



De novo transcriptome assembly of two *Vigna angularis* varieties collected from Korea



Yeonhwa Jo ^{a,1}, Sen Lian ^{b,1}, Jin Kyong Cho ^c, Hoseong Choi ^a, Sang-Min Kim ^d, Sun-Lim Kim ^d, Bong Choon Lee ^d, Won Kyong Cho ^{a,c,*}

^a Department of Agricultural Biotechnology, College of Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Republic of Korea

^b College of Crop Protection and Agronomy, Qingdao Agricultural University, Qingdao, Shandong 266109, China

^c The Taejin Genome Institute, Gadam-gil 61, Hoengseong 25239, Republic of Korea

^d Crop Foundation Division, National Institute of Crop Science, RDA, Wanju 55365, Republic of Korea

ARTICLE INFO

Article history:

Received 28 April 2016

Received in revised form 2 May 2016

Accepted 3 May 2016

Available online 5 May 2016

Keywords:

Adzuki bean

RNA-Seq

Transcriptome

Variety

Vigna angularis

ABSTRACT

The adzuki bean (*Vigna angularis*), a member of the family Fabaceae, is widely grown in Asia, from East Asia to the Himalayas. The adzuki bean is known as an ingredient that adds sweetness to diverse desserts made in Eastern Asian countries. Libraries prepared from two *V. angularis* varieties referred to as Taejin Black and Taejin Red were paired-end sequenced using the Illumina HiSeq 2000 system. The raw data in this study can be available in NCBI SRA database with accession numbers of SRR3406660 and SRR3406553. After *de novo* transcriptome assembly using Trinity, we obtained 324,219 and 280,056 transcripts from Taejin Black and Taejin Red, respectively. We predicted a total of 238,321 proteins and 179,519 proteins for Taejin Black and Taejin Red, respectively, by the TransDecoder program. We carried out BLASTP on the predicted proteins against the Swiss-Prot protein sequence database to predict the putative functions of identified proteins. Taken together, we provide transcriptomes of two adzuki bean varieties by RNA-Seq, which might be usefully applied to generate molecular markers.

© 2016 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/>).

Specifications

Organism/cell line/tissue	Adzuki bean (<i>Vigna angularis</i>)/leaves
Sex	N.A.
Sequencer or array type	HiSeq2000
Data format	Raw and processed
Experimental factors	<i>De novo</i> transcriptome assembly of two adzuki bean varieties
Experimental features	Leaves of ten adzuki bean plants (five of the Taejin Black variety and five of the Taejin Red variety) were harvested for total RNA extraction. Two prepared libraries were paired-end sequenced using the HiSeq 2000 system. The obtained data for each variety were subjected to <i>de novo</i> transcriptome assembly using Trinity, and coding regions were predicted by TransDecoder. We performed BLASTP against the Swiss-Prot protein database to annotate the identified proteins.
Consent	N/A
Sample source location	Hoengseong, South Korea (37°28'49.6"N 127°58'34.3"E)

* Corresponding author at: Department of Agricultural Biotechnology, College of Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Republic of Korea.

E-mail address: wonkyong@gmail.com (W.K. Cho).

¹ These authors contributed equally to this work.

1. Direct link to deposited data

<http://www.ncbi.nlm.nih.gov/sra/SRR3406660> for *Vigna angularis* variety Taejin Black.

<http://www.ncbi.nlm.nih.gov/sra/SRR3406553> for *V. angularis* variety Taejin Red.

2. Introduction

The adzuki bean (*V. angularis*), a member of the family Fabaceae, is widely grown in Asia, from East Asia to the Himalayas [1]. Particularly, adzuki beans are known as an ingredient that adds sweetness to diverse desserts made in Eastern Asian countries, such as China, Japan, Korea, and Taiwan [2]. Most adzuki bean cultivars in Eastern Asian countries are red in color; however, some cultivars display several different colors, such as white, black, and gray. The adzuki bean is relatively small (about 5 mm). According to a previous study, the adzuki bean originated from northeast Asia [3]. The adzuki bean is a diploid legume, and its draft genome contains approximately 26,857 genes [2,4]. In this study, we performed *de novo* transcriptome assembly for two adzuki bean varieties by RNA-Seq.

3. Experimental design, materials, and methods

3.1. Plant materials

Adzuki bean plants for two varieties were grown in a field located in Gadam-ri, Hoengseong-up, South Korea. Leaves from five plants for each variety were harvested and immediately frozen in liquid nitrogen for further experiments.

3.2. RNA isolation, library preparation, and sequencing

Ten leaves collected from five plants were pooled and used for total RNA extraction using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). For mRNA library preparation, we used a TruSeq RNA Library Prep Kit v2 according to the manufacturer's instructions (Illumina, San Diego, U.S.A.). In brief, the poly-A-containing mRNAs were isolated using poly-T oligo-attached magnetic beads. The first strand of cDNA, followed by a second strand of cDNA, was synthesized from purified mRNAs. End repair was performed followed by adenylation of 3' ends. Adapters were ligated, and PCR was conducted to selectively enrich DNA fragments with adapters and to amplify the amount of DNA in the library, respectively. The quality control of generated libraries was conducted using a 2100 Bioanalyzer (Agilent, Santa Clara, U.S.A.). The libraries were paired-end sequenced by Macrogen Co. (Seoul, South Korea) using the HiSeq 2000 platform.

3.3. De novo transcriptome assembly, identification of protein coding regions, and annotation

Libraries prepared from two *V. angularis* varieties referred to as Taejin Black and Taejin Red, respectively, were paired-end sequenced using the Illumina HiSeq 2000 system. We obtained 9.474 GB and 9.116 GB of raw data from Taejin Black and Taejin Red, respectively. We *de novo* assembled transcriptomes of two adzuki bean varieties using Trinity program ver. 2.0.6 [5]. Detailed information on the *de novo* transcriptome assembly of the two *V. angularis* varieties is summarized in Table 1. We obtained 324,219 and 280,056 transcripts from Taejin Black and Taejin Red, respectively. The numbers of components for Taejin Black and Taejin Red were 94,654 and 93,698, respectively. The N50 values for Taejin Black and Taejin Red were 2634 bp and

2461 bp, respectively. The median contig lengths for Taejin Black and Taejin Red were 1504 bp and 1301 bp, respectively. We analyzed coding regions within the assembled transcripts by the TransDecoder program implemented in the Trinity software distribution. As a result, we predicted a total of 238,321 proteins and 179,519 proteins for Taejin Black and Taejin Red, respectively. We carried out BLASTP on the predicted proteins against the Swiss-Prot protein sequence database to predict the putative functions of the identified proteins. The BLAST results showed that most proteins were matched to known proteins in the Swiss-Prot database except 41,619 proteins and 31,720 proteins for Taejin Black and Taejin Red, respectively. Many proteins were derived from eukaryotes (17,274 and 12,123 proteins for Taejin Black and Taejin Red, respectively) followed by bacteria, viruses, and archaea. The transcriptome data in this study will be usefully applied to generate molecular markers for the adzuki bean.

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgments

This work was carried out with the support of the "Cooperative Research Program for Agriculture Science & Technology Development (project no. PJ01186102)" conducted by the Rural Development Administration, Republic of Korea. This work is dedicated to the memory of my father, Tae Jin Cho (1946–2015).

References

- [1] X.X. Zong, A. Kaga, N. Tomooka, X.W. Wang, O.K. Han, D. Vaughan, The genetic diversity of the *Vigna angularis* complex in Asia. *Genome* 46 (2003) 647–658.
- [2] Y.J. Kang, D. Satyawati, S. Shim, T. Lee, J. Lee, W.J. Hwang, S.K. Kim, P. Lestari, K. Laosatit, K.H. Kim, Draft genome sequence of adzuki bean, *Vigna angularis*. *Sci. Rep.* 5 (2015).
- [3] G.-A. Lee, Archaeological perspectives on the origins of adzuki (*Vigna angularis*). *The Holocene* 23 (2013) 453–459.
- [4] A. Parida, S. Raina, R. Narayan, Quantitative DNA variation between and within chromosome complements of *Vigna* species (Fabaceae). *Genetica* 82 (1990) 125–133.
- [5] M.G. Grabherr, B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29 (2011) 644–652.

Table 1

Summary of *de novo* assembled transcriptomes for two adzuki bean varieties (Taejin Black and Taejin Red).

Index	Taejin Black	Taejin Red
Total trinity transcripts	324,219	280,056
Total trinity components	94,654	93,698
Percent GC	40.82	40.73
Contig N50	2634	2461
Median contig length	1504	1301
Average contig	1772.39	1621.81
Total assembled bases	574,641,146	454,197,434