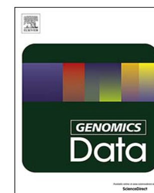
Genomics Data 14 (2017) 5–6

Contents lists available at ScienceDirect

Genomics Data

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



RNA-seq data from different developmental stages of *Rafflesia cantleyi* floral buds

Safoora Amini[a], Halimah Alias[b], Mohd Afiq Aizat-Juhari[c], Mohd-Noor Mat-Isa[b], Jumaat Haji Adam[c], Hoe-Han Goh[d], Kiew-Lian Wan[a,*]

[a] School of Biosciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
[b] Malaysia Genome Institute, Jalan Bangi, 43000 Kajang, Selangor, Malaysia
[c] School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
[d] Institute of Systems Biology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

A R T I C L E   I N F O

Keywords:
Floral bud
Rafflesia
RNA-seq
Transcriptome

A B S T R A C T

Rafflesia cantleyi, known as one of the world's largest flowers, is a specialised holoparasite due to dramatic morphological modifications. It possesses highly reduced vegetative structure and only appears as a flower for sexual reproduction. Moreover, it has an unusual life cycle in that its floral bud development takes up to nine months. In order to fully understand the highly modified floral organ structure and long life cycle of *R. cantleyi*, we used Illumina sequencing technology (HiSeq) for sequence generation followed by *de novo* assembly of sequence reads. We obtained the RNA-seq data from three different stages of floral bud, representing the early, mid and advanced developmental stages. These data are available *via* BioProject accession number PRJNA378435. More than 10.3 Gb raw sequence data were generated, corresponding to 102,203,042 raw reads. Following removal of low-quality reads and trimming of adapter sequences, a total of 91,638,836 reads were obtained. *De novo* assembly of these sequences using Trinity resulted in 89,690 unique transcripts with an N50 of 1653 bp. The obtained transcriptomic data will be useful for further study to understand the molecular interactions that result in *R. cantleyi* floral development.

Specifications	
Organism	*Rafflesia cantleyi*
Sequencer	Illumina HiSeq™ 2000
Data format	Raw
Experimental factors	Floral buds from early, mid and advanced developmental stages
Experimental features	Inner tissue of individual floral buds were sampled for RNA extraction, sequencing and *de novo* transcriptome assembly
Consent	Not applicable
Sample source location	Raub, Pahang, Malaysia (3° 47′ 24″ N, 101° 51′ 25″E)

1. Direct link to deposited data

https://www.ncbi.nlm.nih.gov/bioproject/PRJNA378435

2. Introduction

Rafflesia cantleyi is an endophytic holoparasite plant of the tropical rainforests of South-east Asia belonging to the family *Rafflesiaceae*, subfamily *Rafflesia*. It depends entirely on its specific host, *Tetrastigma* (Vitaceae) [1]. It has no apparent leaves, stems or roots, and only appear as a flower of up to 0.6 m in diameter for sexual reproduction [2]. This large flower includes five highly modified perianth and a petal-derived diaphragm, which together formed a large, bowl-shaped structure called floral chamber [3]. Previous studies have provided insights on its morphological evolution [3,4]; yet, the floral organs development in this giant flower remains poorly understood.

Rafflesia is also known for its long life cycle, which takes an average of 9 months for the flower to be fully opened [1]. The buds of *R. cantleyi* first emerge through the bark of *Tetrastigma* with the bracts still covering them. They then grow progressively and upon maturation, the buds open gradually over a 24 to 48-hour period. Previous work has shed light on the genes potentially involved in the growth and development of the *R. cantleyi* flower [5]. However, knowledge on the

* Corresponding author.
E-mail address: klwan@ukm.edu.my (K.-L. Wan).

http://dx.doi.org/10.1016/j.gdata.2017.07.008
Received 5 April 2017; Received in revised form 6 July 2017; Accepted 13 July 2017
Available online 18 July 2017
2213-5960/ © 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

Table 1
Statistics of *Rafflesia cantleyi* floral bud transcriptome assembly.

Attributes	Floral bud stage 1	Floral bud stage 2	Floral bud stage 3	Merged assembly
Pre-assembly				
Total raw reads	36,776,468	29,428,456	35,998,118	102,203,042
Total processed reads	32,543,575	26,453,373	32,641,888	91,638,836
Post-assembly				
Number of unigenes	48,825	28,950	31,053	70,467
Number of unique transcripts	67,055	44,300	44,827	89,690
N50 (bp)	1765	1972	1943	1653
Size range (bp)	200–15,281	200–15,218	200–16,493	200–16,553

biological processes that occur during *R. cantleyi* flower transition is still scarce. In this study, we carried out sequencing and *de novo* assembly of floral bud transcriptome from three developmental stages to provide essential data for further studies on these aspects of *R. cantleyi*.

3. Experimental design, materials and methods

3.1. Plant materials

Rafflesia cantleyi floral bud samples from different developmental stages were collected from Raub, Pahang, Malaysia. Cross section of the floral buds showed that they represent the early (floral bud stage 1), mid (floral bud stage 2) and advanced (floral bud stage 3) developmental stages. Floral bud stage 1 is made up of undifferentiated masses of cells, floral bud stage 2 contains moderately differentiated and visible internal organs, while floral bud stage 3 consists of more developed internal organs. These *R. cantleyi* floral buds, attached to their host, *Tetrastigma*, were carefully dissected from the host plant. The buds were surface sterilised using 10% (*v*/v) Clorox® solution (1% sodium hypochlorite), followed by three rinses with sterile water.

3.2. RNA isolation and cDNA library construction

Inner tissues of the floral buds were cut and flash-frozen in liquid nitrogen in the field. All samples were then stored at -80 °C until further use. Total RNA was extracted using a modified CTAB extraction protocol based on previously described methods [6–8]. RNA purity and integrity was measured using the ND-1000 Nanodrop spectrophotometer (Thermo Fisher Scientific Inc., USA) and Agilent 2100 bioanalyzer (Agilent Technologies, USA), respectively. cDNA library preparation was performed using the TruSeq® Stranded Total RNA Library Prep Kit (Illumina, USA) according to the manufacturer's protocol. The libraries were sequenced using the Illumina HiSeq™ 2000 platform.

3.3. Transcriptome de novo assembly

Approximately 102.2 million bp paired-end reads were generated using the Illumina HiSeq™ 2000 platform. Raw reads from all the three data sets were filtered using Trimmomatic [9] (default setting SLIDINGWINDOW:4:5, LEADING:5, TRAILING:5, MINLEN:25) to remove adaptor sequence or low quality reads. *De novo* assembly of clean reads was performed using Trinity (v2.0.6) [10] with the default k-mer size of 25. Table 1 shows the transcriptome assembly statistics that were generated utilising the TrinityStats.pl script.

In conclusion, we have generated and *de novo* assembled transcriptome sequences from the floral buds of *R. cantleyi*. These sequences can be used for gene discovery, further studies on the regulation of *Rafflesia* growth and development, and comparative transcriptome analysis with other *Rafflesia* species.

Conflicts of interest

The authors declare no conflicts of interest in this study.

Acknowledgements

This work was supported financially by Research University grants from Universiti Kebangsaan Malaysia [grant numbers LAUREATE-2013-001 and GUP-2016-008].

References

[1] J. Nais, *Rafflesia* of the World, Sabah Parks, Kota Kinabalu, Sabah, (2001).
[2] W. Meijer, Rafflesiaceae, Flora Malesiana, 13 (1997), pp. 1–42.
[3] L.A. Nikolov, P.K. Endress, M. Sugumaran, S. Sasirat, S. Vessabutr, E.M. Kramer, C.C. Davis, Developmental origins of the world's largest flowers, Rafflesiaceae, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 18578–18583.
[4] C.C. Davis, P.K. Endress, D.A. Baum, The evolution of floral gigantism, Curr. Opin. Plant Biol. 11 (2008) 49–57.
[5] X.-W. Lee, M.-N. Mat-Isa, N.-A. Mohd-Elias, M.A. Aizat-Juhari, H.-H. Goh, P.H. Dear, K.-S. Chow, J.H. Adam, R. Mohamed, M. Firdaus-Raih, K.-L. Wan, Perigone lobe transcriptome analysis provides insights into *Rafflesia cantleyi* flower development, PLoS One 11 (2016) e0167958.
[6] J.J. Doyle, J.L. Doyle, A rapid DNA isolation procedure for small quantities of fresh leaf tissue, Phytochem. Bull. 19 (1987) 11–15.
[7] K.E. Reid, N. Olsson, J. Schlosser, F. Peng, S.T. Lund, An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development, BMC Plant Biol. 6 (2006) 27.
[8] R.H. Japelaghi, R. Haddad, G.A. Garoosi, Rapid and efficient isolation of high quality nucleic acids from plant tissues rich in polyphenols and polysaccharides, Mol. Biotechnol. 49 (2011) 129–137.
[9] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (2014) 2114–2120.
[10] B.J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P.D. Blood, J. Bowden, M.B. Couger, D. Eccles, B. Li, M. Lieber, M.D. Macmanes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C.N. Dewey, R. Henschel, R.D. Leduc, N. Friedman, A. Regev, *De novo* transcript sequence reconstruction from RNA-seq using the Ttrinity platform for reference generation and analysis, Nat. Protoc. 8 (2013) 1494–1512.