Lubis *et al. BMC Research Notes* (2023) 16:40
https://doi.org/10.1186/s13104-023-06306-1

BMC Research Notes

DATA NOTE

Open Access

Transcriptome dataset from *Solanum lycopersicum* L. cv. Micro-Tom; wild type and two mutants of INDOLE-ACETIC-ACID (SlIAA9) using long-reads sequencing oxford nanopore technologies

Wahyu Muhammad Yuha Lubis[1], M Adrian[1], Nurul Jadid[2], Ani Widiastuti[3], Hiroshi Ezura[4], Syariful Mubarok[5], Dhika Prita Hapsari[6], Roedhy Poerwanto[6] and Deden Derajat Matra[6*]

Abstract

Objective Tomatoes are the most widely consumed fruit vegetable and are relatively easy to cultivate. However, an increase in temperature causes some plants to respond with a decrease in fruit production. So, it is necessary to develop plants resistant to extreme temperature changes. The tomato cv. Micro-Tom has genetic variations in the gene of INDOLE-ACETIC-ACID, namely SlIAA9-3 and SlIAA9-5. However, the genetic information regarding the full-length transcript of the gene from this type of tomato plant is unknown. Therefore, this study aimed to determine the full-length transcript of the genes of these three types of tomatoes using long-reads sequencing technology from Oxford Nanopore.

Data description The total RNA from three types of Micro-Tom was isolated with the RNeasy PowerPlant Kit. Then, the RNA sequencing process used PCR-cDNA Barcoding kit - SQK-PCB109 and continued with the processing of raw reads based on the protocol from microbepore protocol (https://github.com/felixgrunberger/microbepore). The resulting raw reads were 578 374, 409 905, and 851 948 for wildtype, iaa9-3, and iaa9-5, respectively. After obtaining cleaned reads, each sample was mapped to the tomato reference genome (*S. lycopersicum* ITAG4.0) with the Minimap2 program. In particular, 965 genes were expressed only in the iaa9-3 mutant, and 2332 genes were expressed only in the iaa9-5 mutant. Whereas in the wild type, 1536 genes are specifically expressed. In cluster analysis using the heatmap analysis, separate groups were obtained between the wild type and the two mutants. This proves an overall difference in transcript levels between the wild type and the mutants.

Keywords Climate change, Mutant, Parthenocarpy, RNA-seq

*Correspondence:
Deden Derajat Matra
dedenmatra@apps.ipb.ac.id
[1]Study Program of Agronomy and Horticulture, Graduate School of IPB University, Bogor, Indonesia
[2]Department of Biology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[3]Department of Plant Protection, Faculty of Agriculture, Universitas Gadjah Mada, Yogyakarta, Indonesia
[4]Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan
[5]Department of Agronomy, Faculty of Agriculture, Universitas Padjadjaran, Sumedang, Indonesia
[6]Department of Agronomy and Horticulture, Faculty of Agriculture, IPB University, Bogor, Indonesia

© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

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	Raw RNA-seq reads	Fastq files (.fastq)	https://ddbj.nig.ac.jp/resource/sra-study/DRP009326 [12]
Data file 1	Summary of raw and clean reads and transcriptome assembly	Document file (.docx)	https://doi.org/10.6084/m9.figshare.21701345 [13]
Data file 2	Venn diagram for comparison of the number of expressed genes in wild type, iaa9-3 and iaa9-5 mutants	png file (.png)	https://doi.org/10.6084/m9.figshare.21701354 [14]
Data file 3	Clustering analysis of gene abundance estimation using Heatmaps based on de novo assembled transcript	png file (.png)	https://doi.org/10.6084/m9.figshare.21701357 [15]

Objective

Micro-Tom is a cultivar of a tomato that is affected by a mutation in the IAA9 gene, which IAA9 is a family member of Auxin/IAA (Indole-Acetic-Acid) transcription factors (T.F.) in tomato. The main role of this gene is formation in fruit-set. This antisense technology in the plant using AS-IAA9 shows several developmental defects, including strong parthenocarpy behavior related to IAA [1]. Some advantages of this mutant are higher production or yield in fruits, and it can survive under drought stress conditions [2].

Micro-Tom is small in size, has rapid growth and life cycle, easy transformation, and a short life cycle for fruit harvest [3], making Micro-Tom a convenient model for research in different fields. Several studies have been conducted on tomato genetics, like hormonal functions and interactions, carbohydrate metabolism, amino acids metabolism, and molecular breeding of tomato fruit shelf-life. The phenotype of Micro-Tom is due to at least three mutations, one of them is a dwarf (internode length reduction and smaller, rugose, dark-green leaves production) [4].

Data description

A total of three RNA libraries (wild type, SlIAA9-3, SlIAA9-5) were prepared and sequenced (Data set 1, https://ddbj.nig.ac.jp/resource/sra-study/DRP009326). RNA-seq was performed using MinION ONT (Oxford Nanopore Technologies). Transcriptome sequencing had an estimated read of 578 374, 409 905, and 851 948 for wild type, SlIAA9-3, and SlIAA9-5, respectively. The results of sequencing and pre-processing are summarized in Data file 1 (Table 1, https://doi.org/10.6084/m9.figshare.21701345). After obtaining cleaned reads, each sample was mapped to the tomato reference genome (*S. lycopersicum* ITAG4.0) with the Minimap2 program. In particular, 965 genes were expressed only in the iaa9-3 mutant, and 2332 genes were expressed only in the iaa9-5 (Data file 2, https://doi.org/10.6084/m9.figshare.21701354). In cluster analysis using the heatmap method, separate groups were obtained between the wild type and the two mutants (Data file 3, https://doi.org/10.6084/m9.figshare.21701357).

The total RNA from young leaves was extracted using the RNeasy PowerPlant Kit (Qiagen) following the manufacturer's protocol. The quality and quantity of RNA were checked by Nanophotometer NP-80 (Implen) and Qubit™ RNA Broad Range (B.R.) assay on Qubit® Fluorometer (Invitrogen). Then, the total RNA was subjected to RNA sequencing using PCR-cDNA Barcoding kit - SQK-PCB109 (PCB_9092_v109_revB_10Oct2019) [5]. The sequencing was performed on a Flow Cell R9.4.1 (FLO-MIN106D) on MinION Mk1B. After sequencing, the raw reads were base called using Guppy 6.1.2 with default parameters [6]. Next, data pre-processing followed https://github.com/felixgrunberger/microbepore protocol includes demultiplexing and NanoStat v1.2.1 to assess the reads quality and reads' statistics [5, 7]. Next, full-length reads with remaining SSP (strand-switching primer) and VNP (oligo-dT30VN) primers were identified using pychopper v2.5.0 (https://github.com/nanoporetech/pychopper). Then, polyA-tails and the remaining SSP adapters were removed using Cutadapt [8]. The cleaned reads were mapped to the public tomato reference genome (*S. lycopersicum* ITAG4.0) using Minimap2 [6, 9]. To estimate gene abundance in each sample, the mapped-clean reads were calculated in alignment-based mode using salmon v1.9.0 [10]. Finally, transcripts per million (TPM) from each treatment were compared using clustering analysis by using RStudio 4.1.2 version [11] with some packages; gplots, cluster, and heatmap2.

Limitations

This study had limitations in obtaining a good-quality total RNA without degradation and fragmentation during library construction. In addition, the heat stress treatment which stresses makes the plant difficult to survive in high heat conditions.

Abbreviations

RNA	Ribonucleic acid
RNA-seq	RNA sequencing

Acknowledgements

The authors would like to thank the Seed Center Laboratory, Leuwikopo Experimental Station, Dept. of Agronomy and Horticulture, Faculty of Agriculture, IPB University, for providing sample preparation and sequencing.

Authors' contributions

DDM, NJ, AW, SM: Investigation, resources, writing-original draft. RP, DDM, SM: Conceptualization and funding acquisition. DDM, SM, HE: Writing-reviewing

and editing. WMYL, MA, DPH: Methodology, software, validation, formal analysis, data curation. All authors read and approved the final manuscript.

Funding
This research was funded by the Ministry of Education, Culture, Research, and Technology (KEMENDIKBUDRISTEK), the Republic of Indonesia, through the Program Riset Kolaborasi Indonesia (PRKI) 2022 to Deden Derajat Matra as P.I. with grant number: 3332/IT3.L1/PT.01.03/P/B/2022.

Data availability
Raw FASTQ files were deposited in the DDBJ database under accession number DRP009326 https://ddbj.nig.ac.jp/resource/sra-study/DRP009326 [9].

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 9 December 2022 / Accepted: 6 March 2023
Published online: 20 March 2023

References

1. Wang H, Jones B, Li Z, Frasse P, Delalande C, Regad F, Cha-abouni S, Latche A, Pech J-C, Bouzayen M. The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. Plant Cell. 2005;17:2676–92. https://doi.org/10.1105/tpc.105.033415.
2. Ariizumi T, Shinozaki Y, Ezura H. Genes that influence yield in tomato. Breed Sci. 2013;63:3–13. https://doi.org/10.1270/jsbbs.63.3.
3. Sun HJ, Uchii S, Watanabe S, Ezura H. A highly efficient transformation protocol for Micro-Tom, a model cultivar for tomato functional genomics. Plant Cell Physiology. 2006;47:426–31. https://doi.org/10.1093/pcp/pci251.
4. Gonzales C, Ré MD, Sossi ML, Valle EM, Boggio SB. Tomato cv. 'Micro-Tom' as a model system to study postharvest chilling tolerance. Sci Hort. 2015;184:63–9. https://doi.org/10.1016/j.scienta.2014.12.020.
5. Matra DD, Adrian M, Kusuma J, Duminil J, Poerwanto R. Dataset from de novo transcriptome assembly of Myristica fatua leaves using MinION nanopore sequencer. Data Brief. 2022;46:108838.
6. Adrian M, Poerwanto R, Inoue E, Matra DD. Transcriptome Dataset of Strawberry (*Fragaria* × *ananassa* Duch. Data. 2023;8(2):22. https://doi.org/10.3390/data8020022. Leaves Using Oxford Nanopore Sequencing under LED Irradiation and Application of Methyl Jasmonate and Methyl Salicylate Hormones Treatment.
7. Martin M. Cutadapt removes adapter sequences from highthroughput sequencing reads. EMBnet J. 2011;17:10.
8. de la Rubia I, Srivastava A, Xue W, Indi JA, Carbonell-Sala S, Lagarde J, Albà MM, Eyras ERATTLE. Reference-free reconstruction and quantification of transcriptomes from Nanopore Long-Read Sequencing. BioRxiv 2020, 02, 939–942.
9. Li H, Birol I. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34:3094–100.
10. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14:417–9.
11. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria. 2016. Available online: https://www.R-project.org/ (accessed on 1 December 2022)
12. DNA Data Bank of Japan. 2022. https://ddbj.nig.ac.jp/resource/sra-study/DRP009326. Accessed 12 Des 2022.
13. Matra DD. 2022a. Summary of raw and clean reads and transcriptome assembly. figshare. Dataset. https://doi.org/10.6084/m9.figshare.21701345
14. Matra DD. 2022b. Venn diagram for comparison of the number of expressed genes in wild type, iaa9-3 and iaa9-5 mutants. figshare. Figure. https://doi.org/10.6084/m9.figshare.21701354
15. Matra DD. 2022c. Clustering analysis of gene abundance estimation using Heatmaps based on de novo assembled transcript. figshare. Figure. https://doi.org/10.6084/m9.figshare.21701357.v1

Publisher's note

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