Characterization and genetic variability of coat protein gene of Apple chlorotic leaf spot virus isolates from apple in Kurdistan and western Azarbaijan Provinces, Iran

Kamal Shujaee, Mahdi Azaryar and Mohammad Hajizadeh*

Department of Plant Protection, Faculty of Agriculture, University of Kurdistan, Sanandaj, Iran.

Abstract: In this study, 174 leaf samples from main apple growing areas in Kurdistan and western Azarbaijan Provinces were examined for infection by Apple chlorotic leaf spot virus (ACLSV). The result of RT-PCR showed that 17 samples (9.7%) were infected by ACLSV in both provinces. Subsequently, the genetic structure of ACLSV was determined based on the full coat protein (CP) sequences of eight isolates from this study and those recently published from Iran and worldwide. In phylogenetic analysis, ACLSV isolates from apple clustered into three known phylogenetic types: B6, P205, and SHZ, which are characterized by three amino acids at positions of 40, 75 and 79. The new Iranian ACLSV isolates showed the CP amino acid sequence of Ser40-Tyr75-Glu79 which placed them with the B6 type. Further analyses showed that the Iranian ACLSV isolates (the new and previously published isolates) had lower genetic variability compared to isolates from other countries except for Brazilian population. Most of the CP codons were under negative selection (dN/dS < 1) except for one codon at position 92 which was under positive selection. This study suggests that negative selection was most important evolutionary factor driving the genetic evolution of ACLSV.

Keywords: apple, negative selection, phylogenetic analysis, ACLSV

Introduction

Apple is commercially the most important fruit crop and has the most exporting value in Iran. Iran stands at 4th and 7th positions in orchard area and apple production with 0.24 mHa and 2.45 mt, respectively (FAOSTAT, 2016). Viruses as biotic factor are the main agent in low productivity of apple worldwide (Watpade et al., 2012). Among various viral diseases, ACLSV is widely distributed in Rosaceae fruit trees such as apple, apricot, cherry and pear, usually causing a latent infection in most commercial cultivars (Watpade et al., 2012). However, some apple cultivars are susceptible and may develop severe symptoms. The severity of symptoms induced by ACLSV depends on host plant species and virus strain (Nemeth, 1986). Infection rates of up to 80-100% with yield losses in the order of 30-40% by this virus have been reported (Nemchinov et al., 1995; Wu et al., 1998; Cembali et al., 2003).

ACLSV is the type member of the genus Trichovirus, subfamily Tymivirinae, family Betafelexiviridae, within the order Tymovirales (Adams et al., 2016). ACLSV has a filamentous particle, composed of a positive-sense, single-stranded RNA molecule of about 7.5 kb excluding the polyadenylated tail (Yoshikawa, 1988; Yoshikawa et al., 1992). The genome of ACLSV
Genetic variability of ACLSV

contains three overlapping open reading frames (ORF), which encode a 216-kDa replication-associated protein, a 50-kDa movement protein which overlaps both replicase and CP coding regions, and a 22-kDa coat protein (CP), respectively (Sato et al., 1993; Niu et al., 2012).

Molecular variation and characterization were reported for the ACLSV isolates from India, China and Japan (Al Rwahnih et al., 2004; Yaegashi et al., 2007; Rana et al., 2010; Chen et al., 2014). Considering that Iran is the 7th largest producer of apple around the world, a greater understanding of ACLSV variability is necessary, in particular with regard to isolates from apple. In Iran, ACLSV has been reported from several provinces (Derakhshian et al., 2010; Farmahini et al., 2014; Keshavarz and Shams-Baksh, 2015; Alemzadeh et al., 2016; Abtahi et al., 2019) and previous surveys have indicated that 2.67% of pear (Farmahini et al., 2014) and 18.55% of apple (Keshavarz and Shams-Baksh, 2015) trees are infected with ACLSV. Recently, Abtahi et al. (2019) showed that Iranian ACLSV isolates belonged to B6 and P205 phylogenetic groups based on partial CP gene sequences. So far, all Iranian isolates of ACLSV CP have been partially sequenced (Keshavarz and Shams-Baksh, 2015; Naderpour et al., 2016; Abtahi et al., 2019) and complete CP gene sequence of ACLSV has not been previously determined. Also, despite of the detection of ACLSV in some provices of Iran, no information is available on presence of ACLSV in Kurdistan Province. In this work, we performed an updated analysis of the genetic variation and phylogenetic relationship of Iranian ACLSV isolates by analysis of the recently published coat protein gene sequences (Abtahi et al., 2019) and those investigated in the present study.

Materials and Methods

Source of isolates

During July, August, and September 2017 and 2018, 174 symptomatic and non-symptomatic apple leaf samples were collected from several locations of Kurdistan and Western Azarbaijan Provinces of Iran.

RT-PCR detection of ACLSV

Total nucleic acid (TNA) was extracted from one leaf per sample based on the method described by Foissac et al. (2000) with minor modifications. Reverse transcription (RT) reactions were performed by a cDNA synthesis kit (HyperScriptTM Reverse Transcriptase, GeneAll, South Korea) using random hexamer primer. The synthesised cDNA was subjected to PCR amplification in a 12.5 μl PCR master mix (GeneAll, South Korea) and 10 pmol of each primer, ACLS-MF: 5’ GAGGATCCATGGC GGCAG’ and ACLSV-MR: 5’ TACTCG AGGTAAATGCAAAGATCAG 3’ (designed in this study), In PCR master mix. PCR products were electrophoresed on a 1.2% (w/v) agarose gel (containing 0.5 μg/ml EtBr) in 1X TAE buffer.

Cloning and sequencing of RT-PCR products

RT-PCR products of eight apple trees were purified using a Nucleic Acid Purification Kit (GeneAll, South Korea), ligated into the pTG19 cloning vector (Vivantis, Malaysia), and the ligation mix was transformed into Escherichia coli strain DH5α as described by Chung et al. (1989). The recombinant plasmids were confirmed by restriction analysis and a purified clone from each isolate was subjected to sequencing by Macrogen Inc. (Seoul, South Korea).

Multiple sequence alignment and phylogenetic analysis

Multiple sequence alignment of ACLSV CP sequences, including eight sequences of ACLSV-CP obtained in this study and additional sequences retrieved from GenBank was performed separately using Clustal W (Thompson et al., 1997) implemented in MEGA X (Kumar et al., 2018) with default parameters. To investigate the evolutionary relationship and genetic diversity of these isolates, phylogenetic analysis was done by neighbor joining (NJ) method implemented in MEGA X and branch support was computed with 1000 bootstrap replicates.
Population genetic parameters and checking for recombination
Estimate number of haplotypes (H), polymorphic sites (S), average pairwise nucleotide diversity (\(\pi\)), average number of nucleotide differences (K), and ratio of non-synonymous to synonymous nucleotide diversity (\(dN/dS\)) were obtained using DnaSP V.5.10.01 (Rozas, 2009). To identify individual codon positions evolving under negative (purifying) and positive selections, three different codon-based maximum-likelihood algorithms; single likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), internal fixed effects likelihood (IFEL), fast and unconstrained bayesian approximation (FUBAR) within the HyPhy software package (Pond et al., 2005) were used with significance level set at P-value < 0.1.

Recombination detection program RDP4 (Martin et al., 2015) was applied to detect and analyze recombination signal in the CP gene sequences of ACLSV using several methods (Chimaera, Geneconv, MaxChi, RDP, Siscan, LARD and 3Seq), with default setting, and a P-value threshold of 0.05.

Results
Detection of ACLSV by RT-PCR and sequencing
RT-PCR assays were performed to detect ACLSV in 174 samples. The results showed that 17 samples (9.7%) of both provinces were infected by ACLSV (Fig. 1). Eight isolates chosen from positive samples based on the geographic regions and symptoms were cloned and sequenced. Sequences were submitted to GenBank (accession numbers MK354022-MK354029; Table 1). Alleight ACLSV-CP sequences contained 582 nucleotides (nt) with 193 deduced amino acid residues and had a TAG at the end as stop codon. Identity of ACLSV-CP isolates ranged from 91 to 99.3% at the nucleotide level and 96.8-99.5% at the amino acid level.

Figure 1: Electrophoresis of PCR products amplified by ACLSV-MF and ACLSV-MR primers. M: Marker (Gene RulerTM DNA Ladder Mix 1Kb), 1-3: Coat protein gene of ACLSV amplified from apple samples. 4: Negative control. The picture of leaf sample infected by ACLSV (Line 3) is shown on the right.
Table 1 Data on full coat protein gene sequences of Iranian isolates of ACLSV characterized in this study.

<table>
<thead>
<tr>
<th>Collection site</th>
<th>Virus species</th>
<th>Isolate name</th>
<th>Host</th>
<th>Symptoms</th>
<th>Accession numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kamyaran¹</td>
<td>ACLSV</td>
<td>KSS</td>
<td>Apple</td>
<td>LD, Y</td>
<td>MK354022</td>
</tr>
<tr>
<td>Khoje²</td>
<td>ACLSV</td>
<td>X2</td>
<td>Apple</td>
<td>Y</td>
<td>MK354023</td>
</tr>
<tr>
<td>Khoje³</td>
<td>ACLSV</td>
<td>X</td>
<td>Apple</td>
<td>Y</td>
<td>MK354024</td>
</tr>
<tr>
<td>Divandareh¹</td>
<td>ACLSV</td>
<td>D4</td>
<td>Apple</td>
<td>LD, Y</td>
<td>MK354025</td>
</tr>
<tr>
<td>Sanandaj¹</td>
<td>ACLSV</td>
<td>SSa</td>
<td>Apple</td>
<td>LD, M</td>
<td>MK354026</td>
</tr>
<tr>
<td>Saqqez¹</td>
<td>ACLSV</td>
<td>Sq</td>
<td>Apple</td>
<td>M</td>
<td>MK354027</td>
</tr>
<tr>
<td>Kamyaran¹</td>
<td>ACLSV</td>
<td>KS</td>
<td>Apple</td>
<td>LD, Mo</td>
<td>MK354028</td>
</tr>
<tr>
<td>Marivan¹</td>
<td>ACLSV</td>
<td>MS</td>
<td>Apple</td>
<td>NS</td>
<td>MK354029</td>
</tr>
</tbody>
</table>

¹ Kurdistan Province, Iran, ² Western Azarbaijan, Iran.
³ LD; leaf deformation, Y; yellowing, M; mosaic, mo; mottle, NS; no symptoms, BLM; burnt leaf margin, VC; vein clearing, LC; leaf curl. These symptoms may not necessarily refer to ACLSV and might be caused by other agents.

Phylogenetic analysis

Phylogenetic analysis based on the full-length CP sequences of the eight new ACLSV apple isolates in this study and 40 analogues of apple isolates from Brazil, China, Japan, Germany, India, Latvia, and Lithuania were divided into three major phylogroups; B6 (Yaegashi et al., 2007), SHZ (Chen et al., 2014), and P205 (Yaegashi et al., 2007) (Fig. 2). Phylogroup I (B6 type) with the CP amino acid sequence of Ser⁴⁰-Tyr⁷⁵-Glu⁷⁹ was clustered into two subgroups; subgroup I was composed of the 22 isolates, including the all eight isolates described in this study and subgroup II contained 3 isolates from Brazil with one difference in amino acid at position 32. Phylogroup II (P205 type) with the CP amino acid sequence of Ala⁴⁰-Phe⁷⁵-Glu⁷⁹ was clustered into two subgroups; each subgroup included 6 isolates with two differences in amino acids at positions 72 and 97. Phylogroup III (SHZ type) included 11 isolates mainly from China containing the CP amino acid sequences of Ser⁴⁰-Tyr⁷⁵-Ser⁷⁹ with one exception in the isolate GZL13 with alanine (Ala) at position 40.

Coat protein variability of the Iranian ACLSV sequences

The genetic diversity of ACLSV isolates was studied by sequencing and aligning full-length CP sequences of cDNA clones in west and northwest Iranian isolates. The haplotype diversity for the new Iranian isolates of ACLSV was 1.000, whereas nucleotide diversity was 0.065 ± 0.007. Nucleotide diversity value was lower than 0.1, indicating a low genetic diversity within Iranian ACLSV isolates. dN/dS ratio was lower than 1 (0.037), indicating that CP gene of Iranian ACLSV isolates is under negative selection.

The ACLSV-CP showed the greatest nucleotide diversity at position 351, while showing lower nucleotide diversity close to its 5′-terminal region (Fig. 3). We also found a decrease in the π value from position 494 towards the 3′-terminal region. Accordingly, lowest variability of amino acids was detected in the C and N-terminal regions of the CP (amino acids 2 to 58 and 127 to 193, respectively) with two amino acid changes in positions 40 and 170, while the middle region was substantially more variable (nine amino acid variations).
Figure 2 Neighbor-Joining phylogenetic tree constructed from the nucleotide alignment of coat protein sequences (582 nt) from the *Apple chlorotic leaf spot virus* isolates of this study (8 isolates) and 40 isolates retrieved from GenBank. Numbers over the branches indicate the bootstrap support. Amino acid abbreviations: A, alanine; S, serine; Y, tyrosine; E, glutamate; V, valine; F, phenylalanine; I, isoleucine.

Figure 3 Trend of polymorphism along the coat protein gene sequences in the Iranian ACLSV population, comprising of the new (this study) and published Iranian isolates (Abtahi et al., 2019). The curves were generated by sliding windows with 50 and 10 as the window and step sizes, respectively.
Identifying individual codon positions evolving under negative and positive selections

Most of the codons in CP gene sequences of ACLSV isolates were under negative selection ($d_\text{s}/d_\text{s} < 1$) except for one codon, which was significantly under positive selection at positions 92, detected by SLAC and FEL methods (Table 2). In comparison with positive selection site, the sites under negative selection sites were very high, indicating strong negative selection was accrued on the CP gene.

Table 2 Analysis of selection for the CP of ACLSV Iranian isolates using the in Datamonkey server.

<table>
<thead>
<tr>
<th>Methods</th>
<th>PS</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLAC(^1)</td>
<td>1 (92)</td>
<td>116</td>
</tr>
<tr>
<td>FEL(^2)</td>
<td>1 (92)</td>
<td>120</td>
</tr>
<tr>
<td>IFEL(^3)</td>
<td>0</td>
<td>112</td>
</tr>
<tr>
<td>FUBAR(^4)</td>
<td>0</td>
<td>140</td>
</tr>
</tbody>
</table>

PS: sites under positive selection, NS: sites under negative selection.
Codon-based maximum-likelihood algorithms.
\(^1\) Single likelihood ancestor counting (SLAC).
\(^2\) Fixed effects likelihood (FEL).
\(^3\) Internal Fixed effects likelihood (IFEL).
\(^4\) Fast, Unconstrained Bayesian Approximation (FUBAR).

Discussion

In this study, the infection, genetic variability and phylogenetic relationship of ACLSV from apple in Kurdistan and western Azarbaijan Provinces were studied. The infection level with ACLSV in apple trees in these two provinces was low, while a high infection level with ACLSV in apple cultivars and rootstocks was reported in different countries; up to 86% in Moldova (Verdervskaja and Marinescu, 1985), 73.8% in the Lithuania (Mazeikiene et al., 2018), 70% in the Czech Republic (Grimova’ et al., 2016), 66% in Greece (Mathioudakis et al., 2010) and 60% in the USA (Waterworth, 1993).

In phylogenetic analysis, ACLSV isolates from apple divided into three known phylogroups (B6, P205 and SHZ) by combinations of the three amino acids at positions: 40 (serine or alanine), 75 (tyrosine or phenylalanine), and 79 (glutamate or serine) (Chen et al., 2014) and the new Iranian ACLSV isolates were clustered in the B6 type (Fig. 2). We found that B6 and P205 phylogroups divided into two subgroups, I and II (Fig. 2). Two subgroups in the B6 and P205 had one and two differences in amino acid alignment, respectively. Subgroup I in B6 had alanine instead of valine at position 32 and subgroup I in p205 had valine and serine instead of isoleucine and asparagine at positions 72 and 97, respectively. Although the CP region in ACLSV isolates is relatively conserved, covariations has been reported among phylogroups (Yaegashi et al., 2007, 2014; Chen et al., 2012). Based on the combinations in the five amino acids at positions 40, 59, 75, 130, and 184, ACLSV isolates from Japan were clustered into the two major phylogroups B6 and P205 (Yaegashi et al., 2007). Similarly, ACLSV isolates from apple samples worldwide have been tentatively classified into the three types among a new cluster (SHZ) of China based on the amino acids sequences at three positions 40, 75, 79 (Chen et al., 2014). In the same way, ACLSV isolates from peach samples have been tentatively clustered into Z1 and Ta Tao 5 types (Niu et al., 2012). To understand the role of these amino acid variations in pathogenicity and life cycle needs further data from more isolates, mutation study and ACLSV infection clones of these isolates. The long terminal branches of the isolates belonging to subgroup B of the B6 type (Fig. 1) may suggest that they have accumulated mutations over a long time.

To better characterize the genetic variability of the different Iranian isolates of ACLSV analyzed in this study and those published by Abtahi et al., (2019), the descriptors of variability were determined and compared with the isolates from other countries (Nickel et al., 2018; Table 2). Lower genetic variability in ACLSV-CP (0.0952 ± 0.006) was observed in the Iranian isolates than the isolates from other countries except for Brazilian population (0.0769 ± 0.005). The possible explanation for this result could be that the Iranian isolates originated from...
one host (apple) (Table 1; Abtahi et al., 2019) and therefore host-adaptation as a means of the heterogeneous nature of virus populations (Garcia-Arenal et al., 2003) had not occurred.

The greatest nucleotide variation in Iranian isolates was found at the middle portion of ACLSV-CP (Fig. 3) similar to that reported from China (Chen et al., 2014), Japan (Nickel et al., 2018) and India (Rana et al., 2010). In contrast, the Brazilian isolates showed conservation at the middle of CP (Nickel et al., 2018). However, in the Iranian isolates the lowest nucleotide variation was observed in close to its 5′-terminal region similar to the Chinese, Brazilian, Japanese, and Indian isolates (Rana et al., 2010; Nickel et al., 2018).

The low genetic diversity between the Iranian ACLSV-CP sequences suggests that negative selection restricted the number of molecular variants and that virus populations are stable. Further analysis showed that the dN/dS value for the ACLSV Iranian and worldwide (among our isolates) isolates were 0.047 and 0.066, respectively. This result was also confirmed by evaluation of selective constraints across codon sites by comparing rates of synonymous to non-synonymous substitutions. One positively-selected codon in the ACLSV-CP was detected. In contrast, 112-140 negatively selected sites were found in ACLSV-CP by four methods (Table 2). These data suggest that strong negative selection is acting on the preservation of the ACLSV genome. Due to multiple functions of CP in viruses including genome encapsidation and protection, transmission between plants, host and/or vector interactions, cell to cell movement, suppression of gene silencing, negative selection pressures play a critical role in genome stability. Chare and Holmes (2004) analyzed selection pressures in the capsid genes of plant RNA viruses and found that vector borne viruses are subjected to a greater selection than non-vectored viruses. The ds/ds value for the 10 non-vector-borne viruses was > 0.093 whereas, for ACLSV-CP worldwide isolates was 0.066 (Nickel et al., 2018) suggesting that the value is not absolute and depends on the virus species and virus-host interaction.

In conclusion, the present analysis thus confirms the existence of a new cluster (BR) in phylogeny among ACLSV-CP isolates and a new amino acid sequence combination was suggested for the phylogroup differentiation in this virus. This study also confirms that negative selections are likely the potential drivers of the dynamics of ACLSV molecular evolution.

References


\[40\] and Phe

\[75\] or Ser

\[40\] and Tyr

\[75\]) in the coat protein of apple chlorotic leaf spot virus are crucial for infectivity. Journal of General Virology, 88: 2611-2618.


تعیین مشخصات و تنوع زنتیکی زن پروتئین پوشه‌شی جدایی‌های ویروس لک‌بزرگ کلرتوکی سبب جدا شده از سبیل باغات استان‌های آذربایجان غربی و کردستان

کمال شجاعی، مهدی آذریار و محمد حاجی‌زاده

گروه گیاه‌پزشکی، دانشکده کشاورزی، دانشگاه کردستان، سنندج، ایران.

پست الکترونیکی نویسندگان: m.hajizadeh@uok.ac.ir

دریافت: ۱۳ مرداد ۱۳۹۸؛ پذیرش: ۵ آبان ۱۳۹۸

چکیده: در این پژوهش، ۱۷۴ نمونه برگی سبب از مناطق عمده کشت درختان سبب در غرب (استان کردستان) و شمال غرب (استان آذربایجان غربی) برای آبودیگی به ویروس لک‌بزرگ کلرتوکی سبب مورد بررسی قرار گرفتند. نتایج حاکی از آبودیگی ۱۷ نمونه (۹/۷ درصد از نمونه‌ها) به ویروسACL SV از جمله هوادست بود. این نتایج نشان می‌دهد که تنوع زنتیکی و آنالیز تباضای ان ویروس با نویسنده توانی زن کامل پروتئین پوشه‌شی از هشت جدایی ویروس در این پژوهش و جدایی‌های تعیین ممکن شده از ایران و مورد در باکل زن انجام شد. در بررسی تباضایی جدایی‌های ACLSV با سه گروه نشانه‌های B6، B6، SHZ و P205 براساس ترکیب آمیواسیدی در موقت‌های ۴۰، ۷۵ و ۷۹ نفکیک شدن. جدایی‌های این پژوهش با ترکیب آمیواسیدی سرین، ترورین و اسید گلوتامیک در گروه تباضایی نشانه‌های ACLSV فرار گرفتند. میزان تنوع زنتیکی جدایی‌های ایرانی حاکی از تنوع زنتیکی کم این ویروس در مقایسه با سایر کشورها به دلیل جمعیت بزرگی این ویروس بود. اقلاب کدون های پروتئین پوشه‌شی تحت کلونه متفاوت به دلیل کدون به دلیل کلونه ۹۲ که تحت فشار گزینه مثبت بوده است. این پژوهش نشان داد که فشار انتخاب منفی بیشتر تأثیر را در کامل زن این ویروس داشته است.

واژگان کلیدی: سبب، فشار انتخاب منفی، آنالیز تباضایی، ACLSV