Iranian J Biotech. October 2020;18(4): e2566 DOI: 10.30498/IJB.2020.2566

Research Article

Sequence and Phylogenetic Analysis of the First Complete Genome of Rice tungro spherical virus in Malaysia

Maathavi Kannan[1], Maisarah Mohamad Saad[2], Zamri Zainal[1,3], Hakimi Kassim[1], Ismanizan Ismail[1,3], Noraini Talip[4], Syarul Nataqain Baharum[1], Hamidun Bunawan[1,*]

[1]Institute of Systems Biology, Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia
[2] Institute of Rice Research Centre, MARDI Seberang Perai, Jalan Paya Keladi/ Pinang Tunggal, 13200 Kepala Batas, Penang, Malaysia
[3] School of Biosciences and Biotechnology, Universiti Kebangsaan Malaysia, 43600 Bangi Selangor, Malaysia
[4] School of Environmental and Natural Resource Sciences, Universiti Kebangsaan Malaysia, 43600 Bangi Selangor, Malaysia

*Corresponding author: Hamidun Bunawan, Institute of Systems Biology, Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia; Tel: +60-389214554 Fax: +60-389213398; E-mail: hamidun.bunawan@ukm.edu.my

Background: Rice tungro disease (RTD) is a viral disease mainly affecting rice in Asia. RTD caused by *Rice tungro bacilliform virus* and *Rice tungro spherical virus*. To date, there are only 5 RTSV isolates have been reported.

Objectives: In this study, we aimed to report the complete nucleotide sequence of Malaysian isolate of *Rice tungro spherical virus* Seberang Perai (RTSV-SP) for the first time. RTSV-SP was characterized and its evolutionary relationship with previously reported Indian and Philippines isolates were elucidated.

Materials and Methods: RTSV-SP isolate was isolated from a recent outbreak in a paddy field in Seberang Perai zone of Malaysia. Its complete genome was amplified by RT-PCR, cloned and sequenced.

Results: Sequence analysis indicated that the genome of RTSV-SP consisted of 12,173 nucleotides (nt). Comparative analysis of 6 complete genome sequences using Clustal Omega showed that Seberang Perai isolate shared the highest nucleotide identity (96.04%) with Philippine-A isolate, except that the sORF-2 of RTSV-SP is shorter than RTSV Philippine-A by 27 amino acid residues. RTSV-SP found to cluster in Southeast Asia (SEA) group based on the whole genome sequence phylogenetic analysis using MEGA X software.

Conclusions: Phylogenetic classification of RTSV isolates based on the complete nucleotide sequences showed more distinctive clustering pattern with the addition of RTSV-SP whole genome to the available isolates. Present study described the isolation and molecular characterization of RTSV-SP.

Keywords: Genome; Malaysia; Phylogeny; *Rice tungro spherical virus*

1. Background

The Asian rice, *Oryza sativa* L. is classified in the Poaceae family (Gramineae) (1). Since it is the most vital food crop globally (2), a vast amount of rice is being produced every year, which is about 650 million tons (3). More than 90% of the rice yield are from Asia (4). Rice tungro disease (RTD) is considered to be the most serious threat to rice production among all 22 viral diseases attacking rice plants (5,6). Main outbreaks of RTD have been reported in many Asian countries; Philippines, India, Inonesia, Malaysia, Bangladesh, Nepal, Pakistan, Sri Lanka, Vietnam, China, and Japan(7–10). Tungro virus epidemics can result in 30% to 100% severe yield losses annually (11). Two distinct viruses in the aspect of serology and morphology; *Rice tungro bacilliform virus* (RTBV) and *Rice tungro spherical virus* (RTSV) are the causal agents of this sporadic disease (12–14).

Green paddy leafhopper (*Nephotettix virescens*) transmits RTBV and RTSV as a complex or separately in a semi-persistent way (15). However, spreading of RTBV by the vector is only possible with prior acquisition of RTSV (11). The interaction between both viruses during transmission has been explained in relation to a requirement of a helper factor, possibly a protein component that being produced only in plants infected with RTSV. The component might be essential for RTBV absorption to the leafhopper stylet wall (16–18) but its location on the RTSV genome has not been determined (19). RTBV and RTSV also interact in order

to allow disease development apart from to enable symptoms such as stunting, yellow-orange coloration and reduced tillering to be fully expressed (20). However, the correlation between tungro viruses and symptom severity has not been well studied (21). RTBV can be found in the vascular bundles meanwhile RTSV is limited to the phloem tissues in diseased rice plants. Both viral particles are localized in the cytoplasm of infected cells, either in scattered or aggregated pattern (22).

RTD was first identified in the experimental fields at the International Rice Research Institute, Philippines in 1963 and recognized as an apparent viral disease in 1965 (23). To date, there are only five RTSV whole genomes available in NCBI database. Philippines isolate was first sequenced in 1992 (24), followed by complete sequencing of RTSV isolate Vt6, from Mindanao, Philippines. Compared to the type strain of Philippine, Vt6 found to be more virulent when being tested on TKM6, RTSV resistant cultivar (25). Complete genome sequence of three isolates from India; Orissa (Ori), West Bengal isolate (WB) and Andhra Pradesh (AP) have been reported later (26,27). Since the number and geographical origin of published full-length sequences are very limited, the information on genome wide molecular diversity of RTSV population is still lacking. Tungro disease is called as penyakit merah (PMV) locally (22) and the first incursion was suspected to attack Kerian area in 1933 (11). Major outbreaks of RTD were reported in 1982-1983 in Peninsula Malaysia (28) where it affected more than 20,300 ha of Kedah and Perlis rice fields, causing around US $10 million yield loss (11). Apart from that, in 2012, RTD incidence was detected in 5 regions of Sarawak. Interestingly, the native rice cultivars in the affected areas showed susceptibility towards tungro viruses.

RTSV belongs to genus *Waikavirus* and family *Sequiviridae* (13). The viral genome comprised of a positive sense, single stranded RNA of more than 12 kb in length with a poly(A) track at its 3' end (29). A protein, VPg was suggested to be covalently linked to the 5' end of the RNA (22). RTSV has isometric particles with a diameter of 30nm. Since RTSV features resemble animal picornavirus, the virus has been termed as "plant picornavirus" (30). The viral RNA has a leader sequence of 515 nucleotides and encode a single large ORF1 at 5' end starting after the leader sequence (22). ORF1 encodes a polyprotein, cleaved by a protease resulting in three coat proteins (CPs) organized adjacent to each other (CP1-3) (31).

A leader protein (P1) presents upstream of CP1 (32). The polyprotein also consists of a nucleotide triphosphate (NTP) binding protein, a protease (Pro), and a RNA-dependent RNA polymerase (Rep). Initially, it was reported that 3' tail of RTSV genome contains two small ORFs (sORF-2 and sORF-3), expressed from sub-genomic mRNAs (24,31). The function of these proteins remains unclear and their occurrence were not confirmed in following studies (26,33).

So far, nucleotide sequences of CP1, CP2 and CP3 of a RTSV isolate (MaP1) from Malaysia were reported (34). Knowledge on complete nucleotide sequences of RTSV isolated from Malaysia will assist future analyses on molecular and evolutionary characteristics of RTSV in Malaysia.

2. Objective

Here, we report the first whole genome of RTSV-SP isolate obtained from a paddy field in Seberang Perai, Malaysia in 2018. The nucleotide sequences of RTSV-SP were analyzed in detail to provide a reference data for future genetic studies on RTSV in the particular region as well as in Malaysia generally.

3. Materials and Methods

3.1. Materials

Rice field in Seberang Perai (SP), Malaysia (100°41'39"E, 5°37'00"N) was inspected for samples and rice plants exhibiting RTD symptoms were collected. Plant RNA Mini Kit and PCRBIO HiFi polymerase were purchased from QIAGEN (Kuala Lumpur, Malaysia) and Biomax Scientific (Selangor, Malaysia) respectively. Extraction of total RNA was done using the RNA Mini Kit following the provided instructions.

3.2. cDNA Generation and PCR Amplification of RTSV Whole Genome

cDNA library was synthesized through addition of RNA template (1pg – 5µg) to appropriate components as follows: 4 µL of 5× reaction transcription buffer, 2 µL of dNTPs mixture (10mM), 1 µL of RNase inhibitor, 1 µL of OligodT18 primer (100 µM), 1 µL of RevertAid reverse transcriptase (ThermoFisher Scientific, Malaysia) (20U), and 1 µL of sterile water in 20 µL reaction volume. All constituents were mixed up gently and incubated for an hour at 42 °C. The incubation was terminated at 70 °C in 5 min prior to storage at -20 °C. PCRBIO HiFi polymerase was used to amplify the resulted cDNA through PCR. Nine pairs of overlapping primers were designed in current study based on the aligned complete nucleotide sequences of five RTSV isolates derived from GenBank (https://www.ncbi.nlm.

Table 1. List of overlapping primers used for sequencing of complete genomic cDNA of RTSV

Primers	Sequence (5'-3')	Annealing temperature (°C)	Expected amplicon size (bp)
RTSV1F	TGAAAATTGGGGTATAGAGATACCCC		1,442
RTSV1R	GACGCATGCCATGTTTAGAG	52.0	
RTSV2F	TTGCAGGGATGCCAATGG		1,682
RTSV2R	ACAAGCCCAAGCTCAGTT	55.0	
RTSV3F	CAGGAAGTGAAGGCTCCATAG		1,247
RTSV3R	GAGTGGCTCACGATGTAC	52.3	
RTSV4F	GAGTGGCTCACGATGTAC		1,473
RTSV4R	ACGCGCGATATGGCTCTTATG	55.0	
RTSV5F	ACGCGCGATATGGCTCTTATG		1,358
RTSV5R	GCATTCGCAGTGTAGATATAG	54.0	
RTSV6F	GCATTCGCAGTGTAGATATAG		414
RTSV6R	GCACATTTACATCGGAT	48.7	
RTSV7F	GCACATTTACATCGGATGGT		1,370
RTSV7R	TGTGGGATCTTGGACCTTC	53.7	
RTSV8F	TGTGGGATCTTGGACCTTC		1,470
RTSV8R	GTACATCGGGATGGAATCTGTG	54.7	
RTSV9F	GTACATCGGGATGGAATC		2,062
RTSV9R	GTGTGGGGTAATAGAGTAC	49.4	

nih.gov/) (35) **(Table 1)**. PCRs were conducted in a total volume of 50 µL comprising of 5× buffer (10 µL), cDNA (2 µL), forward primer (10 µM, 2 µL), reverse primer (10 µM, 2 µL), PCRBIO HiFi polymerase (0.5 µL) and sterilized water (34.5 µL). The PCR conditions was set up accordingly; initial denaturation at 95 °C (1 min), followed by 35 cycles of denaturation (95 °C, 15 s), annealing (50 °C -60 °C, 15 s) and extension (72 °C , 1-2 min) and a final extension at 72 °C for 10 min in the Applied Biosystems Veriti 96 Well Thermal cycler.

3.3. Cloning and Sequencing of PCR Products

Expected size of successfully amplified fragments was validated through electrophoresis on 1% (w/v) agarose gels. Purification of PCR products from agarose gel was conducted using QIAquick® Gel Extraction Kit (QIAGEN, Malaysia) and then ligated into PJET1.2 vector (ThermoFisher Scientific, Malaysia) according to manufacturer's instructions. They were consequently transformed into *Escherichia coli* (TOP10) competent cells through calcium chloride transformation method (36). Every positive colony consisting recombinant plasmid was selected and grown in ampicillin supplemented Luria-Bertani broth. The plasmids were isolated using FavorPrep™ Plasmid Extraction Mini Kit from (Biomax Scientific, Malaysia) after being certified through colony PCR. Minimum two recombinant plasmids for every fragment were sent to First BASE Laboratories Sdn. Bhd. Company for sequencing in forward and reverse directions using Sanger sequencing method. A total of 18 bidirectional reads were performed.

3.4. Genomic Sequence and Phylogenetic Analysis

Nucleotide sequence data produced by nine overlapping PCR fragments were assembled using Clustal Omega version 1.2.4 (https://www.ebi.ac.uk/Tools/msa/clustalo/) with default parameters (37). The obtained full-length genome of RTSV-SP was then deposited into NCBI library under accession number: MK655459. BLASTN program (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (38) was accessed to compare the similarity of the RTSV-SP DNA with published sequences. Genome sequence was translated and furthered to ORFs searching. Database searches against the GenBank non-redundant database and SwissProt Protein sequence database were performed using SmartBLAST (https://blast.ncbi.nlm.nih.gov/smartblast/smartBlast.cgi) (39) and BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (38) programs. Nucleotide sequences were aligned

with other available RTSV sequences using Clustal W tool (40). Phylogenetic tree based on the whole nucleotide sequences of RTSV-SP and available RTSV reference isolates as listed in **Table 2** was generated using molecular evolutionary genetics analysis MEGA X software (41). Maximum Likelihood method based on the Tamura 3-parameter model (42) was utilized in the phylogeny construction. For the phylogenetic relationship based on amino acid sequences of ORF1 of Seberang Perai isolate along with previously reported RTSV isolates, Neighbor-Joining method (43) and Poisson correction model (44) was used. The bootstrap values were estimated for 1000 replicates.

4. Results

4.1. Genomic Sequence Analysis of RTSV-SP

The genome sequence of RTSV-SP isolate was 12,173 nucleotides (nt) long, excluding its 3' end poly A tail (GenBank no.: MK655459) with 45.8% of GC content. The 5' portion of the genome contained a 514 nt untranslated region (UTR) followed by a single large ORF-1 of 10,413 nt. ORF-1 has AUG initiation codon at position 515 and give rise to a potential polypeptide of 3,471 amino acid (aa) residues. The polyprotein encoded a leader protein (P1), three coat proteins (CPs); CP1, CP2 and CP3, a nucleotide triphosphate

Table 2. Isolates of RTSV used for genome comparison and phylogenetic analysis

No.	RTSV isolates	Geographical origin	GeneBank accession #	References
1	PhilA	Philippines	NC 001632	(24)
2	Vt6	Philippines	AB064963	(25)
3	Ori	India	AM234048	(26)
4	WB	India	AM234049	(26)
5	AP	India	KC794785	(27)

Table 3. Genome length and nucleotide position of each genome segment of RTSV isolates.

Isolate	Total length (bp)	ORF-1	sORF-2	sORF-3	P1	CP1	CP2	CP3	NTP	Pro	Rep
SP	12,173	515-10930 (3471)*	11125-11343 (72)*	11433-11684 (83)*	515-2437	2438-3061	3062-3670	3671-4546	4547-8083	8084-9061	9062-10927
PhilA	12,175	515-10936 (3473)	11131-11430 (99)	11437-11688 (83)	515-2446	2447-3070	3071-3679	3680-4555	4556-8092	8093-9070	9071-10933
Vt6	12,171	517-10932 (3471)	11127-11243 (38)	11432-11683 (83)	517-2439	2440-3063	3064-3672	3673-4548	4549-8085	8086-9063	9064-10929
Ori	12,174	516-10931 (3471)	11126-11242 (38)	11432-11683 (83)	516-2438	2439-3062	3063-3671	3672-4547	4548-8084	8085-9062	9063-10928
WB	12,174	516-10931 (3471)	11126-11242 (38)	11432-11683 (83)	516-2438	2439-3062	3063-3671	3672-4547	4548-8084	8085-9062	9063-10928
AP	12,171	515-10930 (3471)	11124-11240 (38)	11430-11681 (83)	515-2437	2438-3061	3062-3670	3671-4546	4547-8083	8084-9061	9062-10927

* Values within the parenthesis indicate total number of amino acids for ORF1 and sORFs, CP coat protein, NTP nucleotide triphosphate binding protein, Pro proteinase, Rep RNA polymerase.

binding domain (NTP), a proteinase (Pro) and a RNA polymerase (Rep) in 5' to 3' orientation. Nucleotide position of the domains were presented in **Table 3**. The 3' terminus of RTSV-SP consists of two putative short ORFs (sORF-2 of 72aa and sORF-3 of 83aa). AUG initiation codon of sORF-3 was located at position 11433 nt, 89 nucleotides after stop codon of sORF-2.

4.1. Sequence Identities and Phylogenetic Analysis

RTSV-SP ORF-1 shared 89.54-95.73% nt identity and 96.51-97.85% aa identity to five other available isolates. Although RTSV-PhilA and RTSV-Vt6 are both originated from Philippines, ORF-1 of RTSV-SP showed a higher percentage of identity at amino acid level with PhilA compared to Vt6 isolate. This difference could be clearly observed through phylogenetic analysis conducted based on the ORF-1 amino acid sequences of all six isolates **(Fig. 1)**. Each taxon in the figure is shown as isolate accession number and abbreviated name. Clade 1 consisted of AM234048 Ori: India, AM234049 WB: India, KC794785 AP: India, and AB064963 Vt6: Philippines isolates. Clade 2 comprised of MK655459 SP: Malaysia and NC 001632 PhilA: Philippines isolates.

RTSV-SP sORF-2 is slightly shorter compared to PhilA isolate, by 27 amino acid residues since it is interrupted by a stop codon, UGA at position 11343 nt **(Fig. 2)**. In **Figure 2**, the Andhra Pradesh (AP), Orissa (Ori), West Bengal (WB), Philippines (PhilA and Vt6) isolates are denoted in abbreviated name. Meanwhile, sORF-3 was highly identical to PhilA isolate (96.43% nt and 91.57% aa).

Identities of the complete genome of RTSV-SP with other RTSV isolates were 90.31%-96.04% at nucleotide level, where highest similarity was shared with PhilA isolate (Accession number: NC 001632) **(Table 4).** This identity was also indicated through phylogenetic analysis based on the whole nucleotide sequences of all six isolates. Maximum likelihood tree generated formed two major clades; first consist of SP, PhilA and Vt6 isolates, second includes AP, WB and Ori isolates **(Fig. 3)**. The positing of RTSV-SP in SA group was supported by a bootstrap value of 100%. **Figure 3** shows each taxon as isolate accession number and abbreviated name. The Southeast Asia clade (SEA) comprised of MK655459 SP: Malaysia, NC 001632 PhilA: Philippines, and AB064963 Vt6: Philippines isolates. South Asia clade (SA) comprised of AM234048 Ori: India, AM234049 WB: India, and KC794785 AP: India isolates.

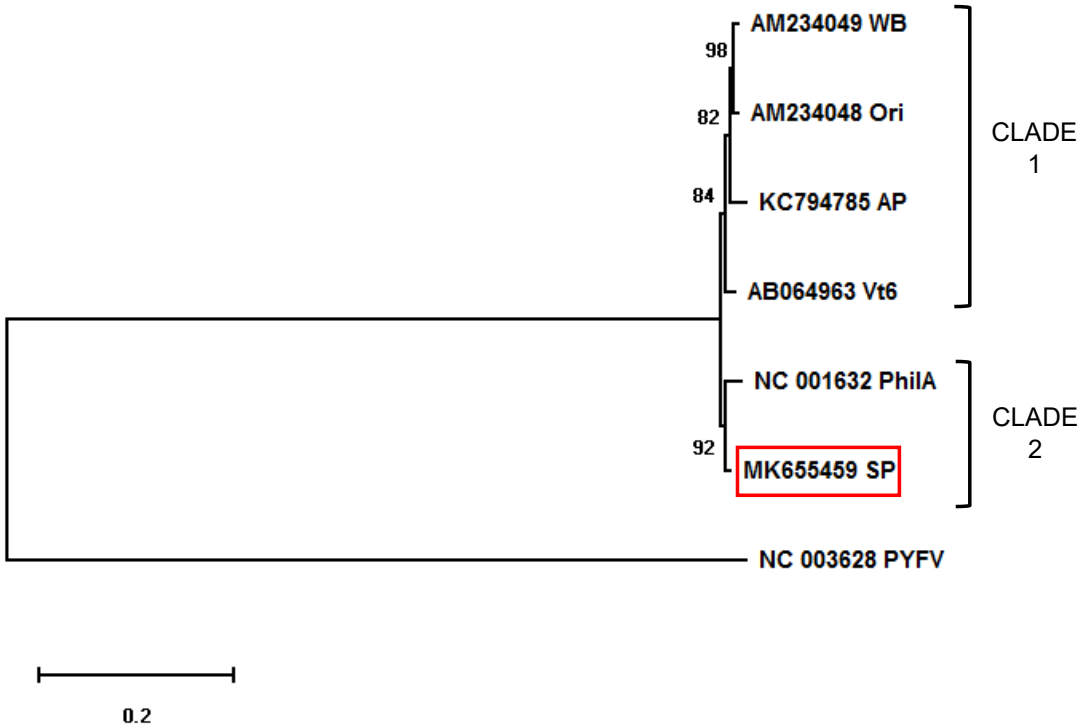

Figure 1. Phylogenetic relationship based on amino acid sequences of ORF1 of RTSV-SP (red box) with previously reported SEA and SA isolates. Amino acid sequence of the single ORF of Parsnip yellow fleck virus (accession no. NC 003628) was included as outgroup. Bootstrap values were calculated for 1000 replicates (cut-off value is 80%).

```
AP      MRFKRSVSEPTAAERSTLNRSNPHTTADIIRECGQHAV---------------------- 38
Ori     MRFKRSVSEPTAAERSTLNRLNPHTTADIIRECGQHAI---------------------- 38
WB      MRFKRSVSEPTAAERSTLNRLNPHTTADIIRECGQHAI---------------------- 38
PhilA   MRFKRSVSEPTAAERSTLNRLNPHTTADIIGECGQHVTRFYASGELFRTDYCTSLEEPGA 60
SP      MRFKRSVSEPTAAERSTLNRLNPHTTADIIGECGQHVTRFYASGELFRTDYCTSLEEPGA 60
Vt6     MRFKRSVSEPTAAERSTLNRLNPHTTADIIGECGQHVI---------------------- 38

AP      ------------------------------------    38
Ori     ------------------------------------    38
WB      ------------------------------------    38
PhilA   TLAVHGHGTQTLTQVLKYRSVYAGLEWHCYLVMYILLRQ    99
SP      TLAVHGHGTHKP--------------------------    72
Vt6     ------------------------------------    38
```

Figure 2. Amino acid sequence alignment of sORF-2 coded by six RTSVs. The grey box indicates 100% similarity of amino acid among six isolates. The yellow box indicates >50% amino acid sequence identity among all isolates, pink represents 100% ID among Indian isolates while green represents 100% similarity among Philippines isolates.

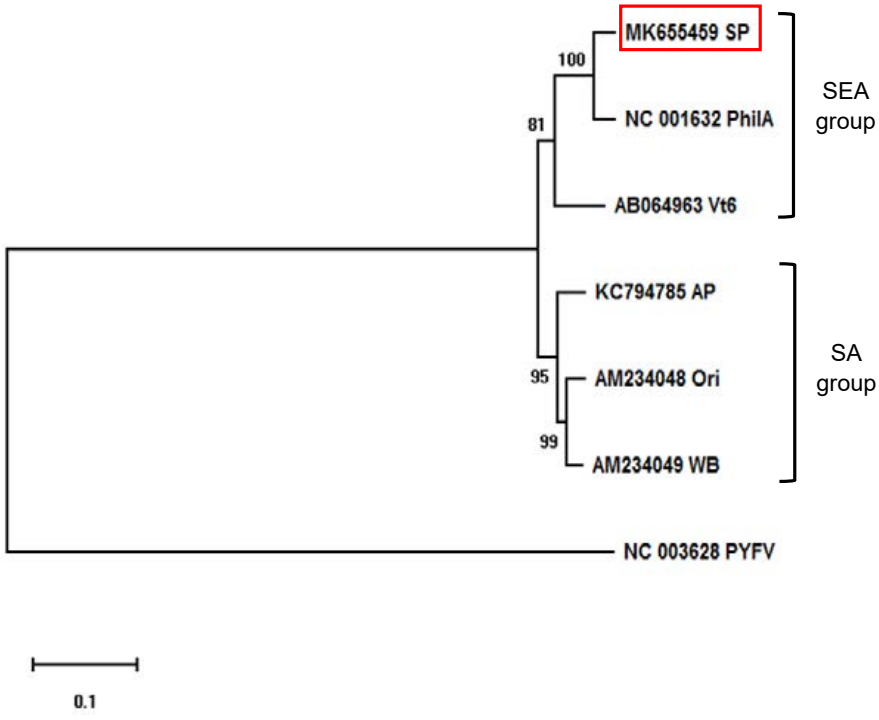

Figure 3. Comparative analysis of complete nucleotide sequences of RTSVs. RTSV-SP is indicated by a red box. The complete genome sequence of Parsnip yellow fleck virus (PYFV, accession no. NC 003628) was included as outgroup. Numbers at the branching points indicate the bootstrap support calculated for 1000 replicates (cut-off value is 80%).

Table 4. Percent identity of nucleotide (NT) and amino acid (AA) sequences between genome segments of RTSV-SP and available RTSV reference isolates.

Genome	PhilA/SP		Vt6/SP		Ori/SP		WB/SP		AP/SP	
segment	nt	aa	nt	aa	nt	aa	nt	aa	nt	aa
ORF-1	95.73	97.85	89.82	97.40	89.54	97.31	89.75	97.45	89.77	96.51
sORF-2	94.52	95.83	96.58	97.37	95.73	92.11	95.73	92.11	94.02	89.47
sORF-3	96.43	91.57	95.63	87.95	95.24	91.57	94.05	87.95	94.44	89.16
Complete sequences	96.04	-	90.59	-	90.31	-	90.51	-	90.45	-

5. Discussion

RTD has been circulating in many regions of Asian countries. Currently, only five RTSV isolates have been sequenced completely. The virus needs to be further explored in the aspect of isolation and molecular characterization of more isolates from different geographical origins in order to better understand the importance of RTSV diversity and manage tungro disease more effectively. In present study, the full genomic sequences of a Penang isolate (RTSV-SP) were analyzed to characterize Malaysian RTSV and the phylogenetic relationship among available RTSV isolates including RTSV-SP was elucidated.

The genome length of RTSV-SP is 12,173 bp, the second shortest among all isolates reported to date. ORF1 of RTSV-SP encoded P1, three CPs, NTP, Pro and a Rep. These features are in agreement with other published isolates; PhilA, Vt6, Ori, WB and AP. P1 of RTSV-SP consists of 640 aa, similarly to Vt6, AP, WB and Ori but slightly shorter than PhilA (643 aa). P1 has a potential of involving in RTSV transmission by the vector (26).

CP1 (624 nt), CP2 (609 nt) and CP3 (876 nt) of SP genome are in consistent with five reported isolates. High similarities between CP1-3 of SP and PhilA 92.0%, 89.3% and 92.6% nt respectively were observed in this study. This is supported by previous study in which coat protein genes of a RTSV isolate (MaP1) from Malaysia were sequenced and compared to the CPs of a Philippine isolate (PhL1) (34). It reported similarly high nucleotide identities of CP1-3 between both isolates; 96.8%, 98.5% and 92.7% respectively. Moreover, they also suggested MaP1 to be closely related to PhL1 isolate since the immunological responses exhibited by proteins of both MaP1 and PhL1 isolates in the western blot test using polyclonal antisera specific to PhL1 RTSV were similar (34). Molecular masses of CP1-3 predicted to be 22.5, 22.0 and 33.0 kDa respectively (45). Among the coat proteins, CP3 gene is probably being the major antigenic determinant on the surface of virus particles that involved in host-virus interaction (46,47).

RTSV-SP consisted of a putative sORF-2 of 72 aa, longer than RTSV isolates of Vt6, WB, Ori and AP by 34 aa (25–27). The sORF-2 in RTSV-SP is interrupted by a stop codon at 11341 nt. This interruption of sORF-2 in RTSV-SP is constant with previous report by Thole & Hull (48), in which the Philippines, Thailand and Indian isolates sequenced were found to contain truncated sORF-2 due to the presence of several stop codons. Thole & Hull (48) suggested that those short ORFs might not translated, thus the virus have an unusually long 3' non-coding region. According to Niazi (26), the Indian isolates were able to overcome the resistance of TKM6 variety as Vt6 while PhilA could not. Hence, as suggested by Isogai et al. (25), sORF-2 is probably responsible for the difference in viral virulence of PhilA isolate with Vt6 and Indian isolates. Construction of a full-length infectious cDNA clone and virulence testing on TKM6 cultivar with SP isolate would be helpful in the identification of respective protein contributing to this hyper-virulence.

Small ORF-3 of RTSV-SP has a length of 83 aa, in agreement with other 5 isolates. The sORF-3 has its own start codon, AUG placed after 92 nucleotides from sORF-2 stop codon. These sORFs were assumed to have encoded movement protein since the protein was not located in the RTSV RNA genome yet. On the other hand, the sORFs might be involving in helper functions needed for RTBV transmission (24).

Phylogenetic analysis based on the whole genomes strongly classified RTSV-SP into SEA group comprised of two Philippines RTSV isolates (RTSV-PhilA and RTSV-Vt6). Previous study by Verma & Dasgupta (26) indicated that Ori and WB clustered closer to Vt6 than to PhilA. Apart from that, Sailaja et al. (27) showed the comparative analysis of RTSV full-length sequences formed three major clades; PhilA in the first branch, Vt6 in second and Indian isolates in third group. In this study, the obtained clustering pattern, separate branching of RTSV isolates into Southeast Asia (SEA) and South Asia (SA) groups clearly reinforced the influence of geographical location on virus genomic evolution. At the same time, RTSV-SP is much closer to PhilA but distant from Vt6 isolate. Previously, the phylogenetic analysis based on non-recombinant regions in RTSV RNA genome; NTP, CPs, Pro, and Rep firmly indicated the divergence of Indian and Philippines isolates. Hence, the recombination within and between Indian and Philippines isolates was suggested to be a dominant factor of RTSV evolution (27).

6. Conclusion

In this study, we report the first full-length genome of RTSV isolate from Malaysia. By comparing the whole genome obtained with other RTSV isolates, a basic knowledge on the genomic structure, diversity and evolution of RTSV-SP has been achieved. Present study revealed that RTSV-SP complete genome shares the highest identity with the PhilA isolate from Philippine and the phylogenetic tree generated confidently classified RTSV-SP into South East Asia cluster. Recombination analysis of RTSV-SP may assist in deeper understanding on the divergence of Indian and Southeast Asian isolates. Moreover, further exploration

of sORF-2 would be helpful in studying its function in the viral virulence. Availability of this complete genomic sequences could also contribute to future research studies on RTSV genetics, epidemiology, and virology.

Conflict of Interest

The authors declare no conflict of interest with this study.

Acknowledgments

The authors would like to thank MARDI Seberang Perai for providing samples for this study.

References

1. Azgar MA, Yonzone R. Studies on the incidence of Rice-Tungro Virus (RTV) and population dynamic of its predominant vector. *Int J Curr Microbiol App Sci*. 2018;**7**(03):778-786. doi: 10.20546/ijcmas.2018.703.091
2. Bunawan H, Dusik L, Bunawan SN, Amin NM. Rice Tungro Disease : From Identification to Disease Control. *World Appl Sci J*. 2014;**31**(6):1221-1226. doi: 10.5829/idosi.wasj.2014.31.06.610
3. Le DT, Netsu O, Uehara-Ichiki T, Shimizu T, Choi IR, Omura T, *et al*. Molecular detection of nine rice viruses by a reverse-transcription loop- mediated isothermal amplification assay. *J Virol Methods*. 2010;**170**(1–2):90-93. doi: 10.1016/j.jviromet.2010.09.004
4. Zeigler RS, Barclay A. The relevance of rice. 2008;**1**(1):3-10. doi: 10.1007/s12284-008-9001-z
5. Ou SH. Rice disease. 2nd ed. Kew: Commonwealth Mycological Institute; 1985. p. 380.
6. Khatun MT, Latif MA, Rahman MM, Hossain M, Ansari TH, Nessa B, *et al*. Recovering ability of upland and rainfed lowland rice varieties against rice tungro disease. *Bangladesh Rice J*. 2017;**21**(1):91-100. doi: 10.3329/brj.v21i1.37390
7. Anjaneyulu A, Satapathy MK, Shukla VD. Rice tungro. Lebanon: Science Publisher Inc; 1995. p. 228.
8. Azzam O, Arboleda M, Umadhay KML, De los Reyes JB, Cruz FS, Mackenzie A, *et al*. Genetic composition and complexity of virus populations at tungro-endemic and outbreak rice sites. *Arch Virol*. 2000;**145**(12):2643-2657. doi: 10.1007/s007050070013
9. Muralidharan, K. Krishnaveni, D. Rajarajeswari, N. V. L. Prasad ASR. Tungro epidemics and yield losses in paddy fields in India. *Curr Sci*. 2003;**85**(8):1143-1147.
10. Dai S, Beachy RN. Genetic engineering of rice to resist rice tungro disease. *Vitr Cell Dev Biol*. 2009;**45**(5):517-524. doi: 10.1007/s11627-009-9241-7
11. Chancellor TCB, Azzam O, Heong KL, editors. Rice tungro disease management. In: Proceedings of the International Workshop on Tungro Disease Management. Makati: Internation Rice Research Institute; 1999. p. 166.
12. Hibino H, Roechan M, Sudarisman S. Association of two types of virus particles with penyakit habang (tungro disease) of rice in Indonesia. *Phytopathology*. 1978;**68**(10):1412-1416. doi: 10.1094/phyto-68-1412
13. Jones MC, Gough IK, Dasgupta I, Rao BS, Cliffe J, Qu R, *et al*. Rice tungro disease is caused by an RNA and a DNA virus. *J Gen Virol*. 1991;**72**(4):757-761. doi: 10.1099/0022-1317-72-4-757
14. Hay JM, Jones MC, Biakebrough ML, Dasgupta I, Davies JW, Hull R. An analysis of the sequence of an infectious clone of rice tungro bacilliform virus , a plant pararetrovirus. *Nucleic Acids Res*. 1991;**19**(10):2615-2621. doi: 10.1093/nar/19.10.2615
15. Blas N, David G. Dynamical roguing model for controlling the spread of tungro virus via Nephotettix Virescens in a rice field. *J Phys Conf Ser*. 2017;**893**(1):012018. doi: 10.1088/1742-6596/893/1/012018
16. Cabauatan PQ, Hibino H. Transmission of rice tungro bacilliform and spherical viruses by Nephotettix virescens Distant. *Philipp Phytopathol*. 1985;**21**:103-109 doi: 10.3186/jjphytopath.49.545
17. Hibino H. Insect-borne viruses of rice. In: Harris KF, editor. Advances in disease vector research. New York: *Springer*; 1990. p. 209-241. doi: 10.1007/978-1-4612-3292-6_8
18. Habibuddin H, Hadzim K, Othman O, Azlan S. Y 1286 is a Balimau Putih-derived rice line resistant to rice tungro bacilliform and spherical viruses. *J Trop Agric food Sci*. 2000;**28**:13-22.
19. Hibino H. Biology and epidemiology of rice viruses. *Annu Rev Phytopathol*. 1996;**34(**1):249-274. doi: 10.1146/annurev.phyto.34.1.249
20. Rivera CT, Ou SH. Transmission studies of the two strains of rice tungro virus. *Plant Dis Rep*. 1967;**51**:877-881.
21. Cruz FS, Hull R, Azzam O. Changes in level of virus accumulation and incidence of infection are critical in the characterization of Rice tungro bacilliform virus (RTBV) resistance in rice. *Arch Virol*. 2003;**148**(8):1465-1483. doi: 10.1007/s00705-003-0117-6
22. Azzam O, Chancellor TCB. The biology epidemiology and managementof RTD in Asia. *Plant Dis*. 2002;**86**(2):88-100. doi: 10.1094/PDIS.2002.86.2.88
23. Hull R. Rice tungro bacilliform virus. In: Webster AG and RG, editor. Encyclopedia of Virology. 2nd ed. Cambridge: Academic Press; 1999. p. 1292-1296. doi: 10.1006/rwvi.1999.0330
24. Shen P, Kaniewska M, Smith C, Beachy RN. Nucleotide sequence and genomic organization of rice tungro spherical virus. *Virology*. 1993;**193**(2):621-630. doi: 10.1006/viro.1993.1170
25. Isogai M, Cabauatan PQ, Masuta C, Uyeda I, Azzam O. Complete nucleotide sequence of the rice tungro spherical virus genome of the highly virulent strain Vt6. *Virus Genes*. 2000;**20**(1):79-85. doi: 10.1023/a:1008116408733
26. Verma V, Dasgupta I. Sequence analysis of the complete genomes of two Rice tungro spherical virus isolates from India. *Arch Virol*. 2007;**152**(3):645-648. doi: 10.1007/s00705-006-0861-5
27. Sailaja B, Anjum N, Patil YK, Agarwal S, Malathi P, Krishnaveni D, *et al*. The complete genome sequence of a south Indian isolate of Rice tungro spherical virus reveals evidence of genetic recombination between distinct isolates. *Virus Genes*. 2013;**47**(3):515-523. doi: 10.1007/s11262-013-0964-5
28. Habibuddin H, Hadzim K, Othman O, Imbe T, Omura T. Selection of rice line Y1036 resistant to the green leafhopper and tungro disease. *MARDI Res J*. 1991;**19**(2):169-175.
29. Suranto S, Arief A, Supyani S. The use of electrophoretic isozymes to detect tungro infected rice. *J Agric Sci*. 2017;**39**(2):145-152. doi: 10.17503/agrivita.v39i2.643
30. Zhang S, Jones MC, Barker P, Davies JW, Hull R. Molecular

cloning and sequencing of coat protein-encoding cDNA of rice tungro spherical virus—a plant picornavirus. *Virus Genes*. 1993;**7**(2):121-132. doi: 10.1007/bf01702392

31. Verma V, Sharma S, Devi SV, Rajasubramaniam S, Dasgupta I. Delay in virus accumulation and low virus transmission from transgenic rice plants expressing Rice tungro spherical virus. *Virus Genes*. 2012;**45**(2):350-359. doi: 10.1007/s11262-012-0787-9
32. Hull R. Molecular biology of rice tungro viruses. *Annu Rev Phytopathol*. 1996;**34**(1):275-297. doi: 10.1146/annurev.phyto.34.1.275
33. Kumar GKD, Dasgupta I. Molecular biology of rice tungro viruses and strategies for their control: Molecular approaches. In: Shamim M, Singh KN, editors. Biotic Stress Management in Rice. 1st ed. Apple Academic Press; 2016. p. 15. doi: 10.1201/9781315365534-1
34. Zhang S, Davies JW, Hull R. Sequences of the three coat protein genes of a Malaysian isolate of rice tungro spherical virus reveal a close relationship to the Philippine isolate. *Virus Genes*. 1997;**15**(1):61-64. doi: 10.1007/s007050050205
35. Rice tungro spherical virus genome (cited 2019 Mar 19); Available from: URL: https://www.ncbi.nlm.nih.gov/
36. Nishimura A, Morita M, Nishimura Y, Sugino Y. A rapid and highly efficient method for preparation of competent Escherichia coli cells. *Nucleic Acids Res*. 1990;**18**(20):6169. doi: 10.1093/nar/18.20.6169
37. Multiple Sequence Alignment [cited 2019 Mar 18]; Available from: URL: https://www.ebi.ac.uk/Tools/msa/clustalo/
38. Basic Local Alignment Search Tool (cited 2019 Mar 20); Available from: URL: https://blast.ncbi.nlm.nih.gov/Blast.cgi
39. Smart Blast (cited 2019 Mar 20); Available from: URL: https://blast.ncbi.nlm.nih.gov/smartblast/smartBlast.cgi
40. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994;**22**(22):4673-4680. doi: 10.1093/nar/22.22.4673
41. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X : Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol*. 2018;**35**(6):1547-1549. doi: 10.1093/molbev/msy096
42. Tamura K. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+ C-content biases. *Mol Biol Evol*. 1992;**9**(4):678-687. doi: 10.1093/oxfordjournals.molbev.a040752
43. Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*. 1987;**4**:406-425. doi: 10.1093/oxfordjournals.molbev.a040454
44. Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. In: BrysonV, Vogel HJ, editors. Evolving Genes and Proteins. New York:Academic Press; 1965. p. 97-166. doi: 10.1016/b978-1-4832-2734-4.50017-6
45. Mangrauthia SK, Malathi P, Agarwal S, Ramkumar G, Krishnaveni D, Neeraja CN, *et al*. Genetic variation of coat protein gene among the isolates of Rice tungro spherical virus from tungro-endemic states of the India. *Virus Genes*. 2012;**44**(3):482-487. doi: 10.1007/s11262-011-0708-3
46. Druka A, Burns T, Zhang S, Hull R. Immunological characterization of rice tungro spherical virus coat proteins and differentiation of isolates from the Philippines and India. *J Gen Virol*. 1996;**77**(8):1975-1983. doi: 10.1099/0022-1317-77-8-1975
47. Mangrauthia SK, Malathi P, Balachandran SM, Reddy CS, Viraktamath BC. Global Analysis of Rice Tungro Spherical Virus Coat Proteins Reveals New Roles in Evolutionary Consequences. *J plant Biochem Biotechnol*. 2010;**19**(2):263-266. doi: 10.1007/bf03263353
48. Thole V, Hull R. Rice tungro spherical virus: nucleotide sequence of the 3′ genomic half and studies on the two small 3′ open reading frames. *Virus Genes*. 1996;**13**(3):239-246. doi: 10.1007/bf00366984