MOLECULAR ECOLOGY

Molecular Ecology (2010) 19, 856–873

doi: 10.1111/j.1365-294X.2009.04511.x

Genetic signature of amphimixis allows for the detection and fine scale localization of sexual reproduction events in a mainly parthenogenetic nematode

LAURE VILLATE,* DANIEL ESMENJAUD,† MAARTEN VAN HELDEN,‡ SOLENN STOECKEL* and OLIVIER PLANTARD*

*INRA, Agrocampus Ouest, University Rennes 1, UMR 1099 BiO3P (Biology of Organisms and Populations applied to Plant Protection), F-35653 Le Rheu, France, †INRA, UMR 1301 IBSV (Interactions Biotiques et Santé Végétale), F-06903 Sophia Antipolis, France, ‡ENITA de Bordeaux, UMR SV (Santé Végétale), F-33170 Gradignan, France

Abstract

Asexuality is an important mode of reproduction in eukaryotic taxa and has a theoretical advantage over sexual reproduction because of the increased ability to propagate genes. Despite this advantage, hidden signs of cryptic sex have been discovered in the genomes of asexual organisms. This has provided an interesting way to address the evolutionary impact of sex in plant and animal populations. However, the identification of rare sexual reproduction events in mainly asexual species has remained a challenging task. We examined the reproductive history in populations of the plant parasitic nematode *Xiphinema index* by genotyping individuals collected from six grapevine fields using seven microsatellite markers. A high level of linkage disequilibrium and heterozygous excess suggested a clonality rate of 95–100%. However, we also detected rare sexual reproduction events within these highly clonal populations. By combining highly polymorphic markers with an appropriate hierarchical sampling, and using both Bayesian and multivariate analysis with phylogenetic reconstructions, we were able to identify a small number of sexually produced individuals at the overlapping zones between different genetic clusters. This suggested that sexual reproduction was favoured when and where two nematode patches came into contact. Among fields, a high degree of genetic differentiation indicated a low level of gene flow between populations. Rare genotypes that were shared by several populations suggested passive dispersal by human activities, possibly through the introduction of infected plants from nurseries. We conclude that our method can be used to detect and locate sexual events in various predominantly asexual species.

Keywords: asexual reproduction, *Caenorhabditis elegans*, clonality, evolution of sex, gene flow, GFLV, negative *Fis*, *Vitis vinifera*, *Xiphinema index*

Received 25 August 2009; revised 12 November 2009; accepted 21 November 2009

Correspondence: Olivier Plantard, Fax: (33) 240 68 78 02; E-mail: olivier.plantard@oniris-nantes.fr

Present address: Laure Villate, Max Planck Institute for Developmental Biology, Department of Evolutionary Biology, Spemannstr. 37, 72076 Tübingen, Germany.

Present address: Olivier Plantard, INRA, ONIRIS, UMR 1300 BioEpAR, Ecole Nationale Vétérinaire, Agroalimentaire et de L'Alimentation Nantes Atlantique, Atlanpôle, La Chantrerie BP 40706, F-44307 Nantes Cedex 03, France.

Introduction

Reproductive modes are major key factors that shape how genes are transmitted through generations (Maynard-Smith 1998; Bengtsson 2003). Asexually is found in numerous eukaryotic groups (Honnay & Bossuyt 2005; De Meeus *et al.* 2007) and can be revealed by several genetic clues. For purely asexual diploid organisms,

theoretical models predict that the absence of genome-wide recombination events should maintain heterozygosity by avoiding homozygote production. Heterozygosity can be further increased by independent accumulations of mutations over both alleles through generations (Lokki 1976; Pamilo 1987; Judson & Normark 1996; Welch & Meselson 2000). This leads to strongly negative F_{IS} values (Delmotte *et al.* 2002; Balloux *et al.* 2003), reflecting an excess of heterozygotes compared to the frequency expected at the Hardy-Weinberg equilibrium. Additionally, in the absence of recombination and segregation, the entire genome behaves as a single cluster of strongly physically linked genes. Therefore, we can expect strong linkage disequilibria among all loci, even those from different chromosomes (Balloux *et al.* 2003; De Meeus *et al.* 2006). Asexual reproduction is thus expected to reduce multilocus genotype diversity and favour allelic diversity at neutral molecular markers (Balloux *et al.* 2003).

When organisms reproduce using partial asexuality, difficulties arise in estimating the contribution of each reproductive mode to the evolution of the species using molecular tools and population genetics indices. For example, when species reproduce using moderate asexuality—defined as more than three or more than 20% of individuals generated by sexual reproduction per generation (Balloux *et al.* 2003; Bengtsson 2003)—population genetics indices show no significant differences from identical but purely sexual populations. The best indicator of clonality in such populations is multilocus linkage disequilibrium, since this quantity is unbiased in panmictic populations and progressively increases as clonality rate increases (De Meeus *et al.* 2006). With approximately 99% clonality, rare events of sexual reproduction will translate into a high variance in F_{IS} over multiple loci (Balloux *et al.* 2003; De Meeus & Balloux 2005), but clusters of genes maintained through asexual reproduction will give rise to a strong linkage disequilibrium (De Meeus & Balloux 2004). Although these theoretical studies suggest that rare sexual reproduction events should have a strong impact on classical population genetics parameters (such as *Fis* and linkage disequilibrium) in species where asexual rate is high, only a few empirical studies have been conducted so far (see however Delmotte *et al.* 2002; Stoeckel *et al.* 2006; Goyeau *et al.* 2007).

The nematode phylum is an excellent system to study the evolutionary biology of mating systems in the metazoans. Their biodiversity is high (26 000 species described to date, Hugot *et al.* 2001) and a variety of mating systems have been described. For non zooparasitic nematodes, the genetic structure of natural populations has rarely been investigated (for the model organism *C. elegans* see however LaMunyon & Ward 1997; Koch *et al.* 2000; Graustein *et al.* 2002; Sivasundar & Hey 2003, 2005; Barrière & Félix 2005, 2007; Haber *et al.* 2005; Cutter 2006; Baïlle *et al.* 2008). Although parthenogenetic species are common among phytoparasitic nematodes, previous investigations have focused only on two amphimictic species (Plantard and Porte 2004; Picard *et al.* 2004, Picard & Plantard 2006).

The plant parasitic nematode *Xiphinema index* has been reported to reproduce solely through asexual reproduction (Dalmasso & Younes 1969; Dalmasso 1970). This nematode is thought to have been introduced from the Middle East into Europe and it has become an invasive agricultural pest. This has great economic impact in European vineyards because it transmits the *Grapevine fanleaf virus* (GFLV) to grapevine crop (Hewit et al. 1958; Taylor & Brown 1997; Brown & Weischer 1998; Esmenjaud 2000). Cytological analyses of its genital apparatus and its ovogenesis sequence demonstrated that females reproduce by meiotic parthenogenesis without recombination (Dalmasso & Younes 1969; Dalmasso 1970). This case of automictic thelythoky is equivalent to clonal reproduction and the offspring are genetically identical to the mother. However, males are rarely found in natural populations (1–2%, Harris 1977), suggesting that sexual reproduction could be occurring in this species. This organismal evidence for sexual reproduction (Schurko *et al.* 2008) does not necessarily imply sexual recombination among genotypes in the wild because these males might be non-functional, as has been shown previously for the parthenogenetic root-knot nematode *Meloidogyne incognita* (Sasser & Carter 1985; Triantaphyllou 1985; Eisenback *et al.* 1991; Karssen 2002; Lunt 2008). Another indication of potential occasional mating is an anecdotal observation of few spermatozoa in the uterus of one *X. index* female (Luc & Cohn 1982).

Many organisms with a major impact on economics or health exhibit mixed asexual and sexual reproduction (Milgroom 1996; Taylor *et al.* 1999; Tibayrenc 1999). Rare events of recombination can have a drastic effect on the reshuffling of genetic material (Burt *et al.* 1996; Bengtsson 2003), and result in the faster spread of virulence genes and emergence of new virulence combinations than through mutation alone (Awadalla 2003). Accurate knowledge of the clonal reproduction rate in pathogens populations is thus crucial for efficient control (Milgroom & Fry 1997; McDonald & Linde 2002). As the use of chemical nematicides is prohibited in an increasing number of countries because of adverse environmental effects (Abawi & Widmer 2000), increasing plant host resistance has become the preferred method of nematode control. A resistant variety of grapevine has been developed, but its success in agricultural programs is still being evaluated (Bouquet *et al.* 2003, 2004;

Riaz *et al.* 2007; Xu *et al.* 2008). This method of nematode control may impose a selective pressure on the nematodes to be able to overcome host plant resistance (Müller 1992; Turner & Fleming 2002).

The patchy spatial pattern of *X. index* in the soil (Villate *et al.* 2008) and its weak ability to disperse actively suggest that there should be a strong genetic structure at the field scale. Moreover, as this nematode lives deeply in the soil (Esmenjaud *et al.*1992; Villate *et al.* 2008) and is associated with the perennial vine crop, we suppose no or only very limited passive dispersal through wind or field machinery among grapevine fields. Indeed, the absence of a conservative stage during the nematode life cycle (such as the cyst for cyst nematodes or egg-masses for root-knot nematodes) renders *X. index* individuals susceptible to mechanical impacts and drought. Passive dispersal can occur via dispersal of the grapevine plants in nurseries, but nematode-infected plants are very rare in nurseries because certified plant material is produced under the severe rules of a prophylactic scheme. In the absence of both active and passive dispersal, combined with the non-native status of this species and probable multiple introductions, a high genetic differentiation of nematode populations between grapevine fields are expected.

To describe the population structure and the mating system of *X. index*, we genotyped 303 geo-referenced individuals collected in six grapevine fields of the Bordeaux vineyard at seven microsatellite loci. First, we characterised the mating system of this species using instantaneous and historical genetic parameters, by the analysis of linkage disequilibrium, F_{IS}, and the ratio of multi- vs. single-copy multilocus genotypes. Second, by combining Bayesian assignment methods, multivariate analysis and phylogenetic reconstructions, we identified some individuals as resulting from sexual reproduction events. Third, by mapping these individuals within the grapevine fields, we located the occurrence of sexual events to specific sites. Finally, we characterised the population differentiation and gene flow at various scales (area, *Château*, grapevine field) using hierarchical F_{ST} and the analysis of isolation by distance patterns.

Materials and methods

Field sampling and isolation of nematodes from the soil

Populations were sampled from six GFLV-infected grapevine fields (populations MA, MB, *C*, *H*, *G*, *T*, located from 1.2 to 45 km apart), in the Bordeaux wine-producing region (France) (Fig. 1). MA and MB fields belong to the same *Château* in the Medoc vineyard area, *H* is in the Pessac area, and *G*, *C* and *T* are in three different *Châteaux* in the Saint-Emilion area. MA and C were sampled in May 2005, *T* in August 2005, *G* in April 2006, *H* in April 2007 and MB in August 2007. At each sampling point, a trench (1.5 m long, 0.3 m wide, 1 m deep) was dug with an excavating machine. A 2-L sample of soil was collected from the side of the trench, at a depth of 0.5–0.8 m. As *X. index* was supposed to be strongly settled, individuals from each grapevine field were sampled from the 2-L sample of soil's scale (lowest) to the plot's scale (highest).

Nematodes were extracted in the laboratory from 2-L soil samples, using the method described in Villate *et al.* (2008). We recovered 1–16 nematodes per sample.

DNA extraction and microsatellite genotyping

DNA was extracted from each nematode individual using the method described in Villate *et al.* (2009). Extracted DNA was then stored at −20 °C until PCR amplification. All individuals were genotyped at seven microsatellites (Villate *et al.* 2009). PCR reactions were carried out following Villate *et al.* (2009). PCR products were run on an ABI PRISM® 3130 XL Genetic Analyzer 16 Capillary system and sized with internal lane standard (500-LIZ) using the program GeneMapper version 3.7 (Applied Biosystems).

Data analysis

Diploid multilocus genotypes (MLGs) represented by several individuals in the data set were considered as 'repeated MLGs'. Because clonal reproduction may affect estimates of genetic diversity and genetic differentiation between populations, all the analyses were performed both with all MLGs (*with*) and keeping only one copy of each repeated MLG (*without*).

Determination of the reproductive mode. We first assessed whether the observed genotypic patterns truly reflected the reproductive mode of the populations or if they may have resulted from a lack of resolution of our genetic markers. To estimate if identical MLGs could be considered as a part of the same asexual lineage, we estimated the probabilities that two individuals sharing the same MLG came from sexual reproduction events. We computed the lower bound of this probability: $P_{(ID)Unbiased}$, under Hardy–Weinberg expectations (Paetkau *et al.* 1998), and its upper bound, $P_{(ID)Sib}$, under strict reproduction between siblings, as recommended by Waits *et al.* (2001). For each population, we calculated the total number of MLGs, the number of repeated MLGs (numbers of clones), the number of MLGs represented by a single copy and the number of individuals per repeated MLG. Immediate asexual reproduction rate in populations

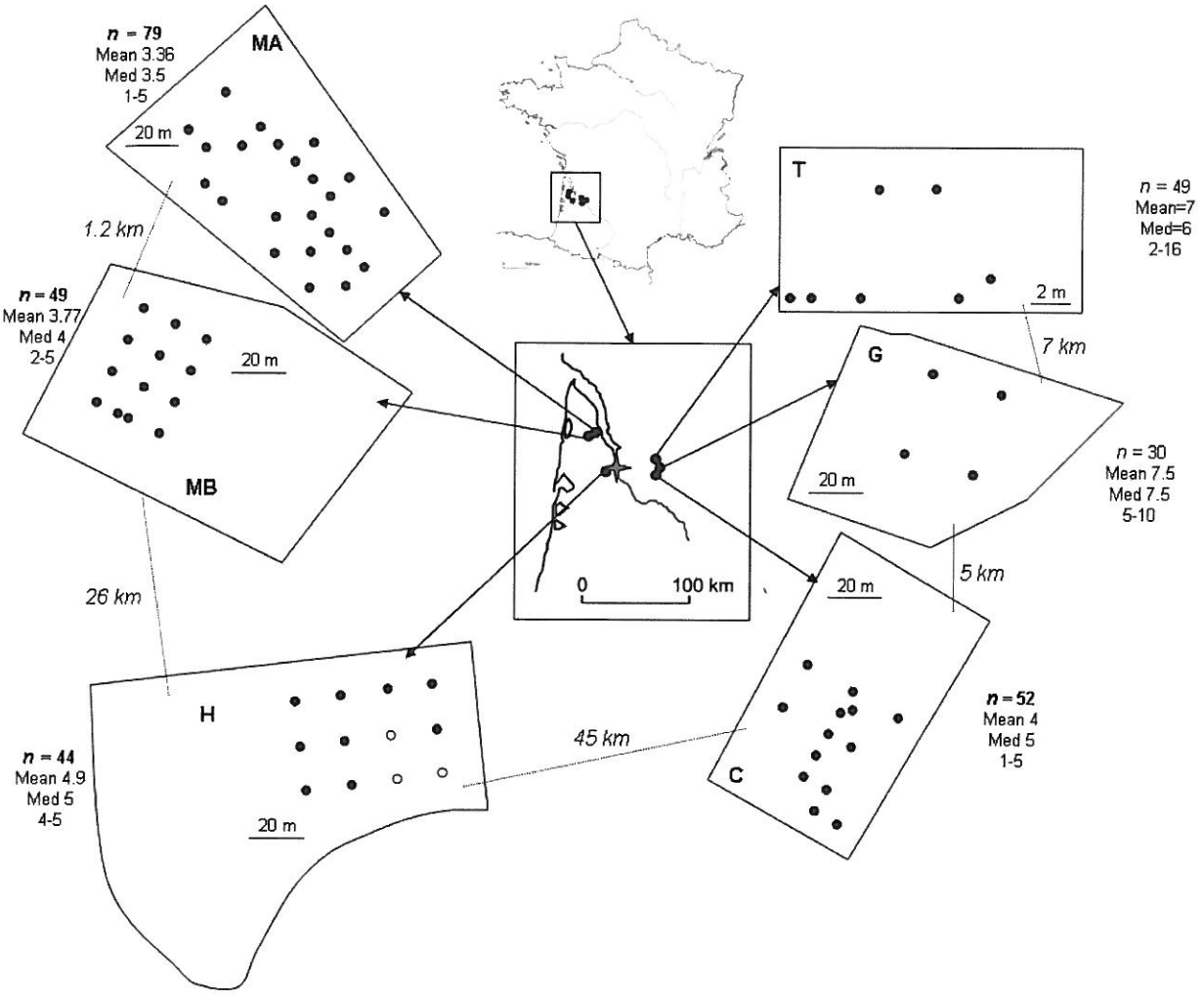

Fig. 1 Localization of the study in France and sampling design of *Xiphinema index* populations in the six grapevine fields. In each grapevine field, black circles indicate the 2-L sampling points; N, the number of nematodes genotyped, with the mean, median and minimum and maximum numbers of nematodes per sampling point. The red star locates Bordeaux city. MA and MB fields are in the Médoc area, H is in the Pessac area, and C, G and T are in the Saint-Emilion area.

was calculated as $1 - (g/N)$, where g is the number of different MLGs and N is the number of individuals. However, the number of different MLGs overestimates the sexual recombination events because new MLGs can also come from mutation events. So $1 - (g/N)$ should be considered as the lower-bound estimate of the asexual rate.

Because the sample sizes among fields varied, we estimated the allelic richness of each population by the rarefaction index (ElMousadik & Petit 1996) using POPULATION software version 1.2.30 (available at http://bioinformatics.org/~tryphon/populations/). In those analyses, all individuals with amplification failure at one locus were excluded because a null amplification may result in an underestimation of allelic richness.

Weir & Cockerham (1984) F_{IS} values were calculated using FSTAT software version 2.9.3 (Goudet 1995). Observed and expected heterozygosity were calculated using GeneClass2 (Piry *et al.* 2004). Hardy–Weinberg exact tests were performed using GENEPOP version 4.0 (Rousset 2008) assuming that the alternative hypothesis H_1 is a heterozygote excess. Genotypic linkage disequilibria among loci were calculated for all populations and within each population, with or without repeated MLGs. We tested linkage disequilibrium between loci after sequential Bonferroni multiple-comparison correction using FSTAT 2.9.3. Finally, we estimated the clonal reproduction rate by comparing the F_{IS} values and associated standard errors observed in our six populations with a plot of F_{IS} and associated standard errors as a

function of the clonal reproduction rate (Balloux *et al.* 2003).

Identification and localisation of individuals generated by sexual reproduction. To identify clusters of genetically related individuals without a priori knowledge, we used a Bayesian method available in the software STRUCTURE version 2.2 (Pritchard *et al.* 2000). Although these assignment methods were originally developed for purely sexual organisms, empirical studies revealed that they could identify clusters of individuals in partially asexual species. This Bayesian approach can thus be applied to species that do not reach all the theoretical expectations requested (Kaeuffer *et al.* 2007) and is robust to deviations from some assumptions, particularly linkage between loci (Falush *et al.* 2003). We performed runs without repeated MLGs and used a number of genetic clusters (K) ranging from 1 to 6. For each K value, STRUCTURE was run 5 times for 20 000 steps after a burning period of 20 000 steps. The 'true K', the most probable number of clusters, was determined with the *ad hoc* method proposed by Evanno *et al.* (2005). We considered an individual as assigned to only one cluster if its probability of assignment is >95%, and an individual as admixed when its probability of assignment is >5% in at least two different clusters. We also used InStruct program (Gao *et al.* 2007), which assumes linkage disequilibrium. We finally confirmed these clusters with a Factorial Correspondence Analysis (FCA) performed without repeated MLGs, using GENETIX Software version 4.05 (Belkhir *et al.* 1996–2004). The FCA is not a model-based approach and can thus be trusted whatever the species and its reproductive mode.

To discriminate if a MLG represented by only one individual resulted from recombination (sexual reproduction) or from somatic mutations (clonality), we computed the allele-sharing genetic distance DAS (Jin & Chakraborty 1994) between MLGs with POPULATION software version 1.2.30. We used those distances to build Neighbour-Joining trees of MLGs within and among grapevine fields. We then considered that the likelihood that several mutations occurred jointly in the same individual, making it different from its related and surrounding individuals, is very low. The genetic distance between a clone and an individual displaying a single MLG resulting from mitotic mutation is thus expected to be lower (i.e. variation at a single microsatellite locus) than between a clone and an individual displaying a single MLG resulting from massive recombination (by meiosis) through sexual reproduction.

For each grapevine field where two clusters and admixed individuals were identified by STRUCTURE, we wrote a Visual Basic module in order to simulate five offspring genotypes coming from sexual reproduction events between individuals of the first cluster and individuals of the second one. The module was written for our own purposes and the source code is available on request. Then, we added these five simulated genotypes to the data set and computed Neighbour-Joining tree on the distance DAS between MLGs.

Finally, we confronted the NJ trees with the spatial repartition of MLGs in the grapevine field and with MLGs admixed by STRUCTURE Software. To visualise the spatial repartition of the genotypes, we coloured each individual on both NJ trees and field maps in accordance with its assignment by STRUCTURE: gray, orange and green for individuals assigned to cluster 1, 2 and 3, respectively, and hatched in corresponding colours for admixed individuals.

Population differentiation. The pairwise F_{ST} values between grapevine fields (Weir & Cockerham 1984) were estimated using GENEPOP version 4.0 (Rousset 2008). To test for isolation by distance at both scales of vineyard area and grapevine plot, we used a standard Mantel test to assess the correlation between genetic distance ($F_{ST}/(1 - F_{ST})$) and geographic distance between fields (calculated with GSP coordinates [GPS GS50 Sensor, Leica, 0.5–5 m precision navigation]) and between sampling points within a grapevine field (measured with a decametre). These analyses were performed using the PASSAGE software (Rosenberg 2002). Because we used microsatellites, because *X. index* is suspected to have low dispersal abilities and because we had no a priori information to choose between allele identity and allele-size-based statistics, we tested for all loci the information content of allele size regarding global population structure (Hardy *et al.* 2003; Kerth & Petit 2005) using SPAGEDI (Hardy &Vekemans 2002). Tests were nonsignificant ($P > 0.125$ for all loci), indicating that allele sizes did not carry information about population structure in this data set, and that allele identity statistics should be used instead (Hardy *et al.* 2003).

The spatial distribution of genetic variance was analysed by a hierarchical analysis of molecular variance in which total genetic variance was partitioned into four hierarchical levels: (i) among vineyard areas, (ii) among *Châteaux* within the same vineyard area, (iii) among grapevine fields within a *Château*, and (iv) among sampling points. These analyses were performed using HIERFSTAT package (Goudet 2005) for R (R Development Core Team, 2007), based on the methods described by Yang *et al.* (1998) and following De Meeus & Goudet (2007). HIERFSTAT analyses were performed with and without MLGs. The significance of the hierarchical F statistics was tested by 1000 randomizations of each

relevant factor among the entities belonging to the next level (e.g. nematodes individuals among sampling point), keeping them separated for other factors (e.g. grapevine plot).

Results

Allelic and genotypic diversities

Among the seven microsatellite loci analysed, the allelic diversity varied between 4 and 7 alleles per locus and between 13 and 32 alleles per population (Table 1, Supporting Information 1). The MA population (for which the sample design covered the largest area) displayed the highest allelic richness, ranging from 22 (for a sample size of 5) to 32 (for a sample size of 70). The *T* population (for which the sample design covered the smallest area) exhibited the lowest allelic richness, ranging from 11.8 (for a sample size of 5) to 13 (for a sample size of 45).

We found 81 distinct MLGs among the 303 individuals genotyped (Fig. 2). There were 53 MLGs represented by only one individual in the data set (hereafter called 'single copy MLG') and the remaining 250 individuals shared only 27 MLGs ('repeated MLGs'). The highest number of individuals sharing an MLG was 40 (MLG 48 in T population: Fig. 2). Each of the 81 MLGs was specific to the field of origin, with the exception of 4 MLGs that were shared by two fields (Fig. 2).

Reproductive mode

Clonal diversity. All populations (except T) showed low $P_{(ID)}$ values (Table 2). This indicates that the probability that two individuals sharing the same MLG came from sexual reproduction events is low. Thus, it is likely that individuals are genetically identical because they belong to the same asexual lineage. The resolution of our markers is thus sufficient to ensure that individuals sharing the same MLG are likely to be asexual clones for all populations except *T*. In the *T* population, $P_{(ID)}$ values indicated a 1–10% probability that two individuals possess the same MLG through sexual reproduction.

The global g/N [ratio of the number of MLGs (g) found over the sample size (N)] was 0.27. By using the immediate rates of asexual reproduction (estimated as $1 - g/N$) to estimate the putative range of effective asexual rate (e.g. c in Balloux *et al.* 2003), we determined that around 53–92% of the individuals came from asexual events (Table 2). Twenty-one to 71% of the MLGs observed in each population were included in a repeated MLG. The highest genotypic diversities (g/N = 0.39 and 0.47) and the highest numbers of repeated MLGs (12 and 5) were found in MA and MB populations, but these repeated MLGs were represented by fewer individuals than the repeated MLGs in the other populations (a maximum of 13 and 16 individuals, respectively).

Linkage disequilibrium and genetic departure from panmixia. All pairwise combinations of loci presented significant linkage disequilibrium when the analysis included all individuals in a repeated MLGs or only one individual per MLG ($P < 0.01$). Over all the 6 populations, the composite measures of LD (R^2) were 0.246 ± 0.114 (Table 1).

The values for F_{IS}, *He*, *Ho*, the standard error of F_{IS} across loci, and the *P*-values for the heterozygote excess test are given in Table 1. All populations exhibited significant negative overall loci F_{IS}. Thus, there was significant excess of heterozygotes in all populations except for the MA population, which displayed a significant deficit in heterozygotes, and the MB population, for which genotypic distributions did not differ from Hardy–Weinberg expectations (Table 1). The populations T, C and G exhibited strong negative F_{IS} values (≈ -1) that did not vary greatly across loci. The population H displayed F_{IS} values that varied among loci from nearly −1 to nearly +1. F_{IS} values in the MB population were close to 0 with little variation across loci. The population MA displayed a significantly positive F_{IS} value with a moderate variance among loci.

When placing our observed values on theoretical curves of *Fis* and standard error as function of the clonal reproduction rate (Balloux *et al.* 2003), we obtained a high corresponding clonality rate ranging from 0.95 for MA to 0.99–1 for other populations (Fig. 3).

Identification of admixed individuals revealing sexual reproduction events by Bayesian assignment (STRUCTURE and InStruct) and by multivariate analysis (FCA with GENETIX)

The likelihood of clustering estimated by STRUCTURE peaked at $K = 3$ (Fig. 4). This value corresponded to the maximum Evanno's criterion ($\Delta K = 238.82$, ln K = −1939; Supporting Information 2). All individuals of the populations *H* and *T* were assigned only to a single cluster (Cluster 2; probability of assignment >95%; Supporting Information 3 and Fig. 4). The G population consisted of 26 individuals displaying MLGs assigned to cluster 2 (probability of assignment >95%) and of three individuals displaying the same MLG (common with the MA population) assigned to cluster 3. The other populations displayed two kinds of individuals: individuals assigned to one cluster (probability of assignment >95%) and admixed individuals (probabil-

Table 1 Number of alleles observed on each locus in each population (Na), sample sizes (N), expected (He) and observed (Ho) heterozygoties, *Fis*, probabilities of the Hardy–Weinberg test for a heterozygote excess (*p-val*), standard error of *Fis* from a locus to the other in each population (SE) and measure of linkage disequilibrium LD (R^2) and standard deviation SD in each population

Locus	Population	MA $N = 79$	MB $N = 49$	H $N = 44$	C $N = 52$	M $N = 30$	T $N = 49$	Σ *pop* $N = 303$
XM3	Na	5	4	2	2	3	2	6
	He	0.603	0.7	0.503	0.489	0.391	0.505	0.532
	Ho	0.027	0.771	0.932	0.82	0.267	1	0.636
	Fis	0.9561	−0.1017	−0.8696	−0.6897	0.3216	−1	−0.1015
	p-val	***	***	***	***	***	***	***
Xi16	Na	4	3	1	2	3	2	4
	He	0.609	0.608	0	0.129	0.188	0.04	0.262
	Ho	1	0.939	0	0.136	0.2	0.041	0.386
	Fis	−0.6494	−0.5522	—	−0.0617	−0.0642	−0.0105	−0.5308
	p-val	***	***	—	NS	NS	NS	***
Xi22	Na	3	2	2	3	3	2	4
	He	0.671	0.41	0.506	0.575	0.598	0.505	0.544
	Ho	0.679	0	1	0.981	0.897	1	0.759
	Fis	−0.0132	1	−1	−0.7168	−0.5135	−1	−0.3535
	p-val	***	***	***	***	***	***	***
Xi24	Na	7	3	4	3	2	1	7
	He	0.639	0.615	0.67	0.511	0.499	0	0.489
	Ho	0.026	0.735	1	0.882	0.862	0	0.584
	Fis	0.9601	−0.1963	−0.5016	−0.7401	−0.75		−0.0156
	p-val	***	***	***	***	***	—	***
Xi27	Na	6	5	2	3	4	2	6
	He	0.806	0.753	0.455	0.553	0.536	0.505	0.601
	Ho	0.671	0.771	0	0.962	0.9	1	0.717
	Fis	0.169	−0.0241	1	−0.7502	−0.6985	−1	−0.1422
	p-val	***	***	***	***	***	***	***
Xi29	Na	3	3	2	2	3	2	4
	He	0.279	0.465	0.505	0.498	0.554	0.151	0.409
	Ho	0.304	0.286	0.955	0.885	1	0.163	0.599
	Fis	−0.0906	0.3883	−0.9111	−0.7895	−0.8297	−0.0787	−0.4091
	p-val	NS	***	***	***	***	NS	***
Xi32	Na	4	3	2	2	4	2	5
	He	0.494	0.634	0.506	0.5	0.6	0.505	0.54
	Ho	0.329	0.939	1	0.904	1	1	0.862
	Fis	0.3357	−0.4874	−1	−0.8214	−0.686	−1	−0.5109
	p-val	***	***	***	***	***	***	***
∑ *loci*	Na	32	23	15	17	22	13	36
	He	0.586	0.598	0.449	0.465	0.481	0.316	0.481
	Ho	0.434	0.634	0.698	0.796	0.732	0.601	0.649
	Fis	0.261	−0.061	−0.564	−0.723	−0.537	−0.918	−0.27119
	p-val	NS	NS	**	**	**	**	***
	SE Fis	0.0655	0.0778	0.119	0.037	0.0795	0.0713	0.0119
Mean LD		0.285	0.129	0.483	0.219	0.214	NA*	0.246
SD		0.273	0.146	0.258	0.149	0.119	NA*	0.114

*Measure not available because of the few different genotypes.
***$P < 0.001$, **$0.001 < P < 0.01$, *$0.01 < P < 0.05$, NS $P > 0.05$.

ity of assignment is >5% in at least two different clusters). The proportion of admixed individuals is 7.6% for the MA population, 14.3% for the MB population and 11.5% for the C population (Supporting Information 3). We found the same clustering with InStruct program (Supporting Information 4).

FCA performed on the whole data set resulting in the partitioning of individuals into three clouds of points corresponding to the three clusters identified with STRUCTURE. Admixed individuals (as estimated using STRUCTURE) are located at intermediate positions among these three groups in the factorial plan (Fig. 5).

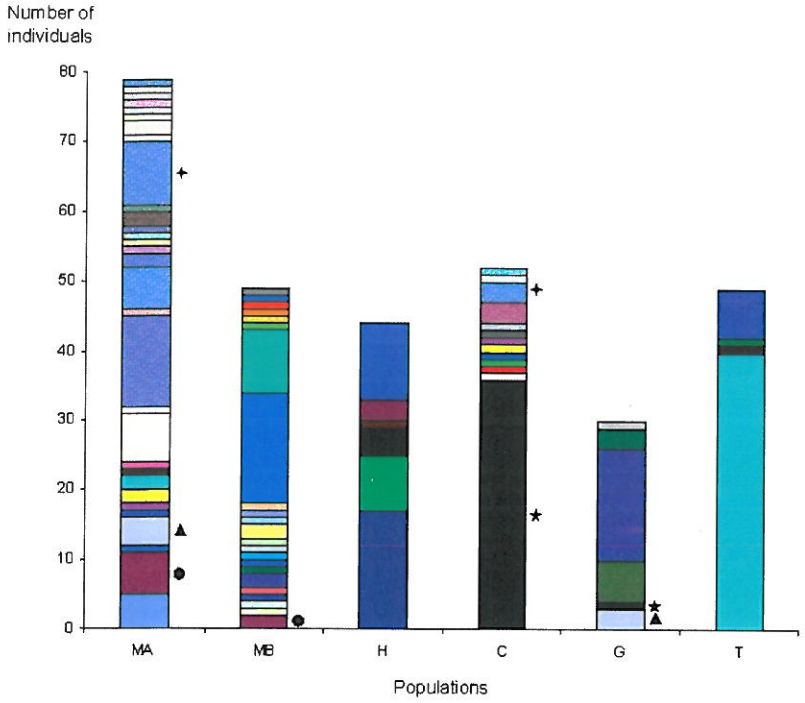

Fig. 2 Genotypic diversity in the six *X. index* populations. Each of the 81 multilocus genotypes (MLGs) is symbolised by one colour. The four MLGs shared by two populations are indicated by circles, triangles, crosses and stars.

Table 2 Genotypic diversity of X. index populations

	MA	MB	C	H	G	T
Sample size (N)	79	49	52	44	30	49
Number of genotypes (g)	31	23	14	7	6	4
(1 – g/N)	*0.61*	*0.53*	*0.73*	*0.84*	*0.80*	*0.92*
Number of clones *(in percentage of genotypes)*	12 (39%)	5 (22%)	3 (21%)	5 (71%)	4 (67%)	2 (50%)
Number of genotypes not included in a clone	19	18	11	2	2	2
Median number of individuals	4,5	2	3	8	5	23,5
per clone (min-max)	2–13	2–16	3–35	3–17	3–16	7–40
$P_{(ID)\ Unbiased}$	1.66e–005	1.9e–005	0.001	0.001	0.0005	0.012
$P_{(ID)\ Sib}$	0.009	0.009	0.033	0.036	0.028	0.102

N, the sample size; *g*, the number of genotypes; $1 - g/N$, the immediate asexual rate; the number of clones and the percentage of genotypes included in a clone; the number of genotypes not included in a clone; the median, minimum and maximum numbers of individuals per clone; and $P_{(ID)Unbiased}$, the upper and $P_{(ID)Sib}$, the lower bounds of the probability of identity calculated from the seven microsatellites genotypes in each population are given for MA, MB, *C*, *H*, *G* and *T* populations.

Identification of admixed individuals revealing sexual reproduction events by NJ trees.

For most populations, the topology of the NJ unrooted trees using DAS distances between MLGs are congruent with the clustering estimated by STRUCTURE (Figs 6–8). The fact that circles of the same colour (coloured according to the cluster assigned by STRUCTURE) are often found in the same clades in the NJ tree supports this interpretation. The green, grey and orange clades will be called 'parental clades' hereafter. The branches corresponding to admixed MLGs identified by STRUCTURE are located in intermediate position between parental clades (for example the location of MLGs 16 through 23 in the MB tree, Figs 6 and 8).

MLGs obtained by simulation of sexual reproduction events between individuals belonging to different parental clusters are found at a similar intermediate position within the trees (Supporting Information 5). For the *H*, *G* and *T* populations, we observed very small genetic distances between the few MLGs observed (see the scales associated to those NJ trees in Fig. 7). This result is also consistent with our observation that these 3 grapevine fields had the lowest allelic diversity (Table 1 and Supporting Information 1). The *H* and *T* populations are composed of one or two major

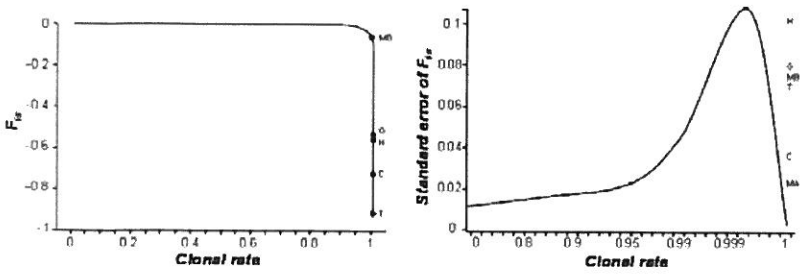

Fig. 3 Theoretical curves of the relationship between *Fis* (left graph) and its standard error (right graph) and the clonal rate *c* (after Balloux *et al.* 2003). Observed values for MA, MB, H, C, G and T populations are figured in black dot and dotted lines.

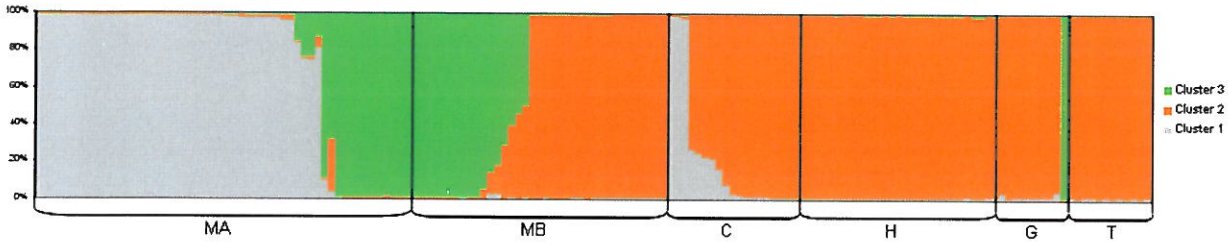

Fig. 4 Estimated structure of the *X. index* populations for $K = 3$. Each nematode is represented by a thin vertical line, which is partitioned into 3 segments representing the individual's estimated membership fractions in the 3 clusters estimated using Structure. Populations MA, MB, C, H, G and T are indicated below.

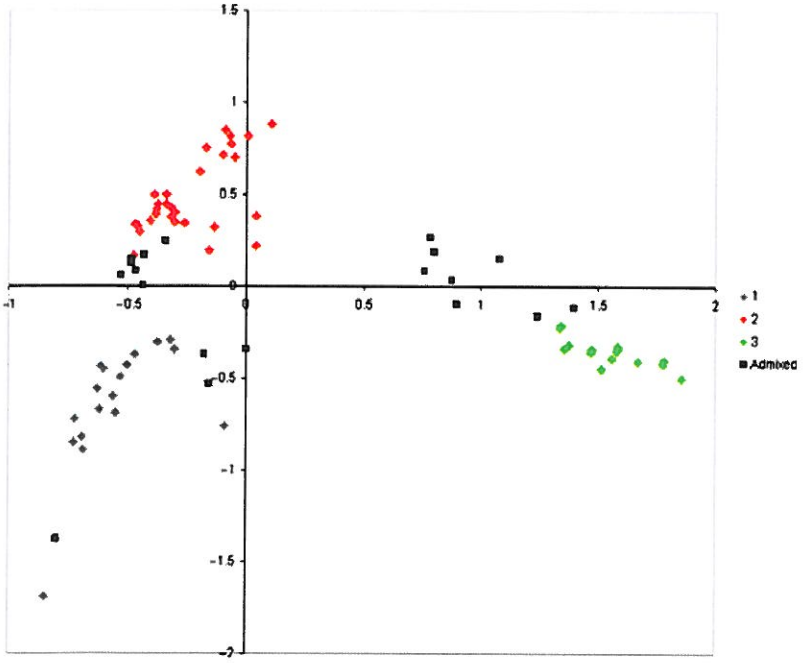

Fig. 5 Factorial Correspondence Analysis (FCA) on the first two axis performed using Genetix. Grey, orange and green diamonds are individuals assigned by STRUCTURE to cluster 1, 2 and 3, respectively; black squares are the admixed individuals.

MLGs and a few additional genetically close MLGs. We observed a different pattern in the G population, however, which had five genetically close MLGs (assigned to cluster 2) and also three individuals with a more genetically distant MLG (assigned to cluster 3) that was shared with MA population (see the NJ tree for field G, Fig. 8).

Localisation of sexual reproduction events in grapevine fields

The three methods we used to identify admixed individuals provided similar results. We then mapped, to the sampling scheme of each grapevine field, the MLGs (numbers in circles, Figs 7 and 8) and their assignment

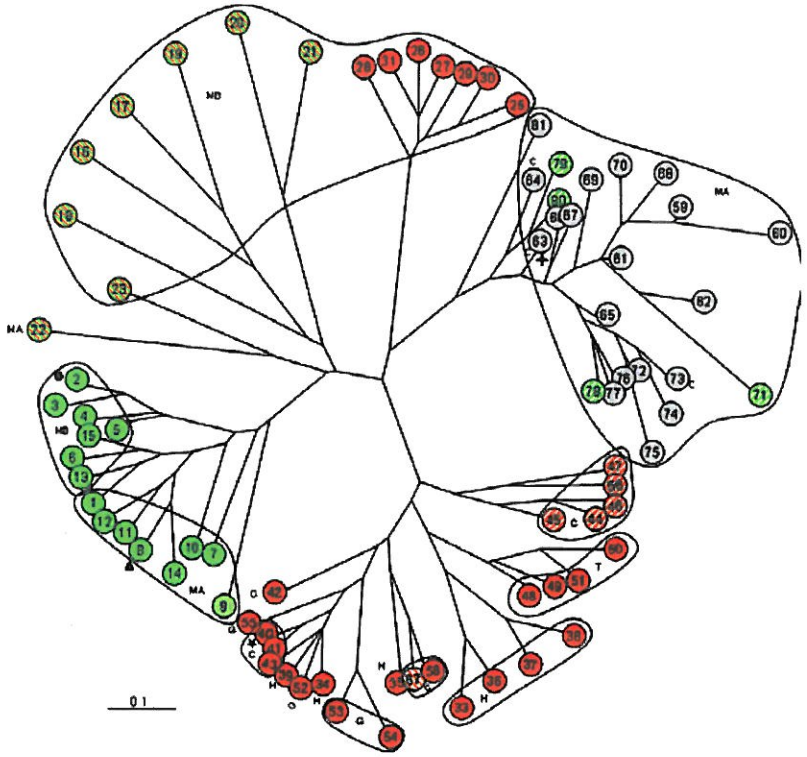

Fig. 6 Neighbour-Joining trees on the distance DAS between multilocus genotypes (MLGs) for the six populations. Each circle represents a MLG (given by its number) and is coloured according to its assignment to clusters estimated by STRUCTURE. Admixed MLGs are represented by circle hatched with the two colours of the two parental clusters. The four MLGs shared by two populations are indicated by circles, triangles, crosses and stars. The geographical origin is given (MA, MB, H, C, G and T).

to one of the three clusters identified by STRUCTURE (orange, grey and green colours) or to admixed MLGs identified by STRUCTURE (hatched circles). Since the clustering obtained by STRUCTURE does not provide any information concerning the genetic relatedness among MLGs, this map cannot completely depict the spatial distribution of genetic variability. Therefore, we included the corresponding NJ tree for each population next to the map of the sampling design of this population. The MLGs appeared spatially clustered according to the group identified by STRUCTURE (i.e. circles of the same colour were not scattered randomly within the grapevine fields). For example, individuals assigned to the green cluster in the MA map were found in one part of the field and those assigned to the grey cluster were found in another part of the field. Admixed individuals were mainly found at the interface between two clusters and did not appear randomly located (for example, the MLGs 44–57 in the C field in Fig. 8). The location of admixed individuals at the interface between two 'parental clades' was observed in most fields where admixed individuals have been identified (Cf. populations MA and MC, Fig. 8).

Population differentiation and genetic structure

Hierarchical F-statistics. The matrixes of hierarchical *F*-statistics are given in Table 3. With or without repeated MLGs, we obtained significant values (*P*-value of 0.001) for $F_{GrapevineField/Château}$, $F_{GrapevineField/Area}$, $F_{GrapevineField/Total}$ (differentiation among grapevine fields within each *Château*, each area and total) and $F_{SoilVolume/GrapevineField}$, $F_{SoilVolume/Château}$, $F_{SoilVolume/Area}$, $F_{SoilVolume/Total}$ (differentiation among soil samples within each grapevine plot, each *Château*, each area and total). No significant differentiation between vineyard areas and between *Château* was detected (Table 3).

Genetic differentiation. Pairwise *Fst* estimates ranged from 0.116 to 0.365 (Table 4). At the grapevine field scale, F_{ST} estimates between soil samples (nematodes grouped by sampling point) ranged from 0 to 0.05 for *T* (mean distance between soil samples = 7.2 m), 0 to 0.08 for *G* (mean distance = 45.7 m), 0 to 0.19 for *H* (mean distance = 38.5 m), 0 to 0.38 for C (mean distance = 33.6 m), 0 to 0.54 for MB (mean distance = 25.8 m) and 0 to 0.76 for MA (mean distance = 44.5 m).

Isolation by distance. At the scale of the vineyard area, no significant isolation by distance was found (r_z = 0.075, t = 0.423, two-tailed P = 0.143). For all populations, Mantel's tests showed very low and no significant correlation between F_{ST} and geographical distances at the scale of the grapevine field (Table 5). However, a

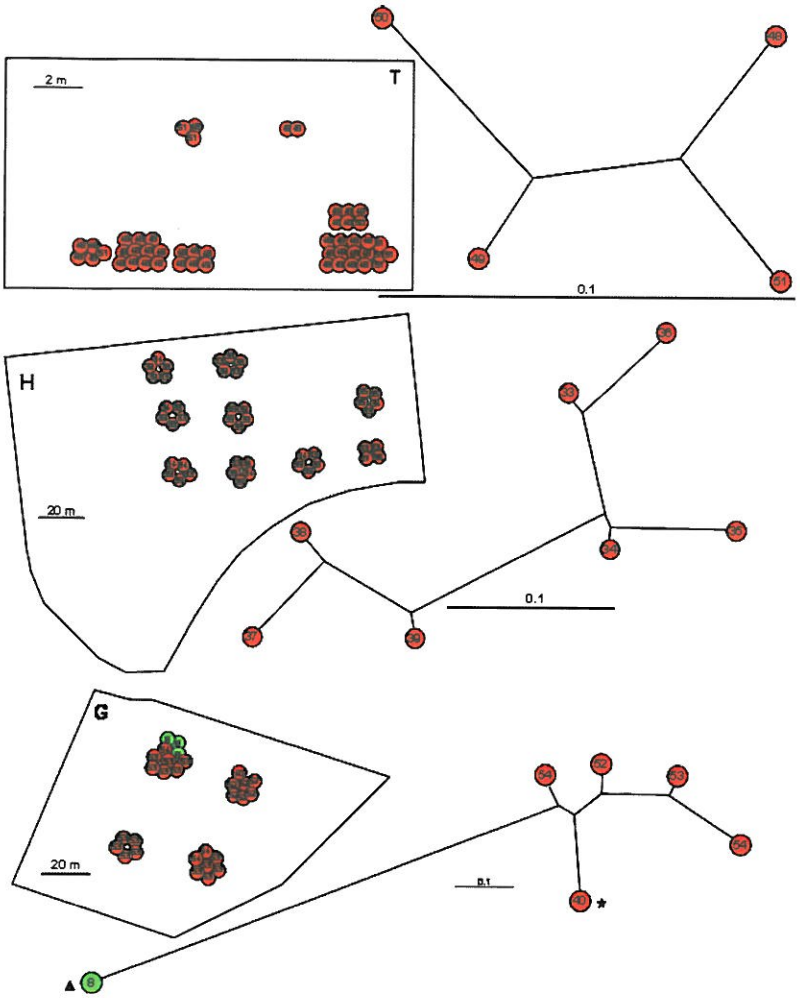

Fig. 7 Neighbour-Joining trees on the distance DAS between multilocus genotypes (MLGs) and spatial repartition of MLGs within each of the grapevine fields H, G and T. On trees, each circle represents a MLG (given by its number) and is coloured according to its assignment to clusters estimated by STRUCTURE. On grapevine field maps, each nematode individual is represented by a circle containing its genotype (number) and coloured according to its assignment to clusters estimated by STRUCTURE. Each set of circles represent nematode individuals from the same 2-L soil sample. Triangle and star represent the two MLGs found in other fields.

significant correlation between those two variables was found in the MA population (two-tailed $P = 0.002$).

Discussion

In mainly parthenogenetic species, the identification of rare sexual reproduction events is a challenging task. The existence of a unique genotype surrounded by numerous individuals of the same clones in a given population might result from immigration or mutation rather than from a rare sexual event. In this study, by taking advantage of (i) the specificities of the organism studied (soil-dwelling habit, poor active dispersal abilities and a slow reproductive cycle), (ii) an accurate and spatially determined sampling scheme and (iii) hypervariable nuclear markers, we have been able to disentangle the confounding effects of various evolutionary forces. We demonstrated that this species has a predominantly asexual mode of reproduction with all the expected genetic signatures of this mating system (strong linkage disequilibrium and heterozygote excess). We used two different methods (F_{IS} analysis and the g/N ratio) for the quantitative estimation of asexual reproduction, which determined a clonality rate from 95–99% or 53–92%, respectively. A few individuals exhibited genetic characteristics suggesting they had resulted from rare sexual reproduction events. These individuals were assigned to several 'parental clades' and were found at intermediate positions between the clades in the Neighbour-Joining tree. These characteristics are expected for individuals resulting from recombination events. Moreover, simulations of MLG resulting from sexual recombination confirmed these expectations. Finally, we mapped the geographical location of all MLGs in each grapevine field. Analysis of the spatial distribution of MLGs and the genetic distances among them allowed us to determine that the hybrid individuals resulting from sexual reproduction events were located at the interface between two groups of individuals of each 'parental clade'.

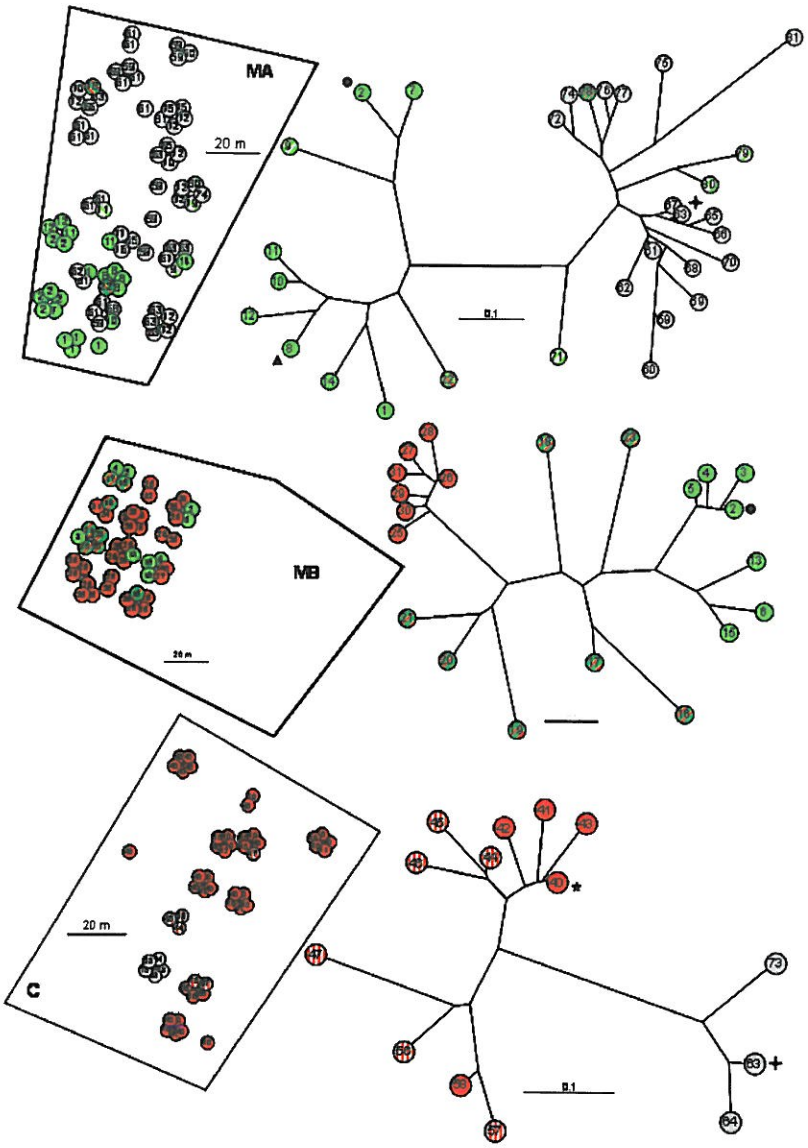

Fig. 8 Neighbour-Joining trees on the distance DAS between multilocus genotypes (MLGs) and spatial repartition of MLGs within each of the grapevine fields MA, MB and C. Star, triangle, crosses and circles represent MLGs found in other fields.

Reproductive mode

Evidence for a mainly clonal reproductive mode. The data observed for all six populations provided evidence that the main reproductive mode of *X. index* in the Bordeaux vineyard is asexual. The estimated immediate rate of asexual reproduction, $1 - g/N$, indicated that between 53% and 92% of the individuals came from asexual events. The MA and MB populations displayed higher genotypic diversities, with clone patches composed of fewer individuals than observed for the other four populations. Individuals exhibited a lower immediate rate of asexual reproduction ($1 - g/N$) and either no divergence from the Hardy–Weinberg expectation (MB), or a non significant deficit in heterozygotes (MA). This indicates a greater rate of sexual reproduction in MA and MB populations. In the four other populations, we found strong significant heterozygote excesses. Two main hypotheses can explain heterozygote excesses: selection processes (overdominance and associative overdominance, Hartl & Clark 1997) or an independent accumulation of mutations in alleles over time as a consequence of asexual reproduction (Judson & Normark 1996; Balloux *et al.* 2003). In our case, negative *Fis* values were associated with genome-wide linkage disequilibrium, which are typical signatures of asexual reproduction (Halkett *et al.* 2005). Consequently, the heterozygote excesses we observed are likely to be due to mutation accumulation over time, as observed in some other organisms like aphids (Delmotte *et al.* 2002).

Table 3 Hierarchical *F*-statistics (in bold) computed by HierFstat and significance of genetic differentiation (in italic) at the different levels (see text for explanation)

	Vineyard area	*Chateau*	Grapevine field	Soil volume
Without repeated MLGs				
Total	0.112*ns*	−0.09*ns*	0.207	0.257
			0.001	*0.001*
Vineyard area	0	−0.235*ns*	0.107	0.164
			0.001	*0.001*
Chateau	0	0	0.277	0.323
			0.001	*0.001*
Grapevine field	0	0	0	0.064
				0.02
Soil volume	0	0	0	0
With repeated MLGs				
Total	0.114*ns*	−0.116*ns*	0.227	0.364
			0.001	*0.001*
Vineyard area	0	−0.260*ns*	0.127	0.282
			0.001	*0.001*
Chateau	0	0	0.307	0.429
			0.001	*0.001*
Grapevine field	0	0	0	0.177
				0.001
Soil volume	0	0	0	0

Table 4 F_{ST} between populations of *X. index*

Fst	MA	MB	H	C	G	T
MA	—					
MB	0.277	—				
H	0.238	0.259	—			
C	0.195	0.262	0.116	—		
G	0.222	0.212	0.175	0.160	—	
T	0.279	0.365	0.224	0.151	0.270	—

Table 5 Mantel's correlation values between $F_{ST}/(1 - F_{ST})$ and geographical distances at the scale of the grapevine field (Mantel's statistic *rz* and two-tailed *P*-value; see text for explanation)

	RZ	two-tailed *P*
MA	0.286	0.002
MB	0.135	0.317
C	−0.075	0.713
H	−0.123	0.508
G	0.211	0.429
T	−0.166	0.337

Without the segregation of alleles due to sexual reproduction, diploid clones become heterozygotes as each allele copy accumulates independent mutation events at all loci over time (Lokki 1976; Pamilo 1987; Tibayrenc & Ayala 2002). This leads to extremely negative values of F_{IS} (Balloux *et al.* 2003). We suggest that a heterozygote excess due to the accumulation of mutations can be observed over a much shorter timescale if rapidly evolving molecular markers such as microsatellite loci are used instead of highly conserved genes.

Rare sexual reproduction events in mainly clonal populations. The few reports of occurrence of males in the field (1–2%, Harris 1977) attest to the rarity of sexual reproduction events in populations. Populations *C*, *H*, *M* and *T* displayed strongly negative F_{IS} with high variability among loci and high linkage disequilibria for all loci pairs. This pattern is consistent with a strongly clonal mode of reproduction with rare events of sexual reproduction in each generation (De Meeus *et al.* 2006). The immediate rate of asexual reproduction ranged from 53% to 92%, but the observed F_{IS} values are consistent with a higher degree of clonality. Indeed, by plotting our observed values of F_{IS} and associated standard errors as functions of the clonal reproduction rate on the theoretical curves of Balloux *el al.* (2003), we found that the rates of clonal reproduction in our 6 populations were very close to 1 (0.95 for MA, 0.99–1 for the others; Fig. 3).

Spatial distribution of genotypes and localisation of sexual reproduction offspring

The spatial distribution of different clones is a fundamental property of the genetic structure of clonal populations. To our knowledge, however, there is no standardized method to characterize the genetic structure of clonal organisms in space. The Bayesian analyses performed with STRUCTURE software assume panmixia and that each locus is at Hardy–Weinberg equilibrium and independent of the others. We verified that this approach is robust to these deviations by performing analyses with only one copy of each MLG and by obtaining the same results with the AFC performed with GENETIX, and the NJ trees.

We assumed that single copy MLGs may result from two different processes: (i) they can result from a punctual mutation and should thus differ from other individuals at only one locus (a common situation due to the high mutation rate of microsatellite loci), or (ii) they can result from a sexual reproduction event and thus they should differ from other individuals at several loci. The two hypotheses can be distinguished by the number of differences between an individual and its nearest relative in phylogenetic trees, with the individuals resulting from sexual reproduction being associated with longer branches (for example, compare the individuals 25–31 and individuals 16–23 in the NJ tree of

population MB in Fig. 8). We found admixed MLGs, suspected to be offspring of sexual reproduction, only in the MA, MB and C populations. Confirmation that these individuals are likely to have resulted from sexual reproduction was achieved through the analysis of a NJ tree in which simulated recombinant individuals are added in observed data sets (see Supporting Information 5. All of the admixed MLGs we identified by STRUCTURE had longer branches on the NJ trees and were thus more genetically distant from the others. For the MB population, MLGs belonging to clusters 2 (individuals 25–31) and 3 (individuals 2–15) are grouped in two sets of genetically close MLGs (Fig. 8). The individuals within each of these clusters only differ by a couple of mutations (small DAS distance). The MLGs of admixed individuals (16–21) shared fewer alleles with the other individuals in the population (larger DAS distance). Thus, we conclude that these genotypes come from genome-wide recombination through infrequent sexual reproduction.

Localisation of sexual reproduction in grapevine fields. By taking advantage of the limited active dispersal of *X. index*, we were able to localise sexual reproduction events in the field. We achieved this goal by creating a map of MLGs resulting from recombination and those assigned to 'parental clades'. This map and the genetic relationship among MLGs determined by the NJ trees confirm our hypothesis concerning the mating system and gene flow within grapevine fields. First, the location of the individuals assigned to each parental clade appeared highly structured in space. This observation is congruent with the aggregative spatial pattern of *X. index* (Villate *et al.* 2008). Moreover, admixed individuals that are likely to have resulted from sexual reproduction events based on the NJ trees were mostly localised to the physical interfaces between two parental clades in the field. Because the genetic signature of recombination is more easily detected when reproduction occurs between genetically more distantly related individuals, individuals resulting from sexual reproduction events can be more easily identified in such interfaces. Moreover, there is an overrepresentation of single copy MLGs in those interface areas. This could be due to the fact that a single sexual reproduction event in those areas may have generated numerous genetically unique individuals. Environmental stress is known to initiate sexual reproduction in a broad range of species that normally undergo asexual reproduction (Bell 1982; Harris 1989; Dubnau 1991; Kleiven *et al.* 1992; Gemmil *et al.* 1997; Dacks & Roger 1999). For example, facultative outcrossing in *C. elegans* has been showed to be favoured in starvation conditions (Morran *et al.* 2009). The overlapping of two *X. index* patches could create stressed conditions that could favour male production and outcrossing.

Genetic differentiation at different spatial scales

Limited gene flow between grapevine fields and between areas through passive dispersal. Analyses performed with HIERFSTAT revealed a strong effect of grapevine fields and of sampling point on the genetic structure of *X. index* (P-value = 0.001). There was a significant differentiation among grapevine fields. However, the observation of a few MLG shared by two populations (Figs 2, 6, 7 and 8) suggests that passive dispersal of nematodes through wine growing activities could have occurred at a low frequency. Biological and ecological traits of *X. index* suggest that passive dispersal may be a rare event. This nematode lives deep in the soil and thus cannot be dispersed by wind or water and can only rarely be disturbed by field machinery. Its relatively large size makes it sensitive to mechanical manipulation and the hydric stress that occurs when earth is turned upside down. Passive dispersal through replanted grapevines from nurseries is a more likely explanation for the observed pattern. Grapevine plants in the Bordeaux vineyard region come from the same nurseries. So if the observed pattern was due to contamination from nursery plants, we would expect a lower genetic differentiation between nematodes populations from different *châteaux* or grapevine fields. But the pattern observed could still be due to the introduction of nematodes from nursery, considering the combined effect of ancestral polymorphism (ancient introduction) and high mutation rates at the microsatellite loci investigated.

An alternative explanation to recent migration is that the MLGs shared between populations reveal ancestral polymorphism that has been conserved independently in those pairs of populations. However, this explanation has a low likelihood as each 'shared MLG' is different for each pair of populations sharing a common MLG.

Strong genetic structure due to the limited active dispersal. No isolation by distance was observed at the grapevine field scale for most populations, but we found a strong differentiation between sampling points within the same grapevine field. This pattern is consistent with limited capacities of active dispersal leading to a patchy spatial distribution within grapevine fields (Villate *et al.* 2008). The clonal reproductive mode may also explain the absence of isolation by distance pattern and the strong genetic differentiation observed (consequences of the Meselson effect at the population scale). We did find a significant correlation between *Fst* and geographical distances at the scale of the grapevine field in the MA

population. It is possible that this is because of the higher rate of sexual reproduction in this population. Finally, the variable number of MLGs observed among populations could also reflect a variation in the age of the populations. In the youngest populations, an insufficient number of generations may have occurred to produce any isolation by distance pattern.

Incidence of the reproductive mode in pathogen populations

Control strategies implemented against parthenogenetic pests might be different from those used for sexual species. Multiple genes underlie the mechanisms involved in adaptation to pesticide or host resistance plants. The occurrence of sexual reproduction in *X. index* populations should have profound consequences for prophylactic policies against this nematode with regard to the development and use of resistant rootstocks. According to McDonald & Linde (2002), pathogen populations with a high evolutionary potential (mixed reproduction system, high potential for genotype gene flow, large effective population size and high mutation rate) are more likely to overcome genetic resistance than pathogen populations with a low evolutionary potential (strict asexual reproduction, low potential for gene flow, small effective population size and low mutation rates). In this way, the rare occurrence of sexual reproduction in *X. index* is expected to favour more rapid adaptation to chemical or host-based strategies for controlling this pest.

By taking advantage of the biology of the nematode *X. index*, combined with highly polymorphic markers and an accurate sampling strategy at a suitable scale, we have been able to identify and locate individuals resulting from sexual reproduction in a mainly parthenogenetic species. The simulation approach we used to generate genotypes of individuals resulting from sexual reproduction confirms our interpretation that individuals located at intermediate positions in unrooted trees can be considered to have resulted from sexual reproduction. Although the combination of both geographical and genetic data is an old quest for population geneticists, this aim has not yet been realized due to a lack of suitable methods. A new method proposed by Guillot *et al.* (2005) or Manel *et al.* (2007) represents significant progress in this area. The combined analyses of phylogenetic trees and a map of sampling designs could be another way to achieve this goal (see for example Biek *et al.* 2007). In our study, we not only have mapped genetic clades onto a sampling design, but we have also been able to accurately locate a particular biological process: rare reproduction events in a mainly parthenogenetic species. We suggest that this method could also be developed for other asexual species. Fungi or plants, in which mating systems are often variable and complex, could also be investigated with this method.

Acknowledgements

The authors would like to thank the estates and their vineyard managers that have funded this research. We thank Bernard Chauvin, Florent Delemarre and Elisa Morin for their help during the sampling, Jean Peccoud for punctual help during analysis, Solène Coedel from the OUEST-Genopole Genotyping platform and the Cornell Computational Biology Service Unit (CBSU). We also thank Thierry De Meeus for helpful comments and advices for the sampling design and analyses, Eric Petit, Philippe Jarne, three anonymous referees and the subject editor François Balloux for relevant comments and advices on earlier versions of the manuscript. We finally thank Beth Rowan (MPI Tübingen, Germany) for providing valuable corrections and comments on the manuscript.

References

Abawi GS, Widmer TL (2000) Impact of soil health management practices on soilborne pathogens, nematodes and root diseases of vegetable crops. *Applied Soil Ecology*, **15**, 37–47.

Awadalla P (2003) The evolutionary genomics of pathogen recombination. *Nature Reviews Genetics*, **4**, 50–60.

Baïlle D, Barrière A, Félix M-A (2008) *Oscheius tipulae*, a widespread hermaphroditic soil nematode, displays a higher genetic diversity and geographical structure than *Caenorhabditis elegans*. *Molecular Ecology*, **17**, 1523–1534.

Balloux F, Lehmann L, De Meeus T (2003) The population genetics of clonal and partially clonal diploids. *Genetics*, **164**, 1635–1644.

Barrière A, Félix M-A (2005) High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. *Current Biology*, **15**, 1176–1184.

Barrière A, Félix M-A (2007) Temporal dynamics and linkage disequilibrium in natural *Caenorhabditis elegans* populations. *Genetics*, **176**, 999–1011.

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

Bell G (1982) *The Masterpiece of Nature: The Evolution and Genetics of Sexuality*. University of California Press, Berkeley, CA.

Bengtsson BO (2003) Genetic variation in organisms with sexual and asexual reproduction. *Journal of Evolutionary Biology*, **16**, 189–199.

Biek R, Henderson C, Weller LA, Rupprecht CE, Real LA (2007) A high-resolution genetic signature of demographic and spatial expansion in epizootic rabies virus. *PNAS*, **104**, 7993–7998.

Bouquet A, Marck G, Pistagna D, Torregrosa L (2003) Transfer of grape fanleaf virus coat protein gene through hybridization with *Xiphinema index* resistant genotypes to obtain rootstocks resistant to virus spread. *Proceedings of the*

8th International Conference on Grape Genetics and Breeding, Vols 1 and 2, 325–334.

Bouquet A, Torregrosa L, Chatelet R (2004) Combination of biotechnological and conventional approaches to rootstock selection presenting a sustainable resistance grape fanleaf disease transmission. *Bulletin de l'OIV*, **77**, 361–376.

Brown DJF, Weischer B (1998) Specificity, exclusivity and complementarity in the transmission of plant virus parasitic nematodes: an annotated terminology. *Fundamental and Applied Nematology*, **21**, 1–11.

Burt A, Carter DA, Koenig GL, White TJ, Taylor JW (1996) Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. *PNAS*, **93**, 770–773.

Cutter AD (2006) Nucleotide polymorphism and linkage disequilibrium in wild populations of the partial selfer *Caenorhabditis elegans*. *Genetics*, **172**, 171–184.

Dacks J, Roger AJ (1999) The first sexual lineage and the relevance of facultative sex. *Journal of Molecular Evolution*, **48**, 779–783.

Dalmasso A (1970) La gamétogénese des genres *Xiphinema* et *Longidorus* (*Nematoda* : *Dorylaimida*). *Comptes Rendus de l'Académie des Sciences, Paris*, **270D**, 824–827.

Dalmasso A, Younes T (1969) Ovogénèse et embryogénèse chez *Xiphinema index* (*Nematoda* : *Dorylaimida*). *Annales de Zoologie et Ecologie Animale*, **1**, 265–272.

De Meeus T, Balloux F (2004) Clonal reproduction and linkage disequilibrium in diploids: a simulation study. *Infection Genetics and Evolution*, **4**, 345–351.

De Meeus T, Balloux F (2005) *F*-statistics of clonal diploids structured in numerous demes. *Molecular Ecology*, **14**, 2695–2702.

De Meeus T, Lehmann L, et al. (2006) Molecular epidemiology of clonal diploids: a quick overview and a short DIY (do it yourself) notice. *Infection Genetics and Evolution*, **6**, 163–170.

De Meeus T, Goudet J (2007) A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels. *Infection Genetics and Evolution*, **7**, 731–735.

De Meeus T, Prugnolle F, Agnew P (2007) Asexual reproduction: genetics and evolutionary aspects. *Cellular and Molecular Life Sciences*, **64**, 1355–1372.

Delmotte F, Leterme N, Gauthier JP, Rispe C, Simon JC (2002) Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. *Molecular Ecology*, **11**, 711–723.

Dubnau D (1991) Genetic competence in *Bacillus subtilis*. *Microbiological Reviews*, **55**, 395–424.

Eisenback JD, Triantaphyllou HH (1991) Root-knot Nematodes: *Meloidogyne* species and races. In:*Manual of Agricultural Nematolog* (ed Nickle WR). pp. 191–274, Marcel Dekker, New York.

ElMousadik A, Petit RJ (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.

Esmenjaud D, Walter B, Valentin G, Guo ZT, Cluzeau D (1992) Vertical-distribution and infectious potential of *Xiphinema index* (Thorne et Allen, 1950) (*Nematoda, Longidoridae*) in fields affected by grapevine fanleaf virus in vineyards in the Champagne region of France. *Agronomie*, **12**, 395–399.

Esmenjaud D (2000) Les nématodes de la vigne. In *Ravageurs de la vigne*. E. Féret. Bordeaux, France, J. Stockel, 17–29.

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.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Gao H, Williamson S, Bustamante CD (2007) An MCMC Approach for joint inference of population structure and inbreeding rate from multi-locus genotype dada. *Genetics* (online).

Gemmil AW, Viney ME, Read AF (1997) Host immune status determines sexuality in a parasitic nematode. *Evolution*, **51**, 393–401.

Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.Available from http://www2.unil.ch/popgen/software/fstat.htm.

Goudet J (2005) HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, **5**, 184–186.

Goyeau H, Halkett F, Zapater MF, Carlier J, Lannoux C (2007) Clonality and host selection in the wheat pathogenic fungus *Puccinia triticina*. *Fungal Genetics and Biology*, **44**, 474–483.

Graustein A, Gaspar JM, Walters JR, Palopoli MF (2002) Levels of DNA polymorphism vary with mating system in the nematode genus *Caenorhabditis*. *Genetics*, **161**, 99–107.

Guillot G, Mortier F, Estoup A (2005) Geneland: a computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 712–715.

Haber M, Schungel M, Putz A, Muller S, Hasert B, Schulenburg H (2005) Evolutionary history of *Caenorhabditis elegans* inferred from microsatellites: evidence for spatial and temporal genetic differentiation and the occurrence of outbreeding. *Molecular Biology and Evolution*, **22**, 160–173.

Halkett F, Simon JC, Balloux F (2005) Tackling the population genetics of clonal and partially clonal organisms. *Trends in Ecology & Evolution*, **20**, 194–201.

Harris AR (1977) A first record of *Tylenchulus semipenitrans* and males of *Xiphinema index* in vineyards in north-eastern Victoria. *Australian Plant Pathology Society*, **6**, 1–2.

Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Hardy OJ, Charbonnel N, Freville H, Heuertz M (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics*, **163**, 1467–1482.

Harris EH (1989) *The Chlamydomonas Sourcebook*. Academic Press, New York.

Hartl DL, Clark AG (1997) *Principles of Population Genetics*, 3rd edn. Sinauer Associates Inc., Sunderland, MA, USA .

Hewitt WB, Raski DJ, Goheen AC (1958) Nematode vector of soil-borne fanleaf virus of grapevines. *Phytopathology*, **48**, 586–595.

Honnay O, Bossuyt B (2005) Prolonged clonal growth: escape route or route to extinction? *Oikos*, **108**, 427–432.

Hugot JP, Baujard P, Morand S (2001) Biodiversity in helminths and nematodes as a field of study: an overview. *Nematology*, **3**, 199–208.

Jin L, Chakraborty R (1994) Estimation of genetic-distance and coefficient of gene diversity from single-probe multilocus DNA-fingerprinting data. *Molecular Biology and Evolution*, **11**, 120–127.

Judson OP, Normark BB (1996) Ancient asexual scandals. *Trends in Ecology & Evolution*, **11**, A41–A46.

Kaeuffer R, Réale D, Coltman DW, Pontier D (2007) Detecting population structure using STRUCTURE software: effect of background linkage disequilibrium. *Heredity*, **99**, 374–380.

Karssen G (2002). *The plant parasitic nematode genus Meloidogyne Goldi, 1892 (Tylenchida) in Europe*. Brill Academic Publishers, Leiden, The Netherlands, 157 pp.

Kerth G, Petit E (2005) Colonization and dispersal in a social species, the Bechstein's bat (*Myotis bechsteinii*). *Molecular Ecology*, **14**, 3943–3950.

Kleiven OT, Larsson P, Hobaek A (1992) Sexual reproduction in *Daphnia magna* requires three stimuli. *Oikos*, **65**, 197–206.

Koch R, van Luenen HGAM, van der Horst M, Thijssen KL, Plasterk RHA (2000) Single nucleotide polymorphisms in wild isolates of *Caenorhabditis elegans*. *Genome Research*, **10**, 1690–1696.

LaMunyon CW, Ward S (1997) Increased competitiveness of nematode sperm bearing the male X chromosome. *PNAS*, **94**, 185–189.

Lokki J (1976) Genetic-polymorphism and evolution in parthenogenetic animals .7. Amount of heterozygosity in diploid populations. *Hereditas*, **8**, 57–63.

Luc M, Cohn E (1982) The male of *Xiphinema index* Nematoda Longidoridae. *Revue de Nematologie*, **5**, 211–216.

Lunt DH (2008) Genetic tests of ancient asexuality in Root Knot Nematodes reveal recent hybrid origins. *BMC Evolutionary Biology*, **8**, 1–16.

Manel S, Berthoud F, Bellemain E, Gaudel M, Luikart G, Swenson JE, Waits P, Taberlet P, Intrabiodiv Consortium (2007) A new individual-based spatial approach for identifying genetic discontinuities in natural populations. *Molecular Ecology*, **16**, 2031–2043.

Maynard-Smith J (1998) *Evolutionary Genetics*, 2nd edn. Oxford University Press, Cary, NC, USA xiv + 330 pp.

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, **40**, 349–379.

Milgroom MG (1996) Recombination and the multilocus structure of fungal populations. *Annual Review of Phytopathology*, **34**, 457–477.

Milgroom MG, Fry WE (1997) Contributions of population genetics to plant disease epidemiology and management. *Advances in Botanical Research Incorporating Advances in Plant Pathology*, **24**, 1–30.

Morran LT, Cappy BJ, Anderson JL, Phillips PC (2009) Sexual partners for the stressed: facultative outcrossing in the self-fertilizing nematode *Caenorhabditis elegans*. *Evolution*, **63**, 1473–1482.

Müller J (1992) Detection of pathotypes by assessing the virulence of *Heterodera schachtii* populations. *Nematologica*, **38**, 50–64.

Paetkau D, Waits LP, Clarkson PL, Craighead L, Vyse E, Ward R, Strobeck C (1998) Variation in genetic diversity across the range of North American brown bears. *Conservation Biology*, **12**, 418–429.

Pamilo P (1987) Heterozygosity in apomictic organisms. *Hereditas*, **107**, 95–101.

Picard D, Plantard O (2006) What constitutes a population for the plant parasitic nematode *Globodera pallida* in its native area (Peru)? *International Journal for Parasitology*, **36**, 115–122.

Picard D, Plantard O, Scurrah M, Mugniery D (2004) Inbreeding and population structure of the potato cyst nematode (*Globodera pallida*) in its native area (Peru). *Molecular Ecology*, **13**, 2899–2908.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, **95**, 536–539.

Plantard O, Porte C (2004) Population genetic structure of the sugar beet cyst nematode *Heterodera schachtii*: a gonochoristic and amphimictic species with highly inbred but weakly differentiated populations. *Molecular Ecology*, **13**, 33–41.

Pritchard JK, Stephens M, Donnely P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Riaz S, Vezzulli S, Harbertson ES, Walker MA (2007) Use of molecular markers to correct grape breeding errors and determine the identity of novel sources of resistance to *Xiphinema index* and Pierce's disease. *American Journal of Enology and Viticulture*, **58**, 494–498.

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Rosenberg MS (2002) *PASSAGE. Pattern Analysis, Spatial Statistics, and Geographic Exegesis*. Department of Biology, Arizona State University, Tempe AZ.

Sasser JN, Carter CC (1985) Overview of the international *Meloidogyne* project 1975–1984. An advanced treatise on Meloidogyne. *Biology and Control*, **1**, 19–24.

Schurko AM, Neiman M Logsdon JM. (2008) Signs of sex: what we know and how we know it. *Trends in Ecology and Evolution*, **24**, 208–217.

Sivasundar A, Hey J (2003) Population genetics of *Caenorhabditis elegans*: the paradox of low polymorphism in a widespread species. *Genetics*, **163**, 147–157.

Sivasundar A, Hey J (2005) Sampling from natural populations with RNAi reveals high outcrossing and population structure in *Caenorhabditis elegans*. *Current Biology*, **15**, 1598–1602.

Stoeckel S, Grange J, Fernandez-Manjarres JF, Bilger I, Frascaria-Lacoste N, Mariette S (2006) Heterozygote excess in a self-incompatible and partially clonal forest tree species - *Prunus avium* L. *Molecular Ecology*, **15**, 2109–2118.

Taylor CE, Brown DJF (1997) Nematode vectors of plant viruses. In: *Nematode Vectors of Plant Viruses*. CAB International, Wallingford, UK.

Taylor JW, Geiser DM, et al. (1999) The evolutionary biology and population genetics underlying fungal strain typing. *Clinical Microbiology Reviews*, **12**, 126–146.

Tibayrenc M (1999) Toward an integrated genetic epidemiology of parasitic protozoa and other pathogens. *Annual Review of Genetics*, **33**, 449–477.

Tibayrenc M, Ayala FJ (2002) The clonal theory of parasitic protozoa: 12 years on. *Trends in Parasitology*, **18**, 405–410.

Triantaphyllou AC (1985) Cytogenetics, cytotaxonomy and phylogeny of root-knot nematodes. *An advanced treatise on Meloidogyne. Biology and Control*, **1**, 113–126.

Turner SJ, Fleming CC (2002) Multiple selection of potato cyst nematode *Globodera pallida* virulence on a range of potato species. I. Serial selection on Solanum-hybrids. *European Journal of Plant Pathology*, **108**, 461–467.

Villate L, Esmenjaud D, Coedel S, Plantard O (2009) Development of nine polymorphic microsatellite markers for the phytoparasitic nematode *Xiphinema index*, the vector of the grapevine fanleaf virus. *Molecular Ecology Resources*, **9**, 229–232.

Villate L, Fievet V, Hanse B, Delemarre F, Plantard O, Esmenjaud D, van Helden M (2008) Spatial distribution of the dagger nematode *Xiphinema index* and its associated Grapevine fanleaf virus in French vineyard. *Phytopathology*, **98**, 942–948.

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, **10**, 249–256.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population-structure. *Evolution*, **38**, 1358–1370.

Welch DM, Meselson M (2000) Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science*, **288**, 1211–1215.

Xu K, Riaz S, Roncoroni NC, Jin Y, Hu R, Zhou R, Walker MA (2008) Genetic and QTL analysis of resistance to *Xiphinema index* in a grapevine cross. *Theoretical and Applied Genetics*, **116**, 305–311.

Yang RC (1998) Estimating hierarchical F-statistics. *Evolution*, **52**, 950–956.

This study is part of the PhD thesis of L. Villate, which focused on population genetics of the nematode *Xiphinema index*. L. Villate is working now as post-doc in population genetics of free living nematodes at the Department of Evolutionary Biology of the Max Planck Institute of Tübingen, Germany. This work was supervised by D. Esmenjaud, O. Plantard and M. van Helden. D. Esmenjaud is a nematologist studying various aspects of the biology of plant-parasitic nematodes associated with fruit trees and grape. O. Plantard is a population geneticist working on dispersal and reproductive mode of plant parasitic nematodes. M. van Helden is interested in sustainable protection of vine crop. S. Stoeckel is a population geneticist interested in asexual reproduction and selection.

Supporting information

Additional supporting information may be found in the online version of this article.

Supporting information 1 Mean allelic richness of each population of *X. index* estimated by the rarefaction index (ElMousadik & Petit 1996).

Supporting information 2 Graphical detection of the true number of clusters K with the maximum Evanno's criterion (ΔK) (Evanno *et al.* 2005).

Supporting information 3 Proportion of individual assigned to the 3 genetic clusters obtained by Bayesian analyses in each population.

Supporting information 4 Estimated structure of the *X. index* populations for $K = 3$ with InStruct program.

Supporting information 5 Neighbour-Joining trees on the distance DAS between multilocus genotypes (MLGs) sampled within MA and MB grapevine fields and simulated MLGs.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.