Asian Journal of Plant Pathology

ISSN 1819-1541

Academic Journals Inc.

www.academicjournals.com

Asian Journal of Plant Pathology 7 (1): 1-14, 2013
ISSN 1819-1541 / DOI: 10.3923/ajppaj.2013.1.14

Complete Nucleotide Sequence of DNA A-like Genome and DNA-β of Monopartite Pepper Yellow Leaf Curl Virus, A Dominant *Begomovirus* Infecting *Capsicum annuum* in West Sumatera Indonesia

J. Jamsari and J. Pedri

Plant Breeding Section, Department of Crop Production, Faculty of Agriculture Andalas University, Campus Limau Manis-25163 Padang, West Sumatera, Indonesia

Corresponding Author: J. Jamsari, Plant Breeding Section, Department of Crop Production, Faculty of Agriculture Andalas University, Campus Limau Manis-25163 Padang, West Sumatera, Indonesia

ABSTRACT

Pepper yellow leaf curl disease which is caused by Pepper Yellow Leaf Curl Virus (PepYLCV) is one of the major problems in chili pepper (*Capsicum annuum*) cultivation in Indonesia. Reducing of the yield could reach 100% in some condition. For this reason, well understanding of virus distribution as well as their genome structure is very crucial for combating the disease. Based on this rationality we characterized genome structure of PepYLCV isolate designated as PepYLCWSV-TD21 which was collected from chili cultivation pepper population in West Sumatera. The result indicated that the PepYLCWSV-TD21 had a monopartite genome and dominantly infected compared to its bipartite counterpart. This was confirmed by analysis of DNA-β presence by specific primer pair Beta01/Beta02. Annotation of both genome structure successfully identified 6 open reading frames designated as V1, V2, C1, C2, C3 and C4 in the DNA-A like genome, whereas only 1 open reading frame designated as C1 was identified in DNA-β. Further characteristics of each open reading frame were further elucidated. These results provided us information on distribution of monopartite PepYLCV in West Sumatera Indonesia, as well as its genome characteristic that in turn could be used as our start point for development of resistant chili pepper cultivar.

Key words: Chili pepper, DNA A like genome, DNAβ genome, monopartite genome, west Sumatera

INTRODUCTION

Pepper yellow leaf curl disease (PepYLCD) is currently an epidemic disease in chili pepper cultivation in West Sumatera Indonesia. The emergence of the disease in Indonesia was first reported in 1999 in West Java by Rusli *et al.* (1999). In 2003, the disease was reported being occurred in central Java (Sulandari, 2004; Sulandari *et al.*, 2006) and became epidemic in Southern and Western Sumatera at 2005. The newest study of reported disease incidence of PepYLCD has reached up to 97% (Trisno *et al.*, 2009).

PepYLCD was caused by geminivirus known as Pepper yellow leaf curl virus (PepYLCV). The virus belongs to *Begomovirus* which mostly have bipartite genomes and only a small number have a monopartite genome. The bipartite *Begomoviruses* have two genomes, termed as DNA-A and

DNA-B. Both DNA-A and DNA-B are essential for virus proliferation (Xie *et al.*, 2010). Detail analysis of both components, showed that DNA-A contains genes required for DNA replication, regulation for gene expression and coat protein formation, whereas DNA-B encompasses proteins involved in intra- and intercellular movement (Sanderfoot and Lazarowitz, 1996; Briddon *et al.*, 2010). Both DNA-A and DNA-B share 85% identity of nucleotide sequence in Common Region (CR) spanning along ~200 nucleotide. This region encompasses hairpin loop structure, the nonanucletide sequence (TAATATTAC) which is also known as origin of virion-strand DNA replication and iterons. Iterons are recognition sequences for binding of DNA-A-encoded replication-associated protein (Rep) (Hanley-Bowdoin *et al.*, 1999; Arguello-Astorga and Ruiz-Medrano, 2001; Briddon *et al.*, 2010).

The monopartite *Begomoviruses* have genome that resembles the DNA-A of bipartite *Begomoviruses* (Khey-Pour *et al.*, 2000; Navot *et al.*, 1991; Dry *et al.*, 1993; Noris *et al.*, 1994; Fauquet *et al.*, 2008) which is known as DNA-A like genome and DNA satellite (Briddon *et al.*, 2008). The DNA satellite is termed also as **β**-component or DNA-**β** (Saunders *et al.*, 2002, 2003; Briddon *et al.*, 2001; Zhou *et al.*, 2003). In some *Begomoviruses*, additional satellite DNA which is termed as alphasatellite can be found (Briddon *et al.*, 2004; Mansoor *et al.*, 1999). Both DNA-**β** and alphasatellite have approximately half of size of the helper *Begomovirus* genome. They play important role for encapsidation (transmission by insect) and systemic infection of the plant host (Briddon and Stanley, 2006; Idris *et al.*, 2011). Intensive analysis of the DNA-**β** molecules concluded that they contain only one major ORF (**β**C1) that is responsible for symptom induction (Briddon and Stanley, 2006). Study in cotton leaf curl disease (CLCuD) successfully proved that **β**C1 gene is responsible for pathogenicity protein (Saeed *et al.*, 2005).

Indonesia as one of the biggest chili pepper consumers belongs to the old word. The group of the old world is believed to be having two groups of *Begomovirus* namely monopartite and bipartite based on the genome constituent. The first complete genome of PepYLCV isolated from tomato and ageratum in Indonesia was reported by Sakata *et al.* (2008). However, the virus was described as bipartite one. The monopartite PepYLCV in Indonesia currently reported by Tsai *et al.* (2006) was isolated from tomato and chili pepper in Bogor-West Java. However, no complete DNA-**β** was reported so far accompanying the DNA-A like genome in monopartite PepYLCV. Since, DNA-**β** is believed as one of the pathogenicity determinant associated with various plants disease exclusively in monopartite *Begomovirus* (Briddon and Stanley, 2006; Nawaz-ul-Rehman and Fauquet, 2005), thus it is important to understand their characteristic.

Here we report a complete nucleotide sequence of DNA-A like genome and DNA-**β** from a single monopartite PepYLCV isolated from West Sumatera Indonesia.

MATERIALS AND METHODS

Sample collection: Chili pepper plants showing typical symptom of PepYLCV infection were collected from three different geographical regions claimed as centers for chili pepper producing area. Low altitude was classified by ≤400 m above sea level (asl). This region was covered by two districts namely: Pesisir Selatan and Pasaman Barat. Medium altitude was classified by regions located between 400-700 m asl, covered by two districts: Payakumbuh and Solok and high altitude represented by regions located on ≥700 asl. This region was represented by Tanah Datar and Agam district. Twenty leaf samples from each high and low altitude and twenty two leaf samples from medium altitude were collected and prepared for the analysis. Plants selection was based on the

infection symptom of PepYLCV, characterized by smaller leaflet, stunting, leaf curling, cupping and yellowing mosaic leaf. All samples collected from the field were maintained in 4°C before used for DNA isolation.

DNA isolation and PCR analysis: DNA isolation principally was carried out according to protocol described by Dellaporta *et al.* (1983) with minor modification. Briefly DNA isolation was carried out as follow. About 100,0 mG of leaf sample was powdered in sterile mortar. Leaf powder was then transferred to 2, 0 mL eppendorf containing 1, 0 mL of extraction buffer (0, 1 M Tris-pH 9.0, 0.1 M EDTA and 1% (w/v) SDS). Suspension was homogenized by vortexing and incubated at 65°C for 30 min. Thirty five microliter of 8 M Potassium acetate was added to the suspension directly after incubation. Precipitation of cell debris was done by mean centrifugation for 15 min at 10,000 rpm. Supernatant was then transferred to the new sterile eppendorf tube prior addition of 500 μL of Phenol: Chloroform (15:1) and then mixed by inverting for 5 min. Supernatant was harvested by centrifugation for 15 min at 10,000 rpm. Precipitation of DNA was performed by adding 2.5x of 96% (v/v) ethanol and subsequently proceed by inverted mixing and incubation for 15 min on ice. Pellet was harvested by additional centrifugation for 15 min at 10,000 rpm. Washing of the DNA was done by adding 500 μL of 70% (v/v) ethanol. Pellet was dried at 55°C for 5 min on a heater block. The DNA was then resuspended with 50 μL of 1x TE and finally stored at -20°C before used. Prior PCR analysis, DNA was set to 5 ng μL^{-1}.

Generally PCR technique applied in this study was carried out in 25 μL of total volume reaction containing 2,5 μL 10x PCR buffer (100 mM Tris-HCl, 500 mM KCl pH 8.3), 2.5 μL 2.5 mM dNTPs, 1 μL (5 pmol μL^{-1}) of each primer, 5 μL DNA (5 ng μL^{-1}), 1 U *Taq* DNA Polymerase (Amersham-USA) and ddH_2O to 25 μL. PCR condition was set depend on primer set used in every analysis. Primers used in this study mainly were published somewhere (Table 1). General PCR reaction was performed using common condition with initial denaturation at 94°C for 5 min; followed by 30 cycles consisting of 94°C for 1 min, 55°C for 2 min and 72°C for 2 min. Final extension was done for 10 min at 72°C. In some cases PCR condition was set as described by the authors in the references.

Sequencing of DNA-A like genome and DNA-β: Sequencing of complete PepYLCV genome, generally was performed by means of primer walking strategy which is principally described by Chinault and Carbon (1979). For the DNA-A like genome, sequencing was started with initial direct sequencing of PCR product generated by primer pair pAR1c715/pALv1978 (Rojas *et al.*,1993; Idris and Brown, 1988). This step could successfully produce 1464 bases. Further sequencing for extending genome coverage was performed by designing additional primer pairs locating on the sequence data from the previous step (TD21-456F/TD21-455R for the second walking and TD21-1010F/WS387R for the third walking). Nucleotide sequence validation was confirmed by comparing two nucleotide sequences overlapping each other. Complete nucleotide sequence of DNA-β was elucidated by cloning of PCR product generated from primer pair Beta01/Beta02 (Paximadis and Rey, 1997). Cloning was done using TOPO-Cloning system (Invitrogen-USA) as recommended by producer. Complete nucleotide sequence was generated from both terminal using M13F and M13R primer. Editing of sequence data was carried out using Bioedit software package tool (Hall, 1999). Homology analysis was performed using BLASTn at NCBI homepage http://blast.ncbi.nlm.nih.gov while alignment was done by Clustal W at http://ebi.ac.uk.

Table 1: Primers used in the study

Primer ID	Sequence 5'----3'	Length (base)	Genome	References
AC1048	GGRTTDGARGCATGWCGTACATG	23	Genome A	Wyatt and Brown (1996)
AV494	GCYATRTYAGRAAGCCMAG	19	Genome A	Wyatt and Brown (1996)
Beta01	GGTACCACTACGCTACGCAGCAGCC	25	β-component	Paximadis and Rey (1997)
Beta02	GGTACCTACCCTCCCAGGGGTACAC	25	β-component	Paximadis and Rey (1997)
CPPROTEIN-C1	GGCCGAATTCTTAATTTTGAACAGAATCA	29	Genome A	Trisno (personal communication)
CPPROTEIN-VI	TAATTCTAGATGTCGAAGCGACCCGCCGA	29	Genome A	Trisno (personal communication)
pAL1c1960	ACNGGNAARACNATGTGGGC	21	Genome A	Paximadis and Rey (1997)
pAL1v-Compl	ACNGGRAAGACRATGTGGGCCTGCAGATGC	30	Genome A	This research
pAL1v1978	GCATCTGCAGGCCCACATYGTCTTYCCNGT	30	Genome A	Rojas *et al*. (1993)
pAR1c715	GATTTCTGCAGTTDATRTTYTCRTCCATCCA	31	Genome A	Idris and Brown (1988)
pAR1c-Compl	TGGATGGAYGARAAYATHAACTGCAGAAATC	31	Genome A	This research
pBL1v2040	GCCTCTGCAGCARTGRTCKATCTTCATACA	30	Genome B	Rojas *et al*. (1993)
pBR1c800	ACGACTGCAGTTVACMGTCCCTTTGAAACG	30	Genome B	Rojas *et al*. (1993)
PCRc1	CTAGCTGCAGCATATTTACRARWATGCCA	29	Genome B	Paximadis and Rey (1997)
PCRc4	GGCCATAGAGCTTTGAGGATCCCGATTCATTTC	33	Genome B	Paximadis and Rey (1997)
pIRc2671	GGGTACCGATATACCAGGAG	20	Genome B	Paximadis *et al*. (1999)
pIRv2672	AATATATAGTGGGTACCGAATGG	23	Genome B	Paximadis *et al*. (1999)
prAC1154	CTSAAYTTCMAAGTYTGGACG	21	Genome A	Paximadis and Rey (1997)
prAC344	CTKGGCTTYCTRTACATRGGC	21	Genome A	Paximadis and Rey (1997)
prAV1134	CGTCCARACTTKGAARTTSAG	21	Genome A	Idris and Brown (1988)
prAV2644	ATTACCGGATGGCCGC	16	Genome A	Paximadis and Rey (1997)
prBC656	TTVACMGTCCCTTTGAAACG	20	Genome B	Paximadis and Rey (1997)
prBV1855	ACRCAARTGRTCKATYTTCAT	21	Genome B	Paximadis and Rey (1997)
TD21-PAL2	CGGTGGTATCACCGTCTTTGTCC	23	Genome A	This research
TD21-PAL2-compl	GGACAAAGACGGTGATACCACCG	23	Genome A	This research
TD21-PAR	TTCGTGAGGGATTTAATTTGTCTG	24	Genome A	This research
TD21-PAR Compl	CAGACAAATTAAATCCCTCACGAA	24	Genome A	This research
TD21-PYLCV-455F	CGTGCAGACGTATTTCCCTTCGAAT	25	Genome A	This research
TD21-PYLCV-455R	CAACAGATTCTTCGACCTGGTAT	23	Genome A	This research
TD24 L	CTTAAAGTCTATCTACCGGCTAG	23	Genome A	This research
TD24 R	CGAGGGACTTACAAGCCTACC	21	Genome A	This research
TD24L-1	GAATTTCAGATAGATGGCCGATC	23	Genome A	This research
ToLCV-A1060F	AAGCTTTATATCATATGAAGTCTGTAC	27	Genome A	Maruthi *et al*. (2005)
ToLCV-A1060R	AAGCTTGATATTTATTAATTTGTTACCG	28	Genome A	Maruthi *et al*. (2005)
ToLCV-A300F	ATGKCSAAGCGWCCRGCAGA	20	Genome A	Maruthi *et al*. (2005)
ToLCV-A300R	AATTTCTGGCCCTGATAAC	19	Genome A	Maruthi *et al*. (2005)
ToLCV-B1100R	GTCGACTTGTGAGGGTACAT	20	Genome B	Maruthi *et al*. (2005)
ToLCV-B1500F	GATGAAGCGATTCTTTCGC	19	Genome B	Maruthi *et al*. (2005)
ToLCV-B1650F	GGATCCAAACCACAAACAAGGC	22	Genome B	Maruthi *et al*. (2005)
ToLCV-B1650R	GGATCCCACAGACCAGAAATC	21	Genome B	Maruthi *et al*. (2005)
WS-387R	GGAGCTAARTCMAGYTCCGAYGTCA	25	Genome A	This research
TD21-456/F	GAATCGGAAACCCAGATTCTA	21	Genome A	This research
TD21-455/R	CAACAGATTCTTCGACCTGGTAT	23	Genome A	This research
TD21-1010F	GCATCCAATCCTGTATACGC	20	Genome A	This research

RESULTS AND DISCUSSION

Distribution of DNA-β in PepYLCV population in chili cultivation in west Sumatera: To get the overview of monopartite PepYLCV distribution in all West Sumatera chili pepper cultivation

Table 2: Distribution of DNA-β presence in the sampling population

Altitude	District	Positive	Negative	Percent. (+)	Percent. (-)	Average ratio (%)	
High	Agam	3	7	30	70	55	45
	Tanah datar	8	2	80	20		
Medium	Solok	1	9	10	90	55	45
	Payakumbuh	11	1	92	8		
Low	Pasaman barat	6	4	60	40	70	30
	Pesisir selatan	8	2	80	20		
Total		37	25			60	40

areas, we checked our isolate collection using specific primer Beta01/Beta02 via PCR technique. The results are listed in Table 2. Thirty samples (60%) out of 62 samples could produce a single PCR product which was approximately 1,500 bp in length. The other 25 samples (40%) did not produce any band. That single band was expected as DNA-**β** fragment as described by Paximadis and Rey (1997).

DNA-**β** is believed to be specifically only contained in monopartite *Begomovirus* only. Their distribution obviously represent distribution of monopartite PepYLCV. Based on this, we deduced that distribution of DNA**β** thus reflect also distribution of monopartite PepYLCV.

Using data from Table 2, we concluded that in general monopartite PepYLCV was dominant in chili pepper cultivation in West Sumatera. This happened in all regions of altitude, where chili pepper is commonly cultivated. Comparing their distribution among the three regions indicated that low altitude, is the most infected by monopartite PepYLCV (70%) compared to medium and high altitude regions (each 55%). Distribution between monopartite and bipartite PepYLCV in every region however seemed not consistent except in low altitude. In high altitude, monopartite PepYLCV was dominant, for instance in Tanah Datar. But on the other hand, in Agam monopartite PepYLCV was dominated by its bipartite counterpart. Similar pattern could be seen in medium altitude, where monopartite PepYLCV was dominant in Payakumbuh but was dominated by its bipartite counterpart in Solok. In the low altitude, domination of monopartite PepYLCV was consistent in both districts. Probably low altitude region is the most suitable for monopartite *Begomovirus* propagation. However, this assumption must be proved by more valid data.

Report on the existence of monopartite *Begomovirus* in Indonesia was first described in tomato and chili-pepper plants by Tsai *et al.* (2006). They have even successfully elucidated the complete sequence of DNA-A like genome which is deposited in NCBI genebank with accession number DQ083764.1 and DQ083765.1. However, their information regarding monopartite *Begomovirus* was not completed by the information of DNA **β** which is believed as specific feature in almost monopartite geminivirus. For this reason, we further investigated genome constituent of our monopartite PepYLCV isolate.

Genome type of PepYLCWSV isolate TD21: In order to investigate genome type of our PepYLCV collections, a series analysis using some sets of primer pair specific for genome A, genome-B and DNA**β** was performed on isolate TD21 and PSS14 by PCR analysis. Twenty one primer pairs in total (Table 1) were tested for their applicability. Twelve primer pairs which were specific for genome-A showed positive single PCR product from both samples but no single PCR product as expected from all 8 primer pairs specific for genome-B was observed. However, primer pair of Beta01/Beta02 clearly produced a positive single PCR product from both isolates which was

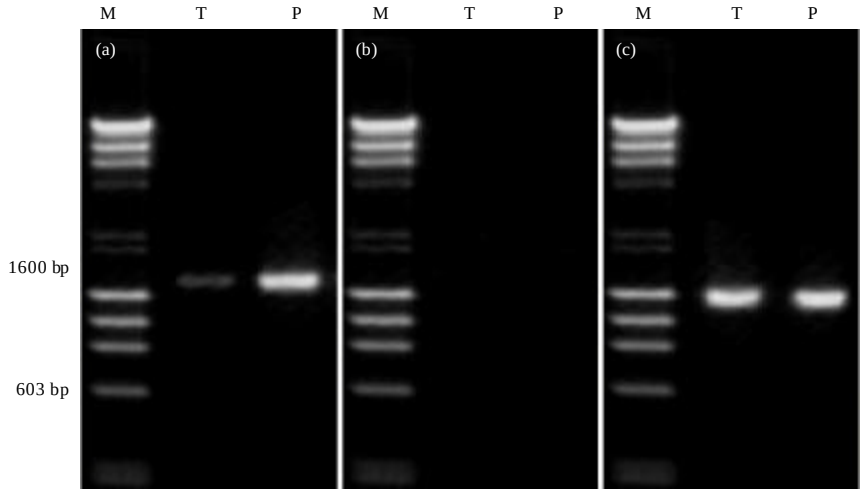

Fig. 1(a-c): (a) Genome type analysis of PepYLCV showed a monopartite feature indicated by the absent of genom-B and existence of genome-A like fragment and DNAβ fragment, (b) A specific primer for genome-B (prBV1855/prBC656) produced no PCR product from two samples collected from highgland (T) and lowland (P) and (c) Single PCR product about 1300 bp in length could be seen after amplification of similar samples with Beta01/Beta02 primer pair, M is 1 kb size marker (Fermentas, USA)

approximately 1300 bp in size. All positive primer pairs were further applied in the subsequent steps using 60 samples as described above. A representative electrophoresis analysis from this step is shown in Fig. 1.

So far majority of *Begomovirus* reported by many authors were dominated by bipartite group Sakata *et al.* (2008). Only little of them were reported to have monopartite genome (Tsai *et al.*, 2006). Furthermore most of the reported monopartite *Begomovirus* was focused on the DNA-A like genome and very rare simultaneously reported together with the presence of DNA-β. Thus, report of monopartite geminivirus containing also DNA-β in one PepYLCV isolate from West Sumatera Indonesia should be very interesting.

Complete sequence of DNA-A like and DNA-β genome: Complete nucleotide sequencing of selected monopartite PepYLCV genome was carried out on isolate TD-21, further the isolate was designated as PepYLCWSV-TD21. This isolate was collected from highland region (Tanah Datar), representing major group of PepYLCV isolates infecting chili pepper in West Sumatera. Sequencing was performed using primer walking strategy, started with universal primer pair PAR1c715/PAL1v1978 (Rojas *et al.*,1993; Idris and Brown, 1988). After controlling and editing of the raw sequence data, a nucleotide sequence of 1464 bp in length was generated from both primer positions exhibiting an overlapping position in 258 bp in length. The region covered Common Region (CR) sequence, stem loop structure harboring conserved nona-nucleotide.

Walking of the primer was continued by designing some internal primers from available sequences (Table 1). After performing three steps of primer walking, overlapping sequence in the upstream and downstream position were identified. Trimming and editing produced a circular nucleotide sequence of 1749 bases in length. The accuracy of sequences was verified by using at

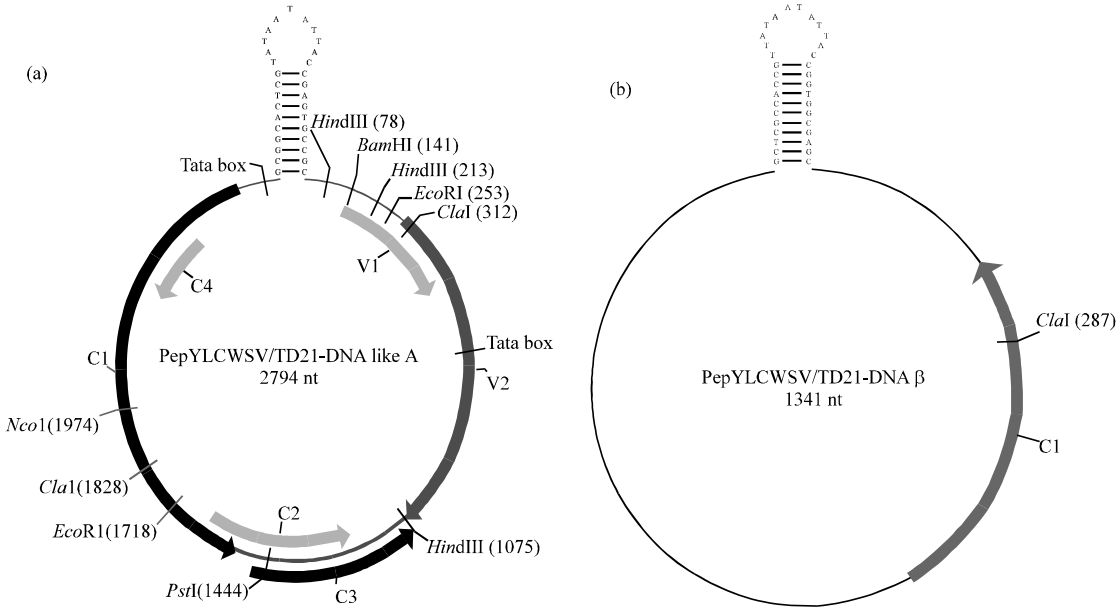

Fig. 2(a-b): (a) Map of DNA-A like genome and (b) DNAβ from a monopartite PepYLCV-TD21

least two contigs from similar region. The last step primer walking brought to the conclusion of final fragment length which spanned of 2749 bp in length (Fig. 2a). The identified DNA-A like genome in fact is 11 bp shorter compared to other DNA-A like genome of monopartite PepYLCV isolated also from Indonesia, for instance a PepYLCV isolated from *Ageratum* sp. (AB267838.1) (Sakata *et al.*, 2008) which is 2760 bp in length. However, compared with previous PepYLCV DNA-A like genome isolated from tomato and chili pepper in Bogor-West Java Indonesia (DQ083765.1 and DQ083765.1), our sequence has 5 bp longer. BLAST searching of both DNA-A like genome exhibited 95% homology with our sequence.

Report on the presence of DNA-A like genome simultaneously together with DNAβ of monopartite geminivirus PepYLCV particularly isolated from *Capsicum annum* cultivation in Indonesia to our knowledge is not exist so far. Therefore this finding is claimed to be the first report informing the sequence of DNA-A like genome and simultaneously DNAβ in one virus particle of monopartite PepYLCV from *Capsicum annuum* in Indonesia particularly in West Sumatera. Despite the existency of monopartite *Begomovirus* infecting some plants for instance chili pepper and tomato (Tsai *et al.*, 2006), tobacco (Li *et al.*, 2005) and some other solanaceae families like tomato (Saeed *et al.*, 2007) and sugar beet (Stanley *et al.*, 1992) were extensively reported by many authors. Almost all reports regarding monopartite features of geminiviruses described only a DNA-A like genome structure.

Complete sequence of DNAβ genome was validated after cloning of complete PCR-product generated from primer pair of Beta01/Beta02 (Paximadis and Rey, 1997) into TOPO plasmid (Invitrogen-USA). After editing and screening of vector sequence and additional nucleotide from Beta01/Beta02 primer a fragment spanned in 1341 bases in length was verified (Fig. 2b). Compared with other DNA-β established in NCBI database exhibited no significant homology. The only one highest homology of PepYLCWSV-TD21 DNA-β sequence was observed with Ludwigia yellow vein virus satellite DNA beta (AJ965541.1). The nucleotide sequence was deposited by Huang *et al.* (2006). The DNA-β sequence of PepYLCWSV-TD21 found in this study is 6 bps shorter than above reported by Huang *et al.* (2006) which is 1347 bp and shared 95% in homology.

Genome organization of DNA-A like genome: Six open reading frames (ORFs) were identified after performing annotation on the DNA-A like genome sequences. Two of them V1, V2 are in the virion sense and four ORFs C1, C2, C3 and C4 are in the complementary sense. Detailed ORFs positions and their putative amino acid residues as well as their predicted molecular weights are listed in Table 3.

Further analysis of six ORFs from DNA-A like genome of PepYLCWSV-TD21 exhibited significant homology (95-97%) for the whole genome and for every ORF or gene compared with other PepYLCV nucleotide sequences available in NCBI public database. BLAST searching of V1-ORF nucleotide exhibited significant homology ranging from 95-99%. The highest homology (99%) of V1-ORF is showed with V1-gene nucleotide sequence from 3 PepYLCV collected from ageratum (AB267838.1) and *Lycopersicon esculentum* (AB189849.1; AB189845.1). However, all the three isolates belong to the bipartite *Begomovirus*. The two V1-gene of monopartite PepYLCV deposited in gene bank (DQ083764.1 and DQ083765.1) showed 97% homology with our V1-gene. BLAST searching of V2-ORF nucleotide sequence showed almost similar result with the V2-gene nucleotide sequences available in NCBI-database. Seemed there is no correlation between V2-sequence with respect of genome type of PepYLCV. The C1-ORF of PepYLCWSV-TD21 which is believed relating to the replication and determine the aggressivity of the strain showed homology with other 9 pepYLCV C1-gene within the range of 94-96%.

A hairpin loop structure containing of 11 conserved nona-nucleotide (TATAATATTAC) could also be identified in DNA-A like genome of PepYLWSV-TD21 (Fig. 3a). The nucleotide sequence

Table 3: Start and stop codon position of ORFs along PepYLCWSV-TD21 DNA-A like genome

Gene/ORF	ORF (start-stop)	No. of amino acid residues	Predicted MW (kDa)
V1 (sense)	135-486	117	29.14
V2 (sense)	295-1067	258	63.79
C1 (non sense)	1523-2609	362	90.07
C2 (non sense)	1216-1621	135	33.75
C3 (non sense)	1071-1479	136	33.75
C4 (non sense)	2257-2452	65	16.10

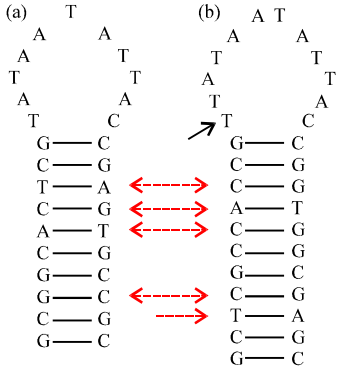

Fig. 3(a-b): (a) Hairpinloop structure of isolate DNA-A like genome and (b) DNA-β from isolat PepYLCWSV-TD21. Additional nucleotide (T) is shown by full arrowhead while variation of dimer structure in the stem of hairpinloop are shown by dashed double arrowhead. Additional nucleotide dimer is also observed in the stem of hairpinloop structure (indicated by dashed arrowhead)

of the hairpin loop structure composed of 31 bases GCGGCACTCGTATAATATTACCGAGTGCCGC. Comparison of this structures with 9 other PepYLCV DNA-A like genome plus additional with TYLCV isolated from *Solanum melongena* (DQ641702.1) showed 100% similarity or identical (Table 4). Seemed, all 31 nucleotides building PepYLCV hairpin loop structure are highly conserved, no matter from any host and genome type they were isolated and no correlation with the type of the genome. Interestingly four additional bases after hairpin loop body structure contain GAAA structure are also identical among all 10 compared PepYLCVs (data not shown). Whether this characteristic is specific for all PepYLCV in general or not? This must be proved furthermore.

Genome organization of DNAβ: Editing of putative DNAβ nucleotide sequences from PepYLCWSV-TD21 brought us to final conclusion of 1341 bases in length. Homology search with other DNAβ sequences isolated also from chili pepper (*Capsicum annuum*) showed that our sequence is shorter than others DNAβ sequences deposited at public nucleotide database which range from 1380-1581 bp (Hussain *et al*., 2009). They shared homology from 81-85%. However, comparing our DNA-β with two other DNA-β isolated from tomato collected also from Indonesia showed only 86-87% homology. Interestingly our DNA-β sequence showed significant homology (96%) with DNA-β sequence isolated from Ludwigia yellow vein virus (AJ965541.1) which normally infects *Ludwigia hyssopifolia* (Huang *et al*., 2006).

The loop structure of DNA-β was composed of 12 nucletotides (Table 5). This was 1 base longer compared with loop structure of DNA-A like genome from similar isolate PepYLCWSV-TD21. The loop contained also highly conserved nona-nucleotide composed of 11 bases (TAATATT/AC) which is 100% identical with nona-nucleotide sequence from DNA-A. This nona-nucleotide could also be found from AJ965541.1; AB113651.1 and AB162142.1. The two later DNAβ sequences were isolated from *Lycopersicon esculentum* and *Ageratum* sp., respectively (Kon *et al*., 2006, 2007). Both were collected in Indonesia (Java) but AJ965541.1 was originated from China (Huang *et al*., 2006). Such conservation in geminivirus is common and well known (Nawaz-ul-Rehman and Fauquet, 2005).

Table 4: Comparison of stem-loop structure sequence of PepYLCV DNA-A like genome with 10 other *Begomoviruses* collected from Indonesia

GB No.	Stem loop structure sequence			Host	References/genome type
-	GCGGCACTCG	TATAATATTAC	CGAGTGCCGC	*Capsicum* sp.	This research/*
AB267834.1	GCGGCACTCG	TATAATATTAC	CGAGTGCCGC	*Ageratum* sp.	Sakata *et al*. (2008))/**
DQ083764.1	GCGGCACTCG	TATAATATTAC	CGAGTGCCGC	*Capsicum* sp.	Tsai *et al*. (2006))/*)/*
AB189850.1	GCGGCACTCG	TATAATATTAC	CGAGTGCCGC	*Capsicum* sp.	Unpublished/**
AB246170.1	GCGGCACTCG	TATAATATTAC	CGAGTGCCGC	*Capsicum* sp.	Unpublished/**
DQ083765.1	GCGGCACTCG	TATAATATTAC	CGAGTGCCGC	*Lycopersicon* sp.	Tsai *et al*. (2006))/*
AB267836.1	GCGGCACTCG	TATAATATTAC	CGAGTGCCGC	*Lycopersicon* sp.	Sakata *et al*. (2008)/**
AB267831.1	GCGGCACTCG	TATAATATTAC	CGAGTGCCGC	*Lycopersicon* sp.	Sakata *et al*. (2008)/**
AB189845.1	GCGGCACTCG	TATAATATTAC	CGAGTGCCGC	*Lycopersicon* sp.	Sukamto, *et al*. (2005)/**
AB267838.1	GCGGCACTCG	TATAATATTAC	CGAGTGCCGC	*Ageratum* sp.	Sakata *et al*. (2008)/**
DQ641702.1	GCGGCCCTCG	TATAATATTAC	CGAGGGCCGC	*Solanum* sp.	Ha *et al*. (2008)/**

All PepYLCV collected from Indonesia showed identical sequence in the stem-loop structure irrespective with their host, only TYLCV collected from *Solanum* sp., showed different nucleotides. Position of nucleotide variation is shown with arrowhead, conserved nona-nucleotide is shown in the box

Table 5: Comparison of stem-loop structure sequence of DNAβ from PepYLCWSV-TD21 with 14 other *Begomoviruses* collected from Indonesia, China and Pakistan

GB No.	Stem-loop structure	Host	References
-	GCTCGCCACCG TTATAATATTAC CGGTGGCGAGC	*C. annuum*	This research
AJ965541	GCTCGCCACCG TTATAATATTAC CGGTGGCGAGC	*L. hyssopifolia*	Huang *et al.* (2006)
AM849549	GCTCGCCCACG TTCTAATATTAC CGTGGGCGAGC	*C. annuum*	Unpublished
FN179279	GCTCGCCCACG TTCTAATATTAC CGTGGGCGAGC	*C. annuum*	Akhter *et al.* (2009)
AM279661	GCTCGCCCACG TTCTAATATTAC CGTGGGCGAGC	*C. annuum*	Hussain *et al.* 2009)
AM279664	GCTCGCCCACG TTCTAATATTAC CGTGGGCGAGC	*C. annuum*	Hussain *et al.* (2009)
AM279671	GCTCGCCCACG TTCTAATATTAC CGTGGGCGAGC	*C. annuum*	Hussain *et al.* (2009)
AM279672	GCTCGCCCACG TTCTAATATTAC CGTGGGCGGGC	*C. annuum*	Hussain *et al.* (2009)
AM258978	GCTCGCCCACG TTCTAATATTAC CGTGGGCGAGC	*C. annuum*	Unpublished
AM260466	GCTCGCCCACG TTCTAATATTAC CGTGGGCGAGC	*C. annuum*	Unpublished
AM279668	GCTCGCCCACG TTCTAATATTAC CGTGGGCGAGC	*C. annuum*	Hussain *et al.* (2009)
AM279663	GCTCGCCCACG TTCTAATATTAC CGTGGGCGAGC	*C. annuum*	Hussain *et al.* (2009)
AM279662	GCTCGCCCACG TTCTAATATTAC CGTGGGCGAGC	*C. annuum*	Hussain *et al.* (2009)
AB113651	GCTCGCCACCG TTATAATATTAC CGGTGGCGAGC	*L. esculentum*	Kon *et al.* (2006)
AB162142	GCTCGCCACCG TTATAATATTAC CGGTGGCGAGC	*Ageratum* sp.	Kon *et al.* (2007)
	* * * *		

Nucletiode variation in the conserved nona-nucleotide is shown with arrowhead while nucleotide variation in body of stem is shown with asterik

Compared this DNA scratch of loop structure with other DNA-β sequences group isolated also from similar species *Capsicum annuum* in Pakistan (AM849549.1; FN179279.1; AM279661.1; AM279664.1; AM279671.1; AM279672.1; AM258978.1; AM260466.1; AM279673.1; AM279668.1; AM279663.1; AM279662.1) differentiated 1 base "C" (TTCTAATATTAC) among those 12 bases in the "top region" of hairpin loop structure. Twenty two nucleotides in total formed secondary structure to build stem supporting hairpin loop structure (Fig. 3b). Compared with the DNA-A like genome, the hairpin loop structure in the DNA-β had one "T" additional nucleotides locating in the neck position between body stem and head of hairpin loop structure (Fig. 3b).

Clustering of stem-loop structure sequence classified our DNAβ loop structure together with other loop structure from DNAβ collected from Indonesia (AB113651 and AB162142) and China (AJ965541) (Table 5). Eventhough their homology level of complete DNAβ was very low, they could form one clade, compared with other group of DNAβ sequence isolated from chili pepper in Pakistan (AM849549.1; FN179279.1; AM279661.1; AM279664.1; AM279671.1; AM279672.1; AM258978.1; AM260466.1; AM279673.1; AM279668.1; AM279663.1; AM279662.1). Two pairs (A/C;C/A)-G/T;T/G) of nucleotide variations before loop forming structure was obviously observed if stem building nucleotides from both groups was compared.

Further annotation of DNAβ sequence identified one ORF resembling C1 gene from DNA B of bipartite geminivirus. The C1-ORF spanned 357 bases in complementary sense, harboring 119 residues of amino acids. The ORF sequence isolated from DNAβ of PepYLCWSV-TD 21 had similar length with other C1-ORFs of DNAβ isolated from three different plant species (*Ludwigia hyssopifolia, Lycopersicon esculentum, Ageratum* sp.). However, compared with 12 C1-ORF sequences isolated from *Capsicum annuum* species, our C1-ORF is six bases shorter. Sequence comparison via., BLAST analysis exhibited that our C1-ORF has homology 91% with C1-ORF isolated from *Ludwigia hyssopifolia* (AJ965541.1). This result is in accordance with cluster

analysis, where our PepYLCWSV-TD21 C1-ORF was grouped together with AJ965541.1 while the other C1-ORF (AB113651.1 and AB162142.1) were grouped in different cluster (data not shown).

C1 gene harbored in the scratch of DNA**β** is believed as pathogenicity determinant associated with various plant disease exclusively by monopartite *Begomovirus* in the old world (Briddon and Stanley, 2006; Nawaz-ul-Rehman and Fauquet, 2005). For this reason, further analysis on the C1 gene of *Begomovirus* could be useful for development of resistant chili pepper.

CONCLUSION

Here, we showed the presence of monopartite Pepper yellow leaf curl virus as dominant *Begomovirus* infecting chili pepper cultivation in West Sumatera Indonesia. The monopartite genome characteristic we elucidated here is in accordance with the most descriptions reported by many authors so far, containing 6 ORFs harbored along its DNA A like genome and 1 ORF located along the sequence in complementary sense of DNA**β** genome. Among them only C1 ORF showed the lowest homology (95%) with other C1 of DNA-A like genome from different monopartite geminiviruses. This data indicated that they probably shared from common ancestor. From practical view point, this information should enrich our understanding in the distribution of mono- and bipartite genome of *Begomoviruses* particularly pepper yellow leaf curl virus infecting chili pepper. Our successful in geminivirus gene sequencing will be very valuable for resistant development of chili pepper against *Begomovirus* via., pathogen derived resistant.

ACKNOWLEDGMENTS

We thank to General Directorate of Indonesia Higher Education for supporting this research via Grant of National Strategic for year 2010-2011. Some parts of this research were also supported by Indonesian Agricultural Research and Development Agency via scheme KKP3T for fiscal year 2009.

REFERENCES

Akhter, A., J. Qazi, M. Saeed and S. Mansoor, 2009. A severe leaf curl disease on chilies in Pakistan is associated with multiple *Begomovirus* components. Plant Dis., 93: 962-962.

Arguello-Astorga, G.R. and R. Ruiz-Medrano, 2001. An iteron-related domain is associated to motif 1 in the replication proteins of geminiviruses: Identification of potential interacting amino acid-base pairs by a comparative approach. Arch Virol., 146: 1465-1485.

Briddon, R.W., S. Mansoor, I.D. Bedford, M.S. Pinner and K. Saunders *et al.*, 2001. Identification of DNA components required for induction of cotton leaf curl disease. Virology, 285: 234-243.

Briddon, R.W., S.E. Bull, I. Amin, S. Mansoor and I.D. Bedford *et al.*, 2004. Diversity of DNA 1: A satellite-like molecule associated with monopartite *Begomovirus*-DNA beta complexes. Virology, 324: 462-474.

Briddon, R.W. and J. Stanley, 2006. Subviral agents associated with plant single-stranded DNA viruses. Virology, 344: 198-210.

Briddon, R.W., J.K. Brown, E. Moriones, J. Stanley, M. Zerbini, X. Zhou and C.M. Fauquet, 2008. Recommendations for the classification and nomenclature of the DNA-**β** satellites of *Begomoviruses*. Arch. Virol., 153: 763-781.

Briddon, R.W., B.L. Patil, B. Bagewadi, M.S. Nawaz-ul-Rehman and C.M. Fauquet, 2010. Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite *Begomoviruses*. BMC Evol. Biol., Vol. 10. 10.1186/1471-2148-10-97

Chinault, A.C. and J. Carbon, 1979. Overlap hybridization screening: Isolation and characterization of overlapping DNA fragments surrounding the *leu*2 gene on yeast chromosome III. Gene, 5: 111-126.

Dellaporta, S.L., J. Wood and J.B. Hicks, 1983. A plant DNA minipreparation: Version II. Plant Mol. Biol. Rep., 1: 19-21.

Dry, I.B., J.E. Rigden, L.R. Krake, P.M. Mullineaux and M.A. Rezaian, 1993. Nucleotide sequence and genome organization of tomato leaf curl geminivirus. J. Gen. Virol., 74: 147-151.

Fauquet, C.M., R.W. Briddon, J.K. Brown, E. Moriones, J. Stanley, M. Zerbini and X. Zhou, 2008. Geminivirus strain demarcation and nomenclature. Arch. Virol., 153: 783-821.

Ha, C., S. Coombs, P. Revill, R. Harding, M. Vu and J. Dale, 2008. Molecular characterization of *Begomoviruses* and DNA satellites from Vietnam: Additional evidence that the New World geminiviruses were present in the Old World prior to continental separation. J. Gen. Virol., 89: 312-326.

Hall, T.A., 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acid Symp. Ser., 41: 95-98.

Hanley-Bowdoin, L., S.B. Settlage, B.M. Orozco, S. Nagar and D. Robertson, 1999. Geminiviruses: Models for plant DNA replication, transcription and cell cycle regulation. Crit. Rev. Plant Sci., 18: 71-106.

Huang, J.F., T. Jiang and X.P. Zhou, 2006. Molecular characterization of two distinct *Begomoviruses* infecting *Ludwigia hyssopifolia*. J. Plant Pathol., 88: 83-88.

Hussain, M., S. Iram, S. Mansoor and R.W. Briddon, 2009. A single species of betasatellite is prevalent in chilli across north central Pakistan and shows phylogeographic segregation. J. Phytopathol., 157: 576-579.

Idris, A.M. and J.K. Brown, 1998. Sinaloa tomato leaf curl geminivirus: Biological and molecular evidence for a new subgroup III virus. Phytopathology, 88: 648-657.

Idris, A.M., M.S. Shahid, R.W. Briddon, A.J. Khan, J.K. Zhu and J.K. Brown, 2011. An unusual alphasatellite associated with monopartite *Begomoviruses* attenuates symptoms and reduces betasatellite accumulation. J. Gen. Virol., 92: 706-717.

Khey-Pour, A., K. Bananej, G.A. Dafalla, P. Caciagli and E. Noris *et al*., 2000. Watermelon chlorotic stunt virus from the Sudan and Iran: Sequence comparisons and identification of a whitefly-transmission determinant. Phytopathology, 90: 629-635.

Kon, T., K. Kuwabara, S.H. Hidayat and M. Ikegami, 2007. A *Begomovirus* associated with ageratum yellow vein disease in Indonesia: Evidence for natural recombination between tomato leaf curl Java virus and Ageratum yellow vein virus-[Java]. Arch. Virol., 152: 1147-1157.

Kon, T., S.H. Hidayat, S. Hase, H. Takahashi and M. Ikegami, 2006. The natural occurrence of two distinct *Begomoviruses* associated with DNAâ and a recombinant DNA in a tomato plant from Indonesia. Phytopathology, 96: 517-525.

Li, Z., Y. Xie and X. Zhou, 2005. Tobacco curly shoot virus DNAbeta is not necessary for infection but intensifies symptoms in a host-dependent manner. Phytopathology, 95: 902-908.

Mansoor, S., S.H. Khan, A. Bashir, M. Saeed and Y. Zafar *et al*., 1999. Identification of a novel circular single-stranded DNA with cotton leaf curl disease in Pakistan. Virology, 259: 190-199.

Maruthi, M.N., S.N. Alam, K.A. Kader, A.R. Rekha, A. Cork and J. Colvin, 2005. Nucleotide sequencing, whitefly transmission, and screening tomato for resistance against two newly described *Begomoviruses* in Banglades. J. Phytopathol., 95: 1472-1481.

Navot, N., E. Pihersky, M. Zeidan, D. Zamir and H. Czosneck, 1991. *Tomato yellow leaf curl virus*: A whitefly-transmitted geminivirus with a single genomic component. Virology, 185: 151-161.

Nawaz-ul-Rehman, M.S. and C.M. Fauquet, 2005. Evolution of geminiviruses and their satellites. FEBS Lett., 583: 1825-1832.

Noris, E., E. Hidalgo, G.P. Accotto and E. Moriones, 1994. High similarity among the tomato yellow leaf curl virus isolates from the West Mediterranean Basin: The nucleotide sequence of an infectious clone from Spain. Arch. Virol., 135: 165-170.

Paximadis, M. and M.E.C. Rey, 1997. Aetiology of tobacco leaf curl in southern Africa. Ann. Appl. Biol., 131: 449-457.

Paximadis, M., A.M. Idris, I. Torres-Jerez, A. Villarreal, M.E.C. Rey and J.K. Brown, 1999. Characterization of tobacco geminiviruses in the old and new world. Arch. Virol., 144: 703-717.

Rojas, M.R., R.L. Gilbertson, D.R. Russell and D.P. Maxwell, 1993. Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. Plant Dis., 77: 340-347.

Rusli, E.S., S.H. Hidayat, R. Suseno and B. Tjahjono, 1999. Geminivirus on pepper: Symptom variation and transmission study. Bul. HPT., 11: 126-131.

Saeed, M., S.A. Behjatnia, S. Mansoor, Y. Zafar, S. Hasnain and M.A. Rezaian, 2005. A single complementary-sense transcript of a geminiviral DNA β satellite is determinant of pathogenicity. Mol. Plant Microbe Interact., 18: 7-14.

Saeed, M., Y. Zafar, J.W. Randles and M.A. Rezaian, 2007. A monopartite *Begomovirus*-associated DNA b satellite substitutes for the DNA B of a bipartite *Begomovirus* to permit systemic infection. J. Gen. Virol., 88: 2881-2889.

Sakata, J.J., Y. Shibuya, P. Sharma and M. Ikegami, 2008. Strains of a new bipartite *Begomovirus*, pepper yellow leaf curl Indonesia virus, in leaf-curl-diseased tomato and yellow-vein-diseased ageratum in Indonesia. Arch. Virol., 153: 2307-2313.

Sanderfoot, A.A. and S.G. Lazarowitz, 1996. Getting it together in plant virus movement: cooperative interactions between bipartite geminivirus movement proteins. Trends Cell Biol., 6: 353-358.

Saunders, K., I.D. Bedford, T. Yahara and J. Stanley, 2003. The earliest recorded plant virus disease. Nature, 422: 831-831.

Saunders, K., N. Salim, V.R. Mali, V.G. Malathi, R. Briddon, P.G. Markham and J. Stanley, 2002. Characterisation of Sri Lankan cassava mosaic virus and Indian cassava mosaic virus: Evidence for acquisition of a DNA B component by a monopartite *Begomovirus*. Virology, 293: 63-74.

Stanley, J., J.R. Latham, M.S. Pinner, I. Bedford and P.G. Markham, 1992. Mutational analysis of the monopartite geminivirus beet curly top virus. Virology, 191: 396-405.

Sukamto, T.K., S.H. Hidayat, K. Ito, S. Hase, H. Takahashi and M. Ikegami, 2005. *Begomoviruses* associated with leaf curl disease of tomato in Java, Indonesia. J. Phytopathol., 153: 562-566.

Sulandari, S., 2004. Biological characterization, serological assay and DNA finger printing analysis of pepper yellow leaf curl virus. M. Sc. Thesis, School of Bogor Agricultural University, Indonesia.

Sulandari, S., R. Suseno, S.H. Hidayat, J. Harjosudarmo and S. Sosromarsono, 2006. Deteksi dan kajian kisaran inang virus penyebab penyakit daun keriting kuning cabai. Hayati, 13: 1-6.

Trisno, J., S.H. Hidayat, T. Habazar, I. Manti and Jamsari, 2009. Detection and sequence diversity of *Begomovirus* associated with yellow leaf curl disease of pepper (*Capsicum annuum*) in West Sumatera, Indonesia. Microbiol. Indonesia., 3: 61-66.

Tsai, W.S., S.L. Shih, S.K. Green, A. Rauf, S.H. Hidayat and F.J. Jan, 2006. Molecular characterization of pepper yellow leaf curl indonesia virus in leaf curl and yellowing diseased tomato and pepper in indonesia. Plant Dis., 90: 247-247.

Wyatt, S.D. and J.K. Brown, 1996. Detection of subgroup III geminivirus isolates in leaf extracts by degenerate primers and polymerase chain reaction. Phytopathology, 86: 1288-1293.

Xie, Y., P. Wu, P. Liu, H. Gong and X. Zhou, 2010. Characterization of alphasatellites associated with monopartite *Begomovirus*/betasatellite complexes in Yunnan, China. Virol. J., Vol. 7.

Zhou, X., Y. Xie, X. Tao, Z. Zhang, Z. Li and C.M. Fauquet, 2003. Characterization of DNAb, associated with *Begomoviruses* in China and evidence for co-evolution with their cognate viral DNA-A. J. Gen. Virol., 84: 237-247.