Microsatellite markers reveal two genetic groups in European populations of the carrot cyst nematode Heterodera carotae

Camille Gautier, Magali Esquibet, Sylvain Fournet, Christophe Piriou, Jean-Claude Yvin, Eric Nguema-Ona, Eric Grenier, Josselin Montarry

Microsatellite markers reveal two genetic groups in European populations of the carrot cyst nematode *Heterodera carotae*

GAUTIER Camille[1,2], ESQUIBET Magali[1], FOURNET Sylvain[1], PIRIOU Christophe[1], YVIN Jean-Claude[2], NGUEMA-ONA Eric[2], GRENIER Eric[1] and MONTARRY Josselin[1]

[1] IGEPP, INRA, Agrocampus-Ouest, Université de Rennes 1, 35650, Le Rheu, France

[2] Centre Mondial de l'Innovation-Laboratoire de Nutrition Végétale Pôle Biocontrôle, Groupe Roullier, Saint Malo, France

*Corresponding author: gautier.camille@orange.fr

SUMMARY

The cyst nematode *Heterodera carotae*, which parasitizes carrot roots, has been recorded in many countries in Europe (Italy, The Netherlands, Switzerland, France, Denmark, …), in South Africa and in North America (Canada, USA). To date, there is a lack of knowledge about the genetic structure of the populations of this economically important nematode. The aim of this work was to study the structuration of the genetic diversity of the carrot cyst nematode at the European scale. We have developed a set of thirteen polymorphic microsatellite markers and used it to genotype seventeen European populations of *H. carotae* coming from France, Switzerland, Italy, Denmark and one non-European population from Canada. As previously showed for other cyst nematode species, the *H. carotae* populations were characterised by a strong heterozygote deficit. A Bayesian clustering analysis revealed two distinct genetic clusters, with one group located in the north of Europe and a second one located in the south of Europe. Moreover, our results highlighted rather limited gene flow at small spatial scale and some events of long distance migration. This first investigation of the genetic diversity of *H. carotae* populations would be useful to develop sustainable control strategies.

Keywords: Daucus carota; Genetic diversity; Heterozygote deficit; Isolation By Distance, Phytoparasitic nematode; SSR markers

HIGHLIGHTS

- A set of 13 polymorphic microsatellite markers was developed for *H. carotae*
- High heterozygote deficit was found among populations
- Genetic diversity was structured into two clusters: a north group and a south group at the European scale
- Significant isolation by distance pattern was highlighted between *H. carotae* populations

1. INTRODUCTION

Population genetics have been applied to diverse plant-parasitic pest and pathogen species (Milgroom, 2015). These studies are of particular importance to understand disease evolution, host parasite interactions and to design efficient control methods against plant parasites. Predicting of where, when and how fast adaptation may occur, is a very challenging scientific question and has also strong practical interest (Lebarbenchon et al., 2008). The extent and speed of adaptation processes depend on species and/or population life history traits including dispersal abilities, population size and reproduction mode. These features strongly contribute to the genetic diversity level in pest populations and therefore, to their ability to adapt to control methods (e.g. phytosanitary products and plant resistances) and environmental conditions. Population genetic approaches allow to estimate the evolutionary potential of a pest and help to predict how and in which conditions, adaptation may appear and spread in populations (McDonald and Linde, 2002). As a result investigating the genetic diversity of plant-parasitic populations, at different spatial scales, could help to predict the potential efficiency and durability of control methods and also to give some new information about the evolutionary history of pest populations.

Plant-parasitic nematodes are harmful pests of cultivated crops causing important economic losses to a wide variety of crops, estimated at $US 100 billion (Bird and Kaloshian, 2003; Nicol et al., 2011). Most of the damages are caused by some species belonging to two major groups, root-knot nematodes such as *Meloidogyne* spp. and cyst nematodes such as *Heterodera* and *Globodera* spp (Molinari, 2011). Among the different solutions to control plant-parasitic nematodes, chemical nematicides, such as the fumigant methyl bromide or the dichloropropane, are the most efficient, but many of them have already been banned in France and Europe and withdrawn from the market due to their harmful effects on human health and the environment (Oka et al., 2000). These regulatory changes promote alternative control

65 solutions, such as biocontrol solutions and an extensive use of plant resistances. There is
66 already some strong evidences of the adaptive potential of nematodes and in particular for
67 cysts nematodes species to the use of resistant plants (Fournet et al., 2013; Niere et al., 2014;
68 Phillips and Blok, 2008), but also some evidences that such adaptive abilities can be different
69 between populations and strongly correlated to their evolutionary history (e.g. Zaheer et al.
70 (1993) and Castagnone-Sereno et al. (2015) for nematodes and Xhaard et al. (2012; 2011) for
71 a fungus).

72 Cyst nematodes are sedentary endoparasites of plants with a survival stage, the cyst,
73 which is the body hardened of dead female containing eggs. Basically, second-stage juvenile
74 (J2) hatch from the cyst thanks to the perception of root exudates released by the host plant
75 (Perry, 1986). The juveniles migrate freely into the soil to the root tip and penetrate inside to
76 establish a feeding site, the syncytium, which is an important nutrient sink for the plant (Jones
77 and Northcote, 1972). Nematodes realize successive moults through the third (J3) and fourth
78 (J4) stages to become either male or female. Adult males leave the root in order to mate
79 females, which are growing through burst on the epidermal layers of the root and become
80 visible on the root surface. After mating, females die and form a cyst (their cuticle turns
81 brown and hardens) that remains in the soil after harvest. The cyst constitutes an effective
82 stage to spread and survive in which juveniles stay viable for several years in the soil (Lilley
83 et al., 2005).

84 The genetic structure and the evolutionary history of several cyst nematodes have been
85 previously studied at different spatial scales, e.g. the potato cyst nematodes *Globodera pallida*
86 (Eves-van den Akker et al., 2015; Picard et al., 2004; Picard et al., 2007; Plantard et al., 2008)
87 and *G. rostochiensis* (Boucher et al., 2013; Mimee et al., 2015), the tobacco cyst nematode *G.*
88 *tabacum* (Alenda et al., 2014; Marché et al., 2001), the beet cyst nematode *Heterodera*
89 *schachtii* (Kim et al., 2018; Plantard and Porte, 2004), the cereal cyst nematode *H. avenae*

(Wang et al., 2018) or the soybean cyst nematode *H. glycines* (St-Marseille et al., 2018; Wang et al., 2015). However, to our knowledge, no study has been conducted to describe the genetic structure of the carrot cyst nematode *Heterodera carotae*, except two recent studies using genic markers with a diagnostic purpose (Escobar-Avila et al., 2018; Madani et al., 2017).

H. carotae was first reported in England in 1931 and described by Jones in 1950 and is distributed worldwide (see Subbotin et al., 2010 for a review). It causes remarkable yield losses to carrot in Europe (Greco et al., 1993; Mugniery and Bossis, 1988) and in the USA (Berney and Bird, 1992). It has been recently described in Canada (Yu et al., 2017). A contrario to its sister species *H. cruciferae*, *H. carotae* is highly specific and develop on the genera of *Daucus* and *Torilis* (Aubert, 1986) and only controlled since decades with chemical nematicides which are banned today. The time frame and geographic regions of the first domestication of its cultivated host, *Daucus carota* L. subsp *sativus* Hoffm., remain still unclear. However, it has been generally accepted that wild carrots, which are indigenous of Europe and Central Asia, are the ancestors of domesticated carrots (i.e. cultivated), and that cultivated carrots are originated from Central Asia during the 10^{th} century (Grzebelus et al., 2014; Iorizzo et al., 2013). Then, carrot crops spread in Europe between the 11^{th} and 15^{th} centuries. The orange-rooted carrots, the most recent evolution of cultivated carrots, are probably the result of selection in the early 17^{th} century. Finally, orange carrots spread from Europe to other continents such as North America (Baranski et al., 2012).

Over the last decades, microsatellite markers have proved to be good candidates for investigating the population genetic structure of nematodes because these markers are highly polymorphic, codominant and broadly neutral (Selkoe and Toonen, 2006). Microsatellites, also known as simple sequence repeats (SSR), consist of motifs of one to six nucleotides tandemly repeated in different frequencies among populations.

The goals of this study were 1) to develop polymorphic microsatellite markers amplifiable in all *H. carotae* populations and 2) to investigate, for the first time, the genetic structure of *H. carotae* populations at the European scale.

2. MATERIAL AND METHODS

2.1 Nematode populations

Eighteen populations of the carrot cyst nematode, *H. carotae*, were collected at the field spatial scale (10 samplings randomized in the field) and multiplied on the cultivar "Carottes nantaises" (Vilmorin) in greenhouse for this study. Thirteen of these populations were sampled in fields in France, one in Italy, one in Switzerland, two in Denmark and the last one was sampled outside Europe in a field in Canada. Cyst were extracted from soil samples by a Kort elutriator and stored at 5°C in moistened sandy soil.

2.2 DNA extraction

One J2 from forty cysts was used for each population (i.e. 720 individuals). An extraction procedure using sodium hydroxide, adapted from Stanton et al. (1998) was conducted. Each J2 was incubated at room temperature overnight in 20 µL of NaOH 0.25 M in microtube. Tubes were then centrifuged at 3,700 rpm during 3 min and incubated in a thermocycler at 99°C for 2 min before the addition of 20 µL of lysis buffer (10 µL HCl 0.25 M, 5 µL Tris HCl 0.5 M pH 8, 5 µL Triton X100 2%) follow by incubation for 2 min at 99°C. Subsequently, 10 µL of a second buffer (5 µL Tris 0.1 M pH 8, 0.5 µL EDTA 0.1 M, 0.5 µL Tergitol® type NP-40, 0.25 µL Proteinase K at 20 mg/mL, adjusted with sterile distilled water) were added and the tubes were incubated at 55°C for 1h and then at 94°C for 10 min. Tubes were centrifuged (3,700 rpm, 30 s) and the supernatants were used for PCR.

139 *2.3 Microsatellite genotyping*

140 In order to assess the genetic variability of *H. carotae* populations, 13 polymorphic
141 microsatellite markers (Hc07, Hc29, Hc35, Hc40, Hc49, Hc55, Hc59, Hc63, Hc72, Hc76,
142 Hc87, Hc91 and Hc94) were developed according to the procedure developed by Malausa et
143 al. (2011) using library enrichment and next generation sequencing. All the details regarding
144 the choice and the development of the microsatellite loci used in this study can be found in the
145 Appendix A. PCR reaction in multiplex contained 5 µL of 2X Type-it Microsatellite PCR kit
146 (Qiagen), 1 µL of 2 µM primer mix, 2 µL of template DNA and 2 µL of sterile water distilled.
147 Cycling conditions included an initial denaturation at 95 °C for 5 min, followed by 30 cycles
148 of denaturation at 95 °C for 30 s, annealing at 57 °C for 90 s and extension at 72 °C for 30 s,
149 followed by a final extension at 60 °C for 30 min. PCR products were then diluted 1:40 in
150 sterile water and 3 µL of this dilution was mixed with 7.5 µL of GeneScan 500 LIZ Size
151 Standard: formamide (Applied Biosystems) = 1:100 (volume ratio). Analysis of PCR products
152 were conducted on ABI 3730XL sequencer (GENTYANE Platform, INRA, Clermont-
153 Ferrand, France). Allele size determinations were achieved using the GeneMapper software
154 v5.0 (Applied Biosystems) by manual identification of peaks. A second round of PCR and
155 electrophoresis was performed for 10% of the global number of individuals. Among the 720
156 individuals, 665 are successfully genotyped and this dataset was used for all genetic analyses
157 except for STRUCTURE analysis.
158

159 *2.4 Data analysis*

160 Linkage disequilibrium between loci was calculated using GENEPOP 4.0.7 (Raymond, 1995),
161 with default Markov chain parameters, to count the number of locus pairs showing significant
162 linkage disequilibrium across all populations. A Bonferroni correction (adjusted $\alpha = 0.0006$
163 for 78 comparisons) was applied to take into account multiple testing. Null allele frequencies

were estimated for each locus across all populations using the likelihood-based method of Chybicki and Burczyk (2009) implemented in the INEst program.

To investigate the genetic diversity, unbiased estimates of gene diversity (Hnb according to Nei (1978)) and allelic richness (Ar) were estimated for each population using respectively GENETIX 4.05.2 (Belkhir et al, 2004) and the rarefaction method (El Mousadik and Petit, 1996) implemented in POPULATION 1.2.32 (Langella, 1999). Deviations from random mating (F_{IS}) were computed for each population using GENETIX 4.05.2, and significances of F_{IS} were tested using the allelic permutation method (10,000 permutations).

To explore the genetic structure of *H. carotae* individuals without taking into account their geographic origin, the Bayesian clustering algorithms implemented in STRUCTURE 2.3.4 (Falush et al., 2003; Pritchard et al., 2000) were run on a dataset without missing data (i.e. 320 individuals), in order to obtain the best assignment for each individuals (Lombaert et al., 2018; Smith and Wang, 2014). Following the recent recommendations of Wang (2017), and because the sizes of samples from the different populations were unbalanced, the alternative ancestry prior, a nondefault ALPHA value (i.e. 1/p with p being the number of populations) and the uncorrelated allele frequency model were used. Simulations were performed using the admixture model. *H. carotae* individuals were assigned to K genetic clusters, with K varying from 1 to 19 (i.e. p + 1). For each assumed K value, thirty independent runs were conducted to assess the consistency of the results across runs, and each run was based on 3×10^6 iterations after a burn-in period of 1×10^6 iterations. The most likely number of genetic clusters (K) was statistically determined using the ad-hoc Evanno statistic ΔK (Evanno et al., 2005). To confirm the accuracy of individual assignments provide by STRUCTURE, a principal component analysis (PCA) was also performed using the procedure available in the package adegenet (Jombart, 2008) for the statistical freeware R

188 version 3.4.3 (R Core Team, 2017). Individuals from each obtained genetic clusters were then
189 used to explore the substructure within each cluster following the same procedure.

190 Genetic differentiation coefficients (F_{ST}) were estimated using GENEPOP 4.5.1
191 according to Weir and Cockerham (1984) for each pair of populations. Patterns of isolation by
192 distance (IBD) were tested by calculating the correlation between the matrices of pairwise
193 genetic distances (F_{ST} / (1 - F_{ST}) (Slatkin, 1995) and the natural logarithm of geographic
194 distance for each pair of populations in the entire dataset and within each genetic cluster
195 (Rousset, 1997). The statistical significance of correlations was assessed with a Mantel test
196 (10,000 permutations) using XLSTAT 2018.7 (Addinsoft SARL, Paris, France).

197

198 **3. RESULTS**

199 *3.1 Genetic diversity of* H. carotae *populations*

200 Among the 665 genotyped individuals from the eighteen *H. carotae* populations, our set of 13
201 microsatellites markers allowed the identification of 63 distinct alleles, with two to nine
202 alleles per locus (Table 1). Only three locus pairs (Hc55 - Hc87, Hc59 – Hc87 and Hc40 -
203 Hc59) among the 78 pairs showed a significant linkage disequilibrium with the Bonferroni's
204 adjustment. The percentage of estimated null allele averaged 8.94% among all microsatellite
205 markers (Table 1). This low number of null alleles will not have an important effect on the
206 estimations of population differentiation and heterozygote deficit (Carlsson, 2008; Wang et
207 al., 2015).

208 The allelic richness (Ar) estimated on a reduced sample of fourteen individuals varied
209 between 1.62 and 2.63 alleles per populations. The unbiased heterozygoty (Hnb) ranged from
210 0.15 to 0.43. The relationship among the unbiased estimates of gene diversity (Hnb) and the
211 allelic richness showed populations with low diversity ($0.1 < Hnb < 0.25$ and $1.5 < Ar < 2.0$),

212 which corresponded to populations from Denmark, Canada, Swiss, France (3001 and 0101),
213 and populations with high diversity (Hnb > 0.3 and Ar > 2.0) for all the other ones (Figure 1).
214 Among the eighteen populations, only four populations were at the Hardy-Weinberg
215 equilibrium (F_{IS} not significantly different to zero), and all the fourteen remaining populations
216 showed a significant positive value of F_{IS} (from 0.11 to 0.38), highlighting heterozygote
217 deficits (Table 2).
218
219 *3.2 Genetic structuration and clustering analysis*
220 Results from the Bayesian clustering analysis showed clearly that, according to the Evanno's
221 ΔK, the best way to explain the structuration of individuals was to group them into K = 2
222 genetic clusters (Figure 2A). The thirty replicate runs gave the same result. STUCTURE
223 results indicated that 94% of individuals were very well assigned (with a percentage of
224 assignation to one or the other cluster higher than 90%), and the genetic differentiation
225 between cluster 1 and 2 was high (F_{ST} = 0.16). Principal component analysis (PCA) supported
226 the existence of two major genetic clusters (1 and 2) with axis 1 and 2 explaining 27.3% and
227 3.1% respectively of total genetic variability (Figure 2B). Regarding both the geographical
228 position of the seventeen European populations with their cluster's membership highlighted
229 two geographical groups, with a south group corresponding to cluster 1 and a north group
230 corresponding to cluster 2, even if some individuals of the cluster 1 were located in the north
231 and vice versa (Figure 3). The Canadian population clustered with populations from the
232 northern group (Figure 3).
233 To see a possible substructure into these two clusters, two other STRUCTURE
234 analyses were performed using individuals from each obtained cluster. The cluster 1, which
235 contained six populations, divided in two sub-clusters (1a and 1b; Figure 4A), which were
236 supported by PCA results (Figure 4B). The cluster 2, which contained twelve populations

237 divided in four sub-clusters (2a, 2b, 2c and 2d; Figure 4C), which were supported by PCA
238 results (Figure 4D).
239
240 *3.3 Genetic differentiation between H. carotae populations*
241 At the large spatial scale explored here, the genetic differentiation was significant for all pairs
242 of *H. carotae* populations. The pairwise F_{ST} values ranged from 0.01 to 0.69 (Figure 5). The
243 lowest pairwise F_{ST} values were observed between populations from the north group (e.g.
244 between populations 50170, 2902, Cre7, 5601, 4402, 7201 – Figure 5). In the genetic cluster
245 1, the highest F_{ST} values were observed for the population ZAP, which is also the most
246 geographically distant. In the genetic cluster 2, the highest F_{ST} values were observed for the
247 population 3001, which is geographically distant from other French populations but far less
248 distant than Danish and Canadian populations. However, the relationships between the genetic
249 distance (F_{ST} / 1 - F_{ST}) and the ln(geographic distance) were significantly correlated, either for
250 the entire dataset ($r^2 = 0.09$; $P < 0.0001$), or within genetic cluster 2 ($r^2 = 0.13$; $P < 0.0001$),
251 except for genetic cluster 1 ($r^2 = 0.20$; $P = 0.131$) (Appendix B).
252
253 **4. DISCUSSION**

254 The present study has provided a set of 13 polymorphic microsatellite markers, multiplexible
255 in four panels, which constitutes a powerful tool to perform population genetics studies in the
256 carrot cyst nematode *Heterodera carotae*. Moreover, we have performed here the first study
257 of the genetic structure of *H. carotae* populations at the European scale, which highlights a
258 high heterozygote deficit and two distinct genetic clusters in this species.
259 Among the eighteen *H. carotae* populations, fourteen deviate from the Hardy-
260 Weinberg equilibrium and show significant positive F_{IS} values. Such a heterozygote deficit
261 appears to be shared between cyst nematode species, as it has been also described for

Globodera pallida (Picard et al., 2004), *Heterodera schachtii* (Plantard and Porte, 2004), *G. tabacum* (Alenda et al., 2014), *H. glycines* (Wang et al., 2015) and *H. avenae* (Wang et al., 2018), which attack respectively potato, sugar-beet, tobacco, soybean and wheat. The cause of this feature is the low active dispersal abilities of nematode juveniles in the soil, which leads either to consanguineous mating (between individuals from the same cyst) or to substructure (Wahlund effect) at the spatial scale of the rhizosphere of a host plant. According to the recent results of Montarry et al. (2015) showing that heterozygote deficits were due to consanguinity for monovoltine species (*G. pallida*) and to substructure for polyvoltine species (*H. schachtii* and *G. tabacum*), we can hypothesize that the heterozygote deficit highlighted in *H. carotae* populations, which performed several generations per year, is mostly due to a Wahlund effect. Samples collected at the plant scale will be helpful now to further disentangle consanguinity from Wahlund effect in the carrot cyst nematode.

Our results showed that *H. carotae* populations are gathered into two genetic clusters. Geographical position of the different genotyped populations showed a clear spatial separation of both clusters, with one group located in the north of Europe and a second one located in the south of Europe, even if some populations included individuals from both genetic clusters. The significant isolation by distance pattern suggests a short-range dispersal among *H. carotae* populations leading to some genetic relationships between the closest populations. Although the Italian population was geographically and genetically distant from the other populations of the cluster 1, the IBD pattern was not significant within this cluster, but this is probably due to the low statistical power (only six populations). Conversely, the IBD pattern was significant within the cluster 2, whereas the most geographically distant populations (i.e. the Danish and the Canadian populations) were not the most genetically distant ones. Altogether, those results reveal i) a gradual migration between the closest populations, mainly due to agricultural practices, which have been showed to contribute to the

287 passive dispersion of cyst nematodes (Alenda et al., 2014), and also ii) some events of long
288 distance migration which are more rare events and probably owing to the plant material
289 transfer within a country or among countries. Because the genetic diversity is much smaller
290 for one French population (3001) and for Danish and Canadian populations than for the other
291 populations of the cluster 2, we can hypothesize that those populations were introduced (i.e.
292 sink populations) and came from the north group (i.e. source populations). Moreover, and
293 because the Canadian population clustered with two French populations in the substructure
294 analysis (sub-cluster 2a), contrary to the Danish populations and the population 3001 which
295 were genetically differentiated in the sub-cluster 2c and 2d, respectively, the event of
296 introduction in Canada could be more recent, at least in this production area. Moreover, at the
297 spatial scale studied here, F_{ST} within each cluster were strong (0.40 for cluster 1 and 0.32 for
298 cluster 2), indicating low gene flow among populations in each cluster. Nevertheless, previous
299 studies on potato, sugar beet and tobacco cyst nematodes, have reported an important gene
300 flow at a fine-scale (Alenda et al., 2014; Plantard et al., 2008; Plantard and Porte, 2004).
301 Alenda et al. (2014) revealed a leading role of the human activities and more specifically of
302 the transport of soil during harvesting in the passive dispersal of plant-parasitic nematodes.
303 Important soil losses from potato, sugar beet and carrot fields, during harvesting, have been
304 notified (Parlak et al., 2016; Ruysschaert et al., 2007a; Ruysschaert et al., 2007b). In the
305 framework of development of sustainable control strategies such as plant resistances or
306 biocontrol products, studying gene flow at a lower spatial scale may be the next essential step.
307 Population genetic studies may also help in understanding the evolutionary history of
308 pest species. Here, the genetic structure analysis revealed two distinct clusters with the same
309 level of genetic diversity. However, to date, we do not know if the structure observed for
310 nematode populations is the result of the history of cultivated carrots or the history of *H.*
311 *carotae* on wild carrots with multiple host switches from wild plants to the cultivated ones.

312 The first scenario is, with no doubt, the most realistic and simple to explore. Indeed, *H.*
313 *carotae* is highly specific of cultivated carrots, wild carrots are not natives of North America
314 and the center of domestication has generally been accepted to be Central Asia (Grzebelus et
315 al., 2014; Iorizzo et al., 2013). These arguments support i) a recent introduction of the
316 Canadian population of *H. carotae* in North America, probably through plant commercial
317 exchanges with European countries and ii) for European populations, a colonization of
318 cultivated areas associated with the domestication process, either between the 11th and 15th
319 from the center of domestication or more recently, with the spread of orange carrots. To
320 explore the origin of our both *H. carotae* clusters, some information argue to identify both
321 Afghanistan and Turkey to be two reliable places of domestication and of diversity (Simon,
322 2000; Stolarczyk and Janick, 2011) where cultivated and wild carrots still cohabit today. It
323 would then require to sample and analyse additional field populations in these regions to
324 determine whether the genetic clusters highlighted here could be found in these potential
325 source populations leading to the description of invasion routes. The second scenario is
326 clearly more difficult to explore. Indeed, even if there is increasing proof that pest species can
327 develop on wild host plants (Gracianne et al., 2014; Monteil et al., 2013; Rouxel et al., 2014),
328 there is no clear evidence of the occurrence of *H. carotae* populations on wild relatives
329 species. Moreover, favouring the scenario of multiple switches from wild host plants to
330 cultivated ones should lead to a more complex genetic structure than the one we observed.
331 Although it may thus represent an extensive and costly effort, sampling wild populations of
332 *H. carotae* on wild relative species may be a first step to strengthen this scenario, which
333 seems to be more speculative. . For both scenarios, new genetic analysis must be done with a
334 new set of populations to determine the possible origin of our two clusters.

335 In the current context of reduction of chemical nematicides, predicting the
336 effectiveness of control means of nematodes, such as resistant crops or biocontrol solutions, is

a key challenge. Hence, the description of the genetic diversity and its structuration in space may help to predict the efficiency of new alternative solutions, to design sustainable control strategies (McDonald and Linde, 2002; Pilet-Nayel et al., 2017; Zhu and Zhan, 2012), and from the farmer point of view, to choose the good control solution at the right place. To do so, and rather than working with randomly chosen populations, representative populations of both genetic clusters have to be used in biotest. Thus, we suggest to use the results of the substructure analysis and to test the different control means on six populations belonging to each sub-clusters (sub-clusters 1a, 1b, 2a, 2b, 2c and 2d). Furthermore, we also know that species with large geographical distribution, such as *H. carotae*, are supposed to comprise populations exhibiting different life history traits in response to their adaptation to different local climate conditions (Fournet et al., 2018). A complementary study may be thus to phenotypically characterize *H. carotae* populations from each genetic cluster under different climatic conditions (e.g. different temperatures).

ACKNOWLEDGMENTS

We gratefully acknowledge Sergei Subbotin for useful comments on a previous version of this paper and Jeremy Clotault for discussion about the history of cultivated carrots. Johannes Hallmann (Institute for Epidemiology and Pathogen Diagnostics – Germany), Lars Møller (GartneriRådgivningen A/S HortiAdvice Scandinavia – Denmark) and Yu Qing (Agriculture and Agri-Food Canada – Ottawa, Canada) who kindly provided some *Heterodera carotae* populations were acknowledged. Camille Gautier is supported by PhD grant from Groupe Roullier and ANRT (Association Nationale Recherche Technologie).

Appendix A. Microsatellite marker development

Appendix B. Relationship between the genetic distance (F_{ST} / 1 - F_{ST}) and ln(geographical distance) in *Heterodera carotae* populations

REFERENCES

Alenda, C., Montarry, J., Grenier, E., 2014. Human influence on the dispersal and genetic structure of French *Globodera tabacum* populations. Infection, Genetics and Evolution 27, 309-317.

Aubert, V., 1986. Contribution à l'étude de la biologie d'un parasite de la carotte. Université de Neuchâtel.

Baranski, R., Maksylewicz-Kaul, A., Nothnagel, T., Cavagnaro, P.F., Simon, P.W., Grzebelus, D., 2012. Genetic diversity of carrot (*Daucus carota* L.) cultivars revealed by analysis of SSR loci. Genetic Resources and Crop Evolution 59, 163-170.

Belkhir K., Borsa P., Chikhi L., Raufaste N. & Bonhomme F. 1996-2004 GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Berney, M., Bird, G., 1992. Distribution of *Heterodera carotae* and *Meloidogyne hapla* in Michigan carrot production. Journal of Nematology 24, 776-778.

Bird, D.M., Kaloshian, I., 2003. Are roots special? Nematodes have their say. Physiological and Molecular Plant Pathology 62, 115-123.

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.

Carlsson, J., 2008. Effects of microsatellite null alleles on assignment testing. Journal of Heredity 99, 616-623.

Castagnone-Sereno, P., Mulet, K., Iachia, C., 2015. Tracking changes in life-history traits related to unnecessary virulence in a plant-parasitic nematode. Ecology and Evolution 5, 3677-3686.

Chybicki, I.J., Burczyk, J., 2009. Simultaneous estimation of null alleles and inbreeding coefficients. Journal of Heredity 100, 106-113.

El Mousadik, A., Petit, R., 1996. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. Theoretical and Applied Genetics 92, 832-839.

Escobar-Avila, I.M., Óliver, E., Subbotin, S.A., Tovar-Soto, A., 2018. First report of carrot cyst nematode *Heterodera carotae* in Mexico: morphological, molecular characterization and host range study. Journal of Nematology 50, 229-242.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14, 2611-2620.

Eves-van den Akker, S., Lilley, C.J., Reid, A., Pickup, J., Anderson, E., Cock, P.J., Blaxter, M., Urwin, P.E., Jones, J.T., Blok, V.C., 2015. A metagenetic approach to determine the diversity and distribution of cyst nematodes at the level of the country, the field and the individual. Molecular Ecology 24, 5842-5851.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164, 1567-1587.

Fournet, S., Kerlan, M.-C., Renault, L., Dantec, J.-P., Rouaux, C., Montarry, J., 2013. Selection of nematodes by resistant plants has implications for local adaptation and cross-virulence. Plant Pathology 62, 184-193.

Fournet, S., Pellan, L., Porte, C., Piriou, C., Grenier, E., Montarry, J., 2018. Populations of the beet cyst nematode *Heterodera schachtii* exhibit strong differences in their life-history traits across changing thermal conditions. Frontiers in Microbiology 9, 2801.

Gracianne, C., Fournet, S., Porte, C., Renault, L., Petit, E.J., Rouaux, C., Fouville, D., Arnaud, J.-F., 2014. Spatial distribution and basic ecology of *Heterodera schachtii* and *H. betae* wild populations developing on sea beet, *Beta vulgaris* ssp. *maritima*. Nematology 16, 797-805.

Greco, N., D'addabbo, T., Brandonisio, A., Elia, F., 1993. Damage to Italian crops caused by cyst-forming nematodes. Journal of Nematology 25, 836-842.

Grzebelus, D., Iorizzo, M., Senalik, D., Ellison, S., Cavagnaro, P., Macko-Podgorni, A., Heller-Uszynska, K., Kilian, A., Nothnagel, T., Allender, C., 2014. Diversity, genetic mapping, and signatures of domestication in the carrot (*Daucus carota* L.) genome, as revealed by Diversity Arrays Technology (DArT) markers. Molecular Breeding 33, 625-637.

Holleley, C.E., Geerts, P.G., 2009. Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. BioTechniques 46, 511-517.

Iorizzo, M., Senalik, D.A., Ellison, S.L., Grzebelus, D., Cavagnaro, P.F., Allender, C., Brunet, J., Spooner, D.M., Van Deynze, A., Simon, P.W., 2013. Genetic structure and domestication of carrot (*Daucus carota* subsp. *sativus*)(Apiaceae). American Journal of Botany 100, 930-938.

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403-1405.

Jones, M., Northcote, D., 1972. Nematode-induced syncytium--a multinucleate transfer cell. Journal of Cell Science 10, 789-809.

Kim, J., Ni, G., Kim, T., Chun, J.Y., Kern, E.M., Park, J.K., 2018. Phylogeography of the highly invasive sugar beet nematode, *Heterodera schachtii* (Schmidt, 1871), based on microsatellites. Evolutionary Applications, 1-13.

Langella, O., 1999. Populations 1.2.31. Population Genetic Software (Individuals or Populations Distances, Phylogenetic Trees) 2012. http://bioinformatics.org/~tryphon/populations/.

Lebarbenchon, C., Brown, S.P., Poulin, R., Gauthier-Clerc, M., Thomas, F., 2008. Evolution of pathogens in a man-made world. Molecular Ecology 17, 475-484.

Lilley, C.J., Atkinson, H.J., Urwin, P.E., 2005. Molecular aspects of cyst nematodes. Molecular Plant Pathology 6, 577-588.

Lombaert, E., Guillemaud, T., Deleury, E., 2018. Biases of STRUCTURE software when exploring introduction routes of invasive species. Heredity 120, 485-499.

438 Madani, M., Palomares-Rius, J.E., Vovlas, N., Castillo, P., Tenuta, M., 2017. Integrative diagnosis of
439 carrot cyst nematode (*Heterodera carotae*) using morphology and several molecular markers
440 for an accurate identification. European Journal of Plant Pathology 150, 1023.
441 Malausa, T., Gilles, A., Meglecz, E., Blanquart, H., Duthoy, S., Costedoat, C., Dubut, V., Pech, N.,
442 Castagnone-Sereno, P., Delye, C., Feau, N., Frey, P., Gauthier, P., Guillemaud, T., Hazard, L., Le
443 Corre, V., Lung-Escarmant, B., Male, P.J., Ferreira, S., Martin, J.F., 2011. High-throughput
444 microsatellite isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA
445 libraries. Mol Ecol Resour 11, 638-644.
446 Marché, L., Valette, S., Grenier, E., Mugniéry, D., 2001. Intra-species DNA polymorphism in the
447 tobacco cyst nematode complex (*Globodera tabacum*) using AFLP. Genome 44, 941-946.
448 McDonald, B.A., Linde, C., 2002. Pathogen population genetics, evolutionary potential, and durable
449 resistance. Annual Review of Phytopathology 40, 349-379.
450 Meglécz, E., Costedoat, C., Dubut, V., Gilles, A., Malausa, T., Pech, N., Martin, J.-F., 2010. QDD: a user-
451 friendly program to select microsatellite markers and design primers from large sequencing
452 projects. Bioinformatics 26, 403-404.
453 Milgroom, M.G., 2015. Population biology of plant pathogens: genetics, ecology and evolution.
454 American Phytopathological Society Press, St. Paul, MN.
455 Mimee, B., Duceppe, M.O., Véronneau, P.Y., Lafond-Lapalme, J., Jean, M., Belzile, F., Bélair, G., 2015.
456 A new method for studying population genetics of cyst nematodes based on Pool-Seq and
457 genomewide allele frequency analysis. Molecular Ecology Resources 15, 1356-1365.
458 Molinari, S., 2011. Natural genetic and induced plant resistance, as a control strategy to plant-
459 parasitic nematodes alternative to pesticides. Plant Cell Reports 30, 311-323.
460 Montarry, J., Jan, P.L., Gracianne, C., Overall, A.D., Bardou-Valette, S., Olivier, E., Fournet, S., Grenier,
461 E., Petit, E.J., 2015. Heterozygote deficits in cyst plant-parasitic nematodes: possible causes
462 and consequences. Molecular Ecology 24, 1654-1667.

463 Monteil, C.L., Cai, R., Liu, H., Llontop, M.E.M., Studholme, D.J., Morris, C.E., Vinatzer, B.A., 2013.
464 Nonagricultural reservoirs contribute to emergence and evolution of *Pseudomonas syringae*
465 crop pathogens. New Phytologist 199, 800-811.
466 Mugniery, D., Bossis, M., 1988. *Heterodera carotae* Jones, 1950. 1. Gamme d'hôtes, vitesse de
467 développement, cycle. Revue de nématologie 11, 307-313.
468 Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of
469 individuals. Genetics 89, 583-590.
470 Nicol, J.M., Turner, S.J., Coyne, D., Den Nijs, L., Hockland, S., Maafi, Z.T., 2011. Current nematode
471 threats to world agriculture, Genomics and molecular genetics of plant-nematode interactions.
472 Springer Science+Business Media B.V. , pp. 21-43.
473 Niere, B., Krüssel, S., Osmers, K., 2014. Auftreten einer außergewöhnlich virulenten Population der
474 Kartoffelzystennematoden. J Kult 66, 426-427.
475 Oka, Y., Koltai, H., Bar-Eyal, M., Mor, M., Sharon, E., Chet, I., Spiegel, Y., 2000. New strategies for the
476 control of plant-parasitic nematodes. Pest Management Science: formerly Pesticide Science
477 56, 983-988.
478 Parlak, M., Palta, Ç., Yokuş, S., Blanco-Canqui, H., Çarkacı, D.A., 2016. Soil losses due to carrot
479 harvesting in south central Turkey. Catena 140, 24-30.
480 Perry, R.N., 1986. Physiology of hatching, Cyst nematodes. Springer, pp. 119-131.
481 Phillips, M., Blok, V., 2008. Selection for reproductive ability in Globodera pallida populations in
482 relation to quantitative resistance from *Solanum vernei* and *S. tuberosum* ssp. *andigena*
483 CPC2802. Plant Pathology 57, 573-580.
484 Picard, D., Plantard, O., Scurrah, M., Mugniery, D., 2004. Inbreeding and population structure of the
485 potato cyst nematode (*Globodera pallida*) in its native area (Peru). Molecular Ecology 13,
486 2899-2908.

487 Picard, D., Sempere, T., Plantard, O., 2007. A northward colonisation of the Andes by the potato cyst
488 nematode during geological times suggests multiple host-shifts from wild to cultivated
489 potatoes. Molecular Phylogenetics and Evolution 42, 308-316.
490 Pilet-Nayel, M.-L., Moury, B., Caffier, V., Montarry, J., Kerlan, M.-C., Fournet, S., Durel, C.-E.,
491 Delourme, R., 2017. Quantitative resistance to plant pathogens in pyramiding strategies for
492 durable crop protection. Frontiers in Plant Science 8, 1838.
493 Plantard, O., Picard, D., Valette, S., Scurrah, M., Grenier, E., Mugniéry, D., 2008. Origin and genetic
494 diversity of Western European populations of the potato cyst nematode (*Globodera pallida*)
495 inferred from mitochondrial sequences and microsatellite loci. Molecular Ecology 17, 2208-
496 2218.
497 Plantard, O., Porte, C., 2004. Population genetic structure of the sugar beet cyst nematode
498 *Heterodera schachtii*: a gonochoristic and amphimictic species with highly inbred but weakly
499 differentiated populations. Molecular Ecology 13, 33-41.
500 Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus
501 genotype data. Genetics 155, 945-959.
502 R Core Team (2017). R: A language and environment for statistical computing. R Foundation
503 for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
504 Raymond, M., 1995. GENEPOP (version 1.2): population genetics software for exact tests and
505 ecumenicism. J. Hered. 86, 248-249.
506 Rousset, F., 1997. Genetic differentiation and estimation of gene flow from F-statistics under
507 isolation by distance. Genetics 145, 1219-1228.
508 Rouxel, M., Mestre, P., Baudoin, A., Carisse, O., Delière, L., Ellis, M.A., Gadoury, D., Lu, J., Nita, M.,
509 Richard-Cervera, S., 2014. Geographic distribution of cryptic species of *Plasmopara viticola*
510 causing downy mildew on wild and cultivated grape in eastern North America. Phytopathology
511 104, 692-701.

512 Ruysschaert, G., Poesen, J., Auerswald, K., Verstraeten, G., Govers, G., 2007a. Soil losses due to
513 potato harvesting at the regional scale in Belgium. Soil Use and Management 23, 156-161.
514 Ruysschaert, G., Poesen, J., Wauters, A., Govers, G., Verstraeten, G., 2007b. Factors controlling soil
515 loss during sugar beet harvesting at the field plot scale in Belgium. European Journal of Soil
516 Science 58, 1400-1409.
517 Selkoe, K.A., Toonen, R.J., 2006. Microsatellites for ecologists: a practical guide to using and
518 evaluating microsatellite markers. Ecology Letters 9, 615-629.
519 Simon, P.W., 2000. Domestication, historical development, and modern breeding of carrot. Plant
520 Breeding Reviews 19, 157-190.
521 Slatkin, M., 1995. A measure of population subdivision based on microsatellite allele frequencies.
522 Genetics 139, 457-462.
523 Smith, O., Wang, J., 2014. When can noninvasive samples provide sufficient information in
524 conservation genetics studies? Molecular Ecology Resources 14, 1011-1023.
525 St-Marseille, A.-F.G., Lord, E., Véronneau, P.-Y., Brodeur, J., Mimee, B., 2018. Genome scans reveal
526 homogenization and local adaptations in populations of the soybean cyst nematode. Frontiers
527 in Plant Science 9, 987.
528 Stanton, J., McNicol, C., Steele, V., 1998. Non-manual lysis of second-stage Meloidogyne juveniles for
529 identification of pure and mixed samples based on the polymerase chain reaction. Australasian
530 Plant Pathology 27, 112-115.
531 Stolarczyk, J., Janick, J., 2011. Carrot: history and iconography. Chronica Horticulturae 51, 13-18.
532 Subbotin, S.A., Mundo-Ocampo, M., Baldwin, J.G., 2010. Systematics of cyst nematodes (Nematodes:
533 Heteroderinae), Volume 8, Part B. Leiden, The Netherlands, Brill.
534 Wang, H.-M., Zhao, H.-h., Chu, D., 2015. Genetic structure analysis of populations of the soybean cyst
535 nematode, *Heterodera glycines*, from north China. Nematology 17, 591-600.
536 Wang, J., 2017. The computer program structure for assigning individuals to populations: easy to use
537 but easier to misuse. Molecular Ecology Resources 17, 981-990.

538 Wang, X., Ma, J., Liu, H., Liu, R., Li, H., 2018. Development and characterization of EST-derived SSR
539 markers in the cereal cyst nematode *Heterodera avenae*. European Journal of Plant Pathology
540 150, 105-113.
541 Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure.
542 Evolution 38, 1358-1370.
543 Xhaard, C., Barres, B., Andrieux, A., Bousset, L., Halkett, F., Frey, P., 2012. Disentangling the genetic
544 origins of a plant pathogen during disease spread using an original molecular epidemiology
545 approach. Molecular Ecology 21, 2383-2398.
546 Xhaard, C., Fabre, B., Andrieux, A., Gladieux, P., Barres, B., Frey, P., Halkett, F., 2011. The genetic
547 structure of the plant pathogenic fungus *Melampsora larici-populina* on its wild host is
548 extensively impacted by host domestication. Molecular Ecology 20, 2739-2755.
549 Yu, Q., Ponomareva, E., Van Dyk, D., McDonald, M., Sun, F., Madani, M., Tenuta, M., 2017. First
550 Report of the Carrot Cyst Nematode (*Heterodera carotae*) From Carrot Fields in Ontario,
551 Canada. Plant Disease 101, 1056-1056.
552 Zaheer, K., Fleming, C., Turner, S., 1993. Distribution and frequency of occurrence of potato cyst
553 nematode pathotypes in Northern Ireland. Plant Pathology 42, 609-616.
554 Zhu, W., Zhan, J., 2012. Population genetics of plant pathogens. eLS. John Wiley & Sons, 1-7.

555

FIGURE LEGENDS

Fig. 1. Graphical representation of populations in term of allelic richness and gene diversity. Allelic richness (Ar), in x-axis was obtained by using de rarefraction method implemented in population software with a reduced sample size of 14 individuals. Gene diversity (Hnb) was estimated by using the unbiased estimate of Nei (1978). The most diverse populations were in top right corner and the less diverse populations were in the bottom left corner of the graph.

Fig. 2. (A) Clustering results of all *H. carotae* individuals by STRUCTURE. Each sample is indicated by a bar divided in K=2 colored section showing the individual's membership for each clusters. Vertical dotted bars separate individuals from different populations which are indicated below. (B) Scatter plot of the first two principal components of the same data used for the Bayesian clustering analyses. Groups corresponding to the different clusters are plotted using the same colours as in (A).

Fig. 3. Geographical location of the eighteen *H. carotae* populations with their membership proportion of clusters. The colours are the same as Figure 2A and correspond to K=2. The diameter of the circle corresponds to the allelic richness.

Fig. 4. Sub-clustering results of cluster 1 (A) and cluster 2 (C) *H. carotae* individuals by STRUCTURE. Each sample is indicated by a bar divided into K clusters represented by different colours. Vertical dotted bars separate individuals from different populations which are indicated below.

579 Scatter plot of the first two principal components of the same data used for the Bayesian
580 clustering analyses for cluster 1 (B) and for cluster 2 (D). Groups corresponding to the
581 different sub-clusters are plotted using the same colours as in (A) for K = 2 and (C) for K = 4.
582
583 **Fig. 5.** Matrix of pairwise F_{ST} between *H. carotae* populations. The distribution of pairwise
584 F_{ST} into cluster 1 and cluster 2 is represented by the box plots.
585

586 **Table 1:** The thirteen microsatellites markers used in this study (primer sequences, motif, size, number of allele per locus, percentage of null

587 allele)

588

Code	Primer left sequence	Primer right sequence	Motif	Size (bp)	Nbr of allele	Null allele (%)
Hc07	GCAGAATAGACGTCCACTAGCA	GAAAGAAAGATATAGCCAAAAGCG	(tgtc)5	140	2	4.07
Hc29	TGTTTGATTGGATTCCCTGG	CAGTTGAATGGTTTTGTGGG	(acag)6	145	4	0.00
Hc35	GCGCCACCTTTTGATGTTAT	CAATTTAAGGAATAAGCGAAAGAA	(ct)8	103	2	0.01
Hc40	CGTCCAGTCTCTTTTCGTTT	ATTTGTTCAGCTTTTATTTGACCG	(ag)8	190	4	11.71
Hc49	ATAATGAAAAGCGAGGGGCT	GCATCACCCATTTCCTTTGT	(ag)6	106	3	0.00
Hc55	GTGGGCGTCGTCAAATCAT	ACATTGTTATCAGAGGCAAATCA	(gt)5	140	3	7.95
Hc59	ACAAGTCGTGTCCACTTCCC	TGTGATTTTGTATGGCATAGGTG	(ct)5	158	7	14.97
Hc63	ATCGTTGAGAAGTTTATTTTGCTTG	CTACGCCCAAAAGGTCAAAA	(ac)8	140	8	9.48
Hc72	CCCTTAATGGTTTTCTCAACTG	AGTATGTGGTTGCCGAAGAA	(ct)7	141	5	17.88
Hc76	AGCTTGCGATGAGTCTCCTG	ATCGCTATGGTGATGCCAA	(tg)7	145	6	0.00
Hc87	TTAATCCTTTTGGATGAGATATTGG	CTTTCGAGTGCACACCCTG	(tc)7	140	9	18.51
Hc91	GCATTATGTTTGCTTTGCCA	TTGATCAAATCGGCATGCTA	(ag)7	133	5	28.70
Hc94	CTGGGGCGAAACTTTTATGA	TTTGGTTAATATTGGAATGAATGC	(ac)9	101	5	2.92

589 **Table 2:** Sampling site, population code, number of genotyped individuals per population (n)
590 and deviation from random mating (F_{IS}). F_{IS} significantly higher than zero are indicated with
591 an asterisk.

592

Country	Code	*n*	F_{is}
France. Ain	0101	37	0.109*
France. Bouches du Rhône	1302	38	0.183*
France. Bouches du Rhône	1303	40	0.224*
France. Finistère	2902	40	0.246*
France. Gard	3001	40	0.039
France. Loire Atlantique	4401	35	0.375*
France. Loire Atlantique	4402	39	0.224*
France. Manche	50170	40	0.223*
France. Morbihan	5601	36	0.045
France. Mayenne	7201	34	0.137*
France. Vaucluse	8401	39	0.252*
Switzerland	FU	40	0.178*
Italy	ZAP	38	0.274*
France. Créance	Cre7	39	-0.022
Canada. Ontarrio	Can1	14	0.327*
Denmark. Odsherred	Dan1	38	0.077
Denmark. Holbaek	Dk	38	0.118*
France. Aisne	Hca02	40	0.285*

593

Figure 1

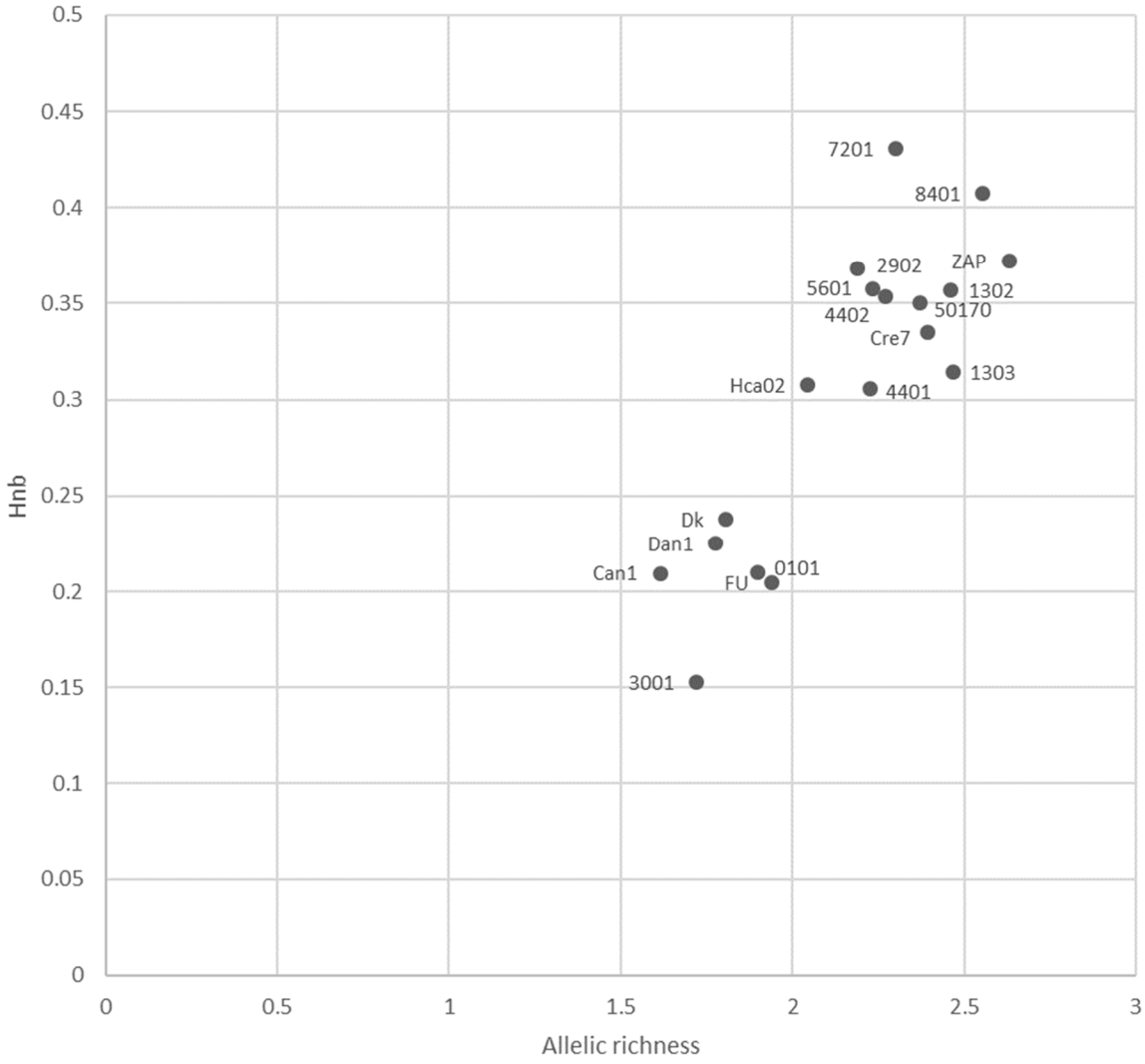
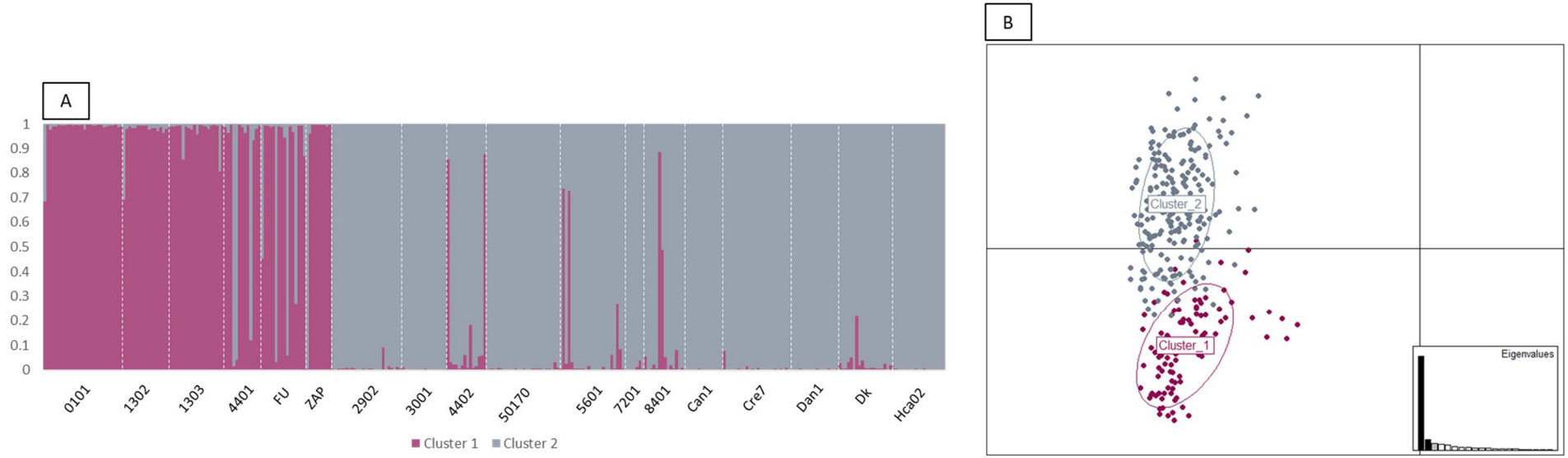
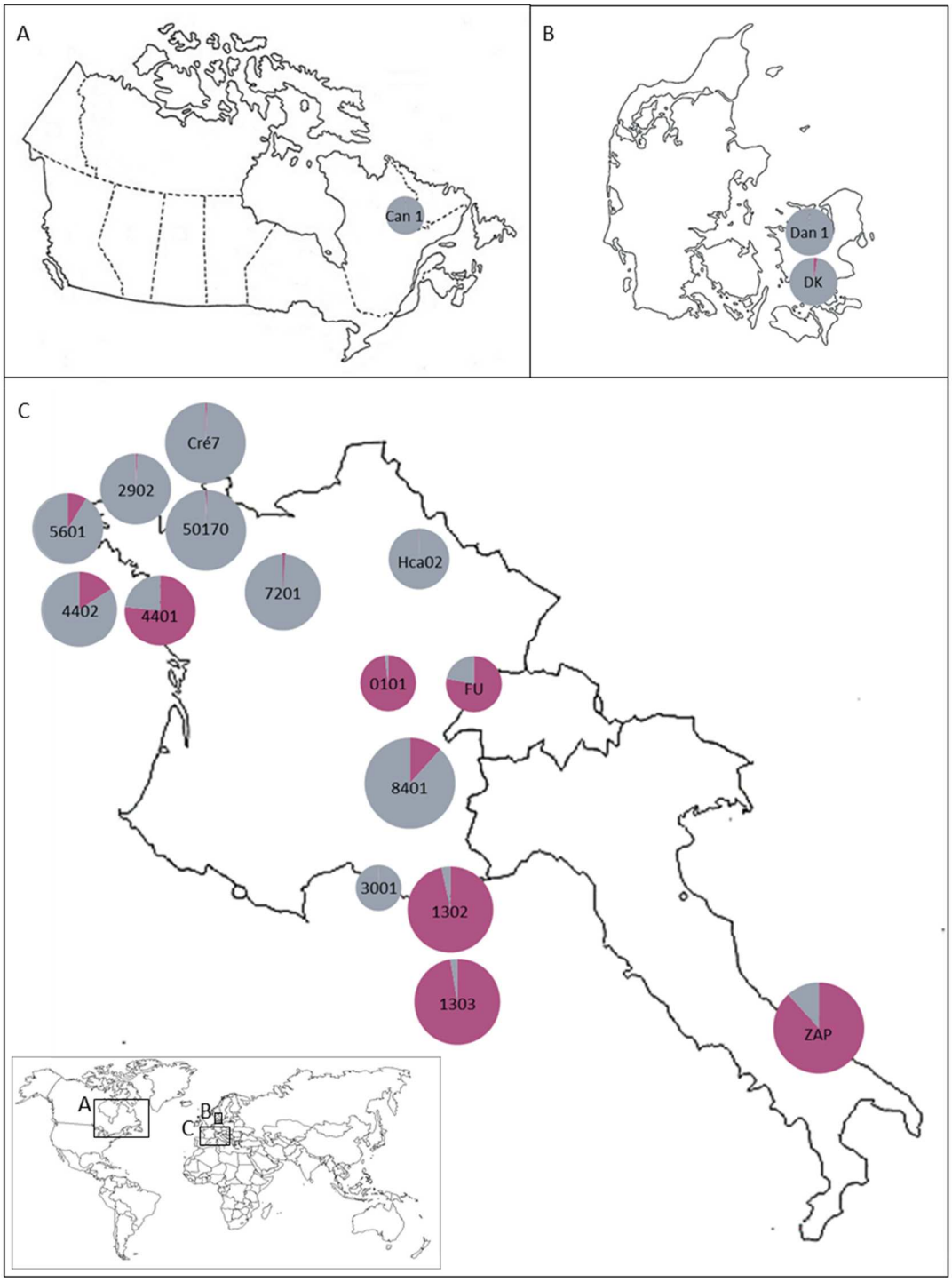

0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Hnb
0
0.5
1
1.5
2
2.5
3
Allelic richness
7201
8401
2902
ZAP
5601
1302
4402
50170
Cre7
1303
Hca02
4401
Dk
Dan1
Can1
0101
FU
3001

Figure 2


A
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0101
1302
1303
4401
FU
ZAP
2902
3001
4402
50170
5601
7201
8401
Can1
Cre7
Dan1
Dk
HcaD2
Cluster 1
Cluster 2
B
Cluster_2
Cluster_1
Eigenvalues

600 **Figure 3**

601



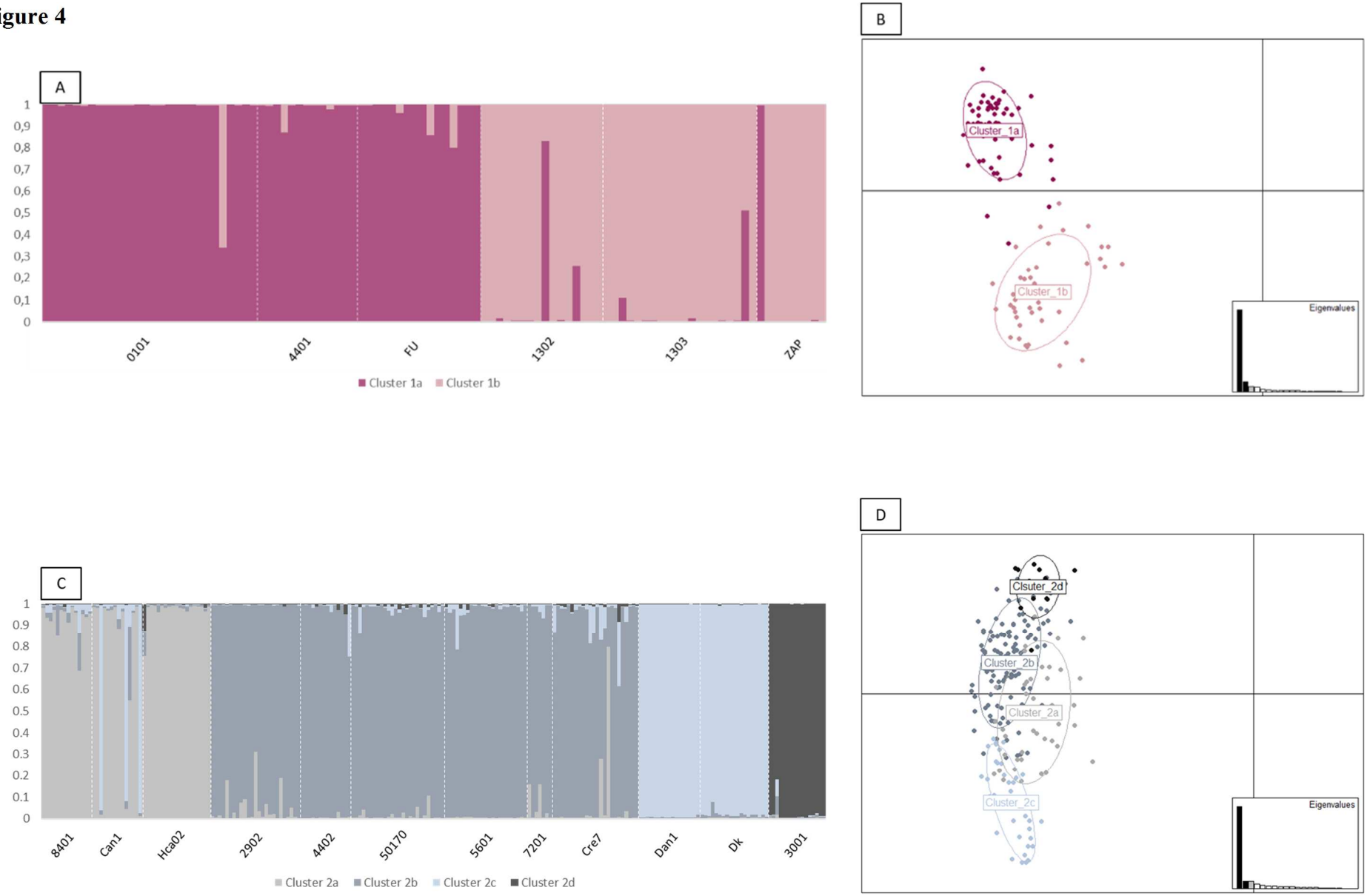
602

603 **Figure 4**

604

605

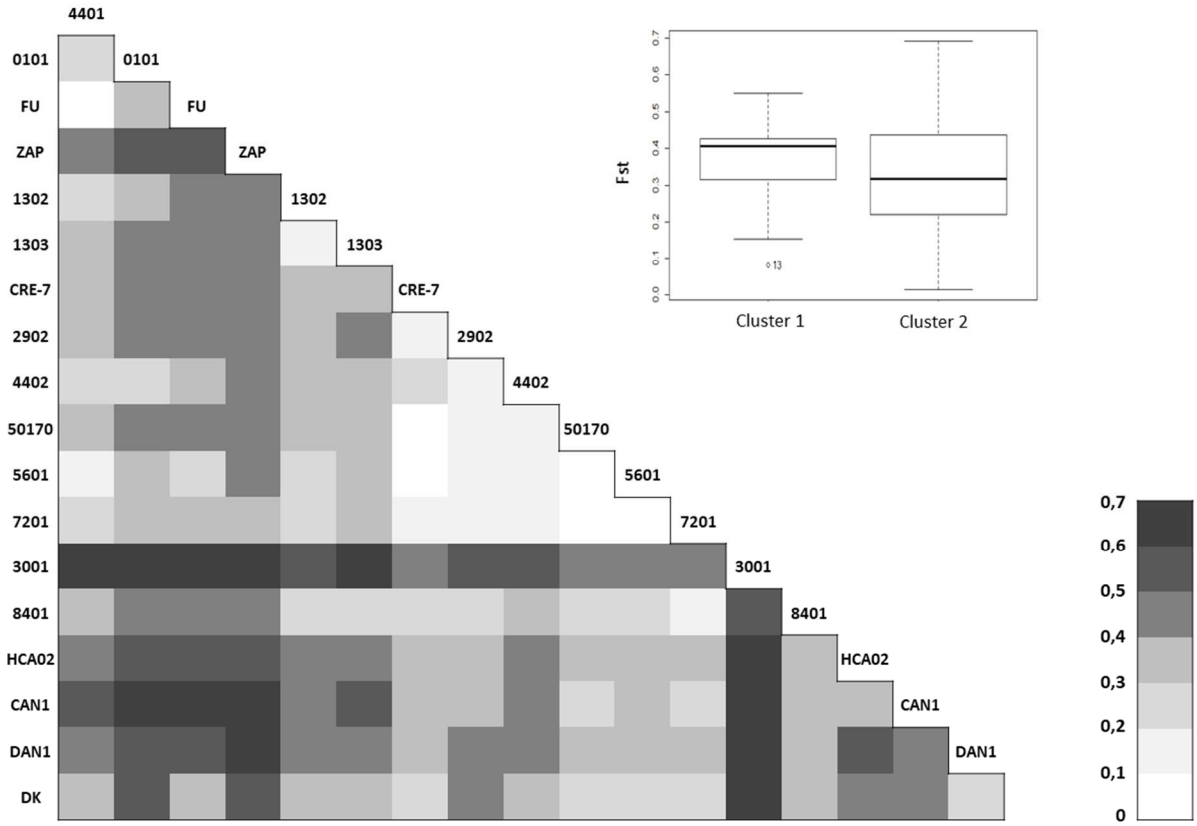
606



607 **Figure 5**

608

609



610 **Appendix A. Microsatellite marker development**

611 Microsatellite markers were developed according to the procedure described by Malausa et al.
612 (2011) using next-generation sequencing and library enrichment. Twelve *Heterodera carotae*
613 populations (0101, 1303, 3001, 4401, 4402, 5601, 7201, 8401, FU, ZAP, Hca02, which were
614 then also used for the genetic structure analysis, and 5001, a population from the north-west
615 of France [Manche], which has been lost before the genetic structure analysis) were used to
616 identify microsatellite markers. DNA of each population was extracted from 100 cysts using a
617 DNeasy Blood & Tissue kit (QIAGEN) and pooled altogether. Enriched libraries were
618 constructed employing eight microsatellites probes ((TG), (TC), (AAC), (AAG), (AGG),
619 (ACG), (ACAT) and (ACTC)) and were sequenced by Genoscreen (Lille, France) using the
620 454 GS-FLX titanium pyrosequencing technology (ROCHE Diagnostics).

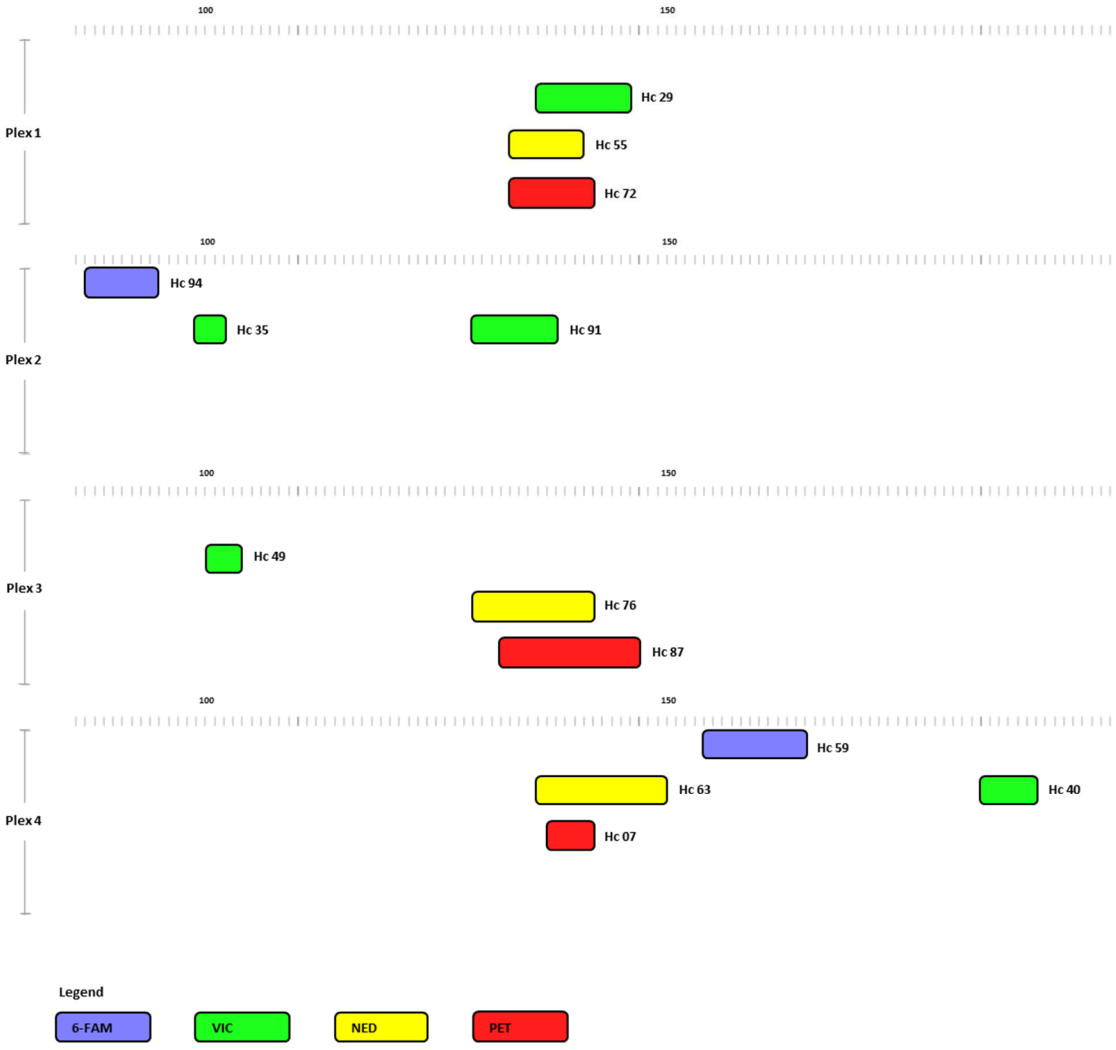
621 To explore the 241,190 reads obtained from next-generation sequencing and to design
622 primers for microsatellite amplification, the QDD program (Meglécz et al., 2010) was used.
623 Among those reads, 1,134 sequences harboured microsatellite motifs and primers were
624 successfully defined for 199 of them. We have then discarded all tri-nucleotide motifs, all di-
625 nucleotide motifs with less than six repetitions and all the loci showing an amplification size
626 below 92 bp, leading to a set of 95 loci with perfect microsatellite motifs.

627 A biological validation of these loci was carried out on single juveniles (J2) from four
628 populations (3001, 5601, ZAP and FU). The procedure of DNA extraction was described in
629 the main text. A PCR was carried out in a final volume of 10 µL using 5 µL of 2X Type-it
630 Microsatellite PCR kit (Qiagen), 0.25 µL of 10 µM forward primer, 0.25 µL of 10 µM reverse
631 primer and 2 µL of template DNA. Volumes were adjusted to 10 µL with sterile water
632 distilled. Cycling conditions included an initial denaturation at 95 °C for 5 min, followed by
633 39 cycles of denaturation at 95 °C for 30 s, annealing at 57 °C for 90 s and extension at 72 °C
634 for 30 s, ended by a final elongation step at 72 °C during 5 min. The amplification of these

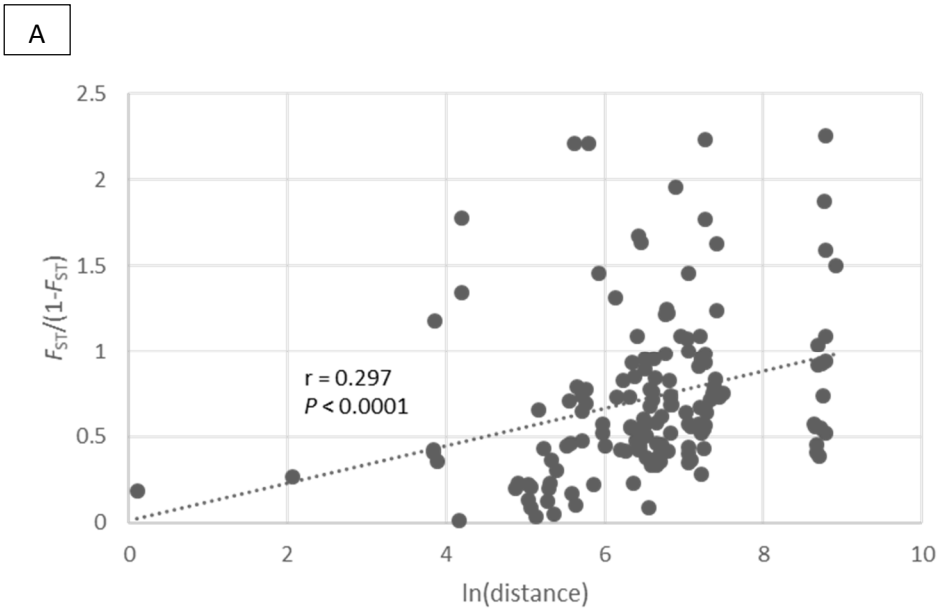
PCR products was revealed by electrophoresis on a 2 % agarose gel. We discarded loci with no or low amplification and multiband amplification and kept 36 pairs of primers with a good amplification of the four populations.

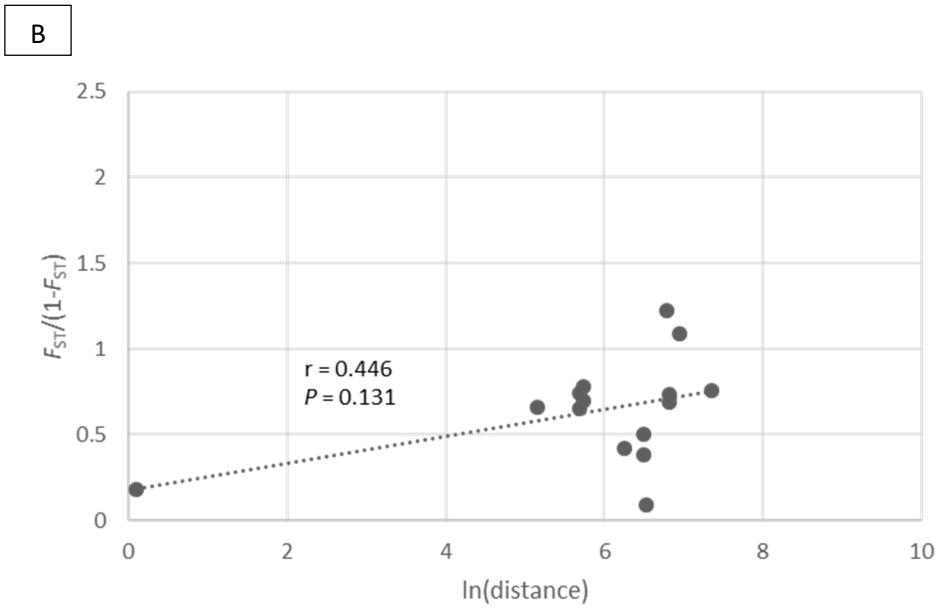
To explore the reproducibility and the polymorphism of the selected loci, two juveniles from two distinct cysts from the four populations (3001, 5601, ZAP and FU) were used. The forward primer of each pair was tailed with M13F [5'- CAC GAC GTT GTA AAA CGA C -3'] to facilitate labelling. A PCR was realized twice in a final volume of 10 µL using 5 µL of 2X Type-it Microsatellite PCR kit (Qiagen), 0.25 µL of 10 µM fluorescent-labeled M13 primer (VIC, Applied biosystem), 0.25 µL of 10 µM forward primer, 0.25 µL of 10 µM reverse primer and 2 µL of template DNA. Volumes were adjusted to 10 µL with sterile water distilled. Cycling conditions included an initial denaturation at 95 °C for 5 min, followed by 20 cycles of denaturation at 95 °C for 30 s, annealing at 57 °C for 90 s and extension at 72 °C for 30 s, followed by 20 cycles of denaturation at 95 °C for 30 s, annealing at 53 °C for 90 s, extension at 72 °C for 30 s and a final extension at 60 °C for 30 min. PCR products were then diluted 1:40 in sterile water and 3 µL of this dilution was mixed with 7.5 µL of GeneScan 500 LIZ Size Standard: formamide (Applied Biosystems) = 1:100 (volume ratio). Analysis of PCR products were conducted on ABI 3730XL sequencer (Applied Biosystems). Allele size determinations were achieved using the GeneMapper software v5.0 (Applied Biosystems) by manual identification of peaks. This step allowed to validate a set of 13 polymorphic microsatellite markers showing a reproducible amplification.

Multiplex manager Software (Holleley and Geerts, 2009) was used to define the best combinations of these 13 microsatellites markers for multiplex reactions. Four multiplex combinations (Fig below) were identified and contained between three to four markers (each 2µM) per panels with a different fluorescent dye (FAM, VIC, NED, PET).


100
150
Plex 1
Hc 29
Hc 55
Hc 72
100
150
Plex 2
Hc 94
Hc 35
Hc 91
100
150
Plex 3
Hc 49
Hc 76
Hc 87
100
150
Plex 4
Hc 59
Hc 63
Hc 40
Hc 07
Legend
6-FAM
VIC
NED
PET

Appendix B. Relationship between the genetic distance (F_{ST} / 1 - F_{ST}) and ln(geographical distance) in *Heterodera carotae* populations for the entire dataset (A), for cluster 1 (B) and for cluster 2 (C).

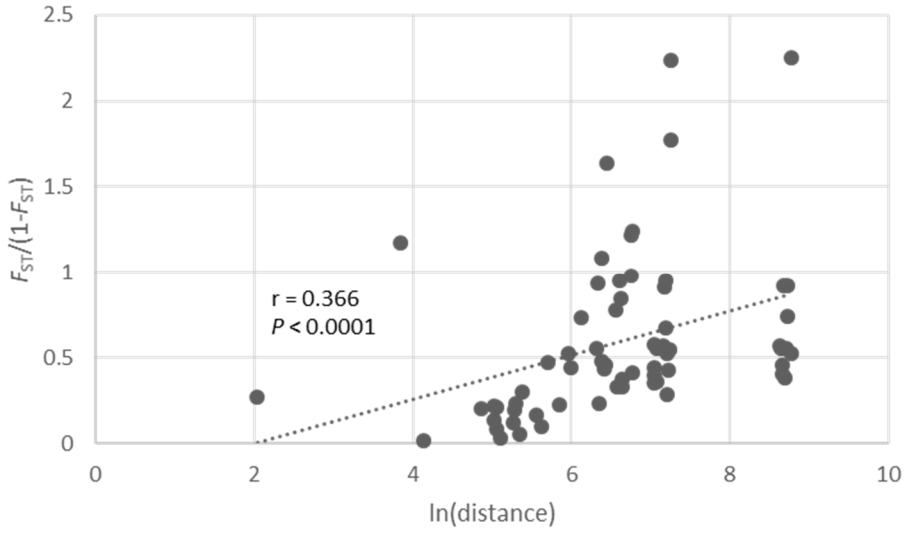




674

675

C



676

Graphical abstract