



Research article

Deep sequencing reveals genetic identity of pepper yellow leaf curl Thailand virus from pepper and whiteflies from Chiang Mai, Thailand

Amisa Laprom^a, Ekachai Chukeatirote^{a,*}, Phoowanarth Maneechoat^b, Somrudee Nilthong^a

^a School of Science, Mae Fah Luang University, Chiang Rai 57100, Thailand

^b Plant Protection Research and Development Office, Department of Agriculture, Ministry of Agriculture and Cooperatives, Bangkok 10900, Thailand

Article Info

Article history:

Received 31 July 2025

Revised 19 September 2025

Accepted 11 October 2025

Available online 28 February 2026

Keywords:

Begomovirus,
Chilli,
Next-generation sequencing (NGS),
Rolling circle amplification (RCA),
Whitefly,
Yellow leaf curl disease

Abstract

Importance of the work: The production of pepper (*Capsicum* spp.) in Thailand is affected greatly by viral diseases, particularly those caused by begomoviruses. Transmitted by the whitefly *Bemisia tabaci* (Gennadius, 1889), these viruses cause severe symptoms, including yellow leaf curl and considerable yield losses. Understanding the diversity and distribution of begomoviruses is essential for developing effective disease management strategies.

Objectives: To identify and characterize begomoviruses associated with yellow leaf curl disease in pepper plants from Chiang Dao District, Chiang Mai Province, Thailand and to assess their genetic relationships to known isolates in Thailand.

Materials and Methods: Symptomatic pepper leaves and viruliferous whiteflies were collected from the field. Viral DNA was extracted, enriched using rolling circle amplification and analyzed using Illumina HiSeq next-generation sequencing. *De novo* genome assembly and sequence comparisons were conducted to identify *Begomovirus* components and to assess genetic variation.

Results: Components (two full-length DNA-A and one DNA-B) were recovered from both plant and whitefly samples. The DNA-A sequences showed 98.0–99.3% identity to *Pepper yellow leaf curl Thailand virus* (PepYLCTHV) isolate SNS-CM5, while DNA-B shared 96.0% identity with isolates WF-SPN-Pep2015 and TMK-KR8. Minor sequence variation was observed in AC1 and AC4 genes. Phylogenetic analysis confirmed that the Chiang Dao isolates formed a distinct clade within the PepYLCTHV group.

Main finding: PepYLCTHV is the predominant *Begomovirus* associated with yellow leaf curl disease in pepper in northern Thailand and the Chiang Dao isolates represent a genetically consistent but regionally distinct cluster, emphasizing the need for localized surveillance in viral disease management.

* Corresponding author.

E-mail address: ekachai@mfu.ac.th (E. Chukeatirote)

online 2452-316X print 2468-1458/Copyright © 2025. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), production and hosting by Kasetsart University Research and Development Institute on behalf of Kasetsart University.

<https://doi.org/10.34044/j.anres.2026.60.1.05>

Introduction

In Thailand, *Capsicum annuum* L. and *C. frutescens* L. are among the most important vegetable crops, cultivated extensively for both domestic consumption and export, with *C. annuum* (commonly known as hot chili pepper) being the dominant species, covering approximately 134,000 ha and yielding an estimated 252,000 t annually (Amekawa et al., 2023). Chili production serves as a vital source of income for many Thai farmers, particularly smallholders. However, in recent years, the cultivated area has been declining steadily, largely due to the impact of viral diseases that considerably reduce both yield and fruit quality, with begomoviruses currently regarded as the most prevalent and destructive viral pathogens affecting chili production in Thailand (Chiemsoombat et al., 2018; Laprom et al., 2019; Charoenvilaisiri et al., 2020). Although there is limited comprehensive field-based data on the yield losses caused by *Begomovirus* infections in chili crops in Thailand, economic losses could reach up to 95% under epiphytotic conditions (Kumsee et al., 2024).

The genus *Begomovirus*, comprises 464 species and, according to International Committee on Taxonomy of Viruses (2025), is the largest and most agriculturally important genus within the family *Geminiviridae*. These plant-infecting viruses cause substantial crop losses, particularly in tropical and subtropical regions (Zerbini et al., 2017). Transmitted exclusively by the whitefly *Bemisia tabaci* (Gennadius, 1889), begomoviruses are characterized by their distinctive geminate (twinned icosahedral) particles and circular single-stranded DNA (ssDNA) genomes (Fiallo-Olive and Navas-Castillo, 2023). The genome may be either monopartite or bipartite, typically in the range 2.5–2.8 kilobases per component (Brown et al., 2015). Bipartite begomoviruses consist of two components: 1) DNA-A, which encodes proteins involved in replication (Rep), transcriptional regulation (TrAP and REN), RNA silencing suppression (AC4) and encapsidation (coat protein, CP); and 2) DNA-B, which encodes movement-related proteins, including the nuclear shuttle protein (NSP) and movement protein (MP) (Briddon et al., 2010). Monopartite begomoviruses carry only a DNA-A-like component but often associate with satellite DNAs, such as betasatellites, that enhance pathogenicity and aid in virus movement (Saeed et al., 2007). Both genome types share a conserved intergenic region containing a stem-loop structure and a nonanucleotide motif (TAATATTAC), which serves as the origin of rolling circle replication (Briddon et al., 2010). This compact genome

architecture allows begomoviruses to efficiently replicate, adapt to diverse hosts and spread rapidly.

Despite the availability of various diagnostic tools, challenges remain in accurately detecting and characterizing the incidence of begomoviruses infecting chili and their whitefly vectors in field conditions. Often, conventional methods rely on multiple sequential techniques, which can be time-consuming. Furthermore, there is a limited amount of comprehensive molecular data on begomoviruses circulating in the chili-growing regions of northern Thailand, hindering effective disease management strategies. Therefore, the present study hypothesized that an integrated molecular approach, combining rolling circle amplification (RCA) and next-generation sequencing (NGS), could provide a more rapid, sensitive and comprehensive detection of begomoviruses in both pepper plants and their insect vectors. This approach aimed to fill existing knowledge gaps by characterizing the viral populations present in the field, thereby supporting improved diagnostics and management of viral diseases in chili production.

Materials and Methods

Sample collection and preliminary virus detection

Approximately 10–20 pepper plants exhibiting typical yellow leaf curl symptoms were collected from chili cultivation fields in Chiang Dao District, Chiang Mai Province, Thailand. Individual symptomatic plants were placed in separate Zip-lock plastic bags and transported in an icebox to the laboratory for further processing. Any whiteflies present on the collected leaves were captured using a manual aspirator and transferred into insect collection vials, which were also kept on ice to preserve DNA integrity. Total DNA was extracted from both plant tissues and pooled whitefly bodies using the same miniprep method, as described by Laprom et al. (2019). Initially, the presence of *Begomovirus* DNA-A components in the field-collected samples was screened using polymerase chain reaction (PCR) with degenerate primers specific to begomoviruses: PAL1v1978 (5'-GCATCTGCAGGCCACATYGTCTTYCCNGT-3') and PAR1c496 (5'-AATACTGCAGGCTTYCTRTACATRG-3'), following protocols established by Laprom et al. (2019). Each 25 µL PCR reaction contained DreamTaq Green PCR Master Mix (Thermo Fisher Scientific; USA), 1.0 µM of each primer and 1 µg of template DNA. Amplification was

carried out in a GeneAmp PCR System 2700 thermal cycler (Applied Biosystems; USA) using the cycling conditions: initial denaturation at 95°C for 2 min; 30 cycles of denaturation at 95°C for 30 s, annealing at 55°C for 30 s and extension at 72°C for 60 s; followed by a final extension at 72°C for 10 min. The PCR products were analyzed using electrophoresis on a 1% agarose gel in 0.5× TBE buffer and visualized under ultraviolet light.

Rolling circle amplification and high-throughput sequencing

Total DNA from two PCR-confirmed samples—S1 (viruliferous whitefly) and S2 (symptomatic pepper plant)—was subjected to RCA for begomoviral genome enrichment, as described by Haible et al. (2006). RCA was performed using the TempliPhi™ DNA Amplification Kit (GE Healthcare; UK), following the manufacturer's protocol. Amplified products were purified using the GeneJET Gel Extraction Kit (Thermo Fisher Scientific; USA) and their concentration and purity were assessed using NanoDrop spectrophotometry and 1% agarose gel electrophoresis. The purified RCA products were concentrated to a final yield of 25 µg and dried using GenTegra-DNA tubes (GENTEGRA; USA), according to the supplier's instructions. Library preparation for next-generation sequencing (NGS) was carried out using the NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (New England Biolabs; USA) and paired-end sequencing (2 × 150 bp) was performed on the Illumina HiSeq platform, yielding approximately 1.0 Gb of raw data per sample (Vishuo Biomedical Thailand, Ltd; Thailand).

Next-generation sequencing data processing and viral genome assembly

Raw Illumina reads obtained from the RCA-enriched DNA of sample S1 (whitefly) contained a mixture of host and viral sequences. These reads were mapped to *Begomovirus* reference genomes retrieved from GenBank and *de novo* viral genome assembly was performed using the Geneious Prime software (version 2021.2; <https://www.geneious.com>; Biomatters; Switzerland). A similar bioinformatics pipeline was applied to reads from sample S2 (pepper plant), which contained both *Capsicum* genomic and begomoviral sequences. Mapping, assembly and genome annotation were conducted to reconstruct full-length viral genomes. A schematic overview of the experimental and analytical workflow is presented in Fig. 1.

Comparative and phylogenetic analysis

Assembled viral genomes from the RCA-NGS datasets of samples S1 and S2 were compared against full-length *Begomovirus* sequences in GenBank, focusing on isolates reported to infect pepper in Thailand. Sequence similarity and genetic distance were calculated and phylogenetic relationships were inferred using the neighbor-joining method implemented in Geneious Prime. These analyses facilitated the identification and classification of viral species and helped to elucidate their evolutionary relationships with known *Begomovirus* isolates.

Results and Discussion

Symptom description and initial detection of begomovirus

The pepper plants collected from Chiang Dao District, Chiang Mai Province, Thailand, exhibited characteristic *Begomovirus* symptoms, including yellowing and curling of leaves, narrowing of young leaves and chlorotic blistering on older foliage. In some cases, affected fruits and stems also showed pale green discoloration and curling of the epidermis (Fig. 2). These symptoms were consistent with those commonly reported in *Begomovirus*-infected chili plants (Chiemsoombat et al., 2018; Laprom et al., 2019). PCR amplification of the total DNA extracted from both the symptomatic pepper plant (sample S2) and the viruliferous whitefly (*Bemisia tabaci*, sample S1) yielded DNA fragments of approximately 1.3 kb, which was within the expected size range of 1.1–1.4 kb for the *Begomovirus* DNA-A component, as reported by Rojas et al. (1993), confirming the presence of *Begomovirus* in both hosts.

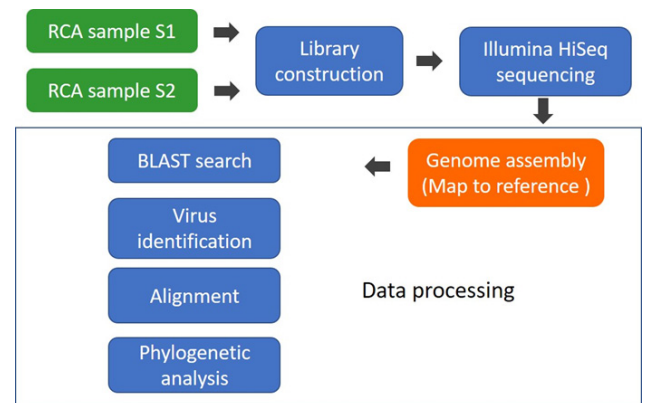


Fig. 1 Schematic overview of key steps involved in assembly and analysis of begomoviral genome sequences from symptomatic pepper plants and viruliferous whitefly samples.



Fig. 2 Symptoms observed in *Begomovirus*-infected pepper plants collected from Chiang Dao District, Chiang Mai Province: (A) yellowing and upward curling of leaves (yellow leaf curl); (B) pale green discoloration and deformation of fruit surface; (C) curly and rough skin on branches and fruit, indicative of severe viral infection.

High-throughput sequencing and genome assembly

Following RCA, high-throughput sequencing was performed using the Illumina HiSeq platform. The processed sequencing data retained 11,604,808 reads for S1 and 10,255,204 reads

for S2. Reads were mapped against reference genomes of *C. annuum* and *B. tabaci* and *Begomovirus* genomes were assembled using Geneious Prime. A large number of contigs were generated from each sample (Table 1) and full-length viral genomes corresponding to the bipartite components (DNA-A and DNA-B) were reconstructed successfully.

The DNA-A and DNA-B sequences for the whitefly sample (S1) were 2,736 and 2,741 nucleotides (nt), respectively. The pepper plant sample (S2) yielded DNA-A and DNA-B sequences of 2,743 and 2,738 nt, respectively (Table 2). These genomes were deposited in GenBank under accession numbers OM677627–OM677630. The BLASTn analysis indicated high sequence identity to *Pepper yellow leaf curl Thailand virus* (PepYLCTHV), with DNA-A from S1 and S2 showing 92.87% and 98.47% identity, respectively, to isolate SNS-CM5 (KT322142) and DNA-B showing 91.48% and 96.43% identity, respectively, to isolate WF-SPN-Pep2015 (KX943290).

Table 1 Reference-guided assembly data for RCA-enriched begomovirus complexes

Begomovirus reference species	Accession No.	Length (nt)	Type	S1 (Chiangdao-whiteflies)				S2 (Chiangdao-pepper)			
				Total reads	Consensus (nt)	% Coverage	% Identity	Total reads	Consensus (nt)	% Coverage	% Identity
Eggplant golden mosaic virus	KU569601	2733	DNA-A	2,784,801	2,735	100	94.90	3,383,823	2,755	100	97.50
Eggplant golden mosaic virus	KU569586	2742	DNA-A	2,790,206	2,739	100	94.70	3,377,465	2,755	100	97.70
Eggplant golden mosaic virus	KU569598	2742	DNA-A	2,790,002	2,756	100	94.70	3,377,433	2,755	100	97.70
Erectites yellow mosaic virus	NC009549	2751	DNA-A	2,487,321	2,752	100	95.40	3,373,218	2,755	100	98.00
Pepper leaf curl virus	AF414287	2739	DNA-A	2,784,852	2,732	100	93.90	3,387,083	2,755	100	97.60
Pepper yellow leaf curl Thailand virus	MK946436	2742	DNA-A	2,786,216	2,737	100	94.70	3,372,224	2,755	100	97.80
Pepper yellow leaf curl Thailand virus	KT322141	2742	DNA-A	2,789,887	2,735	100	94.70	3,377,348	2,755	100	97.70
Pepper yellow leaf curl Thailand virus	KT322143	2742	DNA-A	2,790,349	2,735	100	94.70	3,377,513	2,755	100	97.70
Pepper yellow leaf curl Thailand virus	KT322146	2742	DNA-A	2,789,887	2,744	100	94.70	3,377,348	2,755	100	97.70
Pepper yellow leaf curl Thailand virus	KX943290	2742	DNA-A	2,790,421	2,736	100	94.70	6,683,863	2,755	100	97.40
Pepper yellow leaf curl Thailand virus*	KT322142	2733	DNA-A	2,773,473	2,734	100	95.10	3,383,864	2,743	100	97.40
Pepper leaf curl virus	AF414287	2739	DNA-A	1,086,217	2,737	1	99.20	999,579	2,741	1	99.00
Tomato yellow leaf curl Kanchanaburi virus	KU569584	2753	DNA-A	8,786	362	0.131	99.50	10,050	394	0.141	99.30
Pepper yellow vein Mali virus isolate	KY271077	2786	DNA-A	2,447	296	0.107	99.20	11,651	394	0.144	99.20
Chilli leaf curl virus	MK468696	2761	DNA-A	4,233	688	0.25	99.20	12,338	665	0.238	99.30
Tomato yellow leaf curl Thailand virus	NC000869	2751	DNA-A	24,978	1,303	0.474	99.40	125,459	1,115	0.406	99.10
Tobacco curly shoot virus	NC003722	2743	DNA-A	518,588	1,376	0.498	98.60	195,095	1,361	0.494	99.20
Tomato leaf curl Taiwan virus	NC003898	2739	DNA-A	54,671	1,354	0.482	99.40	307,817	2,092	0.754	95.80
Tomato leaf curl China virus	NC005320	2738	DNA-A	740,849	1,299	0.471	99.10	190,572	1,251	0.458	96.60
Papaya leaf curl China virus	NC005321	2751	DNA-A	283,801	2,285	0.834	98.80	329,909	2,261	0.822	98.90
Ageratum leaf curl virus	NC006384	2735	DNA-A	110,059	1,716	0.627	99.40	132,776	1,683	0.614	99.20
Mimosa yellow leaf curl virus	NC009546	2757	DNA-A	30,866	1,110	0.403	96.10	259,337	1,677	0.605	99.20
Erectites yellow mosaic virus	NC009549	2751	DNA-A	729,998	2,755	1	99.30	1,017,288	2,755	1	98.10
Sunn hemp leaf distortion virus	NC013019	2774	DNA-A	7,119	316	0.114	99.40	–	–	–	–
Pepper leaf curl Lahore virus	NC016984	2747	DNA-A	2	150	0.055	93.30	2	150	0.055	90.70
Pepper yellow leaf curl Thailand virus	MK946435	2732	DNA-B	8,065,590	–	1	98.40	3,372,527	2,754	100	95.70
Pepper yellow leaf curl Thailand virus*	NC040186	2732	DNA-B	8,112,146	2,753	1	98.40	3,418,930	2,738	100	95.50
Pepper yellow leaf curl Thailand virus	KX885224	2731	DNA-B	8,118,795	–	1	98.30	469,006	2,725	100	98.90
Pepper yellow leaf curl Thailand virus	KX885225	2730	DNA-B	8,125,669	–	1	98.20	1,140,148	2,737	100	97.80
Tomato leaf curl New Delhi virus	LC431619	2739	DNA-B	1,422,922	2,223	0.811	99.10	1,752,878	1,644	0.598	98.40
Tomato yellow leaf curl Thailand virus	NC000870	2737	DNA-B	2,432,332	2,120	0.782	98.70	1,484,917	2,060	0.755	96.30
Tomato yellow leaf curl Kanchanaburi virus	KU569593	2753	DNA-B	5,632,431	2,757	1	98.30	1,883,727	2,757	1	98.10
Pepper yellow leaf curl Indonesia virus	MN094865	2739	DNA-B	38,250	495	0.18	99.30	–	–	–	–

Asterisks indicate the selected consensus sequences of begomovirus DNA-A and DNA-B obtained from this study

Table 2 Details of samples (Host species, location, and year of collection), Illumina HiSeq read data generated and processed, and genome assembly contig results obtained using Geneious Prime

Sample	Raw reads	Unmapped reads	Viral reads	Genomic component	BLAST result
S1 (<i>Bemisia tabaci</i> ; Chiang Dao; 2019)	11,604,808	8,831,335	2,773,473	DNA-A	PepYLCTHV- SNS-CM5
		3,492,662	8,112,146	DNA-B	PepYLCTHV- WF-SPN-Pep2015
S2 (<i>Capsicum annum</i> ; Chiang Dao; 2019)	10,255,204	6,871,340	3,383,864	DNA-A	PepYLCTHV- SNS-CM5
		6,836,274	3,418,930	DNA-B	PepYLCTHV- WF-SPN-Pep2015

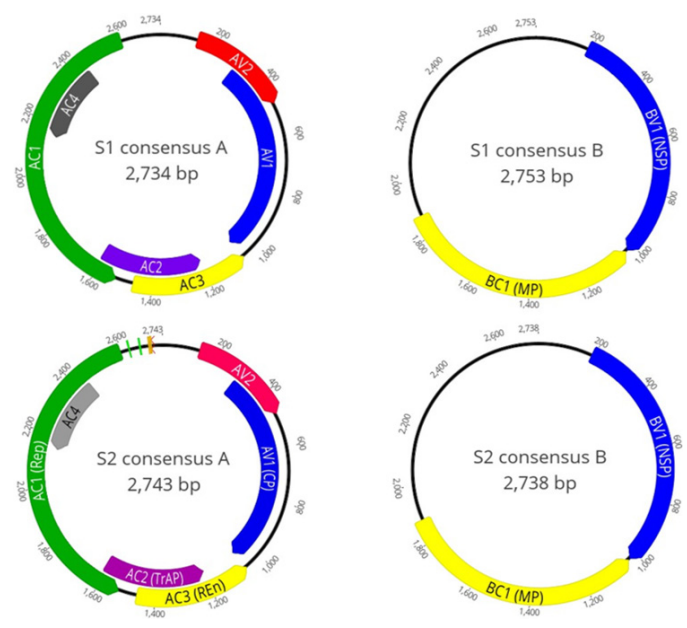
Genome annotation and organization

Annotation of the full-length viral genomes revealed typical bipartite begomovirus architecture. DNA-A encoded six open reading frames (ORFs)—AV1, AV2, AC1, AC2, AC3 and AC4—while DNA-B encoded two ORFs—BC1 (nuclear shuttle protein) and BC2 (movement protein)—consistent with PepYLCTHV and other bipartite *Begomovirus* genomes (Briddon et al., 2010; Lozano et al., 2016). The intergenic regions measured 268 nt in S1 and 275 nt in S2 and contained the conserved nonanucleotide (TAATATTAC) motif within a stem-loop structure, which is the origin of rolling circle replication (Brown et al., 2015). These genomic features confirmed the identity of the virus as PepYLCTHV (Table 2 and Fig. 3).

Sequence analysis and species classification

Comparative analyses of the ORF nucleotide and amino acid sequences revealed identities in the ranges 88–100% and 70–100%, respectively (Table 3). Both samples are considered strains of PepYLCTHV, according to the species demarcation threshold established by International Committee on Taxonomy of Viruses (2025), which defines distinct *Begomovirus* species as having < 91% DNA-A nucleotide identity (Brown et al., 2015). Amino acid comparisons across key proteins showed strong similarity to the PepYLCTHV SNS-CM5 isolate:

93–95% identity for MP, 97–99% for NSP, 97–99% for CP and 91–100% for Rep, among others. These findings were consistent with other reports of PepYLCTHV as a dominant *Begomovirus* infecting peppers in Thailand (Chiemsoombat et al., 2018; Laprom et al., 2019).

**Fig. 3** Genome organization of PepYLCTHV isolates S1 and S2 reconstructed from Illumina read assemblies, indicating open reading frames AV1, AV2, AC1, AC2, AC3, AC4, BV1 and BC1.**Table 3** Percentage pairwise sequence identities among begomovirus isolates from yellow leaf curl pepper plants

Component	ORF name	Gene name	Protein name	Sample S1		Sample S2	
				Nucleotide (%)	Amino acid (%)	Nucleotide (%)	Amino acid (%)
DNA-A	ORF2	AV1	CP	94-96	97-98	94-99	98-99
	ORF1	AV2	V2(MP)	95-98	97-99	96-99	97-100
	ORF4	AC1	Rep	89-96	84-96	88-99	84-99
	ORF6	AC2	TrAP	92-94	89-93	92-99	90-99
	ORF3	AC3	Ren	91-93	85-87	91-98	86-98
	ORF5	AC4	C4	90-97	74-94	88-99	70-97
DNA-B	ORF7	BV1	NSP	94-100	94-100	95-98	96-98
	ORF8	BC1	MP	93-99	95-100	94-99	96-100

The four complete genome sequences obtained in the present study were aligned with the *Begomovirus* sequences available in the NCBI database. Pairwise sequence comparisons of DNA-A (Table 4) and DNA-B (Table 5) from PepYLCTHV isolates S1 and S2, together with other begomoviruses infecting pepper and tomato, showed nucleotide identities in the range 75–98%. The highest identity was observed with the PepYLCTHV isolate SNS-CM5 from Chiang Mai province (KT322142). Neighbor-joining phylogenetic analysis of DNA-A (Fig. 4A) placed

isolates S1 and S2 in separate clusters within the PepYLCTHV clade. Isolate S1 grouped with WF-SPN-Pep2015, SPN-PG1, BRM103, KON-KG5, TMK-KR8, RBI-24, KRI155 and KRI157. In contrast, isolate S2 clustered with SNS-CM5, KKN601, KKN604 and KKN603. Phylogenetic analysis of DNA-B (Fig. 4B) showed that isolates S1 and S2 were grouped together within a cluster containing Thai PepYLCTHV isolates, clearly separated from other *Begomovirus* species infecting pepper, reflecting geographic distribution.

Table 4 Percent nucleotide sequence similarity among begomovirus DNA-A complete genome segments

Begomovirus species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1 AF414287 PepLCV-Malaysia																												
2 KP868762 ChiLCV-TC-Vns	76																											
3 KT322141 PepYLCTHV-KON-KG5	90	75																										
4 KT322142 PepYLCTHV-SNS-CM5	89	76	93																									
5 KT322143 PepYLCTHV-SPN-PG1	90	76	99	93																								
6 KT322146 PepYLCTHV-TMK-KR8	90	75	100	93	99																							
7 KT699194 ChiLCV-Pakistan	77	88	77	78	77	77																						
8 KX943290 PepYLCTHV-WF-SPN-Pep2015	90	76	99	93	99	99	77																					
9 MK468696 ChiLCV-Tob-13	76	82	76	76	76	76	85	76																				
10 MK946436 PepYLCTHV-BRM103	89	75	98	92	98	98	76	98	75																			
11 MW715819 PepYLCTHV-KKN601	90	76	95	97	95	95	78	95	77	94																		
12 MW715821 PepYLCTHV-KKN603	90	76	95	97	94	95	78	95	77	94	100																	
13 MW715822 PepYLCTHV-KKN604	90	76	95	97	95	95	78	95	77	94	100	100																
14 MW715823 PepYLCTHV-RBI-24	90	76	99	93	99	99	77	99	76	98	95	95	95															
15 MW715824 PepYLCTHV-KRI155	90	76	99	93	99	99	77	99	76	98	95	95	95	100														
16 MW715825 PepYLCTHV-KRI157	90	76	99	93	99	99	77	99	76	98	95	95	95	100	100													
17 NC000869 TYLCTHV-Thailand	79	75	79	80	80	79	76	80	75	79	80	80	80	80	80	80												
18 NC000882 PepLCV Thailand	94	74	88	87	88	88	75	88	75	87	88	88	88	88	88	88	78											
19 NC003897 ToLCKV-India	77	83	77	77	77	77	84	77	80	77	78	78	78	77	77	78	77	75										
20 NC003898 ToLCTWV-Taiwan	81	75	80	81	80	80	76	80	75	80	82	81	81	80	80	80	76	80	76									
21 NC004153 ToLCVV-Viet Nam	82	77	82	82	82	82	77	82	77	81	83	82	83	82	82	82	77	80	77	81								
22 NC004613 ToLCLV-Laos	76	74	76	77	77	76	75	77	74	76	77	77	77	76	76	76	75	74	75	78	76							
23 NC005320 ToLCCNV-G32-China	78	76	78	78	78	78	77	78	76	77	79	79	79	78	78	78	79	76	77	78	88	74						
24 NC028046 CHLCAV-India	75	86	74	75	75	74	85	75	85	74	75	75	75	75	75	75	73	80	75	76	74	75						
25 NC028989 PepYLCTHV-KON-KG5	90	75	100	93	99	100	77	99	76	98	95	95	95	99	99	99	79	88	77	80	82	76	78	74				
26 NC029995 TYLCSHV-Y4536	76	78	76	76	76	79	76	76	76	77	77	77	77	76	76	76	76	75	80	76	76	79	75	75	76			
27 S1 consensus A*	89	76	90	92	90	90	77	90	76	90	94	94	94	90	90	90	79	87	77	81	81	77	78	75	90	77		
28 S2 consensus A*	89	76	93	98	93	93	77	93	76	92	96	96	96	93	93	93	80	88	78	81	82	77	78	75	93	77	93	

Asterisks indicate the begomovirus isolate obtained from this study

Table 5 Percent nucleotide sequence similarity among begomovirus DNA-B complete genome segments

Begomovirus species	1	2	3	4	5	6	7	8	9	10	11	12	13
1 DQ169055 TYLCKaV-Viet Nam													
2 KU569593 TYLCKaV-TH12E1	87												
3 KX885224 PepYLCTHV-KON-KG5B	80	81											
4 KX885225 PepYLCTHV-TMK-KR8BÂ	80	80	89										
5 KX943291 PepYLCTHV-WF-SPN-Pep2015	81	81	90	98									
6 MK946435 PepYLCTHV-BRM103	80	81	97	89	89								
7 MT001885 PepYLCTHV-Thailand	80	80	89	96	98	88							
8 MW715820 PepYLCTHV-KKN601	80	80	86	89	89	85	90						
9 NC000870 TYLCTHV-Thailand	75	75	75	74	75	74	74	74					
10 NC005811 TYLCKaV-Thailand:Kanchanaburi	87	98	82	81	81	81	81	76					
11 NC040186 PepYLCTHV-WF-SPN-Pep2015	81	81	90	98	100	89	98	75	81				
12 S1 consensus B*	81	81	87	90	91	87	90	91	75	82	91		
13 S2 consensus B*	81	82	91	96	96	90	95	90	75	82	96	93	

Asterisks indicate the begomovirus isolate obtained from this study

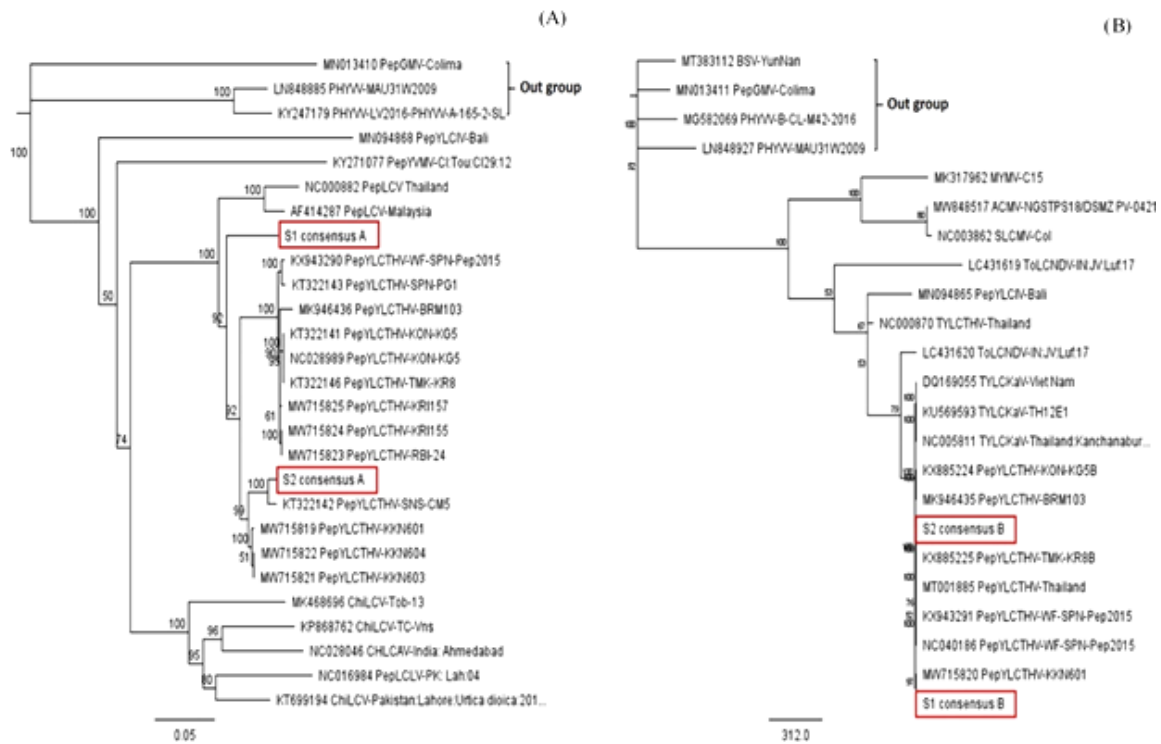


Fig. 4 Phylogenetic trees based on multiple sequence alignments of complete PepYLCTHV DNA-A and DNA-B genomes obtained in this study (highlighted with red boxes) alongside selected isolates from NCBI, where trees were constructed using neighbor-joining method, bootstrap analysis was performed with 1,000 replicates and bootstrap support values are indicated on branches.

Detection of the same *Begomovirus* strain in both symptomatic pepper plants and whiteflies underscores the key epidemiological role of *B. tabaci* as a vector, efficiently transmitting PepYLCTHV across crops (Zerbini et al., 2017). The high sequence identity between virus genomes isolated from plant and insect further confirmed the linkage between field-infected plants and active vector populations, reinforcing other findings on *Begomovirus* transmission dynamics (Lozano et al., 2016).

The use of RCA-NGS enabled enrichment and high-resolution genomic characterization of circular ssDNA viruses directly from complex field samples without prior sequence knowledge. Compared to conventional methods, such as PCR or enzyme-linked immunosorbent assay, RCA-NGS offers a more comprehensive, sensitive and high-throughput approach to detect full viral genomes and associated satellite DNAs in a single workflow (Haible et al., 2006; Mehetre et al., 2021). This strategy not only facilitates virus discovery but also enhances understanding of *Begomovirus* diversity, evolution and population structure (Edwards and Rohwer, 2005; Posada-Céspedes et al., 2017). While the present study successfully identified PepYLCTHV as the causal agent of yellow leaf curl disease in the sampled region, these findings were limited by the geographical scope

(Chiang Dao District) and sample size. Broader sampling across multiple regions and growing seasons is necessary for the full characterization of the diversity and distribution of *Begomovirus* populations in Thailand. Furthermore, follow-up validation is recommended using full-length cloning and Sanger sequencing to confirm genome assemblies and resolve any ambiguities inherent in NGS-based methods (Dayaram et al., 2012).

Genome recovery and methodological justification

Despite the high volume of sequencing data generated via Illumina HiSeq, only three full-length viral genome components (two DNA-A and one DNA-B) were recovered. This may be attributed to the dominance of a single *Begomovirus* species in the collected samples, resulting in high sequence depth for this virus and limiting the detection of mixed infections. In addition, the RCA-based enrichment step may have contributed to selective amplification of the most abundant circular viral genomes, introducing amplification bias. Additional viral diversity may be revealed by future studies using broader sampling and amplification-independent metagenomic approaches.

Comprehensive analysis of the NGS data revealed no evidence of co-infection with other begomoviruses or unrelated plant viruses, indicating a mono-infection scenario dominated by PepYLCTHV in both the plant and whitefly samples. No other *Begomovirus* species or unrelated viral genomes were detected in the dataset, supporting the conclusion that PepYLCTHV was the sole viral agent present in these samples.

This study did not include PCR screening for satellite DNAs (such as betasatellites or alphasatellites) using universal primers. While the NGS data did not reveal sequences corresponding to known satellite DNAs, it is possible that low-abundance satellites may have escaped detection due to amplification or sequencing biases. Future work should include targeted PCR-based detection using established universal primers to confirm the presence or absence of satellite molecules.

The RCA-NGS approach was introduced in the present study to take advantage of its higher sensitivity, throughput and the ability to detect complete viral genomes without prior knowledge of sequence identity. NGS allows for the identification of minor variants, recombination events and co-infecting viral species that may be missed by conventional cloning-based methods (Edwards and Rohwer, 2005; Posada-Céspedes et al., 2017). In the present study, although only three full-length viral genome components were recovered, the RCA-NGS strategy ensured comprehensive and unbiased genome coverage from both the plant and vector samples. Furthermore, the NGS platform enabled simultaneous screening of multiple samples and viruses, providing a scalable approach for future epidemiological surveillance.

Although PCR-based validation is used commonly to confirm NGS findings, it was not performed in the present study because similar PCR-based detection of PepYLCTHV had been demonstrated robustly in other work (Chiemsombat et al., 2018; Laprom et al., 2019). However, the viral sequences obtained here showed high coverage and identity (> 98%) to known PepYLCTHV isolates and typical full-length genome structure with intact ORFs, supporting their authenticity. Furthermore, RCA enriches specifically for circular DNA, reducing host contamination and increasing the reliability of viral genome recovery.

Conclusion

The combined RCA-NGS approach effectively identified and characterized PepYLCTHV strains infecting pepper plants

and their whitefly vectors in northern Thailand, revealing typical bipartite *Begomovirus* genome organizations and high sequence similarity to known isolates. This high-throughput technique enhances *Begomovirus* detection and genomic analysis, providing critical insights into viral diversity and epidemiology in key agricultural regions. Future research is recommended with expanded sampling and validation to deepen understanding of *Begomovirus* population dynamics and to inform sustainable disease control strategies in chili pepper cultivation.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

In loving memory of my beloved mother, Boontiwa Chukeatirote (3 April 2483 B.E.–13 October 2568 B.E.). Mae Fah Luang University and Bayer Crop Science provided funding.

References

- Amekawa, Y., Hongsibsong, S., Sawarng, N., Gebre, G.G. 2023. Chili pepper farmers' pesticide use and residues under Thailand's public good agricultural practices standard: A case study in Chiang Mai Province. *Agriculture*. 13: 1105. doi.org/10.3390/agriculture13051105
- Briddon, R.W., Patil, B.L., Bagewadi, B., Nawaz-ul-Rehman, M.S., Fauquet, C.M. 2010. Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. *BMC Evol. Biol.* 10: 97. doi.org/10.1186/1471-2148-10-97
- Brown, J.K., Zerbini, F.M., Navas-Castillo, J., et al. 2015. Revision of *Begomovirus* taxonomy based on pairwise sequence comparisons. *Arch. Virol.* 160: 1593–1619. doi.org/10.1007/s00705-015-2398-y
- Charoenvilaisiri, S., Seepiban, C., Phironrit, N., Phuangrat, B., Yoohat, K., Deeto, R., Chatchawankanphanich, O., Gajanandana, O. 2020. Occurrence and distribution of begomoviruses infecting tomatoes, peppers and cucurbits in Thailand. *Crop Prot.* 127: 104948. doi.org/10.1016/j.cropro.2019.104948
- Chiemsombat, P., Srikamphung, B., Yule, S. 2018. Begomoviruses associated to [sic] pepper yellow leaf curl disease in Thailand. *J. Agri. Res.* 3: 183. doi.org/10.23880/OAJAR-16000183
- Dayaram, A., Opong, A., Jaschke, A., et al. 2012. Molecular characterisation of a novel cassava associated circular ssDNA virus. *Virus Res.* 166: 130–135. doi.org/10.1016/j.virusres.2012.03.009
- Edwards, R.A., Rohwer, F. 2005. Viral metagenomics. *Nat. Rev. Microbiol.* 3: 504–510. doi.org/10.1038/nrmicro1163

- Fiallo-Olive, E., Navas-Castillo, J. 2023. Begomoviruses: what is the secret(s) of their success? *Trends Plant Sci.* 28: 715–727. doi.org/10.1016/j.tplants.2023.01.012
- Haible, D., Kober, S., Jeske, H. 2006. Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. *J. Virol. Methods* 135: 9–16. doi.org/10.1016/j.jviromet.2006.01.017
- International Committee on Taxonomy of Viruses. 2025. ICTV Report (Family: *Geminiviridae*, Genus: *Begomovirus*). <https://ictv.global/report/>, 19 September 2025.
- Kumsee, N., Suwor, P., Teerarak, M., et al. 2024. Improving pepper inbreds for resistance to *Pepper yellow leaf curl Thailand virus* (PepYLCTHV) through challenged inoculations. *Horticulturae*. 10: 1000. doi.org/10.3390/horticulturae10091000
- Laprom, A., Nilthong, S., Chukeatirote, E. 2019. Incidence of viruses infecting pepper in Thailand. *Biomol. Concepts*. 10: 184–193. doi.org/10.1515/bmc-2019-0021
- Lozano, G., Trenado, H.P., Fiallo-Olive, E., Chirinos, D., Geraud-Pouey, F., Briddon, R.W., Navas-Castillo, J. 2016. Characterization of non-coding DNA satellites associated with sweepoviruses (Genus *Begomovirus*, Geminiviridae) - definition of a distinct class of begomovirus-associated satellites. *Front. Microbiol.* 7: 162. doi.org/10.3389/fmicb.2016.00162
- Mehetre, G.T., Leo, V.V., Singh, G., et al. 2021. Current developments and challenges in plant viral diagnostics: A systematic review. *Viruses*. 13: 412. doi.org/10.3390/v13030412
- Posada-Céspedes, S., Seifert, D., Beerenwinkel, N. 2017. Recent advances in inferring viral diversity from high-throughput sequencing data. *Virus Res.* 239: 17–32. doi.org/10.1016/j.virusres.2016.09.016
- Rojas, M.R., Gilbertson, R.L., Russell, D.R., Maxwell, D.P. 1993. Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. *Plant Dis.* 77: 340–347.
- Saeed, M., Zafar, Y., Randles, J.W., Rezaian, M.A. 2007. A monopartite begomovirus-associated DNA beta satellite substitutes for the DNA B of a bipartite begomovirus to permit systemic infection. *J. Gen. Virol.* 88: 2881–2889. doi.org/10.1099/vir.0.83049-0
- Zerbini, F.M., Briddon, R.W., Idris, A., et al. 2017. ICTV virus taxonomy profile: Geminiviridae. *J. Gen. Virol.* 98: 131–133. doi.org/10.1099/jgv.0.000738