Dispersal constraints and fine-scale spatial genetic structure in two earthworm species

LISE DUPONT1*, YSOLINE GRÉSILLE1, BENOÎT RICHARD2,3, THIBAUD DECAËNS2,4 and JÉRÔME MATHIEU5

1Institut d'écologie et des sciences de l'environnement de Paris, Université Paris Est Créteil, 61 avenue du Général de Gaulle, 94010 Créteil Cedex, France
2EA 1293 ECODIV, SFR SCALE, UFR Sciences et Techniques, Université de Normandie, 76821 Mont Saint Aignan Cedex, France
3Laboratoire BioSol, Unité Agri'Terr, Esitpa – Ecole d’Ingénieurs en Agriculture, 3 rue du Tronquet, F-76134 Mont-Saint-Aignan Cedex, France
4Centre d’Ecologie Fonctionnelle et Evolutive, Université Montpellier 2, UMR 5175 CNRS, 1919 Route de Mende, 34293 Montpellier Cedex 5, France
5Institut d'écologie et des sciences de l'environnement de Paris, Université Pierre et Marie Curie, 7 quai Saint Bernard, 75252 Paris Cedex 05, France

Received 9 July 2014; revised 9 September 2014; accepted for publication 9 September 2014

The limited dispersal ability of earthworms is expected to result in marked genetic isolation by distance and remarkable spatial patterns of genetic variation. To test this hypothesis, we investigated, using microsatellite loci, the spatial genetic structure of two earthworm species, Allolobophora chlorotica and Aporrectodea icterica, in two plots of less than 1 ha where a total of 282 individuals were collected. We used spatial autocorrelation statistics, partial Mantel tests of isolation-by-distance (IBD) and isolation-by-resistance (IBR), and Bayesian test of clustering to explore recent patterns involved in the observed genetic structure. For A. icterica, a low signal of genetic structure was detected, which may be explained by an important dispersal capacity and/or by the low polymorphism of the microsatellite loci. For A. chlorotica, a weak, but significant, pattern of IBD associated with positive autocorrelation was observed in one of the plots. In the other plot, which had been recently ploughed, two genetically differentiated clusters were identified. These results suggest a spatial neighbourhood structure in A. chlorotica, with neighbour individuals that tend to be more genetically similar to one another, and also highlight that habitat perturbation as a result of human activities may deeply alter the genetic structure of earthworm species, even at a very small scale. © 2015 The Linnean Society of London, Biological Journal of the Linnean Society, 2015, 114, 335–347.


INTRODUCTION

Dispersal, which is the main mechanism leading to gene flow within and between populations, directly influences the level of genetic diversity maintained in populations (Clobert, 2001) and the ability of species to expand their range (Holt, 2003). The dispersal capacity of species is thus a fundamental life-history trait that plays a central role in the evolution of populations and their spatio-temporal dynamics. Limited dispersal ability can lead to mating among related individuals and should thus result in marked genetic isolation by distance and remarkable spatial patterns of genetic variation (Arnaud et al., 2001).

Soil invertebrates, such as earthworms, which are known to contribute to the maintenance of soil structure, the regulation of soil organic-matter dynamics,
and the stimulation of soil fertility and plant growth (Edwards, 2004), are believed to have restricted dispersal abilities (Costa et al., 2013) and to have developed original dispersal strategies as a result of the solidity, opacity, and high spatio-temporal heterogeneity of the soil environment (e.g. Mathieu et al., 2010). Determining the scale at which gene flow occurs in the field could indicate the approximate scale of demographic independence of earthworm populations and may inform on connectivity of vulnerable populations (Wilson et al., 2011). In a context of agricultural intensification (e.g. fertilization, pesticide application, and tillage), which results in a decline of soil biodiversity (Liiri et al., 2012), understanding dispersal patterns is crucial for the management of earthworm populations.

Some observations of introduction into earthworm-free habitats suggest that earthworms are able to colonize new areas at distances ranging from 4 to 14 m year\(^{-1}\) (review in Mathieu et al., 2010; Eijsackers, 2011). Mesocosm experiments have also indicated that some species can travel longer distances (26–500 m year\(^{-1}\)) under conditions that trigger dispersal (Mathieu et al., 2010; Caro et al., 2013). Moreover it has been suggested that passive dispersal, either by anthropogenic or natural processes (i.e. surface run-off or phoretic factors), should result in long-distance dispersal (Marinissen & Vandenbosch, 1992, Eijsackers, 2011). Earthworm dispersal has rarely been indirectly estimated using genetic data, and scarce data are available on the level of gene flow within and between earthworm populations. The studies carried out to date have focused on standard techniques, such as estimation of \(F_{st}\) (i.e. proportion of variance in allele frequencies that is among populations) and related statistics to measure gene flow. Most authors who have tested the relationship between genetic and geographical distances between earthworm populations using these population-based measures, whereby species are arbitrarily divided into partially isolated populations, have found significant genetic differentiation and no or a weak isolation-by-distance (IBD) pattern (Enckell et al., 1986; Kautenburger, 2006; Cameron, Bayne & Coltman, 2008; Prasankok et al., 2013; Torres-Leguizamon et al., 2014, but see Novo, Almodóvar & Díaz-Cosín, 2009; Novo et al., 2010a).

All these studies were carried out at the regional/landscape scale, except that of Novo et al. (2010a), which investigated the genetic structure of a homogastrid earthworm at fine spatial scale using \(F\)-statistics (0.064 km\(^2\)). Such unpredictable genetic-differentiation patterns were interpreted as evidence of the key role of anthropogenic or animal-mediated transport for earthworms’ dispersal (Costa et al., 2013). Whilst the IBD model is based on the assumptions that populations are large, equal in size, and stable over time, computer simulations have shown that unpredictable patterns of IBD are obtained when allowing for different population sizes and random fluctuations of population size and when adding environmental noise (Bjorklund et al., 2010). Interpreting an IBD pattern, or a lack thereof, is thus a hard task when studying populations probably subjected to perturbation and fragmentation, such as earthworm populations in agricultural areas. In contrast to population-based approaches, individual-based methods, such as spatial autocorrelation analysis (Hardy & Vekemans, 1999), directly analyse the genotypes of individuals across space, and thus estimates of population structure are not affected by a priori delimitation of populations. These estimates may then be used to infer the biological processes leading to clustering of genotypes (e.g. Carriconde et al., 2008).

Here, we aimed at obtaining a better understanding of gene flow in earthworms at fine spatial scale (within plots measuring less than 1 ha) by analysing patterns of genetic structure using individual-based approaches in addition to traditional population-based measures. Moreover, we tested whether gene flow between individuals was constrained or facilitated by soil properties [isolation by resistance (IBR)], which were described using soil resistivity as a single synthetic variable. Resistivity represents the capacity of the soil to resist the flow of electricity, which is tightly linked to soil physical and chemical properties, such as texture, hydrological properties, or nitrogen content, and can be measured almost continuously in space, offering a much better picture of spatial variations of soil properties than do traditional soil analyses. In several studies, the abundance of endogeic and anecic earthworms was highly related to soil resistivity (Valckx et al., 2009; Joschko et al., 2010; Lardo et al., 2012). However, the correlation was positive or negative depending on the species (Valckx et al., 2009).

In this study, we focused on two endogeic (i.e. species living in the upper organo-mineral soil layers and forming horizontal nonpermanent burrows, Bouché, 1977) earthworm species commonly found in European agricultural soils, the green morph of Allolobophora chlorotica (Savigny, 1826) and Aporrectodea icterica (Savigny, 1826). Allolobophora chlorotica is known to be located in the upper 60-mm soil layer (Sims & Gerard, 1999), is theoretically able to travel more than 167 m year\(^{-1}\) in constant suitable conditions, and is not subject to density-dependent dispersal (Caro et al., 2013). Aporrectodea icterica is found deeper in the soil and is considered to be more mobile, being able to travel up to 500 m year\(^{-1}\) under constant artificial conditions...
and to respond to density (Mathieu et al., 2010; Caro et al., 2013).

By comparing these two earthworm species belonging to the same eco-morphological group but with contrasted dispersal capabilities, we tested the hypothesis that restricted-disperser species should present a higher degree of spatial organization than high-disperser species. Specifically, the objectives of this study were: (1) to estimate the total genetic diversity of these earthworm species in plots of less than 1 ha; (2) to determine the spatial genetic structure and scale of IBD and IBR in these plots; and (3) to estimate gene dispersal within plots if IBD is found.

MATERIAL AND METHODS

Biological Models

Aporrectodea ictera is an abundant diploid and obligatory biparental earthworm species (Casellato, 1987) that is commonly found in agricultural soils (Capowiez et al., 2005). Its taxonomic status is firmly grounded and the species has distinct morphology, making it easy to recognize (Torres-Leguizamon et al., 2012). Conversely, the A. chlorotica aggregate (Dupont et al., 2011) is composed of several sister species. A green colour morph of this aggregate represents a single taxon although composed of two mitochondrial lineages, whilst the taxonomic status of a pink morph (at least five mitochondrial lineages) remains unclear (King, Tibble & Symondson, 2008; Dupont et al., 2011). Although hybridization seems possible between morphs (Dupont et al., 2011), introgression is probably restricted as a result of postzygotic reproductive isolation. Cross experiments indeed revealed: (1) a severely restricted viability of cocoons produced by the green morph in pink–green pairings; and (2) male sterility of the surviving hybrids (Lowe & Butt, 2008). Here, we restricted our study to the species represented by the green morph. This diploid and amphimictic species is common in temperate grassland (Lowe & Butt, 2007).

Study Site and Soil Properties

The sampling was carried out at the ‘Lycée Agricole d’Yvetot’ (Seine Maritime, France), located 200 km north-west of Paris, during March and April 2009. We selected two pastures, ~500 m apart, located in the same topographic situation but with contrasting ages: a 5-year-old pasture (P₅) and a pasture of more than 42 years of age (P₉). In each pasture, sampling was carried out on a 10-m mesh grid of 120 m × 70 m; 104 points per plot were sampled for P₅ but, for logistical reasons, the number of points was reduced to 68 in P₉ (latitude/longitude range in Lambert II étendu: 484496–484634/2513651–2513774 and 484058–484197/2513760–2513882 in P₅ and P₉, respectively).

Soil resistivity was measured in partnership with Geocarta (Geocarta SA, Paris). The automatic resistivity profiling (ARP) (Papadopoulos et al., 2009) technique was chosen for its high accuracy and its reduced sensitivity to superficial geophysical noise. Moreover, this technique is non-invasive and allows soil resistivity to be measured simultaneously at three depths (here: 0–0.5 m, 0–1 m, and 0–1.7 m). Measurements were performed with a mobile device equipped with a differential Global Positioning System (GPS) to retrieve the geographical coordinates of samples with an accuracy of 20 cm. Measurements were made every 50 cm along lines spaced 2 m apart. This sampling scheme allowed for the description of soil resistivity with a very high spatial resolution (Fig. 1).

Earthworm Sampling and DNA Extraction

At each point of the grid, earthworms were sampled using a combination of formaldehyde extraction and hand-sorting. First, 10 l of 4‰ formaldehyde were applied onto a 1 m² surface, and earthworms expelled at the soil surface were collected during a 15-min period. Then, a soil volume of 25 cm × 25 cm × 25 cm and 30-cm depth was dug out in the centre of the square meter and hand sorted in the field. More details on the different species in the plots, sampling methodology, and species morphological identification are given in Richard et al. (2012). Specimens were fixed in pure alcohol until DNA extraction. A fragment of tegument was dissected and total genomic DNA was extracted using the DNeasy 96 Blood & Tissue Kit (Qiagen).

Molecular Identification of A. chlorotica Mitochondrial Lineage

To determine the mitochondrial lineage to which each A. chlorotica individual belonged, we targeted the barcode portion of the cytochrome c oxidase I gene (COI). Some of the sequences have already been published in Dupont et al. (2011) (GenBank accession numbers: HM879975; HM417934–35, 37–41, 43, 45, 47–49, 52, 54, 55; HQ682441–46). For the other samples, a fragment of the COI gene was amplified according to Folmer et al. (1994). After purification using Microclean (Microzone Ltd), sequence reactions were performed using the BigDye Terminator Cycle Sequencing kit V. 1.1 (Applied Biosystems) and sequence data were obtained using a 3130xl Genetic Analyser (Applied Biosystems). Sequences were manually aligned using the BioEdit program (Hall,
The GenBank accession numbers are KC569604–740.

In order to establish the correspondence of the generated sequences to the five lineages of the pink morph and the two lineages of the green morph of *A. chlorotica* uncovered in the studies by King et al. (2008) and Dupont et al. (2011), a phylogenetic analysis was performed using MEGA 6.0 software (Tamura et al., 2013). The sequences were aligned with all the haplotypes of King et al. (2008) and Dupont et al. (2011), and a phylogenetic tree was constructed by selecting the best-fit maximum-likelihood model in MEGA 6.0, based on the lowest Bayesian information criterion (BIC) score. The evolutionary distances were computed using the HKY + G + I as best-fit model with 1000 bootstrap replicate values. In addition, a local BLAST was implemented in the software BioEdit (Hall, 1999) using a nucleotide database formed by all the haplotypes of King et al. (2008) and Dupont et al. (2011). The match with the highest E-value was used to assign a lineage to the query sequence. Individuals that were assigned to the lineages of the pink morph were excluded from the analysis.

**MICROSATELLITE GENOTYPING**

*Allolobophora chlorotica* individuals were genotyped at eight microsatellite loci, as defined by Dupont et al. (2011). *Aporrectodea icterica* individuals were genotyped at the seven microsatellite loci described in Torres-Leguizamon et al. (2012) and one (Ai51) newly identified locus (M. Torres-Leguizamon and L. Dupont, unpubl. data; forward and reverse primer sequences: Ai51F 5’ NED-AATCAACTGAACAGGCGTCC and Ai51R TTCGACAGAATGATTGTCCG). Loci [including Ai51 (annealing temperature 51 °C and 2.5 mM of MgCl₂)] were amplified by the polymerase chain reaction (PCR) following protocols detailed in Dupont et al. (2011) and Torres-Leguizamon et al. (2012). The migration of PCR products was carried out on a 3130xl Genetic Analyser using the LIZ500 size standard, and alleles were scored using GENESCAN V3.7 and GENOTYPER V3.7 software (Applied Biosystems).

**POPULATION GENETIC ANALYSES**

In each pasture plot, the genetic diversity of the two species was analysed by computing allele frequencies, number of alleles (*Nₘₐ*), and expected heterozygosity.
(\(H_e\)) using GENETIX V 4.05 (Belkhir et al., 2004). To take into account variation in sample size, allelic richness (\(A_r\); El Mousadik & Petit, 1996) was estimated using FSTAT v.2.9.3 (Goudet, 2000). The null hypothesis of independence between loci was tested from statistical genotypic disequilibrium analysis using GENEPOP V. 4.1.3 (Rousset, 2008). Evidence of null alleles was examined using the software MICRO-CHECKER (Van Oosterhout et al., 2004). Departure from Hardy–Weinberg expectations within plots were quantified by calculating the Weir & Cockerham’s (1984) estimator of the fixation index, \(F_{st}\), and conformity to Hardy–Weinberg equilibrium (HWE) was assessed with exact tests implemented in GENEPOP V. 4.1.3. To adjust for multiple comparisons, the false discovery rate (FDR) method (Benjamini & Hochberg, 1995), as implemented in the software SGoF (http://webs.ubigo.es/acraaj/SGoF.htm), was applied.

In order to investigate the genetic structure among populations, exact \(G\) tests of allelic differentiation were carried out between plots using GENEPOP V. 4.1.3. We also used a traditional population-differentiation approach based on \(\Phi\) analysis. Weir & Cockerham’s (1984) estimator of \(\Phi\) was calculated using GENEPOP V. 4.1.3.

**Spatial autocorrelation and gene dispersal**

To investigate spatial genetic structure at the individual level, we conducted spatial autocorrelation analysis (Hardy & Vekemans, 1999), which provides a measure of genetic correlation as a function of Euclidean distances. Genetic distances between individuals were estimated using the Rousset’s \(\hat{d}\) estimator (Rousset, 2000) using SPAGEDI V1.3. To visualize the spatial genetic structure, we averaged pairwise genetic distance over a set of distance classes, and plotted it against the distance. Between 10 and 20 distance classes were used to equalize the number of comparisons among each distance class (approximately 200). Permutations (10 000) provided 95% confidence intervals about the null hypothesis of no spatial genetic structure.

If limited dispersal causes a decrease in genetic similarity between individuals at increasing geographical distances, Wright’s IBD model and genetic neighbourhood (Wright, 1943; Wright, 1946) can be used to estimate the dispersal distance. Wright defines a neighbourhood size (\(N_n\)) as \(4d,\pi,\sigma^2\), where \(\sigma^2\) is the effective density and \(\sigma^2\) is the mean squared axial parent–offspring dispersal distance (Wright, 1946). Assuming IBD within populations, \(\sigma^2\) is estimated as \(\sigma^2 = 1/(b,4d,\pi,\alpha)\), where \(b,\alpha\) is the slope of the regression of genetic distance (\(\hat{d}\)) on the logarithm of geographical distance (Rousset, 2000). We estimated gene dispersal using the iterative procedure provided by SPAGeDi v. 1.3 (Hardy & Vekemans, 2002) but as it did not always converge, as it occurs in some cases according to the SPAGeDi manual, we simply used the relation \(\sigma^2 = 1/(b,4d,\pi,\alpha)\).

Because the effective density of adult earthworms (\(d_e\)) is unknown, we used the total earthworm density \(d, d/2, d/5, d/10\) as alternative estimates of \(d_e\). Thus, an upper and lower range of gene dispersal was obtained. We used earthworm densities obtained in Richard et al. (2012) for \(P_A\) (\(d = 22.80\) and \(d = 19.94\) ind m\(^{-2}\) for *A. chlorotica* and *A. icterica*, respectively) and unpublished data of Richard et al. for \(P_B\) (\(d = 8.18\) and \(d = 36.27\) ind m\(^{-2}\) for *A. chlorotica* and *A. icterica*, respectively).

**Isolation by distance and isolation by resistance**

In order to test the hypothesis of IBD and IBR, we used a traditional partial Mantel tests approach on the different explanatory and response distance matrices. The distance matrix of the response corresponded to the Rousset’s \(\hat{d}\) genetic distance between individuals. The distance matrix for IBD was computed as the geographical (Euclidean) distance between each pair of individuals. The distance matrix for IBR was computed as the cumulated cost of the least-cost path between each pair of individuals. Least-cost paths were produced with the Landscape Genetics Toolbox for Arcgis (Perry et al., 2010), which determines movement costs between pairs of points, based on a resistance matrix, represented here by the interpolated soil resistivity (kriged with a spherical model). We considered three scenarios to calculate costs of movements based on soil resistivity. In scenario 1, movement costs increased linearly with soil resistivity; in this case, movement costs at location \(xy\) corresponded to soil-resistivity values at \(xy\): \(\text{cost}_{xy} = R_{xy}\). In scenario 2, movement costs were lowest at intermediate values of soil resistivity; in this case, costs of movements were calculated as the absolute deviation from the average soil resistivity of the plot: \(\text{cost}_{xy} = |R_{xy} - R^*|\), with \(R^*\) = average soil resistivity of the plot. In scenario 3, costs decreased with soil resistivity; in this case, \(\text{cost}_{xy} = 1/R_{xy}\). Once movement costs were calculated at each location according to each scenario, we computed the least-cost path between each pair of points for each scenario. Then, we used partial Mantel tests to assess the likelihood of IBD and IBR following each scenario. These tests can evaluate the relationship between a response variable and an explanatory variable whilst controlling for the effect of a second explanatory variable, when these variables represent matrices of dissimilarity between pairs of locations (Legendre & Legendre, 1998). Tests were performed by permuta-
tion of the data. For instance, we estimated the effect of the geographical distance (D) while controlling the effect of movement costs (scenario 1) and denoted it D/R. We took advantage of this method in order to assess the likelihood of IBD and IBR in our data set, following each IBR scenario.

**Bayesian clustering**

Each study plot could enclose several subpopulations of each species. We investigated the occurrence of such cryptic population structure within plots using the Bayesian model implemented in GENELAND v.4.0.3 (Guillot, Mortier & Estoup, 2005) that simultaneously analyses spatial and genetic information. This model is based on HWE and linkage equilibrium (LE). We proceeded in two steps: a first run to infer the number of genetically distinct clusters (subpopulations) at each locality, K (i.e. true subpopulation number), and a second run with K fixed at the modal value from the first step to estimate the assignment of individuals to the inferred subpopulations. The first step was replicated five times to check for convergence, allowing K to vary from one to five clusters and using $10^6$ Markov Chain Monte Carlo (MCMC) iterations. We used the correlated frequency model, which is predicted to be more powerful at detecting subtle differentiation. Moreover, the putative presence of null allele(s) was taken into account in the model and the spatial coordinates were treated as uncertain in order to allow samples with the same coordinates to be assigned to different subpopulations. In the second step undertaken when K > 1, we ran the MCMC five times again with K fixed, $10^6$ MCMC iterations, and the other parameters unchanged. The runs were then post-processed in order to obtain posterior probabilities of subpopulation membership for each individual.

Because the correlated frequency model is prone to algorithm instabilities and particularly sensitive to departure from model assumptions (e.g. the presence of IBD), the Geneland manual recommends checking ex-post that the inferred groups are significantly differentiated. Thus, genetic differentiation between inferred clusters was checked using the exact G-test available in GENEPOP V4.1.3 (Rousset, 2008).

**RESULTS**

**Genetic diversity**

After excluding 11 individuals belonging to the pink morph and two individuals with ambiguous genotypes, the genotypes of 141 A. chlorotica individuals were analysed (95 and 46 individuals in P_A and P_B, respectively, Fig. 1). The same number of genotypes (141) was analysed in A. icterica (61 and 80 individuals in P_A and P_B, respectively, Fig. 1). All loci were polymorphic in both species, except for locus 2PE40, which was monomorphic in A. icterica in the P_A plot (AiP_A) (Table 1). Higher values of genetic diversity were obtained for A. chlorotica than for A. icterica (Table 1). Values of genetic-diversity indices within species were similar in both plots (Table 1). For example, identical values of $H_e = 0.530$ were obtained in both plots in A. icterica. In A. chlorotica, $H_e$ was in the same range with estimates of 0.752 and 0.791 in P_A and P_B, respectively.

**Hardy–Weinberg equilibrium and linkage equilibrium**

In A. icterica, all loci were unlinked. In contrast, in A. chlorotica one pair of loci (Ac476–Ac419) was out of LE in the P_A plot (AcP_A) and seven pairs departed significantly from LE in the P_B plot (Ac127–Ac418, Ac127–Ac476, Ac127–Ac529, Ac418–Ac170, Ac418–Ac419, Ac418–Ac476, and Ac418–Ac528). In their study of the genetic structure of the A. chlorotica aggregate in Europe, Dupont et al. (2011) showed no linkage among these loci, except for Ac127–Ac476 in one of the populations. Thus, these departures from LE are probably explained by genetic sub-structure within the plots (see Ohta, 1982) rather than by physical linkage between loci. Nevertheless, the locus Ac476 (less informative than Ac419) was excluded from the AcP_A data set, and Ac418 and Ac127 were excluded from the AcP_B data set for the clustering analyses that required LE.

Heterozygote deficiency, indicated by significant deviation from HWE, was observed for several A. chlorotica loci (Table 1). The existence of null alleles was suggested by Microchecker results for some of the A. chlorotica loci (Table 1), but they were different in the two plots. Estimated frequencies of null alleles for these loci were relatively low, ranging from 6.1% to 11.2%. In A. icterica, deviations from HWE were revealed for loci 2PE70 and C4 in both populations (Table 1). Although no null allele was detected in P_B for the 2PE70 locus, the estimated frequency of null alleles was particularly high for the locus C4 in both plots (32.7–36.9%). Thus, this locus was excluded from all other analyses.

**Gene flow within plots**

Significant genetic structure was revealed at the population level, with $F_{st}$ values (0.018 and 0.014 in A. icterica and A. chlorotica, respectively) associated with significant exact tests ($P < 0.001$). The spatial genetic structure at the individual level was investigated through IBD and autocorrelation analyses.
A weak, but significant, relationship between genetic distances and the logarithm of geographical distances was obtained for *A. chlorotica* in PB (slope of the regression = 0.0068, *P* < 0.001). This IBD (Table 2) was confirmed by the autocorrelation analysis (Fig. 2). Considering d as the upper limit of effective density and d/10 as the lower limit of effective density, the lowest estimate of gene dispersal (i.e. parent–offspring dispersal) distance was therefore 3.41 m and 3.78 m in AcPB. Considering two generations of *A. chlorotica* during 1 year in Normandy, an approximate dispersal rate ranging from 6.81 to 7.56 m for 1 year may be estimated for this species. An IBD was also suggested in AcPA when the geographical distances were partialed out by soil resistivity (D/R, *P* < 0.001, Table 2). Whereas no IBD was detected in *A. icterica* populations using all genotypes (Table 2), correlograms suggested a positive relationship between genetic and geographical distance in the restricted range of 20–61 m (Fig. 2). However, the gene-dispersal distance could not be estimated for this species because of the low level of polymorphism (*H* ranged from 0.000 to 0.716). Indeed, Leblois, Estoup & Rousset (2003) recommended using loci with *H* of around 0.7 to maximize the efficiency of the estimation of *σ*².

Results of partial Mantel tests suggest that soil properties might play a role in the structuring of genetic variation at the scale of the plot. Indeed, IBD was shown in AcPB, AcPA, and AiPB (Table 2).

### Genetic sub-structure within plots

The occurrence of cryptic population structure was suggested in both plots for *A. chlorotica* by the Geneland Bayesian analysis that identified two groups of individuals in AcPA (Fig. 2) and five groups in AcPB. The clusters AcPₐ–C₁ (cluster 1 of the *A. chlorotica* Pₐ plot) and AcPₐ–C₂ (cluster 2 of the *A. chlorotica* Pₐ plot) were composed of 42 and 53 individuals, respectively, and were significantly differentiated (exact G-test, *P* < 0.001). The genetic differentiation between AcPₐ–C₁ and AcPₐ–C₂ clusters (*F*ₚ = 0.016) was higher than between AcPₐ–C₁ and the PB plot (*F*ₚ = 0.014, *P* < 0.001) but lower than between AcPₐ–C₂ and PB (*F*ₚ = 0.022, *P* < 0.001).

The five clusters identified in PB were composed of 2, 12, 13, 9, and 10 individuals. These small population sizes prevented robust testing of genetic differentiation. Moreover, the Bayesian spatial correlated model used in Geneland is particularly sensitive to the presence of IBD (see the Geneland manual, Guillot *et al*., 2005; Frantz *et al*., 2009). Hence, the five detected clusters are probably misleading results.
DISCUSSION

The polymorphism of microsatellite DNA sequences varied across loci and across species. In particular, microsatellite markers used for genotyping *A. chlorotica* individuals were more polymorphic than the markers used for *A. icterica*. For instance, the number of alleles was three times higher for *A. chlorotica* than for *A. icterica*. This difference is not a particularity of the study because the level of genetic diversity recorded is similar to the results of Dupont et al. (2011) for *A. chlorotica* and of Torres-Leguizamon et al. (2012) and Torres-Leguizamon et al. (2014) for *A. icterica* at the landscape/region scale. In our study, *A. icterica*...
microsatellites displayed lower polymorphism than all other microsatellites obtained for other earthworm species using similar isolation strategies. For instance, the $N_{\text{all}}$ value of earthworm’s microsatellite loci typically ranges from five to 17 (review in Torres-Leguizamon et al., 2014), whereas A. icterica loci display a maximal $N_{\text{all}}$ of 3.6. Such a low polymorphism of molecular markers decreases the precision in estimates of heterozygote deficiencies (Robertson & Hill, 1984), statistical tests of differentiation (Goudet et al., 1996), and estimates of gene-dispersal distances (Leblois et al., 2003). The low polymorphism of A. icterica microsatellite loci could be reflecting bottlenecks in the evolutionary history of this species (Torres-Leguizamon et al., 2014). More research is needed in order to understand, more clearly, this particularity of A. icterica simple sequence repeats.

Most of the loci were at LE, except for one pair of loci in the $P_A$ plot and several pair of loci in the $P_B$ plot for A. chlorotica. In a single random-mating population, linkage disequilibrium or nonrandom association of alleles between two loci may be produced by epistatic interaction in fitness between the loci concerned (Kimura, 1956) and random genetic drift as a result of to finite population size (Ohta & Kimura, 1969). Because the markers used are supposedly neutral, the hypothesis of epistatic interaction may be ruled out. Random fluctuation of gamete frequencies as a result of genetic drift is enhanced if the population is subdivided or if mating were not random in the population (review in Ohta, 1982).

Thus, linkage disequilibrium within A. chlorotica samples could be explained either by sub-structure within the population (i.e. the Wahlund effect) or by mating among relatives (i.e. inbreeding).

Such phenomena (the Wahlund effect and/or inbreeding) should have resulted in deviation from HWE. Significant deficits of heterozygotes were indeed observed in both plots for both species. In A. icterica, deviations from HWE were largely caused by null alleles at the C4 locus. Null allele existence was also suggested at several A. chlorotica loci, but at a low frequency. Our results showed that sub-structure within plots and/or mating among relatives also contribute to the deviation from HWE in this species.

Sub-structure was suggested by the Bayesian analysis of genetic clustering which revealed two clusters within the $P_A$ plot for A. chlorotica. It is noteworthy that the upper half of $P_A$ was ploughed 1 year before sampling, whereas the lower half was not. It is thus proposed that habitat perturbation because of human activities might be responsible for the A. chlorotica spatial genetic clustering at this fine geographical scale. Ploughing might thus alter the genetic structure of earthworm populations for at least 1 year. This perturbation could probably be even longer for A. chlorotica because of the limited dispersal capacity of the species. This result suggests that natural and artificial (i.e. caused by human activity) habitat spatial heterogeneity can be an important contributor to earthworm population genetic structure.

Figure 2. Correlograms (solid lines) of estimated $\hat{a}$ genetic distance (Rousset, 2000) for both plots and both species. Dotted lines indicate the 95% null hypothesis confidence region. Significant values: *$P < 0.05$. 

© 2015 The Linnean Society of London, Biological Journal of the Linnean Society, 2015, 114, 335–347
Mating among relatives seems likely for *A. chlorotica* in at least one of the plot. In Pb, the spatial autocorrelation analysis revealed a pattern of fine-scale genetic structure with restricted gene dispersal for this species. The estimation of gene-dispersal distance during one generation (approximately 6 months) ranged from 3.41 to 3.78 m. This estimation represents a dispersal ranging from 6.82 to 7.56 m for 1 year. This is close to the estimation of 4 m year⁻¹ previously obtained for the annual dispersal rate of *A. chlorotica* during colonization of worm-free clayey polder soils in the Netherlands (review in Eijssackers, 2011). Such a restricted dispersal confirms that there is a higher probability that individuals mate with individuals born in close proximity to themselves than with individuals born far away, a pattern favouring inbreeding. A similar pattern was revealed by Novo *et al.* (2010a), who investigated the mating strategy of *Hormogaster elisae*, an outcrossing endogeic earthworm endemic to the Central Iberian Peninsula. Their results suggested that individuals of *H. elisae* were rather sedentary and did not relocate over long distances to find mating partners.

In contrast to *A. chlorotica*, neither IBD nor genetic clustering was detected within plots for *A. icterica*. This low signal of genetic structure may be explained by an important dispersal capacity of the species and/or by the low polymorphism of the microsatellite loci that could have prevented the detection of subtle genetic differentiation. A similar absence of relationship between geographical and genetic distance was shown at a regional scale in this species (Torres-Leguizamon *et al.*, 2014). In their review on earthworm genetic structure, Costa *et al.* (2013) asserted that most of the few studies on genetic structure of earthworms found no relationship between genetic and geographical distances. In contrast, Novo *et al.* found a pattern of isolation according to distance at both interpopulation (Novo *et al.*, 2009) and intrapopulation (Novo *et al.*, 2010a) levels in hormogastrid earthworms. Altogether, these studies were not achieved at the same geographical scale, did not use similar molecular markers (e.g. amplified fragment length polymorphisms (AFLPs), microsatellites, or the COI gene), and targeted species belonging to various eco-morphological groups (i.e. anecic and endogeic) and having various reproductive strategies (i.e. amphimixis and parthenogenesis); thus, it is difficult to draw general conclusions about earthworm spatial genetic variation.

Experimental work has shown that environmental properties are strong determinants of dispersal in earthworms and thus predicted that their spatial distribution should be correlated with environmental data (Mathieu *et al.*, 2010). To date, only a few studies have investigated the relationship between environmental data, such as soil characteristics, and the spatial genetic structure of earthworm populations (Lentzsch & Golldack, 2006; Novo *et al.*, 2010b). Here, despite numerous significant partial Mantel tests at the scale of the whole study, little consensus has emerged regarding the direction of the relationship between genetic distances, geographical distances, and soil resistivity because geographical distances and soil resistivity, when not partialed out by the other matrix, were not always positively related to genetic distances. The significant patterns of IBR were thus difficult to interpret, in particular for *A. icterica* in the plot Pb. In previous studies at a larger spatial scale, Lentzsch & Golldack (2006) found no relationship between the distribution of *A. caliginosa* genotypes and soil properties (e.g. pH, soil organic carbon, total nitrogen, and clay content) along a 151-m transect, and Novo *et al.* (2010b) found a weak relationship between soil texture (i.e. coarse sand and total loam content) and genetic distances between populations of hormogastrid earthworms at a similar scale (<100 km²). Altogether, these results reveal that there is no simple relationship between soil properties and earthworm genetic structure, and suggest that other factors, such as demographic events (i.e. population bottleneck events and genetic drift), may be particularly important in shaping the genetic composition of earthworm populations.

**CONCLUSION**

Only scarce data are available on the fine-scale population structure of soil invertebrates. To our knowledge, only two studies describe genetic patterns of soil macro and meso invertebrates using individual-based approaches: the study of Sullivan, Dreyer & Peterson (2009), showing that the collembolan *Folsomia candida* exhibits genetic population structuring over a very fine geographical scale (0.65 km²), and our study. Here, we showed that in a 42-year-old pasture without recent perturbation, *A. chlorotica* displayed a neighbourhood structure of randomly mating earthworms, in which neighbour individuals tend to be more genetically similar to one another, whereas no limit to gene flow was detected for *A. icterica*. In the other plot, where tillage had recently deteriorated the physical conditions of the soil, the expected pattern of IBD in *A. chlorotica* seemed to be erased, whilst a pattern of IBR was revealed for *A. icterica*. Thus, the present study emphasizes that agricultural practices contributing to a fragmentation of the species habitat may durably alter the population genetic structure of earthworms at a very small scale.

ACKNOWLEDGEMENTS

This work was funded by the French National Research Agency (ANR) as a part of the project EDISP no. ANR-08-JCJC-0023. We thank the Conseil Régional de Haute-Normandie for the PhD fellowship allowed to Benoit Richard (VATA network). We are grateful to Gagik Hovhannisian (IRD) and Henri Robain (IRD) for their help with soil resistivity measures and interpretation, to D. Guéroux (Lycée Agricole d’Yvetot) for providing field sampling facilities, to all the participants to EDISP project for their participation in field sampling, and to Naoise Nunan for English language editing. We also thank two anonymous reviewers for their helpful comments.

REFERENCES


Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F. 2004. GENETIX 4.05, logiciel sous Windows pour la génétique des populations. Montpellier: Laboratoire Génome, Population, Interactions, CNRS UMR 5000, Université Montpellier II.


© 2015 The Linnean Society of London, Biological Journal of the Linnean Society, 2015, 114, 335–347


Kautenburger R. 2006. Genetic structure among earthworms (Lumbricus terrestris L.) from different sampling sites in western Germany based on random amplified polymorphic DNA. Pedobiologia 50: 257–266.


Torres-Leguizamon M, Mathieu J, Decaëns T, Dupont L.
Genetic structure of earthworm populations at a regional scale: inferences from mitochondrial and microsatellite molecular markers in *Aporrectodea icterica* (Savigny 1826). *PLoS ONE* 9: e101597.


