Peer-reviewed version available at *European Journal of Plant Pathology* **2020**, *156*, 1005-1014; doi:10.1007/s10658-020-01948-4

Article

Complete Genomic Sequence of *Noni mosaic virus* (NoMV), a Novel Potyvirus Associated with a Mosaic Disease in *Morinda citrifolia* L.

Nai-Tong Yu[1,4]* · Zhi-Ying Cai[2] · Zhongguo Xiong[3] · Yan Yang[2]* · Zhi-Xin Liu[1,4]*

[1] Key Laboratory of Biology and Genetic Resources of Tropical Crops, Ministry of Agriculture and Rural Affairs, Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China
[2] Yunnan Institute of Tropical Crops, Jinhong 666100, China
[3] School of Plant Sciences and BIO5 Institute, University of Arizona, Tucson, AZ 85721, USA
[4] Hainan Key Laboratory of Tropical Microbe Resources, Haikou 571101, China

* Correspondence: yunaitong@163.com (N.Y.); yyzy117@163.com (Y.Y.), liuzhixin@itbb.org.cn (Z. L.).

Abstract: An outbreak of a virus-like disease has caused severe damage to noni plants (*Morinda citrifolia* L.) in Xishuangbanna area of China's southwestern Yunnan province since 2015. The diseased plants displayed typical mosaic symptom with light and dark green patches on leaves. Flexuous filamentous virus particles of about 800 nm in length were observed from the leaf saps by transmission electron microscope. Illumina transcriptomic sequencing further revealed the presence of a potyvirus and its near complete genome was obtained from *de novo* assembly. The complete genome of 9,659 nts was obtained by Sanger sequencing of eight amplicons generate by RT-PCR and 5' and 3' RACE. BLASTp analysis of the polyprotein sequence showed that the virus was most closely related to *Tobacco vein banding mosaic virus* (TVBMV), but these two viruses only shared 50.7% amino acid sequence similarity. Both phylogenetic analyses of the polyprotein and CP amino acid sequences indicated that this virus is a member of genus *Potyvirus*. However, the low sequence homology with all known potyviruses established this virus as a new species in the genus, tentatively named as *Noni mosaic virus* (NoMV). Our field surveys showed that 100% of the symptomatic samples and 28.57% of the asymptomatic samples were infected with this novel potyvirus. Aphids collected from diseased leaves were also detected carrying the virus. In summary, our data indicated that a novel species of potyvirus, NoMV, is prevalent in Yunnan, China and is associated with an emerging mosaic disease on *M. citrifolia*.

Keywords: *Noni mosaic virus;* Potyvirus; Complete genome; *Morinda citrifolia* L; Illumina transcriptomic sequencing

1. Introduction

Plant viruses have seriously threat to crop production and human food safety due to their prevalence and outbreak. According to incomplete statistics, plant virus infection annually causes about 10% reduction of global crop production [1]. The development of Next Generation Sequencing (NGS) has provided a powerful tool for plant pathogens diagnosis, especially for the novel virus identification. Many plant viruses have been discovered by the NGS method, such as *grapevine Muscat rose virus* (GMRV) [2], *Apple rubbery wood virus* (ARWV) [3], *Tea plant necrotic ring blotch virus* (TPNRBV) [4] and *Areca palm necrotic ringspot virus* (APNRV) [5].

Noni (*Morinda citrifolia* L.) is a fruit-bearing tree with its native range extending across tropical or sub-tropical regions of Southeast Asia, Australia and the Pacific Islands. Its fruits are traditionally used as a medicinal herb in many countries [6, 7]. The agricultural planting of noni began as early as 2000 in Hainan and Yunnan Provinces, China. At present, it has been grown in most provinces in South China. With the increased acreages and continuous cropping, diseases such as blight and fruit rot caused by *Phytophthora botryosa* [8] and anthracnose caused by *Colletotrichum* spp. [9] have been reported and may become threats to the development of the noni industry. However, no viral disease of noni has been reported up to now.

In 2015, a virus-like disease was found in noni plants in Xishuangbanna, Yunnan, China. The leaves of diseased plants had striking mosaic symptoms with light and dark green patches. At present, this virus-like disease seriously affects the cultivation and production of noni plant in Xishuangbanna. In such cases, identification of the viral pathogen present in diseased noni plants is significant for scientific and agronomic interest. In this study, NGS technology and Sanger sequencing were used to demonstrate that the putative causal agent of this noni disease is a novel potyvirus with distinctive molecular characteristics. This study is the first report of the novel potyvirus, tentatively named as *Noni mosaic virus* (NoMV) worldwide.

2. Materials and Methods

2.1 Plant materials and electron microscopy observation

In 2015, thirty-one leaf samples of diseased noni plants (*Morinda citrifolia* L.) showing typical mosaic symptoms with light and dark green patches (Fig. 1a) were collected from Xishuangbanna, Yunnan, China. To determine if the diseased plants were infected by virus, leaf-saps were prepared for transmission electron microscope (TEM) examination. Briefly, five fresh leaves of healthy or diseased samples were grounded in 1×PBS at a final concentration of 0.1 g/ml. The grounded samples were centrifuged at 5 000 rpm for 2 minutes to obtain the supernatants, which were then loaded onto copper grids (200 meshes) individually. The copper grids were negatively stained by 1% phosphotungstic acid for 2 minutes, dried under tungsten lamp for about 10 min, and then observed under TEM (HT7700, Hitachi). The width and length of viral particles were measured by using Adobe Photoshop CS3 software.

2.2 Library preparation for transcriptome sequencing

Total RNA was extracted from five diseased noni leaf samples by using a Quick RNA Isolation Kit (Bioteke, Beijing, China) according to the manufacturer's instructions. RNA quality, including purity, concentration, and integrity, were confirmed by using Nanodrop, Qubit 2.0, and Agilent 2100 before processing to cDNA library preparation.

The cDNA library was prepared in the following steps. First, mRNA was enriched from total RNA by Oligo(dT) magnetic beads and then randomly interrupted by adding fragmentation buffer. The fragmented mRNA was primed by random hexamers and reverse-transcribed into first strand cDNA, which was then treated with RNase H to remove RNA and used as a template for second strand cDNA (ds cDNA) synthesis using DNA polymerase I. The ds cDNA was purified by using AMPure XP beads and subjected to end repair and dA-tailing. Subsequently, the adaptors were added into the ds cDNA and the cDNA library was further enriched by PCR amplification. Before high-throughput sequencing, the concentration and insert size of the cDNA library were analyzed by Qubit 2.0 and Agilent 2100, respectively. Finally, the library was sent to Biomarker Biotechnology Corporation (Beijing, China) for deep sequencing, which was performed on HiSeq4000 with paired-end read length at 150 bp.

2.3 Viral genome assembly

Raw data was cleaned by filtering low-quality reads and trimming the adaptors. High-quality clean reads were mapped to viral sequences downloaded from NCBI website (https://www.ncbi.nlm.nih.gov/) using the TopHat software [10]. HTSeq v0.5.4p3 was used to count the number of reads mapped to viral sequences [11]. Contigs were assembled *de novo* from the clean reads using Trinity 2.1.1 [12]. These assembled contigs or unigenes were used to remap the viral reads for the second to obtain more accurate results. The viral contigs were further assembled into the full length potyvirus-like sequence in CodonCode Aligner 6.0.2 (CodonCode, Centerville, MA). The resulting full length potyvirus-like sequence was subjected to BLASTx search against non-redundant protein database (GenBank).

2.4 Sanger genome sequencing of a novel potyvirus

Sanger sequencing was used to confirm the genome sequence of the novel virus, designated as *Noni mosaic virus* (NoMV), arising from deep sequencing and *de novo* assembly. Briefly, seven primer pairs (see Supplementary Table S1) covering the near complete genome of NoMV were designed based on the assembled putative genome. Similarly, two nested primer pairs targeting 5' end were also designed. RNA extraction was conducted as described earlier. Random hexamer and Oligo(dT) were used in reverse transcription. PCR was

carried out using Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific). The 5' end fragment was amplified using a 5' Rapid Amplification of cDNA Ends (RACE) kit (Invitrogen) according to the manufacturer's instructions. Each PCR fragment was cloned into pMD18-T vector (Takara, Beijing, China), and three independent clones of each fragment were subjected to Sanger bidirectional sequencing (Sangon Biotech, Wuhan, China). High quality sequencing results were overlapping-assembled using BioEdit (version 7.0.9.0) to obtain the complete genome of NoMV [13].

2.5 NoMV genome analysis

The complete genome of NoMV obtained from previous step was subjected to sequence analyses. First, putative ORFs of NoMV were identified using ORF finder (https://www.ncbi.nlm.nih.gov/orffinder/) and sequence comparison with closely related viruses. Identified ORFs along with amino acid sequence were then used in BLAST. Sequences with high similarity to NoMV were downloaded and subjected to alignment and pairwise comparison using BioEdit (version 7.0.9.0).

2.6 Phylogenetic analysis

To determine the phylogenetic relationship of NoMV with *Potyvirus*, phylogenetic trees were constructed based on the alignments of polyprotein and coat protein (CP). Twenty potyviruses were chosen for alignment with NoMV because of their close relationship with NoMV revealed by BLAST search. These potyviruses are *Narcissus yellow stripe virus* (NYSV, AFJ92907.2), *Narcissus late season yellows virus* (NLSYV, AFQ95552.1), *Narcissus virus 1* (NV-1, BBE01240.1), *Wild onion symptomless virus* (WOSV, YP_009259366.1), *Scallion mosaic virus* (ScaMV, NP_570725.1), *Turnip mosaic virus* (TuMV, BBA07429.1), *Japanese yam mosaic virus* (JYMV, AJD23399.1), *Asparagus virus 1* (AV-1, AIY55493.1), *Lily mottle virus* (LMoV, ADO34171.1), *Sweet potato latent virus* (SPLV, AJS10748.1), *Jasmine virus T* (JaVT, APZ75429.1), *Yam mosaic virus* (YaMV, AYD60113.1), *Carrot thin leaf virus* (CTLV, AGH25889.1), *Panax virus Y* (PnVY, YP_003725718.1), *Celery mosaic virus* (CeMV, YP_004376199.1), *Apium virus Y* (ApVY, QAA06935.1), *Potato virus A* (PoVA, ADA57721.1), *Pokeweed mosaic virus* (PkMV, AFS28881.1), *Tobacco vein banding mosaic virus* (TVBMV, AEB66864.1), and *Potato virus Y* (PVY, ASI37712.1). All amino acid sequences were aligned using ClustalX, then passed to MEGA 6.0 for tree building using Neighbor-Joining method with 1,000 bootstrap replicates [14].

2.7 Field survey

To investigate the prevalence of the viral disease in Xishuangbanna City, Yunnan Province, three field surveys were conducted from March to May in 2016. Leaf samples of twenty-one asymptomatic and 67 symptomatic noni plants were randomly collected, and 7 asymptomatic and 13 symptomatic samples were selected for RT-PCR detection using a pair of primers specifically targeting NoMV *CP* gene (NMV-F and NMV-R, see in Supplementary Table S1).

During field survey, aphids and whiteflies found on the back of noni plant leaves were also collected for NoMV detection to determine if aphids or whiteflies were carriers of the NoMV. Approximately 50 aphids or 100 whiteflies were pooled to extract total RNA using TRIzol reagent. RT-PCR detection of NoMV was performed as described above.

3. Results

3.1 Potyvirus-like particles were found in disease noni leaf sap

In 2015, leaves of thirty-one noni plants displaying typical mosaic symptoms with light and dark green patches (Fig. 1a) were collected from Xishuangbanna, Yunnan, China. In order to determine if the diseased plants were infected with a virus, leaf-sap of the diseased and healthy leaves were prepared and observed under transmission electron microscope (TEM). The results showed that viral particles of typical potyviruses were observed in the saps of the diseased but not from healthy leaves. These flexuous filamentous particles were at 800±20 nm in length and 20±1 nm in wide (Fig. 1b). Some of the virus particles formed large aggregates (Fig. 1c). No virus particle was found from the healthy samples.

Peer-reviewed version available at European Journal of Plant Pathology 2020, 156, 1005-1014; doi:10.1007/s10658-020-01948-4

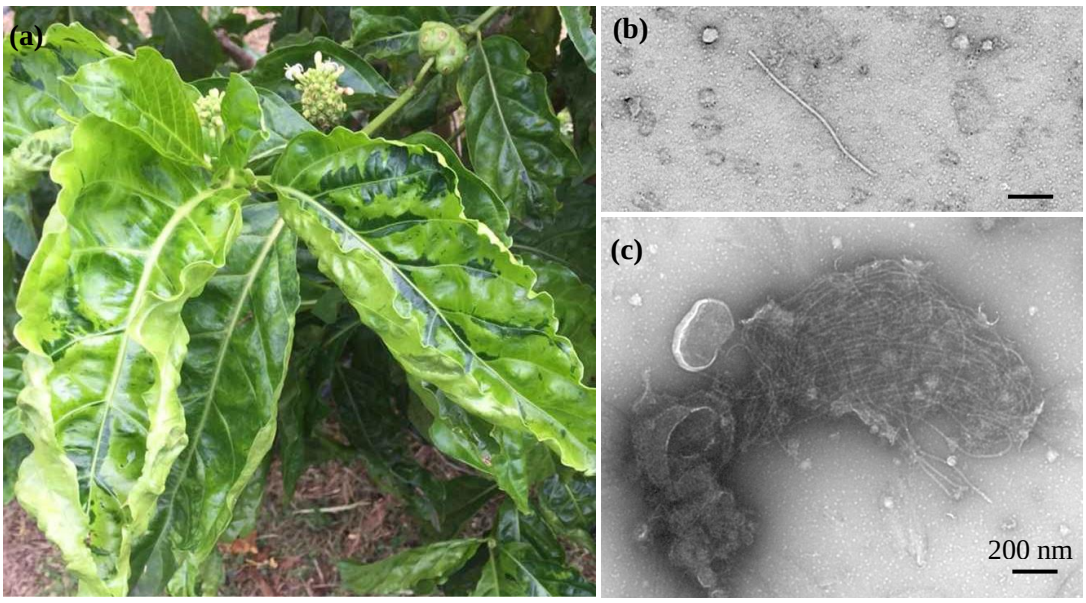

Figure 1. Symptoms and viral particle morphology associated with Noni mosaic disease. **(a),** Noni plants leaves showing mosaic symptoms; (**b**) and (**c**), Transmission electron micrographs of a single virion and aggregated virions from crude extracts of NoMV-infected leaves. Scale bar=200.

3.2 A nearly full length potyvirus-like sequence was assembled from transcriptome sequencing

Raw data were filtered to obtain high-quality clean data, which were then mapped to the virus database in NCBI GenBank. A total of 644,392 viral reads were identified. The mapped viral reads were assembled into contigs or unigenes. A total of 40,911 contigs and 318 unigenes were assembled, respectively (Table 1). In detail, the length range of contigs in 200-300 bp accounted for 99.67%, while the length range of contigs in 300-2000 bp were 0.33%, and only two contig was longer than 2000 bp. Meanwhile, the length ranges for unigenes in 200-300 bp, 300-500 bp, and 500-1000 bp accounted for 60.69%, 23.90% and 12.26%, respectively. The length range of unigenes over 1000 bp was 3.15%, while only two unigenes over 2000 bp were obtained (Table 1). These assembled contigs or unigenes were used to remap the viral reads for the second round and obtain more accurate results.

The assembled viral contigs or unigenes were further extended into the full length potyvirus-like sequence as long as possible in the CodonCode Aligner 6.0.2 (CodonCode, Centerville, MA), and an 8832 bp sequence was finally obtained. Blastn search showed this 8832 bp sequence had significant similarity to potyviruses. Based on the general genome size of potyviruses (~9.7 kb), nearly full length of this potyvirus-like sequence was obtained, leaving approximately 850 bps of 5' terminal sequence to be further sequenced.

Table 1. Summary statistical of the assembly results of potyvirus-like reads

Length (bp)	Contig	Transcript	Unigene
200-300	40,778(99.67%)	222(59.04%)	193(60.69%)
300-500	86(0.21%)	94(25%)	76(23.90%)
500-1000	39(0.10%)	50(13.30%)	39(12.26%)
1000-2000	6(0.01%)	8(2.13%)	8(2.52%)
>2000	2(0.00%)	2(0.53%)	2(0.63%)
Total Number	40,911	376	318
Total Length	2,119,848	138,503	117,038
Mean Length	51.82	368.36	368.04

3.3 Determination of complete genomic sequence of NoMV

To further verify the sequence assembled from transcriptomic sequence and to obtain the missing 5' end sequence, an overlapping amplicon cloning strategy was used with a series of 8 sequential RT-PCR cloning runs. The two ends of the viral genome were obtained by 5' and 3' RACE. These fragments were Sanger

sequenced and assembled. The full length of this potyvirus-like sequence, which was designated as NoMV Yunnan isolate (NoMV-YN), comprises of 9,659 nt. This complete genome sequence was deposited in GenBank under accession number MN158696. NoMV-YN has typical genomic organization and structural characteristic of potyviruses. It has a 291-nt 5' untranslated region (UTR) and a 256-nt 3' UTR. It contains a large open reading frame (ORF) encoding a polyprotein of 3,026 amino acids (aa) residues with a calculated molecular mass of 343 kDa (Fig. 2).

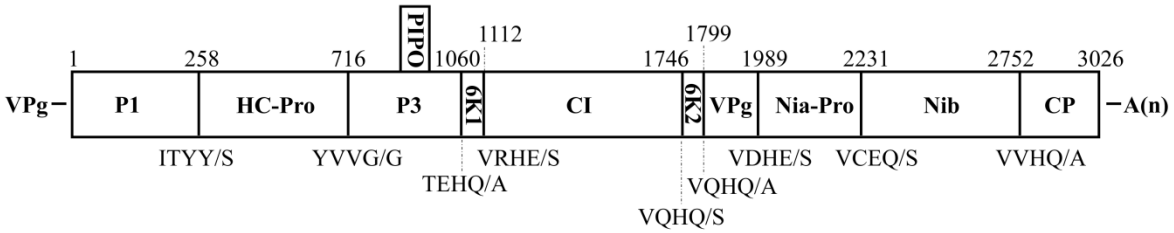

Figure 2. Schematic representation of the genomic structure of *Noni mosaic virus* (NoMV) and the predicted proteolytic cleavage sites of the NoMV polyprotein. Numbers above each part of the polyprotein indicate the total number of amino acids of the mature protein. The cleavage sites are shown below. P1, the first protein; HC-Pro, helper component-proteinase; P3, the third protein; 6K1 and 6K2, 6 kDa protein 1 and 2; CI, cytoplasmic inclusion protein; VPg, viral genome-linked protein; Nia-Pro, 49 kDa proteinase; Nib, nuclear inclusion protein b; CP, coat protein. PIPO (nucleotides 2901 to 3197) derived from RNA polymerase slippage on the P3 cistron.

3.4 NoMV polyprotein analysis

Pairwise comparison of NoMV and 20 closely related members in genus *Potyvirus* showed that this virus shared only 47-50.7% and 53-67% aa sequence identity with the polyproteins and CPs of other potyviruses, respectively (see Table S2). BLASTp analysis of the polyprotein showed that NoMV is most closely related to *Tobacco vein banding mosaic virus* (AEB66864.1), and the two viruses have 50.7% amino acid sequence similarity on polyprotein. Further analysis showed that NoMV polyprotein has typical structure and domains of a characteristic potyvirus. Nine highly conserved potyvirus proteolytic cleavage sites were found in the polyprotein based on potyvirus conservative protease cleavage sites. Ten putative mature proteins from cleavage are P1 protein (P1, 258 aa), helper-component proteinase (HC-Pro, 458 aa), P3 protein (P3, 344 aa), first 6 kDa peptide (6K1, 52 aa), cylindrical inclusion protein (CI, 634 aa), second 6 kDa peptide (6K2, 53 aa), viral protein genome-linked (VPg, 190 aa), nuclear inclusion "a" protease (NIa-Pro, 242 aa), nuclear inclusion "b" protease (NIb, 521 aa) and 274-aa CP (Fig. 2). The putative PIPO ORF (99 aa) was also identified as the conservative motif GAA AAA A which was also found at nt 2896-2902 [15]. In addition, the NoMV CP has 274 aa residuals with a calculated molecular mass of 30.69 kDa. The ^{2759}DAG^{2762} motif associated with aphid transmission was also found at the N terminus of CP (aa 7-9) [16].

Moreover, aa sequence comparison of each mature protein between NoMV and other 20 members of *Potyvirus* was also performed. The result showed that HC-Pro, CI, VPg, NIa-Pro and NIb proteins of NoMV shared relatively high aa sequence identities with those of other 20 members. The similarity ranges from 46% to 66%. However, the similarity of P1 and P3 are as low as 9-14% and 23-36%, respectively (Supplementary Table S3).

3.5 Phylogenetic analysis of NoMV with other potyviruses

To determine if NoMV is a novel potyvirus and how it relates to other potyviruses, a phylogenetic analysis of NoMV and the 20 potyviruses were performed based on the polyprotein amino acid sequences. The results indicated that NoMV together with other potyviruses formed a group, representing the *Potyvirus* genus. Further analysis indicated that NoMV was most closely related to TVBMV within a subgroup (Fig. 3a). The close relatedness between NoMV and TVBMV was confirmed by the phylogenetic analysis based on CP amino acid sequences (Fig. 3b). Although NoMV is closely related to the TVBMV, considering the low sequence similarities between these two viruses, it is clear that NoMV is a distinct species of *Potyvirus*. These results indicated that this virus isolate represents a novel species of potyvirus, tentatively named "*Noni mosaic virus* (NoMV)".

Peer-reviewed version available at European Journal of Plant Pathology 2020, 156, 1005-1014; doi:10.1007/s10658-020-01948-4

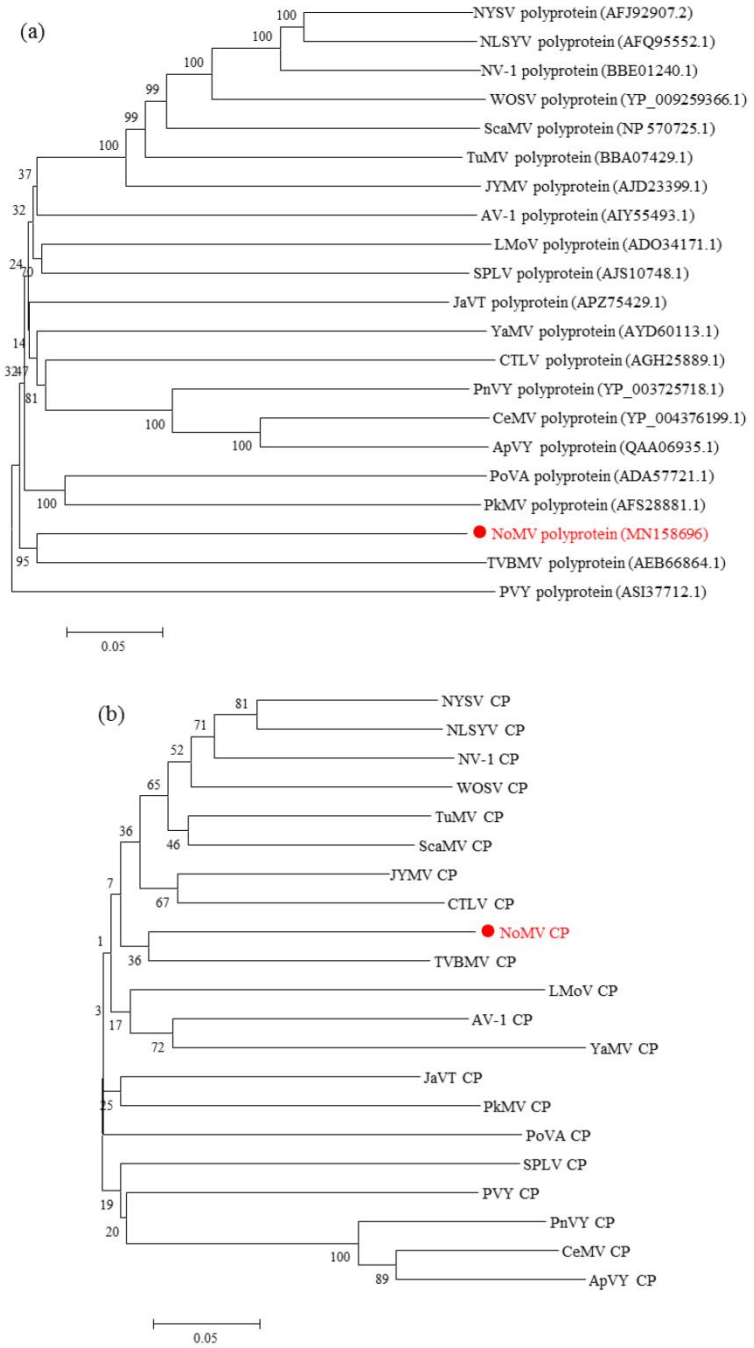

Figure 3. Phylogenetic analysis of deduced amino acid sequences of the polyprotein (**a**) and coat protein (**b**) of NoMV and other potyviruses. These sequences were aligned using ClustalX, and phylogenetic trees were constructed with MEGA 6.0 using the Neighbor-Joining method. The bar represents 0.05 substitutions per site. NoMV is highlighted in red.

3.6 NoMV is prevalent in noni fields

Three field surveys were conducted from March to May in 2016. The surveys showed that about 90% of noni plants were showing typical mosaic symptom at some planting areas. RT-PCR examination of the partial samples showed that 100% of the symptomatic samples and 28.57% of the asymptomatic samples were infected with NoMV (Table 2). These results showed that the widespread Noni mosaic disease (NoMD) is tightly associated with NoMV, and is mostly likely caused by this virus.

Table 2. *Noni mosaic virus* detection on partial noni plants by RT-PCR method

Sample	Collected time	Location	Symptom	RT-PCR
1-H1	2016-3-2	Yunnan, China	Asymptomatic	Negative
1-H2	2016-3-2	Yunnan, China	Asymptomatic	Negative
1-I1	2016-3-2	Yunnan, China	Mosaic	Positive
1-I2	2016-3-2	Yunnan, China	Mosaic	Positive
1-I3	2016-3-2	Yunnan, China	Mosaic	Positive
2-H1	2016-5-3	Yunnan, China	Asymptomatic	Positive*
2-H2	2016-5-3	Yunnan, China	Asymptomatic	Negative
2-H3	2016-5-3	Yunnan, China	Asymptomatic	Negative
2-I1	2016-5-3	Yunnan, China	Mosaic	Positive
2-I2	2016-5-3	Yunnan, China	Mosaic	Positive
2-I3	2016-5-3	Yunnan, China	Mosaic	Positive
2-I4	2016-5-3	Yunnan, China	Mosaic	Positive
2-I5	2016-5-3	Yunnan, China	Mosaic	Positive
3-H1	2016-5-25	Yunnan, China	Asymptomatic	Negative
3-H2	2016-5-25	Yunnan, China	Asymptomatic	Positive*
3-I1	2016-5-25	Yunnan, China	Mosaic	Positive
3-I2	2016-5-25	Yunnan, China	Mosaic	Positive
3-I3	2016-5-25	Yunnan, China	Mosaic	Positive
3-I4	2016-5-25	Yunnan, China	Mosaic	Positive
3-I5	2016-5-25	Yunnan, China	Mosaic	Positive

"*" represents NoMV infection of asymptomatic samples.

3.7 Aphid is a carrier and a potential transmission vector of NoMV

Filed surveys also showed that a large number of aphids (*Aphis atrata* Zhang) and whiteflies were observed on the underside of symptomatic leaves (Figure 4a and 4b). RT-PCR was performed to determine whether aphids or whiteflies could carry the virus and the results showed that a DNA band with expected size was amplified from two aphid RNA preparations, but not from the whitefly RNA preparation (Figure 4c). Sanger sequencing confirmed the amplified band was of NoMV *CP*. These results indicated that the aphid is a carrier and a potential transmission vector of NoMV.

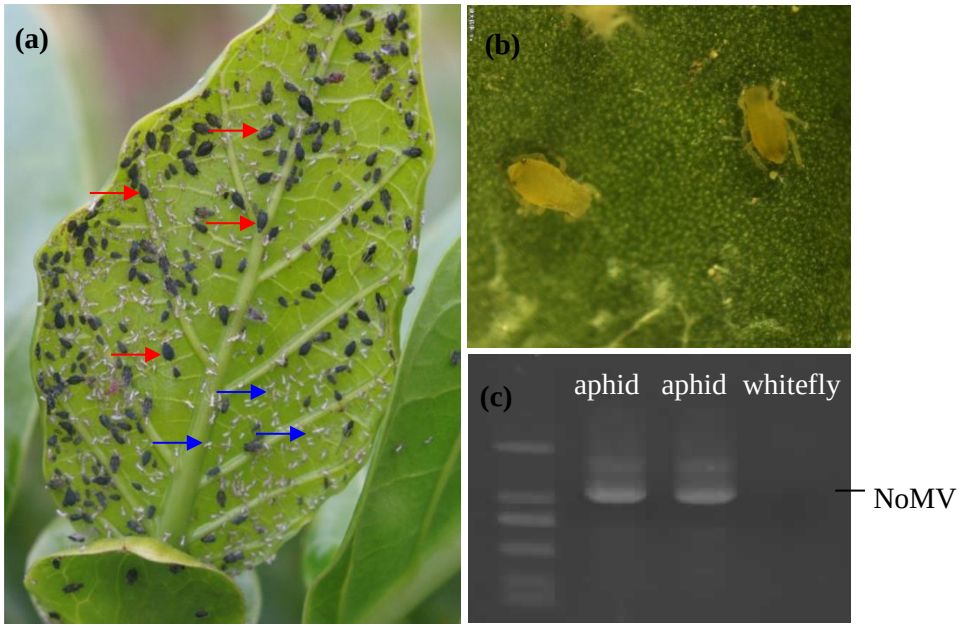

Figure 4. RT-PCR of *Noni mosaic virus* coat protein gene using total RNAs from aphids or whiteflies collected from diseased noni leaves. (**a**), A large number of aphids (red arrows) and whiteflies (blue arrows) at the back of a leaf; (**b**), Aphids larvae; (**c**), RT-PCR detection of NoMV in aphids and whiteflies by NMV-F and NMV-R primers. The left lane contains Trans 2000 bp DNA Marker.

4. Discussion

In this study, we analyzed the transcriptome high-throughput sequencing data of diseased noni plant leaves, which leads to identification of a novel potyvirus in diseased plant and assembly of the complete genome of the potyvirus NoMV. However, the *de novo* assembly from the Illumina reads failed to obtain ~850 bp sequence at the very 5' end of NoMV. This is most likely due to the poor sequence homology at the 5' end of NoMV with any existing viral sequences. Consequently reads containing the 5' end sequences of NoMV were not mapped to any vial genome and were not used for the *de novo* assembly during the subsequent analysis. The draft genomic sequence assembled from the NGS data allowed the design of primers for PCR cloning and 5' and 3' RACE. The complete genome of NoMV was then sequenced by using Sanger sequencing and assembled. When the complete NoMV genome was used as reference to analyze the transcriptomic sequencing data again, we found rare viral reads mapped to the 5' end of NoMV genome. Therefore, the data from the transcriptome high-throughput sequencing would not completely cover this region. Similar results were observed not only in potyvirus [17], but also in babuvirus [18], enamovirus and nucleorhabdovirus [19]. These results also indicated the complete genome of NoMV would likely be obtained by increasing the transcriptomic sequencing data coverage. Considering the cost involved, we chose to fill the flanking region by 5' RACE.

Potyviruses are one of the most important plant-infecting virus groups [20]. At present *Potyvirus*, consisting of about 170 species, is the largest genus in the family *Potyviridae* according to the International Committee on Taxonomy of Virus (ICTV) [21, 22]. The species demarcation criteria of potyviruses, as suggested by ICTV, is based on the complete nucleotide sequence and large ORF amino acid sequence, setting at <76 % nucleotide sequence identity or <82 % amino acid sequence identity [20, 23]. In this study, sequence analysis revealed that the polyprotein and CP aa sequence of NoMV shared less than 51% and 67% sequence identities with those of other potyviruses. These results indicate that NoMV is a novel potyvirus. Although NoMV was more similar with TVBMV and grouped into sub-clades in polygenetic analysis, it is genetically distinct from TVBMV and other potyviruses.

Among the ten proteins produced by cleavage of the potyviral polypeptide, P1 is the most divergent in amino acid sequence. Sequence comparison revealed that the amino acid sequence of NoMV P1 shared less than 14% identities with those of other potyviruses. P1 has been shown to be required for host adaptation in potyviruses [24, 25]. In addition, sequence comparison also revealed that the amino acid sequence of NoMV P3 shared less than 23% identities with those of other potyviruses. P3 has been shown to play a decisive role in the intercellular movement of the potyvirus [26]. These findings suggest that the two genetically divergent proteins, along with other proteins, may allow NoMV to adapt noni plants as a host.

Our study also suggests that aphid is a carrier and a potential transmission vector of NoMV. Some conserved motifs on potyviral proteins have been verified to be responsible for potyvirus aphid transmission [22]. For example, motifs "KITC" and "PIT" on HC-Pro and a motif "DAG" on CP have been identified to be key factors in aphid transmission [16, 27-29]. This finding also arouses the concern that without appropriate management of aphids could accelerate the spread of NoMV. Currently, the disease caused by NoMV has spread to other growing area in Yunnan Province and even some neighboring area of Thailand. It is likely that the disease would further spread without proper management and intervention. Further studies are needed in order to clarify the genetic diversity, biological characterization, and epidemiology of NoMV in different geographic regions.

The data presented in this study provided strong evidence that the novel NoMV is intimately associated with and is the likely cause for the noni mosaic disease prevalent in Yunan, China. Additional experiments are required to conclusively demonstrate the etiology of and the role of NoMV in the noni mosaic disease.

Supplementary Materials: The following are available online, **Table S1**: List of primers used in this study, **Table S2**: The sequence identity of NoMV and other potyviruses based on polyprotein and CP amino acids, **Table S3**: The sequence identity of NoMV and other potyviruses based on P1, HC-Pro, P3, 6K1, CI, 6K2, VPg, Nia-Pro, Nib amino acids.

Author Contributions: Conceptualization, N.Y. and Z.L.; methodology, N.Y.; investigation, Z.C. and Y.Y.; writing—original draft preparation, N.Y.; writing—review and editing, Z.L., Y.Y. and Z.X.; funding acquisition, N.Y.

Funding: This study was funded by the Young Elite Scientists Sponsorship Program by CSTC (Project No. CSTC-QN201704), Guangdong Key Laboratory of Tropical and Subtropical Fruit Tree Research, Institute of

Fruit Tree Research, Guangdong Academy of Agricultural Sciences (No. 2017B030314113) and Central Public-interest Scientific Institution Basal Research Fund for Chinese Academy of Tropical Agricultural Sciences (No.19CXTD-33).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Qian, Y.; Xu, Y.; Zhou, Q.; Zhou, X. Application of next-generation sequencing technology for plant virus identification. *SCIENTIA SINICA Vitae* **2014**, *44*(4), 351–363.
2. Diaz-Lara, A.; Navarro, B.; Di Serio, F.; Stevens, K.; Hwang, M.S.; Kohl, J.; Vu, S.T.; Falk, B.W.; Golino, D.; Al Rwahnih, M. Two novel negative-sense RNA viruses infecting grapevine are members of a newly proposed genus within the family *Phenuiviridae*. *Viruses 2019*, *11*(8), 685.
3. Rott, M.E.; Kesanakurti, P.; Berwarth, C.; Rast, H.; Boyes, I.; Phelan, J.; Jelkmann, W. Discovery of negative-sense RNA viruses in trees infected with apple rubbery wood disease by Next-Generation Sequencing. *Plant Dis*. **2018**, *102*(7), 1254–1263.
4. Hao, X.; Zhang, W.; Zhao, F.; Liu, Y.; Qian, W.; Wang, Y.; Wang, L.; Zeng, J.; Yang, Y.; Wang, X. Discovery of plant viruses from tea plant (*Camellia sinensis* (L.) O. Kuntze) by metagenomic sequencing. *Front Microbiol*. **2018**, *9*, 2175.
5. Yang, K.; Shen, W.T.; Li, Y.; Li, Z.P.; Miao, W.G.; Wang, A.M.; Cui, H.G. *Areca palm necrotic ringspot virus*, classified within a recently proposed genus *Arepavirus* of the family *Potyviridae*, is associated with necrotic ringspot disease in areca palm. *Phytopathology* **2019**, *109*(5), 887–894.
6. Ahmad, A.N.; Daud, Z.A.M.; Ismail, A. Review on potential therapeutic effect of *Morinda Citrifolia*, L. Curr. Opin. *Food Sci.* **2016**, *8*, 62–67.
7. Torres, M.A.O.; Mondêgo-Oliveira, R.; De-Sá, J.C.; Rocha, A.L.; Abreu-Silva, A.L. One plant, many uses: a review of the pharmacological applications of *morinda citrifolia*. *Phytother. Res*. **2017**, *31*(2), 971–979.
8. Gan, B.C.; He, M.J. The planting and use of *Morinda citrifolia*. *Forest By-Product and Speciality in China* **2004**, *2*, 7–9.
9. Wang, Y.F.; Liu, Y.X.; Cai, Z.Y. Biological characteristics of noni anthracnose. *Guangdong Agricultural Sciences* **2015**, *42*(7), 60–63.
10. Trapnell, C.; Pachter, L.; Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **2009**, *25*(9), 1105–1111.
11. Anders, S.; Pyl, P.T.; Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**, *31*(2), 166–169.
12. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M.; MacManes, M.D.; Ott, M.; Orvis, J.; Pochet, N.; Strozzi, F.; Weeks, N.; Westerman, R.; William, T.; Dewey, C.N.; Henschel, R.; LeDuc, R.D.; Friedman, N.; Regev, A. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc*. **2013**, *8*(8), 1494–1512.
13. Hall, T.A. BioEdit, version 7.2.5; Software for Biological Sequence Alignment; Ibis Biosciences: Carlsbad, CA, USA. **2013**.
14. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol*. **2013**, *30*(12), 2725–2729.
15. Gong, D.; Wang, J.H.; Lin, Z.S.; Zhang, S.Y.; Zhang, Y.L. Yu, N.T.; Xiong, Z.; Liu, Z.X. Genomic sequencing and analysis of *Chilli ringspot virus*, a novel potyvirus. *Virus Genes* **2011**, *43*(3), 439–444.
16. López-Moya, J.J.; Wang, R.Y.; Pirone, T.P. Context of the coat protein DAG motif affects potyvirus transmissibility by aphids. *J. Gen. Virol*. **1999**, *80*(Pt 12), 3281–3288.
17. Sheveleva, A.; Kudryavtseva, A.; Speranskaya, A.; Belenikin, M.; Melnikova, N.; Chirkov, S. Complete genome sequence of a novel *Plum pox virus* strain W isolate determined by 454 pyrosequencing. *Virus Genes* **2013**, *47*(2), 385–388.
18. Yu, N.T.; Xie, H.M.; Zhang, Y.L.; Wang, J.H.; Xiong, Z.; Liu, Z.X. Independent modulation of individual genomic component transcription and a cis-acting element related to high transcriptional activity in a multipartite DNA virus. *BMC Genomics* **2019**, *20*(1), 573.
19. Cao, M.; Zhang, S.; Li, M.; Liu, Y.; Dong, P.; Li, S.; Kuang, M.; Li, R.; Zhou, Y. Discovery of four novel viruses associated with flower yellowing disease of green Sichuan pepper (*Zanthoxylum Armatum*) by virome analysis. *Viruses* **2019**, *11*(8), 696.

20. Wylie, S.J.; Adams, M.; Chalam, C.; Kreuze, J.; López-Moya, J.J.; Ohshima, K.; Praveen, S.; Rabenstein, F.; Stenger, D.; Wang, A.; Zerbini, F.M. ICTV Report Consortium, ICTV Virus Taxonomy Profile: *Potyviridae*. *J. Gen. Virol*. **2017**, *98*, 352–354.
21. Revers, F.; Garcã-A, J. A. Molecular biology of potyviruses. *Adv. Virus Res*. **2015**, *92*, 101-199.
22. Huang, C.H.; Tai, C.H.; Lin, R.S.; Chang, C.J.; Jan, F.J. Biological, pathological, and molecular characteristics of a new potyvirus, *Dendrobium chlorotic mosaic virus*, infecting *Dendrobium* Orchid. *Plant Dis*. **2019**, *103*(7), 1605–1612.
23. Adams, M.J.; Antoniw, J.F.; Fauquet, C.M. Molecular criteria for genus and species discrimination within the family *Potyviridae*. *Arch. Virol.* **2005**, *150*, 459–479.
24. Shan, H.; Pasin, F.; Tzanetakis, I.E.; Simón-Mateo, C.; García, J.A.; Rodamilans, B. Truncation of a P1 leader proteinase facilitates potyvirus replication in a non-permissive host. *Mol. Plant Pathol*. **2018**, *19*(6), 1504–1510.
25. Vozárová, Z.; Glasa, M.; Šubr, Z.W. Tracking the potyviral P1 protein in *Nicotiana benthamiana* plants during *plum pox virus* infection. *Acta Virol*. **2017**, *61*(4), 492–494.
26. Cui, X.; Yaghmaiean, H.; Wu, G.; Wu, X.; Chen, X.; Thorn, G.; Wang, A. The C-terminal region of the *Turnip mosaic virus* P3 protein is essential for viral infection via targeting P3 to the viral replication complex. *Virology* **2017**, *510*, 147–155.
27. Plisson, C.; Drucker, M.; Blanc, S.; German-Retana, S.; Le, Gall O.; Thomas, D.; Bron, P. Structural characterization of HC-Pro, a plant virus multifunctional protein. *J. Biol. Chem*. **2003**, *278*(26), 23753–23761.
28. Stenger, D.C.; Hein, G.L.; Gildow, F.E.; Horken, K.M.; French, R. Plant virus HC-Pro is a determinant of eriophyid mite transmission. *J. Virol*. **2005**, *79*(14), 9054–9061.
29. Seo, J.K.; Kang, S.H.; Seo, B.Y.; Jung, J.K.; Kim, K.H. Mutational analysis of interaction between coat protein and helper component-proteinase of *Soybean mosaic virus* involved in aphid transmission. *Mol Plant Pathol*. **2010**, *11*(2), 265–276.