

A genomic resource for the sedentary semi-endoparasitic reniform nematode, *Rotylenchulus reniformis* Linford & Oliveira

Kurt C. Showmaker^{1,2,3,†}, William S. Sanders^{2,3,4,†}, Sebastian Eves-van den Akker⁵, Brigitte E. Martin⁶, Roy N. Platt II⁷, John V. Stokes⁸, Chuan-Yu Hsu³, Benjamin D. Bartlett⁹, Daniel G. Peterson^{3,9} and Martin J. Wubben^{2,10,*}

¹Department of Crop Science, University of Illinois at Urbana-Champaign, Urbana, IL.

²Department of Biochemistry, Molecular Biology, Entomology, and Plant Pathology, Mississippi State University, Starkville, MS 39759.

³Institute for Genomics, Biocomputing, and Biotechnology, Mississippi State University, Starkville, MS 39759.

⁴The Jackson Laboratory, Farmington, CT.

⁵Department of Plant Sciences, University of Cambridge, Cambridge, UK.

⁶Department of Molecular and Cellular Biology, University of Illinois at Urbana-Champaign, Urbana, IL.

⁷Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX.

⁸Department of Basic Sciences, College of Veterinary Medicine, Mississippi State University, Starkville, MS 39759.

⁹Department of Plant and Soil Sciences, Mississippi State University, Starkville, MS 39759.

¹⁰USDA-ARS, Crop Science Research Laboratory, Mississippi State, Starkville, MS 39759.

*Email: Martin.Wubben@ars.usda.gov

This paper was edited by Erik J. Ragsdale

†These authors contributed equally.

Received for publication August 17, 2018

Abstract

The reniform nematode (*Rotylenchulus reniformis*) is a sedentary semi-endoparasitic species that is pathogenic on many row crops, fruits, and vegetables. Here, the authors present a draft genome assembly of *R. reniformis* using small- and large-insert libraries sequenced on the Illumina GAIIx and MiSeq platforms.

The reniform nematode (*Rotylenchulus reniformis* Linford & Oliveira) is a sedentary semi-endoparasitic species that infects 77 plant families, including cotton and soybean, and other high-value crops such as sweet potato and pineapple (Robinson et al., 1997). Here, we report the genome sequencing and assembly of *R. reniformis*.

Biological material for sequencing was a *R. reniformis* stock culture maintained on cotton. Eggs of *R. reniformis* were collected from infected cotton roots (Hussey and Barker, 1973), purified (Jenkins, 1964), surface-sterilized (Baum et al., 2000), pelleted by centrifugation, flash-frozen with liquid nitrogen, and stored at -80°C until DNA extraction. Genomic DNA extraction was performed as described by Blin and Stafford (1976).

Three short-read small-insert libraries (250, 350, and 550bp) and four large-insert Nextera mate-pair libraries (2-3.5-, 4-6-, 8-10-, and 2-10-kb) were constructed from a pooled population of *R. reniformis* eggs using standard protocols provided by Illumina (San Diego, CA). Additionally, a single *R. reniformis* egg was isolated, its DNA was amplified by whole genome amplification (WGA) using the Qiagen Rep-1-G single cell kit (Valencia, CA) and short-read small-insert libraries (400- and 600-bp) were constructed and sequenced using standard protocols provided by Illumina (San Diego, CA). Paired-end sequencing was conducted on Illumina GAIIx and Illumina MiSeq systems. We obtained 19.17 Gb and 29.85 Gb from the pooled population short- and long-read libraries, respectively, and 22.29 Gb from the WGA libraries. Pooled population and WGA sequence data was combined for genome assembly.

Mate-pair reads were processed with NextClip (v1.3; Richard et al., 2014) with default parameters to identify fragments containing linker sequence indicating proper circulation and removal of duplicate reads. All libraries were trimmed for adapters and low-quality base calls with Trimmomatic (v0.32; Bolger et al., 2014) and then assembled with ABySS (v1.5.2; Simpson et al., 2009) with a *k*-mer value of 175. The mate-pair libraries were used for scaffolding the assembly with

SSPACE (v3.0; Boetzer et al., 2011) followed by gap filling of all libraries with GapFiller (v1.10; Nadalin et al., 2012).

The resultant *R. reniformis* genome assembly (RREN1.0, GCA_001026735.1) contained 314 Mb distributed among 100,525 scaffolds with scaffold and contig N50s of 22,705 bp and 5,991 bp, respectively. The genome assembly contained 86.29% of the core eukaryotic proteins utilized by CEGMA (v2.5; Parra et al., 2007). Flow cytometry of propidium iodide stained *R. reniformis* nuclei indicated a genome size of ~190 Mb. The larger genome size (310 Mb) represented by the assembly is likely the result of unresolved haplotypes stemming from heterogeneity within the *R. reniformis* population used for DNA extraction. Repeat prediction with RepeatModeler (Smit and Hubley, 2015) and repeat quantification with RepeatMasker (Smit et al., 2015) identified 245 repeat families within 35.1 Mb of the assembled genome; most elements were DNA transposons or LTR retroelements (20.6 Mb and 11.3 Mb, respectively), while low complexity repeats accounted for 7 Mb. Homologs of known plant-parasitic nematode effector molecules were identified via tBLASTn and included beta-1-4 endogluconase, chitinase, CLE, CEP, chorismate mutase, invertase, ubiquitin extension protein, and venom allergen-like protein as well as cyst nematode pioneers 4D06, GLAND1, GLAND11, GLAND14, GLAND15, 10C02, 20E03, 33A09, 4G05, 7E05, and 10A06 (reviewed by Mitchum et al., 2013). A variant of the DOG (Dorsal Gland) box DNA motif, an enhancer element associated with dorsal gland effectors (Eves-van den Akker et al., 2016), was identified in the assembly.

GenBank accession numbers: The raw DNA sequence data and genome assembly were deposited at GenBank under BioProject No. PRJ-NA214681.

Acknowledgments

Mention of trade names or commercial products in this paper is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. The authors would like to thank Dr. Satish Ganji for his assistance with flow cytometry and whole genome amplification. The authors would also like to thank Drs. William Rutter and Osman Gutierrez (USDA-ARS) for their critical review of the manuscript. This project was funded, in part, by USDA ARS Non-Assistance Cooperative Agreements 58-6066-6-046 and 58-6066-6-059.

References

- Baum, T. J., Wubben, M. J. E., Hardy, K. A., Su, H. and Rodermeel, S. R. 2000. A screen for *Arabidopsis thaliana* mutants with altered susceptibility to *Heterodera schachtii*. *Journal of Nematology* 32:166–73.
- Blin, N. and Stafford, D. W. 1976. A general method for isolation of high molecular weight DNA from eukaryotes. *Nucleic Acids Research* 3:2303–8.
- Boetzer, M., Christiaan, V. H., Hans, J. J., Butler, D. and Pirovano, W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–9.
- Bolger, A. M., Lohse, M. and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–20.
- Eves-van den Akker, S., et al. 2016. The genome of the yellow potato cyst nematode, *Globodera rostochiensis*, reveals insights into the basis of parasitism and virulence. *Genome Biology* 17:124.
- Hussey, R. S. and Barker, K. R. 1973. A comparison of methods of collecting inocula of *Meloidogyne* spp., including a new technique. *Plant Disease Reporter* 57:1025–8.
- Jenkins, W. R. 1964. A rapid centrifugal-flotation technique for separating nematodes from soil. *Plant Disease Reporter* 48:692.
- Mitchum, M. G., Hussey, R. S., Baum, T. J., Wang, X., Elling, A. A., Wubben, M. and Davis, E. L. 2013. Nematode effector proteins: an emerging paradigm of parasitism. *New Phytologist* 199:879–94.
- Nadalin, F., Vezzi, F. and Policriti, A. 2012. Gapfiller: a de novo assembly approach to fill the gap within paired reads. *BMC Bioinformatics* 13 No. S14:S8.
- Parra, G., Bradnam, K. and Korf, I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* 23:1061–7.
- Richard, M. L., Bernardo, J. C., Leah, C., Matthew, D. C. and Mario, C. 2014. Nextclip: an analysis and read preparation tool for Nextera long mate pair libraries. *Bioinformatics* 30:566–8.
- Robinson, A. F., Inserra, R. N., Caswell-Chen, E. P., Vovlas, N. and Troccoli, A. 1997. *Rotylenchulus* species: Identification, distribution, host ranges, and crop plant resistance. *Nematropica* 27:127–80.
- Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J. M. and Birol, I. 2009. ABySS: A parallel assembler for short read sequence data. *Genome Research* 19:1117–23.
- Smit, A. F. A. and Hubley, R. 2015. Repeatmodeler open-1.0, 2008–2015, <http://www.repeatmasker.org/faq.html#faq3>.
- Smit, A. F. A., Hubley, R. and Green, P. 2015. Repeatmasker open-4.0, 2013–2015.