Improving the draft genome of rice root-knot nematode *Meloidogyne graminicola* by long-read sequencing

Vishal Singh Somvanshi[1], Manoranjan Dash[1], Chaitra G. Bhat[1], Roli Budhwar[2], Jeffrey Godwin[2], Rohit N. Shukla[2], Andrea Patrignani[3], Ralph Schlapbach[3], Uma Rao[1]

[1]Division of Nematology, ICAR-Indian Agricultural Research Institute, New Delhi, Delhi, India 110012; [2]Bionivid Technology Private Limited, 209, 4th Cross Rd, B Channasandra, Kasturi Nagar, Bengaluru, Karnataka, India 560043; [3]Functional Genomics Center Zurich, Winterthurerstr. 190, Y32H52, 8057, Zurich, Switzerland

vishal.somvanshi@icar.gov.in

ePoster #3769

Place and Time zone: New Delhi, India; GMT+05:30

Introduction

First draft genome of *Meloidogyne graminicola* was of 38.18 Mb sequenced through short-read Illumina platform

Purpose

Improve previously available genome of *M. graminicola* by using a long read sequencing (PacBio Sequel platform)

Methods

PacBio Sequel raw data
↓
Primary assembly by Falcon and Falcon UnZip
↓
Primary assembly curartion by Purge Haplotig
↓
Removal of contamination by Bolbtools / NCBI
↓
Mapping of clean reads to cleaneqd assembly
↓
Re-run assembly with only the clean reads
↓
Assembly correction using Pilon (Illumina short reads) and Arrow (PacBio reads)
↓
Clustering of PacBio and Illumina Assemblies
↓
Final Draft Assembly
↓
Assembly validation by mapping the reads on to the final assembly and genome completeness assessment by BUSCO / CEGMA

Results

Parameter	Value
Assembled draft genome statistics	
Total scaffolds	514
Total bases	36,860,138
Average scaffold length (bp)	71,712.33
Median scaffold length (bp)	44,375
N25 length (bp)	226,273
N50 length (bp)	105,012
Number of protein coding gene models	14,062

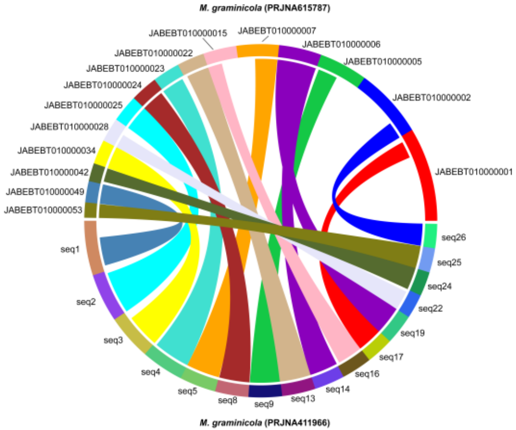

Fig. 1. A comparison of *M. graminicola* IARI strain draft genome assembly (PRJNA411966) to *M. graminicola* VN18 strain (PRJNA615787) by synteny analysis

Publication: Somvanshi et al. 2021. Gene.

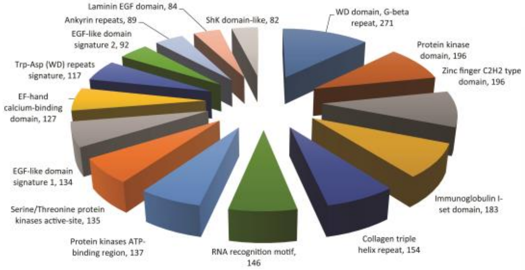

Fig. 2. Topmost fifteen abundant protein domains in the *M. graminicola* genome

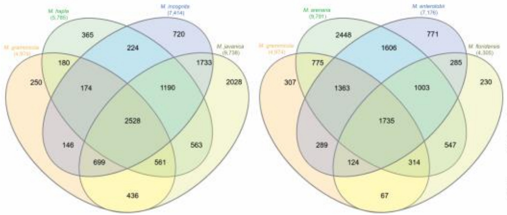

Fig. 3. The orthologous gene groups identified in the genomes of various *Meloidogyne* species

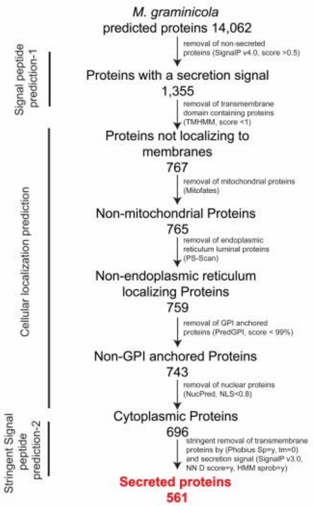

Fig.4. pipeline and tools used for identification of *M. graminicola* secreted proteins

Acknowledgements

We thank the Director, ICAR-IARI, for support and facilities. We acknowledge help from Dr. Etienne Danchin, INRAE, France and Dr. St´ephane Bellafiore for early access to their *M. graminicola* genome assembly.

Funding: This research was supported by the Department of Biotechnology, Ministry of Science and Technology, Government of India Grant No. BT/PR18924/COE/34/48/2017