In vitro and in silico characterization of Solanum lycopersicum wound-inducible proteinase inhibitor-II gene

Faiza MUNIR1, Syed Muhammad Saqlan NAQVI2, Tariq MAHMOOD3,*
1Department of Biotechnology, Quaid-i-Azam University, Islamabad, Pakistan
2Department of Biochemistry, University of Arid Agriculture, Rawalpindi, Pakistan
3Department of Plant Sciences, Quaid-i-Azam University, Islamabad, Pakistan

Abstract: Plant proteinase inhibitors (PIs) are antimetabolic defensive proteins conferring resistance in plants against a variety of competing organisms such as bacteria, viruses, fungi, attacking nematodes, and insects. In the fields of plant biochemistry and molecular genetics research, tremendous success has been achieved in generating transgenic crops that have defensive approaches against biotic challenges. In this study, in vitro and in silico analysis was carried out for a wound-inducible PI-II gene isolated from 4 selected varieties (Roma, Nagina, Moneymaker, and Rio Grande) of Solanum lycopersicum L. Around 684 bp of PI-II gene was amplified, sequenced, translated, modeled to protein structure, and phylogenetically analyzed. The sequence analysis by BLAST showed high similarity scores (99%, 97%, 96%, and 94%) for Moneymaker, Roma, Rio Grande, and Nagina, respectively, with the original PI-II gene sequence from Solanum lycopersicum var. cerasiforme (GenBank accession no. AY007240) selected for primer designing. Sequenced data were translated to protein sequences, and translated sequences were modeled to 3-dimensional structures with iterative threading assembly refinement (I-Tasser) software. Phylogenetic analysis was carried out using Molecular Evolutionary Genetic Analysis software. Comparative phylogenetic analysis with 26 other complete coding sequences of PI from dicotyledonous plants was also done with in vitro analyzed PI-II genes from selected tomato varieties. In silico insight into the phylogenetic evaluation revealed that 30 PIs from different plants share a common root of evolutionary origin. Furthermore, 3-dimensional protein modeling by Ramachandran plot analysis revealed that PI from S. lycopersicum ‘Roma’ has the best quality structure with 85% of residues in most allowed regions.

Key words: Solanum lycopersicum, proteinase inhibitor, wound-inducible, in vitro, phylogenetic tree

1. Introduction

In plants, antimicrobial defense systems are gifts of nature. One of these systems comprises small antimicrobial proteins known as proteinase inhibitors (PIs) (1) that act as plant-defense–mediating proteins and contribute an innate defense against attacking pathogens and encroaching organisms such as infectious fungi, attacking nematodes, and herbivores (2). Koiwa et al. (3) reported 4 major classes of PIs (serine, aspartic, cysteine and metalloproteinase inhibitors) on the basis of target reaction sites. Serine PIs are widespread in the plant kingdom, and most of them have been characterized from family Solanaceae members potato and tomato (4). The second most studied class of PIs is cystatins, and among these, heat-stable rice cystatins are the most important (5). Margis-Pinheiro et al. (6) reported 9 cysteine PI genes (PtCys1–PtCys9) from Populus trichocarpa. PIs play their defensive role by obstructing metabolic proteins, which results in poor digestion in plant pests (7,8).

Plants cope with attackers by generating countless antimetabolic proteins that elicit noxious, revolting, and antimetabolic effects on phytophagous competitors (9). It has been observed that 1%–10% of the total protein content of most storage organs, such as seeds and tubers, are PIs that hinder the activity of different enzymes (10,11). Nonstorage tissues such as leaves, flowers, and roots also contain a large number of PIs (12–14). The tomato and potato PI families have the best studied examples of genes that are systemically expressed upon wounding. In potato, proteinase inhibitor-II (PI-II) is a multigene family, and its constitutive expression in tubers and floral buds and its wound-inducible expression in leaves have been reported (15,16). Plant PIs are developmentally regulated, and distinct regulation patterns have been reported in response to biotic and abiotic stresses. Studies have shown that during seed germination and maturation, and also under cold stress, expression of a wheat cystatin, TaMDC1, can be observed (17). In another study, expression of strawberry
cystatin (Cyf1) was seen in vegetative organs such as leaves and roots (18). Similarly, Lievens et al. (19) reported that during nodulation of Sesbania rostrata, a PI belonging to the Kunitz family (SrPI1) is expressed.

One of the major advances in agriculture has been the introduction of genetically engineered insect-resistant crops. The specificity of PIs in targeting definite groups of insects can help in generating transgenic plants with particular PIs that have inhibitory actions against specific pests (20). For example, transformed white poplar (Populus alba L.) plants developed using the Arabidopsis thaliana cysteine PI gene were resistant against Chrysomela populi beetle (21). Research data demonstrated that biotic stress such as insect chewing results in the expression of plant defensive proteins. For example, approximately 100 genes in lima bean, Phaseolus lunatus L., can be expressed in response to the chewing of the 2-spotted spider mite, Tetranychus urticae (Koch) (22).

PI-I and PI-II type inhibitors are widespread in the family Solanaceae, particularly in potato and tomato; for this study, we selected the PI-II gene from Solanum lycopersicum. We then focused on different tomato varieties of commercial importance from Pakistan to amplify and sequence this PI gene from different tomato varieties. In order to characterize PI genes, the present study was designed with the following main objectives: amplification of a wound-inducible PI-II gene from selected tomato varieties, analysis of the protein structures encoded by these PI genes, sequencing of the amplified genes, and in silico characterization of sequenced PI genes from selected tomato varieties with 26 randomly selected complete coding PI gene sequences from GenBank.

2. Materials and methods

2.1. Plant material

Seeds of 4 selected S. lycopersicum varieties (Roma, Nagina, Moneymaker, and Rio Grande) were acquired from the National Agriculture Research Center (NARC), Islamabad, Pakistan. The seeds were germinated in small pots containing manure soil in a growth room at 27 °C under cool white fluorescent lights (2000 lux), 75% humidity, and a photoperiod regime of 16 h of light and 8 h of dark.

2.2. Genomic DNA isolation and primer designing

DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method illustrated by Richards (23). A pair of primers was designed using the bioinformatics program Primer3 (http://frodo.wi.mit.edu/ primer3/input.htm). A sequence of the PI gene PI-II from the tomato S. lycopersicum var. cerasiforme (accession no. AY007240) was used for primer designing. The sequence of the forward and reverse primers was as follows:

PI-II F: 5’ TATCCATCATGGCTGTCCAC 3’
PI-II R: 5’ AACACACAACCTGTATCCCCACA 3’

2.3. Amplification of PI-II gene from different tomato varieties

To amplify a 684-bp gene, 25 µL of amplification reaction containing 25 pmol of each primer, 2.5 µL of 10X PCR buffer, 1.5 µL of 25 mM MgCl$_2$, 1.5 µL of 2.0 mM dNTPs, 45 ng/µL of genomic DNA, and 1.5 U Taq polymerase (Fermentas) was prepared. The amplification reaction was conducted in a gradient MultiGene Thermal Cycler (Labnet) programmed for 35 cycles of denaturation at 94 °C for 40 s, annealing at 55 °C for 40 s, and extension at 72 °C for 45 s, followed by a single-step final extension at 72 °C for 20 min. The amplification was confirmed on 1.5% agarose gel prepared in 0.5X Tris acetate EDTA (TAE) buffer.

2.4. DNA sequencing

The JETquick (Genomed) PCR Product Purification Spin Kit was used to purify the PCR product. Purified DNA product was used as a template for dye terminator cycle sequencing reaction, and sequencing was done in a Beckman and Coulter CEQ (8800) sequencer.

2.5. Sequence analysis

Different bioinformatics software, databases, and tools were used for data analysis. Phylogenetic trees were built by unweighted pair group method with arithmetic mean (UPGMA) using Molecular Evolutionary Genetic Analysis (MEGA) software version 4.0.02 (24). The sequenced PI genes were translated to protein sequences using the GENSCAN Web Server at MIT (http://www. genes.mit.edu/GENSCAN.html). The protein sequences were aligned using ClustalW (25). Iterative threading assembly refinement (I-Tasser) software was used to predict 3-dimensional protein structures from amino acid sequences. In silico phylogenetic analysis was also performed for a total of 30 PI gene sequences, including the 4 sequenced PI genes from selected tomato varieties used for the present research studies and 26 randomly picked PI gene sequences (complete coding sequences) selected from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov). These 26 PI genes are complete protein coding sequences selected from dicotyledonous plants.

2.6. Protein structure modeling and analysis

Protein 3-dimensional models for the 4 translated PI gene sequences were predicted using an online version of I-Tasser (http://wwwzhanglab.ccmb.med.umich.edu/I-TASSER). The protein structures were authenticated by Ramachandran plot using the PROCHECK program (http://www.ebi.ac.uk/thornton/software.html) (26).
3. Results and discussion

Isolated genomic DNA of 4 *S. lycopersicum* varieties (Roma, Nagina, Moneymaker, and Rio Grande) was qualitatively analyzed by running it on 1% agarose gel. The quantitative measurements were carried out by spectrophotometry. The DNA concentration that resulted in the best amplifications was 45 ng/µL. Amplification of the PI-II gene resulted in a PCR product of approximately 684 bp on 1.5% agarose gel prepared in 0.5X TAE buffer. The sequencing of PI-II genes from selected tomato varieties indicated high quality results signified by sharp peaks. The sequenced data from the tomato varieties Moneymaker Roma, Rio Grande, and Nagina showed 99%, 97%, 96%, and 94% homology, respectively, with the PI-II gene of *S. lycopersicum* var. *cerasiforme* (accession no. AY007240), which was used for primer designing. The sequences were submitted to GenBank and accession numbers were acquired. The similarity scores of the in vitro analyzed PI-II gene from 4 tomato varieties suggest that the nucleotide order for the sequenced PI gene is common.

A phylogenetic tree was constructed for the 4 characterized PI gene sequences to establish the evolutionary relationship with the original PI gene sequence from which the primers were designed. The dendrogram showed 2 major clusters, a and b (Figure 1). The PI genes from *S. lycopersicum* varieties Roma (JN091682) and Rio Grande (JN132113) in cluster a have close phylogenetic affinity with the original PI sequence from *S. lycopersicum* var. *cerasiforme* (AY007240). However, cluster b comprises PI gene sequences from *S. lycopersicum* varieties Nagina (JN132111) and Moneymaker (JN132112), showing a direct evolutionary link with cluster a. The sequenced nucleotides were translated into amino acid sequences by the GENSCAN Web Server. Multiple sequence analysis was performed for the 4 translated amino acid sequences with ClustalW (25). Conserved regions were highlighted with different colors (Figure 2).

3.1. Protein structure analysis

Protein 3-dimensional structures (Figure 3) were generated using I-Tasser version 1.1. The I-Tasser server generates the 5 full length 3-dimensional models of each query sequence along with the confidence score, evaluated TM score (an algorithm that calculates the similarity of topologies of 2 proteins or models), and root mean squared deviation for the evaluations (27). From 5 predicted structures for each PI protein, the best model was selected after Ramachandran plot assessment. The finest model was picked based on highest percentages of residues in most allowed regions and lowest percentage scores in disallowed regions. The confidence score values (C-scores) predicted for the best selected PI protein structures by I-Tasser were −1.320, −1.806, −0.928, and −1.262 for *S. lycopersicum* varieties Roma, Nagina, Moneymaker, and Rio Grande.
Protein 3-dimensional structures are fundamental as the biological activity of a protein is accomplished by its 3-dimensional structure (28).

3.2. Ramachandran plot analysis
The stereochemical quality and exactness of the predicted PI proteins from the 4 tomato varieties (Roma, Nagina, Moneymaker, and Rio Grande) under investigation were analyzed through residue-by-residue geometry and overall geometry of protein structures using the PROCHECK program (http://www.ebi.ac.uk/thornton/software.html) (26). Ramachandran plots were drawn for these protein structures. In Ramachandran plots (Figure 4), the most allowed regions are indicated by red patches, while yellow areas show allowed regions. It was observed that *S. lycopersicum* ‘Roma’ (JN091682) PI protein has 85.0% fully allowed region, 11.7% additional allowed region, 2.5% generously allowed region, and 0.8% disallowed region. In the case of *S. lycopersicum* ‘Nagina’ (JN132111), plot analysis revealed 76.4% fully allowed region, 17.1% additional allowed region, 4.9% generously allowed region, and 1.6% disallowed region. However, in *S. lycopersicum* ‘Moneymaker’ (JN132112), the Ramachandran plot showed 78.3% most favored region, 17.5% additional allowed region, 3.3% generously allowed region, and 0.8% disallowed region. Finally, in *S. lycopersicum* ‘Rio Grande’ (JN132113), there is 79.0% most favored region, 14.3% additional allowed region, 4.2% generously allowed region, and 2.5% disallowed region. Assessment and authentication results from Ramachandran plot analysis showed that the PI protein structure of *S. lycopersicum* ‘Roma’ (JN091682) is a good quality structural model with 85.0% of residues in the most favored region. Similar structure modeling and Ramachandran plot analysis was carried out to validate the structural and functional analysis of cysteine protease and cystatin from tomato (29). In another study, in silico studies were conducted for structural modeling of antioxidant proteins of spinach by Ramachandran plot analysis, and protein models were validated by computational tools PROCHECK and WHAT IF (30).

3.3. Comparative phylogenetic analysis with other PI gene sequences
Phylogenetic analysis was performed using MEGA version 4.0.02. The phylogram generated by cluster analysis of 30 PI gene sequences (Table) (31–43) showed 2 major clusters, I and II (Figure 5). The sequenced PI genes from the 4 tomato varieties in our present research (accession nos. JN091682, JN132111, JN132112, and JN132113) are present in cluster...
I. These PI genes are closely related to the PI gene sequence from *S. lycopersicum* var. *cerasiforme* (accession no. AY007240), from which the primers were designed. Cluster I is divided into 2 subclusters, A and B. Similarly, cluster II has 2 subclusters, C and D. Overall, cluster I represents the PI genes from 4 plant families, Solanaceae, Salicaceae, Fabaceae, and Brassicaceae, and cluster II includes PI genes from 3 plant families, Solanaceae, Fabaceae, and Brassicaceae. In our study, cluster analysis showed that 5 PIs (accession nos. JN091682, JN132111, JN132112, JN132113, and AY007240), all encoding PI-II protein in cluster I, are 99% evolutionarily and genetically linked with a genetic distance of 0.1% as indicated in the rooted neighbor-joining tree (Figure 5). Furthermore, our results revealed that 3 other PI genes (accession nos. AY204562, AY204563, and AY059390) are 100% genetically allied, and these 3 PIs encode trypsin inhibitor protein. Similarly, 2 PIs (accession nos. AY129402 and M15186) have 100% genetic similarity.

The PI genes from the same plant may differ on the basis of function and coding product. In our study there are 3 PI genes (AM162666, AM162667, and AM162668) from the family Brassicaceae; 2 of these (accession nos. AM162666 and AM162668) fall in cluster II, and 1 gene (accession no. AM162667) falls in cluster I. The 2 PI genes in cluster II encode the rapeseed trypsin inhibitor, while that in cluster

**Figure 4.** Ramachandran plot of predicted models of PIs (PI-II) from *S. lycopersicum* varieties. A: Roma (JN091682), B: Nagina (JN132111), C: Moneymaker (JN132112), and D: Rio Grande (JN132113). Plots were generated using PROCHECK program.
Table. The GenBank accession numbers and sizes of 30 PI genes picked from different plants that were phylogenetically analyzed in the current study.

<table>
<thead>
<tr>
<th>Serial no.</th>
<th>Accession no.</th>
<th>Source</th>
<th>Size (bp)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>JN091682</td>
<td><em>Solanum lycopersicum</em></td>
<td>684</td>
<td>Present study</td>
</tr>
<tr>
<td>2</td>
<td>JN132111</td>
<td><em>Solanum lycopersicum</em></td>
<td>684</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>JN132112</td>
<td><em>Solanum lycopersicum</em></td>
<td>682</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>JN132113</td>
<td><em>Solanum lycopersicum</em></td>
<td>677</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>AB110700</td>
<td><em>Solanum lycopersicum</em></td>
<td>559</td>
<td>(31)</td>
</tr>
<tr>
<td>6</td>
<td>AY007240</td>
<td><em>Solanum lycopersicum</em></td>
<td>684</td>
<td>Xie and Wu (unpublished data)</td>
</tr>
<tr>
<td>7</td>
<td>L25128</td>
<td><em>Solanum lycopersicum</em></td>
<td>4272</td>
<td>(32)</td>
</tr>
<tr>
<td>8</td>
<td>AY129402</td>
<td><em>Solanum lycopersicum</em></td>
<td>1670</td>
<td>(33)</td>
</tr>
<tr>
<td>9</td>
<td>M15186</td>
<td><em>Solanum tuberosum</em></td>
<td>1241</td>
<td>(34)</td>
</tr>
<tr>
<td>10</td>
<td>U45450</td>
<td><em>Solanum tuberosum</em></td>
<td>2068</td>
<td>(35)</td>
</tr>
<tr>
<td>11</td>
<td>X04118</td>
<td><em>Solanum tuberosum</em></td>
<td>1914</td>
<td>(36)</td>
</tr>
<tr>
<td>12</td>
<td>X78275</td>
<td><em>Solanum tuberosum</em> (Arran Banner) pin2-CM7 gene</td>
<td>584</td>
<td>(37)</td>
</tr>
<tr>
<td>13</td>
<td>Z12753</td>
<td><em>Solanum tuberosum</em></td>
<td>1695</td>
<td>(38)</td>
</tr>
<tr>
<td>14</td>
<td>Z13992</td>
<td><em>Solanum tuberosum</em></td>
<td>2330</td>
<td>Choi et al. (unpublished data)</td>
</tr>
<tr>
<td>15</td>
<td>AF330701</td>
<td><em>Ipomoea batatas</em></td>
<td>519</td>
<td>(39)</td>
</tr>
<tr>
<td>16</td>
<td>AF330702</td>
<td><em>Ipomoea batatas</em></td>
<td>524</td>
<td>(39)</td>
</tr>
<tr>
<td>17</td>
<td>AY059390</td>
<td><em>Phaseolus vulgaris</em></td>
<td>327</td>
<td>Yang et al. (unpublished data)</td>
</tr>
<tr>
<td>18</td>
<td>AY204563</td>
<td><em>Vigna unguiculata</em> subsp. <em>unguiculata</em></td>
<td>326</td>
<td>Yuan et al. (unpublished data)</td>
</tr>
<tr>
<td>19</td>
<td>AY204562</td>
<td><em>Vigna unguiculata</em> subsp. <em>unguiculata</em></td>
<td>326</td>
<td>Yuan et al. (unpublished data)</td>
</tr>
<tr>
<td>20</td>
<td>AF349441</td>
<td><em>Populus tremuloides</em></td>
<td>682</td>
<td>(40)</td>
</tr>
<tr>
<td>21</td>
<td>AY749108</td>
<td><em>Solanum americanum</em></td>
<td>1943</td>
<td>(41)</td>
</tr>
<tr>
<td>22</td>
<td>Z12824</td>
<td><em>Solanum tuberosum</em></td>
<td>1573</td>
<td>(38)</td>
</tr>
<tr>
<td>23</td>
<td>AM162668</td>
<td><em>Brassica napus</em></td>
<td>642</td>
<td>(42)</td>
</tr>
<tr>
<td>24</td>
<td>AM162667</td>
<td><em>Brassica napus</em></td>
<td>1023</td>
<td>(42)</td>
</tr>
<tr>
<td>25</td>
<td>AM162666</td>
<td><em>Brassica napus</em></td>
<td>1114</td>
<td>(42)</td>
</tr>
<tr>
<td>26</td>
<td>DQ412560</td>
<td><em>Vigna trilobata</em></td>
<td>678</td>
<td>Sinha et al. (unpublished data)</td>
</tr>
<tr>
<td>27</td>
<td>DQ417203</td>
<td><em>Vigna radiata</em></td>
<td>543</td>
<td>Sinha et al. (unpublished data)</td>
</tr>
<tr>
<td>28</td>
<td>DQ417204</td>
<td><em>Vigna unguiculata</em></td>
<td>663</td>
<td>Sinha et al. (unpublished data)</td>
</tr>
<tr>
<td>29</td>
<td>D17331</td>
<td><em>Solanum tuberosum</em></td>
<td>1103</td>
<td>(43)</td>
</tr>
<tr>
<td>30</td>
<td>D17332</td>
<td><em>Solanum tuberosum</em></td>
<td>1036</td>
<td>(43)</td>
</tr>
</tbody>
</table>
I encodes the rapeseed glutamyl endopeptidase inhibitor. Although these PI genes are from the same plant (*Brassica napus*), there is a difference in their function and coding products. Genes closely related in terms of function are evolutionarily linked.

Genes belonging to the same PI and plant family fall into different clusters on the basis of size difference. From the plant family Fabaceae, 2 PI genes (DQ412560 and DQ417204) fall in cluster I, while 3 others (AY204562, AY204563, and AY059390) are in cluster II. Although these 5 PI genes are trypsin PIs from same plant family (Fabaceae), they are in different clusters due to the difference in their sizes (nucleotide length base pairs). The 2 PI genes (DQ412560 and DQ417204) in cluster I are closely related in terms of size at 678 and 663 bp in length, respectively, while 3 other PIs that fall in cluster II (AY204562, AY204563, and AY059390) are 326, 326, and 327 bp, respectively.

In silico phylogenetic evaluations can lead to important insights in terms of evolutionary affinities among investigated protein genes. Martinez et al. (44) phylogenetically analyzed different plant cystatins from rice, arabidopsis, and barley. In their study, 12 cysteine PI genes from rice, 7 from arabidopsis, and 7 from barley were analyzed in silico by constructing a phylogenetic tree by neighbor-joining method using the amino acid sequences of these 26 cystatin proteins. In an earlier study, molecular and phylogenetic analysis of the wound-inducible PI-I gene was carried out for the 7 direct ancestors of *Lycopersicon esculentum*: *L. pennellii, L. chilense, L. hirsutum, L. parviflorum, L. peruvianum var. humifusum, L. cheesmanii*, and *L. peruvianum* (45).

In another report, it was observed that HvCPI-4 from *Hordeum vulgare* and OC-XII protein from *Oryza sativa* are closely allied with the highest similarity scores, while the *Arabidopsis* cystatins were found scattered in the resulting tree and appeared to be functionally distant from rice and barley proteins. In yet another study, Martinez et al. (46) conducted the phylogenetic analysis of 17 cysteine PI proteins from different plants, and it was found that

---

**Figure 5.** A phylogenetic tree indicating relationship of 4 sequenced PI genes from selected tomato varieties in current study with 26 randomly picked complete coding PI gene sequences from different plants.
cysteine PI from *Malus domestica* had a close phylogenetic affinity with the functionally analyzed strawberry FaCPI-1 in terms of antifungal properties. Phytocystatins have not been reported from any other cultivated fruit crop except apple (*Malus domestica*). Among plants, proteinase inhibitors PI-I and PI-II belonging to the serine PI family have been widely investigated. Recently, 9 cysteine PI genes, designated as PtCys1–PtCys9, have been documented from *Populus trichocarpa*. The location of these genes in the genome is as follows: chromosome 1 (PtCys1 and PtCys2), chromosome 2 (PtCys3), chromosome 3 (PtCys4), chromosome 6 (PtCys5 and PtCys6), chromosome 9 (PtCys7), chromosome 14 (PtCys8), and chromosome 16 (PtCys9) (6). Similar phylogenetic analysis was conducted for 15 PIs from the mustard inhibitor family using the neighbor-joining cluster analysis method (42). In that study, 3 PI genes out of 15 mustard PIs from *Brassica napus* were characterized in vitro, and phylogenetic analysis was conducted with 12 other coding PI genes from the mustard inhibitor family selected from GenBank. Baloğlu et al. (47) constructed a neighbor-joining phylogenetic tree to investigate evolutionary association among 14 different Ran binding proteins from different plant species.

In our study, tomato varieties were selected from different regions of Pakistan. Rio Grande, Roma, and Nagina belong to Punjab Province while Moneymaker originates in the northern areas of Pakistan. In Pakistan, tomato varieties are generally categorized into 2 distinct types: determinate tomatoes and indeterminate tomatoes. Determinate varieties flower and set fruit all at once, followed by dropping. Determine varieties are compact plants, and they flower at the ends of shoots, which determines their length. Rio Grande and Roma are good examples of determinate varieties in Pakistan. Indeterminate varieties continue to grow throughout the season. Their flowers grow along vines that do not determine their length; indeterminate varieties require support and pruning. The best example of an indeterminate tomato variety in Pakistan is Moneymaker. Indeterminate varieties have a higher yield potential than determinate varieties.

Postharvest research work conducted on tomato cultivars at different research institutes in Pakistan such as the NARC in Islamabad, the University of Agriculture in Faisalabad, the Sindh Horticulture Agriculture Research Institute in Mirpurkhas, and the Ayub Agriculture Research Institute in Faisalabad reported that the 2 tomato varieties Rio Grande and Roma were high-yielding with a longer postharvest life; hence, these are commercially valuable tomato cultivars in Pakistan.

Specific environmental conditions in a particular region can explain some genetic variations among the 4 genetically analyzed tomato varieties. In another study it was reported that changes in gene structure and function may occur due to transposable elements through insertion, excision, and transposition (48). Various important molecular phenomena reported in the potato type II (pot II) PI family (tandem duplication, domain swapping, and fold circular permutation) can be used for evolutionary studies in the gene family (49,50). According to Mello et al. (51), gene mutations, as internal gene duplications, may be the reason behind the evolution of the family of Bowman–Birk inhibitors (BBIs), revealing great variability in BBIs from monocotyledonous plants. The 2 legume species *Glycine soja* and *Glycine max* were placed under phylogenetic study utilizing the gene sequences from a multigene PI family that illustrated evolutionary propinquity between these 2 legume strains (52).

In the current study, the PI-II gene from tomato varieties was analyzed through extensively studied members of the family Solanaceae. From our phylogenetic results, we conclude that PIs from the same plant family may separate into different clades on the basis of differences in coding products. Genes closely related in terms of translated products/functioning are evolutionally associated. Ramachandran plot analysis of predicted proteins depicted good quality structures. The finest PI protein structural model was from *S. lycopersicum* 'Roma' (JN091682) with the highest percentage (85.0%) of residues in the most allowed region. Our future aim is to use the PI-II gene from *S. lycopersicum* characterized in our present study to generate valuable transformed crop plants with improved defensive chemistry. We are working to express this PI-II gene under the control of a powerful tissue-specific promoter to generate transgenic potato plants.

Acknowledgments

The authors are deeply thankful to the Pakistan Science Foundation, Islamabad, and the Higher Education Commission, Islamabad, Pakisan, for financial assistance.

References


