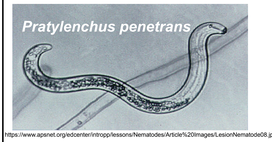
Genome Skimming Reveals Insights into Symbiosis and Chromosome Dynamics in Six Plant-Parasitic Nematode Species

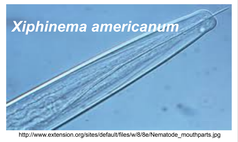
Amanda M.V. Brown[1], Dana K. Howe[1], Levi Dettwyler[1], Amy Peetz[2], Wendy S. Phillips[2], Inga A. Zasada[2], Dee R. Denver[1]

1 Department of Integrative Biology, Oregon State University, Corvallis, OR, USA; 2 USDA-ARS Horticultural Crops Research Laboratory, Corvallis, OR, USA

Abstract: Plant-parasitic nematodes (PPNs) are widespread and cause significant damage to agricultural crops. To date, little is known about the genetic structure and genomic features of most PPNs. Furthermore, some PPNs are known to form stable associations with bacterial endosymbionts thought to affect their reproductive mode (i.e. sexual vs. parthenogenesis), yet no genomes have been sequenced for these symbionts. We have begun a genomic survey of several representative PPNs, using a low-coverage genome skimming approach. PPN species analyzed included *Anguina agrostis*, *Globodera ellingtonae*, *Pratylenchus neglectus*, *Pratylenchus penetrans*, *Pratylenchus thornei*, and *Xiphinema americanum*. We performed Illumina MiSeq paired-end sequencing with 300 base pair reads with a target coverage per nuclear genome of ~15-30X. Illumina reads were de novo assembled and analyzed using a variety of approaches based on patterns of coverage, %GC, and assembled contig length. Our analyses revealed the presence of multiple bacterial symbionts in the genera *Cardinium*, *Janthinobacterium*, *Wolbachia*, and *Xiphinematobacter*. In *Pratylenchus*, we discovered evidence for complex ploidy states in different species. For *G. ellingtonae*, we have also collected Pacific Biosciences RS SMRT data and are optimizing hybrid assembly approaches. This work demonstrates the utility of genome skimming approaches for revealing novel features of eukaryotic genomes and the presence of previously unknown bacterial endosymbionts.



https://www.apsnet.org/edcenter/intropp/lessons/Nematodes/Article%20Images/LesionNematode08.jpg



http://www.extension.org/sites/default/files/w/8/8e/Nematode_mouthparts.jpg

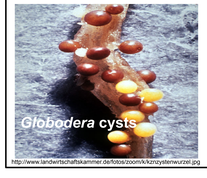
Six Sequenced PPNs:

Nematode	Common Name	Host
Anguina agrostis	Bentgrass nematode	grass
Globodera ellingtonae	Potato cyst nematode	potato
Pratylenchus neglectus	Lesion nematode	wheat
Pratylenchus penetrans	Lesion nematode	raspberry
Pratylenchus thornei	Lesion nematode	wheat
Xiphinema americanum	Dagger nematode	grape

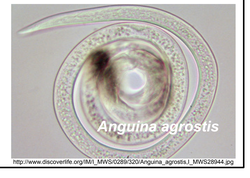
Two MiSeq Runs:

Nematode	Paired reads	N50* (bp)	Total coverage (bp)
Anguina agrostis	16,790,902	16,084	179,321,414
Globodera ellingtonae	18,597,846	12,612	107,194,447
Pratylenchus neglectus	19,772,702	50,199	137,782,631
Pratylenchus penetrans	19,164,652	5,260	445,571,183
Pratylenchus thornei	15,428,194	19,519	204,909,124
Xiphinema americanum	15,133,970	16,084	310,611,870

* Preliminary velvet assemblies with a kmer of 101

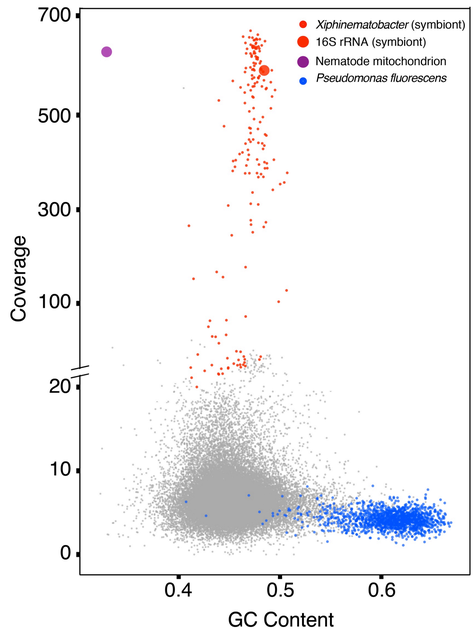


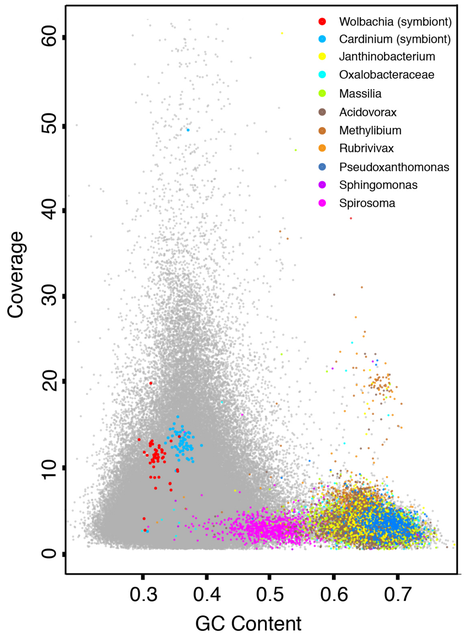
http://www.landwirtschaftskammer.de/fotos/zoom/kzbnzystenwurzel.jpg



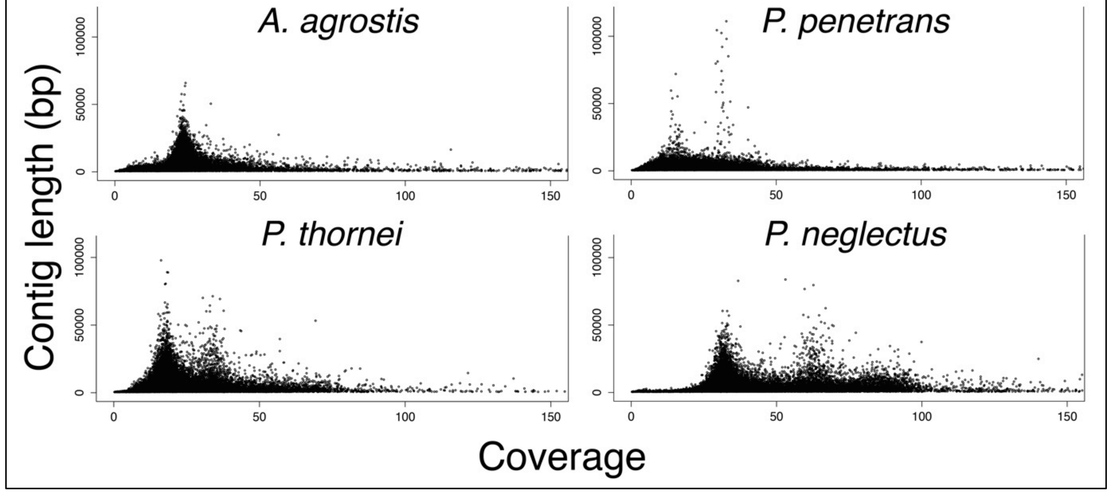
http://www.discoverlife.org/IM/I_MWS/0289/320/Anguina_agrostis,I_MWS28944.jpg

X. americanum and *P. penetrans* %GC vs. Coverage plots revealing symbionts *Xiphinematobacter*, *Cardinium*, *Wolbachia*, *Janthinobacterium*, and other bacteria.

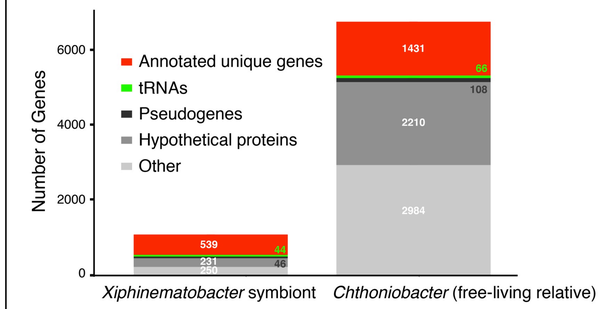


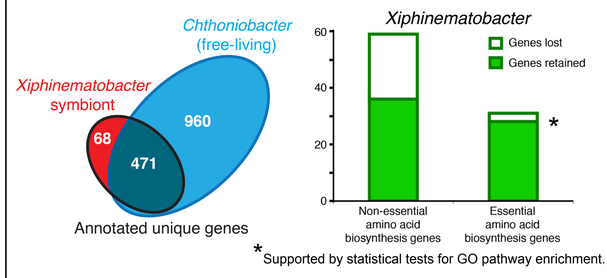


Coverage vs. Length plots showing de novo assembly at multiple coverage peaks, suggesting complex ploidy in *Pratylenchus* species, compared with *A. agrostis*. Within-species blast-analyses (not shown) indicated major regions of duplicated and diverged DNA corresponding to different coverage peaks.



Xiphinematobacter genome annotation suggesting possible nutritional role.





Summary: *Xiphinematobacter* gene loss and retention suggests a role in essential amino acid biosynthesis, consistent with its location in the gut.

Globodera PacBio preliminary assembly:

To address the significant challenge of de novo eukaryote genome assembly, we have been testing pipelines that involve long-read PacBio sequencing technology. Our previous tests showed 4-fold improved N50s using PacBio + Illumina hybrid assembly for *C. elegans*. Thus, we sequenced 2 PacBio SMRT cells for *G. ellingtonae*. Although PacBio sequencing has a much higher error rate than Illumina sequencing, its single-molecule long-reads help join scaffolds across difficult repetitive regions of the genome. So far, our *G. ellingtonae* PacBio data shows ~8X coverage with N50s of ~1550 bp when analyzed on its own. We hope to report details on the success of the hybrid assemblies soon.

Conclusions and Future Directions:

These genomic analyses from six plant-parasitic nematodes demonstrate significant potential of a genome-skimming approach. Our initial assembled contigs were able to reveal previously unknown symbionts, through %GC vs. Coverage plots. The contigs plotted to display Coverage vs. Length revealed an unexpected pattern in *Pratylenchus* that likely reflects complex ploidy similar to that found in *Meloidogyne* species. Our data suggest further genomic annotation of these symbionts and the PPN genomes will be fruitful and may help elucidate further biological features of these important plant pathogens.

Acknowledgements: We thank Mark Dasenko at the CGRB for sequencing assistance and Stephen Alderman and Mariella Ballato for assisting with worm collections. This work was funded by USDA grant SCA 58-5358-3-026.