Genomics Data 14 (2017) 36–39

Contents lists available at ScienceDirect

Genomics Data

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

Data in Brief

RNA-seq data comparisons of wild soybean genotypes in response to soybean cyst nematode (*Heterodera glycines*)

Hengyou Zhang, Bao-Hua Song*

Department of Biological Sciences, University of North Carolina at Charlotte, Charlotte, NC 28223, USA

ARTICLE INFO

Keywords:
Glycine soja
RNA-seq
Soybean cyst nematode (SCN)
Expression
Transcriptome

ABSTRACT

Soybean [*Glycine max* (L.) Merr.] is an important crop rich in vegetable protein and oil, and is a staple food for human and animals worldwide. However, soybean plants have been challenged by soybean cyst nematode (SCN, *Heterodera glycines*), one of the most damaging pests found in soybean fields. Applying SCN-resistant cultivars is the most efficient and environmentally friendly strategy to manage SCN. Currently, soybean breeding and further improvement in soybean agriculture are hindered by severely limited genetic diversity in cultivated soybeans. *G. soja* is a soybean wild progenitor with much higher levels of genetic diversity compared to cultivated soybeans. In this study, transcriptomes of the resistant and susceptible genotypes of the wild soybean, *Glycine soja* Sieb & Zucc, were sequenced to examine the genetic basis of SCN resistance. Seedling roots were treated with infective second-stage juveniles (J2s) of the soybean cyst nematode (HG type 2.5.7) for 3, 5, 8 days and pooled for library construction and RNA sequencing. The transcriptome sequencing generated approximately 245 million (M) high quality (Q > 30) raw sequence reads (125 bp in length) for twelve libraries. The raw sequence reads were deposited in NCBI sequence read archive (SRA) database, with the accession numbers SRR5227314-25. Further analysis of this data would be helpful to improve our understanding of the molecular mechanisms of soybean-SCN interaction and facilitate the development of diverse SCN resistance cultivars.

Specifications	
Organism/cell line/tissue	Wild soybean (*Glycine soja* Sieb. & Zucc.)/whole root tissue
Sex	Not applicable
Sequencer or array type	Illumina HiSeq 2500
Data format	Raw reads in fastq.gz format
Experimental factors	Seedling roots were treated with infective second-stage juveniles (J2s) of the soybean cyst nematode (HG type 2.5.7) for 3, 5, 8 days. Roots from four individual plants were pooled as one biological replicate, and three replicates were prepared for treatment and control at each time point.
Experimental features	This study uses wild soybean that contains higher levels of genetic diversity compared with cultivated soybean as a study system. Two wild soybean genotypes with distinct response to SCN were used for comparative analyses.
Consent	Not applicable
Sample source location	Samples were originally collected from East Asia. Seeds could be requested at USDA Soybean Germplasm Collection, United States (http://www.ars-grin.gov/).

1. Direct link to deposited data

https://www.ncbi.nlm.nih.gov/Traces/study/?acc=SRP098790

2. Introduction

Soybean [*Glycine max* (L.) Merr.] is one of the important food crops grown worldwide providing a variety of soy foods for human and livestock, and potential feedstock for biofuel. Production of soybean in most soybean-producing areas is challenged by various environmental stress, of which soybean cyst nematode (SCN, *Heterodera glycines* Ichinohe) is the leading cause of soybean yield loss [1]. SCN is a sedentary endoparasitic pathogen that infects the soybean roots, causing invisible above-ground symptoms even with yield loss of 15–30% [1]. The value of soybean lost to SCN was estimated at 1.5 billion dollars in the United States [2]. Breeding SCN-resistant soybean varieties is the

* Corresponding author at: Department of Biological Sciences, University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, NC 28223, USA.
E-mail address: bsong5@uncc.edu (B.-H. Song).

http://dx.doi.org/10.1016/j.gdata.2017.08.001
Received 7 June 2017; Received in revised form 2 August 2017; Accepted 9 August 2017
Available online 09 August 2017

most effective and environmentally friendly strategy to reduce the SCN damage compared to crop rotation or nematicide application. Most commercial soybean varieties have originated from limited resistance sources, such as Peking and PI88788 [3]. It has also been demonstrated that SCN race has been evolving, due to long-term use of cultivars originally derived from a single resistance source [4,5]. Thus, new sustainable sources of SCN resistance are urgently needed for effective management of SCN infection. One such alternative is *Glycine soja* Sieb. & Zucc., a wild progenitor of cultivated soybean, which harbors higher levels of genetic diversity and has a demonstrated role in SCN resistance [6–9].

Thus far, the knowledge of the molecular basis of soybean resistance to SCN is mostly from the identification of two major QTLs *rhg1* [10] and *Rhg4* [11], and the transcriptome analysis using microarray [12–17] and RNA-seq technology [18]. Most studies have been focusing on soybean cyst nematode HG type 0, which is the most prevalent HG type in the central US. However, few studies have worked on the HG 2.5.7. (known as race 5), prevalent in Southeast US., and little is known about the molecular mechanisms of plant resistance to this HG type.

With the cost reduction of next-generation sequencing (NGS), the RNAseq-based transcriptome analysis has been one of the most efficient strategies in dissecting the genetic basis of complex traits variation, particularly the interaction between plants and pests [18–20]. Known regulatory genes and a novel set of defense-related genes have been identified as potentially playing a role in SCN resistance [18]. In addition, the RNA-seq data can provide a global view of networks involved in pest resistance [21,22].

In this study, we sequenced the transcriptomes of HG 2.5.7-resistant (S54) and susceptible (S67) wild soybean genotypes, with and without HG 2.5.7 infection, to develop transcriptomic resources for better understanding the genes and pathways involved in SCN resistance in wild soybean.

3. Experimental design, materials and methods

3.1. Plant materials and nematode preparation

G. soja seeds used in this study were requested from USDA Soybean Germplasm Collection (http://www.ars-grin.gov/). S54 and S67 were identified resistant and susceptible to HG2.5.7. respectively in our early study [23]. For seed germination, seed coat was individually sliced to promote germination, surface sterilized with sodium hypochlorite (0.5%) for no > 4 min, followed by a rinse with sterile water, and germinated in Petri dishes with a wet filter paper for two days (kept in dark). The seedlings with approximately 2 cm roots were transplanted into cones (Greenhouse Megastore, IL, USA) filled with sterile sand, with one plant per cone. A randomized complete block design was used. Seeds germination and plant growth were conducted in the environmental chamber (Percival, IA, USA) with growth conditions at 27 °C, 50% relative humidity, and long day photoperiod (16 h L/8 h D) as previously described [24]. All plants were properly watered daily to keep the sand moist.

Soybean cyst nematode HG type 2.5.7 was cultured on SCN susceptible soybean cv. Hutcheson for at least 30 generations in the climate-controlled greenhouse at the University of North Carolina at Charlotte, USA. The growth conditions for SCN stock plants were the same as plant treatment described above.

3.2. Plant treatment and sample collection

The nematode eggs were extracted as previously described [14]. The stock plants with 3-month old SCN culture were used for eggs extraction. Briefly, we dumped soil and roots into a bucket, rinsed out soil from roots, decanted water over nested 850 and 250 µm sieves (Fisher Scientific Inc., US). The female cysts could be collected on the 250 µm sieve. After crashing the cyst coat with a rubber stopper on the 250 µm sieve, the released eggs flowed through the sieve onto a 25 µm mesh sieve. The eggs were purified by sucrose flotation [25]. The clean eggs were placed on folded paper tissue in a plastic tray with a minimal amount of water to keep the tissue moist. The plastic tray was covered with aluminum foil and placed in an incubator at 27 °C for hatching. After 2–3 days, the hatched second stage juveniles (J2s) of SCN were collected and suspended in 0.09% sterile agarose solution (w/v). A final amount of 1800 J2s per plant was added to the roots for treatment, and the same amount of 0.09% agarose solution without J2 s was added on each control plant.

To capture the transcriptomic changes in response to SCN infection at different time points, the whole root was sampled at 3, 5, and 8 days post inoculation (dpi), flash frozen in liquid nitrogen, and stored in a − 80 °C freezer. Roots from four individual plants of each genotype (S54, S67) were powdered and equal amounts of tissue (by weight) were pooled as one biological replicate for RNA preparation; three replicates were prepared for treatment and control at each time point resulting in 12 RNA libraries. A flowchart for sample preparation is shown in Fig. 1. Three extra samples collected at 3 dpi were stained using acid fuchsin [26] to ensure the successful infection.

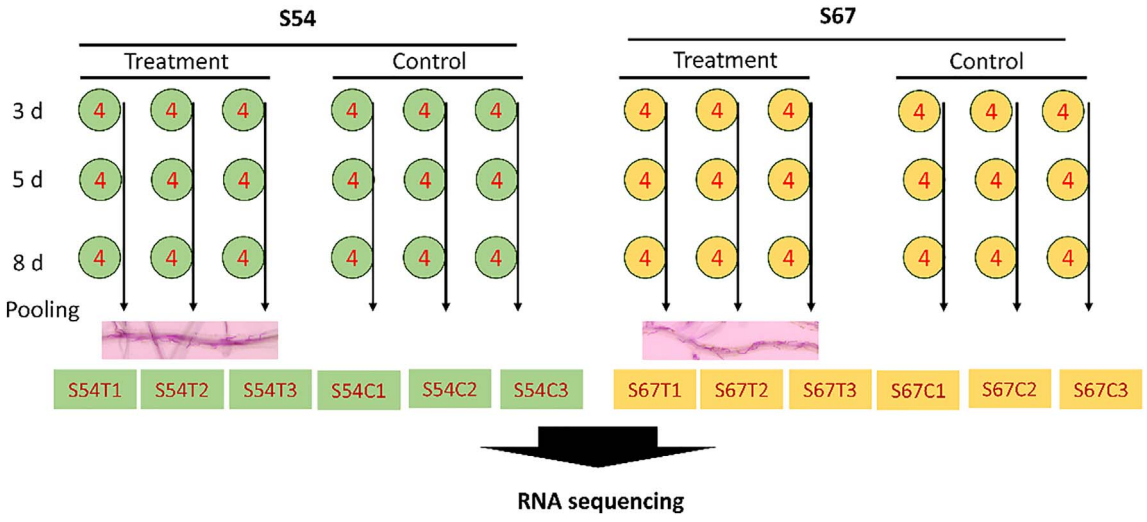

Fig. 1. A flowchart for sample preparation.
Note: Embedded figures indicate the penetrated J2s in S54 and S67 treatment roots.

Table 1
Summary of RNA-Seq for *G. soja* S54 and S67.

Feature	S54_Treated	S54_Control	S67_Treated	S67_Control
Sequencing platform	Hiseq2500	Hiseq2500	Hiseq2500	Hiseq2500
Length of raw reads	125 bp	125 bp	125 bp	125 bp
No. of average reads	18,094,552	22,372,926	20,682,052	20,405,576
No. of clean reads	17,997,709	22,265,698	20,566,719	20,311,775
% of clean reads	99.46%	99.52%	99.44%	99.54%
No. of reads mapped	15,651,020	19,488,275	18,114,667	18,757,405
% of average reads mapped	86.96%	87.53%	88.08%	92.35%
NCBI BioProject ID	PRJNA369554	PRJNA369554	PRJNA369554	PRJNA369554
NCBI BioStudy ID	SRP098790	SRP098790	SRP098790	SRP098790
NCBI BioSample ID	SAMN06290886	SAMN06290886	SAMN06290887	SAMN06290887
NCBI SRA accession number	SRR5227320-22	SRR5227323-25	SRR5227314-16	SRR5227317-19
Total reads	244,665,319			

3.3. Library preparation and transcriptome sequencing

RNA was prepared using the RNeasy mini total RNA isolation kit (Qiagen, US) following the manufacturer's protocol. The concentration and quality of the RNA extracted were evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies, US). Library construction was performed using NEBNext Ultra Directional RNA Library Prep Kit (NEB) and NEBNext Multiplex Oligos for Illumina (NEB) using the manufacturer-specified protocol. The amplified library fragments were purified and checked for quality and final concentration using an Agilent 2200 Tapestation (Agilent Technologies, USA). The qualified libraries were sequenced using the Illumina Hiseq2500 instrument, utilizing a 125 bp read length with v4 sequencing chemistry (Illumina, USA) at the North Carolina State Genomic Sciences Laboratory (Raleigh, NC, USA).

3.4. Data processing

The raw reads were evaluated using FastQC [27], and preprocessed using Trimmomatic (version 0.36) [28] to remove the low-quality reads (quality score < 20) and adapters (options: PE -phred33 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 HEADCROP:8 MINLEN:30). Clean reads were aligned against the soybean reference genome of *G. max* cv. Williams 82 [29] using TopHat2 [30] (Cmd: tophat2 -p 6-i 30-I 15000-G/path/to/annotation.gff3 -o/path/to/tophat.out/path/to/indexed_reference_genome/path/to/clean reads).

3.5. Data description

Nearly 245 million (M) RNA-seq reads were generated, totaling 19 Gb of sequence, from all 12 libraries of the SCN-treated and control roots. The average number of reads were 18.1 M, 22.3 M, 20.7 M and 20.4 M for S54 treated roots, S54 control roots, S67 treated roots, and S67 control roots, respectively. Over 99.0% of the raw reads from each library were retained after quality control. In addition, over 85% of reads in each library were mapped on *G. max* reference genome. The transcriptome sequencing data were summarized in Table 1. In addition to providing data for comprehensive comparative analysis between S54 and S67, our dataset provides the opportunity to investigate alternative splicing and sequence variations of candidate genes.

Conflict of interest

The authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

We thank the funding support by The National Institute of General Medical Sciences of the National Institutes of Health (R15GM122029), North Carolina Biotechnology Center (2014-CFG-8005).

References

[1] T.L. Niblack, Soybean cyst nematode management reconsidered, Plant Dis. 89 (2005) 1020–1026.
[2] A. Wrather, S. Koenning, Effects of diseases on soybean yields in the United States 1996 to 2007. Online, plant health prog. https://www.plantmanagementnetwork.org/pub/php/research/2009/yields/, (2009).
[3] V.C. Concibido, B.W. Diers, P.R. Arelli, A decade of QTL mapping for cyst nematode resistance in soybean, Crop Sci. 44 (2004) 1121–1131.
[4] M.G. Mitchum, J.A. Wrather, R.D. Heinz, J.G. Shannon, G. Danekas, Variability in distribution and virulence phenotypes of *Heterodera glycines* in Missouri during 2005, Plant Dis. 91 (2007) 1473–1476.
[5] T.L. Niblack, A.L. Colgrove, K. Colgrove, J.P. Bond, Shift in virulence of soybean cyst nematode is associated with use of resistance from PI 88788, online, plant health prog. https://www.plantmanagementnetwork.org/pub/php/research/2008/virulence/, (2008).
[6] D.L. Hyten, Q.J. Song, Y.L. Zhu, I.Y. Choi, R.L. Nelson, J.M. Costa, J.E. Specht, R.C. Shoemaker, P.B. Cregan, Impacts of genetic bottlenecks on soybean genome diversity, P. Natl. Acad. Sci. USA 103 (2006) 16666–16671.
[7] H.Y. Zhang, N. Mittal, L.J. Leamy, O. Barazani, B.H. Song, Back into the wild-apply untapped genetic diversity of wild relatives for crop improvement, Evol. Appl. 10 (2017) 5–24.
[8] M. Kim, B.W. Diers, Fine mapping of the SCN resistance QTL *cqSCN-006* and *cqSCN-007* from *Glycine soja* PI 468916, Crop Sci. 53 (2013) 775–785.
[9] M. Kim, D.L. Hyten, T.L. Niblack, B.W. Diers, Stacking resistance alleles from wild and domestic soybean sources improves soybean cyst nematode resistance, Crop Sci. 51 (2011) (2301–2301).
[10] D.E. Cook, T.G. Lee, X.L. Guo, S. Melito, K. Wang, A.M. Bayless, J.P. Wang, T.J. Hughes, D.K. Willis, T.E. Clemente, B.W. Diers, J.M. Jiang, M.E. Hudson, A.F. Bent, Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean, Science 338 (2012) 1206–1209.
[11] S.M. Liu, P.K. Kandoth, S.D. Warren, G. Yeckel, R. Heinz, J. Alden, C.L. Yang, A. Jamai, T. El-Mellouki, P.S. Juvale, J. Hill, T.J. Baum, S. Cianzio, S.A. Whitham, D. Korkin, M.G. Mitchum, K. Meksem, A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens, Nature 492 (2012) 256–260.
[12] N. Ithal, J. Recknor, D. Nettleton, L. Hearne, T. Maier, T.J. Baum, M.G. Mitchum, Parallel genome-wide expression profiling of host and pathogen during soybean cyst nematode infection of soybean, Mol. Plant-Microbe Interact. 20 (2007) 293–305.
[13] P.K. Kandoth, N. Ithal, J. Recknor, T. Maier, D. Nettleton, T.J. Baum, M.G. Mitchum, The soybean *Rhg1* Locus for resistance to the soybean cyst nematode *Heterodera glycines* regulates the expression of a large number of stress- and defense-related genes in degenerating feeding cells, Plant Physiol. 155 (2011) 1960–1975.
[14] V.P. Klink, C.C. Overall, N.W. Alkharouf, M.H. MacDonald, B.F. Matthews, A time-course comparative microarray analysis of an incompatible and compatible response by *Glycine max* (soybean) to *Heterodera glycines* (soybean cyst nematode) infection, Planta 226 (2007) 1423–1447.
[15] V.P. Klink, C.C. Overall, N.W. Alkharouf, M.H. MacDonald, B.F. Matthews, Laser capture microdissection (LCM) and comparative microarray expression analysis of syncytial cells isolated from incompatible and compatible soybean (*Glycine max*) roots infected by the soybean cyst nematode (*Heterodera glycines*), Planta 226 (2007) 1389–1409.
[16] M. Mazarei, W. Liu, H. Al-Ahmad, P.R. Arelli, V.R. Pantalone, C.N. Stewart, Gene expression profiling of resistant and susceptible soybean lines infected with soybean cyst nematode, Theor. Appl. Genet. 123 (2011) 1193–1206.
[17] J.R. Wan, T. Vuong, Y.Q. Jiao, T. Joshi, H.X. Zhang, D. Xu, H.T. Nguyen, Whole-genome gene expression profiling revealed genes and pathways potentially involved in regulating interactions of soybean with cyst nematode (*Heterodera glycines* Ichinohe), BMC Genomics 16 (2015) 148.
[18] P. Hosseini, B.F. Matthews, Regulatory interplay between soybean root and soybean cyst nematode during a resistant and susceptible reaction, BMC Plant Biol. 14

(2014) 300.
[19] X.Y. Li, X. Wang, S.P. Zhang, D.W. Liu, Y.X. Duan, W. Dong, Comparative profiling of the transcriptional response to soybean cyst nematode infection of soybean roots by deep sequencing, Chin. Sci. Bull. 56 (2011) 1904–1911.
[20] X.Y. Li, X. Wang, S.P. Zhang, D.W. Liu, Y.X. Duan, W. Dong, Identification of soybean microRNAs involved in soybean cyst nematode infection by deep sequencing, PLoS One 7 (2012) e39650.
[21] J.J. Chen, W.X. Pang, B. Chen, C.Y. Zhang, Z.Y. Piao, Transcriptome analysis of *Brassica rapa* near-isogenic lines carrying clubroot-resistant and -susceptible alleles in response to *Plasmodiophora brassicae* during early infection, Front. Plant Sci. 6 (2016) 1183.
[22] F. Zhang, G.Z. Zhu, L. Du, X.G. Shang, C.Z. Cheng, B. Yang, Y. Hu, C.P. Cai, W.Z. Guo, Genetic regulation of salt stress tolerance revealed by RNA-Seq in cotton diploid wild species, *Gossypium davidsonii*, Sci Rep 6 (2016) 20582.
[23] H.Y. Zhang, C.Y. Li, E.L. Davis, J.S. Wang, J.D. Griffin, J. Kofsky, B.H. Song, Genome-wide association study of resistance to soybean cyst nematode (*Heterodera glycines*) HG Type 2.5.7 in wild soybean (*Glycine soja*), Front. Plant Sci. 7 (2016).
[24] T.L. Niblack, G.L. Tylka, P. Arelli, J. Bond, B. Diers, P. Donald, J. Faghihi, V.R. Ferris, K. Gallo, R.D. Heinz, H. Lopez-Nicora, R. Von Qualen, T. Welacky, J. Wilcox, A standard greenhouse method of assessing soybean cyst nematode resistance in soybean: SCE08 (standardized cyst evaluation 2008), J. Nematol. 41 (2009) (364–364).
[25] B.F. Matthews, M.H. MacDonald, V.K. Thai, M.L. Tucker, Molecular characterization of argenine kinase in the soybean cyst nematode (*Heterodera glycines*), J. Nematol. 35 (2003) 252–258.
[26] D.W. Bybd, T. Kirkpatrick, K.R. Barker, An improved technique for clearing and staining plant tissues for detection of nematodes, J. Nematol. 15 (1983) 142–143.
[27] S. Andrew, A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/, (2010).
[28] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (2014) 2114–2120.
[29] J. Schmutz, S.B. Cannon, J. Schlueter, J.X. Ma, T. Mitros, W. Nelson, D.L. Hyten, Q.J. Song, J.J. Thelen, J.L. Cheng, D. Xu, U. Hellsten, G.D. May, Y. Yu, T. Sakurai, T. Umezawa, M.K. Bhattacharyya, D. Sandhu, B. Valliyodan, E. Lindquist, M. Peto, D. Grant, S.Q. Shu, D. Goodstein, K. Barry, M. Futrell-Griggs, B. Abernathy, J.C. Du, Z.X. Tian, L.C. Zhu, N. Gill, T. Joshi, M. Libault, A. Sethuraman, X.C. Zhang, K. Shinozaki, H.T. Nguyen, R.A. Wing, P. Cregan, J. Specht, J. Grimwood, D. Rokhsar, G. Stacey, R.C. Shoemaker, S.A. Jackson, Genome sequence of the palaeopolyploid soybean, Nature 463 (2010) 178–183.
[30] C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D.R. Kelley, H. Pimentel, S.L. Salzberg, J.L. Rinn, L. Pachter, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks, Nat. Protoc. 7 (2012) 562–578.