Journal of Oil Palm Research Vol. 30 (4) March 2018 p. 36 – 46
DOI: https://doi.org/10.21894/jopr.2018.0009

VALIDATION OF DIFFERENTIAL GENE EXPRESSION OF TRANSCRIPTOME ASSEMBLY VIA NANOSTRING® TECHNOLOGIES ANALYSIS PLATFORM

NUR QISTINA OTHMAN*; PUVANESWARI PARAVAMSIVAM*; JOON SHEONG TAN*; YANG PING LEE*; YEN YEN KWAN* and SHARIFAH SHAHRUL RABIAH SYED ALWEE*

ABSTRACT

Ganoderma boninense *is a white rot fungus that causes basal stem rot (BSR) disease that contributes to profit loss in the oil palm industry. Transcriptome profiling using RNA-sequencing (RNA-seq) approach had been conducted to discover the differential express transcripts which relate to BSR disease that might potentially be utilised as biomarker for selection of* Ganoderma *tolerant oil palm. To validate the expression of transcripts in medium throughput manner after large scale transcriptome profile, new approaches are available such as reverse transcription quantitative PCR (RT-qPCR) of Fluidigm's BioMark HD system that runs on microfluidic chips, and NanoString's nCounter system that gives absolute counts of the mRNA itself without involving cDNA synthesis and PCR amplification. In this study, differential express transcripts were identified from the comparative transcriptomic of the transcriptome of near-rot section of basal stem tissue of oil palm infected with* G. boninense *against healthy non-infected palms. Then, whether the transcript abundances measured using RNA-seq method was correlated with the nCounter system for the selected transcripts was assessed. Based on the selected 24 transcripts, transcript abundances measured was significantly correlated between RNA-seq and nCounter system ($p < 0.001$). Out of 24 assessed transcripts, three transcripts were highly similar in the measured expression profiles using both platforms. From the three transcripts, only two showed potential as biomarkers for detecting* G. boninense *tolerant trait in oil palm.*

Keywords: *Ganoderma boninense*, expression validation, biomarker.

Date received: 11 April 2017; **Sent for revision:** 17 May 2017; **Received in final form:** 25 October 2017; **Accepted:** 26 January 2018.

INTRODUCTION

African oil palm (*Elaeis guineensis* Jacq.) is one of the main sources of vegetable oil, which is used for both food and non-food industries. Nowadays, the African oil palm is planted mainly in tropical belt, particularly in South-east Asia countries. The top palm oil producing nations are Indonesia, Malaysia, Thailand, Colombia and Papua New Guinea. Oil palm industry is a major contributor to the economic growth of Malaysia. However, incidences of root and stem rot have progressively resulted in huge profit loss (Rees *et al*., 2007). Basal stem rot (BSR) of oil palm has shortened the productive life of the palms and hence constrained the sustainable oil palm production in Asia (Durand-Gasselin *et al*., 2005).

* Felda Global Ventures Research & Development Sdn Bhd, Molecular Research Laboratory, Felda Global Ventures Innovation Centre, PT 23417 Lengkuk Teknologi, 71760 Bandar Enstek, Negeri Sembilan, Malaysia.
E-mail: nqistina.o@feldaglobal.com

Ganoderma boninense, a white-rot fungus, has been identified as the major cause of BSR disease (Khairuddin, 1990; Rao, 1990). *G. boninense* is classified under the division Basidiomycota and the genus *Ganoderma* which causes both BSR and upper stem rot in plant (Seo and Kirk, 2000). *G. boninense* is capable of degrading the plant cell wall and disrupts water and nutrient transport to the upper part of the palm thus causing yellowing of leaves, frond wilting, unopened spear leaves and emergence of basidiocarps on the lower stem (Chung, 2011). Prior to infection, *G. boninense* hyphae attach firmly at the root surface of the palm. Subsequently, the hyphae attachment either remains at the initial point of contact or completely covers the surface of the root. Then penetration of epidermis and exodermis occurs, and followed by internal invasion of the fungus hyphae into the root system which progressively decay the roots and lead to lower stem rot (Rees *et al*., 2009). Successful penetration of *G. boninense* into the outermost tissues and cell wall requires the cell wall-degrading enzymes such as laccase, manganese peroxidase and lignin peroxidase (Cooper, 1984; Janusz *et al*., 2012). An infected oil palm may or may not bear any fruiting bodies. The presence of fruiting body on an infected palm normally shows that the oil palm has been infected for years (Najmie *et al*., 2011). It is almost impossible to maintain a pathogen-free field. Therefore, integrated sanitisation programme is deployed to mitigate the disease incidence in the plantation such as removing diseased material, ploughing, fallowing, planting legume cover crops, chemical and bio-control treatments as well as beneficial microbes such as *Trichoderma* species (Lee *et al*., 2016). Although combination between sanitisation and agronomic management can reduce the impact of the *G. boninense* disease, tolerant planting materials is needed for the long-term solution for BSR (Hushiarian *et al*., 2013).

Seedlings were challenged by artificially inoculated with *Ganoderma* in nursery trial in order to identify tolerant planting material (Idris *et al*., 2004). This process requires tedious work and effort. The easiest and fastest screening method will be utilisation of molecular marker to screen for tolerant palms. Tolerant genes in the oil palm can be detected through the transcriptomics and genomics exploration using next generation sequencing. With the recent advancements in the next generation sequencing technologies, comparative transcriptomics approach using RNA-Seq becomes feasible with the aim to measure the global response of gene expression in specific changes response to abiotic (Cohen *et al*., 2010; Puranik *et al*., 2011) and biotic stresses (Ho *et al*., 2016; Liu *et al*., 2016). Comparative transcriptomics approaches using RNA-Seq method allow researchers to compare and contrast the transcriptomes of two samples to identify differential express transcripts and pathways. To validate a small subset of the differential express transcripts, RT-qPCR method can be used. However, to validate a larger scale of the differential express transcripts, a number of medium throughput PCR-based applications have been developed, such as emulsion-based digital PCR (Kanagal-Shamanna, 2016) technologies employed by Fluidigm Corporation's BioMark HD system, Life Technologies' OpenArray Realtime PCR System, Bio-Rad Laboratories' QX100 ddPCR System and RainDance's RainDrop Digital PCR (Baker, 2012). In addition, direct quantification of target mRNA using hybridisation method such as nCounter Analysis system of NanoString® Technologies (Geiss *et al*., 2008) is becoming popular as it offers multiplex profiling of up to 800 transcripts per sample. It assesses the medium throughput expression analysis via digital molecular barcode technology which gives an absolute count of the mRNA itself without involving the synthesis of cDNA and PCR amplification that might introduce bias (Kulkarni, 2011).

Studies conducted to date have identified a handful of candidate genes responsible for defense mechanisms in host plants. Various potential transcript sequences which are involved in defense mechanism through pathogen related proteins in oil palm such as defensins, glucanases and chitinases have been reported (Yeoh *et al*., 2012; Tan *et al*., 2013; Naher *et al*., 2013). Majority of the gene expression studies were conducted on root or leaf samples of young oil palm seedlings. However, to our best knowledge, no gene expression profile or transcriptomic study using *G. boninense* infected basal stem tissue of matured oil palm has been reported. This is due to difficulty in obtaining high quality total RNA for RT-qPCR application from rot or near-rot tissues of matured oil palm. Therefore, nCounter analysis system from NanoString® Technologies was used as validation method as nCounter analysis system is sensitive to low quality total RNA templates (Geiss *et al*., 2008).

In this study, we tested whether nCounter analysis system was suitable to validate a subset of differential express transcripts that were identified from our initial effort to profile the transcriptome of near-rot section of basal stem tissue of oil palm infected with *G. boninense* (ITIP) or the non-rotten, healthy section of basal stem tissue of the same *G. boninense* infected palm (HTIP) that compared with the transcriptome of healthy section of basal stem tissue of the healthy palm (HTHP). Then, further analysis of the expression pattern of a small set of selected transcripts using RT-qPCR method for samples derive from oil palm seedlings infected with *G. boninense* was done to detect transcript expression signatures which enable us to identify candidate biomarkers of tolerant oil palm.

MATERIALS AND METHODS

Plant Materials

A pair of healthy and infected clonal mature palm was selected from field. The bark was marked according to the health status. The healthy section (HTIP) and near-rot section (ITIP) of infected oil palm were collected from the cross-section of the basal stem. The cross-section of healthy tissue section of healthy oil palm (HTHP) was also collected. The healthy palm was determined by healthy oil palm phenotype with no *G. boninense* fruiting body attached at the basal stem.

All collected basal stem tissues were split into two halves and immediately submerged in liquid nitrogen and kept at -80°C freezer until further processing for ribonucleic acid (RNA) extraction. The first half tissues were used for transcriptome sequencing and the second half tissues were used for nCounter analysis system.

Nine-month old clonal plantlets were artificially infected with *G. boninense* using sitting technique (Idris *et al*., 2004) at a nursery. After nine months of inoculation, phenotypically identified potential tolerant (PT) and potential susceptible (PS) plantlets were sampled (*Figure 1*). Healthy plantlets post infection was classified as PT palms while plants with severe infection symptom were considered as PS palms. Roots, leaves, stems and meristems of 3 PT and 3 PS individuals were sampled. All sampled tissues were immediately submerged into liquid nitrogen and kept at -80°C freezer until RNA extraction.

RNA Extraction

Total RNA of matured basal stem samples; HTIP, ITIP and HTHP were extracted using protocol from Saidi *et al*. (2009). The extracted RNA were used for Illumina's HiSeq 2000 transcriptome sequencing. For nCounter expression analysis and reverse transcription quantitative polymerase chain reaction (RT-qPCR) analysis, total RNA of leaves, roots, stems and meristems of 9-month-old oil palm clonal PT and PS plantlets were extracted using RNeasy Plant Mini Kit (Qiagen, Germany) by following the manufacturer's protocol. The quality of the extracted RNA was determined using Bioanalyzer (Agilent Technologies, USA) and UV absorbance readings (NonoDrop, Thermo Fisher Scientific Inc., USA).

Transcriptome Sequencing

Messenger RNA isolation and cDNA synthesis were performed using TruSeq RNA Sample Preparation Kit (Illumina, USA) and SuperScript II Reverse Transcriptase (Invitrogen, USA) by following the manufacturers' protocol, respectively. The amount of synthesised cDNA was measured using Qubit 2.0 DNA Broad Range Assay (Invitrogen, USA). A minimum of 15 ng cDNA was fragmented using Covaris S220 (Covaris Inc, USA) to a targeted size of 200 - 300 bp. The fragmented cDNA was then end repaired, ligated to Illumina TruSeq adapters, and PCR-enriched using TruSeq RNA Sample Preparation Kit (Illumina, USA) by following the manufacturer's protocol. The final sequencing

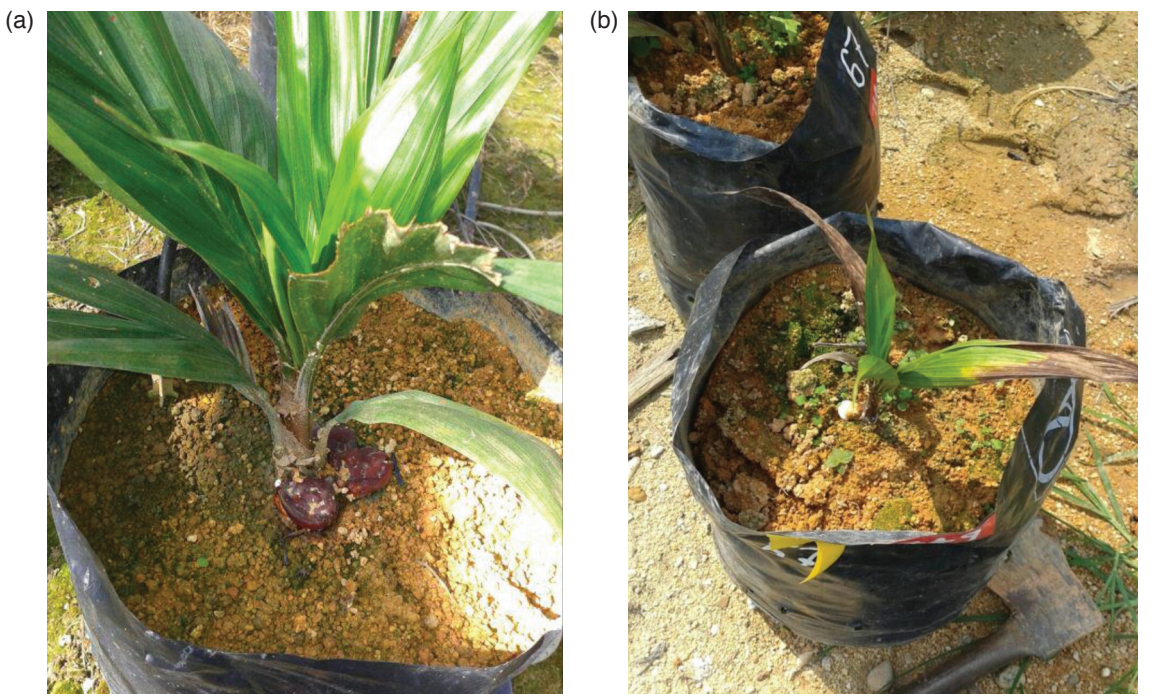


Figure 1. a) Phenotypically identified potential tolerant (PT) and b) potential susceptible (PS) plantlets.

libraries were quantified using KAPA kit (KAPA Biosystem, USA) on Agilent Stratagene Mx-3005p quantitative PCR (Agilent Technologies, USA) and sizes were confirmed using Agilent BioAnalyzer High Sensitivity DNA Chip (Agilent Technologies, USA). The resulting libraries were sequenced using an Illumina flow cell and 209 cycles on the Illumina HiSeq 2000 platform (Illumina, USA).

De novo Assembly, Mapping of Short Reads and Identification of Differentially Expressed Transcripts

Adapter sequences were trimmed using Trimmomatic (Bolger *et al*., 2014). PRINSEQ (Schmieder and Edwards, 2011) was used to remove low quality reads (nucleotide quality score < 20) and sequences with ambiguous bases (N). An oil palm reference unigene was assembled from cleaned illumine paired-end reads of pooled HTHP, ITIP and HTIP sequencing libraries using Trinity package (Grabherr *et al*., 2011). The cleaned paired-end reads were then aligned back to reference unigene (length ≥ 300 bp) using Bowtie (Langmead *et al*., 2009) as part of in-house Perl scripts. Transcripts abundance was estimated using RSEM (Li and Dewey, 2011) with default parameter settings. Differential express analysis between samples was conducted using DESeq (Anders and Huber, 2010) with default parameter settings. Observed differential express transcripts were then scanned for infection-related association or defense-related annotations based on publicly available database (NCBI, nr database).

Counter Expression Analysis

Twenty-four transcripts found to be related to disease resistance and stress-related mechanism were selected from the bioinformatics analysis and submitted in advance to NanoString® Technologies, Inc. (Seattle WA USA) to design and synthesise a CodeSet (multiplex colour coded probe set containing both control and target probes) before RNA analysis. Probes were designed to specific target of desired transcripts. The CodeSet also included probes of six negative and six positive control targets, as well as for seven internal reference genes. Hybridisation was conducted on 100 ng purified RNA according to manufacturer's instructions. The results data consisted of direct molecule counts present in the sample as determined by the number of bar-coded probes detected at the end of the process. These data were compiled in Microsoft Excel and analysed using the nSolver Software (NanoString®, Seattle WA, USA). Fold-change in the expression data obtained was calculated relative to the control sample and averaged over the biological replicates and then was $\log_2$-transformed. Transcripts expression measurement generated by nCounter was then compared against differential expression data generated by bioinformatics analysis. Non-parametric Spearman's rank-order correlation between RNA-Seq data and nCounter data was performed to assess whether any correlation was found between the expression of transcripts measured using RNA-seq and nCounter analysis system and suitability of nCounter analysis platform to validate the expression data. Non-parametric Spearman's rank-order correlation method was selected over the parametric Pearson's linear-correlation method because in our case, the expression data is not normally distributed (data not shown), an assumption that data normally distributed is required for parametric method.

RT-qPCR Analysis

Potential transcripts that had the same expression pattern in transcriptome sequencing and nCounter analysis system were further analysed using RT-qPCR. DNase treated-RNA samples were reverse transcribed using Superscript (III) First-Strand Synthesis System (Invitrogen, USA) by following the manufacturer's protocol. RT-qPCR was performed on Light Cycler® 96 Real-Time PCR (Rosche, USA). A cDNA dilution series was prepared to determine the amplification efficiency. Primer specificity was verified by analysing the melting curve at the end of each PCR. Each PCR reaction (10 μl) comprised of 2 µl of the first-strand DNA, 1.5 μl of each forward and reverse primers and 5 μl of 2X Light Cycler ® 480 SYBR Green I Master Mix. The transcripts ID and primers designed for RT-qPCR are listed in *Table 1*. The PCR parameters were 95°C for 10 min followed by 45 cycles of 95°C for 15 s and 60°C for 30 s. Two endogenous controls transcripts, SLU7 and GRAS were used in parallel for normalisation (Yeap *et al*., 2014). The relative transcript level was calculated based on a comparative C_T method using two reference genes for normalisation.

RESULTS AND DISCUSSION

Reference Transcriptome Assembly and Expression Analysis

With only one matured palm infected with *G. boninense* available at the time of sampling period, we have collected only one biological replicate of infected and non-infected samples. Deep sequencing of the RNA-seq libraries was conducted with Illumina HighSeq 2000 platform to achieve appropriate transcriptomic information for single biological replicate (Liu *et al*., 2013). More than 200 million raw reads were obtained from each of sequencing libraries (*Table 2*). An oil palm reference unigene was assembled based on the

pooled of cleaned Illumina reads of ITIP, HTIP and HTHP sequencing libraries and resulted in 115 613 unigenes (*Table 2*). Out of 115 613 unigenes, about 42.3% and 43.53% transcripts were differential express (Log_2 fold-changed > 0 or < 0, *p*-value > 0.05) under HTIP/HTHP and ITIP/HTHP transcriptome comparison, respectively using DESeq (*Table 3*). These unfiltered differentially expressed transcripts (*p*-value > 0.05) were our initial effort to investigate the transcriptomic changes of near-rot tissues. These unfiltered transcripts were then blast against public database (NCBI, nr database), and scanned for biotic-stress-related and defense-related annotation. Dozens of defense or resistant genes have now been cloned and characterised from a variety of plants (McDowell and Woffenden, 2003) such as *RPS2* gene in *Arabidopsis*, *N* gene in tobacco, *Hm1* gene in maize and *Cf9* gene in tomato (Staskawicz, 2001). Putative defense-related genes from oil palm can be predicted through these HTIP/HTHP and ITIP/HTHP sequence data sets.

Validation of RNA-Seq by NanoSrting® nCounter Analysis System

Comparing the expression data derived from RNA-Seq method with nCounter analysis system become a wide use and acceptable approach for data

TABLE 1. PRIMERS DESIGNED FOR THE QRT-PCR AMPLIFICATION OF SELECTED TRANSCRIPTS

Primer ID	Transcript ID	Putative identity	Sequence (5′ – 3′)
P1	Felda_Reference_Palm_ Transcript_63694_2380	Osmotin-like protein	F: TATCAGCTTCAGGAAATAAGACGG R: TTTGTAGTAGCTCCTTGGATTTGG
P3	Felda_Reference_Palm_ Transcript_59423_2809	Cysteine-rich receptor-like protein kinase	F: CAAAGAAATTGAGGCAGGGA R: ATTGGGTGATGAAGGAACGG
P4	Felda_Reference_Palm_ Transcript_69361_646	Cysteine proteinase inhibitor	F: CGGCATCTGGCAAGATATTGAG R: AACACATAGGCAAAGGCAAGTC

Note: qRt-PCR – quantitative reverse transcription–polymerase chain reaction.

TABLE 2. OVERVIEW OF SEQUENCING AND TRANSCRIPTOME ASSEMBLY OF ILLUMINA READS DERIVED FROM TRUNK TISSUES OF OIL PALM BASAL STEM INFECTED WITH PATHOGENIC *G. boninense* ISOLATE

Sequencing library	Number of raw reads[a]	Read length (bp)	Clean reads[b]	Number of clean paired-end reads	Number of assembled transcripts	N50[c] (bp)	Mean length[d] (bp)	Mapped reads[e]
Felda_Palm_HTHP	224 285 692	100	164 051 867	68 983 603	NA	NA	NA	58 837 807
Felda_Palm_HTIP	276 546 920	100	198 566 601	82 140 764	NA	NA	NA	69 979 926
Felda_Palm_ITIP	214 112 784	100	172 512 674	75 459 523	NA	NA	NA	63 201 291
Felda_Reference_ Transcript[f]	714 945 396	100	535 131 142	226 583 890	115,613	1,350	929	NA

Note: [a]Number of combined pair-end raw reads.
[b]Number of clean pair-end reads after removed low quality bases and adaptor sequences.
[c]N50 length of assembled transcripts.
[d]Mean length of assembled transcripts.
[e]Number of clean paired-end reads mapped to assembled reference transcripts.
[f]The reference transcripts were *de novo* assembled from combined paired-end reads from three oil palm libraries of HTHP, HTIP and ITIP.
NA - not available.
HTHP - healthy section of basal stem tissue of the healthy palm.
HTIP - non-rotten, healthy section of basal stem tissue of the *G. boninense* infected palm.
ITIP - near-rot section of basal stem tissue of oil palm.

TABLE 3. SUMMARY OF REFERENCE TRANSCRIPTOME ASSEMBLY OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS AT DIFFERENT CONDITIONS

Condition 1	Condition 2	Up regulated (condition 1> condition 2)	Down regulated (condition 1< condition 2)
HTHP	HTIP	24 400	24 464
HTHP	ITIP	25 027	25 294

Note: HTHP - healthy section of basal stem tissue of the healthy palm.
HTIP - non-rotten, healthy section of basal stem tissue of the *G. boninense* infected palm.
ITIP - near-rot section of basal stem tissue of oil palm.

validation (Prokopec *et al*., 2013; Tam *et al*., 2014). We decided to assess and validate our expression data of 24 transcripts estimated by RNA-seq method with a probe-based hybridisation method of nCounter analysis system (NanoString® Technologies, Seattle, WA, USA). Given the fact that our RNA samples for RNA-seq experiments was extracted from the near-rotten basal stem tissues and the average RNA Integrity Number (RIN) was 7.6 which is within the accepted range. RIN is a RNA quality metrics used to define sample quality (Gallego Romero *et al*., 2014). The cut-off RIN value between 7.9 and 6.4 is suggested to be used for partially degraded RNA samples (Gallego Romero *et al*., 2014).

Criteria for the 24 transcripts selected for nCounter validation were based on the BLAST-search-based annotations related to composition and degradation of cell wall components, lignin biosynthesis, secondary metabolism, degradation related to hormone metabolism, biotic and abiotic stress related proteins, signaling receptor kinase, RNA regulation of transcription and enzymes such as peroxidases and glucosidases (Janusz *et al*., 2012).

Recent studies have showed that expression values estimated using RNA-seq method was highly correlated with RT-qPCR method (Griffith *et al*., 2010; Shi and He, 2014; Wu *et al*., 2014; Rajkumar *et al*., 2015). RT-qPCR is a method of choice to estimate the gene expression values provided that the input RNA samples are of high RNA integrity (Fleigea and Pfaffl, 2006). NanoString® nCounter system is a preferred method when the input RNA sample is from formalin-fixed paraffin embedded specimens (Saba *et al*., 2015) that the total RNA is partially degraded or not in an ideal quality (Tachibana, 2015). The advantages of nCounter system over the RT-qPCR are that the nCounter system is directly count or quantitate the gene expression levels without involving PCR steps, and only 100 ng of total RNA is needed to accurately measure the expression of a few hundred targeted genes per sample (Geiss *et al*., 2008; Kulkarni, 2011).

A total of 24 transcripts were selected for nCounter expression validation, 19 transcripts were derived from HTIP/HTHP (*Table 4*) comparison and 22 transcripts were derived from ITIP/HTHP comparison (*Table 5*). There were 17 overlap transcripts between the HTIP/HTHP and ITIP/HTHP conditions. The Spearman's rank-order correlation analysis showed that the log2 fold-change values derived from RNA-seq and nCounter system were significantly correlated for both HTIP/HTHP ($rho = 0.9035$, $p < 2.2e\text{-}16$; *Figure 2*) and ITIP/HTHP ($rho = 0.6849$, $p < 0.001$; *Figure 3*).

Although the fold-change direction was consistent, the magnitude of the expression fold-change differed when compared with the expression values derived from RNA-seq method with nCounter system (*Tables 4* and *5*). The observed

TABLE 4. VERIFICATION OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS USING NANOSTRING PLATFORM IN HTIP/HTHP CONDITION

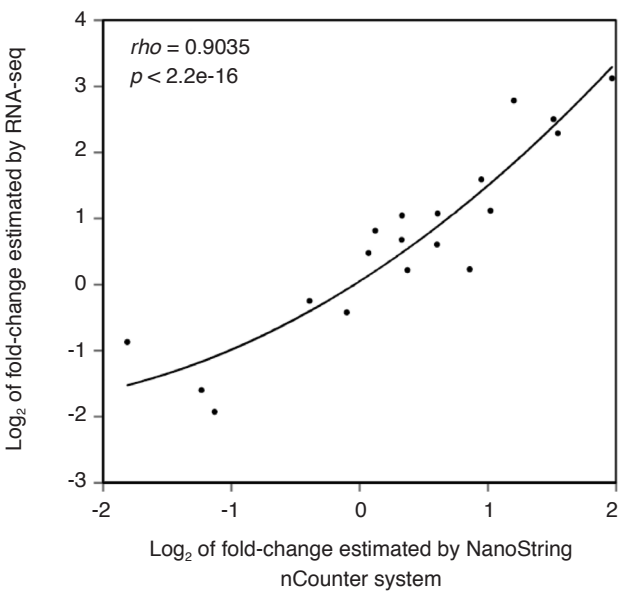
Transcript ID	Putative identity	Fold change of expression using	
		Bioinformatics analysis in HTIP/HTHP	NanoString® platform condition
Felda_Reference_Palm_Transcript_48685_3389	Lipoxygenase family protein	0.23	0.37
Felda_Reference_Palm_Transcript_27311_977	*Trans*-cinnamate 4-monooxygenase (Cinnamic acid 4-hydroxylase)	-0.87	-1.81
Felda_Reference_Palm_Transcript_42818_2273	Membrane-anchored cell wall protein	0.82	0.12
Felda_Reference_Palm_Transcript_69361_646	Cysteine proteinase inhibitor	3.13	1.97
Felda_Reference_Palm_Transcript_56763_3628	Pathogen-induced transcription factor	1.60	0.95
Felda_Reference_Palm_Transcript_33570_1281	Leucine-rich repeat class of receptor-like kinase	-0.24	-0.39
Felda_Reference_Palm_Transcript_49790_630	Leucine-rich repeat (LRR) family protein	1.07	0.61
Felda_Reference_Palm_Transcript_60805_2316	Subfamily of ERF/AP2 transcription factor family	2.79	1.20
Felda_Reference_Palm_Transcript_28152_1762	Leucine-rich repeat transmembrane protein kinase	-1.59	-1.23
Felda_Reference_Palm_Transcript_63698_2058	Beta-glucanase isoenzyme	2.30	1.55
Felda_Reference_Palm_Transcript_59423_2809	Cysteine-rich receptor-like protein kinase	1.13	1.02
Felda_Reference_Palm_Transcript_39245_1096	4-coumarate--CoA ligase 2	-0.42	-0.10
Felda_Reference_Palm_Transcript_54436_1392	ATP binding factor	0.23	0.86
Felda_Reference_Palm_Transcript_30872_3213	Cellulose synthase	-1.92	-1.13
Felda_Reference_Palm_Transcript_40184_1989	WRKY transcription factor family	0.48	0.07
Felda_Reference_Palm_Transcript_46069_878	WRKY transcription factor family	0.68	0.33
Felda_Reference_Palm_Transcript_63694_2380	Osmotin-like protein	2.51	1.51
Felda_Reference_Palm_Transcript_49323_992	Dehydration responsive protein	0.61	0.60
Felda_Reference_Palm_Transcript_47222_907	Xyloglucan endotransglycosylase-related protein	1.05	0.33

Note: HTHP - healthy section of basal stem tissue of the healthy palm.
HTIP - non-rotten, healthy section of basal stem tissue of the *G. boninense* infected palm.

TABLE 5. VERIFICATION OF DIFFERENTIALLY EXPRESSED TRANSCRIPTS USING NANOSTRING PLATFORM IN ITIP/HTHP CONDITION

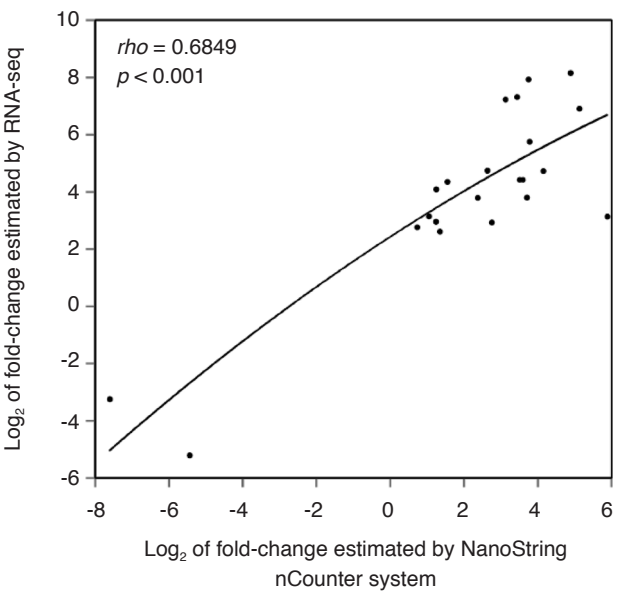
Transcript ID	Putative identity	Fold change of expression using	
		Bioinformatic analysis in ITIP/HTHP	NanoString® platform condition
Felda_Reference_Palm_Transcript_48685_3389	Lipoxygenase family protein	2.77	0.73
Felda_Reference_Palm_Transcript_42818_2273	Membrane-anchored cell wall protein	4.36	1.55
Felda_Reference_Palm_Transcript_69361_646	Cysteine proteinase inhibitor	6.91	5.13
Felda_Reference_Palm_Transcript_56763_3628	Pathogen-induced transcription factor	4.73	4.15
Felda_Reference_Palm_Transcript_49790_630	Leucine-rich repeat (LRR) family protein	4.08	1.24
Felda_Reference_Palm_Transcript_60805_2316	Subfamily of ERF/AP2 transcription factor family	7.32	3.44
Felda_Reference_Palm_Transcript_28152_1762	Leucine-rich repeat transmembrane protein kinase	-5.20	-5.43
Felda_Reference_Palm_Transcript_63698_2058	Beta-glucanase isoenzyme	3.14	5.89
Felda_Reference_Palm_Transcript_70671_2109	Peroxidase	3.80	3.70
Felda_Reference_Palm_Transcript_38860_1652	NADPH/respiratory burst oxidase	5.76	3.78
Felda_Reference_Palm_Transcript_59423_2809	Cysteine-rich receptor-like protein kinase	4.74	2.63
Felda_Reference_Palm_Transcript_39245_1096	4-coumarate--CoA ligase 2	4.42	3.60
Felda_Reference_Palm_Transcript_54436_1392	ATP binding factor	2.96	1.24
Felda_Reference_Palm_Transcript_30872_3213	Cellulose synthase	-3.24	-7.61
Felda_Reference_Palm_Transcript_36764_2705	DREB subfamily of ERF/AP2 transcription factor	7.93	3.76
Felda_Reference_Palm_Transcript_60984_1093	Plant invertase/pectin methylesterase inhibitor	3.15	1.06
Felda_Reference_Palm_Transcript_40184_1989	WRKY transcription factor family	2.61	1.35
Felda_Reference_Palm_Transcript_46069_878	WRKY transcription factor family	4.43	3.52
Felda_Reference_Palm_Transcript_63694_2380	Osmotin-like protein	3.79	2.37
Felda_Reference_Palm_Transcript_47222_907	Xyloglucan endotransglycosylase-related protein	7.24	3.13
Felda_Reference_Palm_Transcript_15522_1221	Pathogenesis-related protein PRB1-3 precursor	2.94	2.75
Felda_Reference_Palm_Transcript_77576_1426	Chalcone synthase	8.15	4.89

Note: HTHP - healthy section of basal stem tissue of the healthy palm.
ITIP - near rot section of basal stem tissue of oil the palm infected with *G. boninense*.



Note: HTIP - non-rotten, healthy section of basal stem tissue of the *G. boninense* infected palm.
HTHP - healthy section of basal stem tissue of the healthy palm.

Figure 2. Spearman's rank-order correlation between RNA-seq derived and nCounter analysis system derived expression fold-change for the HTIP/HTHP transcriptomic comparison.



Note: HTIP - non-rotten, healthy section of basal stem tissue of the *G. boninense* infected palm.
HTHP - healthy section of basal stem tissue of the healthy palm.

Figure 3. Spearman's rank-order correlation between RNA-seq derived and nCounter analysis system derived expression fold-change for the ITIP/HTHP transcriptomic comparison.

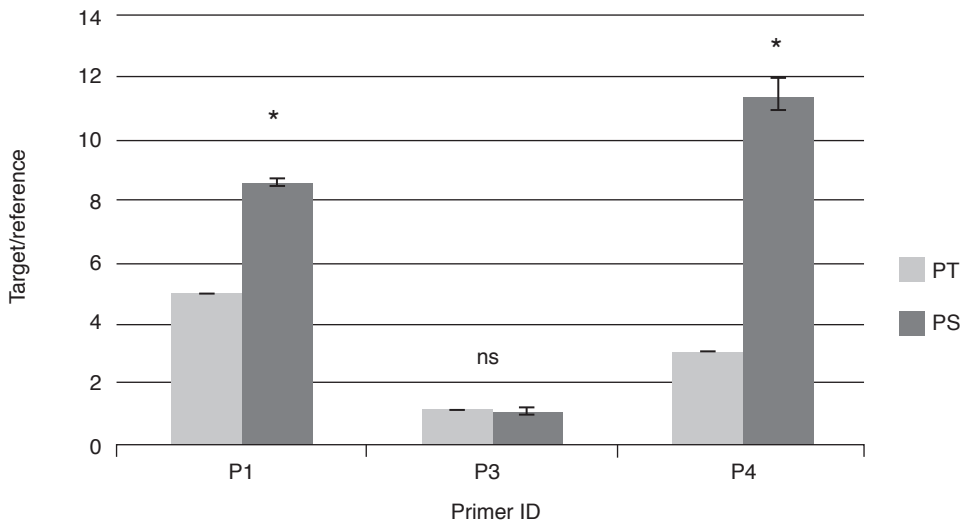


Figure 4. Expression profiles of three selected transcripts in root tissues of potentially tolerant (PT) and potentially susceptible (PS) plantlets. The plantlets were subjected to Ganoderma boninense *inoculation and pre-screened for PT or PS plantlets at the nursery. Error bars indicates standard error. Comparison of expression level was done through Student's t-test and the p-value was obtained as indicated in graph (ns - not significant, *: p, 0.01).*

differences in expression magnitude is most likely contributed by the nature of two different expression measurement platforms (sequencing-based *vs*. hybridization-based) that remain prone to systematic errors (Prokopec *et al*., 2013). Nevertheless, the nCounter analysis platform provides reliable and medium throughput method to validate the expression profiles of transcripts identified from the transcriptomic approach (RNA-seq) that usually generates a large amount of genomic information.

Potential Transcripts Expression by RT-qPCR

Numerous studies have demonstrated that mRNA abundance profiles of two or more contrast conditions or biological states can be potentially developed as biomarkers (Riedmaier and Pfaffl, 2013) for early detection in human diseases (Goodkind and Edwards, 2005), ecotoxicology test (Regier *et al*., 2013), assessing ripening stages in grape (Agudelo-Romero *et al*., 2013) and monitoring tools such as monitoring carbon dioxide injury in apple during storage (Yang *et al*., 2011; Gapper *et al*., 2013). Hence using the above ideas, we further analysed the expression profiles of three selected transcripts in root, leaf, stem, and meristem tissues of 9-month old oil palm plantlets that were predefined as PT and PS to *G. boninense*. Those three transcripts (*Table 1*) showed similar expression patterns using both RNA-seq and nCounter methods (*Tables 4* and *5*). However the expression levels of the three transcripts were not significantly different with p value > 0.05 and could not be used to distinguish between PS and PT plantlets in leaf, stem and meristem tissues (data not shown).

Two transcripts, P1 and P4 were expressed relatively higher in roots of PS than in PT plantlets (*Figure 4*). Student's t-test indicated that the difference in expression was statistically significant at p value < 0.01 in both transcripts. The expression level of P3 transcript in root was almost equal between PT and PS plantlets with p value > 0.01. The RT-qPCR results provide a direction for future experiments to assess this set of transcripts in a larger pool of nursery screened plantlets and potentially develop as biomarkers for *G. boninense* tolerant trait oil palm.

CONCLUSION

In conclusion, a large scale set of transcriptomic data of oil palm basal stem infected with *G. boninense* had been produced. A subset of the gene expression profiles of the basal stem tissues identified from the RNA-seq analysis has been validated with NanoString® nCounter analysis system, a medium-throughput transcript profiling platform with a significant correlation ($p < 0.001$) between the two platforms. nCounter analysis system is a reliable method to validate expression data that were generated from not-an-ideal-quality samples such as near rotten tissues derived from BSR of diseased palm. In addition, the expression profiles of three transcripts in which two out of the three tested transcripts showed potential to discriminate PS palms from PT palms. Thus, the two transcripts can be used for further assessment of potential biomarkers to discriminate between PS and PT palms for the breeding of *G. boninense* tolerant oil.

REFERENCES

AGUDELO-ROMERO, P; ERBAN, A; SOUSA, L; PAIS, M S; KOPKA, J and FORTES, A M (2013). Search for transcriptional and metabolic markers of grape pre-ripening and ripening and insights into specific aroma development in three Portuguese cultivars. *PLoS ONE, 8*: e60422.

ANDERS, S and HUBER, W (2010). Differential expression analysis for sequence count data. *Genome Biology, 11*: R106.

BAKER, M (2012). Digital PCR hits its stride. *Nature Methods, 9*: 541-544.

BOLGER, A M; LOHSE, M and USADEL, B (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics, 30(15)*: 2114-2120.

CHUNG, G F (2011). Management of *Ganoderma* disease in oil palm plantations. *The Planter, 87*: 325-329.

COHEN, D; BOGEAT-TRIBOULOT, M-B; TISSERANT, E; BALZERGUE, S; MARTIN-MAGNIETTE, M-L; LELANDAIS, G; NINGRE, N; RENOU, J-P; TAMBY, J-P; LE THIEC, D and HUMMEL, I (2010). Comparative transcriptomics of drought responses in *Populus*: A meta-analysis of genome-wide expression profiling in mature leaves and root apices across two genotypes. *BMC Genomics, 11*: 630.

COOPER, R M (1984). The role of cell wall-degrading enzymes in infection and damage. *Plant Diseases: Infection, Damage and Loss* (Wood, R K S and Jellis, G J eds.). Blackwell Scientific Publication. Oxford, United Kingdom. p. 13-27.

DURAND-GASSELIN, T; ASMADY, H; FLORI, A; JACQUEMARD, J C; HAYUN, Z; BRETON, F and 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.

FLEIGEA, S and PFAFFL, M W (2006). RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular Aspects of Medicine, 27*: 126-39.

GALLEGO ROMERO, I; PAI, A A; TUNG, J and GILAD, Y (2014). RNA-seq: Impact of RNA degradation on transcript quantification. *BMC Biology, 12*: 42.

GAPPER, N E; RUDELL, D R; GIOVANNONI, J J and WATKINS, C B (2013). Biomarker development for external CO_2 injury prediction in apples through exploration of both transcriptome and DNA methylation changes. *AoB Plants 5*.

GEISS, G K; BUMGAMER, R E; BIRDITT, B; DAHL, T; DOWIDAR, N; DUNAWAY, D L; FELL, H P; FERREE, S; GEORGE, R D and GROGAN, T (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnology, 26*: 317-325.

GOODKIND, J R and EDWARDS, J S (2005). Gene expression measurement technologies: Innovations and ethical considerations. *Computers and Chemical Engineering, 29*: 589-596.

GRABHERR, M G; HAAS, B J; YASSOUR, M; LEVIN, J Z; THOMPSON, D A; AMIT, I; ADICONIS, X; FAN, L; RAYCHOWDHURY, R; ZENG, Q; CHEN, Z; MAUCELI, E; HACOHEN, N; GNIRKE, A; RHIND, N; DI PALMA, F; BIRREN, B W; NUSBAUM, C; LINDBLAD-TOH, K; FRIEDMAN, N and REGEV, A (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nature Biotechnology, 29(7)*: 644-52.

GRIFFITH, M; GRIFFITH, O L; MWENIFUMBO, J; GOYA, R; MORRISSY, A S; MORIN, R D; CORBETT, R; TANG, M J; HOU, Y-C; PUGH, T J; ROBERTSON, G; CHITTARANJAN, S; ALLY, A; ASANO, J K; CHAN, S Y; LI, H I; MCDONALD, H; TEAGUE, K; ZHAO, Y; ZENG, T; DELANEY, A; HIRST, M; MORIN, G B; JONES, S J M; TAI, I T and MARRA, M A (2010). Alternative expression analysis by RNA sequencing. *Nature Methods, 7*: 843-847.

HO, C L; TAN, Y C; YEOH, K A; GHAZALI, A K; YEE, W Y and HOH, C C (2016). *De novo* transcriptome analyses of host-fungal interactions in oil palm (*Elaeis guineensis* Jacq.). *BMC Genomics, 17*: 66-84.

HUSHIARIAN, R; YUSOF, N A and DUTSE, S W (2013). Detection and control of *Ganoderma boninense*: Strategies and perspectives. *SpringerPlus, 2(1)*: 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. Vol. 16(2)*: 12-18.

JANUSZ, G; KUCHARZYK, K H; PAWLIK, A; STASZCZAK, M and PASZCZYNSKI, A J (2012). Fungal laccase, manganese peroxidase and lignin peroxidase: Gene expression and regulation. *Enzyme and Microbial Technology, 52*: 1-12.

KANAGAL-SHAMANNA, R (2016). Digital PCR: principles and applications. *Methods in Molecular Biology, 1392*: 43-50.

KHAIRUDDIN, H (1990). *Basal Stem Rot of Oil Palm: Incidence, Etiology and Control*. Master of Agriculture Science thesis, Universiti Pertanian Malaysia, Selangor, Malaysia.

KULKARNI, M M (2011). Digital multiplexed gene expression analysis using the NanoString nCounter system. *Current Protocols in Molecular Biology, 94*: B25.

LANGMEAD, B; TRAPNELL, C; POP, M and SALZBERG, S L (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology, 10*: R25.

LEE, P L A; YUSOF, M T; ISMAIL, I S; YEN PING, B T; MOHAMED AZNI, I N A; KAMARUDIN, N and SUNDRAM, S (2016). An *in vitro* study of the antifungal activity of *Trichoderma virens* 7b and a profile of its non-polar antifungal components released against *Ganoderma boninense. J. Microbiology, 54(11)*: 732-744.

LI, B and DEWEY, C N (2011). RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics, 12*: 323.

LIU, F; WU, J B; ZHAN, R L and OU, X C (2016). Transcription profiling analysis of mango - *Fusarium mangiferae interaction. Frontiers in Microbiology, 7*: 1443.

LIU, Y; FERGUSON, J F; XUE, C; SILVERMAN, I M; GREGORY, B; REILLY, M P and LI, M (2013). Evaluating the impact of sequencing depth on transcriptome profiling in human adipose. *PLoS ONE, 8(6)*: e66883.

MCDOWELL, J M and WOFFENDEN J (2003). Plant disease resistance genes: Recent insights and potential applications. *Trends in Biotechnology, 21(4)*: 178-183.

NAHER, L; HO, C L; TAN, S G; YUSUF, U K and ABDULLAH, F (2013). Cloning of transcripts encoding chitinases from *Elaeis guineensis* Jacq. and their expression profiles in response to fungal infections. *Physiological and Molecular Plant Pathology, 76(2)*: 96-103.

NAJMIE, M M K; KHALID, K; SIDEK, A and JUSOH, M A (2011). Density and untrasonic characterization of oil palm trunk infected by *Ganoderma boninense* disease. *Meas. Sci. Rev, 11*: 160-164.

PROKOPEC, S D; WATSON, J D; WAGGOTT, D M; SMITH, A B; WU, A H; OKEY, A B; POHJANVIRTA, R and BOUTROS, P C (2013). Systematic evaluation of medium-throughput mRNA abundance platforms. *RNA, 19*: 51-62.

PURANIK, S; JHA, S; SRIVASTAVA, P S; SREENIVASULU, N and PRASAD, M (2011). Comparative transcriptome analysis of contrasting foxtail millet cultivars in response to short-term salinity stress. *J. Plant Physiology, 168*: 280-287.

RAJKUMAR, A P; QVIST, P; LAZARUS, R; LESCAI, F; JU, J; NYEGAARD, M; MORS, O; BØRGLUM, A D; LI, Q and CHRISTENSEN, J H (2015). Experimental validation of methods for differential gene expression analysis and sample pooling in RNA-seq. *BMC Genomics, 16*: 548.

RAO, A K (1990). Basal stem rot (*Ganoderma*) in oil palm smallholdings. *Proc. of the* Ganoderma *Workshop* (Ariffin, D and Jalani, B S eds.). PORIM, Bangi. p. 113-131.

REES, R W; FLOOD, J; HASAN, Y and 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.

REES, R W; FLOOD, J; HASAN, Y; POTTER, U and COOPER, R M (2009). Basal stem rot of oil palm (*Elaeis guineensis*); mode of root infection and lower stem invasion by *Ganoderma boninense. Plant Pathology, 58(5)*: 982-989.

REGIER, N; BAERLOCHER, L; MUNSTERK ÖTTER, M; FARINELLI, L and COSIO, C (2013). Analysis of the *Elodea nuttallii* transcriptome in response to mercury and cadmium pollution: Development of sensitive tools for rapid exotoxicological testing. *Environmental Science Technology, 47*: 8825-8834.

RIEDMAIER, I and PFAFFL, M W (2013). Transcriptional biomarkers - high throughput screening, quantitative verification, and bioinformatical validation methods. *Methods, 59*: 3-9.

SABA, N F; WILSON, M; DOHO, G; DASILVA, J; BENJAMIN ISETT, R; NEWMAN, S; CHEN, Z G; MAGLIOCCA, K and ROSSI, M R (2015). Mutation and transcriptional profiling of formalin-fixed paraffin embedded specimens as companion methods to immunohistochemistry for determining therapeutic targets in oropharyngeal squamous cell carcinoma (OPSCC): A pilot of proof of principle. *Head and Neck Pathology, 9*: 223-235.

SAIDI, M N; GARGOURI-BOUZID; RAYANNI M and DRIRA, N (2009). Optimization of RNA isolation from brittle leaf disease affected date palm leaves and construction of a subtractive cDNA library. *Molecular Biotechnology, 41*: 63.

SCHMIEDER, R and EDWARDS, R (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics, 27*: 863-864.

SHI, Y and HE, M (2014). Differential gene expression identified by RNA-Seq and qPCR in two sizes of pearl oyster (*Pinctada fucata*). *Gene, 538*: 313-322.

SEO, G S and KIRK, P M (2000). Ganodermataceae: Nomenclature and classification. *Ganoderma. Diseases of Perennial Crops* (Flood, J; Bridge, P D and Holderness, M eds.). CABI Bioscience, Egham, United Kingdom. p. 3-22.

STASKAWICZ, B J (2001). Genetics of plant-pathogen interactions specifying plant disease resistance. *Plant Physiology, 125*: 73-76.

TACHIBANA, C (2015). Transcriptomics today: Microarrays, RNA-seq, and more. *Science, 349(6247)*: 544-546.

TAM, S; DE BORJA, R; TSAO, M S and MCPHERSON, J D (2014). Robust global microRNA expression profiling using next generation sequencing technologies. *Laboratory Investigation, 94*: 350-358.

TAN, Y C; YEOH, K A; WONG, M Y and HO, C L (2013). Expression profiles of putative defense-related proteins in oil palm (*Elaeis guineensis*) colonized by *Ganoderma boninense. J. Plant Physiology, 170(16)*: 1455-1460.

WU, A R; NEFF, N F; KALISKY, T; DALERBA, P; TREUTLEIN, B; ROTHENBERG, M E; MBURU, F M; MANTALAS, G L; SIM, S; CLARKE, M F and QUAKE, S R (2014). Quantitative assessment of single-cell RNA-sequencing methods. *Nature Methods, 11*: 41-46.

YANG, X S; WU, J; ZIEGLER, T E; YANG, X; ZAYED, A; RAJANI, M S; ZHOU, D; BASRA, A S; SCHACHTMAN, D P; PENG M; ARMSTRONG, C L; CALDO, R A; MORRELL, J A; LACY, M and STAUB, J M (2011). Gene expression biomarkers provide sensitive indicators of in planta nitrogen status in maize. *Plant Physiology, 157*: 1841-1852.

YEAP, W C; LOO, J M; WONG, Y C and KULAVEERASINGAM, H (2014). Evaluation of suitable reference genes for qRT-PCR gene expression normalization in reproductive, vegetative tissues and during fruit development in oil palm. *Plant Cell, Tissue and Organ Culture, 116*: 55-66.

YEOH, K A; OTHMAN, A; MEON, S; ABDULLAH, F and HO, C L (2012). Sequence analysis and gene expression of putative exo- and endo-glucanases from oil palm (*Elaeis guineensis*) during fungal infection. *J. Plant Physiology, 169*: 1565-1570.