Genomics Data 5 (2015) 151–153

Contents lists available at ScienceDirect

Genomics Data

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

Data in Brief

Identification of genome-specific transcripts in wheat–rye translocation lines

Tong Geon Lee [a],*, Yong Weon Seo [b]

[a] Department of Crop Sciences, University of Illinois, Urbana, IL 61801, USA
[b] Division of Biotechnology, Korea University, Seoul 136-701, Republic of Korea

ARTICLE INFO

Article history:
Received 6 May 2015
Accepted 10 May 2015
Available online 11 June 2015

Keywords:
Wheat–rye translocation
Polyploidy
Homeologous
Synteny
Microarray

ABSTRACT

Studying gene expression in wheat–rye translocation lines is complicated due to the presence of homeologs in hexaploid wheat and high levels of synteny between wheat and rye genomes (Naranjo and Fernandez-Rueda, 1991 [1]; Devos et al., 1995 [2]; Lee et al., 2010 [3]; Lee et al., 2013 [4]). To overcome limitations of current gene expression studies on wheat–rye translocation lines and identify genome-specific transcripts, we developed a custom Roche NimbleGen Gene Expression microarray that contains probes derived from the sequence of hexaploid wheat, diploid rye and diploid progenitors of hexaploid wheat genome (Lee et al., 2014). Using the array developed, we identified genome-specific transcripts in a wheat–rye translocation line (Lee et al., 2014). Expression data are deposited in the NCBI Gene Expression Omnibus (GEO) under accession number GSE58678. Here we report the details of the methods used in the array workflow and data analysis.

Specifications	
Organism/cell line/tissue	*Aegilops speltoides*, *Aegilops squarrosa*, *Aegilops tauschii*, *Secale cereale*, *Triticum aestivum*, *Triticum monococcum*, *Triticum urartu*
Sex	Not applicable
Sequencer or array type	Roche NimbleGen Gene Expression microarray 135K
Data format	normalized data: SOFT, MINiML, and TXT
Experimental factors	Unstressed and drought-stressed plants
Experimental features	A technical possibility for designing genome-specific probes to detect gene expression in wheat–rye translocation lines
Consent	Not applicable
Sample source location	Not applicable

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc = GSE58678.

2. Experimental design, materials and methods

2.1. Plant materials and a custom array design

We use the term wheat–rye translocation line(s) to designate hexaploid wheat (ABD genome; *T. aestivum*) that possesses part of the rye genome (R; *S. cereale*) in the form of chromosome translocations [6,7,8]. Near-isolines (NILs) were developed by backcross introgression to form $BC_3F_{3:4}$ ('Coker 797' *4/'Hamlet') [9] and differed in the presence or absence of the long arm of rye chromosome 2 (2RL) derived from the diploid rye 'Chaupon' [9,10]. We used a NIL carrying 2RL (hereafter, 2BS.2RL) as a material of wheat–rye translocation lines [5]. Details of the sequence preparation for probe design were described in Lee et al. [5]. Sequence data sets used for probe design were as follows: A genome sequence, *T. monococcum* (A genome progenitor of hexaploid wheat, which belongs to the A genome lineage); B, *Ae. speltoides* (B genome progenitor, which belongs to the B genome lineage); D, *Ae. squarrosa* & *Ae. tauschii* (close relative of subgenome D of hexaploid wheat and D genome progenitor, which belong to the D genome lineage); ABD, *T. aestivum*; R, *S. cereale*.

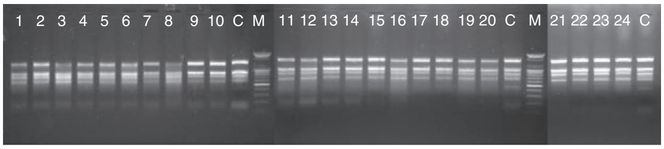

Fig. 1. Agarose gel electrophoresis of the RNA isolated from plant samples. Lanes 1, 2, 3 & 4 for *T. urartu*; 5, 6, 7 & 8 for *Ae. speltoides*; 9, 10, 11 & 12 for *Ae. squarrosa*; 13, 14, 15 & 16 for 'Chinese Spring'; 17, 18, 19 & 20 for 'Chaupon'; 21, 22, 23 & 24 for 2BS.2RL. C, control RNA (3 μg). M, 100 bp size marker.

---

* Corresponding author.
E-mail address: agrospel@illinois.edu (T.G. Lee).

http://dx.doi.org/10.1016/j.gdata.2015.05.045
2213-5960/

Table 1
Experimental metrics report.

Image name[a]	Signal range	Uniformity mean	Uniformity CV	Mean empty	Mean experimental	Mean random
A-rep1	0.324	3420.742	0.056	748.293	3112.410	299.421
A-rep2	0.391	3786.480	0.073	806.584	3431.427	384.572
B-rep1	0.253	3669.083	0.042	699.951	3323.705	388.067
B-rep2	0.317	3906.683	0.065	664.536	3507.074	412.286
D-rep1	0.322	3432.887	0.085	744.644	3157.820	307.147
D-rep2	0.397	3612.947	0.079	651.326	3252.032	358.274
ABD-rep1	0.340	3347.635	0.046	646.016	3057.030	330.758
ABD-rep2	0.265	3868.671	0.059	703.033	3510.957	364.590
2BS.2RL-rep1	0.174	4060.942	0.025	714.850	3647.586	393.915
2BS.2RL-rep2	0.638	3743.465	0.096	744.796	3389.392	392.318
R-rep1	0.202	3423.374	0.032	710.719	3129.917	323.823
R-rep2	0.333	3579.254	0.063	800.672	3287.313	500.004

[a] The name of the analyzed image file. Image name is labelled according to the cDNA probe and replication (replicates 1 or 2) of the array.

2.2. cDNA preparation

cDNAs from diploid progenitors (A, B or D genomes) of hexaploid wheat and diploid rye (R) were used to empirically identify probes that distinguish transcripts derived from distinct genomes in a 2BS.2RL wheat–rye translocation line [5]; *T. urartu* (A genome progenitor of hexaploid wheat, which belongs to the A genome lineage) for A genome cDNA, *Ae. speltoides* for B, *Ae. squarrosa* for D, hexaploid wheat cultivar ‘Chinese Spring’ for ABD, ‘Chaupon’ for R, 2BS.2RL for ABD and genome of 2RL rather than the long arm of wheat chromosome 2B. For the synthesis of double-stranded cDNA, the RevertAid H Minus First Strand cDNA Synthesis Kit (Life Technologies, USA) was used. A mixture of 1 μl of oligo dT primer (100 μM) and 10 μl of total RNA (10 μg; Fig. 1) was denatured at 70 °C for 5 min, then placed on ice. 4 μl of 5× First Strand Buffer, 1 μl RiboLock RNase Inhibitor, 2 μl 10 mM dNTP mix and 1 μl RevertAid H Minus M-MuLV Reverse Transcriptase were added to the mixture to synthesize first strand DNA. The mixture was incubated at 42 °C for 1 h followed by 70 °C to terminate the reaction. 66.7 μl of nuclease free water, 5 μl of 10× reaction buffer for DNA Polymerase I (Life Technologies, USA), 5 μl of 10× T4 DNA ligase buffer (Takara, Japan), 3 μl of 10 U/μl DNA Polymerase I (Life Technologies, USA), 0.2 μl of 5 U/μl RNase H (Life Technologies, USA) and 0.1 μl of 350 U/μl T4 DNA ligase (Takara, Japan) were added to the first strand cDNA mixture for the second strand synthesis, then the reaction was incubated at 15 °C for 2 h. After incubation, the double-stranded cDNA mixture was purified using the MinElute Reaction Cleanup Kit (Qiagen, USA). For the synthesis of Cy3-labeled DNA, 1 μg of double-stranded cDNA was mixed with 30 μl (1 O.D. value) of Cy3-9 mer primers (Sigma-Aldrich, USA), then denatured at 98 °C for 10 min. The reaction was further proceeded by adding 10 μl of 50×

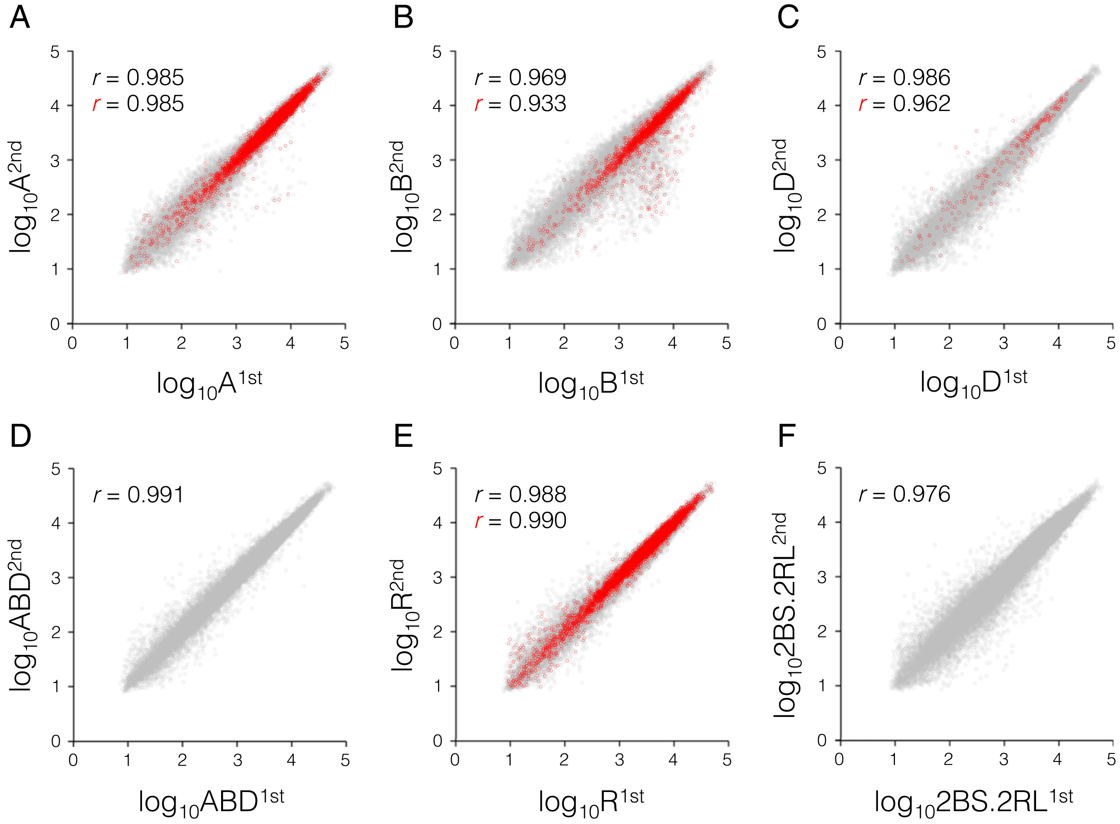

Fig. 2. Scatter plots showing correlation of signal values between two biological replicates. (A) *T. urartu*, (B) *Ae. speltoides*, (C) *Ae. squarrosa*, (D) ‘Chinese Spring’, (E) ‘Chaupon’ & (F) 2BS.2RL. Gray dots represent the entire probes in arrays. Probes derived from diploid genome sequences were shown in red ((A) for A genome-derived probes, (B) for B, (C) for D & (E) for R). Pearson's correlation coefficients were calculated using the log-transformed values for probes in both biological replicates. *r* values are indicated in the plots. *x*- and *y*-axes represent the 1st and 2nd biological replicates, respectively.

dNTP mix (10 mM each), 8 μl of deionized water and 2 μl of Klenow fragment (50 U/μl; Takara, Japan) to the mixture. After incubation at 37 °C for 2 h, 11.5 μl of 5 M NaCl and 110 μl of isopropanol were added to the mixture. DNA was collected by centrifugation at 12,000 g. The products of Cy3-labeled DNA were rehydrated. The concentration of sample was measured using spectrophotometer.

2.3. Hybridization

10 μg of DNA was used for array hybridization. The sample was mixed with 19.5 μl of 2× hybridization buffer (NimbleGen, USA) and finalized to 39 μl with deionized water. Hybridization was performed using the MAUI chamber (Biomicro, USA) at 42 °C for 16–18 h. After hybridization, the array was immediately immersed in 250 ml Wash I (NimbleGen, USA) at 42 °C for 10–15 s. After incubation, the array was transferred to Wash II followed by Wash III. The array was centrifuged at 500 *g* for 1 min.

2.4. Data analysis

The array was scanned using the Genepix 4000 B (Axon, USA) preset with a 5 μm resolution for Cy3 signal. Signals were analyzed by NimbleScan v2.5 (NimbleGen, USA). The grid was aligned to the image with a chip design file (.ndf). Expression analysis was performed: (1) pair reports files (.pair) were generated in which sequence, probe and signal intensity information for Cy3 channel were collected; (2) background subtraction using a local background estimator was performed to improve fold change estimates on arrays with high background signal; (3) the data was normalized and processed with cubic spline normalization using quantiles to adjust signal variations between chips. Experimental metrics report obtained from NimbleScan was listed in Table 1. Probe-level summarization by Robust Multi-array Analysis (RMA) using a median polish algorithm implemented in NimbleScan was used to produce the call files (.calls). Multiple analyses were performed with LIMMA package in R software environment. A threshold of 0.05 (false discovery rate; FDR) was applied. RMA normalized data for each experiment were $\log_{10}$ transformed (Fig. 2) followed by standardization using Z score transformation [11]. The Z ratio was used for calculating differences in hybridization values of probes across different species. A Z ratio of ± 1.96 was deduced as significant ($P < 0.05$); e.g. we identified rye genome sequence-derived probes that hybridized better (Z ratio > 1.96) to cDNA from rye rather than to cDNAs from all other species as rye-specific transcripts [5].

3. Discussion

It's been widely accepted that the hexaploid wheat subgenomes A, B and D were derived from the three diploid species. There is evidence that the sequences of genes were highly conserved between the hexaploid subgenomes and their respective diploid relatives [12,13]. The sequence identity of genes between species in the same diploid lineage is higher than that of between two different diploid lineages for homeologous chromosomes. For both cases of probe sequence design and cDNA preparation, we have used the diploid species that belong to the subgenome lineage or close relatives of each subgenome of the hexaploid wheat [5]. Therefore, subgenome-specific expression profiles in hexaploid wheat are most likely to be detected by their respective genome-specific probes designed in this study. Using the cDNA from a wheat–rye translocation line, we further identified transcripts that showed preferential hybridization to rye chromatin [5]. Hence, our data address an original approach for probing genome-specific transcripts in wheat–rye translocation lines.

References

[1] T. Naranjo, P. Fernandez-Rueda, Homoeology of rye chromosome arms to wheat. Theor. Appl. Genet. 82 (1991) 577–586.
[2] K.M. Devos, G. Moore, M.D. Gale, Conservation of marker synteny during evolution. Euphytica 85 (1995) 367–372.
[3] T.G. Lee, Y.J. Lee, D.Y. Kim, Y.W. Seo, Comparative physical mapping between wheat chromosome arm 2BL and rice chromosome 4. Genetica 138 (2010) 1277–1296.
[4] T.G. Lee, D.Y. Kim, J.W. Johnson, Y.W. Seo, A genome-wide analysis of transcripts in a 2BS.2RL wheat–rye translocation during Hessian fly infestation. Genes Genom. 35 (2013) 795–803.
[5] T.G. Lee, Y.J. Lee, Y.W. Seo, Expression analysis of individual homoeologous wheat genome-and rye genome-specific transcripts in a 2BS.2RL wheat–rye translocation. Genes Genet. Syst. 89 (2014) 159–168.
[6] B. Friebe, J. Jiang, W.J. Raupp, R.A. McIntosh, B.S. Gill, Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status. Euphytica 91 (1996) 59–87.
[7] R.A. Graybosch, Uneasy unions: quality effects of rye chromatin transfers to wheat. J. Cereal Sci. 33 (2001) 3–16.
[8] W.J. Jung, Y.W. Seo, Employment of wheat–rye translocation in wheat improvement and broadening its genetic basis. J. Crop. Sci. Biotechnol. 17 (2014) 305–313.
[9] Y.W. Seo, J.W. Johnson, R.L. Jarret, A molecular marker associated with the *H21* Hessian fly resistance gene in wheat. Mol. Breed. 3 (1997) 177–181.
[10] T.G. Lee, M.J. Hong, J.W. Johnson, D.E. Bland, D.Y. Kim, Y.W. Seo, Development and functional assessment of EST-derived 2RL-specific markers for 2BS.2RL translocations. Theor. Appl. Genet. 119 (2009) 663–673.
[11] C. Cheadle, M.P. Vawter, W.J. Freed, K.G. Becker, Analysis of microarray data using Z score transformation. J. Mol. Diagn. 5 (2003) 73–81.
[12] T. Marcussen, S.R. Sandve, L. Heier, M. Spannagl, M. Pfeifer, International Wheat Genome Sequencing Consortium, K.S. Jakobsen, B.B. Wulff, B. Steuernagel, K.F. Mayer, O.A. Olsen, Ancient hybridizations among the ancestral genomes of bread wheat. Science 345 (2014) 1250092.
[13] International Wheat Genome Sequencing Consortium (IWGSC), A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science 345 (2014) 1251788.