Farm Bill Survey Report

Year:	2017-19
State:	California
Cooperative Agreement Name:	Enhancing diagnostics of cyst forming nematodes
Cooperative Agreement Number:	17-0426-000-FR / AP17PPQFO000C482
Project Funding Period:	20 September 2017-19 March 2019
Project Report:	Annual Accomplishment Report
Project Document Date:	15 June 2019
Cooperators Project Coordinator:	Dr. Sergei A. Subbotin
Name:	California Department of Food & Agriculture
Agency:	Plant Pest Diagnostics Center
Address:	3294 Meadowview Road
City/ Address/ Zip:	Sacramento, CA 95832
Telephone:	916-738-6696
E-mail:	Sergei.subbotin@cdfa.ca.gov

Quarterly Report	
Semi-Annual Accomplishment Report	
Annual Accomplishment Report	X

A. Write a brief narrative of work accomplished. Compare actual accomplishments to objectives established as indicated in the work plan. When the output can be quantified, a computation of cost per unit is required when useful.*

The objective of the project is to generate high-quality sequencing data for potential diagnostic gene markers for quarantine-important groups of the genus *Globodera* and clarify taxonomic status of species complexes of *G. pallida* and *G. tabacum*.

Summary - The genus *Globodera* presently contains thirteen valid and three still undescribed species. Three species, *G. rostochiensis, G. pallida* and *G. ellingtonae* named as potato cyst nematodes (PCNs) cause significant economic losses on potatoes around the world. In our study we provided comprehensive phylogenetic analyses of 384 ITS rRNA, 219 *COI* and 164 *cytb* gene sequences of 11 valid and two undescribed species from the genus *Globodera* using Bayesian inference, maximum likelihood and statistical parsimony. New 208 *COI*, 98 *cytb* and 14 ITS rRNA gene sequences were obtained from 149 populations of these species collected from 23 countries. The phylogenetic analysis revealed that the genus *Globodera* displayed two main clades in trees: *i)* *Globodera* from South and North America parasitizing plants from Solanaceae and *ii)* *Globodera* from Africa, Europe, Asia and New Zealand parasitizing plants from Asteraceae and other families. Based on the results of phylogeographical analysis and age estimation of clades with a molecular clock approach, it was hypothesized that the *Globodera* species originated and diversified from several centres of speciation located in mountain regions and dispersed from these regions across the world during the Pleistocene. High genetic diversity of Bolivian populations was observed for both mtDNA genes. Analysis of phylogenetic relationships of *G. pallida* and *G. rostochiensis* populations revealed incongruence in topology between trees inferred from mtDNA genes, which might be an indication of possible recombination and selective introgression events through gene flow between previously isolated populations. It puts some limitations on the use of the mtDNA marker as universal DNA barcoding identifier for PCNs.

Presently, the genus *Globodera* Skarbilovich, 1959 contains 13 valid species: *G. agulhasensis* Knoetze, Swart, Wentzel & Tiedt, 2017, *G. artemisiae* (Eroshenko & Kazachenko, 1972) Behrens, 1975, *G. capensis* Knoetze, Swart & Tiedt, 2013, *G. ellingtonae* Handoo, Carta, Skantar & Chitwood, 2012, *G. leptonepia* (Cobb & Taylor, 1953) Skarbilovich, 1959, *G. mali* (Kirjanova & Borisenko, 1975) Behrens, 1975, *G. mexicana* Subbotin, Mundo-Ocampo & Baldwin, 2010, *G. millefolii* (Kirjanova & Krall, 1965) Behrens, 1975, *G. pallida* Stone, 1973, *G. rostochiensis* (Wollenweber, 1923) Skarbilovich, 1959, *G. sandveldensis* Knoetze, Swart, Wentzel & Tiedt, 2017, *G. tabacum* (Lownsbery & Lownsbery, 1954) Skarbilovich, 1959 and *G. zelandica* Wouts, 1984 and three undescribed putatively new species, one from Portugal and two species from New Zealand (Subbotin *et al.,* 2010, 2011, Handoo & Subbotin, 2018). The species *G. bravoae* Franco, Cid del Prado & Lamothe-Argumedo, 2000 was described from *Jaltomata procumbens* in Mexico and synonymized with *G. mexicana* in the present publication. Of thirteen *Globodera* species, only the three potato cyst nematodes (PCNs) *G. rostochiensis, G. pallida* and *G. ellingtonae* and the tobacco cyst nematode *G. tabacum* (TCN) are important for agriculture. The golden potato cyst nematode *G. rostochiensis* and the pale potato cyst nematode *G. pallida* are serious pests of potatoes around the world and are targets of strict regulatory actions in many countries.

Identification of the cyst nematodes are essential for the selection of appropriate control measures. Molecular diagnostics using ITS-rRNA and mtDNA gene markers provides an effective solution for

diagnostics of this nematode group. PCR-ITS-RFLP, conventional PCR and Real-Time PCR with species-specific primers and ITS-rRNA and mtDNA gene sequence analysis have been developed and are now becoming useful tools for identification of *Globodera* species (Subbotin *et al.*, 2010). The analyses of South American populations of *G. pallida* revealed high levels of genetic diversity among them (Picard *et al.*, 2007, 2008) and, consequently, Grenier *et al.* (2010) noted the importance of evaluating available molecular diagnostic tools in relation to the new understanding of the wider diversity in South American populations. It has been shown that the populations of PCNs from South America have a different genetic composition from those in Europe. Franco and Evans (1978) showed that most of the South American populations were able to multiply on the potato cultivars containing resistance genes. The introduction of such populations would pose a threat to the use of resistant cultivars as a major tool in reducing the potential spread and damage caused by these species. Analysis of genetic variability and virulence characteristics of a specific nematode population will help to identify the virulence status of intercepted populations for inspection purposes and strengthens the case for using plant health legislation to prevent their introduction (Hockland *et al.*, 2012).

The availability of molecular data for the genus *Globodera* had a significant impact on the systematics of this group, reshaping concepts of its species relationships and origins. Distinct genetic differences among *G. pallida* populations spanning Europe and other regions and those found in South America were revealed by sequences and phylogenetic analyses of the ITS-rRNA (Blok *et al.*, 1998; Subbotin *et al.*, 2000, 2011; Madani *et al.*, 2010; Skantar *et al.*, 2011; Hoolahan et al., 2012b), *cytb* (Picard *et al.*, 2007, 2008; Plantard *et al.*, 2008; Pylypenko *et al.*, 2008; Madani *et al.*, 2010; Geric Stare *et al.*, 2013) and *COI* (Chitambo *et al.*, 2019) genes. The analysis of partial *cytb* gene sequences and microsatellites of *G. pallida* collected in different regions allowed the identification of the origin of western European populations with a high degree of certainty (Picard *et al.*, 2007, 2008; Plantard *et al.* 2008). These analyses showed that all these populations originated from a single restricted area in the extreme south of Peru, located between the north shore of the Lake Titicaca and Cusco. Plantard *et al.* (2008) found that only four *cytb* haplotypes are reported in western Europe, one of them also being found in some populations of this area of southern Peru. Boucher *et al.* (2013) published a study on genetic diversity of *G. rostochiensis* using twelve new microsatellite markers and concluded that the Bolivian populations showed a higher genetic diversity than those originating from Europe and North America and that this PCN species was introduced in Europe from either Peru or Chile. Hockland *et al.* (2012) already noticed that the data on populations from South America was incomplete because: *i*) for *G. pallida* the actual Peruvian phylogeographic pattern observed had to be extended both southward in Peru and northward to Ecuador and Colombia; and *ii*) for *G. rostochiensis* similar investigations needed to be done to reveal the phylogeographic pattern of this species in the Andean region, in countries such as Bolivia, Chile and Argentina.

In the present study, we provide evidence for high diversity of *cytb* and *COI* haplotypes;a phylogeographic pattern of *G. rostochiensis* in Bolivia, as well as new mtDNA sequence data for other *Globodera* species, using the same approach, as applied in the research on DNA barcoding, phylogeny and phylogeography

of the cyst nematodes from the *Avenae* group of the genus *Heterodera* (Subbotin *et al.,* 2018). For the study of relationships, network analysis was used, since Bandelt *et al.* (1995) noticed the network comparing with tree has many advantages. Labeled appropriately, the network can predict haplotypes, indicate most mutated sites and reveal sites where recombination and sequence errors are likely to have occurred. Since the network harbors most trees for the input data, it yields a more concise picture of relationships. The main goals of our study were to: obtain high-quality sequencing data for potential diagnostic gene markers for quarantine-important groups of the genus *Globodera* and clarify taxonomic status of species complexes of *G. pallida* and *G. tabacum*. Other goals were to: *i*) analyse phylogenetic relationships within *Globodera* species using sequences of the ITS rRNA, the partial *COI* and *cytb* genes; *ii*) study genetic diversity of species and populations of *Globodera* using sequences of the partial *COI* and *cytb* genes; *iii*) conduct phylogeographical study for *Globodera* species and analyse their haplotype distribution and *iv*) propose and test the hypotheses of the origin and dispersal of the *Globodera* species.

Materials and methods

Nematode species and populations

Species and populations from different hosts, localities and continents used in this study are listed in Table 1. A total of 149 nematode populations collected in 23 countries were analysed in this study. Of the 13 valid species belonging to the genus *Globodera,* 11 species were studied. Two other valid species *G. mali* and *G. leptonepia* were not included in this study because of lack of live specimens. Cysts were extracted from soil samples using standard flotation and sieving techniques. In the present project, only dead *Globodera* samples in fixed in 70-90% ethanol were used. Delimiting of species of the studied populations was accomplished by integrating results of morphological and morphometric studies, phylogenetic and sequence analysis, as well as analysis of nematode host-plant specificity and geographic distribution of studied samples.

DNA Extraction, PCR and sequencing

DNA was extracted from 50 or more J2 and embryonated eggs released from single cysts. DNA extraction, PCR and sequencing were performed as described by Subbotin *et al.* (2018). Several primer sets were used in the present study: the forward Het-coxiF (5'- TAG TTG ATC GTA ATT TTA ATG G – 3') and the reverse Het-coxiR (5' - CCT AAA ACA TAA TGA AAA TGW GC – 3') primers as described by Subbotin (2015) for amplification of the partial *COI* gene, the forward Het-cytbF2 (5' - CAR TAT TTR ATR TTT GAR GT – 3') and reverse Het-cytbR3 (5' - ACH ARR AAR TTR ATY TCC TC – 3') primers (this study) for amplification of the partial *cytb* gene, the forward TW81 (5' - GTT TCC GTA GGT GAA CCT GC – 3') and the reverse AB28 (5' - ATA TGC TTA AGT TCA GCG GGT – 3') or the reverse primer 5.8SM5 (5' – GGC

GCA ATG TGC ATT CGA -3') as described by Zheng ***et** al.* (2000) for amplification of the ITS1-5.8S-ITS2 rRNA or ITS1 rRNA gene. The PCR products of one sample was cloned into the pGEM-T vector and transformed into JM109 High Efficiency Competent Cells (Promega). Sequencing was performed by Quintara Biosciences (San Diego, USA). New sequences will be deposited in the GenBank database under accession numbers.

Phylogenetic, sequence and phylogeographic analysis

Alignments with the ITS rRNA, *COI* and *cytb* gene sequences were created using ClustalX 1.83 (Chenna *et al.*, 2003) with default parameters. New sequences were aligned with corresponding published gene sequences (Ferris *et al.*, 1995, 1999; Blok *et al.*, 1998; Subbotin *et al.*, 2000, 2001; 2011; Manduric & Andersson, 2004; Sirca & Urek, 2004; Picard *et al.*, 2007, 2008; Plantard *et al.*, 2008; Pylypenko *et al.*, 2008; Grenier *et al.*, 2010; Madani *et al.*, 2010; Skantar *et al.*, 2011; Geric Stare *et al.*, 2013; Lax *et al.*, 2014; Chitambo *et al.*, 2019 and others). Several alignments were created: *i*) ITS rRNA gene alignment containing only reference sequences of each *Globodera* species; *ii*) ITS rRNA gene alignment containing all available sequences of the genus *Globodera*; *iii*) *COI* gene alignment containing sequences of all studied samples; *iv*) *COI* gene sequence alignment containing reference haplotype sequences for all *Globodera* species; *vi*) *cytb* gene alignment containing sequences of all studied samples; *vii*) *cytb* gene sequence alignment containing reference haplotype sequences for all *Globodera* species; and *vii*) several *COI* and *cytb* gene alignments containing sequences of certain species. Sequence alignments were also manually edited using GeneDoc 2.5.0 (Nicholas *et al.*, 1997). Pairwise divergence between taxa was calculated as the absolute distance value and the percent of mean distance, with adjustment for missing data, using PAUP* 4b10 (Swofford, 2003).

The ITS rRNA, *COI* and *cytb* gene sequence alignments were analysed with maximum likelihood (ML), maximum parsimony (MP) using PAUP* or Bayesian inference (BI) using MrBayes 3.1.2 (Ronquist & Huelsenbeck, 2003) as described by Subbotin *et al.* (2018). The best fit models of DNA evolution were obtained using the program jModeltest.0.1.1 (Posada, 2008) with the Akaike Information Criterion. Bootstrap support (BS) values for ML and MP trees were calculated by a heuristic search from 1000 replicates.

The alignments for ITS rRNA, *COI* and *cytb* gene sequences were used to construct phylogenetic network estimation using statistical parsimony (SP) as implemented in POPART software (http://popart.otago.ac.nz) (Bandelt *et al.*, 1999).

Testing for Global Molecular Clock with the *COI* gene sequence alignment was performed with PAUP* under the null hypothesis – the phylogeny was rooted and the branch lengths were constrained such that all of the tips can be drawn at a single time plane. Under the alternative hypothesis, each branch was allowed to vary independently. The Hasegawa, Kishino & Yano (Hasegawa *et al.*, 1985) (HKY85) model of DNA substitution with among site rate variation with a gamma distribution was used.

Divergence times were inferred with BEAST 2.4.5 (Bouckaert *et al.*, 2014) using the *COI* sequence alignment. The tree prior a lognormal relaxed clock with uncorrelated rates were assigned to the Yule model with the mitochondrial substitution genome rate equal to 7.2×10^{-8} per site per generation as calculated by Howe *et al.* (2010) for *Caenorhabditis briggsae*. The life cycle with one generation per year was considered for *Globodera* species (Subbotin *et al.*, 2010). The analyses were run for as described by Subbotin *et al.* (2018).

Ancestral area reconstruction was done with RASP 4.1.2 (Yu *et al.*, 2015), which implement Statistical-Dispersal Vicariance Analysis (S-DIVA). Enhanced S-DIVA method reconstructs the frequencies of ancestral distribution at each node by utilizing all posterior distribution to account for phylogenetic uncertainty and DIVA optimization. The distribution of *Globodera* species was divided into 9 geographical regions (A-I), where these species were naturally found and/or have unique haplotypes: A - South Africa; B - Bolivia and Southern Andes; C - Peru and Northern Andes; D - Mexico; E - Europe; F - North America (USA, Canada); G - East Asia; H - New Zealand and I - Kenya. The number of maximum areas for each node was kept at three. Most likely ancestral regions for each node were mapped on the 50% majority rule consensus BI tree inferred from the analysis of the *COI* sequence alignment.

Results

Species delimiting

The results of the present study suggest that the ITS rRNA gene, *COI* and *cytb* markers are able to differentiate eleven *Globodera* species: *G. agulhasensis, G. artemisiae, G. capensis, G. ellingtonae, G. mexicana, G. millefolii, G. pallida, G. rostochiensis, G. sandveldensis, G. tabacum* and *G. zelandica* from each other. In this study, we synonymize *Globodera bravoae* syn. n. with *G. mexicana* based on molecular and morphological characters.

Phylogenetic and sequence analysis with ITS rRNA gene

Two approaches were applied to analyse the ITS rRNA gene sequences of the *Globodera* species. The first approach included BI, ML and MP analyses of ITS rRNA gene alignment containing only reference sequences of each *Globodera* species and the second approach included SP analysis of two alignments of ITS rRNA gene sequences for two major *Globodera* clades.

Globodera *phylogeny*

The phylogenetic analysis revealed a well-supported topology whereby species generally clustered according to their geographic proximity and host-range. Phylogenetic relationships within 11 valid and three

putative new species as inferred from BI, ML and MP analyses of the ITS rRNA gene sequences are given in Figure 1. *Globodera* species were distributed within two major clades: the solanaceous *Globodera* clade (South America origin) or clade I and the non-solanaceous *Globodera* clade (non-South America origin) or clade II.

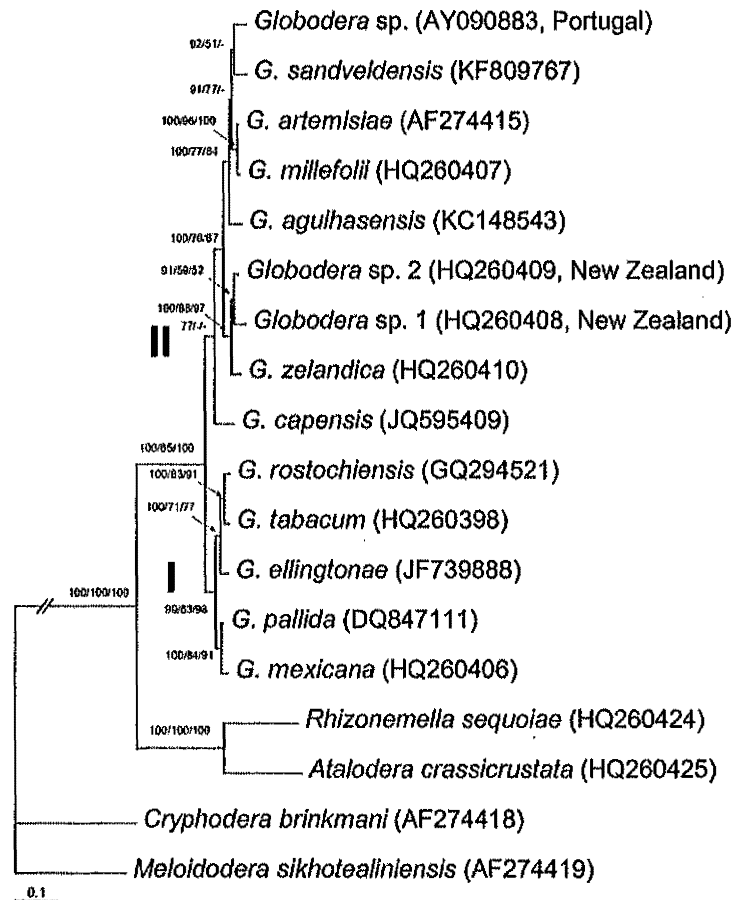


Figure 1. Phylogenetic relationships between *Globodera* species as inferred from Bayesian analysis of the ITS rRNA gene sequences. Posterior probability and bootstrap support values for BI, ML and MP analysis are given for appropriate clades, respectively. Values less than 50% are not indicated.

The solanaceous Globodera *clade*

A total of 353 sequences were used in the analysis. They included 339 sequences downloaded from GenBank, 4 new sequences of *G. rostochiensis*, 9 new sequences of *G. pallida* and one new sequence of *G. mexicana* (= *G. bravoae* syn. n.). The alignment was 977 bp in length and included 128 sequences of *G. pallida*, 113 sequences of *G. rostochiensis*, 73 sequences of *G. ellingtonae*, 33 sequences of *G. tabacum* and 6 sequences of *G. mexicana* (= *G. bravoae* syn. n.). The ITS rRNA gene sequences were represented by following haplotype numbers: *G. pallida* - 22 haplotypes including Hpal1 with 101 sequences, *G.*

rostochiensis – 14 haplotypes including Hros1 with 100 sequences, *G. ellingtonae* – 26 haplotypes including Hell1 with 29 sequences, *G. tabacum* – 11 haplotypes and *G. mexicana* – 3 haplotypes.

A phylogenetic network for the ITS rRNA sequences, reconstructed using SP with POPART software, is given in Figure 2. This method allowed clear discrimination of all studied species. Maximum sequence divergence for species of the solanaceous *Globodera* clade was 6.2%, *G. pallida* – 3.4%; *G. ellingtonae* – 2.4%, *G. rostochiensis* – 2.2%, *G. tabacum* -1.2%, *G. mexicana* – 0.5%. The ITS rRNA gene sequence of the sample identified as *G. bravoae* was different in one nucleotide and two deletions/insertions with those of *G. mexicana* under the GenBank accession number EU006707.

Figure 2. Statistical parsimony network showing the phylogenetic relationships between ITS rRNA gene haplotypes of the solanaceous *Globodera* species. The sequences of each species are marked by different colours. Pies (circles) represent sequences of each species with the same haplotype and their size is proportional to the number of these sequences in the samples. Numbers of nucleotide differences between the sequences are indicated on lines connecting the pies. Small black circles represent missing haplotypes. Newly obtained sequences are given with sample codes.

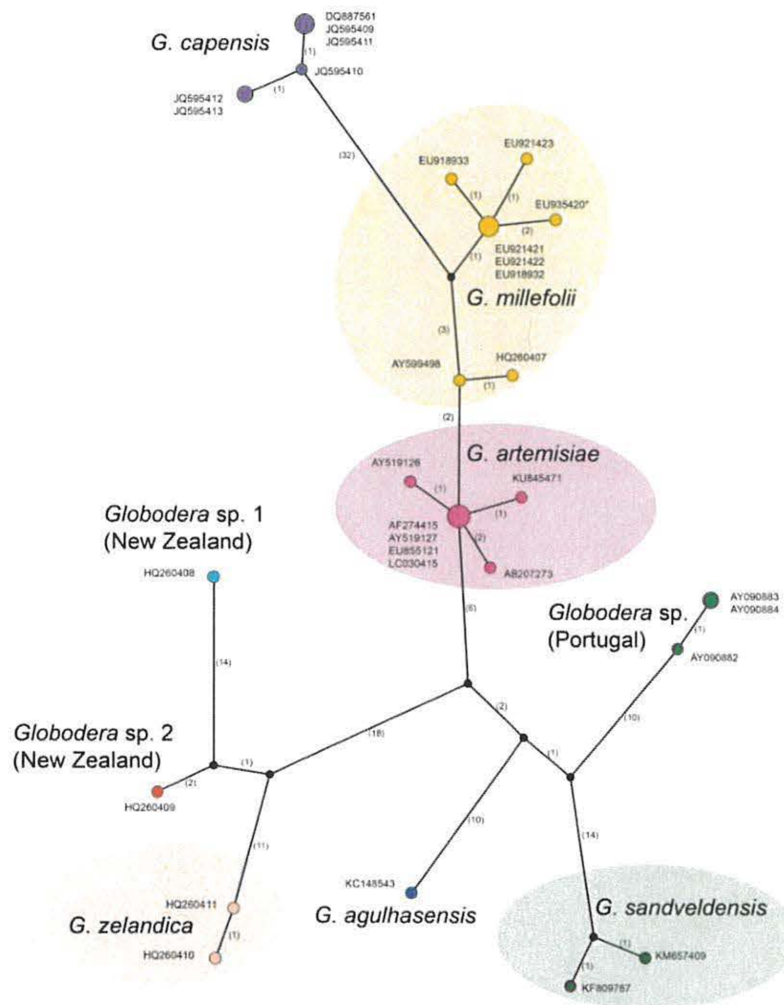


Figure 3. Statistical parsimony network showing the phylogenetic relationships between ITS rRNA gene haplotypes of the non-solanaceous *Globodera* species. The sequences of each species are marked by different colours. Pies (circles) represent sequences of each species with the same haplotype and their size is proportional to the number of these sequences in the samples. Numbers of nucleotide differences between the sequences are indicated on lines connecting the pies. Small black circles represent missing haplotypes. * - identified as. *G. artemisiae* in the GenBank.

The non-solanaceous Globodera *clade*

A total of 31 sequences downloaded from the GenBank was included in the analysis. The alignment was 943 bp in length and contained 8 sequences of *G. millefolii*, 7 sequences of *G. artemisiae*, 6 sequences of *G. capensis*, 2 sequences of *G. sandveldensis*, 2 sequences of *G. zelandica*, 3 sequences of *Globodera* sp. from Portugal, one sequence for *G. agulhasensis*, one sequence for *Globodera* sp. 1, one sequence for *Globodera* sp. 2 from New Zealand. A phylogenetic network for the ITS rRNA sequences, reconstructed

using SP with POPART software, is given in Figure 3. Maximum sequence divergence for species of the non-solanaceous *Globodera* clade was 9.7%.

PHYLOGENETIC AND SEQUENCE ANALYSIS WITH *COI* GENE

Two approaches were applied to analyse the *COI* gene sequences of the *Globodera* species. The first approach included BI and ML analyses of the *COI* gene alignment containing only reference haplotype sequences of species, whilst the second approach included SP analyses of alignments of *COI* sequences for *Globodera rostochiensis, G. pallida, G. mexicana* and *G. tabacum*. A total of 219 *COI* gene *Globodera* sequences, of which 208 were new, was included in this study. The *COI* gene alignment was 443 bp in length.

Globodera *phylogeny*

Phylogenetic relationships within the *Globodera* species, containing 63 reference haplotype sequences and three sequences of outgroup species as inferred from BI and ML analyses, are given in Figure 4. The analysis revealed the *Globodera* species are divided into two major clades: the solanaceous *Globodera* clade or clade I and the non-solanaceous *Globodera* clade or clade II. Maximum *COI* gene sequence divergence for *Globodera* species was 26.6%.

The solanaceous Globodera *clade*

Maximum *COI* gene sequence divergence for the solanaceous *Globodera* clade species was 25.7%
Globodera rostochiensis. A total of 110 sequences of this species were analysed. The haplotype network is given in Figure 5A. Only one haplotype was found in majority populations, however, several populations contained two haplotypes. A total of 29 haplotypes were revealed. All haplotypes were divided into four groups: A, B, C and D. The group A contained 16 haplotypes (CrCOIA1-CrCOIA16). The haplotype CrCOIA1 was mostly distributed across the world (Table 1) and differed from CrCOIA2 from Kenia by one mutation. Only these two haplotypes were found outside of Bolivia. All other groups, the group B with 3 haplotypes (CrCOIB1-CrCOIB3), the group C with 8 haplotypes (CrCOIC1-CrCOIC8), the group D with one haplotype (CrCOID1) were reported from Bolivia (Fig. 6). Maximal intraspecific *COI* gene sequence diversity was 14.2%.

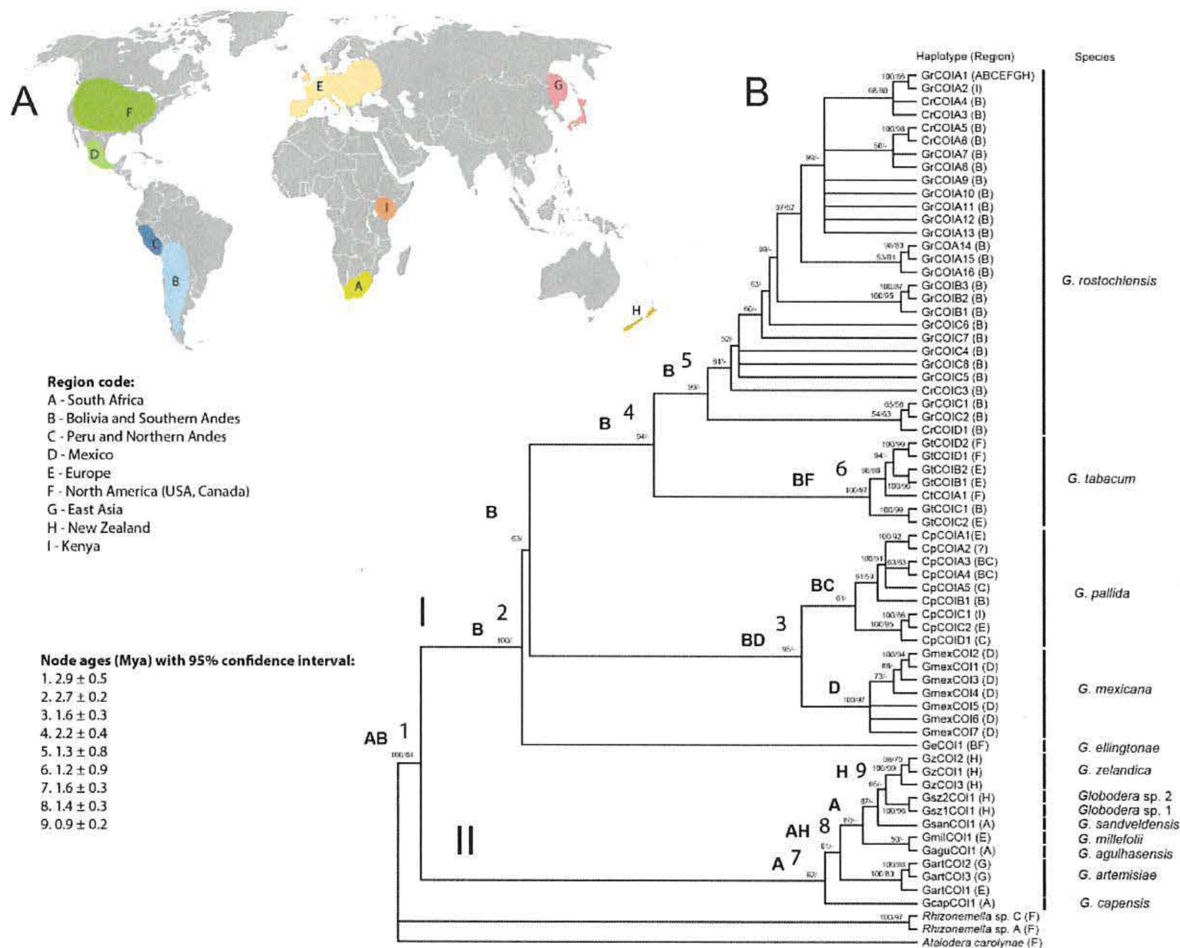


Figure 4. Phylogenetic relationships between *COI* haplotypes of the *Globodera* species as inferred from Bayesian analysis with mapping of regions and indication of node ages. A: World map with regions, where *Globodera* species are naturally distributed and/or have unique haplotypes; B: Phylogenetic tree. Codes with most probable ancestral regions for key nodes obtained from Statistical Dispersal-Vicariance Analysis, posterior probability values for BI analysis and bootstrap values for ML analysis are given to appropriate clades.

Globodera pallida. A total of 43 sequences of this species was analysed. The haplotype network is given in Figure 7A. Only one haplotype was found in majority populations, however, one population contained two haplotypes. A total of 9 haplotypes were revealed. All haplotypes were divided into four groups: A, B, C and D. The group A contained 5 haplotypes (CpCOIA1-CpCOIA5), whereas groups B and D consisted by one haplotype and the group C had two haplotypes. The haplotype CpCOIA1 was reported from several European countries and the USA only (Table 1). The haplotypes from the group C were reported from UK and Kenya. The haplotypes CpCOIA3- CpCOIA5 and the haplotypes from the group B and D were found in South America (Fig. 6). Maximal intraspecific *COI* gene sequence diversity was 20.7%.

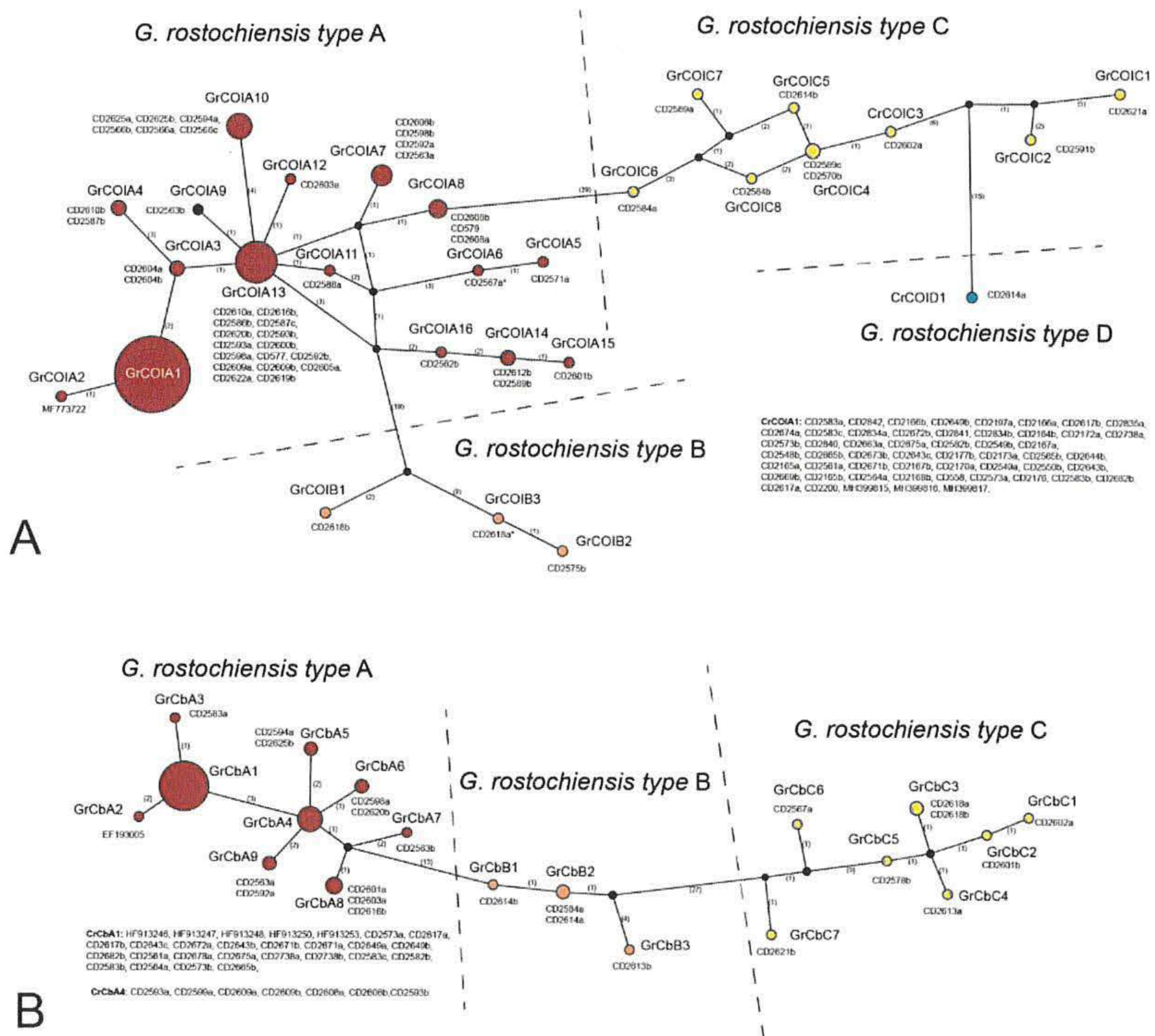


Figure 5. Statistical parsimony network showing the phylogenetic relationships between haplotypes of *Globodera rostochiensis*. Small black circles represent missing haplotypes. Pie chart sizes are proportional to the number of samples with a particular haplotype; A: *COI* gene; B: *cytb* gene. Newly obtained sequences are given with sample codes.

Globodera mexicana. A total of 16 sequences of *G. mexicana* (= *G. bravoae* syn. n.) collected in 8 locations was included in the analysis. The haplotype phylogenetic network is given in Figure 8. Only one haplotype was found in majority populations, however, several populations contained two haplotypes and one population contained three haplotypes. Seven haplotypes (GmCOI1-GmCOI7) were revealed among these sequences. Maximum sequence diversity was 4.7%. The *COI* gene sequence of the sample identified as *G. bravoae* was similar with those of several *G. mexicana* samples (Fig. 8).

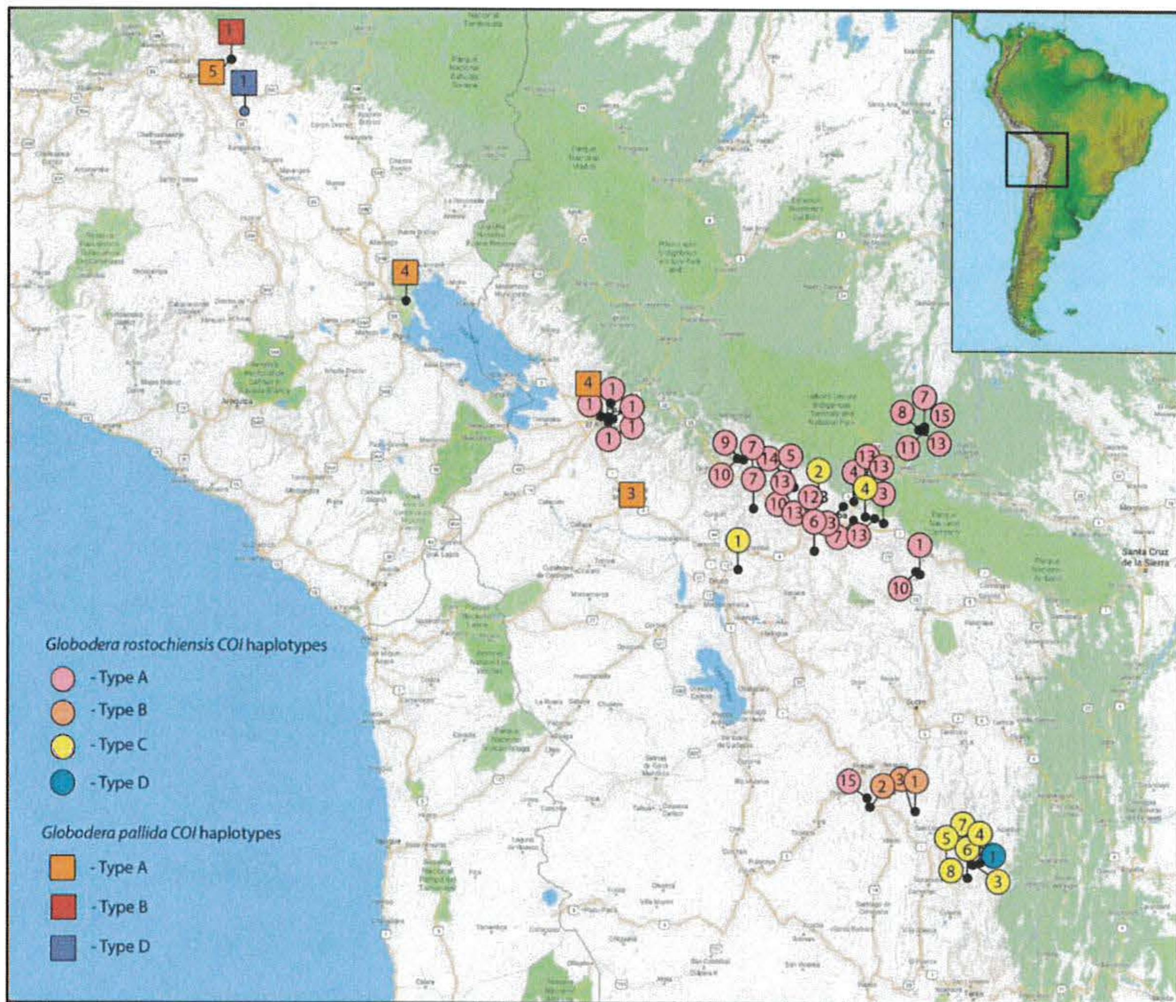
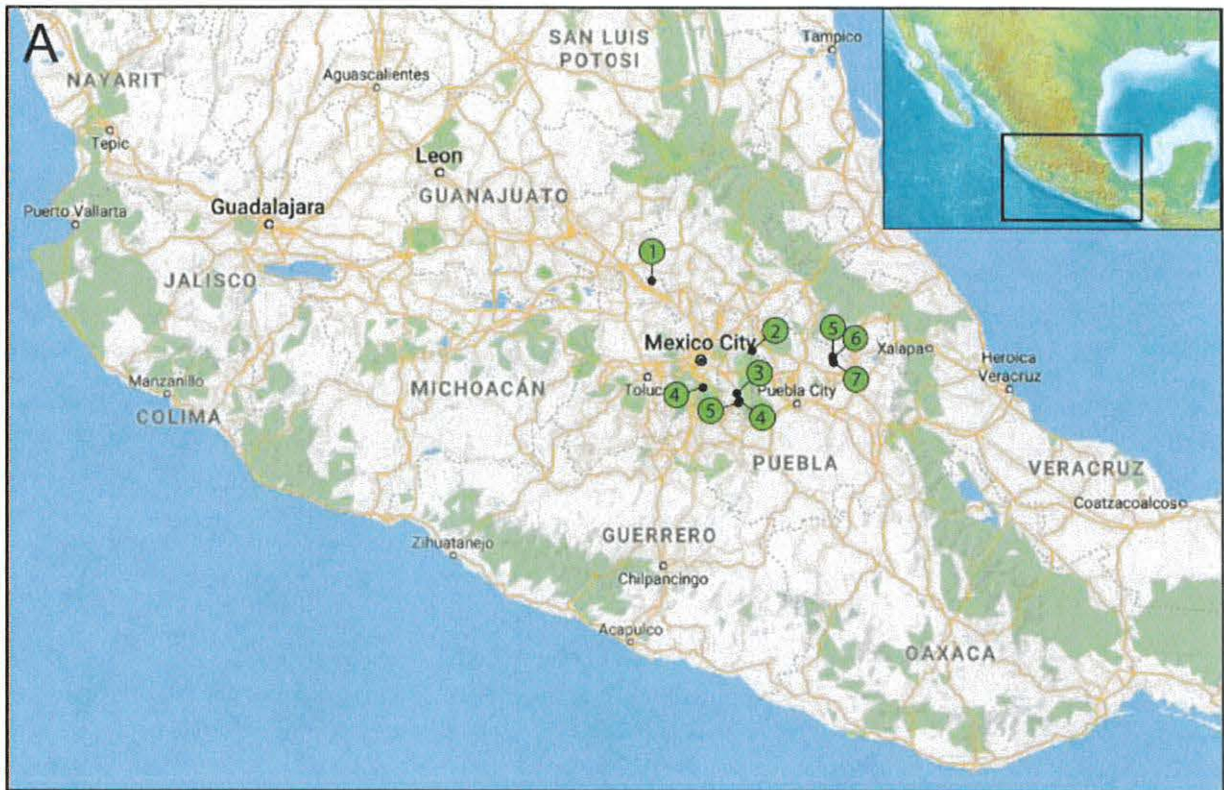


Figure 6. Map of the Andes with distribution of *COI* haplotypes for *Globodera rostochiensis* and *G. pallida* obtained from original samples.

Globodera tabacum. A total of 33 sequences of this species was analysed. The haplotype network is given in Figure 9A. Only one haplotype was found in majority populations, however, one population contained two haplotypes. A total of 7 haplotypes was revealed. All haplotypes were divided into four groups: A, B, C and D. The haplotypes from the group A reported from Virginia, USA and Bolivia, the group B was from France, the group C was from Italy and Argentina and the group D was from Virginia, USA. Maximal intraspecific *COI* gene sequence diversity was 14.1%.

Globodera ellingtonae. One new sequence and one published sequence were included in the study and these sequences were identical.

Figure 7. Statistical parsimony network showing the phylogenetic relationships between haplotypes of *Globodera pallida*. Small black circles represent missing haplotypes. Pie chart sizes are proportional to the number of samples with a particular haplotype. A: *COI* gene; B: *cytb* gene. Newly obtained sequences are given with sample codes.



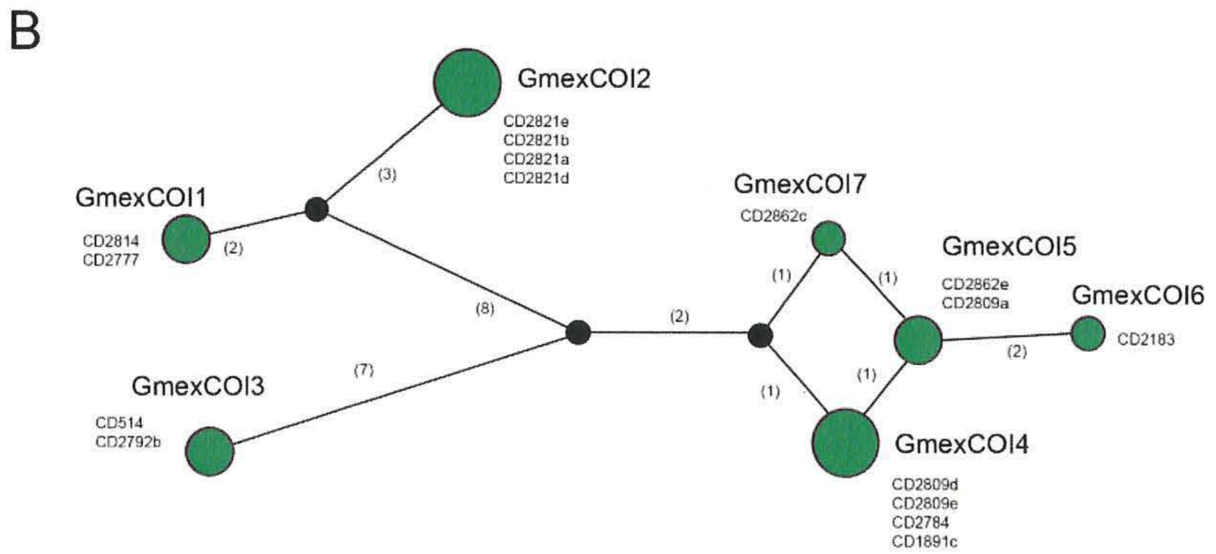


Figure 8. A. Map of the Andes with distribution of *COI* haplotypes for *Globodera mexicana* obtained from original samples; B: Statistical parsimony network showing the phylogenetic relationships between *COI* haplotypes of *G. mexicana*. Small black circles represent missing haplotypes. Pie chart sizes are proportional to the number of samples with a particular haplotype. Newly obtained sequences are given with sample codes.

The non-solanaceous Globodera *clade*

Maximum *COI* gene sequence divergence for the non-solanaceous *Globodera* clade species was 18.9%.
Globodera zelandica. A total of 5 new sequences of this species was obtained. These sequences represented 3 haplotypes. Maximum sequence diversity was 2.2%
Globodera artemisiae. A total of 3 new sequences, each of them belonging to unique haplotype of this species was obtained for this species. Maximum sequence diversity was 2.0%
Globodera millefolii. Two identical sequences were obtained from the sample.
Globodera agulhasensis, G. capensis, G. sandveldensis Globodera sp. 1, *Globodera* sp. 2. Only one sequence was obtained from each species.

PHYLOGENETIC AND SEQUENCE ANALYSIS WITH *CYTB* GENE

Two approaches were also applied to analyse the *cytb* gene sequences of the *Globodera* species. The first approach included BI and ML analyses of the *cytb* gene alignment containing only reference haplotype sequences of species, whilst the second approach included SP analyses of alignments of *cytb* sequences for *G. rostochiensis*, *G. pallida* and *G. mexicana*.
Analysis of chromatogram files revealed some doubly picks for amplified *cytb* gene products for *G. rostochiensis* and *G. pallida*. Because significant differences in the heteroplasmy level were found between sequences generated forward and reverse primers, only the 5'> 3' DNA strand was considered. A total of 164 *cytb* gene *Globodera* sequences and 98 of them were new ones was included in this study. The *cytb* gene alignment was 452 bp in a length.

The solanaceous Globodera *clade*

Phylogenetic relationships within the *Globodera* species, containing 51 reference haplotype sequences of the Solanaceous *Globodera* species and one sequence of *G. capensis* as inferred from BI and ML analyses, are given in Figure 10. Maximum *cytb* gene sequence divergence for the solanaceous *Globodera* species was 26.3%

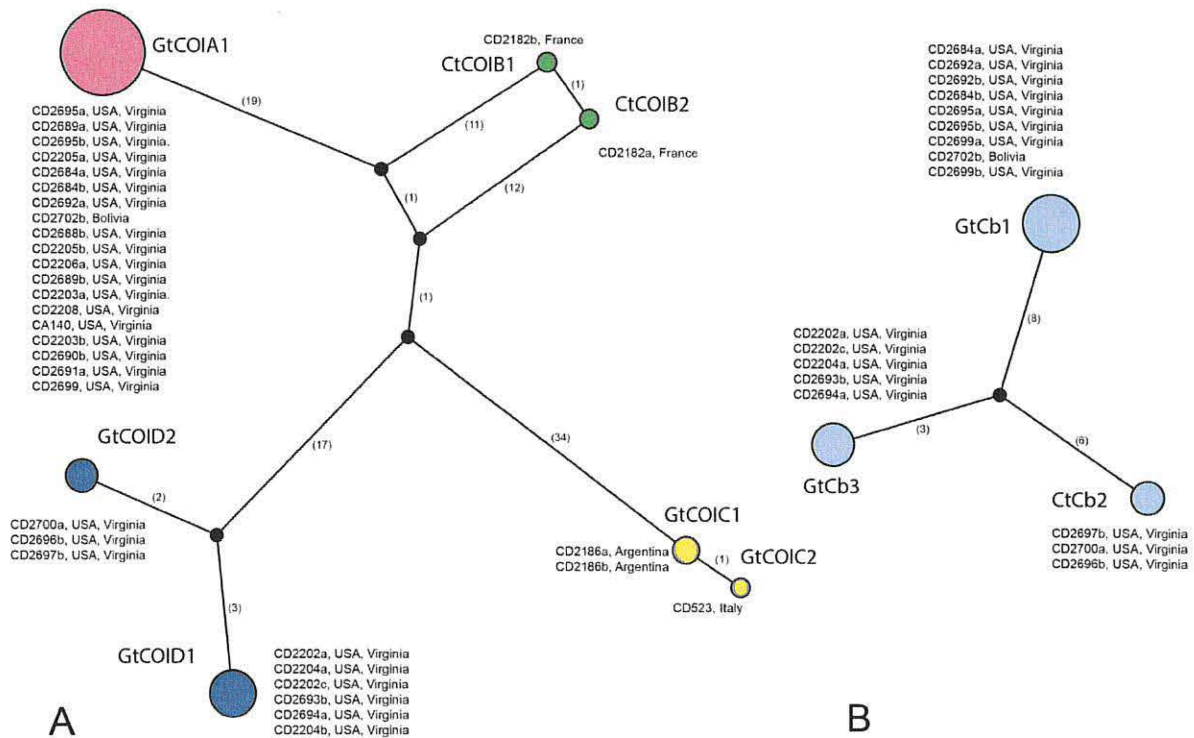


Figure 9. Statistical parsimony network showing the phylogenetic relationships between haplotypes of *Globodera tabacum*. Small black circles represent missing haplotypes. Pie chart sizes are proportional to the number of samples with a particular haplotype. A: *COI* gene; B: *cytb* gene. Newly obtained sequences are given with sample codes.

Globodera rostochiensis. A total of 58 sequences of this species was analysed. The haplotype network is given in Figure 5B. Only one haplotype was found in majority populations, however, four populations contained two haplotypes. The group A contained 9 haplotypes (CrCbA1-CrCbA9). The haplotype CrCbA1 was found in Bolivia, Venezuala, Chile, Canada, Slovenia, Serbia, Croatia and Russia (Table 1). The group B had 3 haplotypes (CrCbB1-CrCbB3) with four sequences and the group C with 7 haplotypes (CrCbC1-CrCbC7) were found in Bolivia only. Maximal intraspecific *cytb* gene sequence diversity was 9.5%.

Globodera pallida. A total of 85 sequences of this species was analysed. The haplotype network is given in Figure 7B. Only one haplotype was found in majority populations, however, two populations contained two haplotypes and one population contained three haplotypes. Thirty-two haplotypes were divided into seven groups: A, B, C, D, E, F and G. The group A contained 16 haplotypes (CpCbA1-CpCbA16). The haplotypes CpCbA1-CpCbA10 were found in Europe or Canada only. The haplotype CpCbA14 was recorded in Spain (Tenerife), UK, New Zealand and Peru. The haplotypes from the B-G groups were found in America. Maximal intraspecific *cytb* sequence diversity was 11.7%.

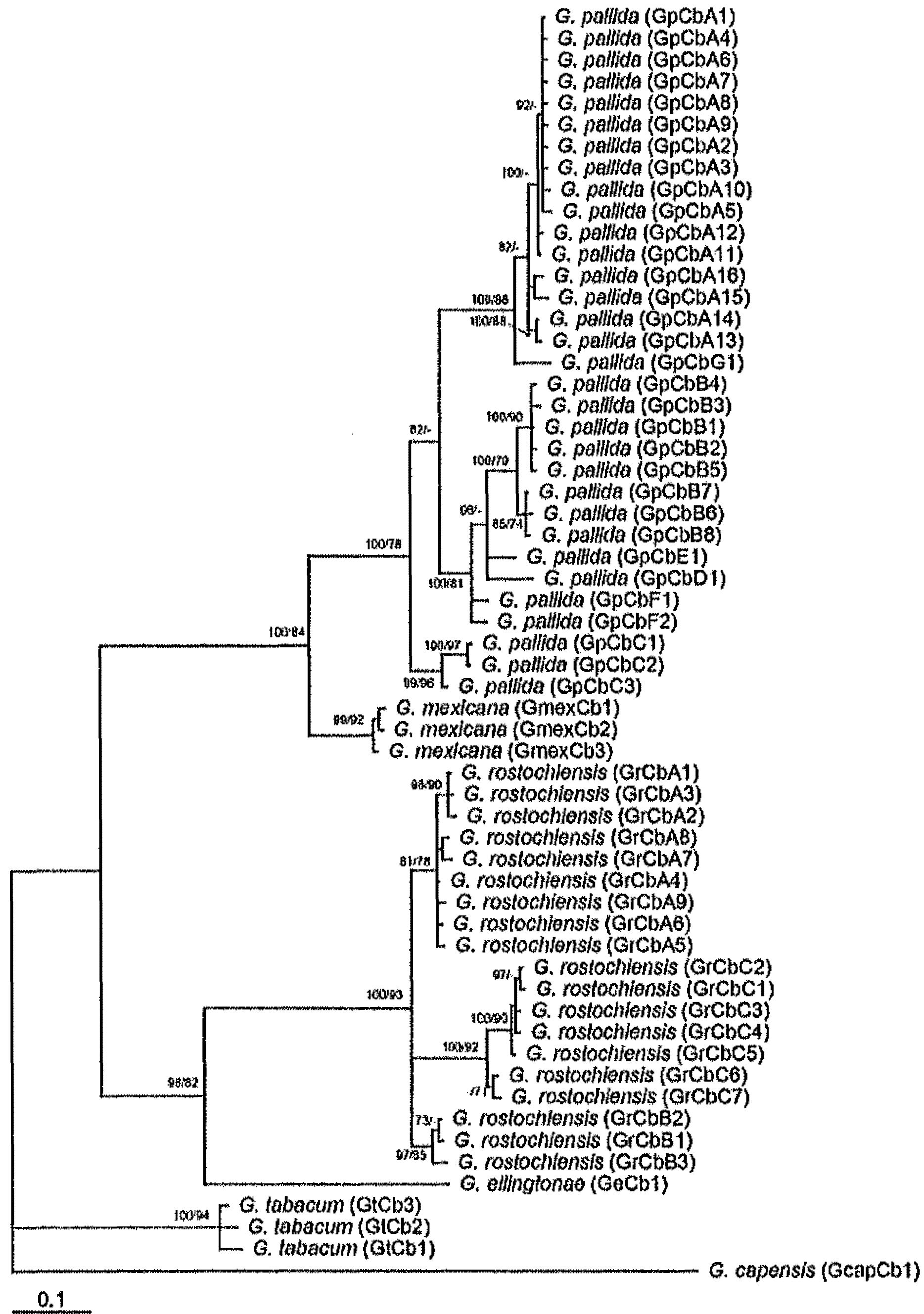


Figure 10. Phylogenetic relationships between *cytb* haplotypes of the solanaceous *Globodera* species as inferred from Bayesian analysis. Posterior probability values for BI analysis and bootstrap values for ML analysis are given to appropriate clades.

Globodera tabacum. A total of 17 sequences of this species was analysed. The haplotype network is given in Figure 9B. A total of 3 haplotypes were revealed. Maximal intraspecific *cytb* gene sequence diversity was 3.1%

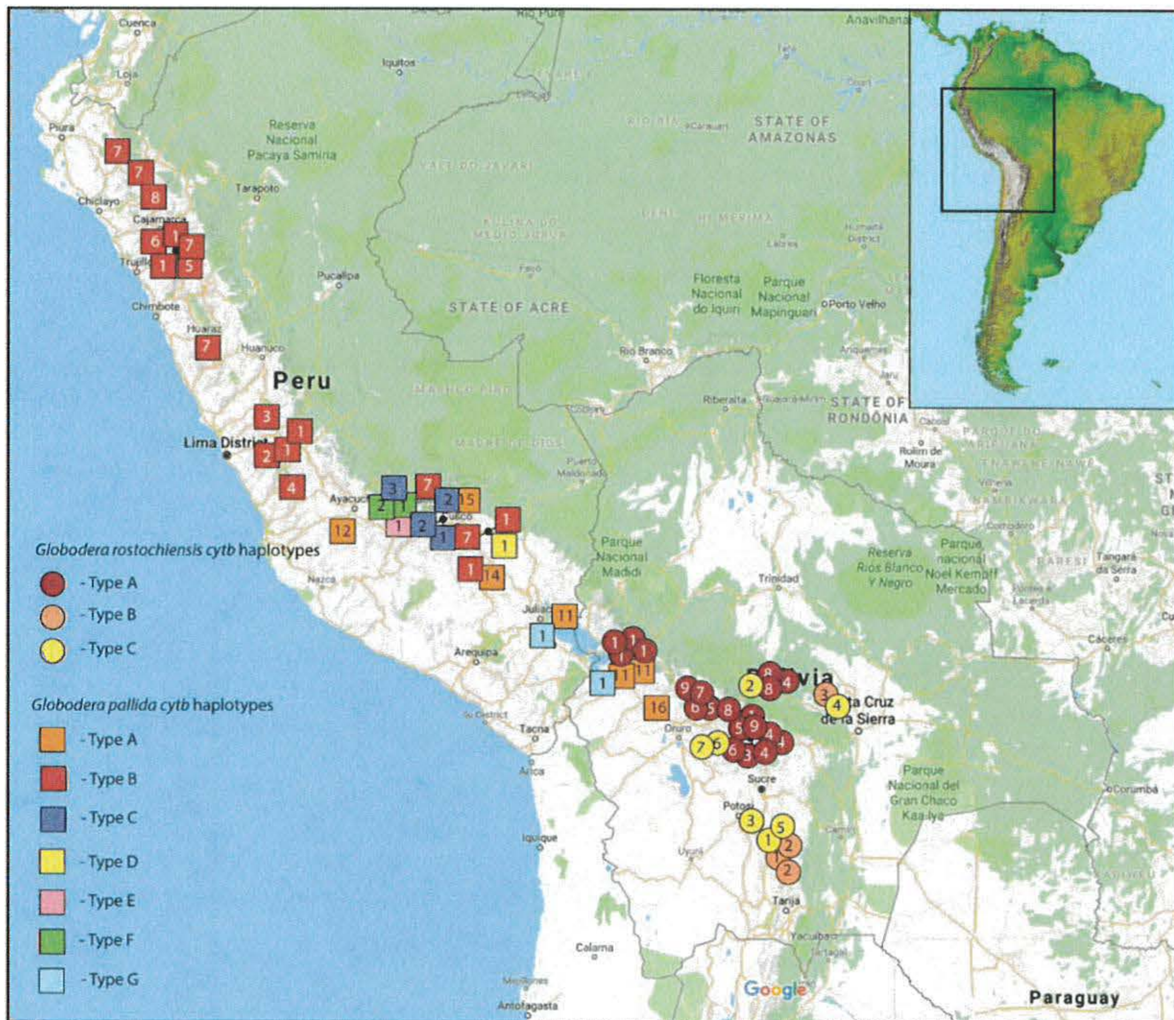


Figure 11. Map of the Andes with distribution of *cytb* haplotypes for *Globodera rostochiensis* and *G. pallida* obtained from original samples and the datasets published by Picard *et al.* (2007) and Plantard *et al.* (2008).

Globodera mexicana. Two new and one published sequences representing three haplotypes were included in the analysis. Maximal intraspecific *cytb* gene sequence diversity was 0.8%.

Globodera ellingtonae. Cytb gene sequence was obtained from the GenBank (KU726971) for this study.

PHYLOGEOGRAPHICAL ANALYSIS AND ESTIMATION OF DIVERGENCE TIMES

The results of phylogeographical analysis of *Globodera* using *COI* mtDNA marker are generally congruent with those based on the ITS rRNA gene, which has been already published by Subbotin *et al.* (2011). In our present study both methods for reconstruction of ancestral areas showed uncertainty in the geographic origin of the genus *Globodera*, since two areas South America and South Africa showed highest probability.

Nevertheless, stronger support was found for the geographic origin of both main clades. The solanaceous *Globodera* has a strong South America origin with colonization of Mexico and North America. The analysis also revealed a strong support for South African origin of the non-solanaceous *Globodera* and a robust reconstruction of dispersal events of several species to New Zealand and Europe-Asia during the Pleistocene.

The likelihood ratio test rejected the hypothesis of the Global Molecular Clock test at $p < 0.001$. The tree topology retrieved from BEAST software contradicted a tree yielded by MrBayes in the position of *G. ellingtonae* and *G. mexicana*. In the BEST tree, *G. ellingtonae* clustered with *G. pallida* and *G. mexicana,* which was nested within *G. pallida*. The estimated divergence times between main clades are shown in Figure 4. Molecular clock estimated that split between solanaceous and non-solanaceous lineages occurred roughly 2.9 ± 0.5 Mya (million years ago). Divergence dates of the solanaceous *Globodera* lineages started 2.7 ± 0.2 Mya, the non-solanaceous *Globodera* lineages – 1.6 ± 0.3 Mya. Dispersals of *Globodera* to Europe and New Zealand occurred 1.4± 0.3 and 0.9 ± 0.2 Mya, respectively.

Discussion

Problems with mtDNA markers for DNA barcoding and phylogeography of *Globodera*

Although many studies showed that *COI* and *cytb* genes are reliable markers for DNA barcoding, identification and phylogeographical studies of many nematode groups (Powers *et al.,* 2010; Palomares-Rius *et al.,* 2017; Subbotin *et al.,* 2018; Nguyen *et al.,* 2019), the published results (Armstrong *et al.,* 2000; Gibson *et al.,* 2007a, b; Hoolahan *et al.,* 2011, 2012a) and the present study provides evidences about heteroplasmy, introgression and recombination events of mtDNA in PCNs. The results inferred from the mtDNA gene fragments should be considered cautiously in the intraspecific studies with this nematode group. Observed differences in relationships of *G. rostochiensis* and *G. pallida* populations reconstructed from *COI* and *cytb* genes might be explained by unusual mitochondrion genome organization for these species.

The structure of metazoan mitochondrial (mt) genomes, including soybean cyst nematode, *H. glycines,* is generally highly conserved as a single compact circle (Gibson *et al.,* 2011). It has been shown that the mtDNA of *G. pallida* and *G. rostochiensis* is multipartite, comprising at least six small circular mtDNA (scmtDNA I-VI) molecules (Armstrong *et al.* 2000; Gibson *et al.* 2007a, b). The *COI* and *cytb* genes with its pseudogenes were found to be located in different scmtDNA molecules. Moreover, two functional copies of *cytb were also present in different* scmtDNA molecules (Gibson *et al.,* 2007a). The *G. ellingtonae* mt genome was found to be comprised of two circles (mtDNA I and mtDNA II) and *COI* and *cytb* genes are located in mtDNA I molecule (Phillips *et al.,* 2016). It has been shown that the different scmtDNA molecules observed at different frequencies in different populations *G. pallida* (Armstrong *et al.* 2007). Evidence for

recombination and the mosaic structure of scmtDNA IV sequences was also revealed in South American populations (Armstrong *et al.* 2007). The subgenomes can be formed by recombination between the various subgenomes (Gibson *et al.*, 2007b). Hoolahan *et al.* (2011, 2012a) also reported about past recombination events detected between a South American population and several UK populations of *G. pallida* and between two South American populations. It has been suggested that these populations may have interbred, paternal mtDNA leakage occurred, and the mtDNA of these populations subsequently recombined. However, progeny from experimental crosses of these populations of *G. pallida* had no evidence of contemporary recombination between the mtDNA of the maternal and paternal populations; i.e. resulting directly from the experimental crosses, which supported arguments that animal mtDNA recombination events are relatively rare (Hoolahan *et al.* 2012a).

SPECIES DELIMITING

The results of the present study showed that the ITS rRNA, *COI* and *cytb* gene markers are able to differentiate eleven studied *Globodera* species: *G. agulhasensis, G. artemisiae, G. capensis, G. ellingtonae, G. mexicana, G. millefolii, G. pallida, G. rostochiensis, G. sandveldensis*, *G. tabacum, G. zelandica* from each other.

Several publications suggested that the pale potato cyst nematode, *G. pallida*, might indeed represent a species complex (Subbotin *et al.*, 2011). Our present analysis of *COI* and *cytb* genes do not provide evidences that additional putative and undescribed species were among studied samples of *G. pallida*. Character of clustering of ITS rRNA, *COI* and *cytb* genes for *G. pallida* and evidences for possible mtDNA gene recombination events suggest that the populations from South America and other parts of the world identified as *G. pallida* should be considered as a single species.

TCN composed of three subspecies, *G. tabacum* subsp. *tabacum*, *G. tabacum* subsp. *virginiae* and *G. tabacum* subsp. *solanacearum,* which display similar morphology and morphometrics but differed in their ability to damage tobacco cultivars and develop on different Nicotianae species (Subbotin *et al.*, 2010). Although, the ITS rRNA gene sequences were not able to identify the different subspecies using the polymorphism revealed by this marker (Subbotin *et al.*, 2011), RAPD and AFLP analysis (Thiéry *et al.*, 1997; Marché *et al.*, 2001) and the analysis of five effectors (Alenda *et al.*, 2013), three genes coding for cell wall degrading enzymes (Pel1, Pel2 and ExpB3) and two genes coding for CLE peptides (CLE1 and CLE4) allowed the molecular discrimination of the subspecies inside the TCN complex. The present analysis using *COI* gene marker revealed a strong intraspecific structure within TCN but did not allow differentiating TCN subspecies unambiguously. More samples should be analyzed to make a conclusion about reliability of *COI* marker for diagnostics of TCN subspecies.

SEQUENCE DATA SUPPORT SYNONYMISATION OF *GLOBODERA BRAVOAE*

In this study, we synonymize *G. bravoae* syn. n. with *G. mexicana* based on molecular and morphological characters. *Globodera bravoae* was described from roots of *Jaltomata procumbens* (Solanaceae) collected in Delegación Magdalena Contreras, Mexico City, Mexico (Franco et al., 2000). The species was considered as different from all others based on some differences in cyst length and DGO of the second stage juveniles. The *Globodera* sample collected in the type locality of *G. bravoae* (Table 1) and morphologically identified as a representative of this species was used for molecular analysis. Because, phylogenetic and sequence analysis of the ITS-rRNA and *COI* gene sequences of *G. bravoae* and *G. mexicana* did not reveal any significant differences between them, and the range of morphometrical characters are overlapping, we consider *G. bravoae* syn. n. as a new synonym of *G. mexicana.*

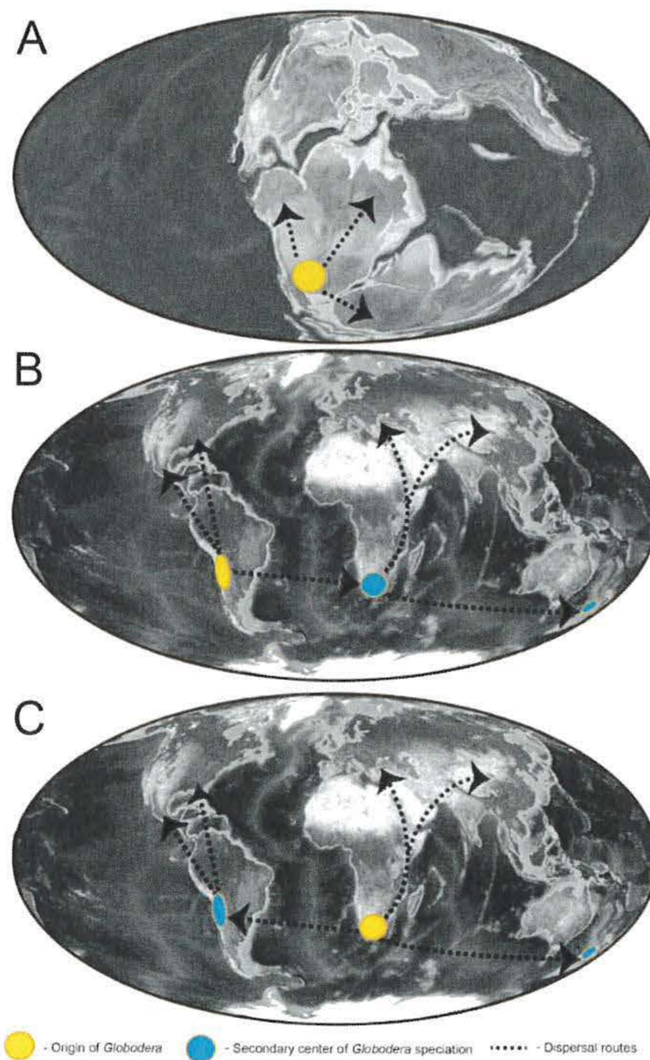


Figure 12. Origin and dispersal of the genus *Globodera* in the world. A: Out of west Gondwana hypothesis; B: Out of South America (Andean origin) hypothesis; C: Out of South Africa hypothesis.

PHYLOGENY OF THE GENUS *GLOBODERA*

Phylogenetic relationships within the genus *Globodera* were studied and discussed in detail using the ITS rRNA gene sequences by Subbotin *et al.* (2000, 2001, 2011), Skantar *et al.* (2011), Knoetze *et al.* (2013, 2017a, b), Lax *et al.* (2014) and other authors. Phylogeny of this genus inferred from *COI* gene in the present study is mainly congruent with that from the ITS rRNA gene, however, it does not give distinct resolution between some species. The genus *Globodera* displays two main clades in phylogenetic trees: *i)* *Globodera* from South and North America parasitizing plants from Solanaceae; and *ii)* *Globodera* from Africa, Europe, Asia and New Zealand parasitizing plants from Asteraceae and other families. The first main clade includes the first subclade with *G. pallida* and *G. mexicana* and the second one with *G. rostochiensis, G. tabacum* and *G. ellingtonae*, although relationships of the last two species with others are not resolved based on *COI* and *cytb* gene analysis. The second main clade consists of three subclades: *i)* *G. capensis* from South Africa occupying a basal position, *ii)* *G. zelandica* and two undescribed species from unknown plants in subalpine vegetation of North and South Islands of New Zealand and *iii)* *G. artemisiae, G. millefolii* from Europe and Asia, *G. agulhasensis* from South Africa, all parasitizing Asteraceae, as well as *G. sandveldensis* from South Africa and one undescribed species from Portugal.

ORIGIN AND PHYLOGEOGRAPHY OF THE GENUS *GLOBODERA*

After analysis of literature and phylogenetic trees, three main hypotheses of the genus *Globodera* origin could be proposed and tested: *i)* out of the west Gondwana hypothesis; *ii)*: out of South America hypothesis and *iii)* out of South Africa hypothesis.

Stone (1979) suggested that *Globodera* may have appeared in Gondwanaland at that time, remaining on the part of the super-continent which became South America while the ancestors of European species were carried northwards when fragments of Gondwanland encountered Laurasia. This hypothesis was partly supported by Picard *et al.* (2008) who considered the age of the *Cactodera* – (*Punctodera* + *Globodera*) divergence to be close to the Heteroderinae–Punctoderinae divergence, which might have occurred after separation of Laurasia and Gondwana, that was, 173–130 Mya. The divergence of the South American and Laurasian *Globodera* lineages was considered to have occurred between 80 and 60 Mya after the break of a temporary connection between North and South America in the Palaeocene. Subbotin *et al.* (2011) also speculated that if the estimated ages of host plants presently associated with non-cyst and cyst nematodes would be considered, slightly younger dates of divergence between nematode lineages could be suggested. The divergence of the two main *Globodera* lineages was associated with the time of origin for the Solanaceae, that would be, 65–51 Mya. Thus, the split of the two *Globodera* lineages might have occurred subsequent to the breakup of Gondwana and the Africa and South America split in the Mid-Cretaceous. Thus, the evolution of Punctoderinae cannot be explained solely by the separation of the continents and

diffusion expansion (Subbotin *et al.,* 2011). However, the molecular clock analysis made in the present study suggested a relatively recent origin of *Globodera,* which could be more than 3 Mya, and thus, rejected 'out of the west Gondwana hypothesis' and indicated recent several long-distance dispersal events from a main center of *Globodera* origin.

It has been hypothesized that centre of origin, diversification and speciation in main *Globodera* clades and subclades occurred in mountainous regions (Subbotin *et al.*, 2011, 2016). The main center of diversification for *Globodera* parasitizing solanaceous plants occurred in the Andes, known as one of the most biodiverse region on earth, from which many Solanaceae plants also occurred (Luebert & Weigend, 2014; Hazzi *et al.*, 2018). This region belongs to one of thirty-six biodiversity hotspot (Tropical Andes Biodiversity Hotspot) containing about one-sixth of all plant life in the world, including 30,000 species of vascular plants.

The *Globodera* belonging to the second main clade have at least two centers of diversification: mountains of the Western Cape in South Africa (Cape Floristic Region) and North and South Island mountains in New Zealand. The Andean region, the Cape and Australian region with New Zealand belong to so-called the Austral Kingdom (Morrone, 2018). The phylogeographic results indicated that New Zealand evidently is the secondary center of speciation and only South America or South Africa could be considered as candidates for the places from which *Globodera* originated. The result of the S-DIVA analysis of species distribution was not able to define with high statistically values the most probable area for *Globodera* origin. The hypotheses 'out of South America' and 'out of South Africa' origin were almost equally supported from the present datasets. Evidently, more taxa from *Globodera* and related genera are needed to be included in the analysis to define a main region of the origin for the genus *Globodera*.

Origin and population structures of PCNs in Andes

It has been shown by many studies that the Andean divide has played an important historical role in origin and creating deep intraspecific structure of different organisms. Uplift of the Andes during the Late Miocene and Pliocene in conjunction with the climatic oscillation of the Pleistocene had a strong effect on the diversification of Andean biota (Luebert & Weigend, (2014). Peruvian and Bolivia Andes have been accepted by many nematologists as the centre of origin of PCNs (Spears, 1968; Evans *et al.*, 1975; Canto Saenz & De Scurrah, 1977; Sosa-Moss, 1987). It has been shown that north of 15.6° S only *G. pallida* was found, but south of this latitude most population examined were *G. rostochiensis* or a mixture of the two species in region of the Lake Titicaca (Evans *et al.,* 1975; Franco, 1977). Our present study using the molecular analysis confirmed this pattern of distribution of two PCN species. Grenier *et al.* (2010) already noticed that the same area was identified by Spooner *et al.* (2005) as the locus of a major cladistic split for wild potatoes of the *Solanum brevicaule* complex. It also corresponds to one of geographic barriers across the Andes named as Yungas Inter-Andean Valleys, which play a role in isolation and diversification events in some endemic organisms in the tropical Andes between the last Miocene to the Pleistocene (Hazzi *et al.*, 2018).

The analysis the mtDNA haplotype distribution suggests two centres of origin for PCNs, one centre for *G. pallida* and another one for *G. rostochiensis*. The center of *G. pallida* origin is likely is in the south of Peru between Cusco and Lake Titicaca. It has been already suggested by Franco (1977) and Stone (1979), that the current distribution indicated that PCN speciation may have occurred north and south of Lake Titicaca, the lake and high ground to its north forming a barrier which have been more effective during parts of the Pleistocene than now. Picard *et al.* (2007) found that the clade containing the southern populations of *G. pallida* is genetically more diverse and forms the most basal lineages. South-to-north pattern of decreasing genetic variability was also clearly observed based on analysis of *cytb* gene and microsatellites. Our results confirmed these conclusions and this distribution pattern of *G. pallida* and moreover it also suggested the centre of origin for *G. rostochiensis*, which may be located in the south of Bolivia or north-west Argentina. Brücher (1959) already proposed that PCNs might have its origin in north-west Argentina. Franco (1977) followed this suggesting that PCNs species evolved in that region in scattered population isolated during the Pleistocene glaciations and then migrated to north (Stone, 1979). The phylogenetic and phylogenetic analysis of the presently available data suggested that South Andes including north-west Argentina might be indeed considered as other and more ancient center of origin for *G. rostochiensis*, *G. elligtonae* and *G. tabacum*. The *COI* haplotypes of *G. rostochiensis* mainly found in south Bolivia occupied the basal lineages in the phylogenetic tree. South-to-north pattern of decreasing genetic variability was also observed for this species. It has been shown that many Andean plant lineages had more recent origin in the northern than in the Central Andes and in the central than in the Southern Andes, which agreed with the assumption that the Andean uplift progressed from South to North (Luebert & Weigend, 2014). The north-west Argentina could be also center for *G. ellingtonae* origin and diversification. This species was described from USA, Oregon (Powel Butte) and Idaho (Teton County) by Handoo *et al.* (2012), but was also found in Argentina (Salta Province)(Lax *et al.*, 2005) and Chile (Antofagasta region)(Grenier *et al.*, 2010; Skantar *et al.*, 2011; Subbotin *et al.* 2011). Comparisons of biological attributes of *G. ellingtonae*, showed that this species is more similar with *G. rostochiensis* than *G. pallida* (Dandurand *et al.*, 2019). In the *COI* phylogenetic tree, *G. tabacum* from Argentina also occupied a basal position, which suggest its origin in same region.

Origin and dispersals of European and North American PCN populations

The PCN distribution in other countries was the result of dispersion of native populations with seed potatoes. Molecular studies using different markers revealed a complex evolutionary history for PCNs. Plantard *et al.* (2008) based on analysis of *cytb* gene and microsatellite markers concluded that the source of European *G. pallida* was from a single restricted area in the south of Peru, located between the north shore of Lake Titicaca and Cusco. Grenier *et al.* (2001, 2010) have provided further evidence that only a fraction of the genetic variability found in South America has been introduced into Europe. Our re-analysis of the *cytb* gene sequences of *G. pallida* with additional data confirmed these conclusions. Our results showed that only four *COI* haplotypes from two groups (A and C) were found outside of South America

where GpCOIA1 was mostly recorded in several European countries and USA. Other two haplotypes, GpCOIC1 and GpCOIC2 were found in UK and Kenya, respectively. However, uniqueness of *COI* Kenyan haplotypes should be carefully tested to determine if it is a recent evolution event, sequence reading mistake or heteroplasmy. High diversity of *cytb* haplotypes of *G. pallida* differing in one or two nucleotides were detected in Europe, however, such difference in *cytb* gene sequences could be explained by problems of heterogeneity of this gene. However, none of European *G. pallida COI* group A haplotype and *cytb* haplotypes does exactly match to any presently known Peruvian haplotypes. Only the exception is the haplotype CpCbA16 detected in UK and Juliaca in southern Peru and La Paz Department of Bolivia.

In the present study, we provided the evidence for high diversity of haplotypes and the structured phylogeographic pattern of *G. rostochiensis* in Bolivia. *Globodera rostochiensis* from Europe and North America and other world widely distributed populations of this species belong to a single *COI* or *cytb* haplotype. These haplotypes are identical to those found mainly in the La Paz Department of Bolivia situated at the western country border, sharing Lake Titicaca with adjacent Peru. Thus, we can hypothesize that *G. rostochiensis* was introduced from this place to other world regions with potato trades. Our hypothesis is different from that by Boucher *et al.* (2013), who proposed that European populations would appear to originate from either Peru or Chile. Evidently, more nematode samples from South America should be analysed to test these hypotheses.

ORIGIN AND POPULATION STRUCTURES OF *GLOBODERA MEXICANA* IN MEXICO

The analysis of *cytb* and *COI* gene sequences revealed that *Globodera pallida* and *G. mexicana* are sister species with divergence date of 1.6 May. *Globodera mexicana* parasitizing wild solonaceous host-plants is likely originated and diversified in Mexico in the result of a long distant dispersal event from the Andes. Our limited sampling results suggests that observed genetic diversity of *G. mexicana* populations may be also associated with zones of local topographical complexity. Mexico is crossed by large mountain systems (Sierra Madre Oriental, Sierra Madre Occidental, Sierra Madre del Sur, Sierra de Chiapas and the Trans-Mexican Volcanic Belt) corresponding to different geological provinces that differ vastly in age. These mountains are hotspots of biodiversity and endemism as a result of local and regional extinction, long-distance colonization, and local recruitment (Sosa *et al.*, 2018). Our work reveals the need to conduct more extensive studies at different elevations in mountain systems to understand factors of diversification of *Globodera mexicana*.

ORIGIN AND SPECIATION OF *GLOBODERA* IN SOUTH AFRICA

No indigenous *Globodera* species has been found in Africa until the discovery of such a species in the Sandveld region on the West Coast of South Africa by Knoetze *et al.* (2006). It gave a rise to new speculations about the origins of the group (Knoetze & Swart, 2014). The Sandveld is situated in the Cape

Floristic Region (CFR), which was identified as a biodiversity hotspot covering an area of 87,892 km^2 and showing exceptionally high diversity and endemism of vascular plants and invertebrates (Cowling & Heijnis, 2001). The rich biodiversity of the Cape Floristic Region Biodiversity Hotspot is due to an extensive and complex array of habitat types derived from topographical and climatic diversity in the region's rugged mountains, fertile lowlands, semi-arid shrublands and coastal dunes (Cowling *et al.,* 2009). Our results showed *Globodera* originated and diversified in South Africa 1.6 Mya and subsequently expanding into Eurasia (1.4 May) and New Zealand (0.9 Maya) *via* several dispersal events.

The results presented here are based on analysis of three genetic markers only. Rapid advancement in genomics technologies and the affordability of high-throughput nucleotide sequencing will allow to provide more detailed analysis of evolutionary relationships between *Globodera* species and give more very precise and detailed investigation of its origin, spread, and species distribution and even predict its potential future patterns.

Acknowledgements

The authors thank Drs N. Africander, V. Blok, A. Chakaeva, M. Christoforou,V. N. Chizov, L.M. Dandurand, N. Dauphinais, M. Docet, E. Grenier, J. Hallmann, G. Karssen, S. Kiewnick, M. Moens, M. Mundo-Ocampo, A. Peetz, H.J. Rumpenhorst, J. Rowe, X. Wang and I.A. Zasada for providing nematode samples. The authors especially acknowledge Dr. J.G. Baldwin for of the *Globodera* materials, which were given by L.I. Miller to the Nematology Department, University of California, Riverside in 1989.

Publication

The results of this project will be published as a manuscript under the title and authorships: "DNA barcoding, phylogeny and phylogeography of the cyst nematode species from the genus *Globodera* (Tylenchida: Heteroderidae)", Sergei A. SUBBOTIN, Javier FRANCO, Rinus KNOETZE, Tatiana V. ROUBTSOVA, Richard M. BOSTOCK and Ignacio CID DEL PRADO VERA.

Reference

Alenda, C. Gallot-Legrand, A., Fouville D. & Grenier, E. (2013). Sequence polymorphism of nematode effectors highlights molecular differences among the subspecies of the tobacco cyst nematode complex. *Physiological and Molecular Plant Pathology* 84, 107-114. DOI: 10.1016/j.pmpp.2013.08.004

Armstrong, M.R., Blok, V.C. & Phillips, M.S. (2000). A multipartite mitochondrial genome in the potato cyst nematode *Globodera pallida*. *Genetics* 154, 181–192.

Armstrong, M.R., Husmeier, D., Phillips, M.S. & Blok, V.C. (2007). Segregation and recombination of a multipartite mitochondrial DNA in populations of the potato cyst nematode *Globodera pallida*. *Journal of Molecular Evolution* 64, 689–701. DOI: 10.1007/s00239-007-0023-8

Bandelt, H.J., Forster, P., Sykes B.C. & Richards, M.B. (1995). Mitochondrial portraits of human populations using median networks. *Genetics* 141, 743-753.

Bandelt, H., Forster, P. & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16, 37-48. DOI: 10.1093/oxfordjournals. molbev.a026036

Blok, V.C., Malloch, G., Harrower, B., Phillips, M.S., & Vrain, T.C. (1998). Intraspecific variation in ribosomal DNA in populations of the potato cyst nematode *Globodera pallida*. *Journal of Nematology* 30, 262–274.

Boucher, A.C., Mimee, B., Montarry, J., Bardou-Valette, S., Bélair G., Moffett, P. & Grenier, E. (2013). Genetic diversity of the golden potato cyst nematode *Globodera rostochiensis* and determination of the origin of populations in Quebec, Canada. *Molecular Phylogenetics and Evolution* 69, 75–82. DOI: 10.1016/j.ympev.2013.05.020

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 10(4), e1003537. DOI: 10.1371/journal.pcbi.1003537

Brücher, E. (1959). Über das Wildvorkommen des Nematodes *Heterodera rostochiensis* in Nord-Argentinien. *Naturwissenschaften* 47, 21.

Canto-Saenz, M. & de Scurrah, M.M. (1977). Races of the potato cyst nematode in the Andean region and a new system of classification. *Nematologica* 23, 340–349.

Cowling, R.M. & Heijnis, C.E. (2001). The identification of broad habitat units as biodiversity entities for a systematic conservation planning in the Cape Floristic Region. *South African Journal of Botany* 67, 15–38.

Cowling R.M., Proches, S. & Partridge T.C. (2009). Explaining the uniqueness of the Cape flora: incorporating geomorphic evolution as a factor for explaining its diversification. *Molecular Phylogenetic and Evolution* 51, 64-74. DOI: 10.1016/j.ympev.2008.05.034

Chitambo, O., Haukeland, S., Fiaboe, K.K.M., Grundler, F.M.W. (2019). African nightshade and African spinach decrease root-knot nematode and potato cyst nematode soil infestation in Kenya. *Plant Disease* (in press). DOI: 10.1094/PDIS-07-18-1193-RE

Dandurand, L.M., Zasada, I.A., Wang, X., Mimee, B., De Jong, W., Novy, R., Whitworth, J. & Kuhl, J.C. (2019). Current status of potato cyst nematodes in North America. *Annual Review of Phytopathology* (in press). DOI: 10.1146/annurev-phyto-082718-100254

Evans, K., Franco, J. & de Scurrah, M.M. (1975). Distribution of species of potato cyst-nematodes in South America. *Nematologica* 21, 365–369.

Ferris, V.R., Miller, L.I., Faghihi, J. & Ferris, J.M. (1995) Ribosomal DNA comparisons of *Globodera* from two continents. *Journal of Nematology* 27, 273–283.

Ferris, V.R., Krall, E., Faghihi, J. & Ferris, J.M. (1999) Phylogenetic relationships of *Globodera millefolii, G. artemisiae,* and *Cactodera salina* based on ITS region of ribosomal DNA. *Journal of Nematology* 31, 498–507.

Franco, J. (1977). *Studies on the taxonomy and biology of potato cyst-nematodes* Globodera *spp. Mulvey and Stone 1976.* PhD. Thesis. University of London.

Franco, J. & Evans, K (1978). Multiplication of some South American and European populations of potato cyst nematodes on potatoes possessing the resistance genes H_1, H_2 and H_3. *Plant Pathology* 27, 1-6.

Franco, N.F., Cid del Prado Vera, I. & Lamothe-Argumedo, R. (2000). *Globodera bravoae* sp. n. (Tylenchida: Heteroderidae) from Mexico. *International Journal of Nematology* 10, 169-176.

Geric Stare, B., Sirca,S. and Urek, G., (2013). Are we ready for the new nematode species of the *Globodera* genus? In book (Trdan, S. and Macek, J. Eds.): *Lectures and papers presented at the at the Slovenian conference on plant protection with international participation (and round table of risks reduction in phyto-pharmaceutical products use in the frame of cropsustain project),* Ljubljana, Društvo za varstvo rastlin Slovenije (Plant Protection Society of Slovenia), 144-150.

Gibson, T., Blok, V.C., Phillips, M.S., Hong, G., Kumarasinghe, D., Riley, I.T. & Dowton, M. (2007a). The mitochondrial subgenomes of the nematode *Globodera pallida* are mosaics: evidence of recombination in an animal mitochondrial genome. *Journal of Molecular Evolution* 64, 463–471. DOI: 10.1007/s00239-006-0187-7

Gibson, T., Blok, V.C. & Dowton, M. (2007b). Sequence and characterization of six mitochondrial subgenomes from *Globodera rostochiensis*: multipartite structure is conserved among close nematode relatives. *Journal of Molecular Evolution* 65, 308-315. DOI: 10.1007/s00239-007-9007-y

Gibson, T., Farrugia, D., Barrett, J., Chitwood, D.J., Rowe, J., Subbotin, S.A, Dowton, M. (2011). The mitochondrial genome of the soybean cyst nematode, *Heterodera glycines*. *Genome* 54. 565–574. DOI: 10.1139/g11-024

Grenier, E., Bossis, M., Fouville, D., Renault, L. & Mugniery, D. (2001). Molecular approaches to the taxonomic position of Peruvian potato cyst nematodes and gene pool similarities in indigenous and imported populations of *Globodera*. *Heredity* 86, 277-290.

Grenier, E., Fournet S., Petit, E & Anthoine, G. (2010). A cyst nematode 'species factory' called the Andes. *Nematology* 12, 163-169.

Handoo, Z., Carta, L., Skantar, A.M. & Chitwood, D.J. (2012). Description of *Globodera ellingtonae* n. sp. (Nematoda: Heteroderidae) from Oregon. *Journal of Nematology* 44, 40-57.

Handoo, Z.A. & Subbotin, S.A. 2018. Taxonomy, Identification and Principal Species. In: *Cyst Nematodes.* (Perry, R.N., Moens, M. and Jones, J.T., Eds). CAB International, UK, pp. 365-397.

Hasegawa, M., Kishino, H. & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 22, 160-174. DOI: 10.1007/BF02101694

Hazzi, N.A., Moreno, J.S., Ortiz-Movliav, C. & Palacio R.D. (2018). Biogeographic regions and events of isolation and diversification of the endemic biota of the tropical Andes. *PNAS* 113, 31, 7985–7990. DOI: 10.1073/pnas.1803908115

Hockland, S., Niere, B., Grenier, E., Blok, V., Phillips, M., den Nij, L., Anthoine, G., Pickup, J. & Viane, N. (2012). An evaluation of the implications of virulence in non-European populations of *Globodera pallida* and *G. rostochiensis* for potato cultivation in Europe. *Nematology* 14, 1-13. DOI: 10.1163/138855411X587112

Hoolahan, A.H., Blok, V.C., Gibson, T. & Dowton, M. (2011). Paternal leakage of mitochondrial DNA in experimental crosses of populations of the potato cyst nematode *Globodera pallida*. *Genetica* 139, 1509-1519.

Hoolahan, A.H., Blok, V.C., Gibson, T. & Dowton, M. (2012a). Evidence of animal mtDNA recombination between divergent populations of the potato cyst nematode *Globodera pallida*. *Genetica* 140, 19–29. DOI: 10.1007/s10709-012-9651-z

Hoolahan, A.H., Blok, V.C., Gibson, T. & Dowton, M. (2012b). A comparison of three molecular markers for the identification of populations of *Globodera pallida*. *Journal of Nematology* 44, 7–17.

Howe, D.K., Baer, C.F. & Denver, D.R. (2010). High rate of large deletions in *Caenorhabditis briggsae* mitochondrial genome mutation processes. *Genome Biology and Evolution* 2, 29-38. DOI: 10.1093/gbe/evp055

Knoetze, R., Malan, A.P. & Mounton, C. (2006). Differentiation of South African potato cyst nematodes (PCN) by analysis of the rDNA internal transcribed spacer region. *African Plant Protection* 12, 103-110.

Knoetze, R., Swart, A. & Tiedt, L.R. (2013). Description of *Globodera capensis* n. sp. (Nematoda: Heteroderidae) from South Africa. *Nematology* 15, 233–250. DOI: 10.1163/15685411-00002673

Knoetze, R. & Swart, A. (2014). A survey of the Cape Floristic Region of South Africa for the presence of cyst nematodes (Nematoda: Heteroderidae). *Zootaxa* 3893, 429-437. DOI: 10.11646/zootaxa.3893.3.7

Knoetze, R., Swart, A. Wentzel, R. & Tiedt, L.R. (2017a). Description of *Globodera agulhasensis* n. sp. (Nematoda: Heteroderidae) from South Africa. *Nematology* 19, 305-322. DOI: 10.1163/15685411-00003050

Knoetze, R., Swart, A. Wentzel, R. & Tiedt, L.R. (2017b). Description of *Globodera sandveldensis* n. sp. (Nematoda: Heteroderidae) from South Africa. *Nematology* 19, 805-816. DOI: 10.1163/15685411-00003087

Lax, P., Manduric, S., Doucet, M.E., Gallardo, C. & de L'Argentier, S.M. (2005). Primera cita del nematodo

blanco del quiste de la papa, *Globodera pallida*, en Argentina continental. XIII Congreso Latinoamericano de Fitopatología. III Taller Argentino de Fitopatología. Córdoba, Argentina.

Lax, P., Rondan Dueñas, J.C., Franco-Ponce, J., Gardenal, C.N. & Doucet, M.E. (2014). Morphology and DNA sequence data reveal the presence of *Globodera ellingtonae* in the Andean region. *Contributions to Zoology* 83, 227-243.

Luebert, F. & Weigend, M. (2014). Phylogenetic insights into Andean plant diversification. *Frontiers in Ecology and Evolution*, 19 DOI: 10.3389/fevo.2014.00027

Madani, M., Subbotin, S.A., Ward, L.J., Li X. & De Boer, S.H. (2010). Molecular characterization of Canadian populations of potato cyst nematodes, *Globodera rostochiensis* and *G. pallida* using ribosomal nuclear RNA and cytochrome *b* genes. *Canadian Journal of Plant Pathology* 32: 252-263. DOI: 10.1080/07060661003740033

Manduric, S. & Andersson, S. (2004). The identity of a Swedish Globodera (Nematoda: Heteroderidae) population, following comparisons with known populations of *G. artemisiae* (Eroshenko & Kazachenko, 1972) Behrens, 1975. *Russian Journal of Nematology* 12, 39-44.

Marché, L. Valette, S., Grenier, E. & Mugniéry, D. (2001). Intra-species DNA polymorphism in the tobacco cyst-nematode complex (*Globodera tabacum*) using AFLP. *Genome* 44, 941-946.

Morrone J.J. (2018). *Evolutionary biography of the Andean region*. CRC press, Taylor & Francis Group, USA. 250 pp.

Nicholas, K.B., Nicholas Jr, H.B. & Deerfield II, D.W. (1997). GeneDoc: analysis and visualization of genetic variation. *EMBNET NEWS* 4, 14.

Nguyen, V.C., Villate, L. Gutierrez-Gutierrez, C., Castillo, P., Van Ghelder, C., Plantard, O. & Esmenjaud, D. (2019). Phylogeography of the soil-borne vector nematode *Xiphinema index* highly suggests Eastern origin and dissemination with domesticated grapevine. *Scientific Report* 9, 7313. DOI: 10.1038/s41598-019-43812-4

Palomares-Rius, J.E., Cantalapiedra-Navarrete, C., Archidona-Yuste, A., Subbotin, S.A. & Castillo, P. (2017). The utility of mtDNA and rDNA for barcoding and phylogeny of plant-parasitic nematodes from Longidoridae (Nematoda, Enoplea). *Scientific Reports* 7, 10905. DOI: 10.1038/s41598-017-11085-4

Phillips, W.S. Brown A.M.V., Howe, D.K., Peetz, A.B., Blok, V.C., Denver, D.R. & Zasada, I.A. (2016). The mitochondrial genome of *Globodera ellingtonae* is composed of two circles with segregated gene content and differential copy numbers. *BMC Genomics* 17, 706. DOI: 10.1186/s12864-016-3047-x

Picard, D., Sempere, T. & Plantard, O. (2007). A northward colonisation of the Andes by the potato cyst nematode during geological times suggests multiple host-shifts from wild to cultivated potatoes. *Molecular Phylogenetics and Evolution* 42, 308-316. DOI: 10.1016/j.ympev.2006.06.018

Picard, D., Sempere, T. & Plantard, O. (2008). Direction and timing of uplift propagation in the Peruvian Andes deduced from molecular phylogenetics of highland biotaxa. *Earth and Planetary Science Letters* 271, 326-336.

Plantard, O., Picard, D., Valette, S., Scurrah, M., Grenier, E. & Mugniéry, D. (2008). Origin and genetic diversity of Western European populations of the potato cyst nematode *Globodera pallida* inferred from mitochondrial sequences and microsatellite loci. *Molecular Ecology* 17, 2208-2218. DOI: 10.1111/j.1365-294X.2008.03718.x

Powers, T.O., Harris, T., Higgins, R., Sutton, L. & Powers, K.S. (2010). Morphological and molecular characterization of *Discocriconemella inarata*, an endemic nematode from north american native tallgrass prairies. *Journal of Nematology* 42, 35-45.

Pylypenko, L.A., Phillips, M.S. & Blok, V.C. (2008). Characterisation of two Ukrainian populations of *Globodera pallida* in terms of their virulence and mtDNA, and the biological assessment of a new resistant cultivar Vales Everest. *Nematology* 10, 585-590.

Ronquist, F. & Huelsenbeck, J.P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572-1574. DOI: 10.1093/bioinformatics/ btg180

Sabo, A., Reis, L.G.L., Krall, E., Mundo-Ocampo, M. & Ferris, V.R. (2002). Phylogenetic relationships of a distinct species of *Globodera* from Portugal and two *Punctodera* species. *Journal of Nematology* 34, 263-266.

Sirca, S. & Urek, G. (2004). Morphometrical and ribosomal DNA sequence analysis of *Globodera rostochiensis* and *Globodera achilleae* from Slovenia. *Russian Journal of Nematology* 12, 161-168.

Skantar, A.M., Handoo, Z.A., Zasada. I. A., Inghram, R.E., Carta, L.K., and Chitwood, D.J. (2011). Morphological and molecular characterization of *Globodera* populations from Oregon and Idaho. *Phytopathology* 101, 480-491. DOI: 10.1094/PHYTO-01-10-0010.

Sosa, V., De-Nova, J.A. & Vásquez-Cruz, M. (2018). Evolutionary history of the flora of Mexico: dry forest cradles and museums of endemism. *Journal of Systematics and Evolution* 56, 523-536. DOI: 10.1111/jse.12416

Sosa-Moss, C. (1987). Cyst nematodes in Mexico, Central and South America. *Nematologia Mediterranea* 15, 1-12.

Spears, J.F. (1968). *The Golden Nematode.* Handbook. United States of America Department of Agricultural Research Service, Washington D.C., USA, 83 pp.

Spooner, D.M., McLean, K., Ramsay, G., Waugh, R. & Bryan, G.J. (2005). A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proceedings of the National Academy of Sciences of the United States of America* 102, 14694-14699.

Stone, A.R. (1979). Co-evolution of nematodes and plants. *Symbolae Botanicae Upsaliensis* 22, 4, 46-61.

Subbotin, S.A., Halford, P.D., Warry, A. & Perry, R. (2000). Variations in ribosomal DNA sequences and phylogeny of *Globodera* parasitising Solanaceae. *Nematology* 2, 591-604.

Subbotin, S.A., Vierstraete, A., De Ley, P., Rowe, J., Waeyenberge, L., Moens, M. & Vanfleteren, J.R. (2001). Phylogenetic relationships within the cyst-forming nematodes (Nematoda, Heteroderidae) based on analysis of sequences from the ITS regions of ribosomal DNA. *Molecular Phylogenetics and Evolution* 21, 1-16. DOI: DOI: 10.1006/mpev.2001.0998

Subbotin, S.A., Mundo-Ocampo, M. & Baldwin, J.G. (2010). *Systematics of cyst nematodes (Nematoda: Heteroderidae). Nematology Monographs and Perspectives 8A* (Series Editors: Hunt, D.J. & Perry, R.N.). Leiden, The Netherlands, Brill.

Subbotin, S.A., Cid Del Prado Vera, I., Mundo-Ocampo, M. & Baldwin, J.G. (2011). Identification, phylogeny and phylogeography of circumfenestrate cyst nematodes (Nematoda: Heteroderidae) as inferred from analysis of ITS-rDNA. *Nematology* 13, 805-824. DOI: 10.1163/15685411-00003170

Subbotin, S.A., Toumi, F., Elekcioglu, I.H., Waeyenberge, L. & Tanha Maafi, Z. (2018). DNA barcoding, phylogeny and phylogeography of the cyst nematodes of the *Avenae* group from the genus Heterodera (Tylenchida: Heteroderidae). *Nematology* 20, 671-702. DOI 10.1163/156

Sworfford, D.L. (2003). *PAUP*: phylogenetic analysis using parsimony (*and other methods)*, version 4.0b 10. Sunderland, Massachusetts, USA, Sinauer Associates.

Thiéry, M., Fouville, D. & Mugniéry D. (1997). Intra and interspecific variability in *Globodera*, parasites of Solanaceous plants, revealed by random amplified polymorphic DNA (RAPD) and correlation with biological features. *Fundamental and Applied Nematology* 20, 495-504.

Yu, Y., Harris, A.J., Blair, C. & He, X.J. (2015). RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and Evolution* 87, 46-49. DOI: 10.1016/j.ympev.2015.03.008

Zheng, J., Subbotin, S.A., Waeyenberge, L. & Moens, M. (2000). Molecular characterisation of Chinese *Heterodera glycines* and *H. avenae* populations based on RFLPs and sequences of rDNA-ITS regions. *Russian Journal of Nematology* 8, 109-113.

Table 1. Species and populations of cyst nematodes from *Globodera* used in the present project.

Species	Location (pathotype)	Host	Sample code	*COI* haplotype	*COI* GenBank accession number	*cytb* haplotype	*Cytb* GenBank accession number	ITS rRNA GenBank accession number	Source and/or reference
G. agulhasensis	South Africa, Western Cape Province, Gansbaai	*Senecio burchelli*	SK18/3, CD1895	GaguCOI1	X	-	-	KC148543	R. Knoetze, Knoetze *et al.* (2017a)
G. artemisiae	Germany, Münster	*Artemisia* sp.	CD2193a	GartCOI1	X	-	-	-	D. Sturhan
G. artemisiae	Russia, Primorskii Krai	*Artemisia* sp.	CD2178a, b	GartCOI2, GartCOI3	X, X	-	-	-	A.S. Eroshenko
G. capensis	South Africa, Western Cape, Sandveld area, Sandberg farm	Unknown	29147, CD1892a	GcapCOI1	X	GcapCb1	X	JQ595413	R. Knoetze
G. ellingtonae	USA, Oregon, Powell Butte	Potato	CD1988a	GeCOI1	X	-	-	-	L.-M. Dandurand, I. Zasada, A. Peetz
G. mexicana	Mexico, State of Mexico, San Miguel de la Victoria, sample 34	*Solanum nigrum*	CD2814	GmexCOI1	X	-	-	-	I. Cid del Prado Vera, S.A. Subbotin
G. mexicana	Mexico, State of Mexico, Texcoco, El jardín de Tequexquinahuc, sample 5	*S. nigrum*	CD2821a, b, d, e	GmexCOI2	X	GmexCb2	X	-	I. Cid del Prado Vera, S.A. Subbotin
G. mexicana	Mexico, State of Mexico, Amecameca, San Diego Huehuecalco, sample 8	*S. stoloniferum*	CD2792b	GmexCOI3	X	GmexCb1	X	-	I. Cid del Prado Vera, S.A. Subbotin
G. mexicana	Mexico, State of Mexico, Amecameca, San Diego Huehuecalco	*S. stoloniferum*	CD514	GmexCOI3	X	-	-	HQ260405, HQ260406	I. Cid del Prado Vera, Subbotin *et al.* (2011)
G. mexicana	Mexico, State of Mexico, Amecameca, San Diego Huehuecalco, sample 5	*S. stoloniferum*	CD2809a, d, e,	GmexCOI4, GmexCOI5	X	-	-	-	I. Cid del Prado Vera, S.A. Subbotin
G. mexicana	Mexico, State of Mexico, Amecameca, San Diego Huhuecalco, sample 6	*S. stoloniferum*	CD2784	GmexCOI4	X	-	-	-	I. Cid del Prado Vera, S.A. Subbotin
G. mexicana (= *G. bravoae* syn. n)	Mexico, La Cañada de Contreras, Delegación Magdalena Contreras, México city	*Jaltomata procumbens*	CD1891b, c	GmexCOI4	X	-	-	X	I. Cid del Prado Vera
G. mexicana	Mexico, Tlaxcala State, Huamantla, Francisco Villa, Tecoac, sample 38	*S. rostratum*	CD2862e, c, CD2183,	GmexCOI5, GmexCOI6, GmexCOI7	X	-	-	-	I. Cid del Prado Vera, S.A. Subbotin
G. millefolii	Estonia	*Achillea millefolium*	CD2181a, b	GmilCOI1	X	-	-	HQ260407	E. Krall, D. Sturhan, Subbotin *et al.* (2001)
G. pallida	Cyprus, eastern part, (Pa2/3)	Potato	CD2739b	GpCOIA1	X	GpCbA1	X	-	N. Dauphinais, M. Christoforou
G. pallida	UK	Potato	CD2169a, CD2664a	GpCOIA1	X	GpCbA1	X	-	J. Rowe
G. pallida	Spain, Tenerife, sample 1	Potato	CD2580a, b	GpCOIA1	X	GpCbA1	X	X, X	J. Franco
G. pallida	Spain, Tenerife, sample 2	Potato	CD2577a, b	GpCOIA1	X	GpCbA1	X	X	J. Franco
G. pallida	Spain, Tenerife, sample 3	Potato	CD2581b, a	GpCOIA1	X	GpCbA1, GpCbA14	X	X	J. Franco
G. pallida	Switzerland (Pa2/3)	Potato	CD2743a	GpCOIA1	X	GpCbA1	X	-	N. Dauphinais, E. Grenier
G. pallida	Netherlands (Pa2)	Potato	CD2740a, b	GpCOIA1	X	GpCbA1	X	-	N. Dauphinais, G. Karssen
G. pallida	Germany, Delmsen	Potato	CD2198a	GpCOIA1	X	-	-	-	
G. pallida	Germany, Kelle	Potato	CD611, CD2199a	GpCOIA1	X	GpCbA1	X	-	H.J. Rumpenhorst
G. pallida	Germany, Lower Saxony	Potato	CD2869a, b	GpCOIA1	X	-	-	-	S. Kiewnick
G. pallida	France, Le Rheu Cedex, Domaine de la Motte (Pa2/3)	Potato	CD2741a, b	GpCOIA1	X	GpCbA1	-	-	N. Dauphinais
G. pallida	USA, Idaho	Potato	CD2744a	GpCOIA1	X	GpCbA1	X	-	N. Dauphinais, L.-M. Dandurand
G. pallida	Bolivia, La Paz, Los Andes, Keruni	Potato	CD2595a	GpCOIA3	X	GpCbA11	X	-	J. Franco
G. pallida	Peru	Potato	CD610	GpCOIA3	X	GpCbA16	X	-	H.J. Rumpenhorst
G. pallida	Bolivia, La Paz, Aroma, Cajani	Potato	CD2597a, b	GpCOIA3	X	GpCbA16	X	X	J. Franco
G. pallida	Bolivia, La Paz, Ingavi, Casachuta	Potato	CD2596a, b	GpCOIA3	X	GpCbA11	X	X	J. Franco
G. pallida	Peru, Capachica	Potato	CD2188a, b	GpCOIA4	X	GpCbA11	X	HQ260426	J. Hallmann, Subbotin *et al.*, (2011)
G. pallida	Bolivia, Las Paz, Los Andes, Seviruyo	Potato	CD2561b	GpCOIA4	X	GpCbG1	X	-	J. Franco
G. pallida	Peru, Cusco, Paucartambo, Cotatoclla	Potato	CD2558a, b, c, d	GpCOIA5, GpCOIB1	X	GpCbC2, GpCbA15	X	X, X	J. Franco

G. pallida	UK, P4A	Potato	CD2706a (cl1, cl2)	GpCOIC2	X	-	-	HQ670257-HQ670262, HQ670269	V. Blok, Hoolahan *et al.* (2011)
G. pallida	Peru, Cusco, Quispicanchis	Potato	CD2553a, CD2553b, CD2553c	GpCOID1	X	GpCbB1, GpCbB7, GpCbD1	X	X	J. Franco
G. pallida	Peru, Cusco, Marangani, Canchis	Potato	CD2554a, b	-	-	GpCbB1	X	-	J. Franco
G. pallida	Peru, Cruz Pampa	Potato	CD2555a, b	-	-	GpCbB1, GpCbB7	X	-	J. Franco
G. pallida	Peru	Potato	CD2557	-	-	GpCbB1	X	-	J. Franco
G. pallida	Costa Rica, Volcan Irazu	Potato	CD2547	-		GpCbB7	X	-	J. Franco
G. pallida	Peru, Cusco, Calca	Potato	CD2552	-		GpCbB7	X	-	J. Franco
G. pallida	UK, East Lothian	Potato	CD2705a	-		GpCb14	X	AM409004	V. Blok, Pylypenko *et al.* (2008)
G. pallida	Panama	Potato	CD2560a, b	-	-	GpCbB1	X	-	J. Franco
G. pallida	Peru, Chocom, Jauja	Potato	CD2559	-	-	GpCbB1	X	-	J. Franco
G. pallida	New Zealand	Potato	CD2742a, b	-	-	GpCbA14	X	-	N. Dauphinais, E. Grenier
G. rostochiensis	Germany, Harmerz (Ro5)	Potato	CD2200	GrCOIA1	X	-	-	-	H.J. Rumpenhorst, J. Hallmann
G. rostochiensis	German, Hannover (Ro1)	Potato	CD2197a	GrCOIA1	X	-	-	-	H.J. Rumpenhorst, J. Hallmann
G. rostochiensis	UK (Ro4)	Potato	CD2166a	GrCOIA1	X	-	-	-	H.J. Rumpenhorst, J. Hallmann
G. rostochiensis	Russia, Belgorod region, sample 1	Potato	CD2682b	GrCOIA1	X	CrCbA1	X	–	S.A. Subbotin
G. rostochiensis	Russia, Kaliningrad region, Pravdinskii district	Potato	CD2176, CD2663a	GrCOIA1	X	-	-	-	S.A. Subbotin
G. rostochiensis	Russia	Potato	CD558	GrCOIA1	X	-	-	-	V. N. Chizov
G. rostochiensis	Russia, Khabarovsk Krai, Komsomolsk-on Amur	Potato	CD2168b	GrCOIA1	X	-	-	-	S.A. Subbotin
G. rostochiensis	Russia, Primirskii Krai	Potato	CD2672b	GrCOIA1	X	CrCbA1	X	-	S.A. Subbotin
G. rostochiensis	Russia, Primirskii Krai, Vladivostok	Potato	CD2674a	GrCOIA1	X	-		-	S.A. Subbotin
G. rostochiensis	Russia, Primirskii Krai, Tavrichenka	Potato	CD2665b, CD2673b, CD2164b	GrCOIA1	X	CrCbA1	X	-	S.A. Subbotin
G. rostochiensis	Russia, Vladimir	Potato	CD2669b	GrCOIA1	X	-	-	-	S.A. Subbotin
G. rostochiensis	Russia, Smolensk	Potato	CD2167a, b	GrCOIA1	X	-	-	-	S.A. Subbotin, Subbotin *et al.* (2000)
G. rostochiensis	Russia, Belgorod region, sample 2	Potato	CD2671a, b	GrCOIA1	X	CrCbA1	X	-	S.A. Subbotin
G. rostochiensis	Russia, Pskov	Potato	CD2173a	GrCOIA1	X	-	-	-	S.A. Subbotin
G. rostochiensis	Russia, Jaroslav region	Potato	CD2675a	GrCOIA1	X	CrCbA1	X	-	S.A. Subbotin
G. rostochiensis	Russia, Moscow region	Potato	CD2172a	GrCOIA1	X	-	-	-	S.A. Subbotin
G. rostochiensis	Russia, Vologda region	Potato	CD2678	-		CrCbA1	X	-	S.A. Subbotin
G. rostochiensis	Kyrgyzstan	Potato	CD2165a, b	GrCOIA1	X	-	-	-	A. Chakaeva
G. rostochiensis	New Zealand	Potato	CD2177b	GrCOIA1	X	-	-	-	J. Rowe
G. rostochiensis	USA, New York (Ro1)	Potato	CD2835a	GrCOIA1	X	-	-	-	X. Wang
G. rostochiensis	USA, New York (Ro2)	Potato	CD2834a, b	GrCOIA1	X	-	-	-	X. Wang
G. rostochiensis	Canada, Quebec	Potato	CD2738a, b	GrCOIA1	X	CrCbA1	X	-	N. Dauphinais
G. rostochiensis	South Africa, Sandveld	Potato	CD2842	GrCOIA1	X	-	-	-	R. Knoetze, N. Africander
G. rostochiensis	South Africa, Gauteng	Potato	CD2840	GrCOIA1	X	-	-	-	R. Knoetze, N. Africander
G. rostochiensis	South Africa, Ceres	Potato	CD2841	GrCOIA1	X	-	-	-	R. Knoetze, N. Africander
G. rostochiensis	Guatemala, sample 32	Potato	CD2548b	GrCOIA1	X	-	-	-	J. Franco
G. rostochiensis	Guatemala, sample 48	Potato	CD2550b	GrCOIA1	X	-	-	-	J. Franco
G. rostochiensis	Guatemala, sample 10	Potato	CD2549a, b	GrCOIA1	X	-	-	-	J. Franco
G. rostochiensis	Venezuela, Marida, Mucuchles	Potato	CD2582b	GrCOIA1	X	CrCbA1	X	-	J. Franco
G. rostochiensis	Chile	Potato	CD2643b, CD2649b, c, CD2644b	GrCOIA1	X	CrCbA1	X	-	N. Viane
G. rostochiensis	Bolivia, La Paz, Los Andes, Lacaya	Potato	CD2617a, b	GrCOIA1	X	CrCbA1	X	-	J. Franco
G. rostochiensis	Bolivia, Las Paz, Los Andes, Seviruyo	Potato	CD2561a	GrCOIA1	X	CrCbA1	X	-	J. Franco
G. rostochiensis	Bolivia, La Paz, Ahijadero	Potato	CD2565b	GrCOIA1	X			-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Mizque, Picosilla	Potato	CD2583a, b, c	GrCOIA1	X	CrCbA1	X	-	J. Franco
G. rostochiensis	Bolivia, La Paz, Ingavi, Hilata Sta Trinidad	Potato	CD2573a, b	GrCOIA1	X	CrCbA1	X	-	J. Franco
G. rostochiensis	Bolivia, La Paz, Collana	Potato	CD2564a	GrCOIA1	X	CrCbA1	X	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Arani, Paredones	Potato	CD2604a, b	GrCOIA3	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Tiraque, Cochimita	Potato	CD2610b	CrCOIA4	X	-	-	-	J. Franco

G. rostochiensis	Bolivia, Cochabamba, Tiraque, Toralapa	Potato	CD2587b	-	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Charca, Pairumani	Potato	CD2571a	GrCOIA5	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Charca, Suragua	Potato	CD2567a	GrCOIA6	X	CrCbC6	X	X	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Alto Chapare, Iluri Grande	Potato	CD2606b	GrCOIA7	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Tapacari, Chuna Chumuni	Potato	CD2598a, b	GrCOIA13, GrCOIA7	X	CrCbA6	X	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Punata, Chaqui Knocha	Potato	CD2592a, b	GrCOIA7, GrCOIA13	X	CrCbA9	X	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Morochata-Plusilla	Potato	CD2563a, b	GrCOIA7, GrCOIA9	X	CrCbA9, CrCbA7	X	X, X	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Alto Chapare, Laracti Chico	Potato	CD2608a, b	GrCOIA8	X	CrCbA4,	X	-	J. Franco
G. rostochiensis	Bolivia	Potato	CD579	GrCOIA8	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Ayopaya, Patamorochata	Potato	CD2625a, b	GrCOIA10	X	CrCbA5	X	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Mizque	Potato	CD2594a	GrCOIA10	X	CrCbA5	X	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Ayopaya, Iglesiani	Potato	CD2566a, b, c	GrCOIA10	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Alto Chapare, Larati	Potato	CD2588a	GrCOIA11	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Tapacari, Pongo	Potato	CD2619b	GrCOIA13	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Alto Chapare, Candelaria	Potato	CD2603a	GrCOIA12	X	CrCbA8	X	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Tiraque, Cochimita	Potato	CD2610a	GrCOIA13	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Tiraque Koari	Potato	CD2616b	GrCOIA13	X	CrCbA8	X	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Tiraque, Waylla Pujru	Potato	CD2586b	GrCOIA13	X	-		-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Tiraque, Toralapa	Potato	CD2587c, CD2600b	GrCOIA13	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Sultora Baja	Potato	CD2620b	GrCOIA13	X	CrCbA6	X	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Arani, Canada Grande	Potato	CD2593a, b	GrCOIA13	X	CrCbA4	X	-	J. Franco
G. rostochiensis	Bolivia	Potato	CD577	GrCOIA13	X	-		-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Arani, Saitay Pampa	Potato	CD2609a, b	GrCOIA13	X	CrCbA4	X	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Arani, Sapolica	Potato	CD2605a	GrCOIA13	X	-		-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Alto Chapare, Chimpa Melga	Potato	CD2622a	GrCOIA13	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Nor Cinti, Arpoja Baja	Potato	CD2612b	GrCOIA14	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Nor Cinti, Jolencia Alta	Potato	CD2589b	GrCOIA14	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Alto Chapare, San Isidro	Potato	CD2601a, b	GrCOIA15	X	CrCbA8, CrCbC2	X	-	J. Franco
G. rostochiensis	Bolivia, Potosi, Chimpapata	Potato	CD2562b	GrCOIA16	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Potosi, Laguna Pampa	Potato	CD2618a	GrCOIB1, GrCOIB3	X	CrCbC3	X	-	J. Franco
G. rostochiensis	Bolivia, Potosi, Santiago	Potato	CD2575b	GrCOIB2	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Potosi, Ibanez, Charca Micani	Potato	CD2621a, b	GrCOIC1	X	CrCbC7	X	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Bolivar, Pampajasi	Potato	CD2591b	GrCOIC2	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Nor Cinti, Hucancorani Centro	Potato	CD2602a	CrCOIC3	X	CrCbC1	X		J. Franco
G. rostochiensis	Bolivia, Nor Cinti, Jolencia Alta	Potato	CD2589c	GrCOIC4	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Arani, Saga Saga	Potato	CD2570b	GrCOIC4	X	-	-	-	J. Franco
G. rostochiensis	Bolivia, Cochabamba, Arani, Laquina	Potato	CD2599	-	-	CrCbA4	X	-	J. Franco
G. rostochiensis	Bolivia, Nor Cinti, Pueblo Bajo	Potato	CD2614a, b	CrCOID1, GrCOIC5	X	CrCbB2, CrCbB1	X	-	J. Franco
G. rostochiensis	Bolivia, Nor Cinti, Pueblo Alto	Potato	CD2584b, a	GrCOIC8, GrCOIC6	X	CrCbB2	X	-	J. Franco
G. rostochiensis	Bolivia, Nor Cinti, Jolencia Alta	Potato	CD2589a	GrCOIC7	X	-	-	X	J. Franco
G. rostochiensis	Bolivia, Saavedra, Mamahota	Potato	CD2613a, b	-		CrCbB3, CrCbC4	X	-	J. Franco
G. rostochiensis	Bolivia, Nor Cinti, Chunchali	Potato	CD2578	-		CrCbC5	X	-	J. Franco
G. sandveldensis	South Africa, Western Cape Province, Sandveld region, Leipoldtville	Unknown	WK1, CD1893	GsanCOI1	X	-		KF809767	R. Knoetze, Knoetze *et al.* (2017b)
G. tabacum	USA, Virginia, 4A	*S. carolinense*	CD2695a, b	GtCOIA1	X	GtCb1	X	-	L.I. Miller, J.G. Baldwin

G. tabacum	USA, Virginia, OCN Standard	*S. carolinense*	CD2205a, b, CD2203a, b, CA140	GtCOIA1	X	-	-	-	L.I. Miller, J.G. Baldwin
G. tabacum	USA, Virginia, Baines	*S. carolinense*	CD2684a, b	GtCOIA1	X	GtCb1	X	-	L.I. Miller, J.G. Baldwin
G. tabacum	USA, Virginia, Anderson	*S. carolinense*	CD2692a	GtCOIA1	X	GtCb1	X	-	L.I. Miller, J.G. Baldwin
G. tabacum	Bolivia	Unknown	CD2702b	GtCOIA1	X	GtCb1	X	-	L.I. Miller, J.G. Baldwin
G. tabacum	USA, Virginia, Lynch	*S. carolinense*	CD2688b	GtCOIA1	X	-	-	-	L.I. Miller, J.G. Baldwin
G. tabacum	USA, Virginia, Horton	*S. carolinense*	CD2206a, CD2699	GtCOIA1	X	GtCb1	X	HQ260387, HQ260394	L.I. Miller, J.G. Baldwin, Subbotin *et al.* (2011)
G. tabacum	USA, Virginia, Paulette	*S. carolinense*	CD2689a, b	GtCOIA1	X	-	-	-	L.I. Miller, J.G. Baldwin
G. tabacum	USA, Virginia, 52A	*S. carolinense*	CD2208	GtCOIA1	X	-	-	-	L.I. Miller, J.G. Baldwin
G. tabacum	USA, Virginia, Inby	*S. carolinense*	CD2690b	GtCOIA1	X	-	-	-	L.I. Miller, J.G. Baldwin
G. tabacum	USA, Virginia, Fisher	*S. carolinense*	CD2691a	GtCOIA1	X	-	-	-	L.I. Miller, J.G. Baldwin
G. tabacum	Argentina, Jujuy Province	Tobacco	CD2186a, b	GtCOIC1	X	-	-	-	M. Docet
G. tabacum	Italy	Tobacco	CD523	GtCOIC2	X	-	-	HQ260403, HQ260404	M. Mundo-Ocampo, Subbotin *et al.* (2011)
G. tabacum	USA, Virginia, HCN Standard	*S. carolinense*	CD2202a, c, CD2204a, b, CD2694a	GtCOID1	X	GtCb3	X	-	L.I. Miller, J. Baldwin
G. tabacum	USA, Virginia, Blanton	*S. carolinense*	CD2693b	GtCOID1	X	-	-	-	L.I. Miller, J.G. Baldwin
G. tabacum	USA, Virginia, Rhodes	*S. carolinense*	CD2700a	GtCOID2	X	GtCb2	X	-	L.I. Miller, J.G. Baldwin
G. tabacum	USA, Virginia, 55A	*S. carolinense*	CD2696b	GtCOID2	X	GtCb2	X	-	L.I. Miller, J.G. Baldwin
G. tabacum	USA, Virginia, 93A	*S. carolinense*	CD2697b	GtCOID2	X	GtCb2	X	-	L.I. Miller, J.G. Baldwin
G. tabacum	France	Tobacco	CD2182a, b	GtCOIB1, GtCOIB2	X	-	-	-	M. Moens
Globodera sp. 1	New Zealand, North Island, Desert Road	Subalpine vegetation	582	Gsz1CO1	X	-	-	HQ260408	D. Sturhan
Globodera sp. 2	New Zealand, South Island, Lake Lyndon area, Canterbury	Subalpine scrub vegetation	CA114	Gsz2CO1	X	-	-	HQ260409	D. Sturhan
G. zelandica	New Zealand, South Island, Banks Peninsula	*Plagianthus regius*	NZ26, CD2195	GzCOI1	X	-	-	HQ260410	D. Sturhan
G. zelandica	New Zealand	Unknown	NZ23, CD2194a, b	GzCOI1, GzCOI2	X	-	-	-	D. Sturhan
G. zelandica	New Zealand	Unknown	Z712, CD2184a, b	GzCOI3	X	-	-	-	D. Sturhan

- - DNA sample was not amplified or sequenced.
X – original sequences will be submitted in the Genbank,

B. If appropriate, explain why objectives were not met.*

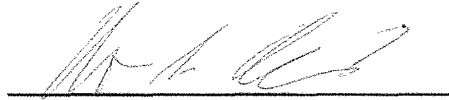
Not applied.

C. Where appropriate, explain any cost overruns or unobligated funds in excess of $1,000. *

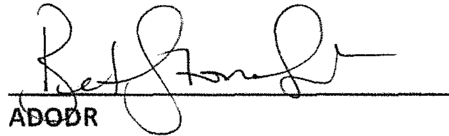
Not applied.

****indicates information is required per 7 CFR 3016.40 and 7 CFR 3019.51***

Approved and signed by



Cooperator **Date:** 6/18/2019



ADODR **Date:** 6.18.19