Identification of Subgenomic DNAs Associated with *Wheat Dwarf Virus* Infection in Iran

Mohamad Hamed Ghodoum Parizipour¹, Amir Ghaffar Shahriari²*

¹Department of Plant Protection, Faculty of Agriculture, Agricultural Sciences and Natural Resources University of Khuzestan, Mollasani, Iran
²Department of Agriculture and Natural Resources, Higher Education Center of Eghlid, Eghlid, Iran.

*Corresponding author: Amir Ghaffar Shahriari, Higher Education Center of Eghlid, Eghlid, Iran. Tel/Fax: +98-7144534056. E-mail: shahriari.ag@eghlid.ac.ir

**Background:** *Wheat dwarf virus* (WDV) is a leafhopper-transmitted DNA virus which causes yellowing and stunting in wheat and barley fields leading to considerable crop loss around the world. Mainly, two host-specific forms of WDV have been characterized in wheat and barley (WDV-Wheat and WDV-Barley, respectively).

**Objectives:** This study was aimed to amplify, sequence and describe subgenomic DNAs (sgDNAs) associated with WDV infection among wheat and barley plants. The nucleotide sequence of sgDNAs were then compared to that of parental genomic DNAs (gDNAs) and the differences were shown.

**Materials and Methods:** A total of 65 symptomatic plants were surveyed for WDV infection using double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) and polymerase chain reaction (PCR). Rolling circle amplification followed by restriction analysis (RCA-RA) was applied to identify both gDNAs and sgDNAs in the infected wheat and barley plants. Nucleotide sequence of eight full-length WDV genomes and five sgDNAs were determined.

**Results:** Genomic sequences of WDV-Wheat and WDV-Barley isolates obtained in this study were identified as WDV-F and WDV-B, respectively, forming a separate cluster in the phylogenetic tree with the highest bootstrap support (100%). Sequence analysis of sgDNA molecules revealed that they have undergone different mutation events including deletions in viral genes, duplication of coding regions, and insertion of host-derived sequences.

**Conclusions:** The association of different types of sgDNAs were found in WDV-infected wheat and barley plants. The sgDNAs exhibited remarkable changes compared to their parental molecules and they might play a role in symptom severity, host genome evolution and emergence of new virus variants/species.

**Keywords:** Mastrevirus, Phylogenetic analysis, Rolling circle amplification

1. **Background**

*Wheat dwarf virus* (WDV, genus *Mastrevirus*, family *Geminiviridae*) cause stunting, mottling, yellowing or reddening in wheat and barley in certain parts of Europe, Asia and Africa (1). WDV has a single-stranded DNA genome, 2.6–2.8 kb in size, encapsidated into a twinned icosahedral particle (2). Its genome has four open reading frames (ORFs), two (*V1* and *V2*) on the virion-sense strand and two (*C1* and *C2*) on the complementary-sense strand encoding four proteins: movement protein (MP; *V1*), coat protein (CP; *V2*), and replication-associated proteins (RepA; *C1* and Rep; *C1*:C2) (3). Rep is expressed through splicing and fusion of *C1* and *C2* ORFs (4, 5). Two non-coding regions are found on the genome including large and short intergenic regions (LIR and SIR, respectively) which are responsible for regulating gene expression and replication. The origin of replication of virion-sense strand is characterized by a highly conserved stem-loop structure positioned within the LIR (6, 7). An oligonucleotide present in the virion anneals to its complementary sequence within the SIR and mediates replication of complementary-sense strand (8-10). The viral genome is replicated through two mechanisms: rolling circle replication (RCR) and recombination-dependent replication (RDR) (11-19).

In addition to the complete viral genome, subgenomic DNAs (sgDNAs) with relatively smaller sizes have...
been found in a number of geminiviral infections around the world (20-25). These molecules were shown to be associated with reduced viral replication, delayed symptom development and/or symptom remission in their hosts (26-34). In terms of size, sgDNAs are typically half the size of their corresponding complete genomes and they are generated through diverse events including recombination, deletion, duplication and insertion of the template viral DNAs (reviewed in 30). To ensure their replication, these molecules mostly contain intergenic regions with cis-acting sequences (25, 26). Particularly, sgDNAs have been detected in WDV-infected hosts originating from Sweden (22) and Germany (25). Although WDV infection has been frequently detected in cereal cultivations of Iran (18), association of sgDNAs with the viral infection has not been reported to date.

2. Objectives
The goal of this study was to identify any sgDNAs associated with WDV infection among wheat and barley plants in Iran. Description of these molecules was also included in the present study.

3. Materials and Methods

3.1. Sampling
Leaf samples collected in 2016 from wheat and barley plants showing typical symptoms of WDV infection in cereal fields of Ahvaz, Dezful and Abadan cities in Khuzestan province, southwest Iran were used to detect WDV infection. A total of 65 plants (43 wheat and 22 barley plants) were surveyed in this study (Table 2).

3.2. Double Antibody Sandwich ELISA (DAS-ELISA)
In order to detect WDV infection in collected samples, DAS-ELISA was applied using polyclonal antibodies of WDV kindly provided by Dr. Antje Habekuß (Julius Kühn Institute, Quedlinburg, Germany). The assay was performed as described by Clark and Adams (35).

3.3. DNA Extraction and PCR
Total DNAs were extracted from leaf samples using ZR Plant/Seed DNA MiniPrep™ kit (ZYMO RESEARCH CORP., USA) according to the protocol provided by the supplier, and subjected to polymerase chain reaction (PCR) in order to detect WDV DNAs. Another DNA extraction method using cetyltrimethylammonium bromide (CTAB) was applied to isolate total DNA from collected leafhoppers (36). PCR was performed on a thermal cycler (BIO-RAD, USA) in a volume of 25 mL containing 2.5 mL of 10X Dream Taq Green Buffer, 0.2 mM of each dNTP, ~1 mL DNA template, 0.2 mM of each primer (Table 2) and 1 U Taq DNA polymerase (Thermo Scientific, Germany). The mixture was heated for 2 min at 95 °C and subjected to a 32-cycle PCR program (30 s at 96 °C, 60 s at 60 °C and 5 min at 72 °C). The final cycle was followed by 10 min incubation at 72 °C. Approximately 5 mL of the PCR products were electrophoresed on a 1.0% agarose gel with TAE buffer (40 mM Tris-acetic acid, 0.1 mM EDTA, pH 8.2-8.4 at 25 °C), pre-stained with DNA Safe Stain (CINACLONE, Iran) and the gel was visualized under UV light. PCR products containing the amplified viral genome were purified and then subjected to nucleotide sequencing at BIONEER corporation (South Korea).

3.4. Rolling Circle Amplification, Cloning, and Sequencing
RCA method, based on the Phi29 DNA polymerase from Illuma™ TempliPhi 100 Amplification kit (GE Life Sciences, USA), was used to amplify the circular DNA molecules of WDV genome from total DNAs of WDV-infected samples as described by Haible et al. (33). RCA products were then digested using HindIII and EcoRI enzymes (Thermo Scientific, Germany) which were expected to be single-cutters according to the results of in silico analysis of Iranian WDV sequences using NEBcutter software (ver. 2) (http://nc2.neb.com/NEBcutter2/). The released DNA fragments with a size equal to or lower than ~2.7 Kb were recovered from
DNA fragment was released from the majority of RCA-isolates of WDV. When release the complete/incomplete gDNAs of the Iranian to RCA followed by digestion analysis to amplify and PCR-positive sample of each viral strain was subjected 4.1. Sequence Analysis of WDV Full-Length Genome

Time Reversible model test (38) with 100 bootstrap the Maximum Likelihood approach using General CLC Main Workbench software (ver. 7.6.2), and alignment of the sequences was performed using a comprehensive phylogenetic analysis. Multiple alignments of full-length genomes were conducted and a Maximum-Likelihood phylogenetic tree was constructed, based on which seven strains of worldwide WDV isolates (A-G) were identified (Fig. 2). These isolates were generally classified into two super-clades: WDV-Wheat and WDV-Barley. WDV-Barley super-clade consisted of strains A, B and G, whereas strains D, C, E, and F were placed in the WDV-Wheat super-clade. WDV sequences obtained in this study were grouped based on a host-dependent classification, thus wheat- and barley-derived WDV isolates were placed in WDV-Wheat and WDV-Barley super-clades, respectively (Fig. 2). Moreover, WDV-Wheat and WDV-Barley isolates characterized in this study formed a separate cluster in the phylogenetic tree with the highest bootstrap support (100%) and divided into WDV-F and WDV-B, respectively (Fig. 2).

3.5. Sequencing Analysis

Base calling was performed using DNASTar Lasergene suite (DNASTAR, Madison, WI, USA). SeqMan Pro software version 8 (DNASTAR, USA) was used to assemble the sequence reads for determination of the full-length sequences of WDV isolates. Complete and incomplete viral DNA sequences were deposited on NCBI database. The resulting genomic and subgenomic sequences were compared to other WDV isolates for which complete genome information was available in GenBank database. A total of 61 WDV full-length sequences, including 8 sequences determined in this study and 53 sequences from different sources and diverse regions previously deposited on GenBank were used to carry out a comprehensive phylogenetic analysis. Multiple alignment of the sequences was performed using CLC Main Workbench software (ver. 7.6.2), and the Maximum Likelihood approach using General Time Reversible model test (38) with 100 bootstrap replicates was used to construct a phylogenetic tree. One isolate of Oat dwarf virus (ODV) was included in the analysis as an outgroup.

4. Results

4.1. Sequence Analysis of WDV Full-Length Genome

PCR-positive sample of each viral strain was subjected to RCA followed by digestion analysis to amplify and release the complete/incomplete gDNAs of the Iranian isolates of WDV. When HindIII was used, a ~2.7 kb DNA fragment was released from the majority of RCA-amplified DNA molecules (Fig. 1A, C). Additionally, DNA fragments with smaller sizes (~1.3 kb) were observed in some HindIII-treated RCA products (Fig. 1). RCA-amplified DNA molecules which were not restricted by HindIII were subsequently subjected to EcoRI digestion, resulting in the release of viral gDNAs and sgDNAs (Fig. 1B, D). Eight full-length (~2.7 kb) and 5 incomplete (~0.75-1.3 kb) fragments from wheat and barley strains of WDV were cloned, sequenced, and designated. The sequences of full-length genomes of WDV were deposited in GenBank. WDV-Wheat was found in both wheat and barley plants while WDV-Barley was detected only in barley plants (data not shown). Multiple alignments of full-length genome sequences were conducted and a Maximum-Likelihood phylogenetic tree was constructed, based on which seven strains of worldwide WDV isolates (A-G) were identified (Fig. 2). These isolates were generally classified into two super-clades: WDV-Wheat and WDV-Barley. WDV-Barley super-clade consisted of strains A, B and G, whereas strains D, C, E, and F were placed in the WDV-Wheat super-clade. WDV sequences obtained in this study were grouped based on a host-dependent classification, thus wheat- and barley-derived WDV isolates were placed in WDV-Wheat and WDV-Barley super-clades, respectively (Fig. 2). Moreover, WDV-Wheat and WDV-Barley isolates characterized in this study formed a separate cluster in the phylogenetic tree with the highest bootstrap support (100%) and divided into WDV-F and WDV-B, respectively (Fig. 2).

Genome organization of WDV isolates obtained in this study was characterized. Each complete viral genome contained four ORFs encoding four proteins: MP (V1) and CP (V2) on the virion-sense strand and RepA (C1) and Rep (C1:C2) on the complementary-sense strand of the genome (Fig. 3). LIR sequence embodying a stem-loop motif which contains the conserved TAATTJAC sequence (7) and SIR sequence were identified in the viral genomes (Fig. 3).

4.3. Description of WDV sgDNAs

Five forms of sgDNAs were found to be associated...
with WDV infection in wheat and barley plants (Fig. 3). In case of WDV-Barley, two sgDNA molecules (MH513493 and MH513492) were described, both containing complete V1 and LIR sequences. Accession No. MH513493 contained a duplicated sequence originating from V1 which was positioned just downstream of the complete V1 (Fig. 3). In addition to complete V1 and LIR, the other subgenomic molecule (MH513492) showed sequence deletion at 3’end of C1 and, interestingly, insertion of a non-viral sequence originating from barley DNA. BLASTn results showed that this host-derived sequence has 89% nucleotide identity (coverage: 76%) to beta-ketoacyl-ACP synthase I (Kas12) gene (M60410) from barley genome. Three forms of sgDNAs (MH513496, MH513494 and MH513495) were identified in wheat plants infected with WDV-Wheat (Fig. 3). The first accession (MH513496) contained the complete sequence of V1 and LIR as well as C1 with a deletion at 3’end of the ORF. The second accession (MH513494) harbored a truncated LIR and an incomplete C1 with deleted sequences at 3’end of the ORF. The third accession (MH513495) consisted of complete C1 and LIR together with an incomplete C1:C2 showing deletion at 3’end of the ORF. None of the sgDNAs contained either wild-type or mutant SIR sequences (Fig. 2).

5. Discussion
Since the first report of WDV in Iran in 2011 (39), the virus has been isolated from diverse sources (19). Application of the RCA technique provided the opportunity for determining more isolates of WDV originating from different sources and various regions in Iran (18). WDV isolates from around the world have been subjected to phylogenetic analysis for more than a decade, revealing genetic variations among these viruses (1, 14-16, 18).

Figure 1. Gel electrophoresis pattern of RCA-RA products from WDV-infected samples collected from symptomatic wheat (A, B) and barley (C, D) plants. The restriction enzymes used to release viral fragments are indicated at the bottom of each gel panel. Marker: 1 Kb GeneRuler™ DNA ladder (Thermo Scientific, Germany)
Due to a long-term co-evolution between virus and host in regions shown to be the origin of cereal cultivation, it can be speculated that new strains of WDV have been emerging over time (18). Fertile Crescent is believed to be the origin of agriculture, and it has been shown to include some regions of Iran (40). It seems that WDV isolates originating from Iran have a remarkable genetic diversity as two novel strains (i.e., WDV-F and WDV-G) have been exclusively reported from Iran (18, 19). In the present study, a total number of eight WDV sequences including four sequences from wheat and four sequences from barley were determined and identified as WDV-F and WDV-B strains, respectively (Fig. 2). Furthermore, two super-clades, WDV-Wheat and WDV-Barley, and seven strains, A, B, C, D, E, F and G, were formed on the cladogram tree which was consistent with the results of previous analyses (18, 19).

Association of sgDNAs with the corresponding viral infections caused by plant DNA viruses has been reported before (20-25, 32). Although these molecules were associated with symptom amelioration in the virus-infected hosts in the majority of the studies (26-29), their presence in cereals naturally infected with WDV did not negatively influence symptom development (25). Similarly, we could not observe any symptom attenuation in WDV-infected wheat and barley plants containing sgDNAs. However, further greenhouse experiments are required to test the effect of sgDNAs on symptom severity in WDV-infected plants. The sgDNAs investigated here showed different mutation events including deletion (in C1, V1 and C1:C2), duplication (of V1) and insertion (of barley DNA) (Fig. 3), conforming to earlier reports (22, 25). All sgDNAs analyzed in this study contained either complete or incomplete LIR sequences (Fig. 3), as has been reported by other researchers (22, 25). Nevertheless, we could not identify any SIR sequence in the sgDNAs, suggesting that LIR sequence alone is adequate for replication of sgDNAs within WDV-infected plant cells. This is the first report of WDV-
derived sgDNAs lacking SIR sequences in natural infections. A similar case was reported about sgDNAs characterized in maize plants infected with *Maize streak mastrevirus* (21). Regarding the rapid emergence of new geminiviruses (41), these molecules might be serving as a genetic pool for producing new variants and/or species of these devastating viruses (30). It has been shown that sgDNAs sometimes contain exogenous sequences originating from another virus simultaneously infecting the plant cell (23) and/or from
the host genome (25). Similarly, we identified a sgDNA containing a barley-derived DNA sequence in WDV-infected barley plants (Fig. 3). The sgDNAs can be encapsidated in viral particles (21) and transmitted to other plants by the insect vector (26). Since these viral particles are not able to cause infection upon entering new host cells, the plant DNA fragments carried by them are released into nuclear space (21). This phenomenon has been considered to play a role in genome evolution of plant hosts of viruses over the long term (25). Further in planta experiments are necessary for identifying the potential effects of these molecules on other aspects of WDV life cycle, such as replication.

The occurrence frequency of sgDNAs are usually low among naturally infected plant hosts (30). In the present survey, of 41 WDV-infected samples (Table 1), only 4 samples were found to contain sgDNAs which showed an occurrence frequency of 9.75%. In a similar study carried out by Schubert et al. (25), a relatively higher frequency (29.62%) was reported for the occurrence of sgDNAs in WDV-infected plants. Patil et al. (30) concluded that these molecules are very rare or are present below the levels of detection in the natural host. Moreover, the frequency of sgDNA can also be influenced by the host plant as different rates of sgDNAs were recorded in various hosts (26, 30).

6. Conclusion

Different types of sgDNAs were found to be associated with WDV infection in wheat and barley plants. These molecules have undergone different mutation events including deletions in viral genes, duplication of coding regions, and insertion of host-derived sequences. The sgDNAs might play a role in symptom severity, host genome evolution and emergence of new virus variants/species.

Acknowledgments

This work was supported by funds from Khuzestan Agricultural Sciences and Natural Resources University, Mollasani, Iran.

References


18. Ghodoum Parizipour MH, Ramazani L, Sardrood BP. Temperature affected transmission, symptom development
27. Frischmuth T, Engel M, Jeske H. Beet curly top virus DI DNA-
26. Frischmuth T, Stanley J. Beet curly top virus symptom
24. Paximadis M, Rey MEC. Genome organization of Tobacco leaf
22. MacDonald H, Coutts RHA, Buck KW. Characterization of
21. Casado CG, Ortiz GJ, Padron E, Bean SJ, McKenna R,
20. Behjatnia SAA, Dry IB, Rezaian MA. Characterization and
19. Ghodoum Parizipour MH, Behjatnia SAA, Afsharifar A,
8
19. Ghodoum Parizipour MH, Behjatnia SAA, Afsharifar A,
Izadpanah K. Natural hosts and efficiency of leafhopper vector in transmission of Wheat dwarf virus. J Plant Pathol. 2016. doi:
10.4454/JPP.V98I3.022
20. Behjatnia SAA, Dry IB, Rezaian MA. Characterization and
006-0929-2
21. Casado CG, Ortiz GJ, Padron E, Bean SJ, McKenna R,
Aghbandje-McKenna M et al. Identification of a defective molecule derived from complete genomes of isolates of the
22. MacDonald H, Coutts RHA, Buck KW. Characterization of
a subgenomic DNA isolated from Triticum aestivum plants
infected with wheat dwarf virus. J Gen Virol. 1988;59(6):1339-
1344. doi: 10.1099/0022-1317-69-6-1339
23. Patil BL, Dutt N, Briddon RW, Bull SE, Rothenstein D, Borah
BK et al. Deletion and recombination events between the
DNA-A and DNA-B components of Indian cassava-infecting
virusres.2006.09.003
24. Paximadis M, Rey MEC. Genome organization of Tobacco leaf
curl Zimbabwe virus, a new, distinct monopartite begomovirus
3091
complete genomes of isolates of the Wheat dwarf virus from
ew geographical locations and descriptions of their defective
s11262-013-0989-9
26. Frischmuth T, Stanley J. Beet curly top virus symptom
amelioration in Nicotiana benthamiana transformed with a
naturally occurring viral subgenomic DNA. Virol. 1994;200:826-
830. doi: 10.1006/viro.1994.1251
27. Frischmuth T, Engel M, Jeske H. Beet curly top virus DI DNA-
mediated resistance is linked to its size. Mol Breed. 1997;3:213-
217.
Fauquet CM. Identification of a defective molecule derived
from DNA-A of the bipartite begomovirus of East African cassava
mosaic virus. Plant Pathol. 2006;55:2-10. doi: 10.1111/j.1365-
3059.2005.01289.x
29. Stanley J, Saunders K, Pinner MS, Wong SM. Novel
defective interfering DNAs associated with ageratum yellow
30. Patil BL, Dasgupta I. Defective interfering DNAs of
plant viruses. Critical Rev Plant Sci. 2006;25:47-64. doi:
10.1080/07352680500391295
31. Stanley J, Townsend R. Characterization of DNA forms
associated with cassava latent virus infection. Nucleic Acids.
32. Al Rwahnih M, Alabi OJ, Westrick NM, Golino D, Rowhani A.
Description of a novel monopartite geminivirus and its defective
subviral genome in grapevine. Phytopathol. 2017;107(2):240-
251. doi: 10.1094/phyto-07-16-0282-r
33. Haible D, Kober S, Jeske H. Rolling circle amplification
2006;135:9-16. doi: 10.1016/j.jviromet.2006.01.017
34. Owor BE, Shepherd DN, Taylor NJ, Edema R, Monjane AL,
Thomson JA et al. Successful application of FTA classic card
technology and use of bacteriophage phi29 DNA polymerase
for largescale field sampling and cloning of complete maize
streak virus genomes. J Virol Methods. 2007;140:100-105. doi:
10.1016/j.jviromet.2006.11.004
35. Clark MF, Adams AN. Characteristics of the microplate method
of enzyme-linked immunosorbent assay for the detection of
1317-34-3-475
36. Maixner M, Ahrens U, Seemuller E. Detection of the German
grapevine yellows (Vergilbungskrankheit) MLO in grapevine,
alternative hosts and a vector by a specific PCR procedure.
bf01874780
38. Tavaré S. Some probabilistic and statistical problems in the analysis
39. Behjatnia SAA, Afsharifar A, Tahan V, Amid Motlagh M,
Eini Gandomani M, Niazi A et al. Widespread occurrence and
molecular characterization of Wheat dwarf virus in Iran. Aust.
40. Harlan JR. Agricultural origins: Centers and noncenters.
41. Kundu JK, Gadiou S, Cervena G. Discrimination and genetic
diversity of Wheat dwarf virus in the Czech Republic. Virus.