



ELSEVIER

Contents lists available at ScienceDirect

Data in brief

journal homepage: www.elsevier.com/locate/dib



Data Article

Transcriptomic data of mature oil palm basal trunk tissue infected with *Ganoderma boninense*



Nur Qistina Othman^{*}, Suhaila Sulaiman, Yang Ping Lee,
Joon Sheong Tan

FGV R&D Sdn. Bhd., FGV Innovation Centre (Biotechnology), PT 23417 Lengku Teknologi, 71760 Bandar Enstek, Negeri Sembilan, Malaysia

ARTICLE INFO

Article history:

Received 3 May 2019

Received in revised form 18 June 2019

Accepted 11 July 2019

Available online 19 July 2019

Keywords:

Basal stem rot

Ganoderma boninense

Oil palm

Infection

ABSTRACT

To date, *Ganoderma boninense* is known to be the causal agent of basal stem rot (BSR) disease in oil palm (*Elaeis guineensis*). This disease causes rotting in the roots, basal and upper stem of oil palm. Infection causes progressive destruction of the basal tissues at the oil palm trunk and internal dry rotting, particularly at the intersection between the bole and trunk. Molecular responses of oil palm during infection are not well study although this information is crucial to strategize effective measures to control or eliminate BSR. Here we report three sets of transcriptome data from samples of near-rot section of basal stem tissue of oil palm tree infected with *G. boninense* (IPIT), healthy section of basal stem tissue of the same *G. boninense* infected palm (IPHT) and the healthy section of basal stem tissue of the healthy palm (HPHT). The raw reads were deposited into NCBI database and can be accessed via BioProject accession number PRJNA530030.

© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

^{*} Corresponding author.

E-mail address: nqistina.o@fgvholdings.com (N.Q. Othman).

Specifications Table

| | |
|----------------------------|---|
| Subject area | Biology |
| More specific subject area | Molecular biology |
| Type of data | Transcriptome data |
| How data was acquired | Paired-end reads of <i>G. boninense</i> transcriptome were sequenced using Illumina HiSeq 2000 |
| Data format | Raw sequence in FASTQ format |
| Experimental factors | Total RNA was extracted from a pair of healthy and infected clonal mature palm |
| Experimental features | Total RNA of infected and non-infected matured oil palm basal stem were extracted using method by Saidi et al., 2009 [1]. The extracted RNAs were used for transcriptome sequencing using Illumina HiSeq 2000 platform, (Illumina, USA). |
| Data source location | Serting Hilir, Negeri Sembilan, Malaysia |
| Data accessibility | Raw reads in FASTQ format were deposited in NCBI SRA database with accession number SRX5608988, SRX5608989 and SRX5608990 (https://www.ncbi.nlm.nih.gov/sra/?term=SRX5608988 , https://www.ncbi.nlm.nih.gov/sra/?term=SRX5608989 , https://www.ncbi.nlm.nih.gov/sra/?term=SRX5608990) |

Value of the data

- The data reported here is beneficial to elucidate the plant-pathogen interaction for matured oil palm tree that is naturally infected with *G. boninense* in an open oil palm estate.
- This data information is crucial to strategized effective measures against BSR in the future.
- The data allows further comparative analysis to identify candidate genes of interest that could play significant roles in oil palm defense system.
- With the accessible data, more subsequent projects can be accelerated to elucidate infection pathway of *G. boninense* in oil palm as well as identification of genes, proteins or metabolites involved during the infection process that may facilitate the development of possible solutions against *G. boninense*.

1. Data

Three sets of transcriptome data generated from the cDNA libraries were prepared from the total RNA extracted from three different samples as mentioned above. The data sets were named as IPHT (the healthy section of infected oil palm), IPIT (the near-root section of infected oil palm) and HPHT (the cross section of healthy tissue section of healthy oil palm tree).

2. Experimental design, materials and methods

2.1. Sampling materials

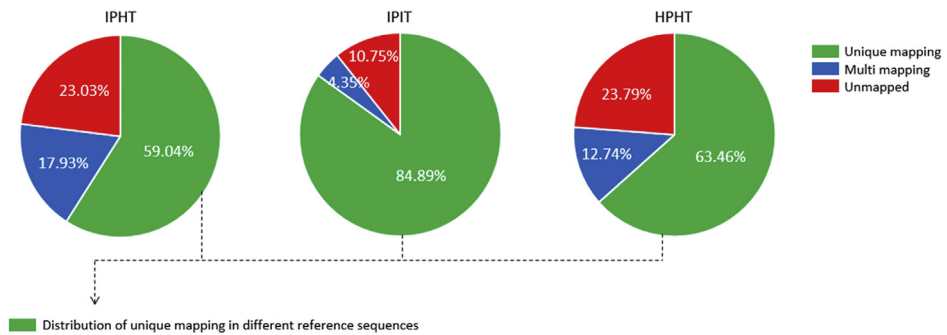
A pair of healthy and infected clonal mature palm was selected from Serting Hilir field research station, Negeri Sembilan, Malaysia owned by Felda Agricultural Services Sdn Bhd. Briefly, the bark was marked according to the health status. The healthy section (IPHT) and near-root section (IPIT) of infected oil palm were collected from the cross section of the basal stem. The cross section of healthy tissue section of healthy oil palm tree (HPHT) was also collected. The healthy palm was determined by healthy oil palm phenotype with no *G. boninense* fruiting body attached at the basal stem and the palm does not exhibit stress-related symptoms phenotypically. All collected basal stem tissues were immediately submerged in liquid nitrogen and kept at -80°C freezer until further processing for RNA extraction.

2.2. RNA extraction and quality assessment of total RNA

Total RNA of matured basal stem samples; IPHT, IPIT and HPHT were extracted using method from Saidi et al., 2009 [1]. The quality of the extracted RNAs was determined using Bioanalyzer (Agilent Technologies, USA) and UV absorbance readings (NanoDrop, Thermo Fisher Scientific Inc., USA). The extracted RNAs were used for Illumina's HiSeq 2000 transcriptome sequencing.

Table 1
 Statistics of raw reads from three samples before and after pre-processing.

| Sample name | Description | Details | Before pre-processing | After pre-processing | | |
|-------------|----------------------------------|----------------------|--------------------------|-------------------------|-------------------------|-------------------|
| | | | | Total reads | Paired reads | Orphan reads |
| IPHT | Infected Palm, Healthy tissue | Total number of read | 276,546,920 (100.00%) | 246,029,758 (88.96%) | 245,994,098 (88.95%) | 35,660 (0.01%) |
| | | Total read size | 27,654,692,000 (100.00%) | 21,610,020,425 (78.14%) | 21,606,913,722 (78.13%) | 3,106,703 (0.01%) |
| IPIT | Infected Palm, Intersection zone | Total number of read | 214,112,784 (100.00%) | 197,773,317 (92.37%) | 197,742,548 (92.35%) | 30,769 (0.01%) |
| | | Total read size | 21,411,278,400 (100.00%) | 18,120,714,405 (84.63%) | 18,117,911,395 (84.62%) | 2,803,010 (0.01%) |
| HPHT | Healthy Palm, Healthy tissue | Total number of read | 224,285,692 (100.00%) | 199,014,398 (88.73%) | 198,985,062 (88.72%) | 29,336 (0.01%) |
| | | Total read size | 22,428,569,200 (100.00%) | 17,674,879,482 (78.81%) | 17,672,289,541 (78.79%) | 2,589,941 (0.01%) |



| Sample / Reference Sequence | Oil palm genome | Percentage (%) | Oil palm chloroplast | Percentage (%) | Date palm mitochondria | Percentage (%) | Ganoderma genome | Percentage (%) |
|-----------------------------|-----------------|----------------|----------------------|----------------|------------------------|----------------|------------------|----------------|
| IPHT | 138,402,486 | 56.26% | 1,697,386 | 0.69% | 2,652,794 | 1.08% | 2,475,626 | 1.01% |
| IPIT | 166,436,864 | 84.17% | 428,833 | 0.22% | 625,292 | 0.32% | 381,702 | 0.19% |
| HPHT | 120,302,279 | 60.46% | 1,438,339 | 0.72% | 2,681,896 | 1.35% | 1,859,256 | 0.93% |

Fig. 1. Statistics of raw data mapping to different reference sequences (oil palm genome, oil palm chloroplast, date palm mitochondria and *Ganoderma* genome) for sample IPHT, IPIT and HPHT.

2.3. 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 200–300bp. 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 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).

2.4. Pre-processing of RNA-seq raw data and transcriptome mapping

The quality of RNA-seq raw data was assessed using FastQC [2] and pre-processed with FASTX-Toolkit [3] to trim off adaptor sequences and low quality bases (<Q20). Reads containing unknown bases or reads that were shorter than 30 bp after trimming were discarded to obtain clean reads (Table 1). The high quality reads of each sample were aligned to reference sequences consisting of *E. guineensis* genome [4], oil palm chloroplast genome [5], date palm mitochondrial genome [6] as well as *Ganoderma boninense* genome [7] (Fig. 1). The read alignment was carried out using TopHat v2.0.12 [8] and Bowtie v2.2.1 [9] as algorithmic core [params: -segment-length 45 -read-mismatches 2].

Acknowledgements

This work was funded by FGV Agri Services Sdn. Bhd. The authors wish to thank the Executive Director/CEO of FGV R&D Sdn Bhd for the given support and consultation in the entire project.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.dib.2019.104288>.

References

- [1] M.N. Saidi, R. Gargouri-Bouزيد, M. Rayanni, N. Drira, Optimization of RNA isolation from Brittle Leaf Disease affected date palm leaves and construction of a subtractive cDNA library, *Mol. Biotechnol.* 41 (2009) 63–68, <https://doi.org/10.1007/s12033-008-9104-1>.
- [2] S. Andrews, FastQC a quality control tool for high throughput sequence data, 2010 citeulike-article-id:11583827, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- [3] H. Lab, FASTX toolkit, http://hannonlab.cshl.edu/fastx_toolkit/index.html. (n.d.). citeulike-article-id:9103573.
- [4] R. Singh, M. Ong-Abdullah, E.L. Low, M.A.A. Manaf, R. Rosli, R. Nookiah, L.C.-L. Ooi, S.-E. Ooi, K. Chan, M.A. Halim, N. Azizi, J. Nagappan, B. Bacher, N. Lakey, S.W. Smith, D. He, M. Hogan, M.A. Budiman, E.K. Lee, R. DeSalle, D. Kudrna, J.L. Goicoechea, R. A. Wing, R.K. Wilson, R.S. Fulton, J.M. Ordway, R.A. Martienssen, R. Sambanthamurthi, Oil palm genome sequence reveals divergence of interfertile species in Old and New worlds, *Nature* 500 (2013) 335–359, <https://doi.org/10.1038/nature12309>.
- [5] P. Uthaipaisanwong, J. Chanprasert, J.R. Shearman, D. Sangsrakru, T. Yoocha, N. Jomchai, C. Jantaturiyarat, S. Tragoonrung, S. Tangphatsornruang, Characterization of the chloroplast genome sequence of oil palm (*Elaeis guineensis* Jacq.), *Gene* 500 (2012) 172–180, <https://doi.org/10.1016/j.gene.2012.03.061>.
- [6] Y. Fang, H. Wu, T. Zhang, M. Yang, Y. Yin, L. Pan, X. Yu, X. Zhang, S. Hu, I.S. Al-Mssallem, J. Yu, A complete sequence and transcriptomic analyses of date palm (*Phoenix dactylifera* L.) mitochondrial genome, *PLoS One* 7 (2012) 1–12, <https://doi.org/10.1371/journal.pone.0037164>.
- [7] M. Mercière, A. Laybats, C. Carasco-Lacombe, J.S. Tan, C. Klopp, T. Durand-Gasselín, S.S.R.S. Alwee, L. Camus-Kulandaivelu, F. Breton, Identification and development of new polymorphic microsatellite markers using genome assembly for *Ganoderma boninense*, causal agent of oil palm basal stem rot disease, *Mycol. Prog.* 14 (2015) 103, <https://doi.org/10.1007/s11557-015-1123-2>.
- [8] D. Kim, G. Perteau, C. Trapnell, H. Pimentel, R. Kelley, S.L. Salzberg, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions, *Genome Biol.* 14 (2013) R36, <https://doi.org/10.1186/gb-2013-14-4-r36>.
- [9] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, *Nat. Methods* 9 (2012) 357, <https://doi.org/10.1038/nmeth.1923>.