Virology 478 (2015) 112–122

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Characterization of the complete genome of *Barley yellow striate mosaic virus* reveals a nested gene encoding a small hydrophobic protein

Teng Yan [a,1], Jing-Rong Zhu [a,1], Dianping Di [b,1], Qiang Gao [a], Yongliang Zhang [a], Aihong Zhang [b], Chong Yan [b], Hongqin Miao [b], Xian-Bing Wang [a,*]

[a] State Key Laboratory of Agro-Biotechnology, College of Biological Sciences, China Agricultural University, Beijing 100193, PR China
[b] Plant Protection Institute, Hebei Academy of Agricultural and Forestry Sciences, IPM Center of Hebei Province, Key Laboratory of Integrated Pest Management on Crops in Northern Region of North China, Ministry of Agriculture, Baoding 071000, Hebei Province, PR China

ARTICLE INFO

Article history:
Received 26 September 2014
Returned to author for revisions
23 December 2014
Accepted 24 December 2014
Available online 7 February 2015

Keywords:
Plant rhabdovirus
Barley yellow striate mosaic virus
Nested gene
Ancillary gene
Endoplasmic reticulum

ABSTRACT

Barley yellow striate mosaic virus (BYSMV), a member of the genus *Cytorhabdovirus*, causes serious crop losses in agriculture. Here, we have cloned the BYSMV-derived small interfering RNAs (siRNAs), assembled the siRNAs and used RT-PCR to reconstruct the BYSMV genome. The genome consists of 12,706 nucleotides and encodes ten predicted genes from the antigenomic strand. The major BYSMV structural proteins share identities ranging from 35% to 62% with northern cereal mosaic virus (NCMV) counterparts. A notable difference is that BYSMV contains three transcriptional units residing between the P and M genes compared with four units in the corresponding region of NCMV. Unexpectedly, the middle mRNA in this region encodes gene5 nested in an alternative frame within gene4 *via* a leaky scanning mechanism. The gene5 encodes a small hydrophobic protein targeting to the endoplasmic reticulum (ER). To our knowledge, this is the first report of nested gene in plant rhabdoviruses.

Introduction

The *Rhabdoviridae* is a large family consisting of diverse members of negative strand RNA viruses with broad host ranges that collectively encompass vertebrates, invertebrates and plants (Kuzmin et al., 2009). To date, nine genera of animal-infecting rhabdoviruses (*Vesiculovirus*, *Ephemerovirus*, *Lyssavirus*, *Perhabdovirus*, *Sigmavirus*, *Tibrovirus*, *Tupavirus, Sprivivirus and Novirhabdovirus*), and two genera of plant-infecting rhabdoviruses (*Cytorhabdovirus* and *Nucleorhabdovirus*), are recognized in the family of *Rhabdoviridae* (Jackson et al., 2005; Kuzmin et al., 2009; Mann and Dietzgen, 2014). These diverse rhabdoviruses cause important impact on human health, agriculture, and wildlife ecology (Kuzmin et al., 2009).

In the plant rhabdoviruses, the cytorhabdoviruses and nucleorhabdoviruses are primarily classified according to their sites of replication, morphogenesis, and virion maturation in the cytoplasm or nucleus of infected cells. More than 100 plant rhabdoviruses members have been described, but there are mostly based on electron microscopic observation of distinctive enveloped bullet-shaped or bacilliform particle (Jackson et al., 2005). Unfortunately, the lack of clear identified replication sites or genome sequences have resulted in an inability to assign more than 75 plant rhabdoviruses to a genus. Plant rhabdoviruses infect a wide range of monocot and dicot species, including some agriculturally important crops including rice, maize, wheat, potato, tomato, eggplant and lettuce (Jackson et al., 2005). Furthermore, most plant rhabdoviruses are transmitted *via* arthropod vectors, usually consisting of leafhoppers, planthoppers and aphids (Jackson et al., 2005; Mann and Dietzgen, 2014). It is therefore imperative to determine the sequence structure of diverse plant rhabdoviruses, and to study functional roles of viral proteins.

Among the ten recognized species of nucleorhabdoviruses, the complete genome sequences of eight viruses have been determined, including sonchus yellow net virus(SYNV) (Choi et al., 1994; Heaton et al., 1989; Zuidema et al., 1986), potato yellow dwarf virus (PYDV) (Bandyopadhyay et al., 2010), eggplant mottled dwarf virus (EMDV) (Pappi et al., 2013), maize fine streak virus (MFSV) (Tsai et al., 2005), maize mosaic virus (MMV) (Reed et al., 2005), maize Iranian mosaic virus (MIMV), rice yellow stunt virus (RYSV) (Huang et al., 2003) and taro vein chlorosis virus (TaVCV) (Revill et al., 2005). The cytorhabdovirus currently contains nine recognized species, of which only the complete genome sequences have been determined for Lettuce necrotic yellows virus (LNYV) (Dietzgen et al., 2006; Wetzel et al., 1994a,

* Correspondence to: State Key Laboratory of Agro-Biotechnology, College of Biological Sciences, China Agricultural University, Beijing 100193, P.R. China. Tel.: +86 10 62733313; fax: +86 10 62732012.
E-mail addresses: yantengforever@gmail.com (T. Yan), jingrongzhu5@gmail.com (J.-R. Zhu), chmrdv@163.com (D. Di), gaoqianngnuli@gmail.com (Q. Gao), zhangyongliang7@gmail.com (Y. Zhang), zhangaihong08@163.com (A. Zhang), yanchong986@126.com (C. Yan), miao5058345@163.com (H. Miao), wangxianbing@cau.edu.cn (X.-B. Wang).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.virol.2014.12.042
0042-6822/

1994b), lettuce yellow mottle virus (LYMoV) (Heim et al., 2008), northern cereal mosaic virus (NCMV) (Tanno et al., 2000), and strawberry crinkle virus (SCV). In addition, the complete genome sequence of a putative rhabdovirus, named as persimmon virus A (PeVA), has been determined and this virus shares 42% identity with LNYV in the conserved L protein(Ito et al., 2013). Generally, the sequenced plant rhabdoviruses have 12–14.5 kb genomes, and encode five major structural proteins [(the nucleocapsid protein (N), phosphoprotein (P), matrix protein (M), glycoprotein(G) and polymerase protein (L)] that are common to all rhabdoviruses and are organized in the conserved order 3′-N-P-M-G-L-5′ (Jackson et al., 2005; Kuzmin et al., 2009; Mann and Dietzgen, 2014). In addition, plant rhabdovirus genomes encode diverse accessory genes between N-P, P-M and/or G-L genes (Walker et al., 2011).

Barley yellow striate mosaic virus (BYSMV) was first detected in *Laodelphax striatellus* (*L. striatellus*) from Italy (Conti, 1969), and subsequently reported in other European countries, Africa, Australia, Iran and Syria (Izadpanah et al., 1991; Makkouk et al., 2004). The symptoms elicited by BYSMV include mosaics, chlorotic striations, stunting and head sterility in cereal and wheat species (Conti, 1969; Izadpanah et al., 1991). BYSMV is transmitted by *L. striatellus* in a propagative manner (Milne and Conti, 1986). In 2010, the 6171 nucleotide sequence of L gene from the BYSMV Iranian isolate was determined and the sequence was shown to have the closest relationship to NCMV (Almasi et al., 2010). Given that the two viruses share similar features including host ranges in the family Poaceae, and transmission by *L. striatellus* planthopper, BYSMV and NCMV may have evolved recently from a common ancestor. However, their serological relationships reveal that BYSMV and NCMV are distinct species (Milne et al., 1986).

In recent years, RNA silencing has been studied extensively as a conserved regulatory process in plants. During antiviral silencing, double-stranded RNA intermediates formed during replication of RNA viruses are recognized and processed into an abundant of overlapping population of viral siRNA (viRNAs) (Ding, 2010). Accordingly, cloned viRNAs can be analyzed and assembled into viral contigs and this provides a powerful method that has been exploited to identify new viruses and determine their genome sequences (Wu et al., 2010).

Recently, BYSMV was isolated from four different wheat fields in northern China (Di et al., 2014). In a prelude to understand the BYSMV genome sequence, BYSMV siRNAs was reassembled into some contigs, which then were used as reference sequences for primers used for RT-PCR of BYSMV genome sequence. The sequence analyses revealed that BYSMV and NCMV have a similar genome organization and the high sequence identity of their major structural proteins indicates that the two viruses are more closely related to each other than to other plant rhabdoviruses. However, some unexpected variation was observed in organization of the ORFs encoding the four ancillary proteins located in the junctions between the P and M genes of BYSMV and the similar NCMV region. The four ORFs of BYSMV are organized into three transcriptional units, in which gene5 is nested within gene4 in an alternative reading frame suggesting that the two genes are expressed from the same mRNA and that gene 5 might be expressed by a leaky ribosome scanning mechanism.

Results

Biological properties of BYSMV

In the spring of 2014, the BYSMV was reported in some wheat fields from Hebei Province in northern China (Di et al., 2014). BYSMV is transmitted specifically by the planthopper (*Laodelphax striatellus Fallen*) in a circulative and propagative manner (Milne and Conti, 1986), and as reported previously the resulting BYSMV-infected wheat (*Triticum aestivum* L.) first developed yellow spot at

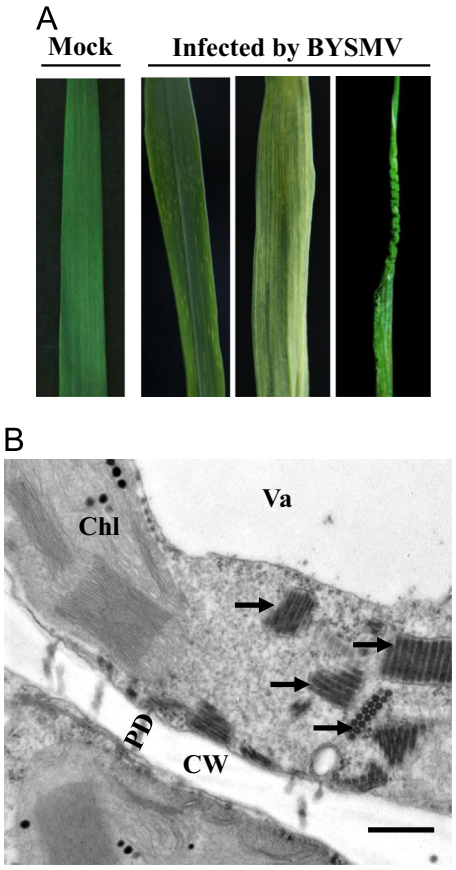

Fig. 1. The symptoms of BYSMV infected wheat plants and cells with virus inclusions. (A) The symptoms of BYSMV-infected wheat plants at 21 dpi. (B) Electron micrograph of thin section of infected wheat cells showing BYSMV particles (black arrow), exclusively in the cytoplasm of infected cells of wheat plants. CW=cell wall, Chl=chloroplast, Va=Vacuole, PD=plasmodesmata. Bar represents 500 nm.

9 dpi, that then turned into chlorotic striations and pronounced twisting of the flag leaves of the dwarfed plants (Fig. 1A). A large number of bacilliform particles with size of 315–353 × 46–57 nm (n=50) were observed by transmission electron microscopy in the cytoplasm of ultra-thin section infected cells (Fig. 1B). Three pairs of primers were designed using the sequence of the L gene of the BYSMV isolate Zanjan-1, and used to amplify the L gene of the Hebei BYSMV isolate. Sequencing results of the resulting polymerase gene (L) fragments were shown to share 96% nucleotide sequence identity with the L gene of BYSMV isolate Zanjan-1 (Fig. 2, fragments G, H and I), so the infected wheat plants were used for small RNA cloning and sequence determination of the remainder of the BYSMV genome.

The complete genome sequence of BYSMV

To rapidly obtain the complete genome sequence of BYSMV, we conducted deep sequencing of small RNAs from healthy and BYSMV infected wheat and assembled thirteen contigs from overlapping viRNAs that were only present in the infected wheat RNAs (data not shown). Our BLAST results showed that ten of the contigs shared about 96% identity with the reported L gene of the BYSMV isolate Zanjan-1 (data not shown). The other three contigs shared 40–50% with nucleotides 114–410, 2650–2930, and 5130–5780 of NCMV, but not with any known host genome sequences (Fig. 2, fragments A, B and C). Given that BYSMV is more closely related to NCMV than other plant rhabdovirus, we speculated that the three contigs were derived from BYSMV gRNA. Therefore, we designed primers based on the assembled