Identification and development of new polymorphic Microsatellite markers for *Ganoderma boninense* main causal agent of oil palm basal stem rot disease

Mercière Maxime[(1)], Laybats Anthony[(1)], Carasco-Lacombe Cathy[(1)], Tan Joon Sheong [(2)], Klopp Christophe[(4)], Tristan Durant-Gasselin [(3)], Sharifah Shahrul Rabiah Syed Alwee [(2)], Breton Fréderic[(1)], Létizia Camus-Kulandaivelu[(1)]

Cirad(1), Felda(2), Palmelit (3), INRA (4)

(1) CIRAD, UMR 108 - AGAPAvenue Agropolis, 34398 Montpellier, Cedex 5 France

(2) FELDA GLOBAL VENTURES RESEARCH & DEVELOPMENT SDN BHD

Level 42, Menara Felda, Platinum Park, No. 11 Persiaran KLCC, 50088 Kuala Lumpur, MALAYSIA

(3) PalmElit SAS, Bat 14, Parc Agropolis, 2214 Bd de la Lironde, 34980 Montferrier sur Lez, France.

(4) INRA Toulouse, CS 52627, 31326 Castanet-Tolosan, FRANCE.

Abstract:

Ganoderma boninense is a telluric lignicolous basidiomycete and the main causal agent of the basal stem rot, one of the most devastating diseases of oil palm (*Elaeis guinensis*). While the fight against *G. boninense* should be a priority in South-East Asia, only scarce information is available about the diversity level of this fungus, and almost nothing is known about its genetic structure and history. In this context, the development of an informative molecular marker set for characterizing *G. boninense* diversity is a key step to understand the biology of this pathogen. A *G. boninense* draft genome sequence assembly of 61.5 Mb (from 454 and Illumina sequencing) has been used to identify and develop a set of microsatellite markers (SSR). A total of 652 SSR were identified of which 145 SSR primer sets were developed. These SSR are characterized by motif from 2 to 6 bases long and 5 to 34 repetitions. A total of 97 SSR were successfully amplified on a first small set of *G .boninense* isolates from Indonesia. Then a population of 48 isolates from several locations in South-East Asia was screened to characterize each locus for allele number, heterozygoty and null allele absence. These results allow us to propose an efficient SSR set to study *G boninense* in infected oil palm plantations in order to better understand the history of this pathogen.

Keywords: *Ganoderma boninense*, oil palm, basal stem rot, microsatellite, genome sequence

1. Introduction

Ganoderma boninense is a telluric basidiomycete and main causal agent of the oil palm basal stem rot (BSR). This pathogen can impact up to 50% of trees in plantation from production loss to the death and tree fall (Corley and tinker, 2003). Main studies about the pathogen were

developed with the aim of improving selection for resistant oil palm progenies (Durand-Gasselin *et al.*, 2005; Idris *et al.*, 2004; Breton *et al.*, 2006) or developing disease management practices to limit disease spreading (Susanto *et al.*, 2005). The little knowledge we have about its physiology (Pilotti, 2005; Rees *et al.*, 2007) or genetic (Pilotti *et al.*, 2003; Miller *et al.*, 1999) is insufficient to explain its biology and diversity across an area as large as South-East Asia.

Microsatellite markers are the most popular and common molecular biology tool used to study species genetic diversity (Jarne and Lagoda, 1996). This is still the most powerful and used tool to study species population genetic and geographical spreading (Presti *et al.*, 2014). The enthusiasm for the use of SSR markers is due their advantageous features. They are widely and randomly distributed, highly polymorphic, co-dominant, reproducible and easy to amplify by PCR.

To improve our knowledge about *G. boninense* biology, it seems to be an important step to determine the actual diversity across its natural area of dispersion. Development of a microsatellite marker set appears to be a compulsory tool to better understand diversity of this pathogen. It can be a first step to improve our crop and disease management practices.

2. MATERIALS AND METHODS

2.1. DNA extraction

Tissues of fruiting bodies tissue and mycelium were crushed in powder and extracted with a classical MATAB protocol adapted from CTAB protocol (JJ. Doyle, 1987) (Tris HCl 1M ph 8,NaCl 5M, EDTA 500mM, MATAB, PEG 6000 and Sodium sulfite) with some modification to improve the DNA yield extraction. A first step of tissue digestion by GLUCANEX buffer (GLUCANEX 0.5g, NaCl 1M 70ml, H2O 30ml, buffer at ph 6) allows higher rate of DNA extraction and is capable of removing majority of the polysaccharide compounds. The MATAB protocol was also modified to 2 ml tubes extraction as compared to the original protocol to facilitate the extraction of a large set of samples. DNA of each sample was determined by a fluorometric method, Qubit® 2.0 Fluorometre from Invitrogen, to avoid measurement bias such as polysaccharides.

2.2. Genome sequencing

The draft genome sequence was produced with the DNA extracted from mycelium of an Indonesian isolate (NJ3) produced by PT SOCFINDO, grown in petri dishes on a PDA-chloramphenicol medium. A Combination of 454-GS-FLX (single reads with mean length of 378 bp) and Illumina HiSeq 2000 (paired reads with insert of 3 Kb and mean length of 50 bp) reads were used for contigs assembly. Assembly was performed using CLC Genomics Workbench 4.6.1 (CLC bio) software with standard parameters.

2.3. SSR development, primers design and validation

Microsatelite markers set were created using the Tandem Repeats Finder (TRF) software (http://tandem.bu.edu/trf/trf.html). The draft genome of *G. boninense* (NJ3) was screened for patterns of 2 to 6 bp long SSR motif with at least 5 repetitions. Primers were designed using the software Primer-BLAST (http://www.ncbi.nlm.nih .gov/tools/primer-blast/) with the following requirements: optimal primer size of 20 bp, product size 100-500 bp, primer melting temperatures Tm 57-63°C with optimal at 60°C, primers must not have more than 5 complementary bp and must be at least 2 bp away from the microsatellite locus. Each primer pair (forward and reverse) was tested with *in silico* PCR, using NJ3 genome sequence, with a local blastn to select primer pairs with a single match only. Primer pairs selected were generated with a M13 sequence added on the 5' extremity of the forward primer. This M13 sequence is complementary of the M13 fluoro-tagged primer.

A first set of 5 isolates from an Indonesian plantation grown on PDAc medium in petri dishes were used to determine potential polymorphism of SSR locus. Screening test was performed on Li-Cor 4300 automated DNA analyser with size standards IRDyes of 50-350 and 50-700 bp (Li-Cor Inc.). Only polymorphic microsatellites were used to screen the 44 fruiting bodies samples. PCR products were run on a 3500xl DNA Analyzer (Applied Biosystems) with LIZ 600 as the size standard.

2.4. Isolates harvesting and conditioning

Forty-four fruiting bodies were collected from Malaysian and Indonesian oil palm plantations. Geographical positions and distributions of samples are summarised in the table below (e.g. Table 1). Each sample was isolated from a single tree. Ten fruiting bodies originated from the mainland of Malaysia spread over two distant FELDA plantations. Meanwhile, 10 other fruiting bodies come from 2 FELDA plantations at Borneo. The 24 remaining fruiting bodies are distributed over 6 plantations in the PT SOCFINDO from North to South of Eastern Sumatra. Only dikaryotic tissue of fruiting bodies has been kept and freeze dried. Finally, all samples were packed with silica gel to avoid rehydration.

Country	Region	Location	Sample Name
Indonesia	Sumatra	Aceh North	SL_1 \ SL_18 \ SL_21 \ SL_27
Indonesia	Sumatra	Aceh South	SL_16 \ SL_22 \ SL_24 \ SL_25
Indonesia	Sumatra	Aek Loba	137B I_1 \ 25B I_1 \ 69B I_2 \ AKII I_1
Indonesia	Sumatra	Mata Pao	316A \ 334A \ 361A \ 371A
Indonesia	Sumatra	PSBB	204A \ 247A \ 284A \ 310A
Indonesia	Sumatra	Tanah Gambus	410A \ 417A \ 424A \ NJ3
Malaysia	Peninsula	Johor	112A \ 78A \ 85A \ 88A \ 96A
Malaysia	Peninsula	Negeri Sembilan	126A \ 135A \ 138A \ 141A \ 154A
Malaysia	Borneo	Sabah North	18A \ 26A \ 31A \ 32A \ 33A
Malaysia	Borneo	Sabah South	10A \ 14A \ 16A \ 5A \ 7A

Table 1. Geographic distribution of samples

2.5. PCR amplification and genetic analysis

Each microsatellite was amplified by simplex PCR in 10 µl reactions containing 5 ng DNA, 100 nM each of fluoro-tagged M13 (FAM, NED, VIC, PET) and reverse primer, 80 nM of

forward primer, 0.2 mM dNTPs, 0.6 mM MgCl2, 1x reaction buffer and 0.2 U QIAGEN Taq DNA polymerase. We added 5% DMSO and BSA at 0.8µg/µL to increase the separation of high GC% DNA region and reduce the effect of pollutant as polysaccharide on the primers hybridizations and abilities of Taq DNA polymerase.

All alleles were scored using the software Genemapper (version 4.1, Applied Biosystems). For each locus, number of alleles, observed and expected heterozygosities and imbreeding coefficient (F_{is}) was calculated using the software Genetix (Belkhir K *et al*, 2004). A test of significance of H_e compared to H_o have been performed (e.g. Table 3)

3. RESULTS AND DISCUSSION

Locus	Repeat motif	Size range (bp)	Primer sequences (5'-3')		T_a (°C)	
			Forward	Reverse	Forward	Reverse
7a	$(CAG)_6$	324-342	TCGGGTAGGCTCGCAGGTGG	GGGCCGCACAGGTCGAGAAA	59.3	59.8
15a	$(TTG)_5$	195-207	GGCGCACCATCCTTGTGGCT	GGCTTATGGCGGGTGTGCGT	60,0	60.0
18a	$(CAC)_6$	254-276	GCGTTGTCGGATGGACGGGG	TGCGGAGGTGGATGTCGGGT	60.1	59.9
19a	$(AG)_6$	201-208	GGCCGAGCCGAAAACGTCACA	GGGCGGAGCGAGTCAGTTGT	60.2	59.1
21a	$(CAG)_5$	150-156	TGCGTGCGTACCGGAGGTTT	CCCGGCCCGTCTCCTCTTCC	59.0	60.7
22a	$(CAG)_6$	198-216	TCGGAAGGCAGACGGACGGA	CTAGGCGTGCGTGGCAGAGC	59.6	60.5
29a	$(CT)_5$	256-272	GATGCCCCCGCCGATGCTAC	CCCTTGGTGGACGGTCTGCG	60,0	59.7
33a	$(CCA)_9$	166-175	AGCFCCCCTCCCAGCTCCAAC	GAATGCGGCGGGGAAACGGA	59.9	60.0
2b	$(CCACAA)_{10}$	331-358	ACACGCGGGCGGTTTCGTTC	TTGCTGTGAGCGAAGGCGGA	60.9	59,0
1c	$(GTC)_3$	218-224	CGCGGGCGAGCCTTTGACT	ACGAAAGCCGAGGCGCAGAA	60.0	59.4
2c	$(TGG)_3$	257-266	CGTTGGCGTGAGCAAGGGGT	CGGAGACGAAGCTGCCCTGC	59.9	60.1
11c	$(CGT)_3$	315-318	CGCTCTGCTCGGCACCTTCT	AGGAGGCTACGACCGCGAGTA	59.2	58.9
26c	$(CCGA)_3$	258-271	TCCCCAACCCGGCCAATCAC	TCGCGCTGTGTCGTGCAGAGT	59.0	61.3
3d	$(CA)_8$	190-232	AGCCCGCAGAGCCAATGTCC	GGCGTGCGACTGCTGGTTCA	59.4	60.3
11d	$(AGG)_8$	266-287	TCTGCGGTGGGGGTTGTGCT	GTCATCGCATCGGGGGTCGG	60.7	59.6
22d	$(GAG)_9$	217-223	GCGGGTGTGAAGGCGGAGG	GAGTCGCTCGCTCGCCCAC	60.0	60.2
24d	$(CAG)_{11}$	210-236	GCACAGGCACAAGCGCAAGG	CGACGACCGCCCCAAAGGAT	59.7	59.1
26d	$(GT)_8$	199-211	AACCCGGTGGTGCCCGACAA	GAACCCGACGAAGAAGCCGGTG	61.0	60.1
36d	$(GA)_8$	289-305	ATGCCTCCTTCGGGGCTTCCA	CGAGGACTTGCTGCCGCGTA	59.9	59.5
37d	$(CGG)_8$	323-338	TCCACGCCTGAGGGCCAGAA	TCGGAGTCGAAGCCACCCTTCA	59.8	59.6
40d	$(GTG)_8$	303-312	TGTCGTGGCTGTTGGCGTGT	CGACGGGATGGTTGGGCAGG	59.5	59.8
44d	$(GGT)_8$	281-305	TATGGGGTACGGGCGGACGG	CGTCCTTCCTCCGCTCTCGCT	60.1	60.4
46d	$(CT)_9$	270-294	GGGCTTCGGGCTCATGGCTG	GGGCGGCGGTGTGTTAGCTTT	60.1	60.2
47d	$(GA)_{10}$	258-288	GGCGACGAGGGCACGAGAGA	CCGCACTTTCGCCAACCACC	60.7	59.8
48d	$(TC)_9$	311-329	GGGGCCTCCTACATACTGCATTCC	GCGGGAGGGGATCGAGGTCA	58.6	59.8
52d	$(GA)_8$	303-313	CCTTGTGGTGGTAGGCGGCA	ATCTCCGTGGCACCGCATCCT	59.0	60.3

Table 2. Primer sequences, locus size range and PCR condition of 26 microsatellite markers developed in *G. boninense*.

3.1. Characterization of microsatellite markers in *G. boninense*

Based on the genome analysis, we had identified a total of 652 SSR. We succeeded to develop 145 SSR primers to test on 5 Indonesian isolates. Of which, only 97 SSR were successfully amplified and produced reproducible mono-locus banding pattern. Forty-eight microsatellite markers were observed as polymorphic. As these 5 isolates came from a reduced geographical area, and obviously represent a very small subset of the wide distribution zone of

G.boninense, it is possible that more markers are polymorphic using samples geographically more distant.

The 48 SSR have been amplified on the larger *G. boninense* isolate collection of 44 samples and genotyped on the ABI sequencer. Only 26 SSR markers have shown good amplifications (Migration patterns with peak over 200 RFU, Relative Fluorescent Unit, and no more than two amplified peak). The twenty-two remaining markers may have not been successfully amplified because of mutation in the genomic sequences corresponding to the markers primers, or due to a locus positioned in GC rich genomic region. A duplication of the genomic region of a marker or a migration artefact could also explain cases of supernumerary peaks. All characteristics of the 26 well amplified and polymorphic SSR markers have been summarized in the table above (e.g. Table 2). The sizes of repeat motifs are distributed between di-nucleotide and hexa-nucleotides. There are 9 di-nucleotides, 15 tri-nucleotides, 1 quadri-nucleotide and 1 hexa-nucleotide. Motifs varied from 3 to 11 repetitions with a size of PCR products ranging from 150 bp to 358 bp. All primers have an annealing temperature between 58.9 and 61.3°C and a difference between forward and reverse of maximum 2.3°C.

3.2. Genotyping and population genetics analysis

Genotype characteristics and genetic variations among the three populations have been

Locus	Indonesia (Sumatra)				Malaysia (Peninsula)				Malaysia (Borneo)				Total			
	Na	H_e	H_o	F_{is}	*Na*	H_e	H_o	F_{is}	*Na*	H_e	H_o	F_{is}	N_a	H_o	H_e	F_{is}
7a	5	0.3378	0.3333	-0.024	3	0.5033	0.4444	0.098	2	0.4945	0.1429	0.830	6	0.3250	0.4263	0.185
15a	4	0.3541	0.1364	0.563	2	0.5333	0.0000	1.143	------	------	------	------	4	0.1000	0.4379	0.623
18a	4	0.3962	0.0500	0.959	5	0.8235	0.2222	0.842	4	0.7879	0.0000	1.200	8	0.0857	0.6762	0.935
19a	5	0.5922	0.3750	0.126	3	0.6275	0.3333	0.328	2	0.2637	0.0000	1.167	5	0.3000	0.5595	0.147
21a	3	0.5877	0.1818	0.642	2	0.4000	0.0000	1.143	2	0.4396	0.0000	1.167	3	0.1081	0.5539	0.809
22a	6	0.6936	0.4500	0.355	5	0.7253	0.5714	0.148	1	0.0000	0.0000	------	6	0.4194	0.7129	0.288
29a	5	0.6774	0.1250	0.987	2	0.5357	0.2500	0.711	2	0.5333	0.0000	1.500	5	0.1304	0.6290	0.969
33a	3	0.5363	0.3333	0.212	4	0.7386	0.6667	0.282	3	0.4394	0.1667	0.589	4	0.3846	0.5771	0.449
2b	9	0.8032	0.3333	0.503	6	0.8791	0.0000	1.167	2	0.6667	0.0000	2.000	10	0.2222	0.8316	0.667
1c	2	0.4220	0.1667	0.631	2	0.2092	0.2222	-0.070	3	0.2417	0.2500	-0.005	3	0.1951	0.3520	0.247
2c	3	0.2339	0.0000	1.067	1	0.0000	0.0000	------	1	0.0000	0.0000	------	3	0.0000	0.1596	1.043
11c	2	0.4034	0.2083	0.505	2	0.1895	0.2000	-0.062	2	0.1895	0.2000	-0.062	2	0.2045	0.3153	0.359
26c	2	0.3369	0.3333	0.011	3	0.3072	0.3333	-0.044	2	0.3660	0.0000	1.125	3	0.2619	0.3313	0.119
3d	6	0.8156	0.7083	0.081	6	0.8039	0.2222	0.670	3	0.6212	0.1667	0.938	7	0.5128	0.8192	0.408
11d	4	0.6011	0.4583	0.118	2	0.1111	0.1111	0.000	4	0.8000	0.6667	0.111	7	0.3889	0.5379	0.079
22d	2	0.4220	0.1667	0.631	2	0.1111	0.1111	0.000	3	0.6071	0.2500	0.633	3	0.1622	0.3850	0.783
24d	8	0.7668	0.2500	0.685	5	0.7974	0.7778	0.062	5	0.6667	0.2500	0.488	8	0.3659	0.8199	0.659
26d	4	0.6028	0.0625	0.980	3	0.6667	0.0000	1.143	1	0.0000	0.0000	------	4	0.0400	0.6310	0.998
36d	6	0.6873	0.2778	0.481	4	0.6417	0.5000	0.101	3	0.7143	0.0000	1.333	7	0.3000	0.7023	0.532
37d	6	0.8016	0.6667	0.228	4	0.7333	0.8000	-0.063	2	1.0000	1.0000	------	6	0.7083	0.7766	0.190
40d	5	0.6924	0.2500	0.420	3	0.6601	0.1111	0.960	2	0.6667	0.0000	2.000	5	0.2000	0.7408	0.546
44d	4	0.6956	0.2632	0.721	3	0.6593	0.0000	1.167	3	0.8000	0.0000	1.500	6	0.1724	0.7508	0.873
46d	6	0.8782	0.1333	0.835	5	0.8485	0.0000	1.200	2	0.5333	0.0000	1.500	8	0.0833	0.8511	0.932
47d	7	0.7130	0.4348	0.245	3	0.6209	0.1111	0.842	4	0.7778	0.2000	0.750	9	0.3243	0.7234	0.354
48d	9	0.8511	0.7500	0.169	6	0.7778	0.4444	0.364	5	0.8000	0.2000	0.935	10	0.5581	0.8438	0.313
52d	3	0.5807	0.1739	0.686	4	0.5425	0.2222	0.594	5	0.7250	0.3750	0.451	5	0.2250	0.6984	0.774

Table 3. Genetic variations at 26 polymorphic microsatellite markers in 3 populations of *G. boninense*. Locus name, number of alleles (N_a), observed (*Ho*) and expected (*He*) heterozygosities and F_{is} calculation based on Weir & Cockerheim (1984) and Robertson & Hill (1984).

described. Each population presents between 1 and 9 alleles per locus, whereas the total alleles per locus varied from 2 to 10 alleles (e.g. Table 3). The low allele number in population, specifically for Borneo and Peninsular Malaysia populations, could be a bias due to the unequal number of samples between populations. Regarding level of global heterozygosity for each locus, the observed heterozygosity (H_o) varied from 0.0 to 0.708 with a mean H_o of 0.2607.The global expected heterozygosity for each locus (H_e including a correction for sampling bias) ranged between 0.1596 and 0.8511 with a mean value of 0.60 and an associated p-value at 95% (P(0.95)) of 1. This result show a significant p-value meaning that there is an evident deficit of heterozygosity in the global population. For the 3 sub-populations, global H_o varied between 0.1547 and 0.2932, and H_e varied between 0.4481 and 0.5810 with significant P(0.95) 1.0 and 0.9615 respectively for the Indonesian and Peninsular Malaysia populations. Complementary test which are not included in this study showed a sub-structure inside those populations which could probably explain the deficit of heterozygocity observed for all locus. A presence of null allele can also be a possible factor which could contribute to this phenomenon. Further analyses have to be proceed with an enlarge sampling, both from numerical and geographical viewpoint, to permit a better comprehension of the allele distribution across the *G. boninense* living area.

4. CONCLUSION

We successfully developed 26 markers from genomic sequence of *G. boninense*. Those markers represent an essential tool for genetic analysis of this main oil palm pathogen. It will facilitate population genetics study of the pathogen across oil palm planting regions and could be a cornerstone to the comprehension of the actual diversity of the pathogen. Furthermore, those SSR markers could become a key tool for breeding programs by facilitating traceability of isolates for oil palm early screening test in nursery.

5. REFERENCES

Breton F, Hasan Y, Hariadi S, Lubis Z, De Franqueville H, 2006. Characterization of parameters for the development of an early screening test for basal stem rot tolerance in oil palm progenies. Journal of Oil Palm Research (Special issue, April 2006) 24–36.

Corley, R.H.V. and P.B. Thinker. 2003. The oil palm, 4th edition. Blackwell Science Ltd, United Kingdom.

Doyle, J. J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem bull, 19, 11-15.

Durand-Gasselin, T., Asmady, H., Flori, A., Jacquemard, J. C., Hayun, Z., Breton, F., & De Franqueville, H. 2005. Possible sources of genetic resistance in oil palm (Elaeis guineensis Jacq.) to basal stem rot caused by Ganoderma boninense–prospects for future breeding. Mycopathologia, 159(1), 93-100.

Idris, A. S., Kushairi, A., Ismail, S., & Ariffin, D. 2004. Selection for partial resistance in oil palm progenies to Ganoderma basal stem rot. Journal of Oil Palm Research, 16(2), 12-18.

Miller, R. N. G., Holderness, M., Bridge, P. D., Chung, G. F., & Zakaria, M. H. 1999. Genetic diversity of Ganoderma in oil palm plantings. Plant Pathology, 48(5), 595-603.

Pilotti, C. A., Sanderson, F. R., & Aitken, E. A. B. 2003. Genetic structure of a population of Ganoderma boninense on oil palm. Plant pathology, 52(4), 455-463.

Pilotti, C. A. 2005. Stem rots of oil palm caused by Ganoderma boninense: Pathogen biology and epidemiology. Mycopathologia, 159(1), 129-137.

Presti, F. T., & Wasko, A. P. 2014. A Review of Microsatellite Markers and their Application on Genetic Diversity Studies in Parrots. Open Journal of Genetics, 4(02), 69.

Rees, R. W., Flood, J., Hasan, Y., & Cooper, R. M. 2007. Effects of inoculum potential, shading and soil temperature on root infection of oil palm seedlings by the basal stem rot pathogen Ganoderma boninense. Plant pathology, 56(5), 862-870.

Susanto, A., Sudharto, P. S., & Purba, R. Y. 2005. Enhancing biological control of basal stem rot disease (Ganoderma boninense) in oil palm plantations. Mycopathologia, 159(1), 153-157.