Genomics Data 3 (2015) 143–145

Contents lists available at ScienceDirect

Genomics Data

journal homepage: http://www.journals.elsevier.com/genomics-data/

Data in Brief

Transcriptome analysis of response to drought in poplar interspecific hybrids

Elena Barghini *, Rosa Maria Cossu, Andrea Cavallini, Tommaso Giordani

Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, I-56124 Pisa, Italy

ARTICLE INFO

Article history:
Received 17 December 2014
Accepted 13 January 2015
Available online 20 January 2015

Keywords:
Poplar
RNAseq
Drought
Gene expression
Illumina

ABSTRACT

To investigate the response of poplar hybrids to drought, leaves were collected from plants to which water was suspended for 8 and 13 days. After measuring the respective relative water content, RNAs were isolated from leaves of moderately and severely droughted plants and from control plants, and Illumina RNA sequencing was performed to analyze RNA synthesis in these tissues. Our data provide a resource (available at Gene Expression Omnibus database under GSE64044) to be employed for comparative analyses of drought response in different poplar species, with the long-term aim of developing strategies to improve plant productivity under drought.

Specifications	
Organism/cell line/tissue	*Populus deltoides* × *Populus nigra* hybrids 661200585 and 661200589
Sex	N/A
Sequencer or array type	Illumina HiSeq2000
Data format	Raw data: FASTQ files, analyzed data: txt files
Experimental factors	Tissues/organs
Experimental features	RNA-seq dataset for gene expression profiling in leaves of poplars exposed to drought
Consent	Full consent
Sample source location	Pisa, Italy

Direct link to deposited data

Deposited data can be found here: http://www.ncbi.nlm.nih.gov/geo/info/linking.html.

Experimental design, materials and methods

Sample collection and RNA isolation

Rooted cuttings (50–70 cm in length) of hybrids between *P. deltoides* (L155-079, female) and *P. nigra* (71077-2-308, male), produced at INRA, Orleans (France), were cultivated in 20 × 20 cm^2 pots in the greenhouse, under natural daylight conditions (750 $\mu m \cdot m^{-2} \cdot s^{-1}$, maximal photon flux density), with air temperature maintained at 17–29 °C, and relative humidity from 55% to 90%. During spring, some hybrids were normally watered and others were treated by suspending watering for 8 days and 13 days. At each of these stages (0, 8, and 13 days of treatment, Fig. 1), one fully expanded leaf per plant, at 6–8 internodes from apex, was collected (at the same time of the day). The leaf was divided into two portions, one was used for RNA isolation, the other was used to measure tissue hydration by determining the relative water content (RWC; RWC = 100 (FW − DW) / (TW − DW), where FW is the fresh weight, DW the dry weight and TW the turgid weight). The experimental design included 2 clones (biological replicates) × 3 treatments (control, C; moderate, D1; and severe drought, D2) × 2 hybrids (obtained from the same parents).

Total RNA was isolated from the leaves of single plants with different RWCs, according to the method described by Logemann et al. [1], followed by DNAse I (Roche) treatments according to the manufacturer's instructions, to completely remove genomic DNA contamination. Finally RNA was purified by phenol/chloroform extraction and precipitated following standard procedures.

Generation of RNA-seq data

RNA-seq libraries were generated using the TruSeq RNA-Seq Sample Prep kit according to the manufacturer's protocol (Illumina Inc., San Diego, CA). Poly-A RNA was isolated from total RNA and chemically fragmented. First and second strand cDNA syntheses were followed by end repair and adenosines were added to the 3′ ends. Adapters were ligated to the cDNA and 200 ± 25 bp fragments were gel purified and enriched by PCR. The libraries were quantified using Bioanalyzer 2100

* Corresponding author. Tel.: +39 50 2216671.
E-mail address: elena.barghini@gmail.com (E. Barghini).

http://dx.doi.org/10.1016/j.gdata.2015.01.004

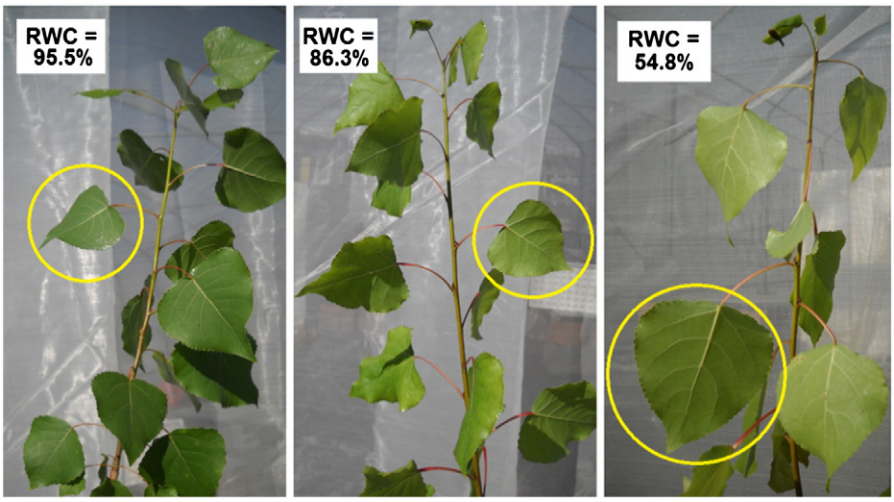

Fig. 1. Leaves (yellow circles) of *P. deltoides* × *P. nigra* plants used in the experiments. From left to right: normally watered, moderately, and severely droughted plants. For each leaf is reported the relative water content.

(Agilent Technologies, Santa Clara, CA) and run on the Illumina HiSeq2000 (Illumina Inc.) using version 3 reagents. Single-read sequences of length 51 bp were collected. Whole RNA-seq data were submitted to NCBI Sequence Read Archive and Gene Expression Omnibus (series accession number GEO64044).

Data processing

Raw single reads (in FASTQ format) were subjected to sequence quality control using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). FastQC performed a series of analysis modules on raw data and created a report with statistics for the data analyzed. For each library, FastQC showed high per base sequence quality, exceeding 30 on Phred scale (less than 1/1000 chance of a base being wrong) and detected adapter contamination, matching the reads to known adapter sequences. Then, raw reads were trimmed using Trimmomatic [2], version 0.32, with the following parameters: ILLUMINACLIP:2:30:10 LEADING:15 TRAILING:15 SLIDINGWINDOW:4:15 MINLEN:51.

After trimming, FastQC was used again to examine the characteristics of the libraries and to verify trimming efficiency, which resulted into removal of 4,906,505 (6.4%) low quality sequences (Table 1).

Differential expression quantification

Reads were aligned to the *Populus trichocarpa* unigene model database version 9.1, available at the Phytozome site (http://www.phytozome.net/poplar) [3], with a tolerance of up to 2 mismatches, using CLC-BIO Genomic Workbench 5.1. The percentage of aligned reads per each sample is reported in Table 1. We used the available *P. trichocarpa* database because this botanical species can be crossed with both *P. deltoides* and *P. nigra* and has diverged from them only between 8 and 13 Ma ago [4].

Gene expression level was calculated and expressed as Reads Per Kilobase per Million reads mapped (RPKM [5]). Expression profiles were evaluated considering RPKM values in C, D1, and D2 plants using Baggerly's test [6]. The weighted proportion fold changes between treatments were considered as significant when the weight of a sample was at least two-fold higher or lower than another, with an FDR corrected p-value ≤ 0.05, according to Baggerly's test. The analysis was limited to genes showing one or more total reads per million in at least one of the two clones of the two hybrids, as used in other studies [7]. Gene expression profiles were subdivided into nine groups, those remaining constant, those increasing their expression in D1 or in D2, or in both drought treatments, those reducing their expression in D1 or D2, or in both treatments, those increasing their expression in D1 and reducing in D2 and vice versa (Fig. 2).

Discussion

The expression of 41,335 poplar genes included in the *P. trichocarpa* Phytozome database was studied. Expressed genes were characterized by gene ontology and by determining the metabolic pathway to which

Table 1
Number of Illumina reads matching to the *P. trichocarpa* unigene database (41,335 CDS sequences) for each library (C, control; D1, moderate drought; D2, severe drought). For each sample the leaf RWC is reported.

Library	RWC	Total reads	Read length	HQ reads	LQ reads	Aligned HQ reads (%)
Hybrid 85, clone 3 (C)	95.51	15,327,554	51	14,790,923	536,631	70.52
Hybrid 85, clone 4 (C)	92.58	4,316,064	51	3,612,582	703,482	72.80
Hybrid 89, clone 6 (C)	95.75	8,927,114	51	8,663,146	263,968	70.37
Hybrid 89, clone 8 (C)	95.40	3,963,712	51	3,816,812	146,900	65.89
Hybrid 85, clone 12 (D1)	86.31	5,487,345	51	4,379,253	1,108,092	72.17
Hybrid 85, clone 24 (D1)	85.64	5,003,314	51	4,798,629	204,685	66.18
Hybrid 89, clone 10 (D1)	84.89	6,123,484	51	5,963,450	160,034	67.00
Hybrid 89, clone 15 (D1)	86.30	7,355,080	51	7,064,731	290,349	68.15
Hybrid 85, clone 42 (D2)	54.78	3,985,186	51	3,515,137	470,049	63.86
Hybrid 85, clone 45 (D2)	61.83	5,715,359	51	5,480,077	235,282	57.96
Hybrid 89, clone 20 (D2)	52.78	4,575,904	51	3,934,155	641,749	59.23
Hybrid 89, clone 35 (D2)	59.69	5,855,333	51	5,710,049	145,284	68.38
Total		76,635,449		71,728,944	4,906,505	70.52

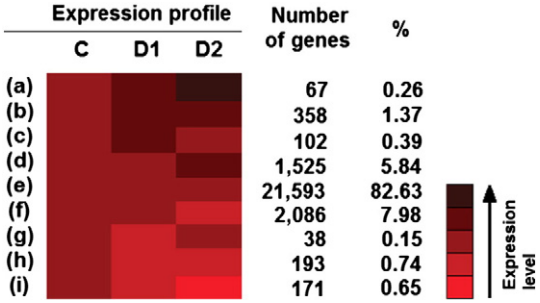

Fig. 2. Schematic representation of nine gene expression patterns (indicated by letters (a)–(i)) observed in leaves of poplar hybrids comparing control conditions (C) to moderate (D1) and to severe drought (D2). For each pattern, the number of genes is reported. Only expressed genes (i.e., with at least one read per million reads aligned in at least one sample) were counted.

they belong. Most genes detected were expressed in control and drought-treated plants, however a number of genes were observed that were significantly induced or repressed by drought [8].

Induction or repression of most genes was more common after severe (relative water content around 55–60%) than after moderate water deficit (around 85%) even for genes that usually respond promptly to changes in environmental conditions, such as those encoding transcription factors. Gene activation or repression might determine the phenotypic differences among individuals and/or species in response to drought. The dataset of expression profiles can help improve our understanding of molecular mechanism regulating drought response and facilitate comparative studies among poplar species.

Conflict of interest

Authors declare no conflict of interest.

Acknowledgments

This research work was supported by PRIN-MIUR, Italy, project "Verso la delucidazione delle basi molecolari dell'eterosi nelle piante coltivate: variazione cis-regolatoria ed espressione genica in ibridi di pioppo". Thanks are due to Dr. Catherine Bastien (INRA-UAGPF, Orleans, France) for ensuring availability of the hybrid plants. EB acknowledges the award of a research fellowship from the Department of Agriculture, Food, and Environment, University of Pisa, Italy.

References

[1] J. Logemann, J. Schell, L. Willmitzer, Improved method for the isolation of RNA from plant tissues. Anal. Biochem. 163 (1987) 16–20.
[2] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30 (2014) 2114–2120.
[3] G.A. Tuskan, S. Difazio, S. Jansson, J. Bohlmann, I. Grigoriev, U. Hellsten, N. Putnam, S. Ralph, S. Rombauts, A. Salamov, et al., The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313 (2006) 1596–1604.
[4] L. Sterck, S. Rombauts, S. Jansson, F. Sterky, P. Rouzé, Y. Van de Peer, EST data suggest that poplar is an ancient polyploid. New Phytol. 167 (2005) 165–170.
[5] A. Mortazavi, B.A. Williams, K. McCue, L. Schaeffer, B. Wold, Mapping and quantifying mammalian transcriptomes by RNA-seq. Nat. Methods 5 (2008) 621–628.
[6] K.A. Baggerly, L. Deng, J.S. Morris, C.M. Aldaz, Differential expression in SAGE: accounting for normal between-library variation. Bioinformatics 19 (2003) 1477–1483.
[7] C. Chen, A.D. Farmer, R.J. Langley, J. Mudge, J.A. Crow, G.D. May, J. Huntley, A.G. Smith, E.F. Retzel, Meiosis-specific gene discovery in plants: RNA-seq applied to isolated *Arabidopsis* male meiocytes. BMC Plant Biol. 10 (2010) 280.
[8] R.M. Cossu, T. Giordani, A. Cavallini, L. Natali, High-throughput analysis of transcriptome variation during water deficit in a poplar hybrid: a general overview. Tree Genet. Genomes 10 (2014) 53–66.