Eco. Env. & Cons. 27 (Special Issue of the 5 IBOC on Sustainability of Natural Resources, December, 2021) : pp. (S84-S89)

ISSN 0971–765X

Characterization of Pathogenesis-Related Transcriptional Activator Genes (Pti5 and Pti6) from *Musa acuminata* (A Genome) and *Musa balbisiana* (B Genome) Using in Silico Approach

Indah Nurharuni, Fenny M. Dwivany* and Karlia Meitha

School of Life Science and Technology, Institut Teknologi Bandung, Jalan Ganesha 10, Bandung 30132, West Java, Indonesia

(Received 18 June, 2021; Accepted 22 Augut, 2021)

ABSTRACT

Bananas plant are perennial monocotyledonous herbs which domestication process started some in Southeast Asia. Most edible bananas belong to the Eumusa (or Musa) section, the majority of these bananas originated from hybridization between *M. acuminata* and *M. balbisiana* species. Pathogenesis-Related Genes Transcriptional Activator (PTI5 and PTI6) are proteins that have function in disease resistant in banana plants. With the availability of *Musa acuminata, Musa balbisiana* genome data, this study aims to identify and compare PTI5 and PTI6 genes in B genome with the annotated model in A genome. The PTI5 and PTI6 genes in A genome were obtained from the National Center of Biotechnology Information (NCBI) and genome B from the analysis of the Basic Local Alignment Search Tool Nuclotide (BLASTN) at Banana Genome Hub. The gene structure, location of exons and introns was predicted using FGENESH+ program. Amino acid sequences were identified for their domains and motifs using the CD-Search Tool and MEMEsuite program. Phylogenetic tree was constructed using protein sequences using MEGAX. The nucleotide sequences and protein sequences *Musa acuminata* and *Musa balbisiana* showed similar nucleotide compositions and gene structures with percentage identity 97% for PTI5 and 89% for PTI6. Protein sequences have conserved domain namely APATELA2/Etylene-Responsive Factor (AP2/ERF) with YRG and RAYD motifs which have function as DNA-binding transcription factor fo Pathogenesis-Related genes. The result of this study can be used as a prelimary data for future studies of disease resistant related genes in *Musa acuminata* and *Musa balbisiana*.

***Key words**: M. acuminata, M. balbisiana, PTI5, PTI6*

Introduction

Banana is one of most important plants in the world as a food crop in Southeast Asia and Indonesia. The annual consumption of bananas in the world reaches more than 100 million tons and has the 4th highest production value after wheat, rice and maize (Nuryanti, 2014). Indonesia ranks the 6th largest banana producer in the world after India, China, the Philippines, Ecuador and Brazil. Banana have various types of cultivars derived from two types of parent genomes, namely *Musa acuminata* (A genome) and *Musa balbisiana* (B genome). The results of hybridization or cross-breeding, a combina-

tion of the two genomes can be produced. Banana type *Musa acuminata* has fruit characters with low starch content, faster fruit ripening time, does not contain seeds and the fruit can be consumed directly. *Musa acuminata* banana is classified as having a weak immune system and is susceptible to pests or diseases, one of which is fusarium wilt (Kumari *et al.*, 2015). Klutuk Wulung (*Musa balbisiana*) is a type of banana plant belonging to the Musaceae family and can grow in the wild (Borborah *et al.*, 2016). This type of banana has thick waxy coating, contain seeds on the fruit, has hard and thick fruit skin and must be processed first to be eaten. *Musa balbisiana* is known to be more resistant to diseases and pests such as *Fusarium*, Black Sigatoka, and Banana Blood Disease compared to other types of bananas (Suryanto *et al.*, 2010). One of the threats of pests and diseases that attack bananas is *Fusarium* wilt. *Fusarium* wilt generally attacks varieties of banana Kepok, Raja and Siem which are species A. Symptoms of *Fusarium* wilt will be seen on the roots, stems and leaves. This disease is caused by *Fusarium oxysporum* which is a pathogen that spreads through the soil and then infects through the roots of banana plants. Pathogens infect plant roots and then spread through plant vessels. Symptoms that appear are a change in leaf color, pseudostem turning into yellow and wilting. Banana plants use a variety of strategies to respond to pathogenic infections. The defense response can be carried out by a single gene or a group of genes that work in coordination to modulate a specific defense response through the signal transduction cascade pathway and transcription activation of multiple genes (Zhang *et al.*, 2019). Pathogenesis-Related Transcriptional Activator (PTI5 & PTI6) is an example of a transcription factor that plays a role in plant resistance to pathogens. The PTI5 and PTI6 genes which are included in the AP2/ERF superfamily have a role in activating Pathogenesis-Related genes that produce proteins which has function in direct defense responses against pathogens. Research related to the PTI5 and PTI6 genes in *Musa acuminata* has been widely carried out, but the annotation of the PTI5 and PTI6 genes in *Musa balbisiana* has not available in database. The annotation and characterization of the PTI5 & PTI6 genes of *Musa balbisiana* (B genome) need to be carried out using information of annotated PTI5 & PTI6 genes of *Musa acuminata* which available in GenBank. This research aim on the characterization of PTI5 and PTI6 genes as it can help complete useful data for the development of disease resistant banana plant breeding, especially *Fusarium* wilt.

Materials and Methodolgy

The annotated PTI5 and PTI6 genes of A genome, Musa acuminata (DH-Pahang) was retrieved from the National Center for Biotechnology Information (NCBI) website (www.ncbi.nlm.nih.gov). The PTI5 and PTI6 genome B, *Musa balbisiana* (PKW) genes were obtained from the results of the Basic Local Alignmet Search Tool (BLAST) on the Banana Genome Hub website (https://banana-genome-hub.southgreen.fr/blast) using the PTI5 and PTI6 of Musa acuminata as query to retrieved the PTI5 and PTI6 genes of B genome using the BLASTN, the sequences retirieved from the highest hit, with a cutoff of 10-10. Similarity analysis was performed using global pairwise alignment using EMBOSS-Needle on the European Bioinformatics Institute (EMBL-EBI) website (https: www.ebi.ac.uk/Tools/psa / emboss_needle /) to get result of similarity percentage, identity percentage and gap. Gene structure were predict using the FGENESH+ program on Softberry website (www.softberry.com) using the nucleotide sequences of PTI5 and PTI6 Musa balbisiana by homology based approach. The prediction including the length of the coding sequence, number of exon and intron, and the position of the exon-introns in the gene. The gene structure visualized using Gene Structure Display Server (GSDS) 2.0 program (http://gsds.gao-lab.org/index.php) by submitting the coding sequence data obtained from the FGENESH+. Protein domain prediction was performed using Conserved Domain Search Tools NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Visualization of protein domains was performed using Illustrator for Biological Sequences (IBS 1.0) software. Motif prediction is performed using MEME-Suite program (https://meme-suite.org/tools/meme). The motif function were identified using InterProScan (https://www.ebi.ac.uk/interpro/search/sequence-search). The upstream sequences (-2000 bp) of PTI5 and PTI6 genes were retrieved from the whole genome of *Musa acuminata* and *Musa balbisiana*. The CAREs analysis was determined using PlantCARE website. Regulatory elements are presented in a matrix of positions, consensus sequences, and individual sites in a specific order on the promoter. The results of the search for cis-acting regulatory elements are

then re-analyzed by grouping the identified motives based on their function.

Results and Discussion

The whole genome sequence *Musa acuminata* cultivar DH-Pahang is available at the GenBank National Center for Biotechnology Information (NCBI) and has been annotated 90% of the total genome in DH-Pahang. *Musa acuminata* (DH-Pahang) has 11 pairs of chromosomes of different sizes. The annotated PTI5 (103978384) and PTI6 (103992507) gene sequences in *Musa acuminata* were obtained from the National Center for Biotechnology Information (NCBI). Information of PTI5 and PTI6 genes is shown in Table 1.

Table 1. PTI5 and PTI6 *Musa acuminata* from geneBank

Genes	Gene ID	Length (bp)	Protein (aa)	Chromosome
PTI5-like	*103978384*	976	183	3
PTI6-like	*103992507*	1703	247	7

The PTI5 (MaPTI5) and PTI6 (MaPTI6) of *Musa acuminata* DH-Pahang genes (A genome) were used to get PTI5 and PTI6 of *Musa balbisiana* genes on the Banana Genome Hub website (https://banana-genome-hub.southgreen.fr/) which provides Whole Genome Sequence of *Musa balbisiana* variety "Klutuk Wulung" (B genome). The result of PTI5 (MbPTI5) and PTI6 (MbPTI6) genes is shown in Table 2.

The PTI5 *Musa acuminata* (MaPTI5) gene which is located on chromosome 3 has length 976 bp with Coding Sequence (CDS) length 552 bp The PTI5 *Musa balbisiana* (MbPTI5) gene from BLASTN has length 976 bp with Coding Sequence (CDS) length 552 bp. The visualization of MaPTI5 and MbPTI5 gene structures were shown in Figure 1.

The MaPTI5 gene has 1 exon which position is 230-782 bp, for MbPTI5 gene also has 1 exon which is located in the sequence 230-782 bp. Both genes have no intron structures, but only exons. The majority of genes in the AP2/ERF superfamily have only 1 intron or no intron at all (Nakano *et al.*, 2006). Identification in the *Arabidopsis thaliana* plant also showed that the PTI5 and PTI6 gene has only 1 intron. The result of Pairwise Sequence Alignment between MaPTI5 and MbPTI5 which shows a comparison of the level of nucleotide similarity between the two genomes, shows the percentage of similarity 97%. The comparison of nucleotide sequences includes the length of the sequence and the composition of the nucleotides. There are several differences in nucleotide sequences in the PTI5 *Musa acuminata* and *Musa balbisiana* genes may caused by the presence of insignificant insertion mutations or deletions. The PTI6 *Musa acuminata* (MaPTI6) gene which is located on chromosome 7 has length 744 bp with Coding Sequence (CDS) located in 733-1476 bp. The PTI5 *Musa balbisiana* (MbPTI5) from BLASTN result has length 1862 with a Coding Sequence (CDS) length 744bp. The results of the visualization of the MaPTI6 and MbPTI6 gene structures are shown in Figure 2.

The MbPTI6 gene has 1 exon with position in 733-1476 bp, for the MbPTI6 has 1 exon which is located in 893-1636 bp. Both MaPTI6 and MbPTI6 genes have 1 exons and have no introns. The comparison of the MaPTI6 and MbPTI6 nucleotide sequences using the Pairwise Sequence Alignment method showed a similarity percentage 89%. The MbPTI6 gene is longer than MaPTI6. The percentage of the similarity of the PTI6 gene between these two denominations was lower than the comparison be-

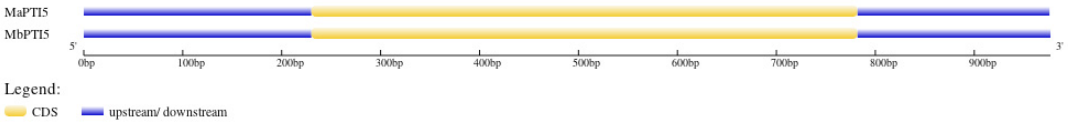

Fig. 1. Structure of the PTI5 *Musa acuminata* (MaPTI5) and *Musa balbisiana* (MbPTI5)

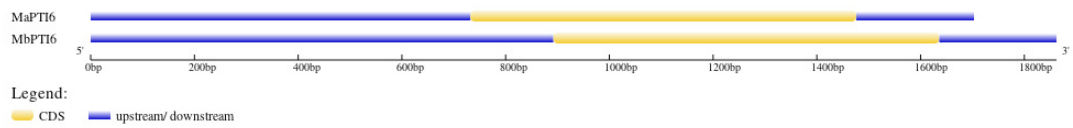

Fig. 2. Structure of PTI5 *Musa acuminata* and *Musa balbisiana*

tween the MaPTI5 and MbPTI5 sequences. The difference in the length of the MaPTI6 and MbPTI6 nucleotides is 159 bp. PTI5 protein of *Musa acuminata* was retrieved from NCBI with accession XP_009392438.1. Putative protein of the PTI5 *Musa balbisiana* gene was taken from prediction using FGENESH+.

The MaPTI5 and MbPTI5 structure of protein domains using IBS 1.0 are shown in Figure 3.

CD-Search Tool were identified the presence of the AP2 / ERF superfamily domains in both protein sequences. The AP2 domain is located in the MaPTI5 protein sequence interval 65-123 of amino acids with the number of amino acid residues 58 aa. In the MbPTI5 protein sequence, the AP2 domain was found at intervals of 62-120 amino acids. PTI6 *Musa acuminata* protein (MaPTI6) obtained from NCBI with accession XP_018684565.1) is composed of 247 amino acids. The putative protein sequence of PTI6 *Musa balbisiana* (MbPTI6) is the same as the MaPTI6 protein consisting of 247 amino acids. The comparison of protein sequences by doing global alignment using EMBOSS-Needle shows an identity percentage of 98% and a similarity percentage of 98%. The protein sequence identity percentage refers to the equation of the length and amino acid composition of the two proteins. The similarity percentage of protein sequences shows the similarity in physiochemical properties of the two proteins. Both MaPTI6 and MbPTI6 proteins contain 1 AP2 domain at the same interval (65-123) with the number of amino acid residues of 58 aa. Proteins belonging to the AP2 / ERF superfamily have one to two domains (He Ping *et al.,* 2001). The AP2 domain is a DNA-binding domain which is one of the domains contained in plant transcription factors. The AP2 domain will specifically bind to the Ethylene Response Element Both MaPTI6 and MbPTI6 proteins contain 1 AP2 domain at the same interval (65-123) with the number of amino acid residues of 58 aa. Proteins belonging to the AP2 / ERF superfamily have one to two domains (He Ping *et al.,* 2001). The AP2 domain is a DNA-binding domain which is one of the domains contained in plant transcription factors. The AP2 domain will specifically bind to the Ethylene Response Element (ERE) or GCC-box on the Pathogenesis-Related gene promoter or PR genes. The PTI5 and PTI6 proteins as trans-acting regulatory elements will bind to the PR gene promoter in the Cis-acting regulatory element (CAREs) region. Pathogen-related genes have a function related to the response to pathogens, namely by producing chitinase and glucanase enzymes to degrade the cell walls of pathogens, including the fungus *Fussarium oxisporum* which causes *Fusarium* wilt disease in banana plants. PR genes that are regulated by transcription factors PTI5 and PTI6 include PR1 which has an anti-fungal role. The PR3 and PR4 genes regulate protein products in the form of chitinase enzymes which play a role in the hydrolysis of the glycosidic bonds in chitin (the constituent of fungal cell walls) (He Ping *et al.,* 2001. Based on the prediction of protein motif using MEME-Suite, it identified 5 conserved motifs. Figure 5 shows the

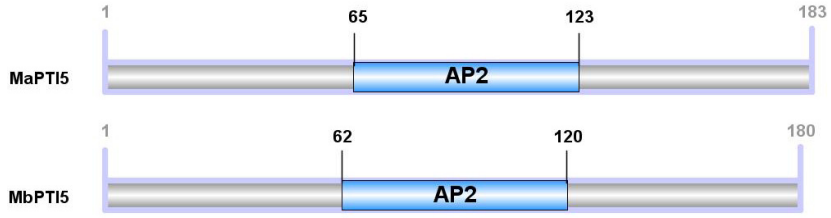

Fig. 3. Domain structure of the PTI5 *Musa acuminata* and *Musa balbisiana* proteins

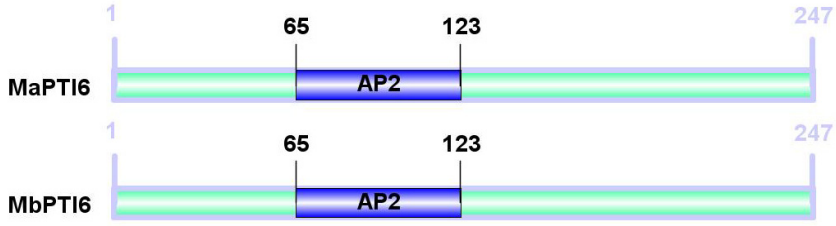

Fig. 4. Structure of PTI6 protein domains *Musa acuminata* and *Musa balbisiana*

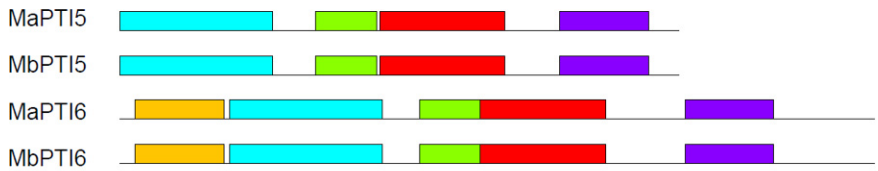

Fig. 5. Motifs sequences of PTI5 and PTI6 proteins *Musa acuminata* and *Musa balbisiana*

motifs that represent the conserved sequences on the protein MaPTI5, MbPTI5, MaPTI6, and MbPTI6.The MaPTI5 and MbPTI5 proteins have protein motives 2,3,1, and 4, respectively, whereas MaPTI6 and MbPTI6 proteins have motives 5,2,3, 1 and 4 respectively.

As for the molecular process motif 1 and motif 3 function as transcription factors that attach to the target DNA to initiate the transcription process. Based on the identification of the gene ontology of PTI5 and PTI6 proteins containing localized DNA-binding motifs in the nucleus. For motifs 2,4, and 5 it could not be predicted for gene ontology because the motive sequence was not found in the database. Motif sequence 1, consisting of 41 amino acid residues and this motif has a sustainable area, namely RAYD in the 20th to 23rd amino acids. The RAYD region has an important role in the structure and function of the domain as a DNA binding area (He Ping *et al.*, 2001). In motif 3 which is 20 amino acids in size, there is a YRG conserved region in the amino acid sequence 2 to 4. The YRG region is located on the N-terminal which is alkaline and hydrophilic, the YRG region will attach directly to the cis-regulatory element region in DNA specific genes. The characteristic possessed by the AP2 / ERF domain is that it has a sustainable area of YRG and RAYD (He Ping *et al.*, 2001). The response mechanism to pathogens in the PTI5 and PTI6 genes is influenced by transcription regulation involving certain transcription factors and cis-acting regulatory elements (CAREs) on specific genes involved in plant defense responses. CAREs consist of short regulatory motifs (5–20 bp) that reside in the promoter region of the gene which is the non-coding region of DNA. The composition of CAREs shown is a regulatory element hormone. The CAREs consist of elements that are responsive to the hormones auxin, gibberellin, ethylene (ET), jasmonic acid (MeJa), and salicylic acid (SA). The most abundant composition is SA regulatory element in both genes and species. This regulatory element is related to the function of the PTI5 and PTI6 genes as transcription factors of the PR protein. The expression of the PTI5 and PTI6 genes is influenced by these hormones and will respond to activation of the target genes. The regulatory region or promoter of the PR gene is known to have motives that are responsive to the hormones SA, ET, and MeJa. The specific region in the PR gene promoter as ethylene responsive is a region called the GCC-box consisting of 11bp (TAAGAGCCGCC) (Okamaru, 1997).

Conclusion

The conclusions of this study are the PTI5 and PTI6 genes in *Musa acuminata* and *Musa balbisiana* have similar gene structures, with a similarity rate of 97% for the PTI5 gene and 89% for the PTI6 gene. All genes have 1 exon in the coding sequence and do not have intron, the dominant cis-acting regulatory element, namely SA responsive. The nucleotide sequences of PTI5 and PTI6 genes in *Musa acuminata* and *Musa balbisiana* are sustainable or conserved. The PTI5 and PTI6 proteins *Musa acuminata* and *Musa balbisiana* all have the AP2/ERF superfamily domain and the sustainable motif areas, namely YRG and RAYD. The AP2 domain functions as a transcription factor that attaches to the Pathogenesis-Related (PR) gene promoter region that plays a role in the response to infection by pathogens that cause *Fusarium* wilt.

Acknowledgments

We gratefully acknowledge to Banana Group, Institut Teknologi Bandung and RISTEK-BRIN for support on this research.

References

Borborah, K., Borthakur, S.K. and Tanti, B. 2016. *Musa balbisiana* Colla-Taxonomy, Traditional Knowledge and Economic Potentialities of The Plant in Assam,

India. *Indian Journal of Traditional Knowledge.* 15(1): 116-120.

He, Ping., Randall F. Warren and Tiehan Zhao, 2001. Overexpression of Pti5 In Tomato Potentiates Pathogen-Induced Defense Gene Expression and Enhances Disease Resistance to *Pseudomonas syringae* pv. tomato. *The American Phytopathological Biosociety,* 14(12) : 1453-1457.

Kumari, Anita and Harsh Kumar, 2015. Banana Genome Sequenced – A Review. *Trends in Bioscience.* 8(4) : 977-983.

Nakano, Toshitsugu., Kaoru Suzuki, Tatsuhito Fujimura, and Hideaki Shinsi, 2006. Genome-Wide Analysis of the ERF Gene Family in Arabidopsis and Rice. *Plant Physiology.* 140 : 411-432.

Nuryanti, L. Novianti, 2014. Outlook Komoditi Pisang. Pusat Data and Informasi Pertanian, Sekretariat Jenderal Kementrian Pertanian.

Okamuro, J.K., Caster, B., Villaroel R., Van Montagu, M. and Jofuku, K.D. 1997. The AP2 domain of APATELA2 defines a large new family of DNA binding proteins in Arabidopsis. *Proc. National Science USA*. 94 : 7076-7081.

Suryanto, Widada, Agus, 2010. Hama dan Penyakit Tanaman. Yogyakarta : Kanisius

Zhang, Lu., Dengbo Zhou, Huigang Hu, Weiming Li, Yulin Hu, Jianghui Xie, Shangzhi Huang and Wei Wang, 2019. Genome-wide characterization of a SRO gene family involved in response to biotic and abiotic stress in banana (*Musa* spp.). *BMC Plant Biology.* 19 : 211.