Virology 524 (2018) 32–44

Contents lists available at ScienceDirect

Virology

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

Whole genome deep sequencing revealed host impact on population structure, variation and evolution of *Rice stripe virus*

Lina Lu[a], Sanling Wu[b], Jun Jiang[c], Jingting Liang[d], Xueping Zhou[a,e,*], Jianxiang Wu[a,**]

[a] State Key Laboratory of Rice Biology, Institute of Biotechnology, Zhejiang University, Hangzhou, Zhejiang 310058, PR China
[b] Analysis Center of Agrobiology and Environmental Sciences, Faculty of Agriculture, Life and Environment Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, PR China
[c] Kaifeng Xiangfu Institute of Agricultural Sciences, Kaifeng, Henan 475100, PR China
[d] Department of Applied Biological Science, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, Zhejiang 310058, PR China
[e] State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, PR China

ARTICLE INFO

Keywords:
Rice stripe virus
High-throughput deep sequencing
Population structure
Genetic variation
O. sativa
N. benthamiana

ABSTRACT

High-throughput deep sequencing and variant detection showed that variations of *Rice stripe virus* (RSV) populations obtained from small brown planthopper-transmitted rice plants and sap-inoculated *N. benthamiana* plants were single nucleotide polymorphisms (SNPs) and insertion-deletions (InDels). The SNPs were more uniform across RSV genome, but InDels occurred mainly in the intergenic regions (IRs) and in the 5′ or 3′ noncoding regions. There were no clear patterns of InDels, although the inserted sequences were all from virus itself. Six, one, and one non-synonymous substitutions were respectively observed in the RdRP ORF, IR and the movement protein ORF. These non-synonymous substitutions were found to be stable, resulting in new consensus sequences in the NBL11 RSV population. Furthermore, the numbers of SNPs and InDels in RSV genome from *N. benthamiana* plants were much higher than that from *O. sativa* plants. These differences are likely caused by selection pressures generated by different host plants.

1. Introduction

Rice stripe virus (RSV) causes severe disease in rice fields and significant losses of rice production in many East Asian countries. RSV is transmitted in fields by *Laodelphax striatellus* Fallen (small brown planthopper [SBPH]) in a persistent, circulative-propagative and transovarial manner (Falk and Tsai, 1998; Xiao et al., 2010). RSV-infected plants often display chlorosis and necrosis in emerging leaves, and stunting of the plant (Huang et al., 2015). RSV is the type member of genus *Tenuivirus* and produces filamentous ribonucleoprotein particles (RNPs). RSV genome consists of four single-stranded RNA segments, designated as RNA1 (8970 nt), 2 (3400–3600 nt), 3 (2400–2600 nt) and 4 (2000–2200 nt) (Toriyama et al., 1994; Takahashi et al., 1993; Zhu et al., 1991, 1992). The 5′- and 3′-terminal 20 nucleotides of each RSV RNA segment are almost 100% complementary to each other. RSV RNAs form panhandle-like structures characteristic of negative-strand RNA viruses. It was suggested that these panhandle-like structures played important roles during RSV RNA transcription and replication, and might act as the recognition site for RNA dependent RNA polymerase (RdRp) (Takahashi et al., 1990). The RSV genome contains seven open reading frames (ORFs) and these ORFs are transcribed through an ambisense coding strategy (Ramirez and Haenni, 1994). The complementary sense of RNA1 encodes a 337 kDa RdRp (Toriyama et al., 1994). RNA2, 3 and 4 contain two ORFs each. In these ambisense RNAs, one ORF is located in the 5′ half of the virion-sense RNA, and the other ORF is located in the 5′ half of the virion-complementary sense RNA, separated by an intergenic region (IR) functioning in terminating RNA transcription (Hamamatsu et al., 1993; Qu et al., 1997; Wu et al., 2013). RNA2 encodes the Ns2 protein (22.8 kDa) from the virion-sense RNA and the Nsvc2 protein (94 kDa) from the virion-complementary sense RNA (Takahashi et al., 1993). RNA3 encodes the Ns3 protein (23.9 kDa) from the virion-sense RNA and the nucleocapsid protein (N, 35 kDa) from the virion-complementary sense RNA (Kakutani et al., 1991; Zhu et al., 1991). RNA4 encodes a non-structural disease-specific protein (SP, 20.5 kDa) from the virion-sense RNA and the Nsvc4 protein (32 kDa) from the virion-complementary

* Corresponding author at: State Key Laboratory of Rice Biology, Institute of Biotechnology, Zhejiang University, Hangzhou, Zhejiang 310058, PR China.
** Corresponding author.
E-mail addresses: lulina2008.cool@163.com (L. Lu), wusanling@zju.edu.cn (S. Wu), kfxnks@163.com (J. Jiang), 1393376050@qq.com (J. Liang), zzhou@zju.edu.cn (X. Zhou), wujx@zju.edu.cn (J. Wu).

https://doi.org/10.1016/j.virol.2018.08.005
Received 13 March 2018; Received in revised form 5 August 2018; Accepted 6 August 2018

sense RNA (Kakutani et al., 1990).

To date, the functions of RSV-encoded proteins remained largely uncharacterized. RSV RdRp is critical for viral replication, and its C-terminal region was shown to interact with a 50 amino acid (aa) region at its N-terminus (Shigemitsu, 1986; Zhao et al., 2015). The *Ns2* gene encodes a membrane-associated protein that may participate in recruiting or manipulating nucleolar functions to promote virus systemic infection through interactions with fibrillarin (Zheng et al., 2015). The *Nsvc2* gene encodes a polyglycoprotein localized exclusively to the ER membrane in *Spodoptera frugiperda* (Sf9) insect cells. This polyglycoprotein could also target Golgi bodies in plant cells and control RSV particle complex maturation during *L. striatellus* transmission (Yao et al., 2014). The Ns3 protein was shown as an RNA silencing suppressor, capable of binding both single-strand and double-strand siRNAs, and suppressing both local and systemic RNA silencing in plant (Xiong et al., 2009). The SP protein is known to play important roles in RSV pathogenesis (Toriyama, 1986; Kong et al., 2014). The RSV SP can form cytoplasmic inclusions in various cell types in viruliferous *L. striatellus*. The RSV ribonucleoprotein particles often closely associated with SP-specific inclusions in cells of viruliferous *L. striatellus*, through a direct interaction with RSV nucleoprotein (Wu et al., 2014). The Nsvc4 protein is known as the viral movement protein (MP) (Xiong et al., 2008).

Elucidation of the extent and structure of virus genetic diversity, and the magnitude and impact of population bottlenecks on virus evolution is crucial for understanding virus infection in nature (McCrone and Lauring, 2018). Plant viruses are known to utilize different mechanisms to generate large amounts of genetic diversities both within the same host species and between different host species (Nagy and Simon, 1997; Simmons et al., 2012). For example, virus replication by error-prone replicase often generates numerous mutant virus in virus population. This phenomenon was regarded as the quasispecies effect or the survival of the flattest (Nimwegen et al., 1999; Elena et al., 2008). Viral quasispecies can also be generated through virus and host plant interactions (Schneider and Roossinck, 2001). Studies have shown that the quasispecies phenomenon is one of the main reasons why plants lose their resistance to virus infections (Schneider and Roossinck, 2001). It has also been suggested that the extent and the structure of viral genetic variation in a specific host plant is determined mainly by point mutations and nucleotide deletions and/or insertions (Roossinck, 1997; Duffy et al., 2008; Acosta-Leal et al., 2011). Like other RNA viruses, RSV uses an error-prone replicase during its replication. Together with a high rate of replication in cells, variants with specific point mutation(s) and deletion/insertion(s) can quickly accumulate in cells, leading to a rapid evolution of RSV population in field.

Numerous studies have been performed in the past two decades to determine the genetic diversity and population structure of RSV (Zhu et al., 1992; Wei et al., 2009; Huang et al., 2013, 2015). Phylogenetic analyses of four RSV genes of RSV isolates obtained from different regions of China, and RSV isolates published elsewhere showed that these isolates could be divided into two subtypes (Wei et al., 2009). Phylogenetic analyses using full length RSV genomic RNA sequences from the tested 33 RSV isolates also indicated that RSV isolates could be divided into two subtypes, except isolate CX. Based on these analyses, all the isolates collected from eastern China, Japan, and Korea formed one monophyletic clade, while all the isolates, except isolate CX, collected from the Yunnan Province, China, formed the other monophyletic clade (Huang et al., 2013). In addition, negative pressures were detected on all seven RSV genes, especially the RdRp gene (Wei et al., 2009; Huang et al., 2013). Deep sequencing technology has been used to study highly variable viral populations and to characterize rapidly evolving viral quasispecies (Henn et al., 2012; Simmons et al., 2012). We have previously analyzed the genetic diversity of RSV in a field-infected rice plant and in RSV-infected first passage *N. benthamiana* plants using high-throughput sequencing technology, and found that the populations from *N. benthamiana* plants had more point mutations than that from the field-infected rice plant (Huang et al., 2015). However, the biological significances of these mutations and whether the RSV population structure is stable remain unknown.

The evolutionary hypothesis tested in the study is whether host and non-host plants have significant impacts on RSV population structure, genetic variation and evolution. Our results showed that RSV genetic variation was due mainly to single nucleotide polymorphisms (SNPs) and insertion-deletions (InDels). The SNPs were more uniform across the RSV genome, but the InDels occurred mainly in the intergenic regions (IRs) and the 5′ or 3′ noncoding regions. Our study also found eight near fixed variants in the RSV-infected *N. benthamiana* plants after eleven successive passages through mechanical inoculations, and the mutations appeared to cause more efficient virus accumulation and an enhanced pathogenicity in *N. benthamiana*. The numbers of SNPs and InDels in RSV genome obtained from *N. benthamiana* were greater than that from *O. sativa,* indicating that selection pressures generated by host and non-host plants were very different. Analyses of the polymorphic patterns of these variations should help us to better understand RSV epidemiology, population structure, and the underlying evolutionary mechanisms that act upon natural RSV populations.

2. Results

2.1. RSV inoculation and high-throughput sequencing

Changes of RSV populations in rice (host) or *N. benthamiana* (non-host) plants were investigated in this study using the same RSV original source (OSL1) (Fig. 1). Results from our preliminary experiments indicated that using the transmission protocol described in the Materials and Methods section, about 90% *L. striatellus* became viruliferous after 2 day feeding on the RSV-infected rice plants. Approximately 40% of *L. striatellus*-inoculated rice plants developed RSV symptoms by 30 dpi, and the symptoms on the OSL2 to OSL11 plants were similar to that on the OSL1 plants in three different experiments (data not shown). Symptoms on the NBL1, NBL2, NBL11 and NBL12 plants and the percentage of RSV-infected plants were monitored at 20 and 36 dpi (Fig. 2A). The NBL1 plants showed mosaic on young leaves by 13 dpi followed by downward leaf curling and plant stunting by 20 dpi (Fig. 2A). The RSV infection rate in the NBL1 plants was 100% (Fig. 2B). Development of RSV symptoms on the NBL2 plants was significantly delayed. Mild mosaic was not seen on the NBL2 plant leaves till 31 dpi followed by mild plant stunting at 36 dpi (Fig. 2A). The percentage of RSV-infected NBL2 plants was only about 15% (Fig. 2B). Mosaic on the NBL3 and NBL4 plants appeared at 20 and 18 dpi, and the percentage of RSV-infected plants gradually increased to 25% and 54% (Fig. 2). Mosaic on NBL5 to NBL12 plants was observed at about 12 dpi and the percentages of RSV-infected plants were all above 85% (Fig. 2). RSV symptoms on the NBL11 plants were slightly milder compared with the NBL1 plants at 20 dpi but became similar to that shown by the NBL1 plants by 36 dpi (Fig. 2A). To determine the accumulation level of RSV in the NBL1 and NBL11 plants, upper systemically infected leaves were collected from these plants at 36 dpi and analyzed for RSV genomic RNAs through Northern blot and qRT-PCR. Results of these assays showed that the accumulation levels of RSV genomic RNAs were similar in the NBL1 and NBL11 plants (Fig. 3A and B). Western blot assays using polyclonal antibodies against RSV Ns2, Ns3, CP and SP showed that the accumulation level of SP was the highest among the four proteins in the NBL1 and NBL11 plants (Fig. 3C). The expression patterns of seven RSV genes in rice and *N. benthamiana* plants were similar and the most abundant transcript in these two plants was Ns3, followed by SP (Fig. 3B and D).