Journal of Oil Palm Research Vol. 35 (4) December 2023 p. 626-638
DOI: https://doi.org/10.21894/jopr.2022.0078

MOLECULAR SCREENING OF BASAL STEM ROT RESISTANCE GENES IN OIL PALM

MOHD AMIN AB-HALIM[1]; ROZANA ROSLI[1]; ENG TI LESLIE LOW[1]*; ZETI-AZURA MOHAMED-HUSSEIN[2] and RAJINDER SINGH[1]

ABSTRACT

Basal stem rot (BSR) is one of the main oil palm diseases that have led to tremendous losses in oil yields for almost a century. Having markers, especially those that are linked to resistance (R) genes could potentially alleviate this problem by providing the tools to select palms that are tolerant or resistant to the disease. This study aimed to develop oil palm genomic markers that can distinguish oil palm plants with different levels of resistance to BSR. We identified 144 homologous R genes in the oil palm genome based on the conserved domain structure of known R proteins. Six simple sequence repeat markers were identified and used to genotype 40 palms with different levels of resistance to BSR. The observed and effective number of alleles ranged from 2.00 to 7.00 and 1.57 to 4.30, respectively. The observed heterozygosity ranged from 0.13 to 0.67, with a mean of 0.46 while the expected heterozygosity ranged from 0.40 to 0.78, with a mean of 0.51. Analysis of genetic distances from the set of markers was able to differentiate susceptible and tolerant palm samples. These results may help in the early selection of durable BSR disease resistant oil palm cultivars.

Keywords: *Elaeis guineensis*, genetic distance, phylogenetic tree, R gene candidates, SSR.

Received: 14 October 2021; **Accepted:** 21 September 2022; **Published online:** 22 December 2022.

INTRODUCTION

Oil palm (*Elaeis guineensis*) is the most important commercial crop in Malaysia and Indonesia. The oil is suitable for the production of edible and non-edible oils, biodiesel, and nutritional and pharmaceutical products (Parveez *et al.*, 2021). The basal stem rot (BSR) disease, which was first described in 1915 in the Republic of Congo, West Africa (Wakefield, 1920), has caused significant losses to palm oil production as trees replanted on previously infected locations often die before reaching the optimum maturity age (Ariffin *et al.*, 2000). About 9.2% of Malaysian oil palm smallholders' replanted areas have been infected by the disease (Mohd Shukri *et al.*, 2020), and the *Ganoderma boninense* fungus was identified as the causal pathogen of the disease (Ho and Nawawi, 1985; Susanto *et al.*, 2005).

There are several methods used for early disease detection, from expert manual observation of basal stem decay and foliar symptoms (Turner and Gillbanks, 1974), isolation of the fungi using *Ganoderma* selective medium (GSM) (Ariffin and Idris, 1991) to remote sensing via spectroscopic image techniques using a sound sensor in GanoSken (Idris *et al.*, 2010), electrical resistance (ER) data (Nurnadiah *et al.*, 2014) and artificial neural networks (Ahmadi *et al.*, 2017). Nevertheless, no effective treatment or control for the BSR disease in oil palm is currently available (Naher *et al.*, 2015; Shamala *et al.*, 2008). The best practices in controlling the disease include field sanitation management, fungicidal treatment, and foliar supplement. But they require intensive labour especially to excavate stumps, followed by ploughing and rotavating the planting areas. A large-scale oil palm plantation requires a rapid and non-complex prevention system for the BSR

[1] Malaysian Palm Oil Board,
6 Persiaran Institusi, Bandar Baru Bangi,
43000 Kajang, Selangor, Malaysia.

[2] Department of Applied Physics,
Faculty of Science and Technology,
Universiti Kebangsaan Malaysia,
43600 UKM Bangi, Selangor, Malaysia.

[2] Centre for Bioinformatics Research,
Institute of System Biology (INBIOSIS),
Universiti Kebangsaan Malaysia,
43600 UKM Bangi, Selangor, Malaysia.

* *Corresponding author e-mail: lowengti@mpob.gov.my*

disease problem. Hence, developing planting materials with resistance to BSR caused by the fungus offers a long-term solution to control the disease (Hushiarian *et al*., 2013). Oil palm resistant to the fungal infection have not been found, but *G. boninense* inoculated oil palm progenies varied in disease severity (Idris *et al*., 2004). These results were used to classify samples to appropriate resistance levels.

In plants, genes that recognise invading pathogens and trigger the plant immune system are known as resistance (R) genes (Hammond-Kosack and Jones, 1997). Abundantly present in plant genomes, they are involved in pathogen interaction mechanisms that result in a signaling cascade to activate the plant's defence mechanism (Ellis and Dodd, 2003; Wanderley-Nogueira *et al.,* 2007). The 'zigzag' model by Jones and Dangl (2006) describes two approaches to the plant immune system. The first involves the transmembrane pattern recognition receptors (PRRs) which act as the first line of defence to arrest pathogen invasion in the extracellular space. These receptors include the Receptor-Like Protein (RLP) and Receptor-Like Kinase (RLK) classes of R genes (Sekhwal *et al.,* 2015). The second defence approach involves the Coiled-Coil (CC) or Toll/Interleukin Receptor (TIR)-Nucleotide-binding Site (NBS)-Leucine-Rich Repeat (LRR) and Kinase classes of R genes which interact with effector molecules from pathogens that successfully penetrate the cell wall (Song *et al.,* 1997). Successful recognition of effector molecules by R genes will trigger the host cell to activate the plant's defence mechanism. However, unrecognised interactions will most likely result in the host being susceptible to the disease. R genes, therefore, play an initial role in protection against pathogen invasion by detecting and initiating the plant's defense mechanisms against elicitor proteins (Ellis and Dodds, 2003).

The oil palm genome sequence was reported previously (Singh *et al.,* 2013), together with information on gene models for putative resistance genes. In this study, we used *GeneThresher*™ methylation filtered oil palm genomic sequence data, and a method for gene prediction and annotation that focuses on identifying resistance genes. The *GeneThresher*™ sequences are gene enriched sequences from the hypomethylated regions of the highly repetitive oil palm genome.

The purpose of this study was to develop a set of R genes associated genomic markers from the oil palm genome sequence. This set of markers were then used to genotype oil palm samples with different levels of resistance to the BSR disease. Clustering of the simple sequence repeat (SSR) profiles allowed identification of closely related samples and clustered profiles associated with BSR resistance levels.

MATERIALS AND METHODS

Plant Materials

Four hybrid crosses in MPOB Kluang Research Station, Johor that were previously classified based on their severity of foliar symptoms following *G. boninense* infection (Idris *et al.,* 2004) were used in this study. The most susceptible hybrid cross is the *dura* x *dura* (DxD) PK 2724 cross, with a death rate of 40% after *G. boninense* infection. The *dura* x *pisifera* (DxP) PK 2567 cross recorded the lowest death rate of 10%, and was thus designated as a partial resistant cross. Another DxP cross, PK2572, with a recorded death rate of 25% was classified as an intermediate group. The *tenera* x *tenera* (TxT) palms (PK 1708), with a death rate of 15%, were also classified as partial resistant palms.

Identification of Oil Palm Resistance Gene Candidates

A total of 139 120 *GeneThresher*™ genomic contigs (Low *et al.,* 2014) from the *E. guineensis* var. *pisifera* were used in this study. GenBank was mined for curated R gene information based on previous reports by Bendahmane *et al.* (1999), Dodds *et al.* (2001), Hammond-Kosack *et al.* (1997) and Martin *et al.* (2003). These protein sequences were then used as reference sequences to identify homologs in the *E. guineensis GeneThresher*™ genomic libraries using TBLASTN (Altschul *et al.,* 1990). The E-value cut-off used was 1^{e-20}, as described by Grzebelus
et al. (2007).

Protein prediction was carried out using a semi-Hidden Markov Model (HMM)-based Nucleic Acid Parser, SNAP (Korf 2004) and GeneSeqer software (Schlueter *et al.,* 2003). The prediction used oil palm and *Oryza sativa* HMM with default parameters. Oil palm R gene candidates (RGC) were compared to the NCBI non-redundant database using BLASTX for translation direction verification and BLASTP for homologous gene comparison. R genes were selected based on domains identified via comparisons to Pfam, PRINTS-S and Superfamily databases using InterProScan (Quevillon *et al.,* 2005) on both the curated R genes and oil palm RGC. The RGC with the same domain ID as the curated R gene domains were grouped into the same class. ClustalW (Thompson *et al.,* 1994) was then used to align RGC sequences, except for the 'Others' class. The alignment was used to build a phylogenetic tree using MEGA6 (Tamura *et al., 2013*).

To improve the identification and annotation of R genes in oil palm, this study utilised information from 43 R genes in 11 plant species, namely *Arabidopsis thaliana* (13), *Beta vulgaris* (1), *Capsicum chacoense* (1), *Hordeum vulgare* (3), *Lactuca sativa* (1),

Linum usitatissimum (3), *Lycopersicon esculentum* (11), *Nicotiana tabacum* (1), *O. sativa* (4), *Solanum acaule* (3) and *Zea mays* (2). The most highly represented class of R genes (~67%) is the NBS-LRR class. This class is divided into two sub-classes based on the 5'-terminal structure (Marone *et al.*, 2013). The Pfam IDs for the 43 R genes are listed in *Table 1*, with their respective predicted protein functions and gene classes.

Analysis of Amplified DNA from SSR Primers

The identification of SSR markers in the RGC sequences was carried out using MicroSAtellite (MiSA) version 2.02 (Thiel *et al.*, 2003) and primers were designed using Primer3 (Rozen and Skaletsky, 2000). The primer pairs were used to genotype 40 samples. DNA extraction and purification were carried out using 5 g of leaf samples according to Doyle and Doyle (1990). The DNA samples were labelled with [γ^{-33}P] dATP (Singh *et al.*, 2009). Polymerase chain reaction (PCR), followed by electrophoresis in 5% acrylamide gel were conducted (Billotte *et al.*, 2001; Singh *et al.*, 2009).

DNA bands were scored for the presence (1) or absence (0) of the amplified bands from each sample group. POPGENE software version 1.32 (Yeh and Boyle, 1997) was used to estimate the genetic distance (GD) between the samples. GD was calculated using the unweighted pair-group method with arithmetic averages (UPGMA) (Nei, 1978) and relationships between samples are shown in a dendrogram constructed using PHYLIP version 3.5c (Felsenstein, 1993).

RESULTS AND DISCUSSION

Identification of Oil Palm RGC

We identified 144 oil palm RGC from the *E. guineensis GeneThresher*™ genomic libraries using the 43 curated R genes as template, and classified them into six classes based on their domain arrangement (*Figure 1*). Each gene class denote the same domain ID as the curated R genes, enabling gene classification (*Table 2*). The *PTO* gene (Kinase class), responsible for the resistance of *L. esculentum* to the bacterial speck disease caused by the *Pseudomonas syringae* (Kim *et al.*, 2002; Martin *et al.*, 1993), forms the basis of the R gene Kinase Class. This gene encodes a single Pkinase (IPR000719) protein domain that encodes a Serine/Threonine Kinase (STK). *PTO* recognises *AVRPTOA*

TABLE 1. PROTEIN DOMAIN INFORMATION OF CURATED RESISTANCE GENES

Class	No. of known R genes from other crops	Domain ID	Domain name	Function
Kinase	1	IPR000719	Pkinase	ATP binding
CNL	19	IPR002182	NB-ARC	Signaling motif to bind and hydrolyse ATP
		IPR000767	DISEASERSIST	4-element fingerprint that provides a signature for disease resistance proteins
TNL	10	IPR002182	NB-ARC	Signaling motif to bind and hydrolyse ATP
		IPR000157	TIR	Involved in protein interaction and signal transductions
		IPR001611	LRR_1	Involved in protein-protein interaction
RLP	4	IPR013210	LRRNT_2	Involved in protein-protein interaction
RLK	1	IPR000719	Pkinase	ATP binding
		IPR013210	LRRNT_2	Involved in protein-protein interaction
		IPR001611	LRR_1	Involved in protein-protein interaction
Others	8	IPR001509	NAD-dependent epimerase/dehydratase	Involved in catalytic activity and coenzyme binding
		IPR002182	NB-ARC	Signaling motif to bind and hydrolyse ATP
		IPR016040	NADP-bd	NADP-binding domain
		IPR004326	Mlo	Trigger a cascade of events that result in plant cell death
		IPR006121	HMA	Involved in bacterial resistance to toxic metals, such as lead and cadmium
		IPR000719	Pkinase	ATP binding
		IPR003657	WRKY	Regulate various physiological programs that are unique to plants, including pathogen defence
		IPR011713	LRR_3	Involved in protein-protein interaction
		IPR000157	Toll/Interleukin receptor TIR domain	Involved in signal transductions

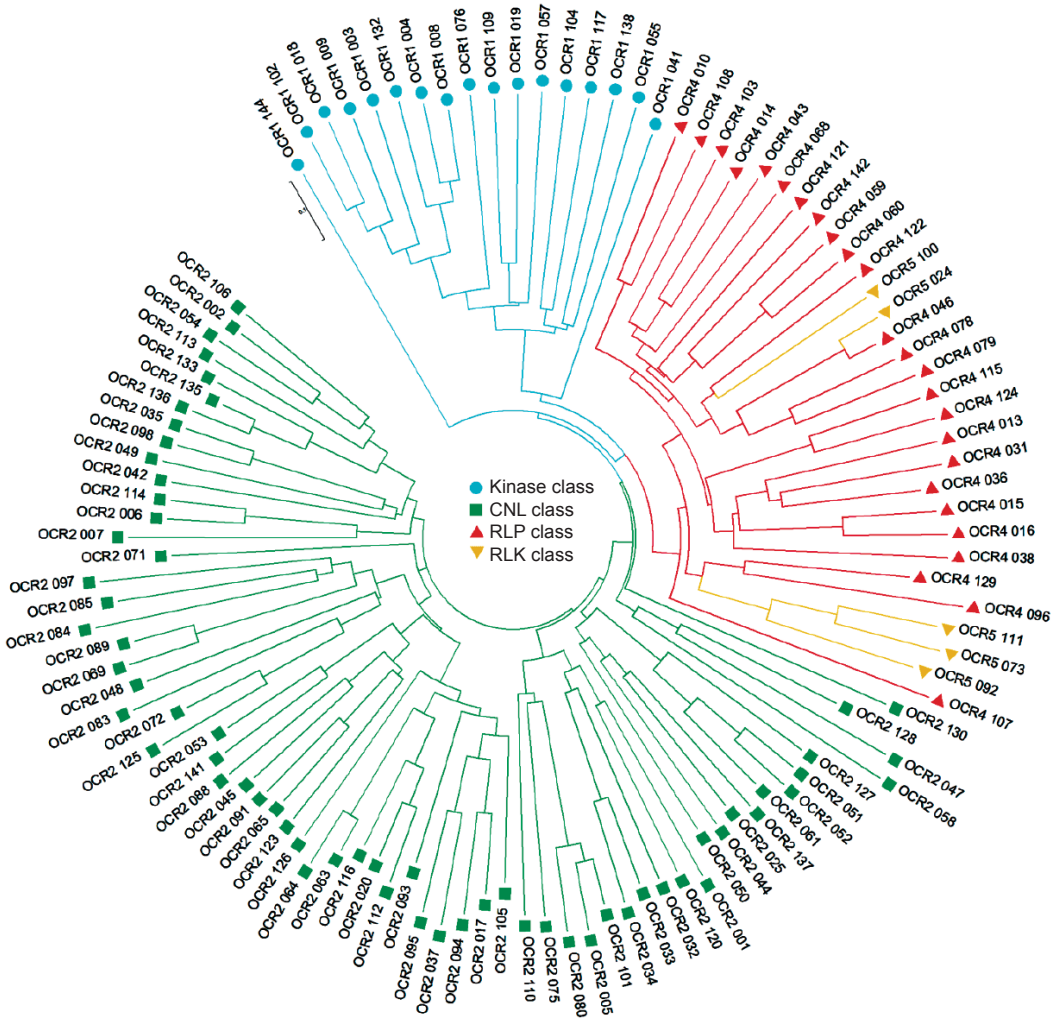

Figure 1. Phylogenetic analysis of the 144 oil palm RGCs identified from hypomethylated or gene-rich regions of the oil palm genome. Kinase, CNL, RLP and RLK classes are indicated in the legend at the centre of the image.

and *AVRPTOB* inside the host cell and halts the colonisation of *P. syringae* (Tang *et al.*, 1996). A total of 17 oil palm RGC were classified into the Kinase class, whereby these genes contain the Pkinase domain (IPR000719) which is similar to the *TaWAK6* genes associated with wheat resistance to leaf rust (Dmochowska-Boguta *et al.*, 2020). Most of the identified oil palm RGC are in the CNL class (65 candidate genes). We were unable to find candidate genes belonging to the TNL class. Analysis of gene models from seven monocot plants, namely *Musa acuminata*, *Setaria italica*, *Sorghum bicolor*, *Z. mays*, *O. sativa*, *Brachypodium distachyon* and *Phyllostachys heterocycle* also showed that the genomes do not have TNL genes, and it is likely that this class of genes were lost after the divergence of dicots and monocots (Shao *et al.*, 2016). Therefore, the TNL gene class may not exist in oil palm since oil palm is also a monocot. However, it is also possible that we could not identify these genes as the oil palm genome is not complete (Singh *et al.*, 2013).

Two important classes of R genes are RLK and RLP. Both contain LRR domains and a hydrophobic protein residue involved in protein-protein interaction (Jones and Jones, 1997). RLK genes also contain an additional Kinase domain. A total of 25 and five oil palm genes were identified as RLP and RLK genes respectively. These genes may be involved in the first stage of oil palm's defence mechanism. Plant defence mechanisms are represented as a four-phase zigzag model (Jones and Dangl, 2006). In phase 1, the pathogen-associated molecular pattern (PAMP) elicitors are recognised by RLK and RLP classes of R genes that contain extracellular domains to interact with these elicitors. These receptors mediate PAMP-triggered immunity (PTI) (Sekhwal *et al.*, 2015). In phase 2 of the infection, successful pathogens secrete effector molecules into the host cell to interfere with the PTI response, resulting in effector-triggered susceptibility. Phase 3 marks the beginning of the second stage of the defense mechanism in the plant immune system, triggered inside the host cell by the Kinase and CNL R gene classes. Once the effectors are recognised by these R genes, colonisation by the pathogen can be halted, usually via a hypersensitive cell death

TABLE 2. CLASSIFICATION OF OIL PALM RGC BASED ON PROTEIN DOMAIN AND BLASTP

Class	Contig	Domain ID	Domain name	Domain E-value	BlastP	Blast E-value
1	OCR1_003			1.90E-18	Pto-like serine/threonine kinase	8.00E-82
1	OCR1_004			1.40E-28	Pto-like serine/threonine kinase	4.00E-74
1	OCR1_008			4.40E-21	Pto	9.00E-79
1	OCR1_009			3.30E-37	Pto resistance protein candidate Tg-65	6.00E-77
1	OCR1_019			3.00E-44	Pto	1.00E-43
1	OCR1_041			8.60E-49	Pto-like protein	1.00E-67
1	OCR1_055			4.20E-34	Pto-like serine/threonine kinase	6.00E-04
1	OCR1_057	PF00069	Pkinase	8.10E-44	Pto-like serine/threonine kinase	1.00E-54
1	OCR1_076			1.60E-07	Pto-like serine/threonine kinase	1.00E-70
1	OCR1_102			1.20E-36	Pto kinase homologs	2.00E-116
1	OCR1_104			9.30E-57	Pto-like protein	2.00E-57
1	OCR1_109			7.80E-29	Putative Pto-like serine/threonine kinase	2.00E-18
1	OCR1_117			9.90E-57	Pto-like protein	4.00E-66
1	OCR1_132			3.10E-18	Pto-like serine/threonine kinase	8.00E-77
1	OCR1_138			3.60E-21	Pto	3.00E-30
1	OCR1_144			5.80E-29	Protein kinase family protein	1.00E-99
2	OCR2_080			2.30E-50	RPS5	1.00E-42
				1.50E-09		
2	OCR2_094			3.90E-36	Disease resistance protein RPS2	2.00E-40
				2.10E-18		
2	OCR2_097			1.70E-65	I2	9.00E-42
				4.10E-17		
2	OCR2_123			1.40E-59	Disease resistance protein RPS2	9.00E-36
				2.50E-07		
2	OCR2_063			1.50E-35	CC-NBS-LRR resistance protein	1.00E-36
				2.40E-10		
2	OCR2_064			7.50E-51	Late blight resistance protein Rpi-pta1	2.00E-67
				1.10E-14		
2	OCR2_128			3.00E-35	Disease resistance protein RPS2, putative	2.00E-77
				7.40E-06		
2	OCR2_098	PF00931 PR00364	NB-ARC Diseasersist	2.50E-23	Disease resistance protein RPM1, putative	4.00E-57
				7.00E-05		
2	OCR2_017			5.60E-58	Disease resistance protein RPS2	1.00E-41
				1.20E-09		
2	OCR2_020			1.80E-91	Disease resistance protein RPS2	2.00E-79
				1.10E-17		
2	OCR2_025			5.10E-55	Disease resistance protein RPS2	6.00E-61
				2.60E-11		
2	OCR2_033			1.40E-35	Disease resistance protein RPS2	8.00E-33
				2.30E-09		
2	OCR2_035			1.10E-20	Disease resistance protein RPM1	8.00E-30
				1.10E-05		
2	OCR2_037			1.70E-92	Disease resistance protein RPS2	9.00E-103
				3.80E-18		
2	OCR2_042			1.10E-34	Disease resistance protein RPP13	7.00E-48
				3.20E-06		

TABLE 2. CLASSIFICATION OF OIL PALM RGC BASED ON PROTEIN DOMAIN AND BLASTP (continued).

Class	Contig	Domain ID	Domain name	Domain E-value	BlastP	Blast E-value
2	OCR2_045			7.90E-81	Disease resistance protein RPS2	3.00E-52
				2.00E-15		
2	OCR2_048			3.50E-40	Disease resistance protein RPM1	2.00E-34
				7.90E-09		
2	OCR2_049			2.40E-27	Disease resistance protein RPM1	7.00E-58
				1.40E-05		
2	OCR2_050			9.80E-44	Disease resistance protein I2	7.00E-53
				2.60E-09		
2	OCR2_053			2.50E-73	Disease resistance protein RPS2	6.00E-52
				5.30E-18		
2	OCR2_061			2.80E-43	Disease resistance protein I2	2.00E-54
				9.20E-08		
2	OCR2_065			1.20E-48	Disease resistance protein RPS2	1.00E-47
				1.60E-07		
2	OCR2_069			3.00E-30	Disease resistance protein RPS2	2.00E-35
				5.50E-11		
2	OCR2_071			5.80E-46	Disease resistance protein RPM1	1.00E-54
				1.20E-10		
2	OCR2_072			3.70E-81	Disease resistance protein RPM1	7.00E-56
				8.90E-12		
2	OCR2_075			6.00E-17	RPS2	1.00E-66
				6.00E-17		
2	OCR2_083			1.40E-67	Disease resistance protein RPP8	2.00E-62
		PF00931 PR00364	NB-ARC Diseasersist	3.70E-16		
2	OCR2_084			9.00E-56	Disease resistance protein RPS2	5.00E-68
				3.10E-17		
2	OCR2_085			9.10E-60	Disease resistance protein RPS2	9.00E-71
				3.80E-17		
2	OCR2_088			8.90E-48	I2	1.00E-25
				1.30E-09		
2	OCR2_089			3.60E-56	Disease resistance protein RPS2	2.00E-61
				1.40E-15		
2	OCR2_091			1.40E-54	CC-NBS-LRR resistance protein	1.00E-55
				6.60E-09		
2	OCR2_093			2.30E-65	Disease resistance protein RPS2	4.00E-56
				1.30E-09		
2	OCR2_095			1.30E-82	Disease resistance protein RPS2	2.00E-83
				5.00E-17		
2	OCR2_101			1.70E-58	RPS5	2.00E-06
				8.70E-11		
2	OCR2_105			1.60E-81	Disease resistance protein RPS2	2.00E-96
				1.50E-17		
2	OCR2_110			1.70E-65	RPS2	9.00E-57
				5.30E-20		
2	OCR2_112			8.50E-38	CC-NBS-LRR resistance protein	3.00E-42
				4.80E-10		
2	OCR2_116			1.70E-36	CC-NBS-LRR resistance protein	1.00E-36
				4.40E-10		

TABLE 2. CLASSIFICATION OF OIL PALM RGC BASED ON PROTEIN DOMAIN AND BLASTP (continued).

Class	Contig	Domain ID	Domain name	Domain E-value	BlastP	Blast E-value
2	OCR2_125			9.50E-52	Disease resistance protein RPP13	1.00E-31
				5.00E-06		
2	OCR2_126			3.40E-43	Disease resistance protein RPS2	6.00E-39
		PF00931 PR00364	NB-ARC Diseasersist	1.10E-05		
2	OCR2_133			2.70E-41	Disease resistance protein RPM1	8.00E-86
				1.80E-05		
2	OCR2_136			1.60E-59	RPM1	3.00E-49
				5.60E-08		
4	OCR4_014			5.00E-12	Cf-2.2	3.00E-54
4	OCR4_043	PF08263	LRRNT_2	1.90E-13	Cf-2.1	4.00E-64
4	OCR4_096			2.50E-06	Cf-2.3	1.00E-40
4	OCR4_010			8.80E-03	Cf-2.3	1.00E-30
4	OCR4_015			2.50E-03	Cf-2.1	1.00E-25
4	OCR4_016			9.10E-04	Putative Cf2/Cf5	4.00E-28
4	OCR4_031			5.10E-03	Cf2/Cf5-like	1.00E-57
4	OCR4_036			2.50E-03	Cf2/Cf5-like	1.00E-122
4	OCR4_038			4.10E-04	Cf2/Cf5-like	2.00E-40
4	OCR4_013			1.30E-02	Putative Cf2/Cf5	1.00E-33
4	OCR4_046			1.60E-02	Cf-2.1 putative	3.00E-38
4	OCR4_059			3.40E-03	Cf-2.1	1.00E-68
4	OCR4_060			2.00E-04	Cf-2.2	4.00E-62
4	OCR4_068	PF00560	LRR_1	3.40E-03	Cf-2.1	2.00E-60
4	OCR4_078			2.10E-02	Cf-2.1	3.00E-55
4	OCR4_079			6.90E-04	Cf-2.1	1.00E-98
4	OCR4_103			1.60E-04	Cf-2.3	5. 00e-77
4	OCR4_107			1.30E-03	Putative Cf2/Cf5	2.00E-22
4	OCR4_108			4.40E-03	Cf-2.1	4.00E-73
4	OCR4_115			3.30E-03	Cf-2.2	5.00E-20
4	OCR4_121			6.60E-03	Cf-2.2	2.00E-67
4	OCR4_122			7.30E-03	Cf-2.1	9.00E-54
4	OCR4_124			4.40E-03	Cf-2.2	6.00E-31
4	OCR4_129			6.90E-04	Cf-2.3	2.00E-09
4	OCR4_142			4.50E-03	Cf-2.3	4.00E-68
5	OCR5_073			6.40E-51		4.00E-104
				3.70E-31		
				6.70E-02		
5	OCR5_092	SSF56112 PF00069 PF00560	Protein kinase-like Pkinase LRR_1	1.50E-53		0
				8.10E-35	Protein kinase Xa21	
				3.90E-01		
5	OCR5_024			2.50E-13		5.00E-121
				1.50E-04		
				1.30E-01		
6	OCR2_001			6.00E-30	RPS5	3.00E-39
6	OCR2_054			1.50E-22	Disease resistance protein RPM1	6.00E-62
6	OCR2_058	PF00931	NB-ARC	6.10E-45	Disease resistance protein RPH8A, putative	3.00E-58
6	OCR2_106			1.70E-36	Disease resistance protein RPM1	2.00E-54
6	OCR2_114			4.80E-64	Disease resistance protein RPP8	4.00E-53
6	OCR2_113			1.30E-41	Disease resistance protein RPM1	6.00E-48

TABLE 2. CLASSIFICATION OF OIL PALM RGC BASED ON PROTEIN DOMAIN AND BLASTP (continued).

Class	Contig	Domain ID	Domain name	Domain E-value	BlastP	Blast E-value
6	OCR2_002			6.90E-43	Disease resistance protein RPM1	9.00E-48
6	OCR2_120			2.70E-32	Putative RPS2	1.00E-47
6	OCR2_006			3.90E-52	Disease resistance protein RPP8	1.00E-46
6	OCR2_044			3.30E-31	Disease resistance protein RPS2	2.00E-40
6	OCR2_034			4.00E-27	Disease resistance protein RPS5	2.00E-36
6	OCR2_127			3.10E-33	Disease resistance protein RPM1	1.00E-34
6	OCR2_032	PF00931	NB-ARC	4.30E-17	RPS5	4.00E-27
6	OCR2_135			6.10E-22	Disease resistance protein RPM1	9.00E-24
6	OCR2_047			3.40E-27	Disease resistance protein RPS2	1.00E-23
6	OCR2_137			3.90E-14	Disease resistance protein RPS2	6.00E-23
6	OCR2_007			5.20E-31	Disease resistance protein Gpa2	2.00E-20
6	OCR2_141			1.00E-29	Disease resistance protein I-2	2.00E-17
6	OCR2_005			2.70E-22	Disease resistance protein RPS5	1.00E-13
6	OCR6_062			3.20E-25	Leucine-rich repeat transmembrane protein kinase	6.00E-96
6	OCR6_029			1.50E-32	Protein kinase Xa21	3.00E-79
6	OCR6_070			1.40E-37	Putative protein kinase Xa21	4.00E-78
6	OCR6_139			2.00E-53	Serine-threonine protein kinase	2.00E-74
6	OCR6_066			1.40E-54	Protein kinase family protein	9.00E-72
6	OCR6_087			6.10E-52	Putative rust resistance kinase Lr10	2.00E-71
6	OCR6_040			1.10E-48	Protein kinase-like resistance protein	1.00E-69
6	OCR6_021			2.10E-50	Resistance protein candidate	1.00E-68
6	OCR6_118			1.70E-50	Resistance protein candidate	3.00E-68
6	OCR6_099			9.10E-36	Putative protein kinase Xa21	8.00E-68
6	OCR6_056	PF00069 PF00069	Pkinase Pkinase	9.70E-41	Protein kinase-coding resistance protein	3.00E-67
6	OCR6_086			2.40E-41	Protein kinase-coding resistance protein	1.00E-62
6	OCR6_119			1.30E-23	Protein kinase family protein	2.00E-61
6	OCR6_030			2.10E-24	Protein kinase Xa21	1.00E-59
6	OCR6_027			1.20E-36	RFO1 (Resistance to *Fusarium Oxysporum* 1)	2.00E-49
6	OCR6_082			4.50E-26	Putative protein kinase Xa21	2.00E-48
6	OCR6_081			1.10E-17	Protein kinase-like resistance protein	8.00E-42
6	OCR6_074			5.80E-16	Putative protein kinase Xa21	2.00E-39
6	OCR6_143			1.60E-31	Bacterial blight resistance protein XA26	1.00E-37
6	OCR6_026			3.00E-44	Cf-2.2	5.00E-23
6	OCR6_011			3.20E-52	Protein kinase-like resistance protein	1.00E-12
6	OCR6_039			2.80E-32	Protein kinase-coding resistance protein	2.30E-02
6	OCR5_100				Protein kinase Xa21	2.00E-170
6	OCR6_134			2.00E-02	Disease resistance protein RPM1	7.00E-56
6	OCR6_012			9.00E-03	Rust resistance protein Rp1	4.00E-36
6	OCR6_077	PF00560	LRR_1	8.30E-03	Putative protein kinase Xa21	5.00E-33
6	OCR6_067			3.40E-03	Putative RPS2	6.00E-25
6	OCR2_130			3.20E-01	RPS2	3.00E-25
6	OCR2_051			5.30E-01	putative disease resistance RPP13-like protein 1-like	5.00E-21
6	OCR6_022	PF07014	Hs1pro-1_C	2.40E-71	Hs1pro-1	7.00E-43
6	OCR6_023			2.40E-56	Hs1pro-1	2.00E-26
6	OCR1_018	PF07714	Pkinase_Tyr	1.80E-33	Resistance gene homologs	5.00E-107
6	OCR6_140	SSF56112	Protein kinase-like	8.00E-14	Protein kinase-coding resistance protein	1.00E-11

TABLE 2. CLASSIFICATION OF OIL PALM RGC BASED ON PROTEIN DOMAIN AND BLASTP (continued).

Class	Contig	Domain ID	Domain name	Domain E-value	BlastP	Blast E-value
6	OCR6_131			7.90E-04	Leucine-rich repeat	0
6	OCR6_028	PF00560 PR00019 PF00069	LRR_1 LeuRichRpt Pkinase	1.30E-06 9.50E-40 7.90E-03 4.90E-06 2.50E-29	Protein kinase-coding resistance protein	6.00E-75
6	OCR2_052	PF00931	NB-ARC	2.00E-23	CC-NBS-LRR resistance protein	3.00E-57
6	OCR6_090	PF00560	LRR_1	2.30E-02	Hs1pro-1	3.00E-48
		PF07231	Hs1pro-1_N	1.00E-07		
		PF07014	Hs1pro-1_C	7.30E-40		
6	OCR5_111	SSF56112	Protein kinase-like	2.60E-01	Protein kinase Xa21	6.00E-16
		PF00069	Pkinase			

response. In phase 4, natural selection may result in pathogens that have gained new effectors. The selection process then favours host plants with new or different R genes that can neutralise the new effectors (Jones and Dangl, 2006). A total of 82 oil palm Kinase and CNL class RGC that are part of the second stage of plant defence mechanism were identified.

Identification of Polymorphic SSRs in Oil Palm RGC

A total of 32 out of the 144 oil palm RGC contain SSR motifs and 32 SSR primer pairs were thus designed. Ten SSR primer pairs that were able to produce amplified products with a Phred quality score of more than 20 were selected to screen 40 palms from the four populations. All primer pairs produced amplicons, of which six were polymorphic (*Table 3*) while four primer pairs, GC2.582.2, GC2.033.1, GC6.887_2 and GC2.033.2 showed monomorphic profiles.

An average of three alleles per locus were detected by the polymorphic SSR markers, with a total of 18 alleles in the progenies (*Table 4*). The highest number of alleles (Na) detected was with the GC6.651 marker, which was seven, and the minimum number of observed polymorphic alleles was two, which was the most common among the marker. The effective number of alleles (Ne) ranged from 1.568 to 4.300. The observed heterozygosity (Ho) and expected heterozygosity (He) ranged from 0.125 to 0.692 and 0.367 to 0.777, respectively. Markers are considered to be polymorphic if the heterozygosity score is more or equal to 0.1, and highly polymorphic if the value is more or equal to 0.7 (Romero *et al.*, 2019). The average Ho and He were 0.462 and 0.512 respectively, indicating that the markers were polymorphic. SSRs are widely used for marker-assisted selection (Lawson and Liqing 2006; Powell *et al.*, 1996; Robinson *et al.*, 2004) and have been applied as DNA markers to screen for RGC in rice (Fjellstrom *et al.*, 2006). Screening with additional markers can help verify the consistency of data obtained in this study.

The developed SSR markers did not include the oil palm RGC classes RLK and RLP. These two classes of R proteins are secreted into the extracellular environment and are not involved in the second stage of the plant defense mechanism.

Genetic Distance and Relationship of the Four Populations with Different BSR Resistance Levels

The amplification patterns of the SSR markers were used to calculate Nei's GD among the 40 oil palm samples with different resistance levels to BSR disease. The results showed that GD ranged from 0 to 0.372. A UPGMA dendrogram showed that the samples were grouped into four clusters (*Figure 2*). Cluster A comprises samples from the susceptible progeny (PK 2724) with a population death rate of 40% to BSR disease. Most (67%) of the intermediate resistant samples (PK 2572) with a population death rate of 25% are grouped into cluster B, which also contains four samples from population PK 2724. The clustering also showed that clusters A and B displayed the lowest genetic similarity coefficient, suggesting that they are closely related. This is in line with observations by Idris *et al.* (2004), where progeny PK 2724 is not significantly different in the severity of its foliar symptoms compared to PK 2572, but was distinctly different when compared to PK 2567 (population death rate of 10%) and PK 1708 (death rate of 15%). Cluster C comprise all of the partial resistance samples from population PK 2567, two samples from population PK 2724 and one sample from population PK 2572. All partial resistance samples from population PK 1708 were separated into a distinct cluster (cluster D). Both Cluster C and D could represent palms partially resistant to the disease.

TABLE 3. POLYMORPHIC SSR MARKERS IN OIL PALM RGC

No.	SSR name	Class	Candidate gene	SSR type	Motif	Size (nt)	Annealing temp. (°C)
1	GC1.365	Kinase	OCR1_057	p1	$(A)_{13}$	160-170	55
2	GC1.979	Kinase	OCR1_144	p3	$(GAG)_{6}$	120-130	51
3	GC2.582	CNL	OCR2_058	p3	$(GCA)_{5}$	270-280	51
4	GC6.398	Others	OCR6_027	p2	$(AT)_{8}$	140-150	51
5	GC6.887	Others	OCR2_114	p1	$(T)_{16}$	230-240	51
6	GC6.651	Others	OCR6_119	c	$(GT)_{10}(AT)_{19}$	250-260	55

TABLE 4. SUMMARY OF GENETIC VARIATIONS AND HETEROZYGOSITY STATISTICS FOR ALL LOCI

SSR name	Na	Ne	Ho	He
GC1.365	2	1.827	0.692	0.459
GC1.979	2	1.853	0.667	0.466
GC2.582	2	1.663	0.350	0.403
GC6.398	3	2.441	0.400	0.598
GC6.887	2	1.568	0.125	0.367
GC6.651	7	4.300	0.539	0.777

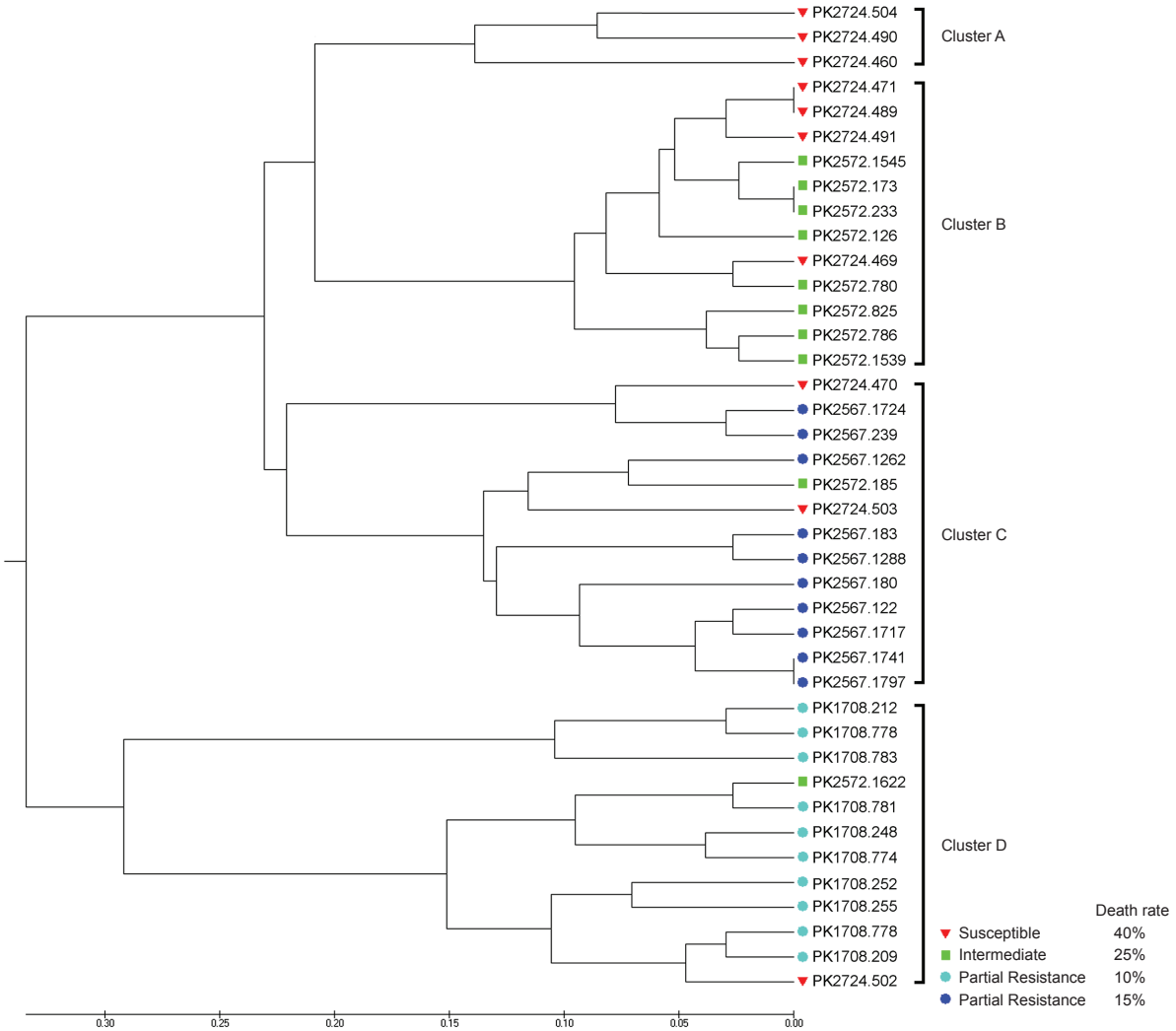

Figure 2. Dendrogram based on UPGMA analysis of 40 oil palm samples with different levels of resistance to Ganoderma boninense *infection.*

The dendrogram shows that the samples in each cluster are from the same population, with some exception that result in mixed samples from different populations. The relationship between sets of samples is shown in *Figure 3*. One of the reasons for the presence of mixed samples in a cluster is that populations will have some levels of spatial variation that can impact its resistance or susceptibility to diseases (Laine *et al.*, 2011). Palms could also be grouped into a cluster based on the similarity of their genetic backgrounds. Progeny PK 2724 originated from crosses of Deli (Elmina) palms while PK 2572 is from a Deli (Banting) x AVROS population. PK 2567 is a hybrid of Congo x Cameroon, and PK 1708 is from a Cameroon x Cameroon population (Idris *et al.*, 2004). The present set of markers were able to differentiate the partial resistant populations from the intermediate and susceptible populations. Using markers associated with resistance to BSR would be beneficial in breeding programs aimed at developing *G. boninense* resistant palms. Nevertheless, this study is limited in the number of samples tested, and needs to be extended to field trials to verify the findings.

CONCLUSION

Identification of resistance gene analogues in the oil palm genome enabled the development of SSR markers related to the BSR disease. These markers have the potential to predict and select oil palm progenies with a higher tolerance level to the disease before planting. Nevertheless, the evaluation of these potential markers needs to be expanded to study their effectiveness in more diverse populations.

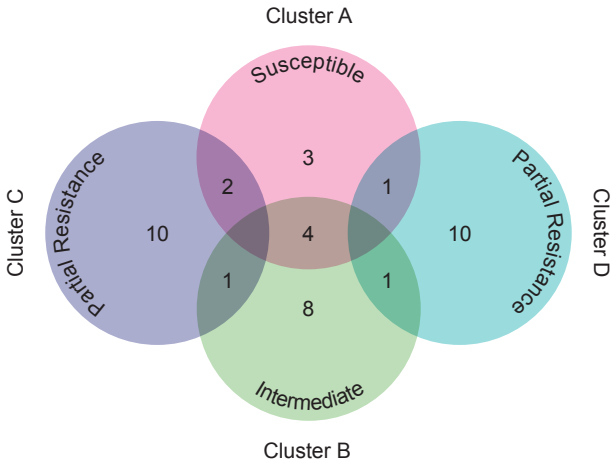

Figure 3. Frequency of oil palm samples with different levels of resistance to BSR disease sets according to UPGMA clusters. All of the partial resistance palm samples are unique and mixed samples are found in intermediate and susceptible sets.

ACKNOWLEDGEMENT

The authors would like to thank the Director-General of MPOB for permission to publish this article.

REFERENCES

Ahmadi, P; Muharam, F M; Khairulmazmi, A; Shattri and M and Idris, A S (2017). Early detection of *Ganoderma* basal stem rot of oil palms using artificial neural network spectral analysis. *Plant Dis., 101(6):* 1009-1016.

Altschul, S F; Gish, W; Miller, W; Myers, E W and Lipman, D J (1990). Basic local alignment search tool. *J. Mol. Boil., 215*: 403-410.

Ariffin, D and Idris, A S (1991). A selective medium for isolation of *Ganoderma* from diseased tissues. *Proc. of the 1991 International Palm Oil Conference, Progress, Prospects and Challenges Towards the 21st Century*. MPOB, Bangi. p. 517-519.

Ariffin, D; Idris, A S and Singh, G (2000). Status of *Ganoderma* in oil palm. *Ganoderma Diseases of Perennial Crops* (Flood, J; Bridge, P D and Holderness, M eds.). New York: Centre for Agricultural Bioscience International Publishing. p. 49-68.

Bendahmane, A; Kanyuka, K and Baulcombe, D C (1999). The *Rx* gene from potato controls separate virus resistance and cell death responses. *Plant Cell, 11*: 781-791.

Billotte, N; Risterucci, A M; Barcelos, E; Noyer, J L; Ambiard, P and Baurens, F C (2001). Development, characterisation, and across-taxa utility of oil palm (*Elaeis guineensis* Jacq.) microsatellite markers. *Genome, 44*: 413-425.

Dmochowska-Boguta, M; Kloc, Y; Zielezinski, A; Werecki, P; Nadolska-Orczyk, A; Karlowski, W M and Orczyk, W (2020). *TaWAK6* encoding wall-associated kinase is involved in wheat resistance to leaf rust similar to adult plant resistance. *PLoS ONE, 15(1)*: e0227713.

Dodds, P N; Lawrence, G J and ElliS, J G (2001). Six amino acid changes confined to the leucine-rich repeat β-strand/β-turn motif determine the difference between the *P* and *P2* rust resistance specificities in flax. *Plant Cell, 13*: 163-178.

Doyle, J J and Doyle, J L (1990). Isolation of plant DNA from fresh tissue. *Focus, 12(12)*: 13-15.

Ellis, J and Dodds, P (2003). Plant pathology: Monitoring a pathogen-targeted host protein. *Curr. Biol., 13*: 400-402.

Felsenstein, J (1993). PHYLIP (Phylogeny Inference Package) version 3.5c. Department of Genome Sciences, University of Washington, Seattle.

Fjellstrom, R; McClung, A M and Shank, A R (2006). SSR markers closely linked to the *Pi-z* locus are useful for selection of blast resistance in a broad array of rice germplasm. *Mol. Bred., 17*: 149-157.

Grzebelus, D; Lasota, S; Gambin, T; Kucherov, G and Gambin, A (2007). Diversity and structure of PIF/Harbinger-like elements in the genome of *Medicago truncatula*. *BMC Genomics, 8*: 409-422.

Hammond-Kosack, K E and Jones, J D G (1997). Plant disease resistance genes. *Annu. Rev. Physiol. Mol. Biol. Plants, 48*: 575-607.

Ho, Y W and Nawawi, A (1985). *Ganoderma boninense* pat. from basal stem rot of oil palm (*Elaeis guineensis*) in Peninsular Malaysia. *Pertanika, 8(3)*: 425-428.

Hushiarian, R; Yusof, N A and Dutse, S W (2013). Detection and control of *Ganoderma boninense*: Strategies and perspectives. *Springerplus 2*: 555.

Idris, A S; Kushairi, A; Ismail, S and Ariffin, D (2004). Selection for partial resistance in oil palm progenies to *Ganoderma* basal stem rot. *J. Oil Palm Res., 16(2)*: 12-18.

Idris, A S; Mazliham, M S; Loonis, P and Mohd Basri, W (2010). GanoSken for early detection of *Ganoderma* infection in oil palm. *MPOB Information Series No. 442*.

Jones, D A and Jones, J D G (1997). The role of leucine-rich repeat proteins in plant defences. *Adv. Bot. Res., 24*: 89-167.

Jones, J D G and Dangl, J L (2006). The plant immune system. *Nature, 444(16)*: 323- 329.

Kim, Y J; Lin, N C and Martin G B (2002). Two distinct *Pseudomonas* effector proteins interact with the *Pto* kinase and activate plant immunity. *Cell, 109(5)*: 589-598.

Korf, I (2004). Gene finding in novel genomes. *BMC Bioinform., 5*: 59.

Laine, A L; Burdon, J J; Dodds, P N and Thrall P H (2011). Spatial variation in disease resistance: From molecules to metapopulations. *J. Ecol., 99(1)*: 96-112.

Lawson, M J and Liqing, Z (2006). Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. *Genome Biol., 7*: R14.

Low, E T L; Rosli, R; Jayanthi, N; Mohd Amin, A H; Azizi, N; Chan, K L; Maqbool, N J; Maclean, P; Brauning, R; Mcculloch, A; Morage, R; Ong-Abdullah, M and Singh, R (2014). Analyses of hypomethylated oil palm gene space. *PLoS ONE, 9(1)*: e86728.

Martin, G B; Bogdanove, A J and Sessa, G (2003). Understanding the functions of plant disease resistance proteins. *Annu. Rev. Plant Biol., 54*: 23-61.

Martin, G B; Brommonschenkel, S H; Chunwongse, J; Frary, A; Ganal, M W; Spivey, R; Wu, T; Earle, E D and Tanksley S D (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science, 262(5138)*: 1432-1436.

Marone, D; Russo, M A; Laidò, G; De Leonardis, A M and Mastrangelo, A M (2013). Plant nucleotide binding site-leucine-rich repeat (NBS-LRR) genes: Active guardians in host defense responses. *Int. J. Mol. Sci., 24*: 7302-7326.

Mohd Shukri, I; Idris, A S; Rusli, M H; Izzuddin, M A; Norman, K; Khairuman, H and Zulkifli, A M (2020). Surveillance of *Ganoderma* disease in oil palm planted by participants of the smallholders replanting incentive scheme in Malaysia. *J. Oil Palm Res., 32(2)*: 237-244.

Naher, L; Siddiquee, S; Yusof, U K and Mondal, M M A (2015). Issues of *Ganoderma spp*. and basal stem rot disease management in oil palm. *Am. J. Agric. Res., 2(3)*: 103-107.

Nei, M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics, 89*: 583-590.

Nurnadiah, E; Aimrun, W; Amin, M S M and Idris, A S (2014). Preliminary study on detection of basal stem rot (BSR) disease at oil palm tree using electrical resistance. *Agric. Agric. Sci. Procedia, 2*: 90-94.

Parveez, G K A; Ahmad Tarmizi A H; Sundram S; Loh S K; Ong-Abdullah, M; Poo Palam, K D, Mohamed Salleh, K; Mohd Ishak, S and Idris, Z (2021). Oil palm economic performance in Malaysia and R&D progress in 2020. *J. Oil Palm Res., 33(2)*: 181-214.

Powell, W; Machray, G C and Provan, J (1996). Polymorphism revealed by simple sequence repeats. *Trends Plant Sci., 1(7)*: 215-222.

Quevillon, E; Silventoinen, V; Pillai, S; Harte, N; Mulder, N; Apweiler, R and Lopez, R (2005). InterProScan: protein domains identifier. *Nucleic Acids Res., 33*: 116-120.

Robinson, A J; Love, C G; Batley, J; Barker, G and Edwards, D (2004). Simple sequence repeat marker loci discovery using SSR primer. *Bioinformatics, 20(9)*: 1475-1476.

Romero, M; Mujica, A; Pineda, E; Ccamapaza, Y and Zavalla, N (2019). Genetic identity based on simple sequence repeat (SSR) markers for Quinoa (*Chenopodium quinoa* Willd.). *Cienc. Investig. Agrar., 46(2)*: 166-178.

Rozen, S and Skaletsky, H (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol., 132*: 365-86.

Schlueter, S D; Qunfeng, D and Brendel, V (2003). GeneSeqer@PlantGDB: Gene structure prediction in plant genomes. *Nucleic Acids Res., 31(13)*: 3597-3600.

Sekhwal, M K; Li, P; Lam, I; Wang, X; Cloutier, S and You, F M (2015). Disease resistance gene analogs (RGAs) in plants. *Int. J. Mol. Sci., 16*: 19248-19290.

Shamala, S; Faridah, A; Zainal Abidin, M A and Umi Kalsom, Y (2008). Efficacy of single and mixed treatments of *Trichoderma harzianum* as biocontrol agents of *Ganoderma* basal stem rot in oil palm. *J. Oil Palm Res., 20*: 470-483.

Shao, Z Q; Xue, J Y; Wu, P; Zhang, Y M; Wu, Y; Hang, Y Y; Wang, B and Chen, J Q (2016). Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat genes reveal three anciently diverged classes with distinct evolutionary patterns. *Plant Physiol., 170(4)*: 2095-2109.

Singh, R; Tan, S G; Panandam, J M; Rahimah, R A; Ooi, L C L; Low, E T L; Sharma, M; Jansen, J and Cheah, S C (2009). Mapping quantitative trait loci (QTLs) for fatty acid composition in an interspecific cross of oil palm. *BMC Plant Biol., 9*: 114.

Singh, R; Ong-Abdullah, M; Low, E T L; Mohammad Arif, A M; Rozana, R; Nookiah, R; Ooi, L C; Ooi, S; Chan, K; Mohd Amin, A H; Norazah, A; Nagappan, J; Bacher, B; Lakey, N; Smith, S W; He, D; Hogan, M; Budiman, M A; Lee, E K; Desalle, R; Kudrna, D; Goicoechea, J L; Wing, R A; Wilson, R K; Fulton, R S; Ordway, J M; Martienssen, R A and Sambanthamurthi, R (2013). Oil palm genome sequence reveals divergence of interfertile species in old and new worlds. *Nature, 500*: 335-339.

Song, W Y; Pi, L Y; Wang, G L; Gardner, J; Holsten, T and Ronald, P C (1997). Evolution of rice *Xa21* disease resistance gene family. *Plant Cell, 9*: 1279-1287.

Susanto, A; Sudharto, P S and Purba, R Y (2005). Enhancing biological control of basal stem rot disease (*Ganoderma boninense*) in oil palm plantations. *Mycopathologia, 159*: 153-157.

Tamura, K; Stecher, G; Peterson, D; Filipski, A and Kumar, S (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol., 30(12)*: 2725-2729.

Tang, X; Frederick, R D; Zhou, J; Halterman, D A; Jia, Y and Martin, G B (1996). Initiation of plant disease resistance by physical interaction of *AvrPto* and *Pto* Kinase. *Science, 274*: 2060-2063.

Thompson, J D; Higgins, D G and Gibson, T J (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence. *Nucleic Acids Res., 22(22)*: 4673-4680.

Thiel, T; Michalek, W; Varshney, R K and Graner, A (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet., 106*: 411-422.

Turner, P D and Gillbanks, R A (1974). *Oil Palm Cultivation and Management*. 2nd edition. The Incorporated Society of Planters, Kuala Lumpur. 672 pp.

Wakefield, E (1920). Diseases of the oil palm in West Africa. *Bull. Misc. Inform. (RBGK), 1920(9)*: 306-308.

Wanderley-Nogueira, A C; Soares-Cavalcanti, N M; Morais, D A L; Belarmino, L C; Barbosa-Silva, A and Benko-Iseppon, A M (2007). Abundance and diversity of resistance genes in the sugarcane transcriptome revealed by *in silico* analysis. *Genet. Mol. Res., 6(4)*: 866-889.

Yeh, F C and Boyle, T J B (1997). Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belg. J. Bot., 129*: 157-163.