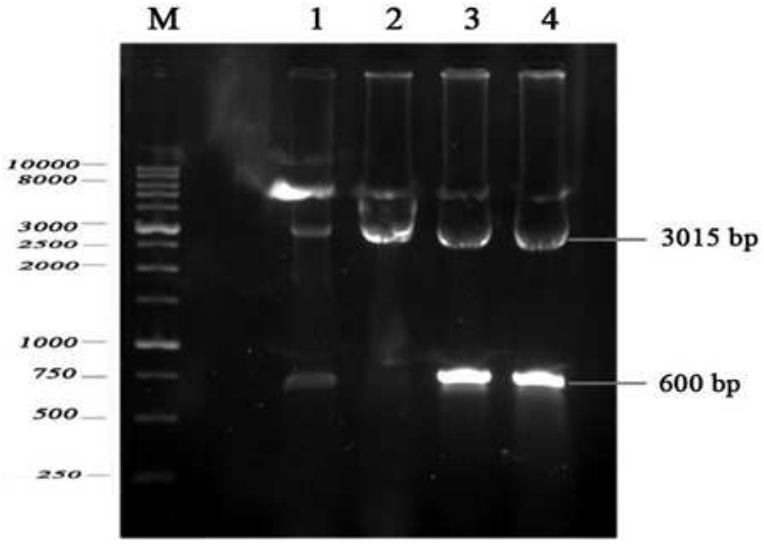


Fig. 5: *EcoR1* restriction digests of pGEMT plasmid (3015 bp) contained the insert (600 bp). Lane 2 undigested clone. Lanes 1, 3 and 4 featured the correct insert orientation digested clones.

```
1   V  M  M  E  F  L  L  L  I  L  S  D  M  P  K  D  F  I  H  H
    1 ATGTCATGATGGAATTCCTATTACTTATTCTTTCTGATATGCCCAAGGACTTTATTCATC
   21   D  K  L  F  D  L  L  A  H  V  G  T  L  T  R  E  V  S  T  L
   61 ATGACAAACTTTTTGATCTCTTGGCTCATGTTGGAACACTTACCAGGGAGGTATCGACTC
   41   V  R  D  L  E  E  K  L  R  N  K  E  G  N  N  Q  T  N  C  A
  121 TTGTACGTGACTTGGAAGAGAAATTAAGGAATAAAGAGGGTAATAACCAAACAAATTGTG
   61   T  L  D  L  L  E  N  I  E  L  L  K  K  D  L  K  H  V  Y  L
  181 CAACCCTAGACTTGCTGGAAAATATTGAACTCCTCAAGAAAGATCTCAAACATGTTTATC
   81   K  A  P  N  S  S  Q  C  C  F  P  M  S  D  G  P  L  F  M  H
  241 TGAAAGCCCCAAATTCATCTCAATGTTGCTTCCCCATGAGTGATGGACCACTCTTCATGC
  101   L  L  H  M  H  L  N  D  L  L  D  S  N  A  Y  S  I  S  L  I
  301 ATCTTCTACACATGCACTTAAATGATTTGCTAGATTCTAATGCTTATTCAATTTCTTTGA
  121   K  E  E  I  E  L  V  S  Q  E  L  E  F  I  R  S  F  F  G  D
  361 TAAAGGAAGAAATCGAGTTGGTGAGTCAAGAACTGGAATTCATAAGATCATTCTTTGGGG
  141   A  A  E  Q  G  L  Y  K  D  I  W  A  R  V  L  D  V  A  Y  E
  421 ATGCTGCTGAGCAAGGATTGTATAAAGATATCTGGGCACGTGTTCTAGATGTGGCTTATG
  161   A  K  D  V  I  D  S  I  I  V  R  D  N  G  L  L  H  L  I  F
  481 AGGCAAAAGATGTCATAGATTCAATTATTGTTCGAGATAATGGTCTCTTACATCTTATTT
  181   S  L  P  I  T  I  K  K  I  K  L  I  K  E  E  I  S  A  L  D
  541 TCTCACTTCCCATTACCATAAAGAAGATCAAACTTATCAAAGAAGAGATCTCTGCTTTAG
  201   E  N  I  P  K  D  R  G  L  I  V  V  N  S  P  K  K  P  V  E
  601 ATGAGAACATTCCCAAGGACAGAGGTCTAATCGTTGTGAACTCTCCCAAGAAACCAGTTG
  221   R  K  S  L  T  T  D  K  I  I  V  G  F  E  E  E  T  N  L  I
  661 AGAGAAAGTCATTGACAACTGATAAAATAATTGTAGGTTTTGAGGAGGAGACAAACTTGA
  241   L  R  K  L  T  S  G  P  A  D  L  D  V  I  S  I  T  G  M  P
  721 TACTTAGAAAGCTCACCAGTGGACCCGCAGATTTAGATGTCATTTCGATCACCGGTATGC
  261   G  S  G  K  T  T  L  A  Y  K  V  Y  N  D  K  S  V  S  R  H
  781 CGGGTTCAGGTAAAACTACTTTGGCATACAAAGTATACAATGATAAGTCAGTTTCTAGAC
  281   F  D  L  R  A  W  C  T  V  D  Q  G  Y  D  E  K  K  L  L  N
  841 ATTTTGACCTTCGTGCATGGTGCACGGTCGACCAAGGATATGATGAGAAGAAGTTGTTGA
  301   K  I  F  N  Q  V  S  D  S  D  S  K  L  S  E  N  I  D  V  P
  901 ATAAAATTTTCAATCAAGTTAGTGACTCAGATTCAAAATTGAGTGAGAATATTGATGTTC
  321   D  K  L  R  K  Q  L  Y  G  K  R  Y  L  I  V  L  D  D  V  W
  961 CTGATAAGCTACGGAAACAACTGTATGGAAAGAGGTATCTTATTGTCTTAGATGACGTGT
  341   E  T  T  T  W  D  E  L  T  R  P  F  P  K  A  K  K  G  S  R
 1021 GGGAGACTACTACATGGGATGAGTTGACAAGACCTTTTCCTAAAGCTAAGAAAGGAAGTA
  361   I  I  L  T  T  R  E  K  E  V  A  L  H  G  K  L  Y  T  D  P
 1081 GAATTATTTTGACAACTCGAGAAAAGGAAGTGGCTTTGCATGGAAAGCTCTACACTGATC
  381   L  D  L  R  L  L  R  P  D  E  S  W  E  L  L  E  K  R  A  F
 1141 CTCTTGACCTTCGATTGCTAAGACCAGATGAAAGTTGGGAATTATTAGAGAAAAGGGCAT
  401   G  N  E  S  C  P  D  E  L  L  D  V  G  K  E  I  A  E  N  C
 1201 TTGGAAATGAGAGTTGCCCTGATGAACTATTAGATGTCGGTAAAGAAATAGCCGAAAATT
  421   K  G  L  P  L  V  A  D  L  I  A  G  V  I  A  G  R  E  K  K
 1261 GTAAAGGGCTTCCTTTGGTGGCTGATCTGATTGCTGGAGTCATTGCTGGGAGGGAAAAGA
  441   R  S  V  W  L  E  V  Q  S  S  L  S  S  F  I  L  N  S  E  V
 1321 AAAGGAGTGTGTGGCTTGAAGTTCAAAGTAGTTTGAGTTCTTTTATTTTGAACAGTGAAG
  461   E  V  M  K  V  V  E  L  S  Y  D  H  L  P  H  H  L  K  P
 1381 TGGAAGTGATGAAAGTTGTAGAATTAAGTTATGACCATTTACCACATCACCTCAAGCCAT
```

Fig. 6: Partial nucleotide sequences of *Mi*$_{1.2}$ gene (1440 bp) and primarily structure of Mi$_{1.2}$ protein under study (479 amino acids residues, translated frame-3).

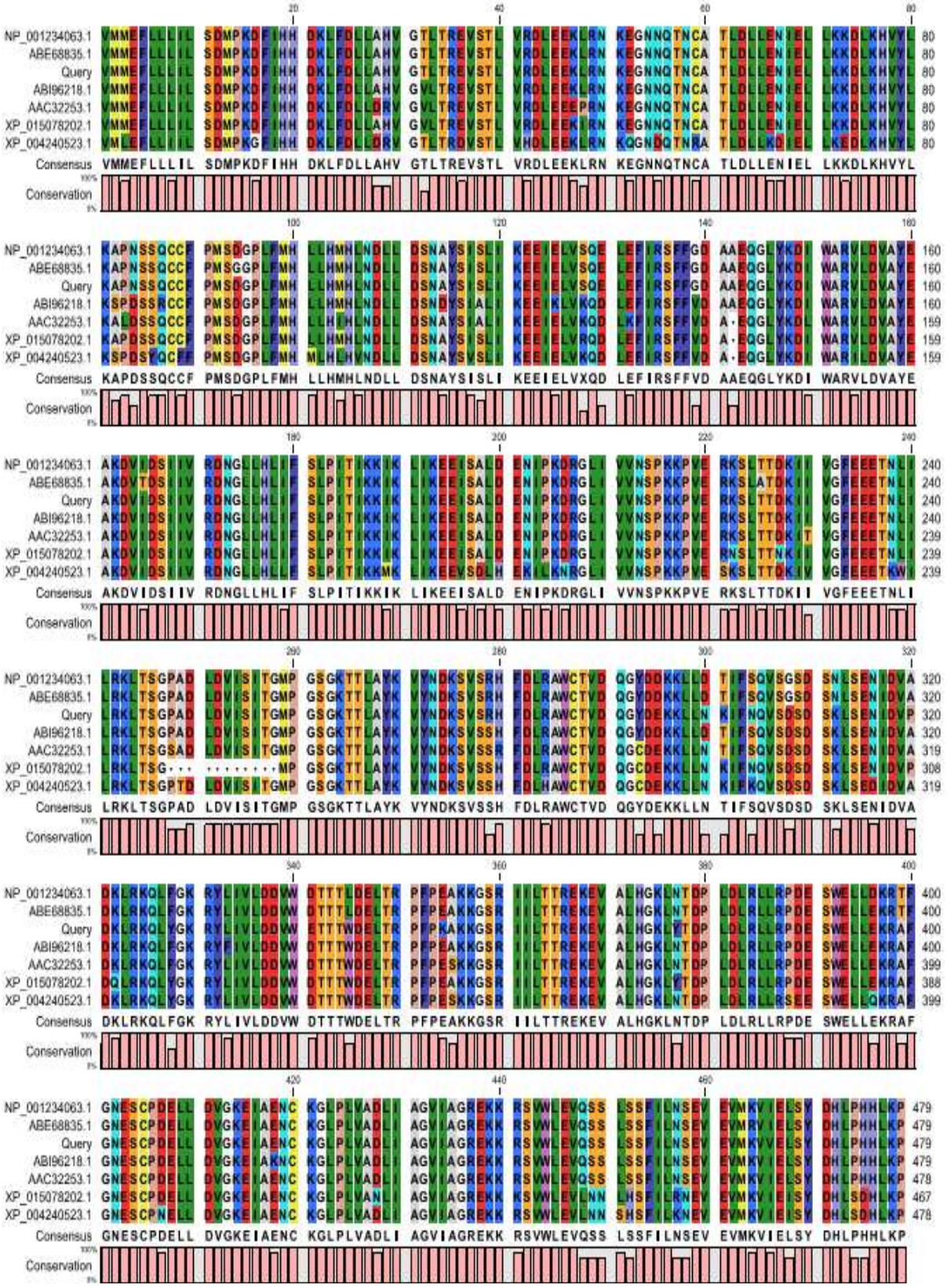


Fig. 7: Alignment of $Mi_{1,2}$ protein from this study with the reference $Mi_{1,2}$ proteins recoded in GenBank.

Amino Acids of $Mi_{1.2}$

The Coding DNA Sequence (CDS) of $Mi_{1.2}$ protein encoded for 479 amino acid residues contained on one stop codon (translate frame-3) (Fig. 6). The molecular weight was 54.59 KDa with isoelectric point (PI) 5.52 and extinction coefficient 49765 $M^{-1}cm^{-1}$ at 280 nm measured in water. $Mi_{1.2}$ protein consisted of 20 amino acids; 177 aliphatic (36.95%); 32 aromatic (6.7%); 14 sulphur (2.9%); 77 basic (16.1%); 105 acidic (22%); 55 aliphatic hydroxyl (11.5%) amino acids. In addition, 235 tRNAsynthetase class I (49.1%) and 244 tRNAsynthetase class II (51%). Furthermore, it contained on 78 negatively charged residues (Asp + Glu) and 64 positively charged residues (Arg + Lys). Besides, Leucine (L) was the main amino acid in sequence (71, 14.8% frequencies); followed by Aspartic Acid (D) and Glutamic Acid (E) (39, 8.1% frequencies). On the contrary, Cysteine (C) and Tryptophan (W) were the lowest amino acids in residue (6, 1.8% frequencies) (Table 4).

Secondary Structure (2D) Model of $Mi_{1.2}$ Protein

Secondary structure predication of $Mi_{1.2}$ protein by using SPOMA server (Geourjon and Deleage, 1995) showed that $Mi_{1.2}$ composed of 232 α-helix (48.4%), 77 β-sheets (16.1%), 39 Beta-turn (8.14%) and 131 coil (27.4%) (Fig. 9).

3D Structure Modeling of $Mi_{1.2}$ Protein

The three-dimensional (3D) structure protein was carried out by using SWISS-MODEL workspace server. The predicated structure of $Mi_{1.2}$ under study was similar to *Solanum lycaopersicum* root-knot nematode resistance protein (NP_001234063.1). Both had monomers, consisted of α-helix and β-sheets with a compact structure, as shown in Fig. 10.

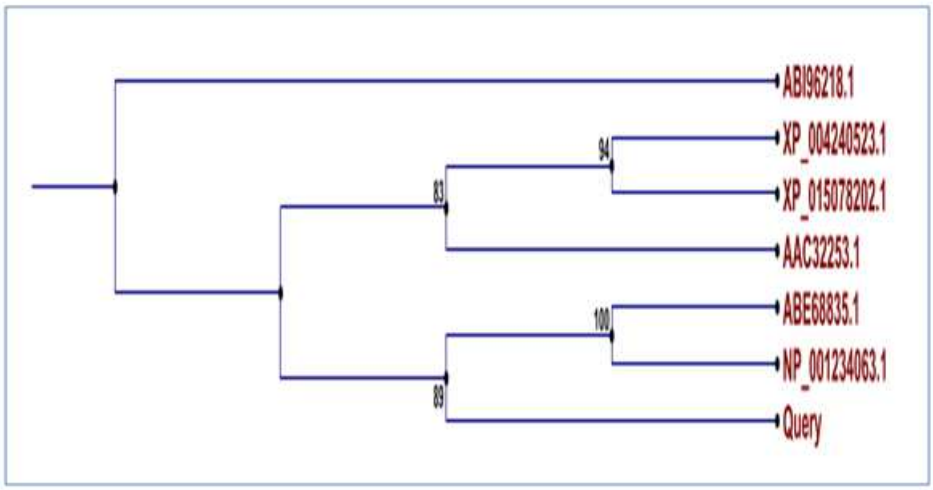
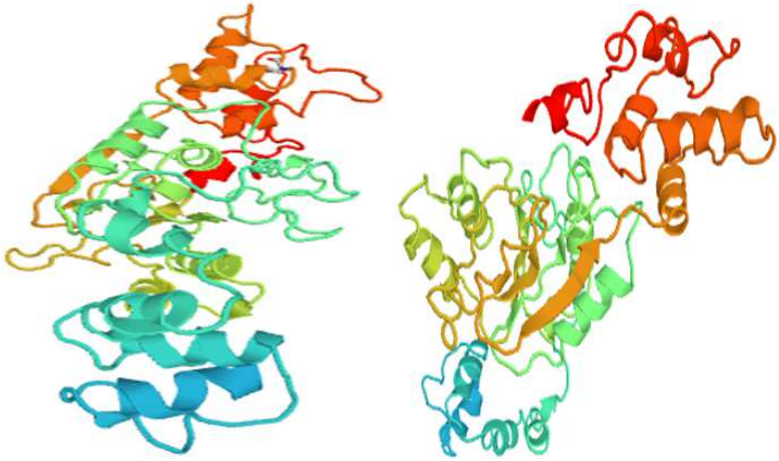


Fig. 8: Phylogenetic tree of $Mi_{1.2}$ protein (Query) using UPGMA with six disease resistance proteins class NBS-LRR in GenBank.

```
         10        20        30        40        50        60        70
          |         |         |         |         |         |         |
VMMEFLLLILSDMPKDFIHHDKLFDLLAHVGTLTREVSTLVRDLEEKLRNKEGNNQTNCATLDLLENIEL
hhhhhhhhhhttccttheehhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhcttccccccchhhhhhhhhhhh

LKKDLKHVYLKAPNSSQCCFPMSDGPLFMHLLHMHLNDLLDSNAYSISLIKEEIELVSQELEFIRSFFGD
hhhhhhheeeeccccccccccccccceeeehhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh

AAEQGLYKDIWARVLDVAYEAKDVIDSIIVRDNGLLHLIFSLPITIKKIKLIKEEISALDENIPKDRGLI
hhhtthhhhhhhhhhhhhhttthhhheeeecttteeeeeeccccchhhhhhhhhhhhhhhttcctttteee

VVNSPKKPVERKSLTTDKIIVGFEEETNLILRKLTSGPADLDVISITGMPGSGKTTLAYKVYNDKSVSRH
eeecccccccccccccceeeeeectthhheehhhcccccceeeeeeecccttcceeeeeeeectteccccc

FDLRAWCTVDQGYDEKKLLNKIFNQVSDSDSKLSENIDVPDKLRKQLYGKRYLIVLDDVWETTTWDELTR
eeehheeeeccccchhhhhhhhhhhhhcccccccccccccchhhhhhhcttteeeeeehhhhhhccchhhhcc

PFPKAKKGSRIILTTREKEVALHGKLYTDPLDLRLLRPDESWELLEKRAFGNESCPDELLDVGKEIAENC
ccccccttceeeeeeecttthheettcccccccchheecttchhhhhhhhhcccccccchhhhhhhhhhhhhc

KGLPLVADLIAGVIAGREKKRSVWLEVQSSLSSFILNSEVEVMKVVELSYDHLPHHLKP
ccccchhhhhhhhhhtccccceeeeehhcthhheehhhhhhhhhhhhhccttcchhhct
```

Fig. 9: Secondary structure protein of $Mi_{1.2}$ protein under study using SPOMA server. H = α-helix, e = β-sheets, t = Beta-turn, c = coil.



(A) (B)

Fig. 10: Predicted 3D structure model of $Mi_{1.2}$ protein under study (A) and compared with *S. lycopersicum* root-knot nematode resistance protein (NP_001234063.1)(B) using a Swiss-model server.

Table 4: Amino acids counts and % frequencies of $Mi_{1.2}$ protein under study.

No.	Amino acid	Count	%Frequency	No.	Amino acid	Count	% Frequency
1	Alanine (A)	20	4.2	15	Proline (P)	19	4.0
2	Arginine (R)	21	4.4	16	Serine (S)	34	7.1
3	Asparagine (N)	19	4.0	17	Threonine (T)	21	4.4
4	Aspartic Acid (D)	39	8.1	18	Tryptophan (W)	6	1.3
5	Cysteine (C)	6	1.3	19	Tyrosine (Y)	11	2.3
6	Glutamine (Q)	8	1.7	20	Valine (V)	32	6.7
7	Glutamic Acid (E)	39	8.1	21	Aliphatic (G,A,V,L,I)	177	37.0
8	Glycine (G)	21	4.4	22	Aromatic (F,W,Y)	32	6.7
9	Histidine (H)	13	2.7	23	Sulphur (C,M)	14	2.9
10	Isoleucine (I)	33	6.9	24	Basic (K,R,H)	77	16.1
11	Leucine (L)	71	14.8	25	Acidic (D,E,N,Q)	105	22.0
12	Lysine (K)	43	9.0	26	Aliphatic hydroxyl (S,T)	55	11.5
13	Methionine (M)	8	1.7	27	tRNAsynthetase class I (E, Q,R,C,M,V,I,L,Y,W)	235	49.1
14	Phenylalanine (F)	15	3.1	28	tRNAsynthetase class II (G, A,P,S,T,H,D,N,K,F)	244	51.0

Total number of negatively charged residues (Aspartic Acid+ Glutamic Acid) = 78
Total number of positively charged residues (Arginine+Lysine) = 64

Discussion

The whitefly is an important insect pest of many crop plants, including tomato. Many wild tomato species contains *Mi* gene, which provides resistance to whitefly (McDaniel *et al*., 2016). The wild tomato (*S. peruvianum*) resistance gene *Mi* encodes a protein with CC-NBS-LRR motifs (Milligan *et al*., 1998). $Mi_{1.2}$ is a single dominant gene in tomato, which provides resistance against certain phloem-feeding herbivores such as whiteflies, aphids, psyllids and root-knot nematodes (Nombela *et al*., 2003; Pallipparambil *et al*., 2014; Chen *et al*., 2015). Schaff *et al*. (2007) mentioned that *Mi* gene provides resistance of tomato, glycosyltransferase and extension may play a main role in the cell wall synthesis, which is a fundamental defence against root knot nematode. The NBS-LRR class of *R* genes could be sub-divided into two main groups depend on existence of domains identical to the *Toll* and interleukin-1 receptor or coiled-coil (CC) domain at the amino terminal (Bhattarai *et al*., 2007). In this study, we designed four pairs of primers for the amplification of $Mi_{1.2}$ gene from GenBank accession number AF039682.1. Primers TomMi1, TomMi2, TomMi3 and Tom Mi4 amplified 620, 600, 3300 and 1993 bp DNA fragments, respectively. Moreover, primers IMOF1 and IMOR1 amplified 998 bp (Bendezu, 2004). We isolated full length DNA and cDNA of $Mi_{1.2}$ gene from *S. peruvianum* tomato leaves using the primers F.P. TomMi3 and R.P. TomMi4. Sequencing results of both assembled fragments (620 bp and 600 bp) joined at the overlap region confirmed the $Mi_{1.2}$ gene sequence.

BLAST analysis showed that the *Mi1.2* gene under study (1440 bp) was homologous to tomato root-knot nematode resistance genes in the GenBank. Nucleotide sequence of *Mi1.2* gene under study was encoded 479 amino acids with molecular weight 54.59 KDa and PI was 5.52, which showed that $Mi_{1,2}$ protein was acidic. The PI is significant in protein purification because it represents the pH where solubility is typically minimal. Here, the protein isoelectric point signifies where mobility in an electro-focusing system is zero and in turn, the point where the protein will aggregate (Geourjon and Deleage, 1995). Chen *et al*., (2006) used the specific primers AM-FW1 and AM-RV1 for the isolation of full length DNA of *Mi1.2* resistance gene of 5.4 kb. They have cloned *Mi1.2* gene into the pDONR201 vector. Recombinant plasmid pDMi was confirmed by digestion by *ApaI* and *NruI* restriction enzymes and by sequencing. The results indicated that the amplicon 5367 bp was long. Also, observed that the DNA fragment had two introns, contained on an Open reading frame (ORF) of 3774 bp encoding 1257 amino acids. The BLAST results found that the predicted ORF of *Mi1.2* gene had 99% identity with tomato root-knot nematode resistance gene *Mi* (AF039682) recorded in GenBank, which is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes.

In the present study, we predicated the secondary and tertiary structures of $Mi_{1,2}$ protein using SPOMA and SWISS-MODEL workspace servers, respectively (Geourjon and Deleage, 1995; Arnold *et al*., 2006). The results showed that $Mi_{1.2}$ protein composed of 232 α-helix (48.4%), 77 β-sheets (16.1%), 39 Beta-turn (8.14%) and 131 coil (27.4%). The 3D model of the $Mi_{1.2}$ is described as consisting of an alpha helix and several beta pleated sheets with a compact structure. It is similar to the 3D crystal structure of *S. Lycopersicum* root-knot nematode resistance protein (NP_001234063.1). This is the first report describing $Mi_{1,2}$ protein isolated from *S. peruvianum*. In this research, we provided information on the three-dimensional structure (3D) of $Mi_{1.2}$ protein that associate directly to the corresponding *R* resistance proteins of the NBS-LRR class. The determination of the crystal structure of Mi proteins will help us to understand the protein-protein interactions between the R protein of the tomato and the Avr protein of the whitefly as confirmed in 'gene for gene' model. In addition, it provides us to comprehend the main role of an intermediary protein complex, which has been visualized in 'guard theory' of plant disease resistance (Chisholm *et al*., 2006; Chattopadhyaya and Pal, 2008). Dorna *et al*., (2014) obtained that the three-dimensional protein structure (3D) by protein crystallography (X-ray) provides to examine folds and motifs in the proteins, molecular folding, phylogenetic and structure/function relationships. One of the principle research troubles in proteomics is the prediction of the tertiary structures (3D). Proteins are long residues composed of 20 different amino acid sequences that in physiological conditions assume an alone 3-D structure. Information on the protein structure provides the study of biological operations with detail. The sequence-protein–structure paradigm (the "lock-and-key" theory) tells that the protein can perform its biological operation only by folding in to a singular, structured shape estimated by its amino acid residue (Anfinsen, 1973). Presently, it has been known that not all protein operations are linked to a folded shape (Tompa and Csermely, 2004; Dunker *et al*., 2008). Some proteins are unfolded or disordered to achieve their functions (Gunasekaran *et al*., 2003). These proteins are known Intrinsically Disordered Proteins (IDP) and act about 30% of the protein sequences.

Conclusion

Up to now, there have been no researches on the secondary (2D) and tertiary (3D) structures of $Mi_{1.2}$ protein. This is the first report describing $Mi_{1.2}$ protein isolated from *S. peruvianum*. The prediction of (3D) modeling is among the research troubles in structural bioinformatics. The 3D structure of a protein that has no templates in the Protein Data Bank (PDB) is a very difficult. A Knowing of the 3D structure of the $Mi_{1.2}$ protein provides very important data on its the biological function in the plant cell.

Acknowledgment

This work was supported by research project (Project reference 5345) founded by the Science and Technology Development Fund-Short Term Fellowship (STDF-STF) Program, Ministry of State for Scientific Research, Egypt.

Author's Contributions

Sherin Amin Mahfouze: Participated in all experiments and contributed to the writing of the manuscript.

Shipra Saxena and Heba Amin Mahfouze: Participated in all experiments**.**

Manchikatla Venkat Rajam: Contributed in reviewing of the manuscript.

Ethics

This research paper is original and contains unpublished data. The corresponding author confirms that all of other authors have read and accepted the manuscript.

References

Alatar, A.A. , M. Faisal, E.M. Abdel-Salam, T. Canto, Q. Saquib, S.B. Javed, M.A. El- Sheikh, and A.A. Al-Khedhairy, 2017. Efficient and reproducible *in vitro* regeneration of *Solanum lycopersicum* and assessment genetic uniformity using flow cytometry and SPAR methods. Saudi J. Biol Sci., 24: 1430-1436.

Anfinsen, C., 1973. Principles that govern the folding of protein chains. Science, 181: 223-223.

Arnold, K., L. Bordoli, J. Kopp and T. Schwede, 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics, 22: 195-201. DOI: 10.1093/bioinformatics/bti770

Bendezu, I.F., 2004. Detection of the tomato $Mi_{1.2}$ gene by PCR using non-organic DNA purification. Nematropica, 34: 23-30.

Bhattarai, K.K., Q. Li, Y. Liu, S.P.D. Kumar and I. Kaloshian, 2007. The *Mi-1*-mediated pest resistance requires Hsp90 and Sgt1. Plant Physiol., 144: 312-323. DOI: 10.1104/pp.107.097246

Chattopadhyaya, R. and A. Pal, 2008. Three-dimensional models of NB-ARC domains of disease resistance proteins in tomato, Arabidopsis and Flax. J. Biomol. Struct. Dyn., 25: 357-71. DOI: 10.1080/07391102.2008.10507184

Chen, H.M., C.Y. Lin, M. Yoshida, P. Hanson and R. Schafleitner, 2015. Multiplex PCR for detection of Tomato yellow leaf curl disease and root-knot nematode resistance genes in tomato (*Solanum lycopersicum* L.). Int. J. Plant Breed Genet., 9: 44-56.

Chen, R.G., L.Y. Zhang, J.H. Zhang, W. Zhang and X. Wang *et al*., 2006. Functional characterization of *Mi*, a root-knot nematode resistance gene from tomato (*Lycopersicon esculentum* L.). J. Integr. Plant Biol., 48: 1458-1465.

Chisholm, S.T., G. Coaker, B. Day and B.J. Staskawicz, 2006. Host-microbe interactions: Shaping the evolution of the plant immune response. Cell, 124: 803-814.

Dangl, J.L. and J.D.G. Jones, 2001. Plant pathogens and integrated defence responses to infection. Nature, 411: 826-833.

Dorna, M, M.B. Silva, L.S. Buriol and L.C. Lamb, 2014. Three-dimensional protein structure prediction: Methods and computational strategies. Comput. Biol. Chem., 53: 251-276.

Dunker, A., I. Silman, V. Uversky and J. Sussman, 2008. Function and structure of inherently disordered proteins. Curr. Opin. Struct. Biol., 18: 756-756.

Fulton, T.M., J. Chunwongse and S.D. Tanksley, 1995. Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol. Biol. Rep., 13: 207-209.

Geourjon, C. and G. Deleage, 1995. SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput. Applied Biosci., 11: 681-684.

Gunasekaran, K., C. Tsai, S. Kumar, D. Zanuy and R. Nussinov, 2003. Extended disordered proteins: Targeting function with less scaffold. Trends Biochem. Sci., 28: 81-81.

Kaur, S., M.S. Dhaliwal, D.S. Cheema, S.K. Jindal and A.K. Gaikwad, 2014. Screening of tomato (*Solanum lycopersicum* L.) germplasm for root-knot nematode resistance using conventional and molecular marker techniques. Indian J. Nematol., 44: 56-61.

Mahfouze, S., A. Heba A. Mahfouze and R.M. Esmail, 2015. Detection of the *Mi-1.2* gene from tomato confers resistance against whitefly (*Bemisia tabaci*). Wulfenia, 22: 136-146

McDaniel, T., C.R. Tosh, A.M.R. Gatehouse, D. George and M. Robson *et al*., 2016. Novel resistance mechanisms of a wild tomato against the glasshouse whitefly. Agron Sustain Dev., 36: 14-14. DOI: 10.1007/s13593-016-0351-4

Milligan, S.B., J. Bodeau, J. Yaghoobi, I. Kaloshian and P. Zabel *et al*., 1998. The root-knot nematode resistance gene *Mi* from tomato is a member of leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. Plant Cell, 10: 1307-1319.

Momotaz, A., J.W. Scott and D.J. Schuster, 2010. Identification of quantitative trait loci conferring resistance to *Bemisia tabaci* in an F_2 population of *Solanum lycopersicum* X *Solanum habrochaites* accession LA1777. J. Am. Soc. Hort. Sci., 135: 134-142.

Nombela, G., V.M. Williamson and M. Muniz, 2003. The root-knot nematode resistance gene *Mi-1.2* of tomato is responsible for resistance against the whitefly *Bemisia tabaci*. Mol. Plant Microbe Interact., 16: 645-649. DOI: 10.1094/MPMI.2003.16.7.645

Pallipparambil, G.R., R.J. Sayler, J.P. Shapiro, J.M.G. Thomas and T.J. Kring *et al*., 2014. *Mi-1.2*, an R gene for aphid resistance in tomato, has direct negative effects on a zoophytophagous biocontrol agent, *Orius insidiosus*. J. Exp. Bot., 66: 549-557. DOI: 10.1093/jxb/eru361

Rossi, M., F. Googin, S. Milligan, I. Kaloshian and D. Ullman *et al*., 1998. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc. Natl. Acad. Sci. USA, 95: 9750-9754.

Sambrook, J., E.F. Fritsch and T. Maniatis, 1989. Molecular Cloning: A Laboratory Manual. 2nd Ed., Cold Spring Harbor Laboratory, New York, ISBN-10: 0879693096, pp: 3.

Schaff, J.E., D.M. Nielsen, C.P. Smith, E.H. Scholl and D.M. Bird, 2007. Comprehensive transcriptome profiling in tomato reveals a role for glycosyltansferase in *Mi*-mediated nematode resistance. Plant Physiol., 144: 1079-1092

Seah, S., J. Yaghoobi, M. Rossi, C.A. Gleason and V.W. Williamson, 2004. The nematode resistance gene, Mi-1, is associated with an inverted chromosomal segment in susceptible compared to resistant tomato. Theor. Applied Genet., 108: 1635-1642

Tompa, P. and P. Csermely, 2004. The role of structural disorder in the function of RNA and protein chaperones. FASEB J., 18: 1169-1175. DOI: 10.1096/fj.04-1584rev