

## Isolation and characterization of *Rep* PepYLCV encoding gene from West Sumatra

Bastian Nova<sup>1</sup>, Elly Syafriani<sup>2</sup>, Lily Syukriani<sup>2</sup>, Alfi Asben<sup>3</sup>, Jamsari Jamsari<sup>2\*</sup>

<sup>1</sup>Biotechnology Department, Magister Program, Andalas University, Padang, West Sumatera, Indonesia

<sup>2</sup>Plant Breeding Department, Faculty of Agriculture, Andalas University, Padang, West Sumatera, Indonesia

<sup>3</sup>Department of Food Technology, Faculty of Agriculture, Andalas University, Padang, West Sumatera, Indonesia

\*Corresponding author: [ajamsari@yahoo.com](mailto:ajamsari@yahoo.com)

### Abstract

Replicase (*Rep*) protein of Geminivirus is known as one of the important components not only for its successful replication in their host but also known to interact with various host plant proteins. However, it is still unclear if those interactions are associated with symptoms level. This research aims to explore the possibility of *Rep* as pathogenic determinant by in silico approach. Here we report the comparison of three *Rep* sequences isolated from Pesisir Selatan and Tanah Datar districts in West Sumatra Indonesia. The PCR-based cloning approach was used in this study to isolate the gene sequences from all isolates. Pathogenic determinant was predicted from phenotype and genotype analysis. Phenotype data showed symptoms appearance after 8 dpi for PSSWS14 and 20 dpi for PSSWS3. Furthermore, genotype showed that the nonconserved region in N-terminal of *Rep* makes different in its putative binding site. It is prospective to be related to the symptoms appearance rates. We predict the differences in N-terminal of *Rep* affecting the symptoms appearance rates of Geminivirus infection.

**Keywords:** Binding site, Geminivirus, Pathogenicity, PepYLCV, Replicase.

**Abbreviations:** dsDNA\_Double stranded DNA; PepYLCV\_Pepper yellow leaf curl virus; PSSWS\_Pesisir Selatan West Sumatra; *Rep*\_Replicase; ssDNA\_Single stranded DNA; TDWS\_Tanah Datar West Sumatra.

### Introduction

Pepper Yellow Leaf Curl Virus (PepYLCV) belongs to the family of Geminiviridae is a major pepper plant pathogen. PepYLCV occurs as single-stranded DNA (ssDNA) and packaged into virions (Bisaro, 1996). The virus genomes consist of one (monopartite) or two (bipartite) DNA components that encode 5–7 proteins involved in viral replication, movement, transmission, and pathogenesis (Jamsari and Pedri, 2013). In the plant cell, the ssDNA of the virus is released from the virion and converted into double-stranded DNA by using plant DNA polymerases (Gutierrez et al., 2004). Short after, the dsDNA is transcribed by host RNA polymerase II to make Replicase protein (also known as C1 protein). This protein is responsible for viral replication via rolling-circle replication and recombination-dependent replication. Later, viral movement proteins, nuclear shuttle protein (NSP) (Carvalho and Lazarowitz, 2004) to transport viral DNA into other cells, causing widespread infection in a plant. A recent study showed transgenic *Nicotiana benthamiana* expressing the virus C2 protein (transcriptional activator) interfere with plant CMT3-2 gene expression (Tu et al., 2017). Thus make the transgenic plant more susceptible to geminivirus infection. Another viral protein, N terminal domain of TYLC Rep protein was responsible for the host transcriptome profile alteration and may also alter the level of infection symptoms (Lucioli et al., 2014).

The level of infection symptom is affected by many factors, include virus accessibility, replication, response to host

defenses, and rate of spreading within the cell (Mandadi and Scholthof, 2013). Among those factors, replication is interesting to evaluate. In the previous study, Jamsari and Pedri (2013) have identified two different PepYLCV isolates in West Sumatera, designated as PepYLCtdV.TDWS21 (TDWS21) and PepYLCpsV.PSSWS14 (PSSWS14). Artificial infection of TDWS21 via whitefly showed apparent symptoms after 20 days post-inoculation (dpi) (Jamsari et al., 2015). Meanwhile, in PSSWS14, the symptoms were observed after 8 dpi, 2-fold faster than TDWS21. The difference in symptoms rate is expected to be related to *Rep* protein. The genomic comparison of two isolates showed high conserved residues in almost all regions, but not in *Rep* residues (Jamsari et al., 2016). In this article, *Rep* gene from infected chili pepper plant designated as PepYLCpsV.PSSWS3 was isolated and compared with PSSWS14 and TDWS21 for pathogenicity determinant analysis.

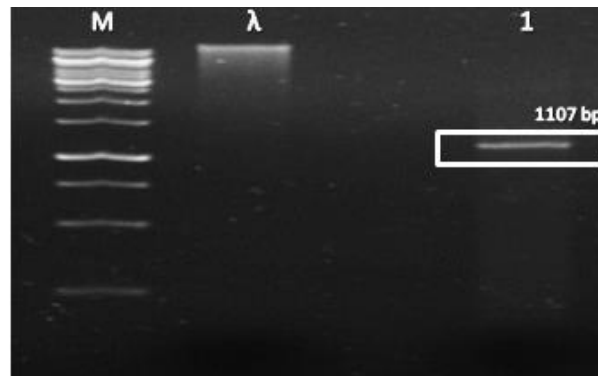
### Results

#### *Rep* gene isolation

The *Rep* gene was successfully isolated and amplified about 1107 bp PCR product as estimated (Fig 1). The primer used to amplify the *Rep* gene of Pesisir Selatan (PSSWS3) was also used to amplify *Rep* gene from Tanah Datar (TDWS21).

**Table 1.** The amino acid sequence similarity with N-terminal replication initiation protein of TYLCV Sardinia.

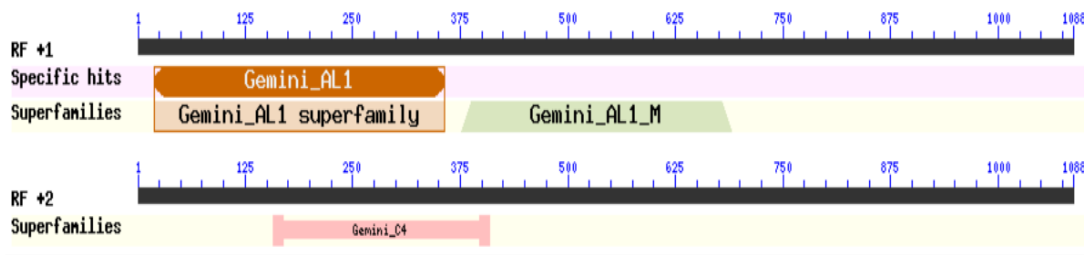
Protein	Similarity (%)
PSSWS3	77.59
PSSWS14	61.02
TDWS21	78.45



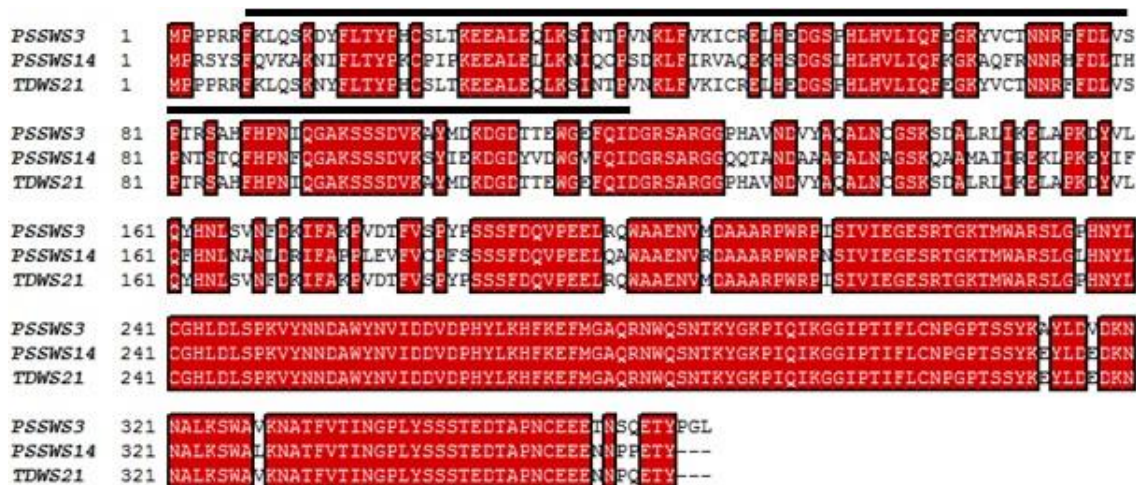
**Fig 1.** Amplification of *Rep* gene. 1107 bp of PCR product was obtained from pepper genomic DNA template by using specific primers. M: 1 Kb DNA ladder; λ: DNA lambda (50 ng/μL); 1: *Rep* gene.

**Table 2.** Evaluation of homology modeling for protein tertiary structure using ProFunc.

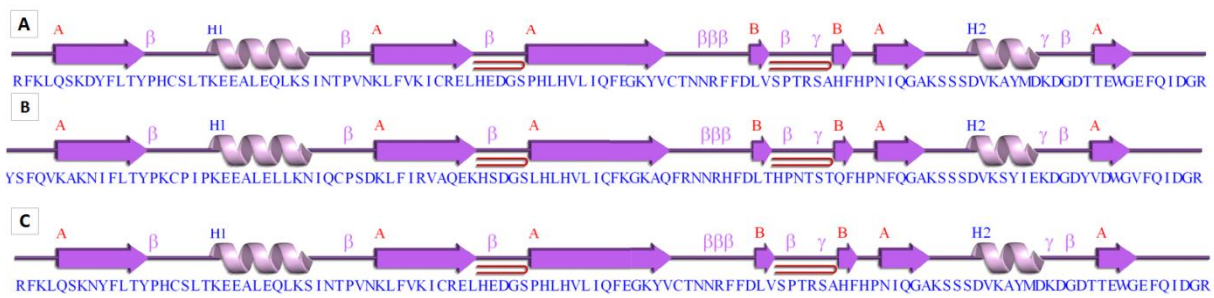
Model	Ramachandran Plot (%)				G-Factor	Z-Score	RMSD (Å)
	Most favourable	Additional allowed	Generously allowed	Disallowed			
PSSWS3	86.40	11.70	1.90	0.00	-0.31	12.60	0.10
PSSWS14	86.50	11.50	1.90	0.00	-0.24	12.60	0.13
TDWS21	86.40	11.70	1.90	0.00	-0.31	11.60	0.10



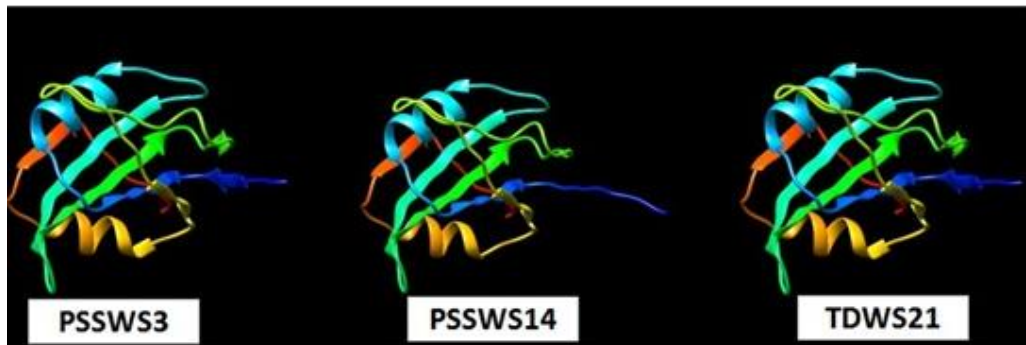
**Fig 2.** CD search analysis result of C1 (*Rep*) gene sequence. AL1 catalytic domain was observed at N-terminal.



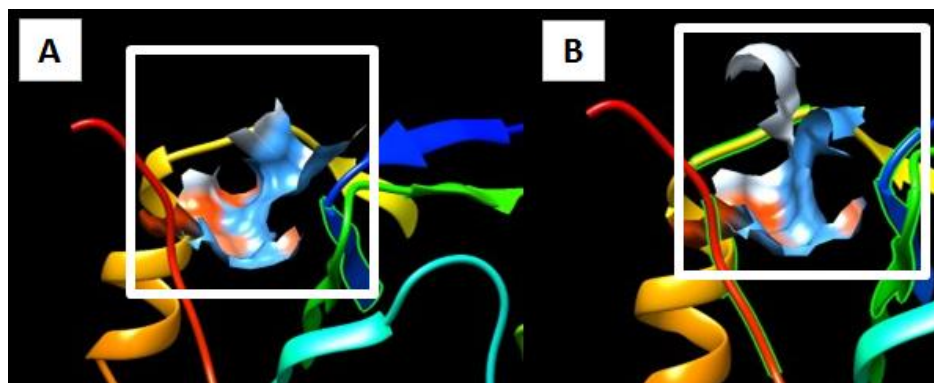
**Fig 3.** The full-length amino acid and alignment sequence of three PepTYLCV strains. The conserved residues are indicated by red fill box. The catalytic domain is represented by a black line.



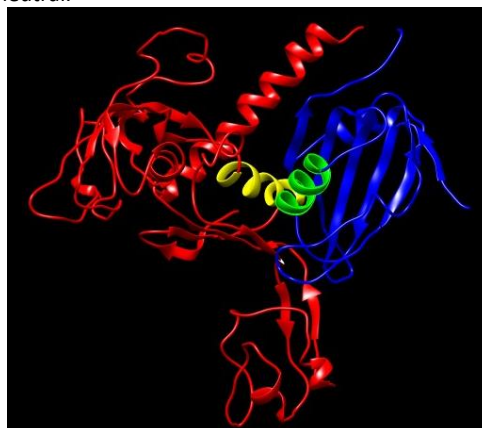
**Fig 4.** The secondary structure alignment of PepYLCV strains. (A) PSSWS3, (B) PSSWS14, (C) TDWS21. H: alpha helix;  $\beta$ : beta sheet.



**Fig 5.** Modelling of three PYLCV strains. N-terminal shown in blue ribbon. C-terminal shown in red ribbon.



**Fig 6.** The *Rep* binding site prediction using CASTp. (A) PSSWS3, (B) PSSWS14. Red colour shows hydrophobic area. Blue colour shows hydrophilic area. White colour is neutral.



**Fig 7.** The *Rep* protein binding site prediction of PSSWS3 strain with DNA-directed RNA polymerase II. The *Rep* protein is indicated by red color. DNA-directed RNA polymerase II is indicated in blue color. The predicted site is represented by yellow color for DNA-directed RNA polymerase II and green color for *Rep* protein.

### Sequencing analysis

The BLAST showed that PSSWS3 has a high sequence similarity (99%) to *Rep* gene from TDWS21 (KT809346.1) with query coverage is 98%. This data showed that despite the difference of location, our PSSWS3 has high homology to *Rep* gene from Tanah Datar.

The full length of PSSWS3 *Rep* residues was 364 aa containing a Geminivirus *Rep* catalytic domain begin from residue 7 to 119 (Fig 2). Multiple alignments of three isolates showed conserved in all region but not in N terminal (Fig 3). The percentage of the amino acid sequence similarities of PepYLCV strains with N-terminal Replication initiation protein of Tomato yellow leaf curl virus-Sardinia (TYLCV) are shown in Table 1. Rep of TYLCV was used as a comparison because its tertiary structure has been and verified by X-ray crystallography known (Campos-Olivas et al., 2002). Also, its biological functions have determined experimentally (Lucioli et al., 2014).

### Bioinformatics analysis

The secondary structure showed similar protein motifs for all three isolates (Fig 4) Further, the predicted tertiary structure of *Rep* protein also shown similar structure among three isolates (Fig 5). The ProFunc program was used to evaluate the model's reliability and model's evaluation scores are shown in Table 2. Since the PSSWS3 and TDWS21 have high similarity amino acid sequence and tertiary structure similarity, only PSSWS3, and PSSWS14 were used for further analysis.

### Rep binding site prediction

The CASTp program was used to predict Rep protein binding pocket of PepYLCV strains (Fig 6). PSSWS3 and PSSWS14 shown the difference in pocket shape. ProBis program predicts that PSSWS3 provided the binding site for the DNA-directed RNA polymerase II (PDB id. 2PMZ) (Fig 7). In contrast, DNA-directed RNA polymerase II was excluded from the PSSWS14 binding prediction list.

### Discussion

Geminivirus contains viral components with multifunctional activities. *Rep* protein is responsible not only for replication steps but also interacting with other host plant components and occasionally causing severe symptoms (Arguello-astorga et al., 2004).

Our previous study showed that different strains show different symptoms rate. This is predicted to be associated with *Rep* activity. Thus, genomic analysis of *Rep* promotes novel and alternative insight to Geminivirus pathogenicity determinant.

In experimental result, PSSWS14 was observed to be more aggressive than TDWS21. Our new isolate PSSWS3 was obtained from the same location of PSSWS14. Interestingly, catalytic domain alignment between PSSWS3 and PSSWS14 showed significant differences in amino acid residues. In contrast, PSSWS3 shared high identity with TDWS21 which obtained from a different location. The similar structure between PSSWS3 and TDWS21 suggest both of isolates would have similar *Rep* activity.

PSSWS3 and PSSWS14 were used for further analysis. Secondary and tertiary structure analysis showed no difference in domain and structure. The CASTp program was able to show the difference in predicted pocket site. The difference would have affected the binding affinity. Both of strains show pocket sites in N-terminal region suggest this terminal is important for biological activity.

Further analysis with ProBis program showed PSSWS3 binding site with RNA-directed binding RNA polymerase II. This binding support the idea PSSWS3 is able to take over the transcriptional machine for expression other viral proteins. In contrast, PSSWS14 have no predicted RNA polymerase II binding site. The difference may explain the difference in severity level of infection.

### Materials and Methods

#### Plant and virus material

About 60 days old *C. annum* cv Kopay plant showing PepYLCV infection, grown in Pesisir Selatan was used in this study. The infected leaf samples were washed with sterilized distilled water and cut into small pieces. Detail description of plant and virus material used in this study is described in Jamsari et al., (2016).

#### DNA isolation

Genomic DNA was extracted from 100 mg fresh infected pepper leaves using the CTAB extraction method described by Delaporta et al., (1983). DNA concentration was assessed by electrophoresis on 2% agarose gel.

#### Amplification of the replicase gene

A pair of specific primer was designed with C1-TD21*Bam*HI-NT as forward primer (5'- CATGGGATCCATGCC-TCCACCACGT -3') and C1-TD21*Sma*I-NT as reverse primer (5'- CTAATCCCGGGTACGTCTCCTGCGA-3'). PCR was done by using genomic DNA as templates. Total PCR mixture of 25  $\mu$ L contained 3  $\mu$ L of genomic DNA (125 ng/ $\mu$ L), 3  $\mu$ L primer C1-TD21*Bam*HI/*Sma*I-NT (5 pmol/ $\mu$ L), 12.5  $\mu$ L of KAPA2G, and PCR-grade H<sub>2</sub>O. The thermal program was set up as follows: 35 cycles composing of 94  $^{\circ}$ C for 60 s, 60  $^{\circ}$ C for 2 min, 72  $^{\circ}$ C for 2 min. Initial denaturation was done at 94  $^{\circ}$ C for 5 min and followed by a final extension at 72  $^{\circ}$ C for 10 min. The amplified product was assessed by electrophoresis on 2% agarose gel.

#### Molecular cloning and sequencing

The amplified PCR products were purified using Promega Purification Kit (Promega, USA), cloned into pGEM-T Easy Vector (Promega, USA). Purified PCR products were transformed into *Escherichia coli* (DH5 $\alpha$ ) and screened by using white-blue colonies. The white colonies were assessed by colony PCR using specific primers. The positive transformant was verified by direct sequencing.

#### Bioinformatics analysis

The sequences were used in BLAST search against GenBank database for similarity analysis. The CD-Search was used for domain analysis (<https://www.ncbi.nlm.nih.gov/>

[Structure/cdd/](#)). The amino acid sequences and alignments were generated using BioEdit version 7.2.5. Homology modeling of Rep was constructed using SWISS-MODEL web-server (<http://swissmodel.expasy.org/>) by using N-terminal Replication initiation protein of Tomato yellow leaf curl virus-Sardinia (PDB id: 1I2M) as a modeling template. All models were confirmed by using Ramachandran plot in ProFunc (<https://www.ebi.ac.uk/thornton-srv/databases/profunc/>) and visualized by using UCSF Chimera 1.11.2 software. The binding site prediction was performed by using CASTp web server (<http://sts.bioe.uic.edu/castp/>).

## Conclusion

Our experimental and bioinformatics analysis of PepYLCV Rep suggests that the differences of amino acid sequences in N-terminal region are prospective to be a pathogenicity determinant of Geminivirus infection.

## Acknowledgement

This research was fully funded by Ministry of Research, Technology and Higher Education through Professorship Cluster Research Grants, the fiscal year 2017. Contract number: 43/UN.16.17/PP.HGBL/LPPM 2017.

## References

- Arguello-Astorga G, Lopez-Ochoa L, Kong L, Orozco BM, Settlege SB, Hanley-Bowdoin L (2004) A novel motif in geminivirus replication proteins interacts with the plant retinoblastoma-related protein a novel motif in geminivirus replication proteins interacts with the plant retinoblastoma-related protein. *J Virol.* 78:4817–4826.
- Bisaro DM (1996) Geminivirus DNA replication. In: DePamphilis ML (ed) DNA replication in eukaryotic cells. Cold Spring Harbor, New York.
- Campos-Olivas R, Louis JM, Clerot D, Gronenborn B, Gronenborn AM (2002) The structure of a replication initiator unites diverse aspects of nucleic acid metabolism. *Proc Natl Acad Sci USA.* 99:10310–10315.
- Carvalho MF, Lazarowitz SG (2004) Interaction of the movement protein NSP and the arabidopsis acetyltransferase AtNSI is necessary for cabbage leaf curl geminivirus infection and pathogenicity. *J Virol.* 78:11161–11171.
- Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA mini preparation: version II. *Plant Mol Biol Rep.* 1:19-21.
- Gutierrez C, Ramirez-Parra E, Mar Castellano M, Sanz-Burgos AP, Luque A, Missich R (2004) Geminivirus DNA replication and cell cycle interactions. *Vet Microbiol.* 98:111–119.
- Jamsari J, Ferita I, Noverta A, Husada ED, Herberg FW, Nellen W, Syukriani L (2016) A pathogenic isolate of monopartite PePYLCV DNA a-like genome differs significantly in C1 gene and CR sequence, but not in their other genes. *Plant Pathology J.* 15:124–134.
- Jamsari J, Syukriani L, Utami HP, Herberg F, Nellen W, Ferita, I (2015) Injection technique could as a new promising method for artificial infection of geminivirus particles in chili pepper. *Asian J Agr Res.* 9:23–32.
- Jamsari J, Pedri, J (2013) Complete nucleotide sequence of DNA A-like genome and DNA-β of monopartite pepper yellow leaf curl virus, a dominant begomovirus infecting *Capsicum annum* in West Sumatera Indonesia. *Asian J Plant Pathol.* 7:1–14.
- Lucioli A, Berardi A, Gatti F, Tavazza R, Pizzichini D, Tavazza M (2014) Tomato yellow leaf curl Sardinia virus-resistant tomato plants expressing the multifunctional N-terminal domain of the replication-associated protein show transcriptional changes resembling stress-related responses. *Mol Plant Pathol.* 15:31–43.
- Mandadi KK, Scholthof KG (2013) Plant immune responses against viruses: how does a virus cause disease? *Plant Cell.* 25:1489–1505.
- Tu YC, Tsai WS, Wei JY, Chang KY, Tien CC, Hsiao HY, Fu SF (2017) The C2 protein of tomato leaf curl Taiwan virus is a pathogenicity determinant that interferes with expression of host genes encoding chromomethylases. *Physiol Plant.* 161:515–531.