Genomics Data 9 (2016) 35–36

Contents lists available at ScienceDirect

Genomics Data

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

Data in Brief

De novo transcriptome assembly of 'Angeleno' and 'Lamoon' Japanese plum cultivars (*Prunus salicina*)

Máximo González [a,*], Jonathan Maldonado [b], Erika Salazar [c], Herman Silva [b], Basilio Carrasco [a,*]

[a] Facultad de Agronomía e Ingeniería Forestal, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Casilla 306, Macul, Santiago, Chile
[b] Universidad de Chile, Facultad de Ciencias Agronómicas, Departamento de Producción Agrícola, Laboratorio de Genómica Funcional & Bioinformática. Av. Santa Rosa 11315, La Pintana, Santiago, Chile
[c] Centro Regional de Investigación La Platina, Unidad de Recursos Genéticos y Banco de Germoplasma, Instituto de Investigaciones Agropecuarias. Av. Santa Rosa 11610, La Pintana, Santiago, Chile

ARTICLE INFO

Article history:
Received 16 June 2016
Accepted 18 June 2016
Available online 19 June 2016

Keywords:
Japanese plum
De novo transcriptome
Skin
Color

ABSTRACT

Japanese plum (*Prunus salicina* L.) is a fruit tree of the Rosaceae family, which is an economically important stone fruit around the world. Currently, Japanese plum breeding programs combine traditional breeding and plant physiology strategies with genetic and genomic analysis. In order to understand the flavonoid pathway regulation and to develop molecular markers associated to the fuit skin color (EST-SSRs), we performed a next generation sequencing based on Illumina Hiseq2000 platform. A total of 22.4 GB and 21 GB raw data were obtained from 'Lamoon' and 'Angeleno' respectively, corresponding to 85,404,726 raw reads to 'Lamoon' and 79,781,666 to 'Angeleno'. A total of 139,775,975 reads were filtered after removing low-quality reads and trimming the adapter sequences.

De novo transcriptome assembly was performed using CLC Genome Workbench software and a total of 54,584 unique contigs were generated, with an N50 of 1343 base pair (bp) and a mean length of 829 bp. This work contributed with a specific Japanese plum skin transcriptome, providing two libraries of contrasting fruit skin color phenotype (yellow and red) and increasing substantially the GB of raw data available until now for this specie.

Specifications	
Organism/cell line/tissue	*Prunus salicina* L./Fruit skin
Sex	N/A
Sequencer or array type	Illumina HiSeq2000
Data format	Raw data: FASTAQ file
Experimental factors	Transcriptome profiling of two different *Prunus salicina* cultivars.
Experimental features	RNAs were isolated from the skin of fruits with contrasting skin colors, where 'Angeleno' and 'Lamoon' have red and yellow skin, respectively.
Consent	N/A
Sample source location	Pirque, Región Metropolitana, Chile (33°40′12″S–70°35′06″W)

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/sra/SRR3657496 for *Prunus salicina* cultivar 'Lamoon'.

http://www.ncbi.nlm.nih.gov/sra/SRR3657497 for *Prunus salicina* cultivar 'Angeleno'.

2. Introduction

Japanese plum (*Prunus salicina* L.) is one of the most economically important stone fruit of the Rosaceae family worldwide. Currently, several plant-breeding programs are incorporating, onto their conventional breeding plant breeding and plant physiology strategies, innovative biotechnological methods including functional genomics, mutations, and molecular markers in order to increase their efficiency in the development of new varieties [1]. Despite of the importance of Japanese plum, the availability of genetic information is scare. Carrasco et al. [2] analyzed the genetic relationships among Japanese plums cultivars using SSR and Inter Simple Sequence Repeat. Recently, Jo et al. [3] has published the first transcriptome to Japanese plum, reporting 17.72 GB of raw data.

Transcriptome studies facilitate the understanding of metabolic pathways, opening the possibility to associate transcription factors with traits of agronomic interest. For the analysis of the anthocyanins biosynthesis in Japanese plum fruits (unpublished data), we performed *the novo* transcriptome based on two contrasting cultivars for skin anthocyanin accumulation. Previously, our group used this information to developed 44 EST-SSRs from genes involved on citric acid

* Corresponding authors.
E-mail addresses: mpgonza2@uc.cl (M. González), bcarrasco@uc.cl (B. Carrasco).

http://dx.doi.org/10.1016/j.gdata.2016.06.010

metabolism, carbohydrate metabolism and the flavonoid pathway. Three EST-SSR markers from the putative flavonoid pathway transcription factors, *PsMYB10*, *PsMYB1* and *PsbHLH35*, which allow separate colored from uncolored fruit skins were developed [4]. Here we provide the transcriptome and the two libraries of contrasting skin color phenotype, used in both studies.

3. Experimental design, materials and methods

3.1. Plant material

Fruits at harvest maturity of cultivars 'Lamoon' (yellow skin) and 'Angeleno' (red skin) were collected from a Japanese plum orchard at the Experimental Field "Pirque" (33°40′12″S–70°35′06″W) that belong to the Faculty of Agronomy and Forest Engineering of the Pontificia Universidad Católica de Chile. Skins were separated from the pulp and immediately frozen in liquid nitrogen and stored at −80 °C until use.

3.2. RNA extraction, library construction and deep sequencing

Skins of 30 fruits from five to ten trees per cultivar were pooled and used for total RNA extraction using the protocol of Meisel et al. [5]. Library preparation and paired-end high-throughput sequencing for each sample was performed at a contract sequencing facility (Macrogen, Inc. Seoul, South Korea). The cDNA library was sequenced using an Illumina HiSeq2000 platform.

3.3. De novo transcriptome assembly and annotation

We obtained a total of 22.4 GB and 21 GB raw data from 'Lamoon' and 'Angeleno' respectively, corresponding 85,404,726 raw reads to 'Lamoon' and 79,781,666 to 'Angeleno' (Table 1). A total of 139,775,975 reads were filtered after removing low-quality reads and trimming the adapter sequences.

De novo transcriptome assembly was performed using CLC Genome Workbench software (version 4.8). We used a hybrid assembly strategy, where the reads are assembled into the contigs from a pool of all the paired end short-read data, using the following parameters: similarity = 0.95; length fraction = 0.7; insertion/deletion cost = 3 and mismatch cost = 3.

The assembly generated 54,584 unique contigs, with an N50 of 1343 bp and mean length of 829 bp (Table 1), which are available in Supplementary File 1. While Jo et al. [3] published a leaf transcriptome, we contributed with the specific transcriptome of the fruit skin, providing two libraries of contrasting skin color phenotype and doubling, to our knowledge, the GB of raw data available to Japanese plum.

Table 1
Summary of the sequencing and *de novo* assembly.

	'Angeleno'	'Lamoon'	*De novo* assembly
No. of pb	8,057,948,266	8,625,877,326	13,546,478,792
No. of reads	79,781,666	85,404,726	139,775,975
%GC	45.54	45.16	40.1
Q20	96.92	95.6	
No. of contigs			54,584
Length of N50			1343
Average length			829

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgments

This work was funded by the Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) through Scholarships Doctorado en Chile 2009 No 21090118, Apoyo a la Tesis Doctoral 2011 No. 24110179, Tesis en la Industria 2013 No. 781211008 and FONDECYT/RegularNo. 1120261; Consorcio Tecnológico de Industria Hortofruticula S.A. and Fondo Instituto de Investigaciones Agropecuarias — Ministerio de Agricultura (INIA-MINAGRI) Project No 501453-70.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gdata.2016.06.010.

References

[1] B. Carrasco, L. Meisel, M. Gebauer, R. Garcia-Gonzales, H. Silva, Breeding in peach, cherry and plum: from a tissue culture, genetic, transcriptomic and genomic perspective. Biol. Res. 46 (2013) 219–230.
[2] B. Carrasco, C. Díaz, M. Moya, M. Gebauer, R. García-Gonzáles, Variability and genetic relationships among Japanese plums cultivars (*Prunus salicina* Lindl.) using SSR and inter simple sequence repeat. Rev. Cs. Inv. Agr. 39 (2012) 533–543.
[3] Y. Jo, S. Lian, J.K. Cho, H. Choi, H. Chu, W.K. Cho, De novo transcriptome assembly of two different *Prunus salicina* cultivars. Genomics Data 6 (2015) 262–263.
[4] M. Gonzalez, E. Salazar, J. Castillo, P. Morales, I. Mura-Jornet, J. Maldonado, H. Silva, B. Carrasco, Genetic structure based on EST-SSR: a putative tool for fruit color selection in Japanese plum (*Prunus salicina* L.) breeding programs. Mol. Breed. 36 (6) (2016) 1–15, http://dx.doi.org/10.1007/s11032-016-0491-x.
[5] L. Meisel, B. Fonseca, S. González, R. Baeza-Yates, V. Cambiazo, R. Campos, M. Gonzalez, A. Orellana, J. Retamales, H. Silva, A rapid and efficient method for purifying high quality total RNA from peaches (*Prunus persica*) for functional genomics analyses. Biol. Res. 38 (2005) 83–88.