



**HAL**  
open science

## Europe as a secondary distribution hub in the worldwide invasion of the potato cyst nematode *Globodera rostochiensis*

Magali Esquibet, James M. Mwangi, Sebastian Kiewnick, Xiaohong Wang, Benjamin Mimee, Nurul Dwi Handayani, Wim Bert, Johannes Helder, John Wainer, Itaru Sakata, et al.

### ► To cite this version:

Magali Esquibet, James M. Mwangi, Sebastian Kiewnick, Xiaohong Wang, Benjamin Mimee, et al.. Europe as a secondary distribution hub in the worldwide invasion of the potato cyst nematode *Globodera rostochiensis*. *Scientific Reports*, 2024, 14, pp.13915. 10.1038/s41598-024-64617-0 . hal-04640037

**HAL Id: hal-04640037**

**<https://hal.inrae.fr/hal-04640037>**

Submitted on 9 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



OPEN

## Europe as a secondary distribution hub in the worldwide invasion of the potato cyst nematode *Globodera rostochiensis*

Magali Esquibet<sup>1</sup>, James M. Mwangi<sup>2,3</sup>, Sebastian Kiewnick<sup>3</sup>, Xiaohong Wang<sup>4</sup>, Benjamin Mimee<sup>5</sup>, Nurul Dwi Handayani<sup>6,11</sup>, Wim Bert<sup>6</sup>, Johannes Helder<sup>7</sup>, John Wainer<sup>8</sup>, Itaru Sakata<sup>9</sup>, Nathan Garcia<sup>10</sup>, Eric Grenier<sup>1</sup> & Josselin Montarry<sup>1</sup>✉

The potato cyst nematode *Globodera rostochiensis* originates from the Andean Mountain region in South America and has unintentionally been introduced to all inhabited continents. Several studies have examined the population genetic structure of this pest in various countries by using microsatellite markers. However, merging microsatellite data produced from different laboratories is challenging and can introduce uncertainty when interpreting the results. To overcome this challenge and to explore invasion routes of this pest, we have genotyped 22 *G. rostochiensis* populations from all continents. Within populations, the highest genetic diversity was observed in the South American populations, the European populations showed an intermediate level of genetic diversity and the remaining populations were the less diverse. This confirmed pre-existing knowledge such as a first introduction event from South America to Europe, but the less diverse populations could originate either from South America or from Europe. At the continental scale, STRUCTURE genetic clustering output indicated that North America and Asia have experienced at least two introduction events. Comparing different evolutionary scenarios, the Approximate Bayesian Computation analysis showed that Europe served as a secondary distribution centre for the invasion of *G. rostochiensis* into all other continents (North America, Africa, Asia and Oceania).

Nematoda constitutes a Metazoan phylum that includes free-living species, such as the model organism *Caenorhabditis elegans*, as well as many animal- and plant-parasitic species. Plant-parasitic nematodes cause considerable economic losses in agriculture. Every crop can be infested by at least one plant-parasitic nematode species, and the worldwide crop losses due to plant-parasitic nematodes have been estimated around \$173 billion per year<sup>1</sup>. The most important plant-parasitic nematodes are root-knot nematodes (genus *Meloidogyne*) and cyst nematodes (genera *Heterodera* and *Globodera*)<sup>2</sup>. The latter are particularly difficult to eradicate because they form cysts that can persist many years in soils. Cyst nematodes are sedentary endoparasitic nematodes. They enter plant roots as second stage juveniles (J2) and establish a specialized feeding structure. Subsequently, adult males and females develop and the males leave the root to mate with females. The females continue to feed and when egg development is completed, they die and form a cyst containing hundreds of eggs, with one juvenile per egg. Being relatively small soil-borne parasites, the active dispersal ability of juveniles is limited to short distances<sup>3</sup>. However, their survival form, the cyst, can be dispersed over large spatial scale, either through natural means such as wind, water, and wildlife or through the transportation of soil and tubers due to human activities<sup>4-6</sup>.

<sup>1</sup>IGEPP, INRAE, Institut Agro, Univ Rennes, Le Rheu, France. <sup>2</sup>Department of Biological Sciences, Chuka University, Chuka, Kenya. <sup>3</sup>Julius Kühn-Institut, 38104 Braunschweig, Germany. <sup>4</sup>US Department of Agriculture, Agricultural Research Service, Robert W. Holley Center for Agriculture and Health and School of Integrative Plant Science, Cornell University, Ithaca, NY, USA. <sup>5</sup>Agriculture and Agri-Food Canada, Saint-Jean-sur-Richelieu, QC, Canada. <sup>6</sup>Nematology Research Unit, Department of Biology, Ghent University, Ghent, Belgium. <sup>7</sup>Laboratory of Nematology, Wageningen University, Wageningen, The Netherlands. <sup>8</sup>AgriBio Centre, Agriculture Victoria Research, Melbourne, Australia. <sup>9</sup>Hokkaido Agricultural Research Centre (HARC), National Agriculture and Food Research Organization (NARO), Hokkaido, Japan. <sup>10</sup>Nematology Unit, Plant Health Laboratory, ANSES, 35653 Le Rheu, France. <sup>11</sup>Indonesian Agricultural Quarantine Agency, Ministry of Agriculture, E Building 5th Floor, Jl. Harsono RM, 3 Ragunan, Jakarta 12550, Indonesia. ✉email: josselin.montarry@inrae.fr

Potato cyst nematodes (PCN) comprise two major species, *Globodera rostochiensis* and *G. pallida*, which are economically important pests of potatoes and are subject to strict regulations in many countries. *Globodera rostochiensis* was reported by EPPO (European and Mediterranean Plant Protection Organisation) as present in 76 countries around the world: 42 in Europe, 13 in Asia, 10 in America (4 in North America and 6 in South America), 8 in Africa and 3 in Oceania (EPPO Global Database, 08 September 2023). The Andean region is not only the centre of origin of potatoes but also the home to many potato pathogens, including plant-parasitic nematodes that attack this crop. *Globodera rostochiensis* is widely distributed in the Andean region: Peru, Ecuador, Colombia, Bolivia, Venezuela and Chile<sup>7</sup>. It was introduced around 1850 to Europe, probably through the importation of tubers from South America<sup>8,9</sup>. In North America, *G. rostochiensis* was first detected in New York state in 1941<sup>10</sup> and in the western part of Canada in 1962<sup>11</sup>. The pest was detected more recently, in 2006, in Quebec, the eastern part of Canada<sup>12</sup>. In Japan, it was first identified in 1972<sup>13</sup>. This nematode was also reported in Lebanon<sup>14</sup> and more recently, in several China provinces<sup>15</sup>. In Australia, *G. rostochiensis* was detected for the first time near Perth in 1986<sup>16</sup>, and this population was subsequently successfully eradicated<sup>17</sup>. However, it was detected in Victoria, near Melbourne (i.e., more than 3000 km from Perth) in 1991<sup>18</sup>. *Globodera rostochiensis* was also identified in Indonesia in 2003<sup>19</sup>. In Africa, *G. rostochiensis* was already known in the North (e.g., Algeria, Egypt and Tunisia) and in South-Africa<sup>20</sup>. In 2015, *G. rostochiensis* was first reported in East-Africa in Kenya<sup>21</sup> and later in Rwanda and Uganda<sup>22,23</sup>.

Over the past 10 years, several population genetic studies were performed utilizing microsatellite markers (polymorphic neutral DNA markers that consists of short DNA repeat units) to explore genetic diversity and relationships among *G. rostochiensis* populations. Boucher et al.<sup>24</sup> developed a set of 12 polymorphic microsatellite markers, and their results, based on the genotyping of 15 populations from South America, North America and Europe, showed that *G. rostochiensis* was introduced to North America at least twice. A few years later, three studies explored the origin of *G. rostochiensis* populations sampled in Kenya<sup>25</sup>, Australia<sup>26</sup> and Indonesia<sup>27</sup>. Using six Kenyan populations and the same 12 microsatellite markers, Mwangi<sup>25</sup> showed that the Kenyan populations formed a distinct genetic cluster separate from the populations studied by Boucher et al.<sup>24</sup>. Using nine of the 12 microsatellite markers, the results of Blacket et al.<sup>26</sup> supported a probable single introduction and indicated that the Australian populations were genetically distinct from populations previously sampled worldwide. Ten markers of the 12 developed by Boucher et al.<sup>24</sup> were found to be polymorphic in Indonesian *G. rostochiensis* populations (five populations from North Sumatra and two populations from East Java). The resulting genetic diversity and structure analyses suggested that the origin of North Sumatra cysts was East Java<sup>27</sup>, but the origin of the latter was not explored. Using whole genome resequencing data, a more recent study confirmed that *G. rostochiensis* was first introduced to Java and soon after dispersed from Java to Sumatra<sup>28</sup>.

Merging microsatellites data produced in different laboratories, i.e., using distinct fluorescent dyes, PCR conditions, *Taq* Polymerases, sequencers, and protocols to score alleles, could introduce uncertainty regarding the interpretation of the results (e.g.,<sup>29–32</sup>). When a newly sampled population was found genetically distant from previous ones, it can be challenging to distinguish the true genetic divergence from effects resulting from technical differences between laboratories. Moreover, the number of microsatellite markers used varied among the studies mentioned above.

To generate an accurate and extensive dataset on a global scale, *G. rostochiensis* populations were collected from different parts of the world and all populations were genotyped, from DNA extraction to allele-size reading, in a single laboratory. To minimize reading errors, allele sizes were identified using the automatic calling and binning procedure and completed by a manual examination by only one reader. This study aimed to explore the genetic features, genetic relationships, and the origin and routes of *G. rostochiensis* invasion worldwide using 22 populations from North and South America, Africa, Europe, Asia, and Oceania.

## Results

### Genetic features of *G. rostochiensis* populations

Using a set of 11 microsatellite markers, we identified 77 alleles among the 22 *G. rostochiensis* populations that were genotyped (i.e., among the 793 individuals without missing data), ranging from four (for Gp116, Gp118 and Gp135) to 16 (for Gr90) alleles per locus. Forty private alleles (alleles present in only one population) were detected, 37 being from South American populations.

The median number of individuals per population was 35 and the population having the lowest number of individuals was the Jap-Ho population ( $n = 29$ ) (Table 1). The South American populations were the most diverse: the allelic richness ( $A_r$ ), estimated on a reduced sample of 29 individuals, ranged from 1.72 to 3.97 alleles per locus, and the unbiased expected heterozygosity ( $H_{nb}$ ) ranged from 0.235 to 0.548 (Table 1). The highest genetic diversity was found in populations B2 and B4 from Bolivia. A low genetic diversity, with  $H_{nb}$  lower than 0.1, was observed within Kenyan and Australian populations, as well as within the two Indonesian populations from North Sumatra, the US population from New York State, the Japanese and one European population (CZ). The remaining European populations exhibited an intermediate genetic diversity: the allelic richness ranged from 1.54 to 1.70 alleles per locus and the unbiased expected heterozygosity from 0.149 to 0.155 (Table 1).

Regarding the deviation from the random mating hypothesis, 20 populations were at the Hardy–Weinberg equilibrium (i.e.,  $F_{IS}$  not significantly different to zero). Only one Kenyan population, HAR2, and the Jap-Ho population showed strong heterozygote deficits, with significant positive values of  $F_{IS}$ : 0.277 and 0.667, respectively (Table 1).

### Genetic differentiation between *G. rostochiensis* populations

Among the 231 pairwise comparisons, the  $F_{ST}$  values, ranging from 0 to 0.91 (Fig. 1), were significant except for nine population pairs comparing populations sampled in the same country: Kenya, Indonesia or Australia.

Country	Continent <sup>1</sup>	Code	n <sup>2</sup>	H <sub>nb</sub>	Ar (n = 29)	F <sub>IS</sub> <sup>3</sup>
Bolivia (Capaña)	S-A	B2	32	0.548	3.97	0.078
Bolivia (Tiraque)	S-A	B4	33	0.396	3.16	0.056
Peru (Huancane)	S-A	267	37	0.235	1.96	0.081
Chile (La Serena)	S-A	3346	40	0.250	1.72	0.002
Czech Republic (Svojše)	Europe	CZ	36	0.066	1.49	0.041
France (Dunkerque)	Europe	Dunk	46	0.155	1.61	0.032
Netherlands (Wageningen)	Europe	NL	35	0.151	1.54	0.002
Portugal (Montalegre)	Europe	Port	47	0.149	1.70	0.029
Canada, Quebec (St Amable)	N-A	Ama	38	0.147	1.43	0.043
United States (New York)	N-A	US	39	0.005	1.13	-0.000
Kenya, (Haraka)	Africa	HAR2	31	0.097	1.62	0.277 *
Kenya, (Kinangop)	Africa	KIN2	31	0.098	1.54	0.070
Kenya, (Tigoni)	Africa	TGN	32	0.065	1.35	0.129
Kenya, (Rironi)	Africa	RIR	32	0.023	1.17	-0.112
Indonesia, Sumatra (Lingga Julu)	Asia	NRK4	39	0.073	1.52	-0.025
Indonesia, Sumatra (Suka Ndebi)	Asia	NRK6	38	0.078	1.36	-0.046
Indonesia, Java (Sumber Brantas)	Asia	NRM1	39	0.100	1.41	-0.046
Indonesia, Java (Krajan)	Asia	NRM2	33	0.161	1.72	-0.046
Lebanon (Beyrouth)	Asia	Leb-Be	34	0.120	1.36	-0.067
Japan, Hokkaido (Kutchan)	Asia	Jap-Ho	29	0.009	1.18	0.667 *
Australia, Victoria (Cora Lynn)	Oceania	CL	35	0.089	1.26	0.212
Australia, Victoria (Thorpdale)	Oceania	TH	37	0.052	1.34	0.051

**Table 1.** Genetic diversity indices (H<sub>nb</sub> and Ar) and deviation from random mating (F<sub>IS</sub>) for each *Globodera rostochiensis* population. <sup>1</sup>S-A for South-America and N-A for North-America. <sup>2</sup>n is the number of genotyped individuals per population. <sup>3</sup>Stars indicate that F<sub>IS</sub> is significantly different from zero.

Overall, the genetic differentiation between the nine different defined groups, i.e., the two putative sources (South America and Europe) and the seven target populations or pools of populations (Ama, US, Kenya, Indonesia, Lebanon, Japan and Australia) was high and significant. Whereas for the F<sub>ST</sub> calculated within each defined group, two cases were distinguishable: (i) the genetic differentiation was significant among populations from South America and from Europe, (ii) the genetic differentiation was low and mainly not significant among populations from Kenya, Indonesia and Australia. This result clearly indicated that the nematode dispersion within South America and the introduction of *G. rostochiensis* into Europe were much older events than the introduction to the other parts of the world.

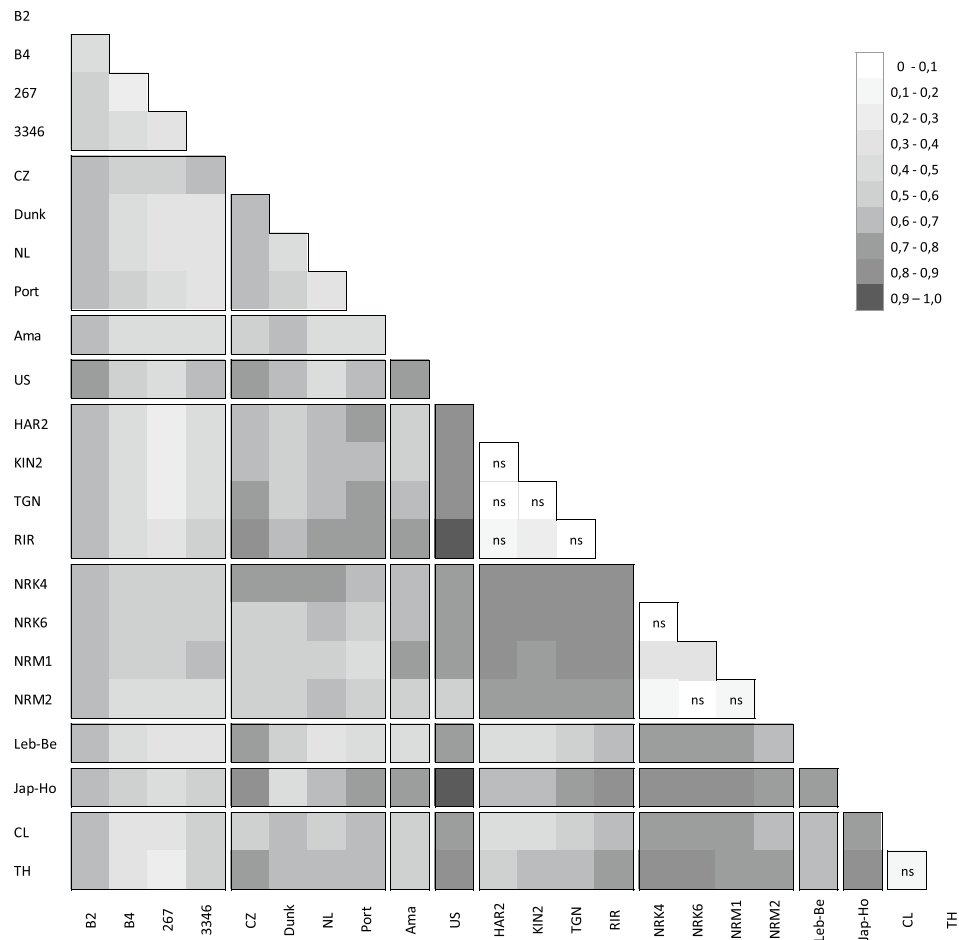
### Genetic structure among *G. rostochiensis* populations

The Bayesian clustering analysis identified K = 3 as the optimal number of genetic clusters, with a high proportion of individuals well assigned to one cluster. Seventy-six percent of individuals were assigned to one cluster with a percentage of assignment higher than 95% (54% to cluster 1, 18% to cluster 2 and 4% to cluster 3). The four Kenyan, the two Australian, the Canadian, the Lebanese, the Japanese and the Chilean populations were entirely assigned to cluster 1 (Fig. 2). All the Indonesian populations and the US population were clearly assigned to cluster 2, even if the percentage of assignment was a little smaller for individuals of the US population (Fig. 2). The European populations (CZ, Dunk, NL and Port) and the Peruvian population were mainly assigned to cluster 1 but the number of individuals with a percentage of assignment higher than 95% was quite low (e.g., only 11% of the individuals from population NL). The remaining individuals were assigned to both cluster 1 and cluster 2 (Fig. 2). The two Bolivian populations had a distinct assignment from all other populations and corresponded alone to cluster 3. B2 was completely assigned to cluster 3 (31 out 32 individuals with a percentage of assignment higher than 95%), while all the individuals of B4 were assigned either to cluster 1 or 3 (Fig. 2).

### Inferences of invasion scenarios

The Approximate Bayesian Computation (ABC) was used here to compared four evolutionary scenarios that can explain observed data (Fig. 3). The main objective was to determine whether the origin of each target population is from South America or Europe. The DIYABC results indicated, with the posterior probability of scenario S2 higher than those for the other scenarios, that all the studied populations (Ama, US, Kenya, Indonesia, Leb-Be, Jap-Ho and Australia) were originated from Europe after the initial introduction into this continent (Table 2). The 95% confidence intervals (CIs) for scenario S2 did not overlap with those for the other scenarios (Table 2). These results also ruled out the possibility that a given population diverged from a ghost population which itself originated from Europe earlier in time.

Before comparing possible scenarios, we confirmed that at least one combination of scenarios and priors can produce simulated data sets that are close enough to the observed data using the pre-evaluation of scenarios and



**Figure 1.** Matrix of pairwise  $F_{ST}$  between the 22 *Globodera rostochiensis* populations. Populations were separated according to the different groups constituted for the ABC analyses. ns indicates non-significant differentiation between two populations.

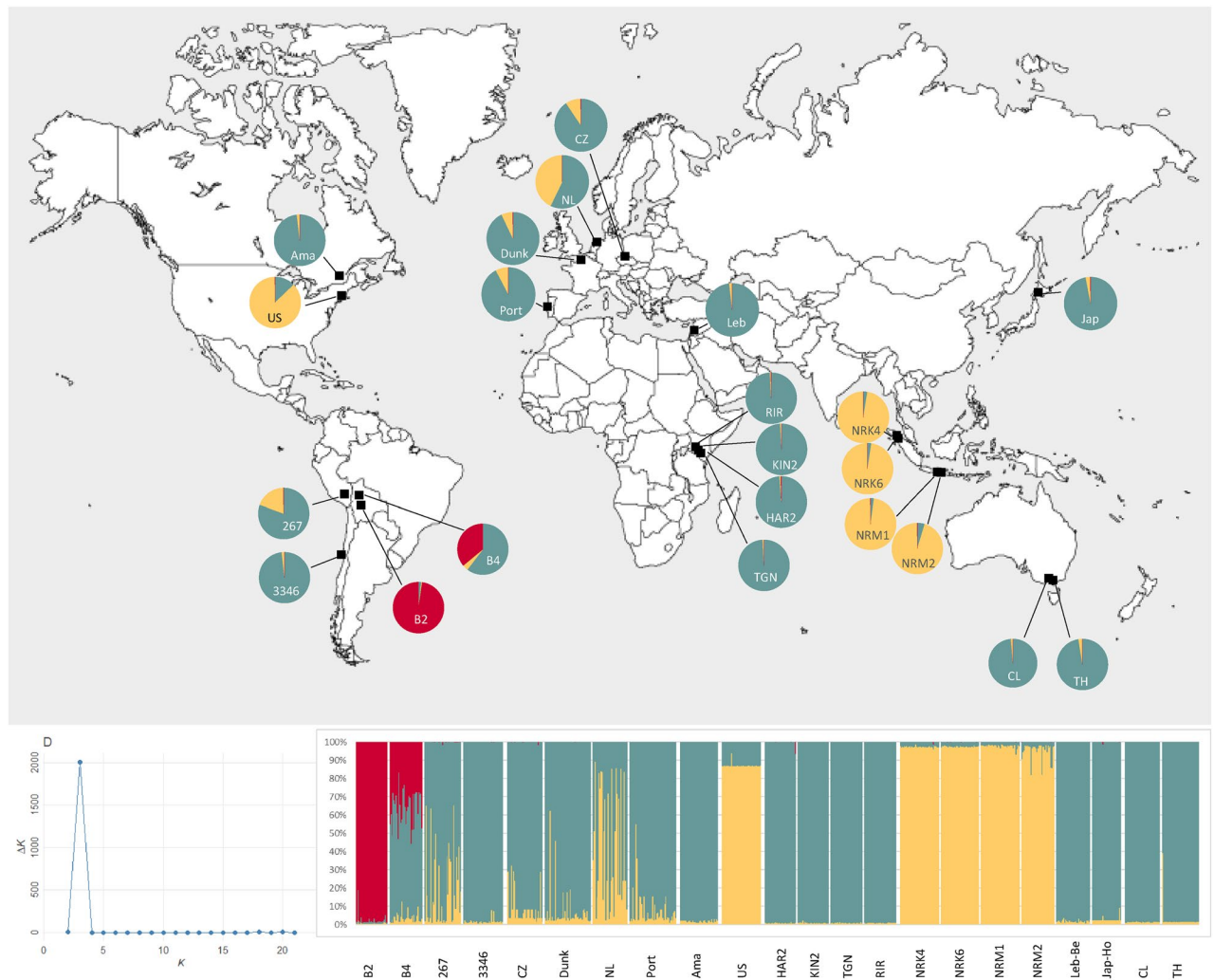
prior distributions option. This was performed through a principal component analysis (Supporting information S1a to S7a). Once the best scenario (S2) was selected according to the logistic regression test (Supporting information S1b to S7b), the discrepancy between the model-posterior combination and the observed data was evaluated for each target population. Although we obtained many summary statistics on the tails of distributions, all the seven principal component analyses (PCA) showed that posterior values were close to the observed dataset (Supporting information S1c to S7c). We concluded that the adequacy of the model-posterior combination sufficed to correctly explain the observed dataset.

## Discussion

Because it is not always possible to compare/merge microsatellite data produced by different laboratories, due to subtle differences in the protocols used<sup>31</sup>, we decided to collect and genotype, using a fully standardized protocol, 22 populations of the potato cyst nematode *G. rostochiensis* originating from all inhabited continents. Our results confirmed the South American origin for this species and its initial introduction into Europe. Moreover, our analyses provided support for the hypothesis that populations from North America, Africa, Asia and Oceania most likely originated from Europe. This suggests that Europe served as a bridgehead in the worldwide invasion history of *G. rostochiensis*.

Among the 22 genotyped *G. rostochiensis* populations, only two (HAR2 from Kenya and Jap-Ho from Japan) showed a heterozygote deficit ( $F_{IS} > 0$ ). This finding is surprising in the context of cyst nematodes, as heterozygote deficits were frequently found in populations of *G. pallida*<sup>4</sup>, *Heterodera schachtii*<sup>5,33</sup>, *H. glycines*<sup>34</sup> and *H. carotae*<sup>35,36</sup>. The heterozygote deficit in cyst nematodes was attributed to the low active dispersal ability of juveniles leading to inbreeding<sup>37</sup>. Future research is needed to explain this particular genetic feature of *G. rostochiensis* populations, which suggests the possibility of a significant cost of consanguineous mating or capacity of the females of this species to preferentially mate with non-sibling males.

In species introduction events, the level of genetic diversity allows distinguishing source and sink populations: an introduced population being less diverse than a native one due to the bottleneck effect resulting from the introduction of a limited number of individuals. Our results confirm that the potato cyst nematode *G. rostochiensis* is

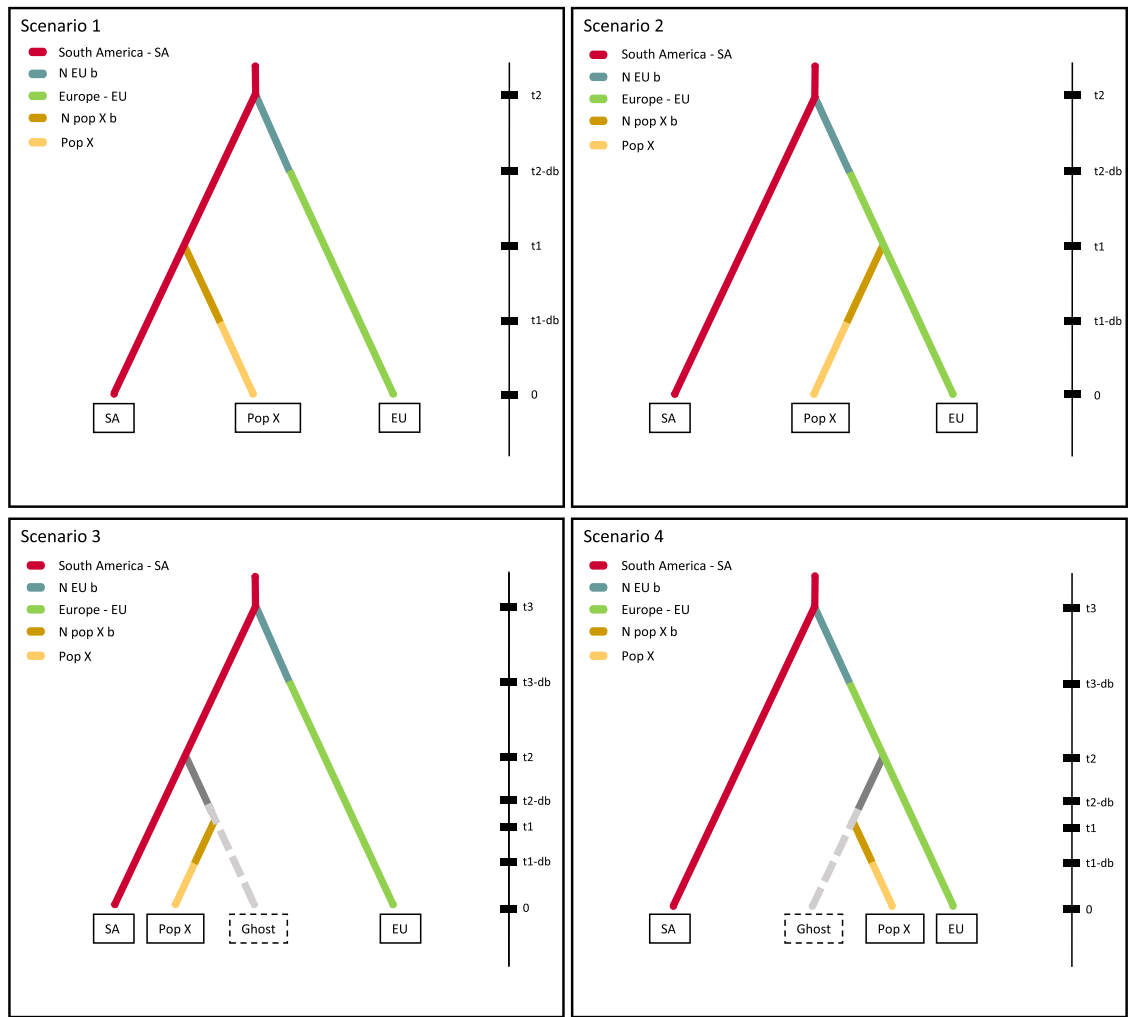


**Figure 2.** Bayesian clustering analysis (STRUCTURE) of the 793 *Globodera rostochiensis* individuals from the 22 populations collected in different continents. Assignment probabilities of individuals are presented for  $K = 3$ , i.e., the most likely number of clusters statistically determined using  $\Delta K$  method (bottom-left graph). On the bottom-right graph, each vertical line represents an individual for which the genetic assignment is partitioned into three clusters, represented by three colors (cluster 1 in green, cluster 2 in yellow and cluster 3 in red), and vertical white lines separate each of the 22 populations. The map at the top shows the geographical location of the 22 *G. rostochiensis* populations and their membership proportion of clusters.

originated from South America<sup>8,38,39</sup>, as the four *G. rostochiensis* populations from South America were the most genetically diverse ( $H_{nb} > 0.235$ ; Table 1). All the other populations were introduced/sink populations, but the less genetically diverse populations could originate either from South America or from Europe, as the European populations showed an intermediate level of genetic diversity (Table 1).

Based on the matrix of pairwise  $F_{ST}$  (Fig. 1) and the genetic clustering output (Fig. 2), we can hypothesise that Indonesian, Kenyan and Australian *G. rostochiensis* populations, which were characterized by a strong genetic homogeneity and low  $F_{ST}$  among populations from the same geographic area, were probably originated respectively from a single introduction event and subsequently spread within each respective country. Moreover, the *G. rostochiensis* populations from North America were shown to belong to two distinct genetic clusters, confirming that the nematode was introduced twice into this continent<sup>24</sup>. The same pattern was observed in Asia, with at least two distinct introduction events as all four Indonesian populations were assigned to cluster 2 while the remaining Asian populations (Jap-Ho from Japan and Leb-Be from Lebanon) were assigned to the genetic cluster 1 (Fig. 2).

The global distribution of the nematode leads to a large number of possible evolutionary scenarios among *G. rostochiensis* populations. To reduce the number of scenarios to be tested, we grouped populations according to their geographical origins and pre-existing historical knowledge. The European populations were unintentionally introduced from South America, after the mid-nineteenth century, due to the importation of potato tubers<sup>8,9</sup>. Since this period, and due to increased trade between Europe and the rest of the world, Europe likely became a second significant source for the dispersion of *G. rostochiensis* worldwide. Consequently, South American populations (Bolivia, Peru and Chile) were grouped together to represent the native area of *G. rostochiensis*, as well as European populations to represent a likely second significant region of origin for its worldwide dispersion.



**Figure 3.** The four population divergence scenarios tested by the Approximate Bayesian Computation (ABC) approach using each target population (Canada, United States, Kenya, Indonesia, Lebanon, Japan and Australia). Since an invasive population generally starts with a few individuals, an initial size reduction (noted here ‘N pop X b’ for each target population and ‘N EU b’ for population Europe) was added. Times are not to scale, 0 indicates the present day and t on top of the bar indicates the oldest split, bd illustrates the duration of the bottleneck period. A ghost population is a possible unsampled source population.

Competing scenario	Origin	Canada (1 population)	USA (1 population)	Kenya (4 populations)	Indonesia (4 populations)
S1	South America	0.00 [0.00, 0.01]	0.01 [0.00, 0.01]	0.02 [0.00, 0.48]	0.00 [0.00, 0.01]
S2	Europe	<b>0.95 [0.94, 0.96]</b>	<b>0.95 [0.94, 0.96]</b>	<b>0.87 [0.82, 0.93]</b>	<b>0.92 [0.89, 0.94]</b>
S3	South America and unsampled source population	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	0.01 [0.00, 0.43]	0.00 [0.00, 0.00]
S4	Europe and unsampled source population	0.04 [0.03, 0.05]	0.04 [0.03, 0.05]	0.10 [0.00, 0.49]	0.08 [0.05, 0.10]
Competing scenario	Origin	Lebanon (1 population)	Japan (1 population)	Australia (2 populations)	
S1	South America	0.01 [0.01, 0.02]	0.01 [0.00, 0.01]	0.01 [0.00, 0.32]	
S2	Europe	<b>0.94 [0.93, 0.95]</b>	<b>0.95 [0.92, 0.97]</b>	<b>0.94 [0.92, 0.96]</b>	
S3	South America and unsampled source population	0.00 [0.00, 0.01]	0.00 [0.00, 0.00]	0.00 [0.00, 0.30]	
S4	Europe and unsampled source population	0.04 [0.03, 0.05]	0.04 [0.02, 0.07]	0.05 [0.00, 0.34]	

**Table 2.** Results of the ABC analyses to infer the origin of each target genetic unit. Scenarios S1 and S3 tested a South American origin and scenarios S2 and S4 a European origin. The effect of a ghost population was tested by scenarios S3 and S4. The table shows the posterior probability of each scenario, with 95% confidence intervals in brackets, and bold numbers indicate the most highly supported scenario for each target unit.

Due to low  $F_{ST}$  values observed among populations from the same country, the four Kenyan populations were pooled together, as well as the four Indonesian populations and the two Australian populations. This allowed us to elucidate the invasion history of *G. rostochiensis* by evaluating two alternative hypotheses for each target population: the population was introduced from South America, the centre of origin of PCN, or from Europe after its initial establishment into this continent. Our results clearly showed that all the target populations, collected in Canada, the USA, Kenya, Indonesia, Lebanon, Japan and Australia, were all introduced from Europe (Table 2 and Supporting information). Means of transportation of the nematode are largely unknown and could be very diverse. In the case of introduction from South America to Europe it is highly possible that nematode cysts were distributed with the imported potato tubers from South America<sup>9,38</sup>. Spears<sup>38</sup> also suggested that in the case of the introduction in the USA, military equipment brought back from Europe after World War I may have been a possible means of nematode invasion. Our results support the likelihood of such a scenario, although we still lack specific evidence about the actual means of nematode transportation. Clearly other means of long-distance transportation exist, including contaminated soil within shipments of bulbs, ornamental plants, etc.<sup>40</sup>. In Japan, an unexpected pathway was discovered through imported Peruvian guano contaminated with viable nematode cysts and subsequently used as a fertilizer<sup>41</sup>. This discovery led the authors to suggest that *G. rostochiensis* was introduced to Japan from Peru. Our results, pointing to a European origin, suggested that various transportation means might have occurred in Japan. Since we have only examined one Japanese population here, further research is necessary to investigate the genetic diversity among different *G. rostochiensis* populations in Japan.

A strong bottleneck occurred from South America to Europe, leading to much lower genetic diversity in Europe than in the native area of *G. rostochiensis*, and Europe subsequently served as the source for waves of invasions worldwide. The important role of Europe in worldwide dispersal has been reported for many species, pathogenic or not. For instance, the grapevine downy mildew, *Plasmopara viticola*, which originated in North America, was first introduced into Europe from where it further invaded different continents<sup>42</sup>. Furthermore, the mealybug *Pseudococcus viburni*, an insect infecting a large number of plant families, had a similar mode of global spread to that of *G. rostochiensis*: Europe was initially invaded from South America and later became the main source of worldwide spread<sup>43</sup>. Generally, the bridgehead invasion scenario, where an introduced population serves as the source of other invasive populations, applies to many cases of agricultural pest introductions<sup>44</sup>.

The low  $F_{ST}$  measured among *G. rostochiensis* populations from the same geographical area (Indonesia, Kenya and Australia) suggested that in each country these populations recently diverged from their common ancestor (a recent introduction event) or that gene flow still remain strong between these populations in each country. The latter is less probable due to the low dispersal abilities of nematodes and the strict regulatory measures applied in many countries. Although introductions in Indonesia, Kenya and Australia occurred more recently compared to that in Europe, we cannot determine the timing of introduction events for the populations Ama, US, Leb-Be and Jap-Ho populations since we have genotyped only one population in these countries. However, the different dates of first detection also suggested that the introduction was more recent than the initial one in Europe. ABC analyses could have allowed us to estimate the number of generations since the introduction event. However, due to the variations in the collection times of different populations and the variable number of generations per year for *G. rostochiensis* based on environmental conditions, we were unable to perform this type of estimation accurately.

Quarantine status of some plant pests is part of crop protection strategies and has generally not received the scientific support it deserves. Our results support the idea that the quarantine status of *G. rostochiensis*, a pathogen with a narrow host range, has effectively prevented its further spread from South America to countries outside Europe. Using sequences of mitochondrial genes, Subbotin et al.<sup>45</sup> suggested that the centre of origin for *G. rostochiensis* could be in the south of Bolivia or north-west Argentina. Our results, which show that Bolivian populations are the most diverse and the only ones assigned to cluster 3, confirm this putative centre of origin and may be used to guide quarantine regulations in efforts to prevent further invasions by genetically distant and diverse *G. rostochiensis* populations. New invasions from its native area could significantly increase diversity and facilitate nematode adaptation to chemical nematicides and resistant potato cultivars (e.g.,<sup>46</sup>). Regarding the last recent introductions, we cannot rule out the possibility that in some cases the introduction occurred before the implementation of quarantine measures in the country. Moreover, each introduction of an exotic pathogen is not always successful due to the absence of suitable establishment conditions. But, as the frequency of introduction events increases, the likelihood of getting the right conditions also rises. As growers worldwide seek for the best cultivars to compete globally, the international transport of seed tubers is expected to increase. Importing material from approved suppliers with processes to minimize pests and diseases incidence in seed tuber fields could be an effective way to further reduce the risk of new introductions in addition to the quarantine status and sole proof of pest absence in the field of origin of the plant material.

## Conclusion

Microsatellite-based genotyping of *Globodera rostochiensis* populations from all continents allowed for a reconstruction of invasion routes at a worldwide scale. This collaborative population genetic study confirms that this plant-parasitic nematode species evolved in the Andean region of South America and was first introduced into Europe. Genetically less diverse populations sampled in North America, Africa, Asia and Oceania, most likely originate from Europe, which served as a secondary distribution hub for the worldwide invasion of *G. rostochiensis*. These insights are not only essential to prevent further introductions of genetically diverse populations but also to enable the development of effective control strategies. Control methods proven successful within the European secondary hub could be extrapolated to effectively manage *G. rostochiensis* populations in other global regions stemming from this hub.



## Methods

### *Globodera rostochiensis* populations

In this study, 22 *G. rostochiensis* populations collected from different parts of the world (Supporting information Table S1) were genotyped using microsatellite markers. Seventeen populations were selected based on four previously published studies, and five new populations were added. We selected seven populations among the 15 used by Boucher et al.<sup>24</sup>: populations B2 and B4 from Bolivia, 267 from Peru, 3346 from Chile, Ama from Canada, Dunk from France, and Port from Portugal. Two Australian populations were chosen and obtained from Blacket et al.<sup>26</sup>: Cora Lynn (CL) and Thorpdale (TH). Four Indonesian *G. rostochiensis* populations were also chosen and obtained from Handayani et al.<sup>27</sup>: two from North Sumatra (NRK-4 and NRK-6) and two from East Java (NRM-1 and NRM-2). In Kenya, we obtained four populations from Mwangi et al.<sup>25</sup>: three from Nyandarua (HAR2, KIN2, RIR) and one from Kiambu (TGN). Finally, the remaining five new populations of the dataset included: a Lebanese population (Leb-Be), a Czech population (CZ), a Dutch population (NL), a North American population from New York state (US), and a Japanese population from Hokkaido (Jap-Ho).

### Microsatellite genotyping

Total genomic DNA was extracted from single individual juveniles as described by Boucher et al.<sup>24</sup>. For each population, one second-stage juvenile (J2) was isolated per cyst and a total of 35–50 J2s from 35–50 distinct cysts were used for DNA extraction. DNA quality was validated by PCR amplification of an ITS fragment according to Thiéry and Mugniéry<sup>47</sup>. DNA samples showing a positive amplification of the ITS marker were then processed for microsatellite PCR amplification and genotyping. A set of 11 microsatellite markers developed by Boucher et al.<sup>24</sup> was used in three multiplex combinations. The marker Gp116 used by Boucher et al.<sup>24</sup> was discarded because of the presence of a second microsatellite motif in the sequence<sup>26</sup>. All the primers were synthesized at Thermo Fisher Scientific. The panel 1 included Gr50 (6FAM), Gp109 (NED), Gp126 (PET) and Gp135 (VIC). The panel 2 included Gr85 (VIC), Gr96 (6FAM) and Gp118 (PET). The panel 3 included Gr67 (6FAM), Gr75 (NED), Gr90 (VIC) and Gr91 (PET). PCR reagents, volume and cycling conditions used were as described by Boucher et al.<sup>24</sup>. PCR multiplex was performed on a 96-well Thermal Cycler (Applied Biosystems). PCR products were then diluted to 1:40 in sterile water, and 3  $\mu$ L of this dilution was mixed with 0.05  $\mu$ L of GeneScan 500 LIZ Size Standard (Applied Biosystems) and 5  $\mu$ L of formamide (Applied Biosystems). Genotyping was performed on ABI 3730XL sequencer (Applied Biosystems) at the Gentyane INRAE platform. Allele sizes were identified using the automatic calling and binning procedure of GeneMapper<sup>®</sup> v 5.0 (Thermo Fisher Scientific) and completed by a manual examination.

### Analyses of population genetic structure and diversity

To explore the genetic diversity and structure of the 22 *G. rostochiensis* populations, 840 individual juveniles were genotyped. Data analyses were done on a reduced dataset free of any missing data (i.e., on 793 individuals). Allelic richness ( $A_r$ ) was estimated using the rarefaction method implemented in POPULATIONS 1.2.32<sup>48</sup>, which estimated the mean number of alleles per locus for a reduced sample size. An unbiased estimate of gene diversity ( $H_{nb}$  according to Nei<sup>49</sup>) and deviation from random mating ( $F_{IS}$ ) were computed using GENETIX 4.05.2<sup>50</sup>. The statistical significances of  $F_{IS}$  were estimated using the allelic permutation method (10,000 permutations) implemented in GENETIX.

The differentiation coefficients between each pair of populations ( $F_{ST}$ ) were computed using GENEPOP 4.5.1 according to Weir and Cockerham<sup>51</sup>, and their statistical significances were estimated by 5000 random permutations of individuals among populations. A Bonferroni adjustment was applied to take into account multiple testing, i.e.,  $\alpha = 0.05$  was lowered to  $\alpha = 0.00022$  for 231 comparisons ( $22 \times 21/2$ ).

Genetic structure of the 22 *G. rostochiensis* populations was explored using the Bayesian clustering algorithms implemented in STRUCTURE 2.3.4<sup>52,53</sup>. For each number of genetic clusters  $K$  (from 1 to 22), ten independent runs were executed using the admixture model, uncorrelated allele frequency and the default priors except for alpha, for which the value was set to 0.0454 (i.e.,  $1/p$ ,  $p$  being the number of populations) following the recommendations of Wang<sup>54</sup>. The initial burn-in period consisted of 1,000,000 iterations and the number of Markov Chain Monte Carlo (MCMC) repetitions was 3,000,000. Structure Harvester Web ver.0.6.94<sup>55</sup> was applied to determine the most likely number of clusters statistically determined using the ad-hoc Evanno statistic  $\Delta K$ <sup>56</sup>. The ten independent replicates for the optimal value of  $K$  were then merged using CLUMPP version 1.1.2<sup>57</sup>.

### Inferences of global invasion history

To investigate the most likely pathways for establishment of *G. rostochiensis* populations in the different parts of the world, Approximate Bayesian Computation (ABC) methods implemented in DIYABC 2.1.0<sup>58</sup> were used. Based on historical knowledge of the origins of *G. rostochiensis*, we excluded the hypothesis that some populations were native except for the South American populations (Bolivia, Peru and Chile) and we draw for each target population (Canada, United States, Kenya, Indonesia, Lebanon, Japan and Australia) two origins and main route hypotheses. The first hypothesis is a direct South American origin, the second hypothesis is an origin from Europe after the initial introduction into this continent.

To reduce the number of scenarios, we grouped populations into pools. Five pools of populations were constituted: South America (populations B2, B4, 267 and 3346), Europe (populations Dunk, Port, NL and CZ), Kenya (populations HAR2, KIN2, TGN and RIR), Indonesia (populations NRK4, NRK6, NRM2 and NRM1) and Australia (populations CL and TH). It is noteworthy that North American populations Ama (Canada) and US (USA) were not pooled together because the results of Boucher et al.<sup>24</sup> showed that they were probably resulting from two distinct introduction events in North America.

For each of the seven target populations, Ama (Canada, North America), US (USA, North America), Kenya (East Africa), Indonesia (Southeast Asia), Leb-Be (Lebanon, West Asia), Jap-Ho (Japan, East Asia) and Australia (Oceania), four scenarios were constructed reflecting its possible origin. Scenarios S1 and S3 tested a direct South American origin while the European origin was tested by scenarios S2 and S4. All scenarios were constrained by the fact that the oldest demographic event corresponds to the split between South America and Europe. Both possible origins were tested in a tree topology integrating (scenarios S3 and S4) or not integrating (scenarios S1 and S2) the effect of a possible unsampled source population (i.e., a ghost population), itself introduced from South America or from Europe. Divergence events were followed by a bottleneck representing the signal of an introduction event (Fig. 3).

We formalized the four scenarios, prior distributions and computed summary statistics for each target unit. Seven reference tables were built under the Generalized Stepwise Mutation model with (i) mean mutation rate ranging from 0.0001 to 0.001 and uniform prior distribution, (ii) mean parameter of the geometric distribution (mean P) ranging from 0.1 to 0.3 and uniform prior distribution and (iii)  $t_2 > t_1$  and  $t_3 > t_2$  (Fig. 3). Three single sample statistics were computed: the mean number of alleles per locus, the mean gene diversity across loci<sup>49</sup> and the mean allele size variance across loci. For each pair of populations, the pairwise  $F_{ST}$  values<sup>51</sup>, the mean index of classification<sup>59,60</sup> and the distance between two samples<sup>61</sup> were computed. 4,000,000 simulations datasets, i.e., 1,000,000 simulations datasets per scenario, were run.

Possible scenarios were compared through the computation of the posterior probabilities of each scenario. The relative posterior probabilities of the competing scenarios were assessed by a logistic regression estimate. We also used the option ‘model checking’ with PCA in DIYABC, using all summary statistics, to evaluate how well the best scenario and parameter posterior distributions combination fit the observed data.

### Additional information

The different *Globodera rostochiensis* populations were collected by the different co-authors who obtained the required authorizations. They provided to INRAE the biological material following official endorsed Letter Of Authorization necessary for shipment of quarantine organisms between laboratories. The INRAE laboratory is accredited to hold and perform experiments on this quarantine organism. Authors declare that the cyst nematode species, *Globodera rostochiensis*, sampled in the present work is not concerned by both the IUCN Policy Statement on Research Involving Species at Risk of Extinction and the Convention on the Trade in Endangered Species of Wild Fauna and Flora.

### Data availability

The file (Esquibet\_Gros.txt) containing the genotypic data (Genepop format) for the 22 *Globodera rostochiensis* populations is available at data.inrae.fr (<https://doi.org/10.57745/3INH20>).

Received: 18 October 2023; Accepted: 11 June 2024

Published online: 17 June 2024

### References

1. Elling, A. A. Major emerging problems with minor *Meloidogyne* species. *Phytopathology* **103**, 1092–1102 (2013).
2. Jones, J. T. *et al.* Top 10 plant-parasitic nematodes in molecular plant pathology. *Mol. Plant Pathol.* **14**, 946–961 (2013).
3. Wallace, H. R. The dynamics of nematode movement. *Annu. Rev. Phytopathol.* **6**, 91–114 (1968).
4. Picard, D., Plantard, O., Scurrah, M. & Mugniéry, D. Inbreeding and population structure of the potato cyst nematode (*Globodera pallida*) in its native area (Peru). *Mol. Ecol.* **13**, 2899–2908 (2004).
5. Plantard, O. & Porte, C. Population genetic structure of the sugar beet cyst nematode *Heterodera schachtii*: A gonochoric and amphimictic species with highly inbred but weakly differentiated populations. *Mol. Ecol.* **13**, 33–41 (2004).
6. Alenda, C., Montarry, J. & Grenier, E. Human influence on the dispersal and genetic structure of French *Globodera tabacum* populations. *Infect. Genet. Evol.* **27**, 309–317 (2014).
7. Silvestre, R., Dandurand, L. M., Zasada, I. A., Franco, J. & Kuhl, J. C. An assessment of potato cyst nematode (*Globodera* spp.) research from the Andean region of South America. Part 1: Occurrence and impact. *Nematropica* **51**, 85–105 (2021).
8. Evans, K., Franco, J. & De Scurrah, M. M. Distribution of species of potato cyst-nematodes in South America. *Nematologica* **21**, 365–369 (1975).
9. Grenier, E., Fournet, S., Petit, E. & Anthoine, G. A cyst nematode ‘species factory’ called the Andes. *Nematology* **12**, 163–169 (2010).
10. Cannon, O. S. *Heterodera schachtii* found in a Long Island potato field. *Plant Dis. Rep.* **25**, 408 (1941).
11. Olsenand, O. A. & Mulvey, R. H. The discovery of golden nematode in Newfoundland. *Can. Plant Dis. Surv.* **42**, 253 (1962).
12. Mahran, A. *et al.* The golden potato cyst nematode *Globodera rostochiensis* pathotype Ro1 in the Saint-Amable regulated area in Quebec, Canada. *Plant Dis.* **94**, 1510 (2010).
13. Yamada, E., Takakura, S. & Tezuka, H. On the occurrence of the potato cyst-nematode, *Heterodera rostochiensis* Wollenweber, in Hokkaido, Japan. *Jpn. J. Nematol.* **2**, 12–15 (1972).
14. Ibrahim, S. K., Saad, A. T., Haydock, P. P. J. & Al-Masri, Y. Occurrence of the potato cyst nematode *Globodera rostochiensis* in Lebanon. *Nematology* **2**, 125–128 (2000).
15. Jiang, R. *et al.* First record of the golden potato nematode *Globodera rostochiensis* in Yunnan and Sichuan provinces of China. *J. Integr. Agric.* **21**, 898–899 (2022).
16. Stanton, J. M. First record of potato cyst nematode, *Globodera rostochiensis*, in Australia. *Australas. Plant Pathol.* **15**, 87 (1986).
17. Collins, S. J., Marshall, J. M., Zhang, X. H. & Vanstone, V. A. Area freedom from *Globodera rostochiensis* in Western Australia. *Appl. Biol.* **103**, 55–62 (2010).
18. Guy, G., Woodward, J. & Hinch, J. M. *Globodera rostochiensis* and possibly *G. pallida* in Australia. *J. Nematol.* **24**, 594 (1992).
19. Mulyadi, M., Rahayu, B., Triman, B. & Indarti, S. Identification of golden potato cyst nematode (*Globodera rostochiensis*) in Batu, East Java. *J. Perlind. Tanam. Indones.* **9**, 46–53 (2003).
20. Kleynhans, K. P. N. Potato cyst nematodes (*Globodera* species) in Africa. In *Potato Cyst Nematodes. Biology, Distribution and Control* (eds Marks, R. J. & Brodie, B. B.) 347–351 (CABI Publishing, 1998).
21. Mwangi, J. M., Kariuki, G. M., Waceke, J. W. & Grundler, F. M. First report of *Globodera rostochiensis* infesting potatoes in Kenya. *New Dis. Rep.* **31**, 18 (2015).

22. Niragire, I., Couvreur, M., Karssen, G., Uwumukiza, B. & Bert, W. First report of potato cyst nematode (*Globodera rostochiensis*) infecting potato (*Solanum tuberosum* L.) in Rwanda. *Plant Dis.* **409**, 19–21 (2019).
23. Cortada, L. *et al.* First report of potato cyst Nematode, *Globodera rostochiensis*, infecting potato (*Solanum tuberosum*) in Uganda. *Plant Dis.* **104**, 3082 (2020).
24. Boucher, A. C. *et al.* Genetic diversity of the golden potato cyst nematode *Globodera rostochiensis* and determination of the origin of populations in Quebec, Canada. *Mol. Phylogenet. Evol.* **69**, 75–82 (2013).
25. Mwangi, J. M. Resistance based integrated pest management strategy for *Globodera rostochiensis* and *Globodera pallida* in potato cropping systems. PhD Thesis, University of Kassel. <https://doi.org/10.17170/kobra-20191212863> (2019).
26. Blacket, M. J. *et al.* Molecular assessment of the introduction and spread of potato cyst nematode, *Globodera rostochiensis*, in Victoria, Australia. *Phytopathology* **109**, 659–669 (2019).
27. Handayani, N. D. *et al.* Distribution, DNA barcoding and genetic diversity of potato cyst nematodes in Indonesia. *Eur. J. Plant Pathol.* **158**, 363–380 (2020).
28. Handayani, N. D. *et al.* Genomic reconstruction of the introduction and diversification of golden potato cyst nematode populations in Indonesia. *Phytopathology* **112**, 396–403 (2022).
29. Delmotte, F., Leterme, N. & Simon, J. C. Microsatellite allele sizing: Difference between automated capillary electrophoresis and manual technique. *BioTechniques* **31**, 810–818 (2001).
30. Hahn, M., Wilhelm, J. & Pingoud, A. Influence of fluorophore dye labels on the migration behaviour of polymerase chain reaction—Amplified short tandem repeats during denaturing capillary electrophoresis. *Electrophoresis* **22**, 2691–2700 (2001).
31. Vignal, A., Milan, D., San Cristobal, M. & Eggen, A. A review on SNP and other types of molecular markers and their use in animal genetics. *Genet. Sel. Evol.* **34**, 275–305 (2002).
32. Selkoe, A. K. & Toonen, R. J. Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* **9**, 615–629 (2006).
33. Kim, J. *et al.* Phylogeography of the highly invasive sugar beet nematode, *Heterodera schachtii* (Schmidt, 1871), based on microsatellites. *Evol. Appl.* **12**, 324–336 (2019).
34. Wang, H. M., Zhao, H. H. & Chu, D. Genetic structure analysis of populations of the soybean cyst nematode, *Heterodera glycines*, from north China. *Nematology* **17**, 591–600 (2015).
35. Gautier, C. *et al.* Microsatellite markers reveal two genetic groups in European populations of the carrot cyst nematode *Heterodera carotae*. *Infect. Genet. Evol.* **73**, 81–92 (2019).
36. Esquibet, M. *et al.* Evidence of strong gene flow among French populations of the carrot cyst nematode *Heterodera carotae*. *Plant Pathol.* **69**, 168–176 (2020).
37. Montarry, J. *et al.* Heterozygote deficits in cyst plant-parasitic nematodes: Possible causes and consequences. *Mol. Ecol.* **24**, 1654–1667 (2015).
38. Spears, J. F. The golden nematode. Handbook. United States Department of Agriculture, Agriculture Handbook No. 353 (1968).
39. Evans, K. & Stone, A. R. A review of the distribution and biology of the potato cyst-nematodes *Globodera rostochiensis* and *G. pallida*. *Trop. Pest Manag.* **23**, 178–189 (1977).
40. Steiner, G., Taylor, A. L. & Cobb, G. S. Cyst forming plant parasitic nematodes and their spread in commerce. *Proc. Helm. Soc. Wash.* **18**, 13–18 (1951).
41. Inagaki, H. & Kegasawa, K. Discovery of the potato cyst nematode, *Heterodera rostochiensis* Wollenweber, 1923, (Tylenchida: Heteroderidae) from Peru guano. *Appl. Entomol. Zool.* **8**, 97–102 (1973).
42. Fontaine, M. C. *et al.* Europe as a bridgehead in the worldwide invasion history of the grapevine downy mildew, *Plasmopara viticola*. *Curr. Biol.* **31**, 2155–2166 (2021).
43. Correa, M. C. G. *et al.* European bridgehead effect in the worldwide invasion of the obscure mealybug. *Biol. Invas.* **21**, 123–136 (2019).
44. Guillemaud, T., Ciosi, M., Lombaert, E. & Estoup, A. Biological invasions in agricultural settings: Insights from evolutionary biology and population genetics. *C. R. Biol.* **334**, 237–246 (2011).
45. Subbotin, S. A. *et al.* DNA barcoding, phylogeny and phylogeography of cyst nematode species from the genus *Globodera* (Tylenchida: Heteroderidae). *Nematology* **22**, 269–297 (2020).
46. McDonald, B. A. & Linde, C. Pathogen population genetics, evolutionary potential, and durable resistance. *Annu. Rev. Phytopathol.* **40**, 349–379 (2002).
47. Thiéry, M. & Mugniery, D. Interspecific rDNA restriction fragment length polymorphism in *Globodera* species, parasites of Solanaceous plants. *Fundam. Appl. Nematol.* **19**, 471–479 (1996).
48. Langella, O. Populations 1.2.31. Population Genetic Software (Individuals or Populations Distances, Phylogenetic Trees) 2012. <http://bioinformatics.org/~tryphon/populations/> (1999).
49. Nei, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583–590 (1978).
50. Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. *GENETIX 405, logiciel sous Windows™ pour la génétique des populations* (Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, 2004).
51. Weir, B. S. & Cockerham, C. C. Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370 (1984).
52. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
53. Falush, D., Stephens, M. & Pritchard, J. K. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587 (2003).
54. Wang, J. The computer program structure for assigning individuals to populations: Easy to use but easier to misuse. *Mol. Ecol. Resour.* **17**, 981–990 (2017).
55. Earl, D. A. & von Holdt, B. M. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361 (2012).
56. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).
57. Jakobsson, M. & Rosenberg, N. A. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806 (2007).
58. Cornuet, J. M. *et al.* DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **30**, 1187–1189 (2014).
59. Rannala, B. & Mountain, J. L. Detecting immigration by using multilocus genotypes. *Proc. Natl. Acad. Sci. USA* **94**, 9197–9201 (1997).
60. Pascual, M. *et al.* Introduction history of *Drosophila subobscura* in the New World: A microsatellite based survey using ABC methods. *Mol. Ecol.* **16**, 3069–3083 (2007).
61. Goldstein, D. B., Linares, A. R., Cavalli-Sforza, L. L. & Feldman, M. W. An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**, 463–471 (1995).

## Acknowledgements

We acknowledge Gentyane INRAE platform for running the genotyping from the PCRs and providing the raw data for microsatellite genetic analyses.

## Author contributions

ME performed the experiments according to a protocol elaborated jointly by JM and EG and approved by all authors. JMM and SK described and provided the Kenyan populations. XW described and provided the US population. BM described and provided the population from Quebec, Canada. NDH, WB and JH described and provided the Indonesian populations. JW described and provided the Australian populations, IS provided the Japanese population, NG provided the CZ population. EG and JM provided the remaining populations from South America, Europe and Lebanon. ME and JM analysed the data and prepared the figures. All the authors wrote the text and have made substantial contributions to the interpretation of data. All authors have approved the current version.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-64617-0>.

**Correspondence** and requests for materials should be addressed to J.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024