Eco. Env. & Cons. 27 (Special Issue of the 5 IBOC on Sustainability of Natural Resources, December, 2021) : pp. (S49-S58)
Copyright@ EM International
ISSN 0971–765X

Comparative Genomics Analysis of the Pathogenesis-related (PR) Genes on Two Banana Genomes (A and B Genomes) in Response to Banana Blood Disease Infection

Kenzie Sachisiva Anindita, Popi Septiani and Fenny Martha Dwivany*

Institut Teknologi Bandung, School of Life Science and Technology Bandung, 40132, Indonesia.

(Received 2 May, 2021; Accepted 15 June, 2021)

ABSTRACT

Banana *(Musa* spp.) is one the most important food commodity in many tropical and subtropical countries. In the recent decades, banana production has been severely hindered by several diseases. Banana Blood Disease caused by *Ralstonia syzygii*, has become one of the main threats for banana *(Musa* spp.) production in Indonesia. An alternative to overcome the problem arising from plant disease is to develop a plant which is tolerant to a pathogen infection. Tolerant plant can be developed by utilizing genes which are responsible in defense mechanism to response pathogen infection. Some studies reported that Pathogenesis Related (PR) proteins encoded by *PR* gene family were accumulated in plant tissue during pathogen infection. The presence of these protein indicate its role in self-defense response in some plant model organisms. In this study, we aim to characterize some *PR* gene family including *PR1, PR2, PR3, PR4*, and *PR5* in two different banana, *Musa acuminata* cv. Pahang (AA group) and *Musa balbisiana* cv. Klutuk Wulung (BB group), possessing different genomes (A and B genomes) by comparative genomic analysis. We reported that each corresponding *PR* genes from those species have been predicted to have the same number of exons and introns with high level of similarity ranging from 76.5-96.2% and its protein > 79%. The abundance of Cis-acting regulatory elements whose role as a light-responsive are dominant in each *PR* gene. The protein sequences of each *PR* genes shared the same domain associated to defense mechanism. The result from gene ontology search, showed that each *PR* genes have biological activities and molecular function in defense against invading pathogens. The Phylogenetic analysis showed that *Musa acuminata* and *Musa balbisiana* taxa is closely related to each other and clustered in a big clade with other monocots species. The result of this study suggested that *PR* genes are important to be further investigated as potential markers in developing resistant banana against Blood Disease.

Key words*: Pathogenesis related protein, Musa acuminata, Musa balbisiana, Banana blood disease, Comparative genomics*

Introduction

As primary perennial crops, bananas are the main staple food crops in Indonesia. In general, according to a genome-based system, bananas have been classified into four groups called A, B, S, and T respectively represented by *Musa acuminata* (AA; 2n = 22), *Musa balbisiana* (BB; 2n = 22), *Musa schizocarpa* (SS; 2n = 22) and *Musa australimusa* (TT = 2n = 20) (Pua, 2007). In addition, banana cultivars such as Siem (ABB), Nangka (AAB), and Raja Sereh (AAB) which triploid bananas are the result of polyploidization and hybridization of two diploid species derived from *M. acuminata* and *M. balbisiana.*

Banana production is potentially limited by blood disease which is caused by *Ralstonia syzygii subsp. Celebesensis* (Rsc). The bacteria was first isolated in 1921 by Gäumann and remained confined to Sulawesi-Indonesia (Eden-Green, 1994). The disease spread with a spreading speed of 100 km/year and became endemic to almost all islands in Indonesia (Kusumoto, 2004). Symptoms arising from the infection of Rsc are wilt plants, typical yellow-pigmented and dry leaves, bark sections and appearance of red fungus with brown spots and red mucus like blood that smells rotten on transversely sliced pseudostems (Jeger, 1995). There are several attempts to prevent and treat blood diseases in banana. However, the most effective way is using banana cultivars with genetic resistance to blood disease bacterium (BDB).

In this study, two different banana cultivars with different genomes, *Musa acuminata* cv. Pahang (AA group) representing the A genome and *Musa balbisiana* cv. Klutuk (BB group) representing the B genome were compared to characterize the genes which were associated to defense mechanism from pathogen infection, in particular Rsc infection. According to previous study, we reported that *pathogenesis related* (PR) genes were significantly upregulated during Rsc infection in both banana genomes, confirming the importance of this family genes on defense mechanism in response to Rsc (Rahmawati, 2018). In this study we investigate the characteristics of *PR* genes family consisting of *PR1, PR2, PR3, PR4* and *PR5* genes in A and B genomes by genomic comparative analysis. PR proteins, encoded by *PR* genes are group of molecules induced by pathogens and signalling molecules related to defense response such as salicylic and jasmonic acid. The important role of PR protein in plant innate immune system, especially for systemic acquired resistance (SAR), cause the usage of these protein as diagnostic molecular markers in defense mechanism to pathogen infection (Ali *et al.*, 2018).

The uncharacterized B genome of *Musa balbisiana* is one of the obstacles related to the improvement of genetic quality in bananas. Therefore, in this study, *in-silico* comparative genomic analysis was carried out on *PR1, PR2, PR3, PR4, and PR5* genes from *Musa acuminata* cv. Pahang (A genome) and *Musa balbisiana* cv. Klutuk (B genome). The purpose of this study was to determine and compare the gene and protein characters of *PR1, PR2, PR3, PR4*, and *PR5* in A and B genome including gene structure, nucleotide composition, cis-acting regulatory element (CAREs), gene ontology (GO), motives and domains, and phylogenetic relationships with other species. This research is expected to provide basic information regarding the structural characteristics and function of pathogenesis-related genes that have the potential to develop banana plant which tolerant to biotic stress, particularly to blood disease infection.

Materials and Methodology

Gene sequences of *PR1, PR2, PR3, PR4, and PR5* of Pahang *(Musa acuminata,* A genome) and Klutuk Wulung *(Musa balbisiana,* B genome) were retrieved from Banana Genome Hub (https://banana-genomehub.southgreen.fr) by TBLASTN (Gerts et al., 2006). The dataset used for TBLASTN were protein sequences from *Arabidopsis thaliana* as model organisms. The protein sequences of *Arabidopsis thaliana* were retrieved from The Arabidopsis Information Resource (www.arabidopsis.org). To find the percentage similarity of PR sequences between the two species of banana were performed using EMBOSS-Needle alignment (https://www.ebi.ac.uk/Tools/psa/emboss_needle/).

The intron and exon organization of *PR* genes of Klutuk Wulung (B genome) was predicted using FGENESH+ (http://www.softberry.com/berry.phtml?topic=fgenes_plusandgroup=programsandsubgroup=gfs. In this analysis, we used PR sequence from Pahang (A genome) as reference sequence to predict the structure of Klutuk Wulung (B genome). Illustrator for Biological Sequences (IBS) ver. 1.0 was used to visualize the intron and exon organization of *PR* genes (Liu *et al.*, 2015). The *PR* genes structure were compared with each other between two species of banana used in this research.

The cis-acting regulatory elements (CAREs) in promoter sequences were identified within 1.5 kbp upstream from each *PR* genes of Pahang (A genome) and Klutuk Wulung (B genome) using PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Sequence up to 1.5 kbp upstream at 5'UTR was used because the promoter sequence was found at -2000 up to +200 bp from the transcription start site (TTS) (Yu *et al.*, 2016). The CAREs motifs which are associated to the stress response obtained from each PR gene were counted for the frequency of abundance and compared be-

tween two bananas (A and B genome).

The conserved domain of each PR proteins was identified by CD-Search Tools (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi). The motif in each PR protein sequence were identified using dataset of PR protein sequences from another species which retrieved by BLASTP from NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The protein sequences were filtered from different species with query coverage > 90% and identity > 60%. The motifs were searched through MEME-Suite (http://meme-suite.org/tools/meme). The motif with e-value<0.05 will be checked in Interpro database (https://www.ebi.ac.uk/interpro/search/sequence/). Furthermore, to obtain better information, we searched the gene ontology description of each PR protein from *Musa acuminata* database provided by PLAZA Monocots 4.5 (https://bioinformatics.psb.ugent.be/plaza/versio ns/plaza_v4_5_monocots/) (Van Bel *et al.,* 2018).

A dataset of PR protein sequences from another species which have been retrieved from BLASTP on NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi) with query coverage > 90% and identity > 60%, were used to construct the phylogenetic tree. Multiple sequence alignment were performed using MUSCLE 3.1.1. (Edgar, 2004). Subsequently, the quality of alignment was enhanced using Gblocks(http://phylogeny.lirmm.fr/phylo_cgi/one_task.cgi?task_type=gblocks). Phylogenetic trees for PR proteins were constructed by Mr.Bayes 3.2.7 (Ronquist *et al.,* 2012). Bayesian analysis for each PR proteins were computed for different numbers of generations at the point when the standard deviation of split frequency reaches less than 0.01. TreeView was used to visualize the topology of the tree (Page, 1996).

Results and Discussion

Comparison of *PR* genes sequences between banana A genome (*Musa acuminata* cv. Pahang) and B genome (*Musa balbisiana* cv. Klutuk Wulung) revealed percentage similarity of 55.5%, 71.6%, 80.2%, 52.8%, and 10.9% for *PR1, PR2, PR3, PR4* and *PR5*, respectively. Percentage similarity value for each *PR* genes were higher compared to previous result in which *PR1* is 86.8%, *PR2* is 94.3%, *PR3* is 76,5%, *PR4* is 76,5%, and *PR5* is 96,2%. The percentage similarity provide information about the similarity and conserve region between two sequences (Xiong, 2006). This analysis showed that the *PR* genes from both genomes have high level of similarity.

The predictive structure of *PR* genes from B genome has the same amount of intron and exon as A genome (Figure 1). The distribution of exon and intron will determine the architecture of protein. Changes occurring in gene structure may affect the function or activity of the corresponding protein (Ruvinsky *et al.,* 2007). The similar amount and distribution of exon and intron on A and B genomes of banana indicate the evolution of those two species are closely related to each other and have the same common ancestor which explains the same genus for both species.

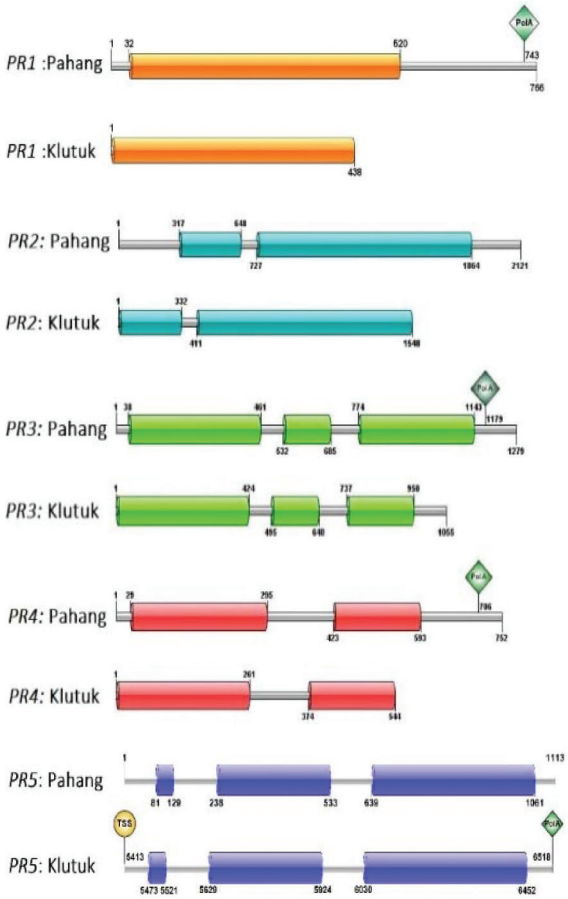

Fig. 1. Predicted structure of PR1, PR2, PR3, PR4, and PR5 genes from A (*Musa acuminata* cv. Pahang) and B genomes (*Musa balbisiana* cv. Klutuk Wulung).

Cis-acting regulatory elements (CAREs) found from each *PR* genes were grouped into different categories based on their function. Regarding defense

response against Blood Disease Bacteria (BDB), we used CAREs motif which are associated with stress response and hormonal regulation (Table 1). The CAREs obtained from *PR* genes were counted for the frequency of abundance and compare to each other between A and B genomes based on their functional category. The result showed that all the *PR* genes from the two genomes have both category in stress and hormonal response which were associated with defense mechanism in response to pathogen infection (Figure 2).

Light-responsive motifs (chsCMA2a, G-box, Gap-box, 3-AF1, GT1-motif, TCCC-motif, ATCT-motif, TCT-motif, GATA-motif, I-box, and Box4) were presented in *PR1, PR2, PR3, PR4,* and *PR5* in A and B genomes. The number of light-responsive motifs were higher than other CAREs element, except for *PR5* which was the second-highest after drought, salt, and abscisic acid responsive element. This result indicates light-responsive motifs are dominant on each *PR* genes. Light is the major factor affecting many biological processes on plants, including defense response to biotic stress. Following the pathogen infection, there will be a change in energy and signaling pathway which is affected by alteration in the photosynthetic energy pathway (Kangasjärvi *et al.,* 2012).

Wound motifs (WUN-motif, WRE3, and W-box) were observed in PR1, PR2, PR3, and PR4 in A and B genomes. W-box as one of the wound-motif are capable to interact with WRKY70 transcription factor. This interaction will regulate the expression of genes that depend on the amount of salicylic acid and jasmonic acid products (Ciolkowski *et al.,* 2008; Kaur *et al.,* 2017; Pandey and Somssich, 2009). However, wound motifs were not present in PR5. Generally, the expression of gene regulated by WRKY occurs after binding to W-box, but some of WRKYs recognize CAREs without W-box motif (Chen *et al.,* 2019). Based on this analysis, we hypothesized there is a probability that the expression of PR5 is activated without interaction between WRKY and W-box. This result indicates that there are some differences in the biological process and activity tailored by PR5 protein from A and B genomes.

Motif involved in hormonal regulation such as salicylic acid (TCA) and methyl-jasmonic (TGACG and CGTCA) responsive element were present in *PR1, PR4,* and *PR5* in A and B genomes. The presence of these motifs is important, considering salicylic acid and jasmonic acid are key signaling molecules that determine the expression of *PR* genes. A methyl-jasmonic (MeJA) and salicylic acid (SA) responsive element has a role in improving the ability to induce *PR* genes (Kim *et al.,* 1993). The presence of SA and MeJA responsive element in *PR* genes reveals that SA and MeJA products can interact with other motifs (Caarls *et al.,* 2015).

The value of PR1 protein percentage similarity between A genome and B was 82.8% and the percentage identity was 85.9%. Comparative analysis of domain architecture for PR1 between A genome and B reveals the presence of cysteine-rich secretory protein, antigen 5, dan pathogenesis related-1 (CAP/

Table 1. The CAREs motif classification on stress response and hormonal regulation category based on their function.

Category	Function	CAREs motif
Stress response	Light responsive element	G-box, Gap-box, chs-CMA2a, 3AF1 binding site, TCCC-motif, ATCT-motif, TCT-motif, GATA-motif, I-box, Box4, AE box, MRE.
	Defense and stress	DRE motif
	Participate in drought, low temperature, salt, and ABA	MBS, MYB. MYC
	Anaerobic induction and anoxic specific inducibility	ARE, GC-motif
	Wound responsive element	WRE3, WUN-motif, W-box
	Activation sequence-1	As-1
Hormonal response	Gibberellin-responsive element	p-box, GARE-motif, TATC-box
	Abscisic acid response	ABRE3a, ABRE4, ABRE
	MeJA-responsiveness	GCTCA-motif.TGACG-motif, CGTCA-motif
	Salicylic acid responsiveness	TCA-motif
	Auxin responsiveness	AuxRR-core
	Elicitor responsive element	ERE-motif

PR1) domain. The structure of CAP domain is a distinct characteristic for PR1 protein (Breen *et al.,* 2017). Identification of motif protein in PR1 shows only two motif consensus are reported in Interpro database as shown in Table 2.

The identification for *PR1* gene ontology was conducted to obtain better information regarding PR1 function in defense response against biotic stress, especially from Blood Disease Bacteria. The result show PR1 has a role related to the biological process which is the response to injury (GO:0009611), response to salicylic acid (GO:0009751), jasmonic acid (GO:0009751), and response to bacteria (GO:0042742). Related to the cellular component, PR1 present in the extracellular area (GO:0005576) and cell wall (GO:0005618). PR1 protein is secreted and accumulated in extracellular space or apoplastic (Breen *et al.,* 2017). The mechanism of PR1 against bacteria is through interaction between the positive charge of amino acid residues of PR1 with the negative charge of phospholipid in bacterial cell membranes. This interaction can inhibit the growth and colonization of bacteria (Gamir *et al.,* 2017; Pecenkova *et al.,* 2017).

The domain architecture analysis for PR2 or β-1,3-glucanases (β-1,3-Gs) in A genome and B reveals the presence of domain X8 and glycosyl hydrolase 17 (GH17). The percentage similarity of PR2 protein between A genome and B is 98.8% and the percentage identity is 97.3%. Domain X8 and GH17 are distinct characteristics of PR2 protein (Doxey *et al.,* 2007). Identification of motif protein in MEME-Suites shows only three motif consensus found in PR2 are reported in Interpro database as shown in Table 3.

The result of gene ontology gave us information regarding PR2 response against bacteria. The PR2 protein is involved in carbohydrate metabolism (GO:0005975) as a biological process and has a role in hydrolase activity (GO:0004553) as a molecular function. The PR2 protein present in cell membrane as an anchored component (GO:0031225). PR2 mechanism against bacteria is through the breakdown of glycosidic bond which involved in the formation of bacterial cell wall. This process mediates by X8 domain that consists of GPI-anchored which are capable to recognize and bind with carbohydrates. The interaction between GPI-anchored and carbohydrate facilitates the hydrolyze β-1,3-Gs by GH17. This cause bacterial cell wall to become unstable and initiate production of signaling molecule which induces self-defense response in plant (Calderan-Rodrigues *et al.,* 2018; Doxey *et al.,* 2007).

The percentage similarity of PR3 between A genome and B is 80.7% and the percentage identity is 79,4%. The analysis of PR3 domain in A genome and B reveals the presence of chitin-binding domain type 1 (CHTBD1) and glycosyl hydrolase 19 (GH19). The domain CHTBD1 and GH19 are distinct characteristics for chitinase I in plants (Ferreira *et al.,* 2007; Landim *et al.,* 2017). Chitin is polysaccharides of β-(1,4)-linked N-acetyl-D-glucosamine (GlcNAc) residues which important structural components in cell wall of bacteria, fungi, an exoskeleton of an insect, and many more (Medeiros *et al.,* 2018). Identification of motif protein using MEME-Suites shows only five motif consensus found in PR3 are reported in Interpro database (Table 4).

The results obtained from *PR3* gene ontology search showed descriptions of PR3 protein mechanism against bacterial attack. The PR3 protein has a role in biological processes such as the catabolic process of chitin (GO:0006032), cell wall macromolecules (GO:0016998), carbohydrates metabolism (GO:0005975), and induction of systemic resistance by jasmonic acid and ethylene (GO:0009871). The

Table 2. The motif consensus of PR1 which detected in Inter Pro database

Motif *PR1*	E-value	Sequences
1	1.6e-286	HNTARAAVGVGPVSWDDTVAAYAQNYANQRIGDCQLVHSGGPYGENLF WG
2	3.5e-163	VCGHYTQVVWRDSTTIGCARVKCNNGAIFII

Table 3. The motif consensus of PR2 which detected in Inter Pro database

Motif *PR2*	E-value	Sequences
1	1.9e-277	SINYACSFSDCTSLGYGSSCNHLDLEGNASYAFNMYYQVRNQKAGACDFS
2	4.1e-273	ANFDTLVWALKKAGYPDMPIIVGEVGWPTDGDKNANIEMAKEFNQGLJQH
3	2.1e-240	KIDVYLFSLIDEDAKSIAPGNFERHWGIFEYDGKPKYELDLS

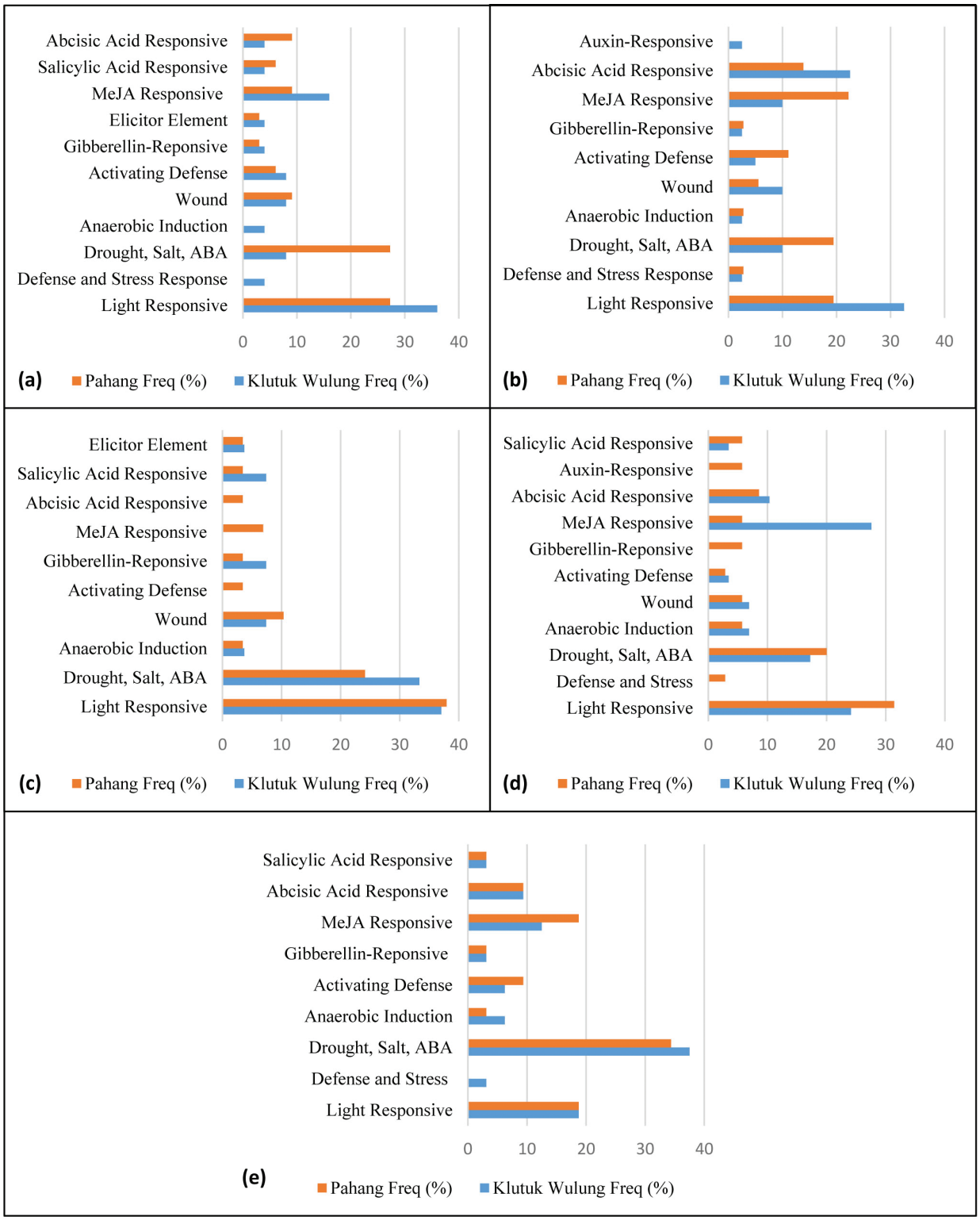

Fig. 2. Frequency of CAREs motif identified on (a) PR1, (b) PR2, (c) PR3, (d) PR4, (e) PR5 from A genome (Pahang) and B genome (Klutuk Wulung).

PR3 has molecular functions as chitinase activity (GO:0005829) and chitin-binding (GO:0005774). The PR3 protein present in the cytosol (GO:0005829) and vacuole membrane (GO:0005774). The antibacterial mechanism of PR3 occurs through lysozyme activity which causes by the catalytic domain of GH19 chitinases that has lysozyme-like fold. Peptidoglycan in bacterial cell wall composed of glucan chains formed between GlcNAc and N-acetyl muramic acid by β-1,4 bonds. GH19 performs hydrolytic action on β-1,4 glucan which causes bacterial cell lysis (Medeiros *et al.,* 2018). The role of CHTBD1 domain in this process is to facilitate attachment of GH19 with chitin (Ubhayasekera, 2011).

The percentage similarity of PR4 between A genome and B is 89% and the percentage identity is 82.9%. The domain analysis of PR4 protein reveals both A genome and B had Barwin domain. The presence of domain Barwin is mostly associated with chitin-binding domain. PR4 is classified into two classes based on the occurrence of chitin-binding domain. PR4 class I has chitin-binding domain and exhibits chitinase activity. PR4 class II only has Barwin domain and exhibits RNase activity. The presence of two conserved histidine residues (at position 11 and 111 relative to CARWIN structure) in Barwin domain has been correlated with RNAse activity (Bertini *et al.,* 2012; Franco *et al.,* 2019). Identification of motif protein in MEME-Suites shows only three motif consensus found in PR3 are reported in Interpro database (Table 5).

The result obtained from gene ontology searching gives information about the antimicrobial activity of PR4. Regarding defense response from pathogen attack, PR4 has a role in defense response to bacterial (GO:0042742) and systemic acquired resistance (GO:0009627) as biological processes. The PR4 exhibits molecular function as ribonuclease (GO:0004540) and chitinase activity (GO:0004568). The PR4 protein present in plant vacuole (GO:0000325). These results showed PR4 protein is important for antimicrobial activity, which hydrolysis phosphodiester bond in RNA as ribonuclease. This activity affects RNA integrity and inhibits colonization of bacteria in plant tissues (Franco *et al.,* 2019).

The percentage similarity of PR5 between A genome and B is 98% and the percentage identity is 96.5%. Our domain analysis of PR5 in A genome and B shows that both have thaumatin-like protein domain (TLP) and glycosyl hydrolase 64 (GH64). GH64 and TLP domain are conserved domain for protein in PR5 family (Salazar and Fernando, 2019). Identification of motif protein in MEME-Suites

Table 4. The motif consensus of PR3 which detected in InterPro database

Motif *PR3*	E-value	Sequences
1	3.3e-349	DLLNNPDLVATDPVISFKTAJWFWMTPQSPKPSCHDVITGRWTPSAADRA
2	3.1e-331	DTATRKREIAAFLAQTSHETTGGWATAPDGPYAWGYCFKQEQGNPPDYCV
3	6.3e-277	RLPGYGVITNIINGGJECGKGSDSRVADRIGFYKRYCDILGVSYGDNLDC
4	3.2e-265	GGGVASJISSSLFDQMLKHRNDAACPAKGFYTYNAFIAAANSFSGFGTTG
5	6.1e-159	AEQCGSQAGGALCPGGLCCSQFGWCGSTSPYC

Table 5. The motif consensus of PR4 which detected in InterPro database

Motif PR4	E-value	Sequences
1	4.7e-343	ASNVRATYHYYNPEQNNWDLNAVSAYCATWDADKPLEWRSKYGWTAFCGP
2	1.1e-287	TGTQATVRIVDQCSNGGLDLDQGVFNQJDTDGKGYAQGHLIVNYQFVBCG
3	3.8e-085	VGPTGQAACGKCLRVTNTA
4	1.9e-018	MKRRVSIVVAVLLCLAAAA

Table 6. The motif consensus of PR5 which detected in InterPro database

Motif *PR5*	E-value	Sequences
1	6.8e-291	VVACKSACLAFBLDVFCCRNSYGKPEKCKPSMYSKMFKDACPSYFSYAYD
2	1.5e-271	CQTGDCQGLLSCNGTIGLPPATLVEVALQEDKSKPSFYDVSLVDGYNLPI
3	8.4e-149	FHISNKCPFPIWPAAAPNAGHPVIADGG
4	1.9e-134	GQTKRVHAPPTWNGRFWARTGCNFNS

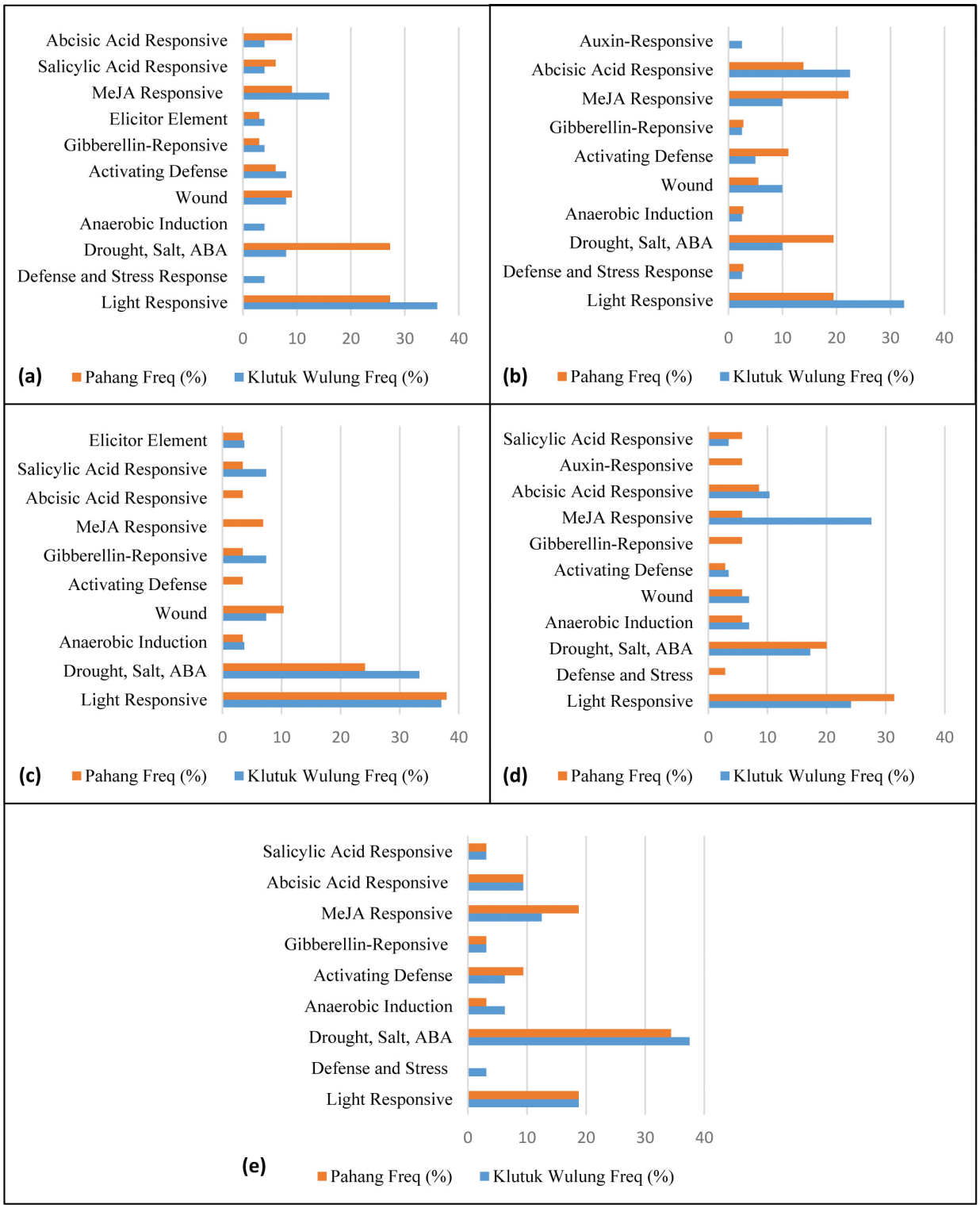

Fig. 3. The phylogenetic tree of (a) PR1, (b) PR2, (c) PR3, (d) PR4, and (e) PR5 between banana A genome (Musa acuminata cv. Pahang), B genome (Musa balbisiana cv. Klutuk Wulung), and other species.

shows only four motif consensus found in PR5 are reported in Interpro database as shown in (Table 6).

There were no GO Terms found from gene ontology searching for PR5. Nonetheless, the presence of GH64 and TLP domain has a role in the protection of plants against pathogen attacks, including bacterial. The GH64 domain is capable to hydrolyze β-1,3 glucan on the bacterial cell wall. In addition, TLP domain is capable to interact with β-1,3 glucan through hydrogen bond and cause the formation of transmembrane pores (Ferreira *et al.,* 2007). These interactions may disturb the integrity of bacterial cell wall and cause cell lysis. The occurrence of conserved cysteine residue showed that PR5 is acidic which secreted and accumulated in the apoplast, cell wall, and intercellular space of the plant (Wang *et al.,* 2010).

Analysis of the phylogenetic tree was done to defined evolutionary relationships for *PR* genes between two banana species with other species based on genetic approach. In this research, we used molecular data that are in the form of protein sequences to construct phylogenetic trees. The Bayesian analysis showed that the PR1, PR2, PR3, PR4, and PR5 protein from two species of bananas are closely related to each other which is supported by a high branch support value not less than 95 (Figure 3). Some branch points in PR4 phylogenetic tree resulting in a multifurcating node. The phylogeny with multifurcating nodes may cause by lack of data so the program was not capable to decide the exact order of the branch (Xiong, 2006). Furthermore, the relationship of banana A genome and B with different species are cluster together in the same clade consist of other monocots species. Our results confirm that PR protein sequence of banana A genome and B which is considered in the same genus has the most similar motif composition for each other. Thus, quite similar to other monocots species from other taxa. The presence of motif as conserve sequence in protein during the evolution of plants reveal the importance of that motif in carrying out the function of the protein (Zhu *et al.,* 2012).

Conclusion

The *PR1, PR2, PR3, PR4*, and *PR5* genes in two banana genomes have different numbers of exons, introns and nucleotide compositions. The Cis-Acting regulatory elements in response to light are tend to be dominant in the promoter of any *PR* genes. In addition, the motive for response elements to jasmonic acid (JA) and salicylic acid (SA) is also found which is needed for the activation of the PR gene. The gene ontology search of each *PR* gene shows different biological and molecular activities, but with the main function of defending itself from invading pathogens. The corresponding PR protein sequence between A and B genomes have the same domain and plays a role in the mechanism of self-defense response against pathogens, including bacteria. In addition, the two bananas in this study have a close evolutionary relationship and belong to a clade with other monocot plant species. The result of this study suggested that *PR* genes are important to be further investigated as potential markers in developing resistant banana against Blood Disease. In the future, understanding the *PR* genes that come into play with the stress signaling pathway and transduction mechanism will provide opportunities for enhanced resistance engineering in crop plants.

Acknowledgments

We would like to thank School of Life Science and Technology, Institut Teknologi Bandung, and RISTEK-BRIN for material, facility, and financial support.

References

Ali, S., Ahmad, B., Kamili, A. N., Ali, A., Ahmad, Z., Akhter, J. and Grover, A. 2018. Pathogenesis-related proteins and peptides as promising tools for engineering plants with multiple stress tolerance. *Microbiological Research*. 212–213 (April) : 29–37.

Breen, S., Williams, S. J., Outram, M., Kobe, B. and Solomon, P. S. 2017. Emerging Insights into the Functions of Pathogenesis-Related Protein 1. *Trends in Plant Science*. 22(10): 871–879.

Calderan-Rodrigues, M. J., Fonseca, J. G., Clemente, H. S., Labate, C. A. and Jamet, E. 2018. Glycoside Hydrolases in Plant Cell Wall Proteomes: Predicting Functions That Could Be Relevant for Improving Biomass Transformation Processes. In: *Advances in Biofuels and Bioenergy*. *Intech Open*. 166–182.

Caarls, L., Pieterse, C. M. J. and Van Wees, S. C. M. 2015. How salicylic acid takes transcriptional control over jasmonic acid signaling. *Frontiers in Plant Science*. 6 (MAR) : 1–11.

Ciolkowski, I., Wanke, D., Birkenbihl, R. P. and Somssich, I. E. 2008. Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant Molecular Biol-*

ogy. 68 (1–2) : 81–92.
Chen, X., Li, C., Wang, H. and Guo, Z. 2019. WRKY transcription factors: evolution, binding, and action. *Phytopathology Research*. 1(1) : 1–15.
Doxey, A. C., Yaish, M. W. F., Moffatt, B. A., Griffith, M., and McConkey, B. J. 2007. Functional divergence in the Arabidopsis β-1,3-glucanase gene family inferred by phylogenetic reconstruction of expression states. *Molecular Biology and Evolution*. 24 (4) : 1045–1055.
Eden-Green, S. J. 1994. Diversity of Pseudomonas solanacearum and related bacteria in Southeast Asia: new directions for Moko disease. In: A. C. Hayward and G. L. Hartman (Eds.), *Bacterial wilt: the disease and its causative agent Pseudomonas solanacearum*. Wallingford, UK: CAB International.
Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. 32 (5) : 1792–1797.
Ferreira, R. B., Monteiro, S., Freitas, R., Santos, C. N., Chen, Z., Batista, L. M. and Teixeira, A. R. 2007. The role of plant defense proteins in fungal pathogenesis. *Molecular Plant Pathology*. 8 (5) : 677–700.
Franco, F. P., Dias, R. O., Toyama, D., Henrique-Silva, F., Moura, D. S. and Silva-Filho, M. C. 2019. Structural and functional characterization of PR-4 SUGARWINs from sugarcane and their role in plant defense. *Frontiers in Plant Science*. 9(January) : 1–11.
Gamir, J., Darwiche, R., van't Hof, P., Choudhary, V., Stumpe, M., Schneiter, R. and Mauch, F. 2017. *The Plant Journal*. 89 : 502-509.
Jeger, M. 1995. Banana diseases. In: *S. Gowen (Ed.)*, *Bananas and Plantains* (pp. 317–381). London: Springer Science and Bussuniess Media LLC.
Kangasjärvi, S., Neukermans, J., Li, S., Aro, E. M. and Noctor, G. 2012. Photosynthesis, photorespiration, and light signalling in defense responses. *Journal of Experimental Botany*. 63(4) : 1619–1636.
Kaur, A., Pati, P. K., Pati, A. M. and Nagpal, A. K. 2017. In-silico analysis of cis-acting regulatory elements of pathogenesis-related proteins of *Arabidopsis thaliana* and *Oryza sativa*. *PLoS ONE*. 12 (9).
Kim, S. R., Kim, Y. and An, G. 1993. Identification of methyl jasmonate and salicylic acid response elements from the nopaline synthase (nos) promoter. *Plant Physiology*. 103(1) : 97–103.
Kusumoto, S. 2004. Occurrence of blood disease of banana in Sumatra, Indonesia. *J Gen Plant Pathol*. *70* : 45–49.
Landim, P. G. C., Correia, T. O., Silva, F. D. A., Nepomuceno, D. R., Costa, H. P. S., Pereira, H. M. and Grangeiro, T. B. 2017. Production in Pichia pastoris, antifungal activity and crystal structure of a class I chitinase from cowpea (*Vigna unguiculata*): Insights into sugar binding mode and hydrolytic action. *Biochimie*. 135 : 89–103.
Liu, W., Xie, Y., Ma, J., Luo, X., Nie, P., Zuo, Z. and Ren, J. 2015. IBS: An illustrator for the presentation and visualization of biological sequences. *Bioinformatics*. 31(20) : 3359–3361.
Medeiros, S.C., Monteiro-Junior, J.E., Sales, G.W.P., Grangeiro, T.B. and Accioly, N. 2018. Chitinases as Antibacterial Proteins. *J Young Pharm.* 10(2) : 144-148.
Page, Roderic D.M. 1996. Tree View: An application to display phylogenetic trees on personal computers, *Bioinformatics*. 12 (4) : 357–358.
Pandey, S. P. and Somssich, I. E. 2009. The role of WRKY transcription factors in plant immunity. *Plant Physiology*. 150(4) : 1648–1655.
Pecenkova, T., Pleskot, R. and Zarsky, V. 2017. Subcellular Localization of *Arabidopsis thaliana* Pathogenesis-Related 1 (PR1) Protein. *Int. J. Mol. Sci*.18 (825).
Pua, E.C. 2007. *Banana. In Biotechnology in agriculture and forestry, vol. 60, trans-genic crops V , edited by E.C. Pua and M.R. Davey*. (pp. 3–34). Berlin, Germany : Springer-Verlag.
Rahmawati, A. R. 2018. *Studi Ekspresi Gen Terkait Ketahanan Terhadap Penyakit Darah (Blood Disease) Pada Tanaman Pisang (Musa spp.).* Bandung : Institut Teknologi Bandung.
Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S. and Huelsenbeck, J. P. 2012. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*. 61 (3) : 539–542.
Ruvinsky, A. and Watson, C. 2007. Intron Phase Patterns in Genes: Preservation and Evolutionary Changes. *The Open Evolution Journal*. 1 : 1–14.
Salazar, A. M. and Fernando, D. D. 2019. Predicted functions, subcellular localizations, and expression patterns of genes encoding secretory proteins associated with pine pollen germination. *Tree Genetics and Genomes*. 15(39) : 1–15.
Ubhayasekera, W. 2011. Structure and function of chitinases from glycoside hydrolase family 19. *Polymer International*. 60(6): 890–896.
Van Bel, M., Diels, T., Vancaester, E., Kreft, L., Botzki, A., Van De Peer, Y. and Vandepoele, K. 2018. PLAZA 4.0: An integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Research*. 46 (D1) : D1190–D1196.
Wang, X., Tang, C., Deng, L., Cai, G., Liu, X., Liu, B. and Han, Q. 2010. Characterization of a pathogenesis-related thaumatin-like protein gene TaPR5 from wheat induced by stripe rust. *Physiologica Plantarum*. 139 (1) : 27–38.
Xiong, J. 2006. *Essential Bioinformatics*. Cambridge University Press, Texas. U.A.
Yu, C. P., Lin, J. J. and Li, W. H. 2016. Positional distribution of transcription factor binding sites in Arabidopsis thaliana. *Scientific Reports*. 6 (January): 1–7.