



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Shoot transcriptome of the giant reed,
Arundo donax

Roberto A. Barrero^a, Felix D. Guerrero^b, Paula Moolhuijzen^a,
John A. Goolsby^c, Jason Tidwell^c, Stanley E. Bellgard^d,
Matthew I. Bellgard^{a,*}

^a Centre for Comparative Genomics, Murdoch University, Murdoch, Australia

^b United States Department of Agriculture, Agricultural Research Service, Knippling-Bushland U. S. Livestock Insects Research Laboratory, Kerrville, USA

^c United States Department of Agriculture, Agricultural Research Service, Cattle Fever Tick Research Laboratory, Edinburg, USA

^d Landcare Research Ltd, Private Bag 92170, Auckland 1142, New Zealand

ARTICLE INFO

Article history:

Received 6 October 2014

Received in revised form

11 December 2014

Accepted 19 December 2014

Available online 22 January 2015

Keywords:

Arundo donax

Giant reed

Shoot

Transcriptome

RNA *de novo* assembly

RNA-Seq

ABSTRACT

The giant reed, *Arundo donax*, is a perennial grass species that has become an invasive plant in many countries. Expansive stands of *A. donax* have significant negative impacts on available water resources and efforts are underway to identify biological control agents against this species. The giant reed grows under adverse environmental conditions, displaying insensitivity to drought stress, flooding, heavy metals, salinity and herbaceous competition, thus hampering control programs. To establish a foundational molecular dataset, we used an Illumina Hi-Seq protocol to sequence the transcriptome of actively growing shoots from an invasive genotype collected along the Rio Grande River, bordering Texas and Mexico. We report the assembly of 27,491 high confidence transcripts (≥ 200 bp) with at least 70% coverage of known genes in other Poaceae species. Of these 13,080 (47.58%), 6165 (22.43%) and 8246 (30.0%) transcripts have sequence similarity to known, domain-containing and conserved hypothetical proteins, respectively. We also report 75,590 low confidence transcripts supported by both trans-ABBySS and Velvet-Oases *de novo* assembly pipelines. Within the low confidence subset of transcripts we identified partial hits to known (19,021; 25.16%), domain-containing (7093; 9.38%) and conserved hypothetical (16,647; 22.02%) proteins. Additionally 32,829 (43.43%) transcripts encode putative hypothetical proteins unique to *A. donax*. Functional annotation resulted in

* Corresponding author. Tel.: +61 8 9360 6088.

E-mail address: mbellgard@cgc.murdoch.edu.au (M.I. Bellgard).

5,550 and 6,070 transcripts with assigned Gene Ontology and KEGG pathway information, respectively. The most abundant KEGG pathways are spliceosome, ribosome, ubiquitin mediated proteolysis, plant–pathogen interaction, RNA degradation and oxidative phosphorylation metabolic pathway. Furthermore, we also found 12, 9, and 4 transcripts annotated as stress-related, heat stress, and water stress proteins, respectively. We envisage that these resources will promote and facilitate studies of the abiotic stress capabilities of this exotic plant species, which facilitates its invasive capacity.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Specifications table

Subject area	Biology
More specific subject area	RNA-seq transcriptome data of <i>Arundo donax</i>
Type of data	Table, figure
How data was acquired	2 × 100 HiSeq (single lane of 100 bases pair-end approach)
Data format	Raw FASTQ and processed FASTA
Experimental factors	10 g of actively growing shoot, excised approximately 20 cm above soil level
Experimental features	Assembled transcriptome of actively growing shoot tissue excised from <i>A. donax</i> grown in field plots
Data source location	Laredo, TX, USA
Data accessibility	Data is with this article and also available at http://www.ncbi.nlm.nih.gov/GBRH01000000 The assembled and annotated <i>A. donax</i> USA genotype Rio Grande RNA transcriptome has been deposited at DDBJ/EMBL/GenBank under the project accession PRJNA256910

Value of data

- First transcriptome sequence data made available in GenBank/DDBJ/Embl for the *A. donax* invasive Rio Grande basin genotype.
- The *A. donax* shoot transcriptome dataset provides insights into one of the fastest growing terrestrial plants [1].
- *A. donax* has high tolerance to abiotic stresses and its high invasive nature threatens many natural environments and ecosystems.
- The abundant biomass of *A. donax* plants makes it an ideal candidate for biofuel programs [2].

1. Experimental design, materials and methods

1.1 Plant tissue

Approximately 10 g of *A. donax* shoot tissue was excised from an actively growing shoot, approximately 20 cm above the soil surface of a field plot at the Cattle Fever Tick Research Laboratory, Edinburg, TX, USA. The plants were propagated from plants collected at Laredo, TX in 2008 and designated the Invasive Rio Grande Basin genotype. Excised shoot tissue was taken under natural non-stressed growth conditions and quickly transferred to small vials and placed in dry ice and maintained frozen at -80°C until transferred into liquid N_2 during the RNA purification steps.

Download English Version:

<https://daneshyari.com/en/article/175182>

Download Persian Version:

<https://daneshyari.com/article/175182>

[Daneshyari.com](https://daneshyari.com)