Virus Research xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Virus Research

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

Effects of genetic changes to the begomovirus/betasatellite complex causing cotton leaf curl disease in South Asia post-resistance breaking

Rob W. Briddon*, Fazal Akbar[1], Zafar Iqbal, Luqman Amrao[2], Imran Amin, Muhammad Saeed, Shahid Mansoor

National Institute for Biotechnology and Genetic Engineering, Jhang Road, Faisalabad, Pakistan

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Geminivirus
Begomovirus
Betasatellite
Cotton leaf curl disease
Resistance breaking

ABSTRACT

Cotton leaf curl disease (CLCuD) has been a problem for cotton production across Pakistan and north-eastern India since the early 1990s. The appearance of the disease has been attributed to the introduction, and near monoculture of highly susceptible cotton varieties. During the intervening period the genetic make-up of the virus(es) causing the disease has changed dramatically. The most prominent of these changes has been in response to the introduction of CLCuD-resistant cotton varieties in the late 1990s, which provided a brief respite from the losses due to the disease. During the 1990s the disease was shown to be caused by multiple begomoviruses and a single, disease-specific betasatellite. Post-resistance breaking the complex encompassed only a single begomovirus, *Cotton leaf curl Burewala virus* (CLCuBuV), and a recombinant version of the betasatellite. Surprisingly CLCuBuV lacks an intact transcriptional-activator protein (TrAP) gene. The TrAP gene is found in all begomoviruses and encodes a product of ~134 amino acids that is important in virus–host interactions; being a suppressor of post-transcriptional gene silencing (host defence) and a transcription factor that modulates host gene expression, including microRNA genes. Recent studies have highlighted the differences between CLCuBuV and the earlier viruses that are part of on-going efforts to define the molecular basis for resistance breaking in cotton.

1. Introduction

Viruses of the family *Geminiviridae* have genomes consisting of circular, single-stranded (ss)DNA. They are classified into seven genera (*Begomovirus*, *Curtovirus*, *Mastrevirus*, *Topocuvirus*, *Becurtovirus*, *Eragrovirus* and *Turncurtovirus*) according to their host range, insect vector and genome organization (Adams et al., 2013; Brown et al., 2012). They are widely distributed throughout the world and infect either monocotyledonous or dicotyledonous hosts. Geminiviruses of the genus *Begomovirus* are transmitted by the whitefly *Bemisia tabaci* and have genomes that consist of either a single or two ssDNA components. The two components of bipartite begomoviruses are known as DNA A and DNA B and both are, in most cases, essential for symptomatic infection of host plants. Monopartite begomoviruses are often associated with DNA satellites known as alphasatellites and betasatellites (Briddon and Stanley, 2006). The satellite-associated begomoviruses are widespread in the Old World and represent the largest group of begomoviruses, outnumbering both the truly monopartite and the bipartite begomoviruses (Brown et al., 2012).

Geminiviruses replication involves a double-stranded (ds) DNA intermediate that is also used as a template for transcription. Transcription is bidirectional to generate mRNAs diverging from an intergenic region. Transcripts initiate downstream of either consensus TATA box motifs or initiator elements, suggesting that host RNA polymerase II is involved in transcription of viral mRNAs. The viral RNAs are polyadenylated and composed of multiple RNA species, indicating the complexity of geminiviral transcription (Hanley-Bowdoin et al., 1999). The genomes of monopartite begomoviruses encode six genes, two in the virion-sense orientation (encoding the coat protein [CP] and V2 protein) and four in the complementary-sense orientation (encoding the replication-associated protein [Rep], the transcriptional-activator protein [TrAP], the replication enhancer protein [REn] and the C4 protein). The CP is involved in virus movement within and between plants (Briddon et al., 1990; Rojas et al., 2001) and the V2 protein is involved in virus movement in plants as well as, for some virus species, in overcoming RNA silencing; a host defence triggered

* Corresponding author. Tel.: +92 41 2651475; fax: +92 41 2651472.
E-mail addresses: rob.briddon@gmail.com, robbriddon@nibge.org (R.W. Briddon), mala_442000@yahoo.com (F. Akbar), zafariqbal2009@gmail.comluq (Z. Iqbal), raoluqman@gmail.com (L. Amrao), imranamin1@yahoo.com (I. Amin), saeed_hafeez@yahoo.com (M. Saeed), shahidmansoor7@gmail.com (S. Mansoor).
[1] Present address: Center for Biotechnology and Microbiology, University of Swat, Pakistan.
[2] Present address: Department of Plant Pathology, Faculty of Agriculture, University of Agriculture, Faisalabad, Pakistan.

0168-1702/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.virusres.2013.12.008

Fig. 1. Symptoms typical of CLCuD in cotton. The disease induces vein darkening, vein swelling and upward or downward leaf curling. For severe infections, enations on the main veins on the undersides of leaves develop into cup-shaped, leaf-like structures.

by dsRNA (known as RNA interference; Amin et al., 2011a; Glick et al., 2008; Hohn and Vazquez, 2011). The Rep protein is a rolling-circle replication initiator protein that also interferes with host cell-cycle, suppresses transcriptional gene silencing by indirectly reducing DNA methylation and is the only virus encoded protein required for virus replication (Hanley-Bowdoin et al., 1999, 2004; Rodríguez-Negrete et al., 2013). TrAP is involved in up-regulating late, virion-sense encoded genes (in some cases) as well as host genes and is a RNA silencing suppressor (Amin et al., 2011a; Dry et al., 2000; Sunter and Bisaro, 1992; Trinks et al., 2005). The REn protein is involved in virus DNA replication (by interacting with Rep) and also interacts with host components (Settlage et al., 2005). The product of the C4 gene may be a symptom determinant and also exhibits suppressor of RNA silencing activity (Amin et al., 2011a).

The betasatellites are a recently identified class of ssDNA satellites (Briddon and Stanley, 2006). They are, in many cases, required by their helper begomoviruses to symptomatically infect the hosts from which they were isolated (Briddon et al., 2001; Saunders et al., 2000). Betasatellites encode a dominant symptom/pathogenicity determinant (known as βC1) which is a suppressor of PTGS and may be involved in virus movement in plants and enhance virus DNA levels in plants (Amin et al., 2011a; Qazi et al., 2007; Saeed et al., 2007).

Cotton leaf curl disease (CLCuD; Fig. 1) is the most significant biotic constraint to cotton production across most of Pakistan and northwestern India (Sattar et al., 2013). The disease first appeared in an epidemic form in 1991–92 and at that time was shown to be caused by several monopartite begomoviruses (seven distinct species were identified), often as multiple infections (Mansoor et al., 2003b), and a specific betasatellite – Cotton leaf curl Multan betasatellite (CLCuMB; Briddon et al., 2001; Sattar et al., 2013). In the late 1990s the introduction of resistant cotton varieties restored cotton production in Pakistan to pre-epidemic levels. However, from 2001 to 2002 onwards, the disease appeared on all previously resistant varieties (Mansoor et al., 2003a), an indication of changes in the virus complex causing the disease that lead to the ability to break resistance. This complex came to be known as the "Burewala" strain – named after the city where the outbreak initiated.

2. The resistance breaking CLCuD complex

In contrast to the situation pre-resistance breaking, only a single begomovirus, a distinct species now known as *Cotton leaf curl Burewala virus* (CLCuBuV), has been identified in cotton across the Punjab, Pakistan. Unusually, CLCuBuV lacks an intact TrAP gene (Amrao et al., 2010). Despite 30 years of sequencing of geminiviruses only one other geminivirus, the curtovirus *Horseradish curly top virus* (HrCTV; Klute et al., 1996), has been shown to lack one of the usual complement of genes. However, only one isolate of HrCTV was shown to lack the REn gene, although the cloned virus was shown to be infectious to plants. In contrast, CLCuBuV was shown to occur across a wide area of Pakistan and, more recently, India (Rajagopalan et al., 2012; Zaffalon et al., 2011) and all isolates originating from *G. hirsutum* lacked the full TrAP gene. The study of Amrao et al. (2010) showed that although a small number of CLCuBuV isolates contained an in-frame stop codon in the TrAP gene, the majority of isolates contained either two stop codons (separated by 3 codons) or a stop codon and a frame shift. The first stop codon was common to all isolates, truncating the open

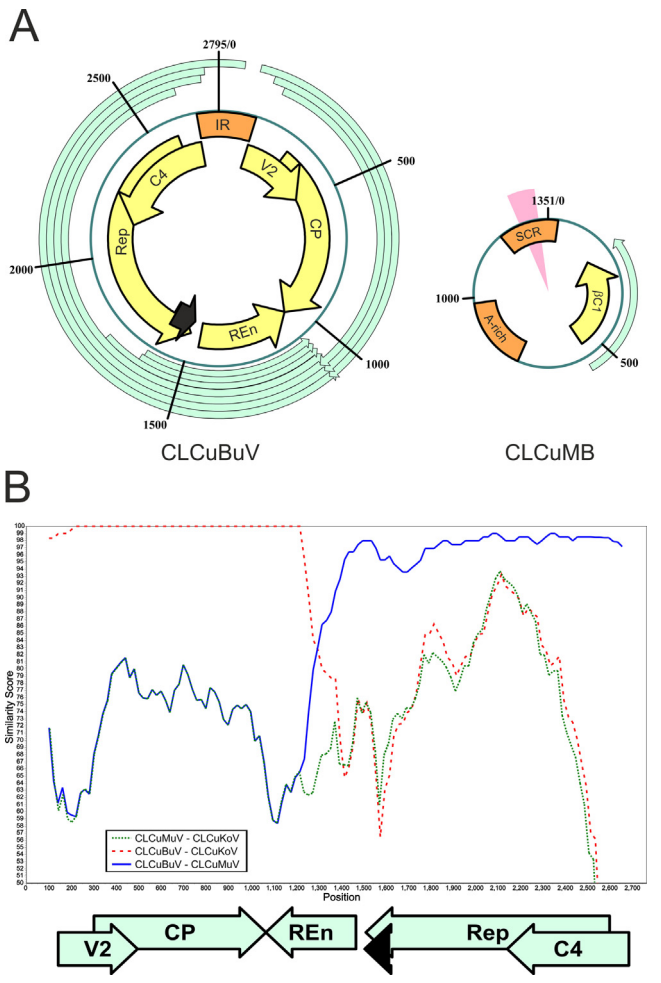

Fig. 2. (A) Structure of the genome of *Cotton leaf curl Burewala virus* (CLCuBuV) and its associated betasatellite, Cotton leaf curl Multan betasatellite (CLCuMB). The positions and orientations of genes are shown by arrows and are the V2 gene, the coat protein gene (CP), the replication-associated gene (Rep), the replication-enhancer gene (REn) and the βC1 gene. The truncated TrAP gene is shown in black. The positions and orientations of transcripts identified by Akbar et al. (2012) are shown as arrows around the outside of each figure. The positions of the non-coding intergenic region (IR) of CLCuBuV and the satellite conserved region (SCR) and adenine-rich region (A-rich) of CLCuMB are shown. The recombinant fragment of CLCuBuV associated with CLCuBuV within the SCR originating from Tomato leaf curl betasatellite is shown as a slice. (B) Plotcon analysis of CLCuBuV for recombination. Shown are Plotcon traces of percentage nucleotide sequence identities for comparisons of the genome CLCuBuV with *Cotton leaf curl Multan virus* (CLCuMuV) and *Cotton leaf curl Kokhran virus* (CLCuKoV). The approximate positions of genes are shown below the diagram.