Virus Research 169 (2012) 272–275

Contents lists available at SciVerse ScienceDirect

Virus Research

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

Short communication

Population structure of *Blackberry yellow vein associated virus*, an emerging crinivirus☆

Bindu Poudel[a], Sead Sabanadzovic[b], Jozef Bujarski[c,d], Ioannis E. Tzanetakis[a,*]

[a] Department of Plant Pathology, Division of Agriculture, University of Arkansas, Fayetteville, AR 72701, USA
[b] Department of Biochemistry, Molecular Biology, Entomology, and Plant Pathology, Mississippi State University, Mississippi State, MS 39762, USA
[c] Department of Biological Sciences, Northern Illinois University, DeKalb, IL 60115, USA
[d] Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland

ARTICLE INFO

Article history:
Received 23 May 2012
Received in revised form 9 July 2012
Accepted 12 July 2012
Available online 20 July 2012

Keywords:
Closteroviridae
Recombination
Selection pressure
Evolution

ABSTRACT

Blackberry yellow vein disease (BYVD), a disorder caused by virus complexes, has become a major threat to fresh market blackberry production in the United States. *Blackberry yellow vein associated virus* (BYVaV) is the most prevalent virus in the BYVD complexes; detected in about 50% of samples exhibiting typical disease symptoms. Thirty-four virus isolates infecting wild and cultivated blackberries were collected from several areas with high BYVD incidence. Sequence variability and virus evolution predictions were calculated for four genomic regions coding for six proteins and accounting for about 30% of the virus genome. Nucleotide diversity ranged between 7 and 12%, and all proteins studied were under negative selection. Several isolates were identified as potential recombinants suggesting that recombination might be a driving force behind BYVaV evolution.

Recent increase in blackberry acreage and expansion to new areas has led to an increased incidence of viruses and the emergence of a new disorder designated as blackberry yellow vein disease (BYVD; Martin et al., 2012). Martin et al. (2004) discovered a new crinivirus named *Blackberry yellow vein associated virus* (BYVaV) in all symptomatic BYVD plants. Further study determined that BYVaV is not the causal agent of the disorder and that BYVD is rather caused by diverse virus complexes (Susaimuthu et al., 2008). Notwithstanding, BYVaV is the most prevalent virus in the complexes; detected in several commercial nurseries scattered around the United States, all fresh market blackberry growing areas, and several wild blackberry populations (Susaimuthu et al., 2007; Poudel, 2011).

BYVaV is a member of the genus *Crinivirus*, family *Closteroviridae* and has a bipartite genome of 15.7 kb in size. The 7.8 kb RNA1 codes for a papain-like protease, methyltransferase, helicase and polymerase, all involved in virus replication; whereas the 7.9 kb RNA2 codes for proteins involved in encapsidation, transmission and movement (Tzanetakis et al., 2006). An isolate from an 'Apache' blackberry from South Carolina has been sequenced (Tzanetakis et al., 2006) and represents the virus type isolate; however, knowledge on additional BYVaV isolates is limited. Although asymptomatic in single infections (Susaimuthu et al., 2007), BYVaV interactions with other viruses indicate that disease symptoms are elicited upon infection with different viruses (Susaimuthu et al., 2008). It is therefore important to study the virus population structure and set the foundation to better understand BYVaV isolates/strains interactions with other viruses as a part of BYVD development. This communication provides a comprehensive analysis of the population structure of BYVaV with information on 34 isolates collected from vast geographical area that spans several hundred thousand km^2 (Table 1).

Young, fully developed leaves showing BYVD symptoms were collected from the wild in Arkansas and commercial fields in Arkansas, Georgia, Mississippi, North Carolina, and South Carolina. The Arkansas isolates were studied in more detail as the University of Arkansas has a leading breeding program for fresh market blackberries and infected material from the program could be the virus source of material grown around the world. Samples were processed within 24 h of collection. Total nucleic acid preparations were extracted as described by Tzanetakis et al. (2007) and subjected to reverse-transcription polymerase chain reaction (RT-PCR) with two sets of detection primers (Susaimuthu et al., 2006, 2008), so as to identify the highest possible number of BYVaV isolates. Positive samples were subjected to dsRNA extractions as described in Laney et al. (2011). All regions were amplified from reverse transcribed dsRNA templates essentially as described in Tzanetakis et al. (2005). The designation of the different regions studied, along with nucleotide coordinates are: 1 [polyprotein (1a): RNA1 nt 2384–3583], 2 [RNA-dependent RNA polymerase (RdRp);

☆ Sequences reported in this communication have been deposited in Genbank and have been assigned the accession numbers JQ56161–JQ56296.

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

0168-1702/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.virusres.2012.07.009

Table 1
Isolates used in population structure study of *Blackberry yellow vein associated virus*.

Isolate name	Wild/cultivated	Collection state
AR1	Cultivated	Arkansas
AR2	Cultivated	Arkansas
AR3	Cultivated	Arkansas
AR4	Cultivated	Arkansas
AR5	Cultivated	Arkansas
AR6	Cultivated	Arkansas
AR7	Cultivated	Arkansas
AR8	Wild	Arkansas
AR9	Wild	Arkansas
AR10	Wild	Arkansas
AR11	Wild	Arkansas
AR12	Wild	Arkansas
AR13	Wild	Arkansas
GA1	Cultivated	Georgia
GA2	Cultivated	Georgia
GA3	Cultivated	Georgia
MS1	Cultivated	Mississippi
MS2	Cultivated	Mississippi
MS3	Cultivated	Mississippi
NC1	Cultivated	North Carolina
NC2	Cultivated	North Carolina
NC3	Cultivated	North Carolina
NC4	Cultivated	North Carolina
NC5	Cultivated	North Carolina
NC6	Cultivated	North Carolina
NC7	Cultivated	North Carolina
SC1	Cultivated	South Carolina
SC2	Cultivated	South Carolina
SC3	Cultivated	South Carolina
SC4	Cultivated	South Carolina
SC5	Cultivated	South Carolina
SC6	Cultivated	South Carolina
SC7	Cultivated	South Carolina
SC8	Cultivated	South Carolina
SC0	Type isolate	South Carolina

RNA1 nt 6319–7518], 3 [p4 and part of heat shock protein 70 homolog (HSP70 h); RNA2 nt 309–1103] and 4 [part of the coat protein (CP) and minor coat protein (CPm); RNA2 nt 4709–6024] (Fig. 1).

Purified PCR products were cloned onto the TOPO-TA pCR2.1 vector (Invitrogen) and transformed into α-select chemically competent cells (Bioline). Selected recombinant plasmids were sequenced by Functional Biosciences, Inc. (Madison, WI) using universal M13 forward and reverse primers. A minimum of three independent clones were sequenced in both directions to obtain the consensus sequence for each region studied.

Sequence contigs were obtained with CAP3 (Huang and Madan, 1999). Diversity among isolates was studied using Bioedit (Hall, 1999). Selection pressure was calculated using the Synonymous/Non-Synonymous Analysis Program (SNAP; Korber, 2000). The program calculates synonymous/non-synonymous rates based on Nei–Gojobori method (Nei and Gojobori, 1986). The results were further validated by using the codon based *Z*-test, also implementing the Nei–Gojobori method (Nei and Gojobori, 1986)

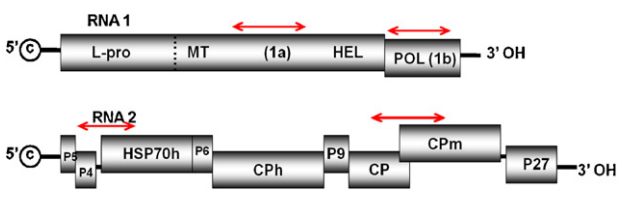

Fig. 1. Schematic representation of the BYVaV genomic regions used for the diversity study indicated by left-right arrows. *Abbreviations*: Pro: papain-like protease; MT: methyltransferase; HEL: helicase; Pol: RNA-dependent RNA polymerase; HSP70h: heat shock protein homolog; CPh: coat protein homolog; CP: coat protein; CPm: minor coat protein.

on MEGA v.5 (Tamura et al., 2011). Phylogenetic and molecular evolutionary analyses also were conducted with MEGA. The maximum likelihood method was used to construct the trees using the nucleotide sequences for the four regions studied implementing the Kimura 2-parameter model (Kimura, 1980). Trees were also constructed for six coding regions using the Poisson model for amino acid substitution (Zuckerland and Pauling, 1965). In both analyses, grouping robustness was determined after 1000 pseudoreplicates. Branches supported by <50% were collapsed as they were considered untrustworthy (Felsenstein, 1985). Recombination analysis was performed on RDP4 which implements eight algorithms to identify recombination sites (Martin et al., 2010). Recombination patterns were considered statistically significant events when detected by four or more programs (Martin et al., 2010).

Four regions, two from each genomic RNA, were analyzed, accounting for about 30% of the BYVaV genome (4512 nt/isolate). It is generally accepted that replication-associated proteins as those coded in RNA1 show a slower evolutionary rate compared to those coded by RNA2, involved in movement and genome protection (Dolja et al., 2006). On the other hand isolates of the monopartite closteroviruses and ampeloviruses tend to show higher nucleotide diversity toward the 5′ end of the genome (Roy et al., 2005; Herrera-Isidrón et al., 2009; Aguilar et al., 2003) but this is grossly understudied in criniviruses, a gap in knowledge this study aims to address.

The positioning of region 1 (polyprotein) was based on the presence of additions/deletions in the orthologous region of *Beet pseudo-yellows virus* (BPYV), the closest relative to BYVaV identified to date (Tzanetakis and Martin, 2004). Region 2 codes for the RdRp and as in the case of the polyprotein, is less studied in crinivirus population studies. Additional rationale for choosing RdRp is that our data on *Strawberry pallidosis associated virus* indicate significant diversity in this genomic region (Tzanetakis et al., unpublished). BYVaV code for two small hydrophobic peptides with identifiable transmembrane domains near the 5′ terminus of RNA2. This has not been observed for any other crinivirus, thus region 3 was designed to evaluate the presence of the peptides and the transmembrane domains in all isolates. The C-termini of the crinivirus coat proteins contain the conserved Arg and Asp residues, found in all filamentous RNA viruses, and are more conserved compared to the N-termini of the proteins (Tzanetakis et al., 2006). For this reason, region 4 was designed to include the CP/CPm junction and determine whether the two areas are under different selection pressures. Several primer set combinations were evaluated to determine the most efficient PCR amplification of the studied regions. Based on those experiments the regions analyzed are slightly altered from the original design. The data in this communication represents the minimum sequence length as amplified from all isolates.

The ratio of non-synonymous substitutions per non-synonymous site (d_{Ns}) over synonymous substitutions per synonymous site (d_S) was used as the indicator of protein selection pressure. Coding areas are under positive or diversifying selection when d_{Ns}/d_S is >1, neutral when $d_{Ns}/d_S = 1$, and negative (purifying) when d_{Ns}/d_S is <1. The d_{Ns}/d_S ratio was calculated pairwise for isolates and averaged for a particular region. Region 1, showed stability with maximum nucleotide and amino acid diversity of 8% and 6%, respectively, and the $d_{Ns}/d_S = 0.042$. Apparently, region 2 was under stringent purification selection. Although the nucleotide diversity among studied isolates reached 12%, the maximum amino acid diversity was only 5% and $d_{Ns}/d_S = 0.021$ (Table 2).

Region 3 covered two coding regions, p4 and the N-terminus of the HSP70 h. The maximum nucleotide and amino acid diversity of p4 reached 12% and 11%, respectively. Accordingly, the d_{Ns}/d_S value of 0.155 was highest of all regions studied. The presence of the p4 transmembrane domain was verified in all isolates. The N-terminus