

SCIENTIFIC REPORTS



OPEN

Genetic diversity and phylogeny of South African *Meloidogyne* populations using genotyping by sequencing

Milad Rashidifard¹, Hendrika Fourie¹, Pierre-Yves Véronneau², Mariette Marais³, Mieke Stefanie Daneel⁴ & Benjamin Mimee¹ ²

Meloidogyne species cause great crop losses worldwide. Although genetic host plant resistance is an effective control strategy to minimize damage caused by *Meloidogyne*, some resistant genes are ineffective against virulent species such as *Meloidogyne enterolobii*. Detailed knowledge about the genetic composition of *Meloidogyne* species is thus essential. This study focused on genotyping-by-sequencing (GBS) and Pool-Seq to elucidate the genetic relation between South African *M. enterolobii*, *M. incognita* and *M. javanica* populations. Hence, 653 common single nucleotide polymorphisms (SNPs) were identified and used to compare these species at genetic level. Allele frequencies of 34 SNPs consistently differed between the three *Meloidogyne* species studied. Principal component and phylogenetic analyses grouped the *M. enterolobii* populations in one clade, showing a distant relation to the *M. javanica* populations. These two species also shared genetic links with the *M. incognita* populations studied. GBS has been used successfully in this study to identify SNPs that discriminated among the three *Meloidogyne* species investigated. Alleles, only occurring in the genome of *M. enterolobii* and located in genes involved in virulence in other animal species (e.g. a serine/threonine phosphatase and zinc finger) have also been identified, accentuating the value of GBS in future studies of this nature.

Root-knot nematodes (*Meloidogyne*) are polyphagous, obligate pests that are distributed worldwide and parasitize almost all the higher plant species, resulting in great economic losses¹. *Meloidogyne incognita* is generally considered as the most damaging root-knot nematode species worldwide². Since this species can infect *Arabidopsis thaliana*, it is also a key model system to study metazoan adaptations to plant parasitism, hence its genome has already been elucidated³. However, *Meloidogyne enterolobii* listed as a threat species, can be confused with *M. incognita* and other thermophilic species due to it exhibiting similar morphological characteristics⁴. Of more significance is that *M. enterolobii* has the ability to overcome resistance genes that are effective against its thermophilic counterparts *Meloidogyne arenaria*, *M. incognita* and *Meloidogyne javanica*^{5–8}. *Meloidogyne enterolobii* for example had been reported to reproduce optimally on tomato and pepper that exhibit the *Mi-1*, *N* and *Tabasco* resistance genes, respectively⁸, while *M. arenaria* failed to reproduce on such resistant plants⁹. This phenomenon has far reaching implications for the management of this species.

Management of root-knot nematodes has been done traditionally by means of chemical control. This approach generally keeps the nematode population under the economic threshold level since eradicating these pests is considered impossible^{4,10}. However, the development of resistance against the different chemical compounds and the progressive withdrawal of synthetically-derived nematicides due to animal, human and environmental concerns^{11,12} are the main drives for the exploitation and use of environmentally-friendly strategies. Currently, genetic host plant resistance is a very effective and viable strategy to control root-knot nematodes in various agricultural cropping systems¹³. Nonetheless, the existence of virulent root-knot nematode populations reduces

¹Unit for Environmental Sciences and Management, North-West University, Private Bag X6001, 2520, Potchefstroom, South Africa. ²Agriculture and Agri-Food Canada, St-Jean-sur-Richelieu Research and Development Centre, 430 boul. Gouin, St-Jean-sur-Richelieu, J3B 3E6, Québec, Canada. ³Nematology Unit, Biosystematics, Agricultural Research Council - Plant Health and Protection (ARC-PHP), Private Bag X134, 0121, Queenswood, South Africa. ⁴Agricultural Research Council - Tropical and Subtropical Crops (ARC - TSC), Private Bag X11208, 1200, Mbombela, South Africa. Correspondence and requests for materials should be addressed to M.R. (email: 27216179@nwu.ac.za)

	mnC	minCov		
		5	20	50
11 populations	0.8	929	542	140
	1.0	653	277	59
<i>M. enterolobii</i> only	0.7	7,534	2,683	1,324
	1.0	5,572	2,092	1,032

Table 1. Influence of the minimum call rate (mnC) and the minimum coverage at each locus (minCov) on the number of SNPs identified by the UNEAK pipeline for *Meloidogyne enterolobii*, *M. incognita* and *M. javanica* populations from South Africa.

the efficacy of this strategy¹⁴. *Meloidogyne enterolobii* occurs in many countries and has initially been reported from the Mpumalanga Province in South Africa during the 1990s from guava (*Psidium guajava*) orchards¹⁵. Its established occurrence in South Africa fits the hypothesis of the late Dr Kent Kleynhans and Mr Piet Willers that the occurrence and host range is wider than the initial localities and hosts around Mbombela, Mpumalanga (personal communication, Dr Kent Keynhans, Agricultural Research Council-Plant Protection Research Institute, Pretoria, 1998). It has hence been reported from other crop production areas, infecting green pepper (*Capsicum annum*); potato (*Solanum tuberosum*) and tomato (*Solanum lycopersicum*)^{16,17}. This scenario justified a more detailed genetic study of South African *Meloidogyne* populations to determine if genomic differences linked with virulence exhibited by *M. enterolobii* could be found between *M. incognita* and *M. javanica*.

Several studies have been conducted to elucidate the genetic diversity of *Meloidogyne* populations by using different molecular techniques, e.g. random amplified polymorphic DNA (RAPD), restriction fragment length polymorphisms (RFLP), PCR based on sequences of rDNA, mtDNA, ITS and IGS, or satellite DNA probe markers^{18–26}. However, most of these methods are expensive, time-consuming, require several PCR analyses and many nematode individuals²⁷. The ultimate drawback of these techniques are that they are targeting only a small part of the genome of a nematode and are hence not optimal for pan-genomic comparison. It is therefore necessary to apply novel and rapid molecular genotyping tools to obtain more detailed information about the genetic diversity between *Meloidogyne* species.

Single nucleotide polymorphisms (SNPs) are popular and common molecular markers used to study the entire genome of nematodes^{28,29}. Significant advances in sequencing technologies are providing lots of information at relatively low cost³⁰. Genotyping by sequencing (GBS)³¹ is a simple protocol based on next generation-sequencing (NGS) of genomic fragments of organisms (e.g. nematodes) obtained by specific restriction enzymes followed by a bioinformatics pipeline³⁰. This enzyme-based reduction of complexity, combined with the use of barcodes for multiplexing considerably reduces sequencing cost while providing genome-wide information³². This technique has proven to be useful and accurate to characterize nematode species even when no information about the genome was available²⁷. Currently, good reference genomes are available for *M. hapla*³³ and *M. incognita*^{3,34}. However, no annotated reference genomes exist for *M. enterolobii* and *M. javanica*, although some assemblies from whole genome sequencing (WGS) data were published recently^{34,35}. Genotyping by sequencing has, for example successfully been applied in combination with Pool-Seq by Mimeo *et al.*²⁷, to investigate the genetic diversity among populations of the golden cyst nematode *Globodera rostochiensis*. Pool-Seq is a method described by Futschik and Schlötterer³⁶, which instead of sequencing isolated individuals directly uses a population (several individuals pooled together). When using a sufficiently big pool size, Pool-Seq even showed to be more appropriate for estimating allele frequencies and is more cost effective than sequencing the DNA of individuals³⁶.

This study aimed to investigate the genetic diversity of three different *Meloidogyne* species *viz.* *M. enterolobii*, *M. incognita* and *M. javanica* using GBS in order to highlight relationships among these species and loci putatively involved in virulence.

Results

SNP calling. The sequencing of 11 *Meloidogyne* populations digested with *PstI/MspI* restriction enzymes generated 83 038 291 reads. After initial quality control, 77 095 925 good barcoded reads were kept for further analysis. The UNEAK pipeline identified 2,786 SNPs before filtering. The final dataset contained from 59 to 929 SNPs depending on filtering stringency (Table 1). In order to ensure a good accuracy, the dataset without missing data containing 277 SNPs (minimum call rate = 1.0 and, minimum coverage = 20) was kept for phylogenetic and Principal Component Analyses (PCA). The dataset, containing 653 SNPs, was explored to find interesting markers. The allele frequencies at these loci in the 11 *Meloidogyne* populations are presented in Supplementary Table 1. When the pipeline was run on the *M. enterolobii* populations only, 13,047 SNPs were identified. Of these, 2,092 were present in all populations and supported by a minimum coverage of 20 reads (Table 1).

A total of 1,016 variants were called after alignment of the raw reads to the *M. incognita* genome and 419 of them were in predicted genes. However, very few remained after filtering. For a minimum call rate of 1.0, only 95, 84 and 64 SNPs had a minimum coverage of 5, 20 and 50 reads, respectively, and 148, 122 and 93 SNPs when a minimum call rate of 0.8 was used with the same coverage.

Population genetics. PCA of allele frequencies of the 11 *Meloidogyne* populations using the 277 SNPs dataset separated the populations in two main clusters (Fig. 1). All the *M. enterolobii* populations (R1, R4, R5 and R6) grouped in a first cluster, while *M. javanica* populations (R24, R27, R28, R30 and R31) were placed in a second. This distinction was very clear and the differences between these two species (first dimension in the PCA)

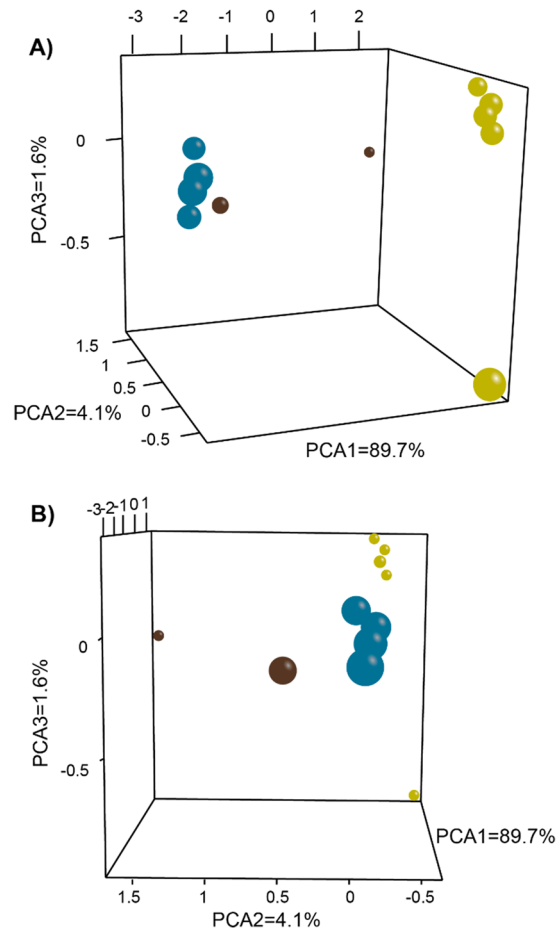


Figure 1. Principal Component Analysis (PCA) of 11 South African *Meloidogyne* populations based on allele frequencies at 277 loci and showing the genetic relations among two *M. incognita* (brown spheres), four *M. enterolobii* (blue spheres), and five *M. javanica* (beige spheres) populations. Panels show (A) First dimension on x axis and (B) second dimension on x axis.

explained 89.7% of the variation in the dataset (Fig. 1A). Although the two populations of *M. incognita* (R25 and R34) exhibited different genetic relation on this first dimension, this species was clearly separated from *M. enterolobii* and *M. javanica* in the second dimension of the PCA explaining 4.1% of the total variation (Fig. 1B).

The phylogenetic tree, using the dataset of 277 SNPs, also revealed two main clusters separating *M. enterolobii* populations from those of *M. javanica*, with *M. incognita* populations being intermediate (Fig. 2).

A direct comparison of allele frequencies between species highlighted several SNPs clearly differentiating the three species. Among these, three had different zygosity status between the species in all the populations tested and 31 were homozygous for an allele only in *M. enterolobii* (Table 2). The sequence surrounding each SNP has been retrieved from the *M. incognita* reference genome. Out of these, 19 were located in predicted genes. When screening the variants obtained by aligning the raw reads to the *M. incognita* genome, 14 were located in genes coding for 10 different proteins and had allele frequencies specific to *M. enterolobii* (Table 3).

Discussion

This study represents a baseline investigation of the genetic diversity of South African *Meloidogyne* populations. Root-knot nematodes are reported to infect various crop hosts and to cause great damage and economic losses in South Africa. Since *M. enterolobii* is known to be highly virulent, and because resistant cultivars are not equally effective against the different *Meloidogyne* species, accurate and reliable species identification is crucial to use appropriate management strategies.

The GBS method used in this study proved to be useful in identifying diagnostic SNPs for the discrimination of three of the highly damaging thermophilic *Meloidogyne* species occurring in South Africa. Accurate distinction of *M. enterolobii*, *M. javanica* and *M. incognita* was still challenging using various molecular techniques, except for SCAR-PCR²⁰. Although *M. enterolobii* can be separated from other *Meloidogyne* species by the use of various universal markers e.g. 28S, COI, 16S and IGS^{24,37,38}, distinguishing between *M. javanica* and *M. incognita* has been unsuccessful in various studies due to the high genetic similarity among the latter species and *M. arenaria*^{37–41}. In this study, however, several putatively discriminative SNPs to *M. javanica*, *M. incognita* and *M. enterolobii* were identified and will enable the accurate distinction between these species through the development of allele-specific PCR. The good number of SNPs obtained with UNEAK on the *M. enterolobii* populations alone

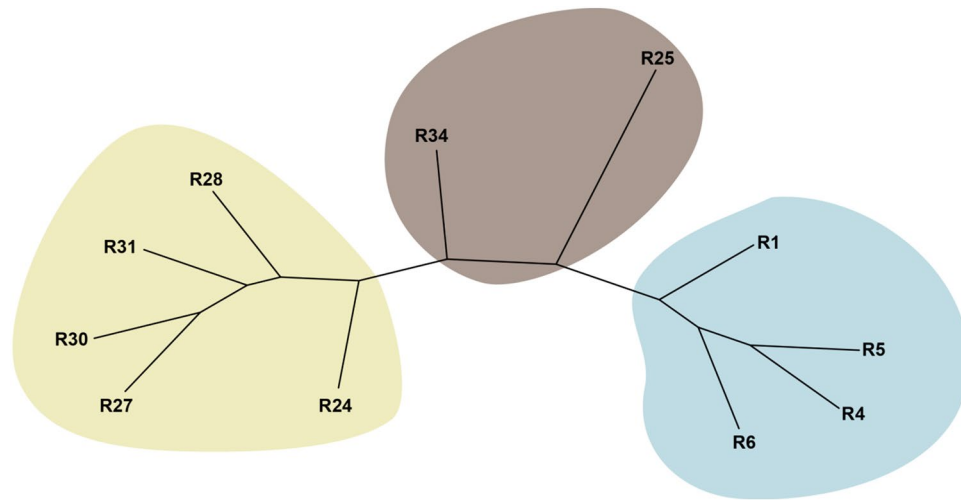


Figure 2. Neighbor-joining phylogenetic tree of 11 South African *Meloidogyne* populations, representing two *M. incognita* (brown), four *M. enterolobii* (blue) and five *M. javanica* (beige) populations, based on genome-wide allele frequencies at 277 loci.

revealed the high potential of this method and indicated that the technique will be very useful to compare more populations of this species and to include co-variables like virulence or origin for a more in-depth characterization. Ultimately, this kind of comparison should be done using whole genome sequencing. However, when dealing with numerous samples, the cost associated with WGS is still prohibitive. Thus, GBS represents an interesting alternative. This molecular approach was initially described for plants by Elshire *et al.*³¹, and modified specifically for nematodes by Mimee *et al.*²⁷. It was used for the first time during this study to investigate the genetic diversity in *Meloidogyne* species. The technique also takes advantage of Pool-Seq (sequencing of composite samples) which removed the fastidious step of isolating and extracting DNA from single juveniles. The bioinformatics pipeline allowed the rapid *de novo* identification of SNPs, without the need of a reference genome. When aligning the sequencing reads to the closest reference species, *M. incognita*, we found less good quality SNPs (no missing data and good coverage) than when using the GBS pipeline. This indicates that significant differences exist between the two species. This was confirmed by the eight-fold increase in the number of SNPs when the pipeline was run with *M. enterolobii* populations alone. As UNEAK will only keep the SNPs that are present in all populations (at minimum call rate = 1.0), this approach indicates that the three species are probably more different than anticipated. On the other hand, some genetic variants were found upon alignment with the *M. incognita* genome and not by the UNEAK pipeline. This is explained by the high stringency of the pipeline that only tolerates one mismatch by read and rejects sequences with multiple SNPs or other kind of variants. Therefore, combining the two approaches will maximize the discovery rate of SNPs in the genome of *Meloidogyne* species. Also, it was hypothesized that several mitotic parthenogens *Meloidogyne* species had acquired pairs of divergent gene copies during past hybridisation event³⁵. This will result in an excess of heterozygosity that could complicate classical phylogenetic analyses. One of the advantage of GBS is that it is not affected by ploidy as the UNEAK pipeline compare all the sequences. Thus, if the same sequence is present in many versions in an organism due to ploidy or gene duplication, the pipeline will compute all versions together and output the allele frequency for that sequence and not for each physical locus in the genome. Furthermore, using Pool-Seq, we theoretically captured all the allelic diversity of each population.

PCA analyses confirmed the genetic separation between *M. enterolobii*, *M. javanica* and *M. incognita*. This result was expected since the distinction of these species has been reported using different markers^{24,37}. The analysis also revealed a close genetic proximity of the four *M. enterolobii* populations when the 277 SNPs dataset (present in all four populations) was used. This is in agreement with results obtained by Tigano *et al.*²² that described *M. enterolobii* as a geographic homogenous species with low diversity between populations. Similarly, Onkendi and Moleleki³⁸ showed 100% homology between sequences of IGS and COII from South African *M. enterolobii* populations. However, when the GBS approach was used for *M. enterolobii* only, more variation was observed and suggests that this species could be more diverse than initially thought. As genetic diversity is an important driver of adaptation to new environmental conditions or host plants, it will be interesting to explore this diversity by studying more *M. enterolobii* populations.

When comparing allele frequencies of 277 SNPs distributed all across the genome, the two populations of *M. incognita* investigated in this study showed substantial genetic difference, one being related to *M. enterolobii* and the other to *M. javanica*. A possible explanation for this phenomenon may be that in this analysis only loci sequenced in all *Meloidogyne* populations were conserved, meaning that the sequences that were unique to one or two species only (not present or not sequenced in the other) were discarded even if they contain SNPs. Thus, this place the emphasis on the differences that exist in terms of SNPs contained in the genomes of these three *Meloidogyne* species.

Locus	Alleles	Allele frequency ^a			Localization or gene function
		<i>M. enterolobii</i>	<i>M. incognita</i>	<i>M. javanica</i>	
<i>Discriminating all 3 species</i>					
TP40058	G/A	1	0.37	0	Muscle M-line assembly protein unc-89
TP37437	T/C	1	0.56	0	n.a. ^b
TP30509	T/A	0	0.69	1	n.a.
<i>Homozygous in Me^c and different in other species</i>					
TP11533	T/A	1	0.50	0.06	uncharacterized protein (CBN07616) <i>Caenorhabditis bremeri</i>
TP1752	T/C	0	0.60	1	Histone deacetylase
TP17550	A/T	0	0.08	0.08	n.a.
TP22081	T/C	0	0.55	1	Zinc finger, C2H2
TP27938	T/C	0	0.58	1	intergenic
TP27939	T/C	0	0.94	1	intergenic
TP27941	T/C	0	0.55	0.99	n.a.
TP28969	G/A	1	0.43	0	vacuolar protein sorting-associated protein VTA1
TP30499	T/A	0	0.56	1	n.a.
TP34022	T/G	1	0.38	0.20	n.a.
TP34025	G/T	1	0.33	0.25	cytoplasmic polyadenylation element-binding protein
TP35049	G/A	0	0.53	0.63	n.a.
TP35052	G/A	0	0.49	0.46	spectrin beta, non-erythrocytic 4 ^d
TP35086	G/A	0	0.47	0.45	n.a.
TP35088	G/A	0	0.47	0.53	Plectin repeat ^e
TP36654	G/A	0	0.60	1	Valine—tRNA ligase
TP36738	G/A	0	0.54	0.99	Valine—tRNA ligase
TP37436	T/C	1	0.85	0.03	Unknown
TP4522	T/C	1	0.48	0.02	Intergenic
TP555	G/A	0	0.14	0.81	n.a.
TP558	G/A	0	0.23	0.82	n.a.
TP7297	G/C	1	0.67	0.26	intron
TP9881	G/A	0	0.65	1	n.a.

Table 2. Homozygous loci in *Meloidogyne enterolobii*, obtained by genotyping by sequencing, that putatively discriminate the species from two other *Meloidogyne* species from South Africa. ^aAllele frequencies represent the mean of four populations for *M. enterolobii*, two for *M. incognita* and five for *M. javanica*. ^bn.a. indicated that this sequence was not retrieved from the reference genome. ^c*Meloidogyne enterolobii*. ^dTwo similar SNPs (TP35055 and TP35079) were found in this protein and probably represent gene duplication or alignment artefacts. ^eSix similar SNPs (TP35091, TP35096, TP35115, TP35118, TP35125 and TP35129) were found in this protein and probably represent gene duplication or alignment artefacts.

This study also highlighted allelic differences between the species that exists in specific genes that could modify protein sequence and function. Ultimately, in depth characterization of the *M. enterolobii* genome will give us a better understanding of why the species is more aggressive and overcome resistance genes that are effective against *M. incognita* and *M. javanica*. For now, 19 SNPs identified by UNEAK and 14 by the alignment on the *M. incognita* reference genome were located in exons of predicted genes. These variants were specific to *M. enterolobii* and could affect nematode-host interactions. Some of these genes have already been reported to be involved in parasitism. For example, two SNPs were located in genes coding for a serine/threonine phosphatase. A protein with a similar sequence is a known effector in the Hessian fly, *Mayetiola destructor*. This plant-galling arthropod uses effectors to modify host cells in a way that is superficially similar to root-knot nematodes. The type-2 serine/threonine protein phosphatase (PP2C) domain was shown to be associated with the ability of the fly to survive and parasitize wheat seedlings in susceptible plants. This protein is also recognized as an avirulence factor in cultivars carrying the *H24* resistant genes. Interestingly, a single loss-of-function mutation in that gene is sufficient to overcome resistance in virulent population⁴².

Results from this study also highlighted differences between *M. enterolobii* and the other species in a sequence coding for a zinc finger, C2H2 domain. These small protein motifs are known to interact with different molecules and are involved in multiple functions from gene transcription and mRNA trafficking to protein folding and apoptosis⁴³. Wang *et al.*⁴⁴, reported that the C2H2 zinc finger *PsCZF1* was involved in pathogenesis in *Phytophthora sojae* by demonstrating that *PsCZF1*-deficient mutants lost virulence on different soybean cultivars. Another study on *Alternaria brassicicola* showed that a knockdown mutation in the zinc finger transcription factor *AbVf19* resulted in a 90% decrease in virulence⁴⁵. Interestingly, Gross and Williamson⁴⁶ identified a novel transposable element in root-knot nematodes that contained a C2H2 zinc-finger motif and that could be involved in the ability of some populations to bypass resistance in tomato mediated by the *Mi-1* gene.

Locus	Alleles	Allele frequency ^a			Description	E value
		<i>M. enterolobii</i>	<i>M. javanica</i>	<i>M. incognita</i>		
MiV1ctg11_241891	G/A	0	1	1	fibronectin type III domain	5.71e-23
MiV1ctg17_134782	T/G	0	1	0.52	Ubiquitin carboxyl-terminal hydrolase 4	2.95e-07
MiV1ctg2_298920	T/C	0	0.29	0.67	Major facilitator superfamily MFS-1 domain containing	4.31e-16
MiV1ctg27_10072	A/T	0	0.97	0.52	Serine threonine- phosphatase 2B catalytic subunit	2.25e-05
MiV1ctg27_9897	T/C	0	0.99	0.52	Serine threonine- phosphatase 2B catalytic subunit	2.25e-05
MiV1ctg39_178906	A/G	0	0.58	0.59	Cytoskeleton-associated 5	1.43e-11
MiV1ctg39_178981	T/A	1	0.04	0.35	Cytoskeleton-associated 5	1.43e-11
MiV1ctg59_26757	A/G	0	1	0.53	Glycogenin-1	1.95e-25
MiV1ctg59_26792	A/G	0	1	0.53	Glycogenin-1	2.45e-18
MiV1ctg61_132221	ATCAA/ACAG	0	1	0.67	sodium- and chloride-dependent glycine transporter 1	2.46e-16
MiV1ctg7_255261	C/T	1	0.71	0.22	choline Carnitine O-acyltransferase	1.07e-12
MiV1ctg7_255276	G/T	0	1	0.57	choline Carnitine O-acyltransferase	1.23e-12
MiV1ctg7_56934	T/C	0	1	0.89	Valyl-tRNA synthetase	3.06e-49
MiV1ctg75_84149	A/C	0	1	0.83	adenylate kinase isoenzyme 5	1.13e-16

Table 3. Homozygous loci in *Meloidogyne enterolobii*, obtained by alignment of reads on the *M. incognita* reference genome that discriminate the species from the other selected *Meloidogyne* species from South Africa, and that are located in annotated genes. ^aAllele frequencies represent the mean of four populations for *M. enterolobii*, two for *M. incognita* and five for *M. javanica*.

Finally, the plant cytoskeleton is hypothesized to play a crucial role in host defense response and a target for virulence⁴⁷. Root-knot nematodes are actively remodeling the cytoskeleton of infected plants during feeding cell and gall formation⁴⁸. This reprogramming is induced by secreted effectors, several being homologous to plant proteins⁴⁹. In this study, we have identified mutations in several *M. enterolobii* genes that code for proteins required for microtubule and cytoskeleton formation: Cytoskeleton-associated 5, spectrin and plectin repeat.

In this paper, the GBS technique proved to be a powerful and accurate technique to obtain detailed information about the diversity that exists among root-knot nematode species. The bioinformatic pipeline allowed the identification of high-quality diagnostic SNPs from the South African *Meloidogyne* species. The phylogenetic comparison of these variants generated valuable and novel knowledge about the genetic diversity of *Meloidogyne* species. Candidate genes associated with virulence were also highlighted and should be further explored to evaluate whether they are involved in *M. enterolobii* pathogenicity.

Methods

Sampling and DNA extraction. Eleven *Meloidogyne* populations, representing *M. enterolobii*, *M. incognita* and *M. javanica* (Table 4), were obtained from root and rhizosphere soil of different host plants in the Mpumalanga and Limpopo provinces of South Africa during 2015 and 2016. Single egg masses of each population were inoculated on roots of two-leaf stage tomato seedlings of a root-knot nematode susceptible cultivar (Floradade)⁵⁰ to ensure species purity. Subsequent mass rearing of pure populations was done in the glasshouse on tomato seedlings (cultivar Floradade) planted in individual pots that were filled with sandy-loam soil (5.3% clay, 93.6% sand, 1.1% silt, 0.47% organic matter and pH (H₂O) of 7.47) previously fumigated with Telone II (as 1,3-dichloropropane; dosage of 150 l/ha). An ambient temperature range of 21 (min) – 26 °C (max) and 14L:10D photoperiod were maintained in the glasshouse. Sixty days after inoculation, approximately 100 females of each population were randomly dissected from infected tomato roots of individual plants and DNA extracted by using DNeasy Blood & Tissue kit (Qiagen, Germany) according to the protocol supplied. The DNA of each *Meloidogyne* population was quantified using the Invitrogen Quant-iT Qubit dsDNA HS Assay Kit (Invitrogen, USA), and 30 ng of each sample was used for sequencing library construction.

Species identification. All populations of *Meloidogyne* species used in this study were identified using morphological and molecular approaches. Morphological identification was done based on female oesophagus and perineal patterns. The molecular confirmation was done by sequencing the 28S rDNA (D2-D3)⁵¹, COI⁵² and NADH5³⁷ mtDNA, regions and comparing to NCBI database, and by the amplification of a species specific SCAR marker^{20,53}.

Library preparation and sequencing. Sample preparation and sequencing were done by the Genomics Analysis Core Facility at the Institute for Integrative and Systems Biology (IBIS; Université Laval, Quebec, Canada) according to the GBS method developed by Elshire *et al.*³¹, using *PstI/MspI* restriction enzymes. The library was sequenced on one Ion Proton chip (ThermoFisher Scientific) at IBIS.

Sample ID	Nematode Species	Locality of origin	Host Plant
R1	<i>M. enterolobii</i>	Mbombela 1 (Mpumalanga)	Guava: <i>Psidium guajava</i>
R4	<i>M. enterolobii</i>	Hoedspruit (Limpopo)	Guava: <i>Psidium guajava</i>
R5	<i>M. enterolobii</i>	Erik Boerdery (Mpumalanga)	Guava: <i>Psidium guajava</i>
R6	<i>M. enterolobii</i>	Erik Boerdery (Mpumalanga)	Guava: <i>Psidium guajava</i>
R25	<i>M. incognita</i>	Mooketsi 1 (Limpopo)	Tomato: <i>Solanum lycopersicum</i>
R34	<i>M. incognita</i>	Pont Drift (Limpopo)	Tomato: <i>Solanum lycopersicum</i>
R24	<i>M. javanica</i>	Mbombela 2 (Mpumalanga)	Tomato: <i>Solanum lycopersicum</i>
R27	<i>M. javanica</i>	Mooketsi 2 (Limpopo)	Tomato: <i>Solanum lycopersicum</i>
R28	<i>M. javanica</i>	Mooketsi 3 (Limpopo)	Tomato: <i>Solanum lycopersicum</i>
R30	<i>M. javanica</i>	Polokwane (Limpopo)	Tomato: <i>Solanum lycopersicum</i>
R31	<i>M. javanica</i>	Mbombela 3 (Mpumalanga)	Spinach: <i>Spinacia oleracea</i>

Table 4. Root-knot nematode species used in this study as well as the origin of the species and host plants which it infected.

SNP calling and processing of the raw sequences. The Universal Network Enabled Analysis Kit (UNEAK) pipeline⁵⁴ was used to analyse Pool-Seq data and create the SNP list for the 11 *Meloidogyne* populations. To eliminate sequencing errors, a tolerance rate of 0.03 was used for UNEAK running. Filtering of SNPs was done by setting the minimum call rate (number of populations in which a locus must have been scored) at 0.8 (i.e. <20% missing data) or 1.0 (no missing data). The minimum minor allele frequency (MAF) threshold was set to 0.01. The effect of locus coverage on the final SNP number was also compared by setting the minimum number of reads at a given locus in each population required to accept a SNP (minCov) to 5, 20, or 50. The UNEAK pipeline was also run on the *M. enterolobii* populations alone with the minimum call rate set at 0.7 or 1.0 and the minimum coverage at 5, 20 and 50.

Raw sequences were also aligned on the *M. incognita* genome³ by using burrows-wheeler aligner (BWA)⁵⁵ after trimming with Trimmomatic (TRAILING = 20, MINLEN = 20) (<http://www.usadellab.org/cms/?page=trimmomatic>) and demultiplexing with Sabre (<https://github.com/najoshi/sabre>). Then, variants and annotations were called with freebayes⁵⁶ and SnpEff⁵⁷ respectively using the gene-finding format (GFF) of *M. incognita*.

Population genetics. A neighbor-joining phylogenetic tree was elaborated by using Gendist and Neighbor modules in PHYLIP v3.695 with the genome-wide allele frequencies obtained from UNEAK. PCA was done in R by using the same sets of allele frequencies with the *prcomp()* function from the *stats* package.

The SNPs of interest from UNEAK were retrieved from the *M. incognita* genome by means of BLASTN with the default parameters, except for a smaller word size of 4, with the Blast2GO application⁵⁸.

Data Availability

The data are submitted to NCBI SRA Portal with the following information. BioProject (PRJNA485255) and accession number (SAMN09786892, SAMN09786893, SAMN09786893, SAMN09786895, SAMN09786896, SAMN09786897, SAMN09786898, SAMN09786899, SAMN09786900, SAMN09786901 and SAMN09786902).

References

- Jones, J. T. *et al.* Top 10 plant-parasitic nematodes in molecular plant pathology. *Mol. Plant Pathol.* **14**, 946–961 (2013).
- Trudgill, D. L. & Blok, V. C. Apomictic, polyphagous root-knot nematodes: exceptionally successful and damaging biotrophic root pathogens. *Annu. Rev. Phytopathol.* **39**, 53–77 (2001).
- Abad, P. *et al.* Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nat. Biotechnol.* **26**, 909–915 (2008).
- Moens, M., Perry, R. N. & Starr, J. L. *Meloidogyne* species – a diverse group of novel and important plant parasites in *Root-knot Nematodes* (eds Perry, R. N., Moens, M. & Starr, J. L.) 1–13 (CAB International, 2009).
- Williamson, V. M. Plant nematode resistance genes. *Curr. Opin. Plant Biol.* **2**, 327–331 (1999).
- Brito, J. A. *et al.* Effects of the *Mi-1*, *N* and *Tabasco* genes on infection and reproduction of *Meloidogyne mayaguensis* on tomato and pepper genotypes. *J. Nematol.* **39**, 327–332 (2007).
- Brito, J. A., Stanley, J., Mendes, M. L., Cetintas, R. & Dickson, D. W. Host status of selected cultivated plants to *Meloidogyne mayaguensis* in Florida. *Nematropica* **37**, 65–72 (2007).
- Thies, J. A., Dickson, D. W. & Fery, R. L. Stability of resistance to root-knot nematodes in ‘Charleston Belle’ and ‘Carolina Wonder’ bell peppers in a sub-tropical environment. *HortScience* **43**, 188–190 (2008).
- Kiewnick, S., Dessimoz, M. & Franck, L. Effects of the *Mi-1* and the *N* root-knot nematode-resistance gene on infection and reproduction of *Meloidogyne enterolobii* on tomato and pepper cultivars. *J. Nematol.* **41**, 134–139 (2009).
- Karssen, G., Wesemae, W. M. L. & Moens, M. Root-knot nematodes in *Plant Nematology* (ed. Perry, R. N. & Moens, M.) 73–108 (CAB International, 2013).
- WHO (World Health Organization). Public health impact of pesticides used in agriculture (1990).
- Kishi, M. *et al.* Relationship of pesticide spraying to signs and symptoms in Indonesian farmers. *Scand. J. Work Environ. Health* **21**, 124–133 (1995).
- Molinari, S. Natural genetic and induced plant resistance, as a control strategy to plant-parasitic nematodes alternative to pesticides. *Plant Cell Rep.* **30**, 311–323 (2011).
- Bakker, J., Folkertsma, R. T., Rouppe van der Voort, J. N. A. M., De Boer, J. M. & Gommers, F. J. Changing concepts and molecular approaches in the management of virulence genes in potato cyst nematodes. *Annu. Rev. Phytopathol.* **31**, 169–190 (1993).
- Willers, P. First record of *Meloidogyne mayaguensis* Rammah and Hirschmann, 1988: Heteroderidae on commercial crops in the Mpumalanga province, South Africa. *Inligtingsbulletin-ITSG*. 19–20 (1997).

16. Marais, M., Swart, A. & Buckley, N. Overview of the South African plant-parasitic nematode survey (SAPPNS) in *Nematology in South Africa: A View from the 21st Century* (eds Fourie, H., Spaull, V. W., Jones, R. K., Daneel, M. S. & De Wale, D.) 73–117 (Springer, 2017).
17. Visagie, M. *et al.* Identification of *Meloidogyne* species associated with agri- and horticultural crops in South Africa. *Nematology* **20**, 397–401 (2018).
18. Zijlstra, C. Identification of *Meloidogyne chitwoodi*, *M. fallax* and *M. hapla* based on SCAR-PCR: a powerful way of enabling reliable identification of populations or individuals that share common traits. *Eur. J. Plant Pathol.* **106**, 283–290 (2000).
19. Onkendi, E. M. & Moleleki, L. N. Distribution and genetic diversity of root-knot nematodes (*Meloidogyne* species) in potatoes from South Africa. *Plant. Pathol.* **62**, 1184–1192 (2013).
20. Zijlstra, C., Donkers-Venne, D. T. H. M. & Fargette, M. Identification of *Meloidogyne incognita*, *M. javanica* and *M. arenaria* using sequence characterised amplified region (SCAR) based PCR assays. *Nematology* **2**, 847–853 (2000).
21. Cenis, J. L., Opperman, C. H. & Triantaphyllou, A. C. Cytogenetic, enzymatic, and restriction fragment length polymorphism variation of *Meloidogyne* species from Spain. *Phytopathology* **82**, 527–531 (1992).
22. Tigano, M. *et al.* Genetic diversity of the root-knot nematode *Meloidogyne enterolobii* and development of a SCAR marker for this guava-damaging species. *Plant Pathol.* **59**, 1054–1061 (2010).
23. Stanton, J., Jugall, A. & Moritz, C. Nucleotide polymorphisms and an improved PCR-based mtDNA diagnostic for parthenogenetic root-knot nematodes (*Meloidogyne* species). *Fundam. Appl. Nematol.* **20**, 261–268 (1997).
24. Bekker, S. *et al.* Discriminating between the eggs of two egg-mass-producing nematode genera using morphometric and molecular techniques. *Nematology* **18**, 1119–1123 (2016).
25. Castagnone-Sereno, P., Esparrago, G., Abad, P., Leroy, F. & Bongiovanni, M. Satellite DNA as a target for PCR-specific detection of the plant-parasitic nematode *Meloidogyne hapla*. *Current Genetics* **28**, 566–570 (1995).
26. Castagnone-Sereno, P., Semblat, J.-P., Leroy, F. & Abad, P. A new AluI satellite DNA in the root-knot nematode *Meloidogyne fallax*: relationships with satellites from the sympatric species *M. hapla* and *M. chitwoodi*. *Mol. Biol. Evol.* **15**, 1115–1122 (1998).
27. Mimee, B. *et al.* A new method for studying population genetics of cyst nematodes based on Pool-Seq and genome-wide allele frequency analysis. *Mol. Ecol. Resour.* **15**, 1356–1365 (2015).
28. Wicks, S. R., Yeh, R. T., Gish, W. R., Waterston, R. H. & Plasterk, R. H. A. Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nat. Genet.* **28**, 160–164 (2001).
29. Davis, M. W. *et al.* Rapid single nucleotide polymorphism mapping in *C. elegans*. *BMC Genomics* **6**, 118, <https://doi.org/10.1186/1471-2164-6-118> (2005).
30. Jarquín, D. *et al.* Genotyping by sequencing for genomic prediction in a soybean breeding population. *BMC Genomics* **15**, 740 (2014).
31. Elshire, R. J. *et al.* A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **6**, e19379, <https://doi.org/10.1371/journal.pone.0019379> (2011).
32. Poland, J. A. & Rife, T. W. Genotyping-by-sequencing for plant breeding and genetics. *The Plant Genome* **5**, 92–102 (2012).
33. Opperman, C. H. *et al.* Sequence and genetic map of *Meloidogyne hapla*: A compact nematode genome for plant parasitism. *Proc. Natl. Acad. Sci.* **105**, 14802–14807 (2008).
34. Blanc Mathieu, R. *et al.* Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematode. *PLoS Genet.* **13**, e1006777, <https://doi.org/10.1371/journal.pgen.1006777> (2017).
35. Sztitenberg, A. *et al.* Comparative genomics of apomictic root-knot nematodes: hybridization, ploidy, and dynamic genome change. *Genome Biol. Evol.* **9**, 2844–2861 (2017).
36. Futschik, A. & Schlötterer, C. The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics* **186**, 207–218 (2010).
37. Janssen, T., Karssen, G., Verhaeven, M., Coyne, D. & Bert, W. Mitochondrial coding genome analysis of tropical root-knot nematodes (*Meloidogyne*) supports haplotype based diagnostics and reveals evidence of recent reticulate evolution. *Sci. Rep.* **6**, 1–13 (2016).
38. Onkendi, E. M. & Moleleki, L. N. Detection of *Meloidogyne enterolobii* in potatoes in South Africa and phylogenetic analysis based on intergenic region and the mitochondrial DNA sequences. *Eur. J. Plant Pathol.* **136**, 1–5 (2013).
39. Zeng, Y., Ye, W. M. & Kerns, J. First report and morphological and molecular characterization of *Meloidogyne incognita* from *Radermachera sinica* in China. *Nematropica* **44**, 118–129 (2014).
40. De Ley, I. *et al.* Phylogenetic analyses of internal transcribed spacer region sequences within *Meloidogyne*. *J. Nematol.* **31**, 530–531 (1999).
41. Blok, V. C. Achievements in and future prospects for molecular diagnostics of plant-parasitic nematodes. *Can. J. Plant. Pathol.* **27**, 176–185 (2005).
42. Zhao, C. *et al.* Avirulence gene mapping in the Hessian fly (*Mayetiola destructor*) reveals a protein phosphatase 2C effector gene family. *J. Insect. Physiol.* **84**, 22–31 (2016).
43. Laity, J. H., Lee, B. M. & Wright, P. E. Zinc finger proteins: new insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* **11**, 39–46 (2001).
44. Wang, Y. *et al.* The *PsCZF1* gene encoding a C2H2 zinc finger protein is required for growth, development and pathogenesis in *Phytophthora sojae*. *Microb. Pathog.* **47**, 78–86 (2009).
45. Srivastava, A. *et al.* A zinc-finger-family transcription factor, *AbVf19*, is required for the induction of a gene subset important for virulence in *Alternaria brassicicola*. *Mol. Plant. Microbe. Interact.* **25**, 443–452 (2012).
46. Gross, S. M. & Williamson, V. M. Tm1: a mutator/foldback transposable element family in root-knot nematodes. *PLoS One* **6**, e24534, <https://doi.org/10.1371/journal.pone.0024534> (2011).
47. De Almeida Engler, J. & Favery, B. The plant cytoskeleton remodelling in nematode induced feeding sites in *Genomics molecular genomics plant-nematodes interaction* (eds Jones, J., Gheysen, G. & Fenoll, C.) 369–393 (Springer, 2011).
48. De Almeida Engler, J., De Siqueira, K. M. S., Do Nascimento, D. C., Da Costa, T. G. & Englera, G. A cellular outlook of galls induced by root-knot nematodes in the model host. *Arabidopsis thaliana. Nematoda* **3**, e062016, <https://doi.org/10.4322/nematoda.00616> (2016).
49. Bellafiore, S. *et al.* Direct identification of the *Meloidogyne incognita* secretome reveals proteins with host cell reprogramming potential. *PLoS Pathog.* **4**, e1000192, <https://doi.org/10.1371/journal.ppat.1000192> (2008).
50. Fourie, H., Mc Donald, A. H., Mothata, T. S., Ntidi, K. N. & De Waele, D. Indications of variation in host suitability to root-knot nematode populations in commercial tomato varieties. *Afr. J. Agric. Res.* **7**, 2344–2352 (2011).
51. Subbotin, S. A., Sturhan, D., Chizhov, V. N., Vovlas, N. & Baldwin, J. G. Phylogenetic analysis of Tylenchida Thorne, 1949 as inferred from D2 and D3 expansion fragments of the 28S rRNA gene sequences. *Nematology* **8**, 455–474 (2006).
52. Derycke, S., Vanaverbeke, J., Rigaux, A., Backeljau, T. & Moens, T. Exploring the use of cytochrome oxidase c subunit 1 (COI) for DNA barcoding of free-living marine nematodes. *PLoS One* **5**, e13716, <https://doi.org/10.1371/journal.pone.0013716> (2010).
53. Long, H., LIU, H. & Xu, J. H. Development of a PCR Diagnostic for the Root-knot Nematode *Meloidogyne enterolobii*. *Acta Phytopathol. Sinica* **2**, 109–115 (2006).
54. Lu, F. *et al.* Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genetics* **9**, e1003215, <https://doi.org/10.1371/journal.pgen.1003215> (2013).
55. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

56. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. arXiv: preprint arXiv:1207.3907 [q-bio.GN] (2012).
57. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w¹¹¹⁸; iso-2; iso-3. *Fly* **6**, 80–92 (2012).
58. Conesa, A. *et al.* Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).

Acknowledgements

This work is based on the research supported by the National Research Foundation (NRF) under Grant SUR20110707000020249. Any opinion, finding and conclusion or recommendation expressed in this material is that of the author(s) and the NRF does not accept any liability in this regard.

Author Contributions

M.R. designed the project, established single-egg mass *Meloidogyne* populations and extracted the DNA of the populations studied, H.F. and M.M. supervised the project. M.R. and M.S.D. performed sampling, P.Y.V. and B.M. analysed GBS results. M.R. wrote the manuscript and all co-authors revised it.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-31963-9>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018