Virus Research 201 (2015) 79–84

Contents lists available at ScienceDirect

Virus Research

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

Population structure of blueberry mosaic associated virus: Evidence of reassortment in geographically distinct isolates

Thanuja Thekke-Veetil[a], James J. Polashock[b], Mojca V. Marn[c], Irena M. Plesko[c], Annemiek C. Schilder[d], Karen E. Keller[e], Robert R. Martin[e], Ioannis E. Tzanetakis[a,*]

[a] Department of Plant Pathology, Division of Agriculture, University of Arkansas System, Fayetteville, AR 72701, United States
[b] USDA-ARS, GIFVL, Chatsworth, NJ 08019, United States
[c] Agricultural Institute of Slovenia, Hacquetova 17, Ljubljana, Slovenia
[d] Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI 48824, United States
[e] USDA-ARS, Corvallis, OR 97330, United States

ARTICLE INFO

Article history:
Received 15 October 2014
Received in revised form 19 February 2015
Accepted 22 February 2015
Available online 28 February 2015

Keywords:
Diversity
Genetic variation
Selection
Ophiovirus
Reassortment

ABSTRACT

The population structure of blueberry mosaic associated virus (BlMaV), a putative member of the family *Ophioviridae*, was examined using 61 isolates collected from North America and Slovenia. The studied isolates displayed low diversity in the movement and nucleocapsid proteins and low ratios of non-synonymous to synonymous nucleotide substitutions, indicative of strong purifying selection. Phylogenetic analyses revealed grouping primarily based on geography with some isolates deviating from this rule. Phylogenetic incongruence in the two regions, coupled with detection of reassortment events, indicated the possible role of genetic exchange in the evolution of BlMaV.

1. Introduction

Blueberry mosaic has been reported from several regions around the world (Martin et al., 2012). Diseased plants display mild to brilliant mosaic and mottling on the foliage with colors of yellow, yellow–green and sometimes pink (Varney, 1957) as well as delayed fruit ripening, reduced yield and fruit quality (Ramsdell and Stretch, 1987). Recently, a new member of the *Ophioviridae*, blueberry mosaic associated virus (BlMaV), was detected in all symptomatic material tested (Thekke-Veetil et al., 2014). Ophioviruses are a group of negative-strand RNA viruses with segmented genomes comprised of three or four RNAs. The genome of BlMaV is comprised of three RNAs (RNAs 1–3) encoding four proteins in the viral complementary strand (Fig. 1). RNA 1 encodes a 272 kDa replicase and a 23 kDa protein of unknown function. RNAs 2 and 3 code for a 58 kDa movement protein (MP) and 40 kDa nucleocapsid protein (NP) respectively (Thekke-Veetil et al., 2014).

Processes that increase genetic variability are key drivers of evolution as they provide viruses the tools to mediate the ever evolving arms race between pathogen and host. RNA viruses generally display elevated genetic variation because of their error-prone replication and genetic exchange through recombination or segment reassortment (Domingo and Holland, 1994; Nagy, 2008). Distinct genetic variants in a virus population may differ in severity of symptom expression, in their host range or in the transmission efficiency by their biological vectors. Examples are the mild and severe strains of citrus tristeza virus, cucumber mosaic virus and tomato spotted wilt virus which cause symptoms ranging from mild mosaic to plant death (Aramburu et al., 2010; Lin et al., 2003; Roistacher and Moreno, 1991). Analysis of virus population structure allows the grouping of isolates with unique molecular and possibly biological characteristics and provides information that could be used to predict the durability of genetic or transgenic resistance. Therefore, knowledge of the virus population structure is important in understanding disease development and implementing effective management strategies including development of reliable diagnostic assays for screening propagation material. In addition, information on the virus population structure provides insights on the virus dynamics and factors influencing virus evolution (Garcia-Arenal and McDonald, 2003).

There have been several reports on the population structure of positive-strand RNA viruses (Alabi et al., 2011; Fargette et al., 2004;

* Corresponding author. Tel.: +1 479 575 3180; fax: +1 479 575 7601.
E-mail address: itzaneta@uark.edu (I.E. Tzanetakis).

http://dx.doi.org/10.1016/j.virusres.2015.02.022
0168-1702/

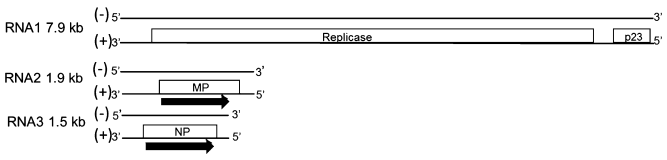

Fig. 1. Genome organization of blueberry mosaic associated virus. (−): negative-sense RNA (viral RNA); (+): positive-sense RNA; rectangles represent open reading frames and protein products encoded in each RNA are indicated. MP: movement protein; NP: nucleocapsid protein. The regions sequenced and analyzed are indicated by arrows.

Lopez et al., 1998; Poudel et al., 2012; Poudel and Tzanetakis, 2013; Rubio et al., 2001; Thekke-Veetil et al., 2013; Tomimura et al., 2004; Turturo et al., 2005; Vigne et al., 2004; Wang et al., 2011). On the other hand, research on negative-strand plant RNA viruses is limited (Callaghan and Dietzgen, 2005; Miranda et al., 2000; Tentchev et al., 2011; Wei et al., 2009) and in the case of ophioviruses they are limited to mirafiori lettuce big-vein virus (MiLBVV) and citrus psorosis virus (CPsV) (Alioto et al., 2003; Martin et al., 2006; Navarro et al., 2005). This communication reports on the population structure of 61 BlMaV isolates obtained from Canada, Slovenia and the United States and analyzes the potential role of evolutionary forces shaping the genetic structure of the population.

2. Methods

2.1. Virus isolates, amplification and sequencing

Isolates were collected from symptomatic plants between 2009 and 2013 (Supplementary Table 1) from British Columbia in Canada (BC1 to 2), Slovenia (Slo1 to 9) and the United States (Arkansas, AR1 to 5; Kentucky, KY1 to 4; Michigan, MI1 to 6; New Jersey, NJ1 to 11; Oregon, OR1 to 22). Presence of BlMaV in the samples was confirmed by RT-PCR as described (Poudel et al., 2013) using the virus specific primer set [(+) 5′-CCCCGGATTGTTCCCGAACCTT-3′; (−) 5′-ATTCTGACCACCCCGGTCTTGT-3′] designed from the MP region, with the PCR program consisting of initial denaturation at 94 °C for 3 min followed by 40 cycles of 94 °C for 20 s, 54 °C for 20 s and 72 °C for 40 s.

The amplification of the two regions used in the study was done as follows: cDNAs were diluted 1:2 in water to minimize the effect of potential inhibitors in downstream reactions. One thousand two hundred (1200) nucleotide (nt) regions in the MP and NP genes were amplified (Supplementary Table 2) using either *Taq* DNA polymerase (Genscript Corporation, NJ) or Phire Hot Start II DNA Polymerase enzymes (Thermo Fisher Scientific Inc., MA) according to manufacturers' instructions. More than one primer pairs were required to amplify both regions in all isolates due to the sequence variations at the priming sites. Amplicons were cloned and at least three recombinant clones were sequenced in both directions and assembled essentially as described earlier (Thekke-Veetil et al., 2013). The sequences were deposited in GenBank under accession numbers KJ849092-KJ849150, KP793911-2 (MP region) and KJ849151-KJ849209, KP793913-4 (NP region).

2.2. Phylogenetic analysis, selection pressure and detection of recombination events

Nucleotide sequences were aligned using Clustal W (Thompson et al., 1994). Variations in nt and predicted amino acid (aa) sequences were determined using BioEdit (Hall, 1999). The phylogenetic trees were constructed using the maximum likelihood method with 1000 bootstrap replicates in MEGA v.6 (Tamura et al., 2013). The trees were redrawn after collapsing all the branches with <70% bootstrap support as those were considered unreliable.

SNAP (Korber, 2000) was used to estimate d_{Ns}/d_S where d_{Ns} represents the average number of non-synonymous substitutions per non-synonymous site and d_S represents the average number of synonymous substitutions per synonymous site (Nei and Gojobori, 1986). Depending on the d_{Ns}/d_S ratio, the selection pressure on the genes was considered negative or purifying ($d_{Ns}/d_S < 1$), neutral ($d_{Ns}/d_S = 1$), or positive or diversifying ($d_{Ns}/d_S > 1$). Cumulative behavior of codon mutations and distribution of mutations in the regions of analysis were also obtained from SNAP. Screening for positive selection on codons was conducted using three different codon-based maximum likelihood methods (SLAC, FEL and REL) (http://www.datamonkey.org; Kosakovsky Pond and Frost, 2005). Synonymous codon usage bias was measured by calculating the effective number of codons (ENC; Wright, 1990) used in a gene by applying DnaSP v5 (Librado and Rozas, 2009). ENC values range from 20 (the number of amino acids), indicative of maximum bias as only one codon is used from each synonymous-codon group, to 61 (the number of sense codons), which indicates no codon-usage bias where all synonymous codons for each amino acid are equally used.

Putative recombination/reassortment events in the MP/RNA2 and NP/RNA3 were investigated using nine algorithms employed in RDP4 (Martin et al., 2010). Analyses were conducted on the two regions separately and also after concatenating the sequences of both regions into one for each isolate. Only events supported by at least five RDP4-implemented algorithms were considered plausible recombination/reassortment events. The recombinants/reassortants were compared with the parental sequences using SimPlot software (Lole et al., 1999). All potential reassortants were reanalyzed for the presence of mixed infection after sequencing at least 10 additional plasmids for the respective regions where reassortments were reported. When sequence diversity indicated infection by multiple isolates, an additional 10 clones were sequenced as confirmation.

3. Results

3.1. Population structure

Analysis of over 20% of the virus genome (2400 nt per isolate; regions indicated in Fig. 1) revealed significant sequence conservation. For the MP region, overall sequence variation among isolates peaked at 9% and 4% in the nt and aa sequences respectively (Supplementary Tables 3 and 4). Variation reached 13% and 5%, respectively, in the nt and aa sequences of the NP (Supplementary Tables 5 and 6). Two isolates from New Jersey (NJ10 and NJ11) were unique in the NP region with a Gly_{95} insertion and exhibited 87–90% nt identity to the rest of the isolates.

Phylogenetic trees were generated from both the nt and aa sequences to assess the evolutionary relationships of the isolates. Trees derived from nt sequences were more informative compared to those derived from aa and are discussed here. Isolates mostly clustered together based on their geographical origin with some deviating from this structure (Fig. 2). Examples are Slo1, Slo8, OR13, OR14, OR15 and KY4 for MP region, and Slo8, OR14 and OR15 for the NP region. A few NJ isolates were phylogenetically distant from most other NJ isolates and formed separate clades; NJ8-1 and NJ9 for the MP region and NJ10 and NJ11 for the NP region. All analyzed geographical populations represented isolates collected from different cultivars indicating that time or genetic makeup of the host were not related to the grouping of the BlMaV isolates. In addition, topologies of MP and NP trees were different which suggested phylogenetic incongruence of the isolates.

3.2. Selection constraints, recombination and reassortment

The rate of molecular evolution differs between genes and therefore d_{Ns}/d_S was calculated, showing stringent purifying