Journal of Virological Methods 221 (2015) 15–21

Contents lists available at ScienceDirect

Journal of Virological Methods

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

Detection, discrimination and discovery of a new *Tobacco streak virus* strain

M. Dutta[a], A. Ali[b], U. Melcher[a,*]

[a] Department of Biochemistry and Molecular Biology, Oklahoma State University, Stillwater, OK, 74078-3035 USA
[b] Department of Biological Science, The University of Tulsa, OK, 74104, USA

Article history:
Received 25 March 2015
Accepted 29 March 2015
Available online 20 April 2015

Keywords:
Microbial forensic methods
454 sequencing
RT-qPCR
SYBR green

A B S T R A C T

Soybean plants that exhibited symptoms of virus infection were sampled from different counties of Oklahoma. These plants were tested serologically for 15 viruses known to infect soybean plants. Fifty-seven samples that exhibited typical virus-like symptoms did not test positive for any of the 15 viruses used in a dot-immunobinding assay (DIBA). Four samples were pooled and used for next generation sequencing using the 454-Roche protocol. Sequence and phylogenetic analysis of the sequences obtained revealed infection with a distinct strain of *Tobacco streak virus* (TSV). TSV was one of the 15 viruses initially tested for using DIBA and had tested negative. TSV belongs to the genus *Ilarvirus* and has been reported as a causal agent of bud blight in soybean crops in Brazil and the United States. Out of 10 reported primer pairs for TSV reverse transcription-polymerase chain reaction (RT-PCR), only two had the potential, based on sequence similarity, to amplify part of the genome of the distinct strain of TSV found in Oklahoma and only one was actually able to amplify the region. In this study, a new primer pair, specific to all known TSV and capable of amplifying the Oklahoma strain (TSV-OK), was designed from a highly conserved region of coat protein (CP) sequences and end-point PCR and quantitative RT-PCR detection methods were developed and their sensitivity assayed. This is the first report of specific primers designed from this highly conserved region in the CP of TSV for detection of TSV. Twenty-three of the 57 DIBA soybean samples that initially tested negative were retested with the new specific end-point PCR method and found positive for TSV infection.

1. Introduction

Study of diseases that may affect soybean is important to ensure good commercial crop yields. Mosaic, mottling, rugosity of leaves, mottling of seed coats, discoloration of seeds of soybean can be caused by strains of several viruses [*Tobacco streak virus* (TSV), *Soybean mosaic virus* (SMV), and *Bean pod mottle virus* (BPMV)] (Weiden and Ginsberg, 1994; Fagbenle and Ford, 1970; Ghanekar and Schwenk, 1974). During a survey of soybean fields in Oklahoma a number of soybean plant samples were tested serologically by dot-immunobinding assay (DIBA) against the antisera of 15 different viruses including TSV (Ali et al., 2015, in preparation). More than 50% of the collected samples were negative to all 15 tested viruses. In this work, some of the previously collected DIBA negative samples (Ali et al., 2015, in preparation) were tested by NGS, which identified a novel strain of TSV that infects soybean plants in Oklahoma. TSV has been reported to cause bud blight in Brazil and the United States (Rabedeaux et al., 2005). Bud blight causes pod necrosis and dieback of terminal portions of stems of a plant and has been observed in soybean in the past and in recent years. In the U.S., TSV was previously isolated from plants expressing symptoms of bud blight in commercial soybean fields in Wisconsin and Oklahoma (Rabedeaux et al., 2005). Bud blight has several additional known causal agents, including *Tobacco ringspot virus* (TRSV), *Tomato ringspot virus* (TomRSV), and SMV (Rabedeaux et al., 2005). Almost all soybean varieties tested have been susceptible (Ghanekar and Schwenk, 1974; Wang et al., 2005). Not much is known about the impact of TSV on soybean health and productivity and hence TSV infection merits further investigation (Rabedeaux et al., 2005).

Viruses within genus *Ilarvirus* family *Bromoviridae*, of which TSV is the type member, have positive-sense, single-stranded RNA

Abbreviations: TSV, tobacco streak virus; SMV, soybean mosaic virus; BPMV, bean pod mottle virus; DIBA, dot-immunobinding assay; TRSV, tobacco ringspot virus; TomRSV, tomato ringspot virus; MP, movement protei; CP, coat protein; SNSV, strawberry necrotic shock virus; PMoV, parietaria mottle virus; BCRV, blackberry chlorotic ringspot virus; EM, electron microscopy; NGS, next-generation sequencing; RdRp, RNA-dependent RNA polymerase; UTR, untranslated region.

* Corresponding author at: Department of Biochemistry and Molecular Biology, Oklahoma State University, 246 NRC, Stillwater, Oklahoma 74078-3035, USA. Tel.: +1 405 744 6210; fax: +1 405 744 7799.
E-mail address: ulrich.melcher@okstate.edu (U. Melcher).

http://dx.doi.org/10.1016/j.jviromet.2015.03.025
0166-0934/

tripartite genomes (RNA 1 to 3). RNA 1 and RNA 2 encode for the viral replication protein and RNA-dependent RNA polymerase (RdRp), respectively (Pallas et al., 2013). RNA 3 encodes the movement protein (MP) and a subgenomic RNA 4 that is translated into the coat protein (CP) (Pallas et al., 2013).

Currently, there is no specified level of sequence similarity to distinguish among species, strains or subgroups of ilarvirus members (Bujarski et al., 2012). Serological relatedness has been used to assign species to subgroups, and serological differences, host specificity or geographical location have been used to distinguish different species belonging to the same genus (Halk and Franke, 1983; Clement and Converse, 1986). RNA 2 of members of subgroups 1 and 2 have a unique second open reading frame absent in other ilarvirus subgroups (Pallas et al., 2013). The type isolate of TSV was originally isolated from white clover (TSV-WC). Species within subgroup 1 include TSV, *Strawberry necrotic shock virus* (SNSV), *Parietaria mottle virus* (PMoV), *Blackberry chlorotic ringspot virus* (BCRV) and the proposed species "Bacopa chlorosis virus" (Pallas et al., 2013).

TSV has a broad host range that includes members of *Solanaceae* and *Leguminosae*. TSV has been detected frequently amongst weeds bordering agricultural fields from where it serves as a source of inoculum that could infect crops. Thrips *Microcephalothrips abdominalis* and *Thrips tabaci* are thought to play roles in the transmission of TSV (Jones, 2005). The mechanism of transmission is described mainly as mechanical in nature, where TSV-contaminated pollen grains are introduced into the plant through wounds (Sharman and Thomas, 2013; Klose et al., 1996; Adams et al., 2009).

In the diagnosis of novel, unidentified or unusual viral plant diseases, causative agents may go undetected when using methods such as PCR or ELISA, because these tests are very specific to a particular species or even strain of a virus. Also, given the fast rate of evolution of many RNA viruses, methods to be developed need to detect and identify viral variants in a broader sense (Zheng et al., 2008). Methods such as electron microscopy (EM) or sap inoculation of plant virus indicator species do not allow accurate species level diagnosis. This research explores Next-generation sequencing (NGS) which offers an alternative solution where sequences are generated in a non-specific fashion and identification is made based on similarity alignment against GenBank data (Adams et al., 2009; Ho and Tzanetakis, 2014).

TSV is well known to be genetically diverse (Clement and Converse, 1986). This diversity could contribute to failure of TSV identification by conventional methods like ELISA. In such cases, detection of genetically diverse viruses can be facilitated by NGS. Rapid diagnosis can be difficult in cases where a virus is not well recognized as infecting a particular plant host or geographical region and is further complicated where the virus differs from known strains. Currently, most methods in use for virus detection are species-specific (Arif et al., 2014) (e.g. PCR or ELISA). A few detection techniques have been developed for genera and families (Sharman and Thomas, 2013). Also, sap mechanical inoculation and PCR with degenerate primers, can detect viruses in a species-specific manner but are limited by the fact that they can be applied to a relatively small range of potential viruses. NGS technology has proved to be a powerful, sensitive technology (Adams et al., 2009; Al Rwahnih et al., 2009; Kreuze et al., 2009) especially suited to detecting unknown or unsuspected viruses.

2. Methodology

2.1. Sample collection

Symptomatic leaf samples from 57 soybean plants, used in this study, were collected in a soybean survey in July 2012 made in Noble and Kay counties of Oklahoma USA and were negative by dot-immunobinding assay (DIBA) to 15 viruses known to infect soybeans (Ali et al., 2015, in preparation). Soybean plants were at the young stage before flowering. Field symptoms of a typical virus-infected plant included mild mottling, rugosity and chlorotic lesions on the third tri-foliate compound leaf. However, all symptomatic plants tested negative for TSV by DIBA. To further investigate the cause of the symptoms and whether an unknown virus was causing the disease symptoms, NGS was performed.

2.2. 454-sequencing

Total RNA was extracted from leaf samples using Tri Reagent (Molecular Research Center, Cincinnati, OH). The RNA concentrations of the samples were estimated by Nanodrop (Thermo scientific, Waltham, MA). Nanodrop concentrations ranged from 600 ng/μl to 1700 ng/μl. Four samples of RNA concentrations greater than 1 μg/μl were pooled for sequencing. These samples were chosen because of their relatively high RNA yields (>1 μg/μl) and purity as evidenced by bright bands on gel electrophoretic separation on a 1.5% agarose gel.

The pooled RNA extract mixtures were used to synthesize double-stranded cDNA using the cDNA Synthesis System kit (Roche, Indianapolis, IN) following the manufacturer's protocols. A library was then produced from the RNA extract, using kits supplied by 454-Roche. Sequencing was performed following the manufacturer's protocols.

2.3. Sequence analysis

The sequence reads obtained were assembled using NEWBLER v. 2.7 (Roche). The resulting contigs and unassembled sequences were aligned to the NCBI nr GenBank viruses database (taxid: 10239) using BLAST 2.7. Three contigs (00539, 00542 and 00545) were chosen for further sequence and phylogenetic analyses.

Open reading frame prediction and protein translation analyses were performed with ExPASY, a translating bioinformatics tool (Gasteiger et al., 2003). Sequence alignments and phylogenetic trees were constructed (using a maximum-likelihood method with 500 bootstrap replicates) using Clustal W and MEGA (v. 6) (Tamura et al., 2013). Phylogenetic trees were constructed using the amino acid sequences encoded by RNA 1, RNA 2 and RNA 3 of TSV isolates (Table 1) and by contigs 00539, 00542 and 00545. Methods of phylogenetic analyses used were neighbor-joining, maximum likelihood and maximum parsimony. Evolutionary distances (number of base substitutions per site) were computed using the Maximum Composite Likelihood method and were used to determine branch lengths on trees. Codon positions included were 1st + 2nd + 3rd + noncoding. Initial tree(s) for the heuristic searches were obtained automatically by maximum parsimony when the number of common sites was below 100 or less than one-fourth of the total number of sites, or by the BIONJ method with a MCL distance matrix.

Estimates of evolutionary divergence between sequences was carried out based on the pair-wise amino acid distances per site between contigs 00539 and isolates of TSV (Table 1) that had produced BLAST 2.7 hits with high statistical significance (*E*-value cut off 10^{-3}). Standard error estimate(s) were obtained by a bootstrap procedure (500 replicates). Analyses were conducted using the Dayhoff matrix based model (Tamura et al., 2013; Schwarz, 1979).

Recombination analysis was made by RDP (Martin et al., 2010) with nucleotide sequence alignments of contigs 00539, 00542 and 00545 separately with the corresponding segments of other TSV isolates (Table 1). When a recombination event was supported by at least three different algorithms out of the seven incorporated in the RDP software the recombination event was considered to