

DATA NOTE

Open Access



Transcriptome of *Aquilaria malaccensis* containing agarwood formed naturally and induced artificially

Farah Adibah Abdul Kadir¹, Kamalrul Azlan Azizan² and Roohaida Othman^{1,2*} 

Abstract

Objectives: Agarwood is the aromatic heartwood formed upon wounding of *Aquilaria* trees either naturally formed due to physical wound sustained from natural phenomena followed by microbial infection, or artificially induced using different inoculation methods. Different induction methods produce agarwoods with different aromas which have impacts on their commercial values. In lieu of elucidating the molecular mechanisms of agarwood formation under different treatment conditions, the transcriptome profiles of trunk tissues from healthy *A. malaccensis* tree, and naturally and artificially induced trees were obtained.

Data description: The transcriptome of trunk tissues from healthy *A. malaccensis*, and naturally and artificially induced trees were sequenced using Illumina HiSeq™ 4000 platform which resulted in a total of 38.4 Gb clean reads with Q30 rate of at least 91%. The transcriptome consists of 85,986 unigenes containing 1305 bases on average which were annotated against several databases. From this, 44,654 unigenes were mapped to 290 metabolic pathways in the Kyoto Encyclopedia of Genes and Genomes database. These transcriptome data represent considerable contribution towards *Aquilaria* transcriptome data and enhance current knowledge in comprehending the molecular mechanisms underlying agarwood formation in *Aquilaria* spp.

Keywords: *Aquilaria malaccensis*, Agarwood, Transcriptome, Trunk tissues, Natural formation, Artificial induction

Objective

The valuable agarwood is the fragrant resinous heartwood which are important ingredients in fragrances, incense and medicines [1, 2]. Nine *Aquilaria* spp. belonging to the Thymelaeaceae family are agarwood producers with *A. malaccensis* among the primary ones [3, 4]. Agarwood formation is a slow defense response upon wounding of *Aquilaria* trees due to natural phenomena such as gale and insect bites followed by microbial infection [5, 6]. The wood tissues die to avoid damage expansion and

form agarwood accompanied by the release of secondary metabolites [7]. The high agarwood demand has caused natural resource depletion of all agarwood-producing *Aquilaria* species and listed them in the Convention on International Trade in Endangered Species of Wild Fauna and Flora Appendix II [8] where stringent jurisdiction controls their trades [9]. *A. malaccensis* is among the eight *Aquilaria* species categorised on the International Union for Conservation of Nature red list as endangered species [10]. Strategies for sufficient supply of agarwood include reforestation and artificial induction method development. Artificial inductions may involve stem wounding alone or coupled with inoculations using microbial cultures and/or chemical stimulants [11]. However, different induction techniques produce agarwoods

*Correspondence: roohaida@ukm.edu.my; roohaida.othman@gmail.com

¹ Department of Biological Sciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, UKM, 43600 Bangi, Selangor, Malaysia

Full list of author information is available at the end of the article



with different aromas [12, 13] due to different compositions of metabolites released [14–17].

Previous transcriptomic studies have focused on mechanically wounded [18] and chemically induced [19] *A. sinensis*, and *A. malaccensis* senescing calli [20]. Studies on artificially induced *A. malaccensis* had focused on terpene synthase gene expression analyses [21, 22]. There has been no report on transcriptome of *Aquilaria* containing naturally formed agarwood since such trees are rare. To understand the mechanisms of agarwood formation from the different treatments, here we present *A. malaccensis* transcriptomes from healthy, and naturally formed and artificially induced trees. The annotated transcriptomes presented here provide a valuable resource for researchers interested in agarwood formation in agarwood-producing species.

Data description

Identification of *A. malaccensis* Lam. [23] was performed by A. Damanhuri (Curator for Universiti Kebangsaan Malaysia Herbarium; UKMB) for voucher specimen labelled as M. H. Azhari 1. Trunk samples from uninjured trees were used as control, whilst agarwood-containing trunk samples were obtained from naturally and artificially induced trees. Naturally formed agarwood was found in broken tree trunks resulting from natural phenomena combined with microbial infection. Artificially induced agarwood was formed 5 years after nail wounding combined with honey-containing inoculum injection.

RNA extraction, library construction and sequencing

Total RNA extracted following the Trizol protocol [24] was evaluated for quantity and quality using NanoDrop spectrophotometer (Thermo Fisher Scientific Inc., USA) and Agilent 2100 Bioanalyzer (Agilent Technologies, USA). Total RNAs having RNA integrity number (RIN) values of at least 8.0 were used for construction three cDNA libraries using commercial service provided by Macrogen, Inc. (Seoul, Republic of Korea) on Illumina HiSeq 4000 platform (San Diego, USA) (Datasets 1–3). A schematic overview of this study is shown in Data file 1.

Transcriptome assembly

Raw reads (271,072,542) obtained through Illumina HiSeq sequencing were filtered to remove reads with low quality (Data file 5). Adapter sequences in the raw reads were eliminated using Cutadapt software (version 2.3–0) and raw reads quality was examined using FastQC version 0.11.2 [25]. High quality reads (Data file 5) were acquired by removing adapters and other undesirable

sequences using Trimmomatic [26]. Trinity version 2.1.1 [27] was utilised for the assembly of the reads de novo [28] while the reads were clustered into non-redundant unigenes set using TIGR Gene Indices clustering tools version 2.1 [29]. Transcript abundance was calculated by mapping the trimmed reads on to the assembled transcripts and measured with RNA-Seq by employing Expectation Maximization (RSEM) version 1.2.196. From the output, transcripts with the values of Fragments Per Kilobase of exon Per Million fragments mapped (FPKM) less than one were filtered from the original transcripts file to produce the final unigenes consisting of 85,986 contigs (Data file 5). The contig lengths of *A. malaccensis* ranged from 201 to 25,238 bp with 1305 bp average length. BLAST searches were conducted for gene functional annotations against public databases such as Swiss-Prot, GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomics) (Data files 2, 3 and 6). The annotated unigenes were mapped to KEGG database to hit a total of 290 metabolic pathways (Data file 4).

Quality, completeness and depth of the *A. malaccensis* transcriptome

The *A. malaccensis* transcriptome was appraised using BUSCO (Benchmarking Universal Single-Copy Orthologs) to determine the completeness of the unigene assembly [30]. Embryophyta and eudicotyledon gene sets were used which have 1375 and 212 near-universal single-copy orthologs, respectively (Data file 7). Comparison of the statistics for *A. malaccensis* transcriptome and the previously published *A. sinensis* transcriptomes [18, 19] showed 238 and fivefold increases of the former over the 454-based [18] and Illumina-based *A. sinensis* transcriptomes [19], respectively (Data file 8). The *A. malaccensis* high quality reads also have the highest average transcript lengths of 1305 bp (see Table 1).

Limitations

All the samples used were of different ages. The uninjured healthy tree was much younger than the trees containing agarwood. The trees containing naturally formed agarwood were much older since the process of agarwood formation took much longer than 5 years which was the time period for the formation of agarwood in the artificially induced trees.

Table 1 Overview of data files/data sets

Label	Name of data file/data set	File types (file extension)	Data repository and identifier (DOI or accession number)
Data set 1	RNA-Seq of healthy trunk tissue of <i>Aquilaria malaccensis</i>	fastq file (.fastq)	NCBI sequence read archive (https://identifiers.org/insdc.sra:SRR8863603) [31]
Data set 2	RNA-Seq of naturally induced agarwood from trunk tissue of <i>Aquilaria malaccensis</i>	fastq file (.fastq)	NCBI sequence read archive (https://identifiers.org/insdc.sra:SRR8863602) [32]
Data set 3	RNA-Seq of synthetically induced agarwood from trunk tissue of <i>Aquilaria malaccensis</i>	fastq file (.fastq)	NCBI sequence read archive (https://identifiers.org/insdc.sra:SRR8863604) [33]
Data set 4	TSA: <i>Aquilaria malaccensis</i> , transcriptome shotgun assembly	Entry note	NCBI Sequence Read Archive https://identifiers.org/ncbi/insdc:GHJS000000000.1 [34]
Data file 1	Schematic overview of the study	Image file (.tiff)	Figshare (https://doi.org/10.6084/m9.figshare.13417400.v2) [35]
Data file 2	Comparison of functional annotations for <i>Aquilaria malaccensis</i> unigenes against different databases	Image file (.tiff)	Figshare (https://doi.org/10.6084/m9.figshare.13417400.v2) [35]
Data file 3	Functional classifications of <i>Aquilaria malaccensis</i> unigenes	Image file (.tiff)	Figshare (https://doi.org/10.6084/m9.figshare.13417400.v2) [35]
Data file 4	Most enriched KEGG pathways for <i>A. malaccensis</i> differentially expressed genes	Document file (.docx)	Figshare (https://doi.org/10.6084/m9.figshare.13417400.v2) [35]
Data file 5	Summary statistics for transcriptome assembly and annotations sequencing results for <i>Aquilaria malaccensis</i>	Document file (.docx)	Figshare (https://doi.org/10.6084/m9.figshare.13417400.v2) [35]
Data file 6	Summary of gene functional annotations for <i>Aquilaria malaccensis</i> unigenes	Document file (.docx)	Figshare (https://doi.org/10.6084/m9.figshare.13417400.v2) [35]
Data file 7	BUSCO analysis of assembly completeness	Document file (.docx)	Figshare (https://doi.org/10.6084/m9.figshare.13417400.v2) [35]
Data file 8	Comparison of transcriptome assembly and annotation statistics for <i>Aquilaria</i> spp.	Document file (.docx)	Figshare (https://doi.org/10.6084/m9.figshare.13417400.v2) [35]

Abbreviations

IUCN: International Union for Conservation of Nature; CITES: Convention on International Trade in Endangered Species of Wild Fauna and Flora; FPKM: Fragments Per Kilobase of exon Per Million fragments mapped; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomics; BUSCO: Bench-marking universal single-copy ortholog; SRA: Short read archive; TSA: Transcriptome shotgun assembly.

Acknowledgements

The authors thank Mr. Nor Mohamad Zazali Ali from from Oudh Malizie Resources for providing all the *A. malaccensis* trunk samples for this study and to Mr. Afiq Adham Abd Rasib for his assistance in data analysis.

Authors' contributions

RO and KAA designed the project and devised the experiment; FDAK performed the experimental work; FDAK and RO performed the data analyses; and RO wrote the manuscript. All authors read and approved the final manuscript.

Funding

The grant for this study was awarded by the Ministry of Higher Education, Malaysia under FRGS/1/2015/SG05/UKM/02/1.

Availability of data materials

The raw fastq files were deposited in the National Center for Biotechnology Information and are available with accession numbers (SRR8863602–SRR8863604) under Bioproject PRJNA497968 (Datasets 1–3); SRR8863602, SRR8863603, SRR8863604 [31–33]. Assembly of non-redundant ORF unigene sequences are available from NCBI transcriptome shotgun assembly (TSA) database (<https://identifiers.org/ncbi/insdc:GHJS000000000.1>) [34]. The supplementary materials (Figures S1–S3, Tables S1–S5) can be accessed on Figshare (<https://doi.org/10.6084/m9.figshare.13417400.v2>) [35]. Please see Table 1 and reference list for details and links to the data.

Declarations

Ethics approval and consent to participate

Collection of plant materials for scientific research has been conducted in compliance with the Convention on the Trade in Endangered Species of Wild Fauna and Flora [8].

Consent for publication

Not applicable.

Competing interests

The authors declare that there are no competing interests.

Author details

¹ Department of Biological Sciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, UKM, 43600 Bangi, Selangor, Malaysia. ² Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia, UKM, 43600 Bangi, Selangor, Malaysia.

Received: 23 December 2020 Accepted: 15 March 2021

Published online: 25 March 2021

References

1. Akter S, Islam MT, Zulkefeli M, Khan SI. Agarwood production—a multidisciplinary field to be explored in Bangladesh. *Int J Pharm Life Sci.* 2013;2:22–32. <https://doi.org/10.3329/ijpls.v2i1.15132>.
2. Naziz PS, Das R, Sen S. The Scent of Stress: evidence from the unique fragrance of agarwood. *Front Plant Sci.* 2019;10:840. <https://doi.org/10.3389/fpls.2019.00840>.
3. Ng LT, Chang YS, Kadir AA. A review on agar (gaharu) producing *Aquilaria* species. *J Trop Forest Prod.* 1997;2(2):272–85.

4. Rasool S, Mohamed R. Understanding agarwood formation and its challenges. In: Mohamed R, editor. Agarwood: science behind the fragrance. Singapore: Springer; 2016. p. 39–56.
5. Pojanagaroon S, Kaewrak C. Mechanical methods to stimulate aloes wood formation in *Aquiliria crassna* Pierre ex H Lec (kritsana) trees. ISHS Acta Hort. 2006;676:161–6. <https://doi.org/10.17660/ActaHortic.2005.676.20>.
6. Zhang Z, Yang Y, Wei JH, Meng H, Sui C, Chen HQ. Advances in studies on mechanism of agarwood formation in *Aquilaria sinensis* and its hypothesis of agarwood formation induced by defense response. Chin Tradit Herb Drugs. 2010;41(1):156–9.
7. Naef R. Volatile and semi volatile constituents of agarwood, the infected heartwood of *Aquilaria* species: a review. Flavour Fragr J. 2011;26(2):73–87. <https://doi.org/10.1002/ffj.2034>.
8. CITES. Convention on International Trade in Endangered Species of Wild Fauna and Flora. <https://cites.org/eng/disc/text.php>. Accessed 27 Jan 2020.
9. Ito M, Honda G. Agarwood-its sedative effect on mice and current state in the production sites. Aroma Res. 2008;34:122–7.
10. IUCN. IUCN Red List of Threatened Species. <http://www.iucnredlist.org>. Accessed 28 Jan 2020.
11. Tan CS, Isa NM, Ismail I, Zainal Z. Agarwood induction: current developments and future perspectives. Front Plant Sci. 2019;10:122. <https://doi.org/10.3389/fpls.2019.00122>.
12. Chen H, Yang Y, Xue J, Wei J, Zhang Z, Chen H. Comparison of compositions and antimicrobial activities of essential oils from chemically stimulated agarwood, wild agarwood and healthy *Aquilaria sinensis* (Lour.) Gilg trees. Molecules. 2011;16:4884–96. <https://doi.org/10.3390/molecules16064884>.
13. Chen X, Liu Y, Yang Y, Feng J, Liu P, Sui C, et al. Trunk surface agarwood-inducing technique with *Rigidoporus vinctus*: an efficient novel method for agarwood production. PLoS ONE. 2018;13(6):e0198111. <https://doi.org/10.1371/journal.pone.0198111>.
14. Ismail N, Azah MAN, Jamil M, Rahiman MHF, Tajuddin SN, Taib MN. Analysis of high quality agarwood oil chemical compounds by means of SPME/GC-MS and Z-score techniques. Malays J Anal Sci. 2013;17:403–13.
15. Jayachandran K, Sekar I, Parthiban KT, Amirtham D, Suresh KK. Analysis of different grades of agarwood (*Aquilaria malaccensis* Lamk.) oil through GC-MS. Indian J Nat Prod Resour. 2014;5:44–7.
16. Wong YF, Chin ST, Perlmutter P, Marriott PJ. Evaluation of comprehensive two-dimensional gas chromatography with accurate mass time-of-flight mass spectrometry for the metabolic profiling of plant-fungus interaction in *Aquilaria malaccensis*. J Chromatogr A. 2015;1387:104–15. <https://doi.org/10.1016/j.chroma.2015.01.096>.
17. Abdul Kadir FA, Azizan KA, Othman R. Datasets of essential oils from naturally formed and artificially induced *Aquilaria malaccensis* agarwoods. Data Brief. 2020;28:104987. <https://doi.org/10.1016/j.dib.2019.104987>.
18. Xu YH, Zhang Z, Wang MX, Wei JH, Chen HJ, Gao ZH, et al. Identification of genes related to agarwood formation: transcriptome analysis of healthy and wounded tissues of *Aquilaria sinensis*. BMC Genom. 2013;14:227. <https://doi.org/10.1186/1471-2164-14-227>.
19. Ye W, Wu H, He X, Wang L, Zhang W, Li H, et al. Transcriptome sequencing of chemically induced *Aquilaria sinensis* to identify genes related to agarwood formation. PLoS ONE. 2016;11:e0155505. <https://doi.org/10.1371/journal.pone.0155505>.
20. Siah CH, Namasivayam P, Mohamed R. Transcriptome reveals senescing callus tissue of *Aquilaria malaccensis*, an endangered tropical tree, triggers similar response as wounding with respect to terpenoid biosynthesis. Tree Genet Genomes. 2016;12:33. <https://doi.org/10.1007/s11295-016-0993-z>.
21. Azzarina AB, Mohamed R, Lee SY, Nazre M. Temporal and spatial expression of terpene synthase genes associated with agarwood formation in *Aquilaria malaccensis* Lam. NZ J For Sci. 2016;46:12. <https://doi.org/10.1186/s40490-016-0068-9>.
22. Afiq Adham AR, Tong FX, Zeti Azura MH, Othman R. Sequence analysis of terpene synthase cDNA from transcriptome profile of infected *Aquilaria malaccensis*. Malay J Biochem Mol Biol. 2018;21(1):71–2.
23. IPNI. International Plant Names Index. The Royal Botanic Gardens, Kew, Harvard University Herbaria & Libraries and Australian National Botanic Gardens. 2020. <http://www.ipni.org>. Accessed 10 Dec 2020.
24. Rio DC, Ares M, Hannon GJ, Nilsen TW. Purification of RNA using TRIzol (TRI Reagent). Cold Spring Harb Prot. 2010;5:1–4. <https://doi.org/10.1101/pdb.prot5439>.
25. Andrews S. FastQC: A quality control tool for high throughput sequence data [online]. 2010. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 27 Jan 2020.
26. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(14):2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
27. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29:644. <https://doi.org/10.1038/nbt.1883>.
28. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013;8(8):1494–512. <https://doi.org/10.1038/nprot.2013.084>.
29. Perteã G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, et al. TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics. 2003;19:651–2. <https://doi.org/10.1093/bioinformatics/btg034>.
30. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31(19):3210–2. <https://doi.org/10.1093/molbev/msx319>.
31. Abdul Kadir FA, Azizan KA, Othman R. RNA-Seq of healthy trunk tissue of *Aquilaria malaccensis*. NCBI sequence read archive. 2019. <https://identifiers.org/insdc.sra:SRR8863603>.
32. Abdul Kadir FA, Azizan KA, Othman R. RNA-Seq of naturally induced agarwood from trunk tissue of *Aquilaria malaccensis*. NCBI sequence read archive. 2019. <https://identifiers.org/insdc.sra:SRR8863602>.
33. Abdul Kadir FA, Azizan KA, Othman R. RNA-Seq of synthetically induced agarwood from trunk tissue of *Aquilaria malaccensis*. NCBI Sequence Read Archive. 2019. <https://identifiers.org/insdc.sra:SRR8863604>.
34. Abdul Kadir FA, Azizan KA, Othman R. TSA: *Aquilaria malaccensis*, transcriptome shotgun assembly. GenBank. 2019. <https://identifiers.org/ncbi/insdc:GHJS000000000.1>.
35. Abdul Kadir FA, Azizan KA, Othman R. Supplementary files for transcriptome of *Aquilaria malaccensis* containing naturally formed and artificially induced agarwoods. figshare. 2020. <https://doi.org/10.6084/m9.figshare.13417400.v2>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.