Virus Genes (2016) 52:152–155
DOI 10.1007/s11262-015-1279-5

Detection by next generation sequencing of a multi-segmented viral genome from sugarcane associated with Ramu stunt disease

Dimitre Mollov[1] · Clarissa Maroon-Lango[2] · Lastus Kuniata[3]

Received: 24 September 2015 / Accepted: 18 December 2015 / Published online: 4 January 2016
© Springer Science+Business Media New York (Outside USA) 2016

Abstract Ramu stunt disease of sugarcane was first reported in Papua New Guinea in the mid 1980s. The disease can reduce sugarcane yields significantly and causes severe stunting and mortality in highly susceptible cultivars. The causal agent of Ramu stunt has been investigated but its characterization has not been completed. Sugarcane cv. Ragnar from Papua New Guinea with symptoms of Ramu stunt was analyzed by next generation sequencing. Total RNA was extracted and whole transcriptome shotgun sequencing was performed using an Illumina platform. Over thirty-seven million reads with an average length of 100 nucleotides were obtained. More than eighteen thousand contigs were assembled and subjected to BLASTX analysis. Twenty-one contigs were virus related and six were associated with plant viruses. The BLAST algorithms revealed sequence similarity to *Tenuivirus* and *Phlebovirus*, genera of viruses whose members contain genomes consisting of multiple RNA segments. The six contigs derived from the RNA sequencing data correspond to six RNAs that compose the Ramu stunt virus genome. Primers were designed for each of the six RNAs and RT-PCR amplicons were obtained only from the symptomatic sugarcane. There was concordance between the sequence data of the contigs obtained from the NGS and that of the amplicons obtained by RT-PCR. The NGS approach allowed us to determine the complete genomic sequence of Ramu stunt virus. It is likely that this virus is the causal agent of Ramu stunt disease.

Keywords Ramu stunt disease · Ramu stunt virus · Tenuivirus · Multi-segmented virus genome · Negative strand ssRNA virus · Next generation sequencing

Ramu stunt, a destructive disease of sugarcane, was first reported in Papua New Guinea in the mid-1980s [1–3]. Symptoms include pale green to yellow streaks that vary in length from small specks to chlorotic striping through the entire leaf. Infected canes also appear thinner, with abnormal proliferation from the nodes and reduced internode length. The overall effect of the disease is plants with severely reduced growth rates, and that fail to ratoon. Initial reports suggested the causal organism to be phytoplasma [4, 5] but was later described as a virus, possibly belonging to the genus *Tenuivirus* [6]. The disease is still known to occur only in Papua New Guinea, where it can be effectively transmitted by the island sugarcane planthopper, *Eumetopina flavipes* Muir (Hemiptera) [2]. Therefore, Ramu stunt disease poses a major quarantine risk to other sugarcane producing countries, especially nearby Australia [3, 6, 7]. This communication reports the full genomic sequence of a virus provisionally named Ramu stunt virus (RmSV), associated with the Ramu stunt disease of sugarcane. RmSV is most closely related to members of the genus *Tenuivirus*. Tenuiviruses (from Latin *tenuis* meaning thin) are transmitted by plant hoppers in a circulative manner and have negative and/or ambi-sense ssRNA-

Edited by Thomas Hohn.

✉ Dimitre Mollov
dimitre.mollov@ars.usda.gov

[1] USDA-ARS, National Germplasm Resources Laboratory, Beltsville, MD, USA

[2] USDA-APHIS PPQ Plant Health Programs, Riverdale, MD, USA

[3] New Britain Palm Oil, Ramu Agri Industries, Lae, Morobe Province, Papua New Guinea

segmented genomes that contain four to six RNAs [8, 9]. Each RNA encodes one to two proteins. Members of the genus *Tenuivirus* are more closely related to the vertebrate-infecting viruses in the genus *Phlebovirus* of the family *Bunyaviridae* than they are to other plant viruses.

Tenuiviruses have host ranges limited to the family *Poaceae*, affecting crops such as rice, maize, and wheat. The International Committee on Taxonomy of Viruses recognizes seven members in the genus *Tenuivirus*: *rice stripe virus* (RSV) [10], *rice grassy stunt virus* (RGSV) [11], *rice hoja blanca virus* (RHBV) [12, 13] *maize stripe virus* (MSV) [14], *Iranian wheat stripe virus* [15], *echinochloa hoja blanca virus* (EHBV) [16], and *urochloa hoja blanca virus* (UHBV) [17]. Brazilian wheat spike virus, European wheat striate mosaic, rice wilted stunt virus, winter wheat mosaic virus, maize yellow stripe virus (MYSV) [18], and winter yellow head virus [19] have been tentatively placed in the genus *Tenuivirus* but have not been approved as species [9].

Sugarcane (*Saccharum officinarum*) cultivar Ragnar exhibiting symptoms of Ramu stunt disease was imported from the Ramu Madang Province in Papua New Guinea through the USDA APHIS Plant Germplasm Quarantine Program in Beltsville, Maryland. Clonally propagated plants were established, grown in a quarantine greenhouse under standard conditions, and used as a source for RNA extraction. Total RNA was extracted from symptomatic leaf tissue from two plants using Qiagen RNeasy® kit and used for the RNA Seq analysis.

The RNA from the Ramu stunt infected source was sent to Macrogen (Seoul, Korea) for cDNA library preparation and sequencing as pair-end 100 bp reads on an Illumina HiSeq 2500 system. Sequences were assembled into contigs using CLC Genomics Workbench (Qiagen). BLASTN and BLASTX queries to NCBI GenBank were used to detect virus-related contigs. To verify virus-like sequences, overlapping primer pairs were designed from the putative RNAs and RT-PCR was carried out using the illustra Ready-To-Go RT-PCR Beads (GE Healthcare). To obtain or verify the sequences at the 5′- and 3′-termini, a primer with a sequence identical to the 5′ conserved 10 nt [20] end of Tenuiviruses RNAs and a GC rich anchor (TenuiD, 5′-CCC GGG CGG CCG CAC ACA WAG TC) was paired to sequence specific primers at both ends for all predicted genomic RNAs. The TenuiD primer was modified after de Miranda et al. [12] to account for the A/U substitution at the 5′- or 3′ UTR-conserved genomic sequences [8, 20]. All RT-PCR products were cloned using pGEM®-T Easy Vector System (Promega) and a minimum of three clones per reaction were sequenced in both directions. Sequences were initially assembled in Genious v. 7 but final mapping of the reads to each of the RNAs were performed using CLC Genomics Workbench (Qiagen). The RdRp proteins of RmSV and those of selected members of the *Tenuivirus*, *Phlebovirus*, and *Tospovirus* were used in phylogenetic analysis by the Neighbor-joining algorithm with 1000 bootstrap replicates on MEGA v. 6 [21]. *Mirafiori lettuce virus,* an *Ophiovirus*, was used as an out group.

A total of 37,606,204 reads were obtained from sample one and 38,990,344 reads were obtained from sample two from the RNA sequencing experiments. These reads assembled into 676,591 and 746,495 contigs from sample one and two, respectively. In order to optimize the BLAST searches, a subset of 18,175 (sample 1) and 18,510 (sample 2) contigs were used. The subset included contigs 500–10,000 nt in length. Out of these contigs 21 were virus related, and six were plant virus related. These six contigs were assigned as RNA 1 through RNA 6 based on length, and assembled as 16,312 nucleotides. Both samples yielded the same contigs and had 100 % sequence identity. All sequences were deposited in the NCBI GenBank under accession numbers: KR094115, KR094116, KR698381, KR094117, KR094118, and KR094119.

Tentative annotation of the genomic segments of RmSV is presented in Table 1.

RNA 1 is 8902 nt long containing one ORF with predicted protein of 2921 amino acids (aa) with M_r of 338 kDa. The predicted translation product is similar in

Table 1 Ramu stunt virus (RmSV) RNA segments, ORF position, protein size, and significant amino acid sequence identity

RNA segment	NT position	Protein size		Comparative maximal % aa sequence identity, virus and GenBank accession number
		aa size	kDa	
RNA 1	34–8799	2921	338	38 %, rice stripe virus RdRp, ADE60694
RNA 2	54–953	299	33	28 %, Kismayo virus nucleoprotein, AIU95035
	481–74	135	18	No significant similarity
RNA 3	50–658	202	14	25 %, MYSV hypothetical protein, CAI94658
	1524–1141	127	17	27 % MYSV hypothetical protein, CAI94656
RNA 4	22–1176	384	43	24 %, MYSV hypothetical protein, CAI94659
RNA 5	37–900	287	33	28 %, Punique virus nucleocapsid, ACZ43797
RNA 6	60–590	176	21	27 %, MYSV hypothetical protein, CAI94656

size to the RNA polymerases of RSV and RGSV. BLAST analysis showed that the RdRp protein has the highest identity of 38 % to RSV RdRp. RNA 2 is 1675 nt long and codes for a putative nucleoprotein of 299 aa with M_r of 33 kDa (nt 54–953). On the complementary strand of RNA 2, there is a 135 aa hypothetical protein (nt 481–74) that has no similarity to known viral proteins. RNA 3 encodes a 202 aa protein (nt 50–658) and on the complementary strand a 127 aa protein (nt 1524–1141). Both of these predicted peptides have similarities to MYSV hypothetical proteins [22] of 25 and 27 %, respectively. RNA 4 is 1575 nt long and encodes a 384 aa, 43 kDa predicted protein (nt 22–1176). BLAST results showed a 24 % similarity to a MYSV hypothetical protein [22]. RNA 5 is 1377 nt long and has predicted ORF of 287 aa (nt 37–900). This protein has a M_r of 33 kDa and is mostly similar (28 %) to a number of Phlebovirus nucleocapsid proteins. RNA 6 is composed of 1201 nt and encodes a hypothetical protein of 176 aa (nt 60–590) similar to a MYSV hypothetical protein [22].

Comparison of the 5′ UTR sequences of all six RNAs showed a conserved segment 5′-ACACAAAGUC-3′ typical of the Tenuiviruses [8, 20]. The 3′ UTR is also conserved with RNA 1 through 5 having the 10 nt segment 5′-

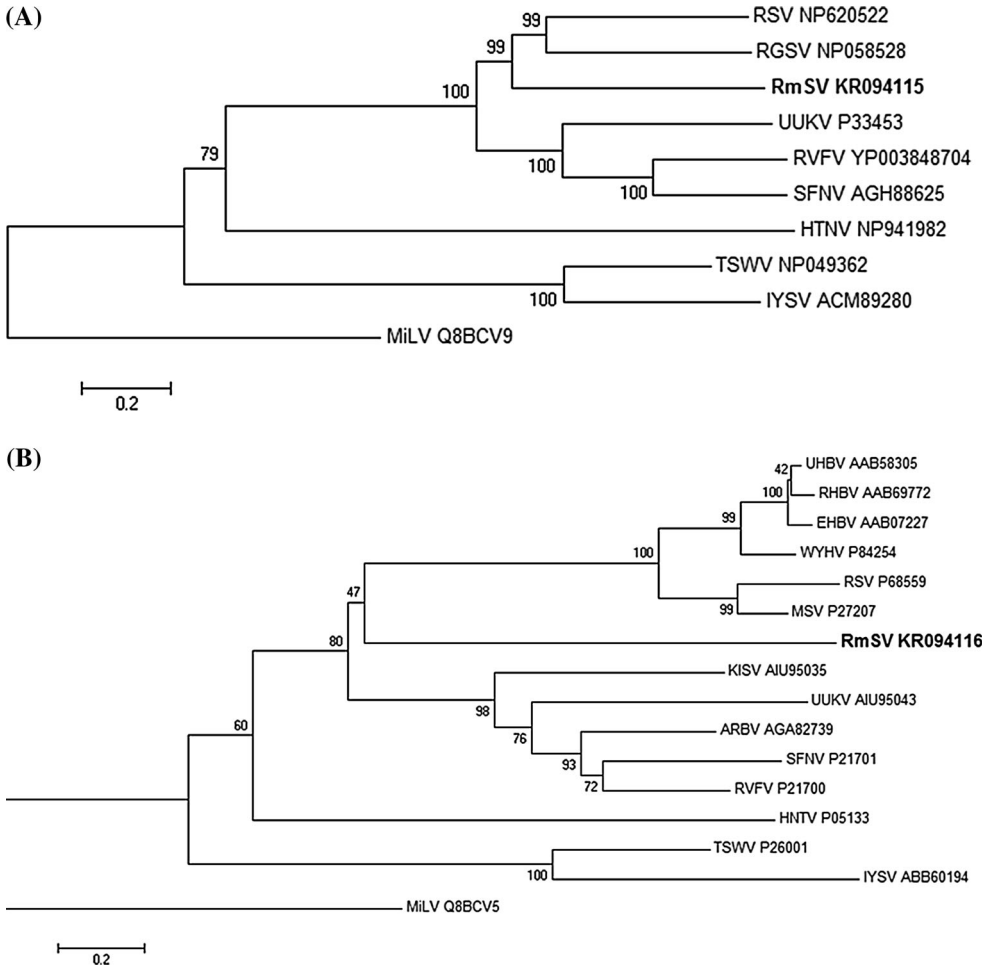

Fig. 1 **a** Phylogenetic analysis using the RNA dependant RNA polymerase (RdRp) amino acid sequences of Ramu stunt virus (RmSV) and RdRp sequences of Tenuiviruses: rice stripe virus (RSV) and rice grassy stunt virus (RGSV); Phleboviruses: Uukuniemi virus (UUKV), Rift Valley fever virus (RVFV), Sandfly fever Naples virus (SFNV); Hantavirus: Hantaan virus (HNTV); Tospoviruses: tomato-spotted wilt virus (TSWV) and Iris yellow spot virus (IYSV). Ophiovirus: Mirafiori lettuce virus (MiLV) used as an out-group. **b** Phylogenetic analysis using the nucleoprotein amino acid sequences of Ramu stunt virus (RmSV) and nucleoprotein sequences of Tenuiviruses: urochloa hoja blanca virus (UHBV), rice hoja blanca virus (RHBV), echinochloa hoja blanca virus (EHBV), Wheat yellow head virus (WYHV), rice stripe virus (RSV) and maize stripe virus (MSV); Phleboviruses: Kismayo virus (KISV), Uukuniemi virus (UUKV), Arbia virus (ARBV), Sandfly fever Naples virus (SFNV), and Rift Valley fever virus (RVFV); Hantavirus: Hantaan virus (HNTV); Tospoviruses: tomato-spotted wilt virus (TSWV) and Iris yellow spot virus (IYSV); and Ophiovirus: Mirafiori lettuce virus (MiLV) used as an out-group. Both trees were generated by the neighbor-joining algorithm and bootstrap values indicated at each node were calculated using 1000 bootstrap replicates. *Tree branches* are proportional to genetic distances between sequences and the scale bar represents substitutions per amino acid site

GACUUUGUGU-3′ while RNA 6 had one base pair difference 5′-GACUAUGUGU-3′ [8]. The 3′ terminus of RSV RNA 1 has the identical 10 nt segment found in RmSV [20].

Once the full length of all RNAs was finalized, remapping results showed different frequencies of read maps to the final product: RNA 1—340,701 (0.91 %); RNA 2—870,223 (2.31 %); RNA 3—285,005 (0.76 %); RNA 4—681,795 (1.81 %); RNA 5—420,168 (1.12 %); RNA 6—2,103,686 (5.59 %). The minimum sequence depth was at least 2000 reads but in some cases was over 240,000. The mapping confirmed that almost all of the 12 termini of the six RNAs were fully completed, with a few exceptions of three to four nucleotides missing at some of the ends. These missing termini were completed by RT-PCR using TenuiD and sequence specific primers.

After the RmSV complete sequence was finalized, primers were designed to amplify products from each RNA. RT-PCR was performed from the Ramu stunt-infected sugarcane and from two apparently healthy cultivars. Amplicons were obtained from Ramu infected RNA from all of the six RNAs primer pairs. No amplification products were obtained from the healthy sugarcane. These analyses indicate that the six RNAs obtained from the RNA Seq experiments likely comprise RmSV, particularly because no other plant virus-related contigs were identified.

Phylogenetic analysis of the RdRp of RmSV and the RdRp proteins of selected viruses from the genera *Tenuivirus*, *Phlebovirus*, and *Tospovirus* placed RmSV closely to *Tenuivirus* (Fig. 1a). Similar results were obtained when the nucleoprotein was analyzed (Fig. 1b).

Next generation sequencing techniques facilitated the assembly and molecular characterization of RmSV, the likely cause of Ramu stunt disease of sugarcane. RmSV comprises six segments of ssRNA that total 16,312 nucleotides, with a genome organization that shows degree of similarity to tenuiviruses. This information will help develop improved diagnostic methods for Ramu stunt virus and will aid in global quarantine and disease management efforts.

References

1. J.M. Waller, B.T. Egan, D. Eastwood, Ramu stunt, an important new sugarcane disease in Papua New Guinea. Trop. Pest Manag. **33**, 347–349 (1987)
2. L.S. Kuniata, G.R. Young, E. Pais, Nagaraja, Preliminary observations on Eumetopina sp. (Hemiptera: Delphacidae) as a vector of Ramu stunt disease of sugarcane in Papua New Guinea. J. Aust. Entomol. Soc. **33**, 185–186 (1994)
3. K.L. Anderson, N. Sallam, B.C. Congdon, The effect of host structure on the distribution and abundance of the island sugarcane planthopper, Eumetopina flavipes Muir, vector of Ramu stunt disease of sugarcane. Virus Res. **141**, 247–257 (2009)
4. C.P.R. Cronje, R.A. Bailey, P. Jones, S. Suma, The phytoplasma asociated with Ramu stunt disease of sugarcane is closely related to white leaf phytoplasma group. Plant Dis. **83**, 588 (1999)
5. S. Suma, P. Jones, in *A Guide to Sugarcane Diseases*, ed. by P. Rott, R.A. Bailey, J.C. Comstock, B.J. Croft, A.S. Saumtally (CIRAD and ISSCT, Montpellier, 2000), pp. 226–230
6. K.S. Braithwaite, B.J. Croft, R.C. Magarey, Progress in identifying the cause of Ramu stunt disease of sugarcane. Proc. Aust. Soc. Sugar Cane Technol. **29**, 235–241 (2007)
7. K.S. Braithwaite, R. Kombukon, L.S. Kuniata, and R.C. Magarey, Ramu stunt: resistance screening and validation of the diagnostic test. Proc. Aust. Soc. Sugar Cane Technol 34 (2012)
8. B.W. Falk, H. Tsai, Biology and molecular biology of viruses in the genus Tenuivirus. Annu. Rev. Phytopathol. **36**, 139–163 (1998)
9. A.M.Q. King, E. Lefkowitz, M.J. Adams, E.B. Castens, *Virus taxonomy: ninth report of the international committee on taxonomy of viruses* (Academic Press, San Diego, 2012)
10. S. Toriyama, M. Takahashi, Y. Sano, T. Shimizu, Nucleotide sequence of RNA 1, the largest genomic segment of rice stripe virus, the prototype of the tenuiviruses. J. Gen. Virol. **75**, 3569–3579 (1994)
11. S. Toriyama, T. Kimishima, M. Takahashi, T. Shimizu, N. Minaka, K. Akutsu, The complete nucleotide sequence of the rice grassy stunt virus genome and genomic comparisons with viruses of the genus Tenuivirus. J. Gen. Virol. **79**, 2051–2058 (1998)
12. J. de Miranda, M. Hernandez, R. Hull, A.M. Espinoza, Sequence analysis of rice hoja blanca virus RNA 3. J. Gen. Virol. **75**, 2127–2132 (1994)
13. E. Shikata, G.E. Calvez-E, Fine flexuous threadlike particles in cells of plants and insect hosts infected with rice hoja blanca virus. Virology **39**, 635–641 (1969)
14. R.E. Gingery, L.R. Nault, O.E. Bradfute, Maize stripe virus: characteristics of a member of a new virus class. Virology **112**, 99–108 (1981)
15. J. Heydarnejad, W.S. Barclay, K. Izadpanah, F.R. Hunter, M.J. Gooding, Molecular characterization of Iranian wheat stripe virus shows its taxonomic position as a distinct species in the genus Tenuivirus. Arch. Virol. **151**, 217–227 (2006)
16. J. de Miranda, M. Munoz, R. Wu, A.M. Espinoza, Sequence of Echinochloa hoja blanca tenuivirus RNA-5. Virus Genes **12**, 131–134 (1996)
17. J.R. de Miranda, M. Munoz, R. Wu, A.M. Espinoza, Phylogenetic placement of a novel tenuivirus from the grass Urochloa plantaginea. Virus Genes **22**, 329–333 (2001)
18. E.D. Ammar, R.E. Gingery, D.T. Gordon, A.E. Aboul-Ata, Tubular helical structures and fine filaments associated with the leafhopper-borne mayze yellow stripe virus. Phytopathology **80**, 303–309 (1990)
19. D.L. Seifers, T.L. Harvey, T.J. Martin, S. Haber, Y.M. She, W. Ens, K.G. Standing, R. Salomon, A. Gera, Association of a virus with wheat displaying yellow head disease symptoms in the great plains. Plant Dis. **89**, 888–895 (2005)
20. M. Takahashi, S. Toriyama, Y. Kikuchi, T. Hayakawa, A. Ishihama, Complementarity between the 5′- and 3′-terminal sequences of rice stripe virus RNAs. J. Gen. Virol. **71**, 2817–2821 (1990)
21. K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. **30**, 2725–2729 (2013)
22. A. Mahmoud, M. Royer, M. Granier, E.D. Ammar, J.C. Thouvenel, M. Peterschmitt, Evidence for a segmented genome and partial nucleotide sequences of maize yellow stripe virus, a proposed new tenuivirus. Arch. Virol. **152**, 1757–1762 (2007)