Landscape features impact connectivity between soil populations: a comparative study of gene flow in earthworms

L. Dupont, M. Torres-Leguizamon, P. René-Corail, J. Mathieu

Landscape features impact connectivity between soil populations: a comparative study of gene flow in earthworms

L. Dupont[1*], M. Torres-Leguizamon[1], P. René-Corail[1] and J. Mathieu[2]

[1] Université Paris Est Créteil (UPEC), UPMC, Paris 7, CNRS, INRA, IRD, Institut d'écologie et des sciences de l'environnement de Paris, 94010 Créteil Cedex, France

[2] Sorbonne Universités, UPMC Univ Paris 06, UPEC, Paris 7, CNRS, INRA, IRD, Institut d'Ecologie et des Sciences de l'Environnement de Paris, 75005, Paris, France

* Corresponding author. Tel : +33(0)145171664 ; Fax : +33(0)145171999 ; e-mail address : lise.dupont@u-pec.fr

Key-Words : Dispersal, genetic diversity, genetic structure, landscape connectivity, local landscape structure, microsatellites

Running Title : Landscape genetics of earthworms

Abstract

Landscape features are known to alter the spatial genetic variation of above ground organisms. Here, we tested the hypothesis that the genetic structure of below ground organisms also responds to landscape structure. Microsatellite markers were used to carry out a landscape genetic study of two endogeic earthworm species, *Allolobophora chlorotica* (N = 440, 8 microsatellites) and *Aporrectodea icterica* (N = 519, 7 microsatellites), in an agricultural landscape in the North of France, where landscape features were characterised with high accuracy. We found that habitat fragmentation impacted genetic variation of earthworm populations at the local scale. A significant relationship was observed between genetic diversity (H_e, A_r) and several landscape features in *A. icterica* populations but not in the *A. chlorotica* ones. Moreover, a strong genetic differentiation between sites was observed in both species, with a low degree of genetic admixture and high F_{st} values. The landscape connectivity analysis (MLPE) at the regional scale, including Isolation By Distance (IBD), Least Cost Path (LCP) and Cost Weighted Distance (CWD) approaches, showed that genetic distances were linked to landscape connectivity in both species. This indicates that the fragmentation of natural habitats has shaped their dispersal patterns and local effective population sizes. Landscape connectivity analysis confirmed that *a priori* favourable habitats such as grasslands may constitute dispersal corridors for these species.

Introduction

A number of studies showed that spatial genetic variations of aboveground organisms respond to changes in their landscape, through mechanisms involving movements of organisms (review in Storfer *et al.* 2010; Manel & Holderegger 2013; Hall & Beissinger 2014). It is now well-grounded that landscape features alter aboveground organisms' genetic structure. Comparatively little is known about the impact of landscape-scale habitat heterogeneity on belowground organisms, such as soil invertebrates, whose mobility is more restricted (Vanbergen et al. 2007). Despite their importance in ecosystem functioning and in the delivery of many ecosystem services (Lavelle et al. 1997; Jouquet et al. 2006; Lavelle et al. 2006; Blouin et al. 2013), we still do not have a grasp of even basic information about population genetic structure of soil organisms. For instance, Costa et al. (2013) found only sixteen different species among the collembolans, earthworms and isopods groups of soil invertebrates for which a population genetics study was carried out. Some of these papers investigated the spatial genetic structure of soil organisms at a fine-scale (Sullivan et al. 2009; Novo et al. 2010; Dupont et al. 2015), but none addressed the effect of landscape features on genetic variation. However, terrestrial habitat heterogeneity is known to affect the diversity of soil species' assemblages by producing variation in the diversity of plant and litter (Vanbergen et al. 2007). It is therefore assumed that aboveground structure and diversity could profoundly impact the population genetic structure of belowground organisms.

The methodology of landscape genetics allows one to test the influence of the landscape and environmental characteristics on microevolutionary processes and metapopulation dynamics, including gene flow and local adaptation (Manel *et al.* 2003; Storfer *et al.* 2007;

Holderegger & Wagner 2008). Landscape connectivity is a twofold parameter made up of structural connectivity and functional connectivity. Structural connectivity refers to the physical relationship between landscape elements while functional connectivity may be defined as the ease with which a lanscape can be crossed by an organism (Taylor et al. 2006). Depending on the organisms, the permeability of the landscape will differ and some constituent elements of the landscape can facilitate dispersal (i.e. “corridors”) while others can impede or reduce the passage of dispersers (i.e. “barriers”) (Taylor et al. 1993). Landscape structure can also have an important effect on passive dispersers by altering the abiotic and biotic conditions that affect movement (Matthysen 2012). In order to understand how landscape characteristics influence functional connectivity, resistance surfaces are usually computed and translated into measures of inter-population connectivity principally using two kinds of models. Least-cost path models (Adriaensen et al. 2003) assume that movement or gene flow rates between each pair of sites is related to the total cumulative resistance or ‘cost’ (sum of per-pixel resistance values) along a single optimal path, while circuit-theory based models (McRae 2006) incorporate all possible pathways across landscapes, and their parameters and predictions can be expressed in terms of random walk probabilities (Cost Weighted Distance "CWD", or "resistance approach"). Spear *et al.* (2010) highlighted that both models provide complementary indices of connectivity with Least-cost-path distances being more informative at local scales and circuit theoretic models being particularly useful for incorporating effects of gene flow over multiple generations.

Here, we were interested in landscape features impacting genetic variation and functional connectivity in earthworms. Dispersal by passive mechanisms, such as zoochory, wind, water and human activities is believed to be implicated in their long-distance movement (Eijsackers

2011; Costa et al. 2013; Dupont et al. 2015), whilst their active dispersal is dependent on habitat quality, conspecific density, and habitat modification by conspecifics in endogeic (i.e. species living in the upper organo-mineral soil layers and forming horizontal non-permanent burrows, Bouché 1977) and anecic (i.e. species forming permanent or semi-permanent vertical burrows in the soil which open at the surface where the earthworm emerges to feed, Bouché 1977) species (Mathieu et al. 2010; Caro et al. 2012; Caro et al. 2013). The distribution of these restricted dispersers is known to be controlled by soil parameters at the field scale and by land use (forest, grassland and agricultural field), soil management, soil type and climatic conditions at larger scales; studies at the landscape scale are thus challenging since small-scale heterogeneity as well as gradients affecting large-scale patterns have to be accounted for (Palm et al. 2013).

In order to investigate whether landscape features impact the genetic structure of earthworm populations, we carried out a regional-scale comparative survey of genetic variation in two species commonly found in European agricultural landscapes, the green morph of *Allolobophora chlorotica* (Savigny, 1826) and *Aporrectodea icterica* (Savigny, 1826). Both species are endogeic but present several ecological differences. *A. chlorotica* typically lives between soil surface and the upper 60 mm soil layer (Sims & Gerard 1999) and is theoretically able to travel over 167 m.year^{-1} in constant suitable conditions (Caro et al. 2013). However, in the field, dispersal distances ranging from 6.82 to 7.56m per year were estimated in a recent population genetics study at fine spatial scales (Dupont et al. 2015). *A. icterica* is found deeper in soils and is considered to be more mobile, being theoretically able to travel up to 500 m.year^{-1} under constant artificial conditions (Mathieu et al. 2010; Caro et al. 2013). An extremely low signal of genetic structure was obtained for this species in a fine-scale population genetics study at the within plot scale (100x80m). This result was explained

by the great dispersal capacity of the species (Dupont et al. 2015). Moreover, no pattern of isolation by distance (IBD, i.e. decrease in the genetic similarity among populations as the geographic distance between them increases) was observed among six *A. icterica* populations separated by less than 13km (Torres-Leguizamon et al. 2014).

We analysed the relationship between landscape features and genetic variation in these two common earthworm species, at fine and regional scales (Fig. 1), in an agricultural landscape in North of France, where both species are common (e.g. Richard et al. 2012). First, we tested the hypothesis that the mosaic of habitats created by anthropogenic drivers alters the genetic diversity in earthworm populations using a buffer approach. It consists in assessing the correlation between the genetic variation of the earthworm population and the local landscape structure. Second, we tested the hypothesis that the different elements of the landscape could act either as dispersal barriers or corridors for earthworms with a landscape connectivity analysis at the regional scale (Zeller et al. 2012), encompassing Isolation by Distance (IBD), Least Cost Path (LCP) and resistance (CWD) approaches. The role of three elements, i.e. grasslands, crops and roads, was specifically tested. Grassland represents a suitable habitat that could be easily crossed by endogeic species (Bouché 1972; Decaens et al. 2008) while soil tillage and the use of pesticides in cultivated soils are known to have a detrimental effect on earthworms (Bertrand et al. 2015) and roads have been shown to represent dispersal corridors for invasive earthworms (Cameron & Bayne 2009).

Material and Methods

Study Area and Sampling

Earthworms were collected in Normandy (northern France, Fig. 2) in 2009 and 2010. The first sampling campaign was carried out in March and April 2009 in two pastures (P_A and P_B) ~ 500m apart at the local agricultural school "Lycée Agricole d'Yvetot". Details of these sampling sites and methods are given in Dupont *et al*. (2015). The second sampling campaign was carried out in April 2010, during which 39 other pastures were prospected. *A. chlorotica* and *A. icterica* were found in 14 and 19 pastures, respectively (Fig. 2 and Table 1). We selected pastures that had similar management histories, in order to reduce the effect of local environmental variations. The location of the plots was chosen in order to maximise the normality of the pairwise distance between plots. All pastures were at least 5 years old, and the great majority was grazed by cattle. Within each plot of 10 x 10m, 30 individuals were captured by sampling five monoliths of soil (25 x 25cm x 30 cm deep). If a species was present in the samples of a plot, but less than 30 individuals were captured, we sampled other monoliths - less than 10 m apart from the others- until the target of 30 individuals per population was reached. Individuals were preserved in pure ethanol for DNA analysis.

DNA extraction, microsatellite genotyping and basic genetic statistics

Total genomic DNA of *A. icterica* was extracted using either the CTAB extraction protocol, as described in Torres-Leguizamon *et al.* (2014) or the DNeasy 96 Blood & Tissue Kit (Quiagen). The latter was also used for *A. chlorotica*.

A. chlorotica individuals were genotyped at the eight microsatellite loci described in Dupont *et al*. (2011) while *A. icterica* individuals were genotyped at seven microsatellite loci

described in Torres-Leguizamon *et al*. (2012) and Dupont et al. (2015). Loci were amplified by polymerase chain reaction (PCR) following protocols detailed in Dupont *et al*. (2011), Torres-Leguizamon *et al*. (2012) and Dupont et al. (2015). The migration of PCR products was carried out on a 3130xl Genetic Analyser using the LIZ500 size standard, alleles were scored using GENESCAN V3.7 and GENOTYPER V3.7 software (Applied Biosystems, Foster City, CA, USA).

Individuals missing 3 or more loci (e.g. failed PCR, poor-quality DNA extract) were excluded from our dataset and mean genotyping error rates per locus and per allele (Pompanon et al. 2005) were estimated from repeat genotyping of 5% of samples (24 individuals per species). The null hypothesis of independence between loci was tested from statistical genotypic disequilibrium analysis using GENEPOP v. 4.4 (Rousset 2008). Evidence of null alleles was examined using the software MICRO-CHECKER (Van Oosterhout et al. 2004) and from the frequency of null homozygote within populations. The statistical power to detect genetic divergence was measured for all the samples and markers using POWSIM 4.0 to evaluate the hypothesis of genetic homogeneity under Fisher's exact tests (Ryman & Palm 2006). Microsatellite loci were tested for departure from Hardy–Weinberg equilibrium (HWE) within each sampling population using exact tests implemented in GENEPOP v. 4.4. To adjust for multiple comparisons, the FDR method (Benjamini & Hochberg 1995) as implemented in the software SGoF (http://webs.uvigo.es/acraaj/SGoF.htm) was applied.

Genetic variation of earthworm populations

For each population, the genetic diversity was analysed by computing allelic richness standardized for sample size (A_r; N= 26 and N = 9 for *A. chlorotica* and *A. icterica* respectively) using the program FSTAT v2.9.3.2 (Goudet 2000) and expected heterozygosity

(*He*) using Genetix v 4.05 (Belkhir et al. 2004). Weir and Cockerham's (1984) estimator of the inbreeding coefficient F_{is} was calculated using GENEPOP v. 4.4 (Rousset 2008). The distribution of the genetic diversity within populations can diverge from equilibrium models due to demographic changes. We tested whether the populations recently experienced a reduction of their effective size using the approach detailed in Cornuet & Luikart (1996) and implemented in their software BOTTLENECK v. 1.2.02. Using a Wilcoxon test, the observed heterozygosity was compared with the heterozygosity expected under equilibrium, considering a two-phase mutation model (TPM) recommended for microsatellite data (Piry et al. 1999) with 90% single-step mutations and 10% multiple-step mutations (and a variance among multiple step of 12). Populations exhibiting a significant heterozygosity excess would be considered as having experienced a recent genetic bottleneck whereas expanding populations (e.g. recovering from a bottleneck) are characterized by loci exhibiting a heterozygosity deficiency (Cornuet & Luikart 1996).

We estimated genetic differentiation between populations by calculating Weir and Cockerham's (1984) estimator of pairwise F_{st} values and carrying out exact tests of allelic differentiation between populations using GENEPOP v.4.0. To adjust for multiple comparisons, the FDR correction was used. Due to the frequent presence of null alleles, we used the program FREENA to calculate pairwise F_{st} estimates corrected for null alleles (F_{st_COR}) using the so-called ENA method (Chapuis & Estoup 2007). This software was also used to estimate the Cavalli-Sforza and Edwards (1967) genetic distance for each pair of populations (D_c) and this distance was also estimated using the INA correction described in Chapuis and Estoup ($D_{c\_COR,}$ 2007). Matrices of pairwise genetic distances were compared with Mantel tests (Mantel 1967) using the R program (R Development Core Team 2012).

We used the program BAPS v.6 (Corander & Marttinen 2006; Corander *et al.* 2008) to detect clusters of genetically similar populations and to estimate individual coefficients of ancestry with regard to the detected clusters. When testing for population clusters, we ran 5 replicates for k = 5, k= 10, k= 15, k= 20, k= 25 and k=30, where k is the maximum number of genetically divergent groups (populations). When estimating individual ancestry coefficients *via* admixture analysis we used recommended values of (i) the number of iterations used to estimate the admixture coefficients for the individuals (100), (ii) the number of reference individuals from each population (200) and (iii) the number of iterations used to estimate the admixture coefficients for the reference individuals (20).

Landscape genetics

Landscape elements were mapped at high resolution (precision ~2m) over the whole area. Land use cover and linear elements such as roads and rivers were obtained by merging different sources of data. As background data, we used databases from the French National Geographic Institute (IGN), encompassing shapefiles (BD TOPO, accuracy ~1m), and raster (BD Ortho, resolution = 0.5m) of the year 2010. We crossed this information with field work with a differential GPS with 10 cm real time accuracy, in order to check the boundaries of plots and their management. We also compared our data with Corine Land cover 2006 to identify any inconsistencies. Historical and management information was gained with google maps, from interviews with the farmers, and checked with the different version of Corine Land Cover. The data base and the different geographical layers were built up in ArcGis 10.1 (ESRI) in the projection Lambert 93 (EPSG: 2154). Data were stored in a vector format and rasterized at 10m resolution in order to perform the landscape analysis. Polygons and linear elements were rasterized separately and merged in raster format. Linear elements were

buffered before rasterization in order to avoid artefact gaps. Landscape structure variables were computed in Fragstats (McGarigal et al. 2012) and were computed at patch scale or at the buffer scale (500m of radius) depending on the metric. Landscape descriptors were then normalized (centred and reduced) and selected for the statistical modelling process based on their Variance Inflation Factor value (VIF), in order to avoid collinearity. There is no theoretical base to choose the threshold of the VIF value to exclude variables, and it is usually recommended to use a predictors with a VIF below ten (Montgomery & Peck 1992; Zuur *et al.* 2010; Dormann *et al.* 2013). We used a threshold of six in this study. Landscape structure descriptors were correlated to genetic diversity indices (A_r, H_e) with a forward multiple regression with AIC criterion. This approach produces r2 goodness of fit and avoid over-fitting, thanks to the AIC criteria. It assesses the significance of the separate effects, but the corresponding coefficient needs to be interpreted with caution. In particular, because to inherent correlation between landscape descriptors, the sign of the parameters cannot be interpreted straightforwardly: a positive parameter does not imply a positive net effect of the variable on the genetic variation.

Landscape connectivity was performed by defining different scenarios of cost of movements within landscape elements, based on species ecology. Elements were classified in three categories: Barrier, Neutral or Corridors, which corresponded to decreasing movement cost (50, 20, 1 respectively). The result is called a resistance surface map. In all scenarios, urban areas were considered to be barriers; forested areas, hedges and permanent water bodies were considered neutral and temporary water bodies were considered to be corridors. The other elements – grasslands, crops, roads - were considered differently according to the scenario (for details see Table 2). Combining all these possibilities yielded 27 scenarios of resistance surface. In order to test the robustness of our results we also run the analyses by

multiplying the costs by 100 in each scenario (giving costs of 5000, 2000, 100). The results were well congruent with initial costs. Connectivity was assessed in three ways. First, simple geographical distance along a straight line between all localities (Euclidian distance) was used to estimate the distance between localities. This scenario makes the assumption that landscape elements do not play a role in dispersal, and is usually referred to as Isolation by Distance (IBD). Second we calculated the least cost path between each pair of site for each of the 27 scenarios. This approach makes the assumption that individuals disperse optimally regarding landscape structure, and is usually referred to as Least Cost Path (LCP). Last, we calculated for all the 27 scenarios all paths between each pair of site, weighted by their cumulative cost, to produce 27 corresponding cost weighted distance matrices (CWD), which are usually referred to as resistance distances in circuit theory (McRae & Nürnberger 2006). All these spatial analyses were performed in R with the package {gdistance}. Once all pairwise distances were computed, we looked for the ones that best matched to the (logit transformed) genetic differentiation between populations (F_{st}/1-F_{st} and F_{st_COR}/1-F_{st_COR}). This was done using Maximum Likelihood Population Effect (MLPE, Clark et al. 2010; Van Strien et al. 2012), a type of linear mixed model that takes into account the non-independence of values within pairwise distance matrices. For this we adapted an R script supplied by Marteen J. Van Strien.

Results

Microsatellite data

All microsatellite markers were polymorphic across all populations, with 4–19 and 4 - 21 alleles per locus for *A. chlorotica* and *A. icterica*, respectively (Supplementary data Tables S1 and S2 respectively). We did not find any evidence of genotypic linkage disequilibrium at any pair of loci in any species. The mean genotyping error rate per locus was 3.12 % and 4.65 % in *A. chlorotica* and *A. icterica*, respectively (ranging from 0% to 8.33% and from 0% to 14.29%, respectively), mainly due to allelic dropouts. The mean genotyping error rate per allele was 1.56 % and 3.11 % in *A. chlorotica* and *A. icterica*, respectively (ranging from 0% to 4.17% and from 0% to 9.52%, respectively). Significant departures from HWE were observed in 39 of 112 and in 33 of 118 single-locus exact tests after FDR correction in *A. chlorotica* and *A. icterica*, respectively. Across all populations, the presence of null alleles was suggested by MICRO-CHECKER for all *A. chlorotica* loci except Ac 476, with frequencies ranging from 0.08 to 0.34 (Supplementary data Table S1) and for PB10, 2PE70 and C4 *A. icterica* loci, with frequencies ranging from 0.13 to 0.41 (Supplementary data Table S2). However, no locus showed null alleles in all populations. A few failures of amplification could be interpreted as null homozygotes that would confirm the presence of null alleles at some loci (Supplementary data Tables S1 and S2). However, amplification failures observed at loci that did not present heterozygote deficit, highlighted that the lack of amplification may be due to causes other than null alleles such as degraded DNA.

Genetic variation within populations

Higher values of genetic diversity were obtained for *A. chlorotica* than for *A. icterica* (Table 1). For example, standardized allelic richness (A_r) ranged from 6.30 to 9.28 and from 1.73 to 3.69 in *A. chlorotica* and *A. icterica*, respectively. A significant heterozygosity excess (Wilcoxon test, P<0.05) was detected in 1 *A. chlorotica* and 5 *A. icterica* populations (Table 1). Significant F_{is} was observed in almost all populations except in P_B and I36 for *A. chlorotica* and in I03, I10 and I31 for *A. icterica* (Table 1).

Relationship between genetic diversity and local landscape structure

Landscape features in buffers were correlated to genetic diversity in *A. icterica* but not in *A. chlorotica* (Table 3). In *A. icterica*, the r2 between A_r and H_e, and landscape features were respectively 0.41 and 0.56 and both were significant ($p < 0.05$). In this species, patch diversity was significantly linked to A_r and H_e, and patch richness was linked to A_r.

Genetic structure at regional scale

The statistical power for both sets of microsatellite loci to detect various levels of true population differentiation (F_{st}) between populations is presented Table 4. Both sets of markers will detect a true F_{st} of 0.005 or larger with a probability of 96% or more. The alpha error (corresponding to the probability of obtaining false significances when true $F_{st} = 0$) was close to 5% in all cases.

All genetic distances matrices ($F_{st}/1\text{-}F_{st}$, $F_{st\_COR}/1\text{-}F_{st\_COR}$, D_c and D_{c_COR}) were significantly correlated with Mantel r value ranging from 0.77 ($p = 0.001$) to 0.99 ($p = 0.001$) in *A. icterica* and from 0.88 ($p = 0.001$) to 0.99 ($p = 0.001$) in *A. chlorotica* (Supplementary data Tables S3, S4 and S5).

F_{st} analysis showed significant genetic structure at the level of the whole study for both species (F_{st} = 0.059, F_{st_COR} = 0.055, P<0.001 and Fst = 0.152, F_{st_COR} = 0.138, P< 0.001 for *A. chlorotica* and *A. icterica*, respectively). Pairwise F_{st} estimates ranged from 0.008 to 0.116 (F_{st_COR} ranged from 0.009 to 0.105) and 0.005 to 0.430 (F_{st_COR} ranged from 0.004 to 0.412) for *A. chlorotica* and *A. icterica*, respectively (Supplementary data Tables S3 and S4 respectively). All exact tests of allelic differentiation were significant (P ≤ 0.005). Analyses using BAPS identified 8 and 12 genetic clusters in *A. chlorotica* and *A. icterica*, respectively (P = 0.99 and P = 1 respectively). For *A. chlorotica*, one cluster was composed of 4 populations that were close geographically to one another (P_A, P_B, I32 and I33), another cluster was composed of the I07, I10, I15 and I18 populations and all other populations corresponded to a different cluster. For *A. icterica*, 5 clusters were composed of two geographically close populations (P_A and P_B, I02 and I03, I04 and I32, I07 and I08, I11 and I25) while all other clusters were composed of only one population. Low levels of admixture were observed among the clusters (Supplementary data Fig S1).

Relationship between genetic differentiation and landscape connectivity

Genetic variations were linked to landscape connectivity in both species (Table 5, Fig 3). In both species resistance distance (CWD) had the most explanatory power, followed by least cost path (LCP) and finally by isolation by distance (IBD). In *A. chlorotica*, the best scenarios were those in which grasslands were considered to be corridors, whereas crops and roads were considered to be barriers (Table 5). In *A. icterica*, no significant isolation by distance was found except with the non-corrected Dc distance and the only common point between the several likely scenarios was that crops were considered as barriers (Table 5). The best congruent models between the two species, taking into account the results from the

different genetic indices, were scenarios 9, 10 and 11. The most frequent role of the different landscape element in these scenarios was corridor for grasslands and barrier for crops and roads.

Discussion

Microsatellite markers in earthworms

Microsatellites markers have been developed for only a few earthworm species (i.e. 7 species so far, review in Torres-Leguizamon et al. 2012; Souleman et al. 2016) and these markers have rarely been used for population genetics studies (but see Velavan et al. 2009; Novo et al. 2010; Dupont et al. 2015). Two different research groups have tried to developed reliable markers for one of the most emblematic European earthworm species, *Lumbricus terrestris*. Of the ten markers obtained in this species by Velavan *et al*. (2007), only three were used in a subsequent study (Velavan et al. 2009) suggesting difficulties in genotyping the samples with the other ones. Souleman et al. (2016) couldn't obtain reliable results with these markers. Thus, they developed eight new markers for which they obtained a low amplification success and a significant heterozygote deficit, suggesting null alleles. In our study, null alleles were suspected at seven out of eight loci in *Allobophora chlorotica* and at four out of eight loci in *Aporrectodea icterica*. It is already known that the development of microsatellite molecular markers can be problematic in some taxa (e.g. in molluscs, McInerney et al. 2011). It was proposed that such methodological difficulties may have been caused by genomic complexities contained within microsatellite flanking regions. In particular, unstable flanking regions may arise when indels or mutations occur at PCR primer binding sites, thereby causing null alleles (McInerney et al. 2011). We therefore believe that microsatellite flanking regions are particularly variable in earthworm species. This could be verified by gathering more genomics data on these taxa. Nevertheless, we showed that the estimation of genetic divergence was not significantly altered by the presence of null alleles in

the dataset. Indeed, similar results were obtained with all indices of genetic divergence (F_{st}, F_{st_COR}, D_c, D_{c_COR}), and correction for the presence of null alleles did not change the results.

Landscape structure and population genetic diversity

Agriculture and urbanization result in habitat loss and fragmentation that variously impact many animal groups. Anthropogenic landscape fragmentation results in reduced size and increased isolation of habitat patches. Fragmented populations are thus expected to experience increased genetic drift and reduced gene flow, which result in the erosion of genetic diversity and the increase of genetic differentiation among local populations (Keyghobadi 2007). Moreover, small populations isolated by surrounding inhospitable landscapes are more vulnerable to demographic variability, environmental stochasticity and genetic processes including inbreeding depression, the random fixation of deleterious alleles and the loss of adaptive potential (Frankham 1995).

In this study, we tested how landscape structure in a man-made environment impacted genetic diversity of earthworm populations by characterizing landscape at the buffer scale. A significant relationship was observed between genetic diversity indices (H_e and A_r) and two landscape features (i.e. patch diversity and patch richness, Table 3) in *A. icterica* while no correlation was detected for *A. chlorotica*. We thus confirmed that geographic isolation of *A. icterica* populations due to natural and artificial barriers to gene flow probably accentuate the loss of genetic variability through genetic drift, such as already suggested in a previous population genetic study of this species (Torres-Leguizamon et al. 2014). Interestingly, one quarter of the *A. icterica* populations seemed recently founded, such as revealed by the heterozygosity excess in these populations. Overall, these results suggest that demographic

changes occur more frequently in *A. icterica* than in *A. chlorotica* and that these demographic changes can be explained by the local landscape structure.

Genetic differentiation between populations

A local decline of effective population size may be explained by the disruption of historical patterns of gene flow in a fragmented habitat (Keyghobadi 2007). Analyses of spatial patterns of genetic structure showed the presence of a strong genetic differentiation in both species, with a low degree of genetic admixture and high F_{st} values. F_{st} values were higher for *A. icterica*, highlighting that these populations are more genetically isolated than the ones of *A. chlorotica*. This was not expected because *A. icterica* has a higher potential for active dispersal. Caro *et al*. (2013) indeed demonstrated in a mesocosm study that *A. icterica* had a higher dispersal rate than two other endogeic species, namely *A. chlorotica* and *Aporrectodea caliginosa*. Moreover, in a recent population genetic study at very fine scale, Dupont *et al*. (2015) showed a low signal of genetic structure within two *A. icterica* populations sampled in two plots of less than 1 ha separated by ~ 500m while *A. chlorotica* populations showed spatial neighbourhood structure in the same sites. This difference was interpreted as a higher dispersal capacity of *A. icterica*. In the light of the results at very fine scale (Dupont et al. 2015) and at landscape scale (this study), we can assume that *A. chlorotica* essentially disperse through passive mechanisms over larger distance while passive dispersal might be more restricted for *A. icterica*. *A. chlorotica* is a small bodied species and lives near the soil surface in the upper 60 mm soil layer (Sims & Gerard 1999), two features probably facilitating dispersal *via* various vectors (e.g. zoochory, wind, water and soil transfer *via* human activities) while *A. icterica* is found deeper in the soil and is bigger.

Landscape connectivity at regional scale

The landscape connectivity analysis revealed that genetic structure was linked to landscape connectivity in both species, with resistance distance (cwd) having the most explanatory power. Thus, landscape features better explain genetic structure than Euclidian distances. We specifically tested the hypothesis that linear features such as roads may function as dispersal corridors (see for instance Tyser & Worley 1992; Cameron & Bayne 2009) for these species. It has indeed been shown that European earthworms that are invasive in Canada and the northern USA were introduced and spread along road networks (review in Cameron & Bayne 2009). It was however not clear whether the spread of earthworms along roads is more likely to occur *via* initial transport of earthworms or their cocoons in soil or gravel during road construction or *via* transport by vehicles after the road has been built (Cameron & Bayne 2009). In addition, roads and sidewalks could also function as dispersal corridors when earthworms crawl out of the soil and disperse at night after heavy rain, as is often observed in some species (e.g. Chuang & Chen 2008). Our results rather suggested that roads constitute obstacles for earthworm dispersal. Using MLPE, we indeed showed that the majority of the most likely landscape connectivity scenarios considered roads as barriers (Table 5).

The second hypothesis tested was that grasslands represent a suitable habitat that could be easily crossed by endogeic species and thus represent dispersal corridors, while soil under crops has a detrimental effect on earthworms. These expectations were confirmed by the MLPE analysis, the most likely landscape connectivity scenarios generally considered grasslands as corridors and crops as barriers.

Conclusion

Simultaneously investigating two ecologically similar species highlighted several common features in the response of each species to the landscape. We showed that functional connectivity was impacted by landscape features and that a favourable habitat could act as a corridor for the dispersal of earthworms. We thus confirmed that the aboveground landscape has a fundamental role in dispersal and gene flow of below-ground organisms. However, we also observed some differences between species which could be linked to the dispersal and life history attributes of each species. Indeed, population genetic diversity was significantly influenced by the local landscape structure in *A. icterica* but not in *A. chlorotica*. This result highlights that the exact effect of each habitat type on genetic variation over space and time and of agricultural practices on earthworm dispersal should be studied using specific sampling strategies.

Acknowledgements

This work was funded by the French National Research Agency (ANR) as a part of the project EDISP no. ANR-08-JCJC-0023 leaded by Jerome Mathieu. We thank all the participants to the EDISP project for their participation to field sampling. We also thank Roger Hijmans and Jacob van Etten for their advice on GIS analysis in R, Marteen J. Van Strien for the R script to perform MLPE and Naoise Nunan for English language editing.

References

Adriaensen F, Chardon JP, De Blust G*, et al.* (2003) The application of "least-cost" modelling as a functional landscape model. *Landscape and Urban Planning*, **64**, 233-247.

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

Benjamini Y Hochberg Y (1995) Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological*, **57**, 289-300.

Bertrand M, Barot S, Blouin M*, et al.* (2015) Earthworm services for cropping systems. A review. *Agronomy for Sustainable Development*, **35**, 553-567.

Blouin M, Hodson ME, Delgado EA*, et al.* (2013) A review of earthworm impact on soil function and ecosystem services. *European Journal of Soil Science*, **64**, 161-182.

Bouché MB (1972) Lombriciens de France. Ecologie et systématique INRA, Paris.

Bouché MB (1977) Stratégies lombriciennes. In: Soil organisms as components of ecosystems. (eds. Lohm, U., Persson, T.), pp. 122-132. Ecological Bulletin, Stockholm.

Cameron EK Bayne EM (2009) Road age and its importance in earthworm invasion of northern boreal forests. *Journal of Applied Ecology*, **46**, 28-36.

Caro G, Abourachid A, Decaëns T, Buono L Mathieu J (2012) Is earthworms' dispersal facilitated by the ecosystem engineering activities of conspecifics? *Biology and Fertility of Soils*, **48**, 961-965.

Caro G, Decaëns T, Lecarpentier C Mathieu J (2013) Are dispersal behaviours of earthworms related to their functional group? *Soil Biology & Biochemistry*, **58**, 181-187.

Cavalli-Sforza LL Edwards AWF (1967) Phylogenetic analysis. Models and estimation procedures. *American Journal of Human Genetics*, **19**, 233-257.

Chapuis MP Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621-631.

Chuang SC Chen JH (2008) Role of diurnal rhythm of oxygen consumption in emergence from soil at night after heavy rain by earthworms. *Invertebrate Biology*, **127**, 80-86.

Clark M, Tanguy A, Jollivet D*, et al.* (2010) Populations and Pathways: Genomic approaches to understanding population structure and environmental adaptation. In: Introduction to marine genomics (eds. Cock, M.J., Tessmar-Raible, K., Boyen, C., Viard, F.), pp. 73-118. Springer.

Corander J Marttinen P (2006) Bayesian identification of admixture events using multilocus molecular markers. *Molecular Ecology*, **15**, 2833-2843.

Corander J, Sirén J Arjas E (2008) Bayesian spatial modelling of genetic population structure. *Computational Statistics*, **23**, 111-129.

Cornuet J-M Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001-2014.

Costa D, Timmermans MJTN, Sousa JP*, et al.* (2013) Genetic structure of soil invertebrate populations: Collembolans, earthworms and isopods. *Applied Soil Ecology*, **68**, 61-66.

Decaens T, Margerie P, Aubert M, Hedde M Bureau F (2008) Assembly rules within earthworm communities in North-Western France - A regional analysis. *Applied Soil Ecology*, **39**, 321-335.

Dormann CF, Elith J, Bacher S*, et al.* (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 27-46.

Dupont L, Grésille Y, Richard B, Decaëns T Mathieu J (2015) Fine-scale spatial genetic structure and dispersal constraints in two earthworm species. *Biological Journal of the Linnean Society*, **114**, 335-347.

Dupont L, Lazrek F, Porco D*, et al.* (2011) New insight into the genetic structure of the *Allolobophora chlorotica* aggregate in Europe using microsatellite and mitochondrial data. *Pedobiologia*, **54**, 217-224.

Eijsackers H (2011) Earthworms as colonizers of natural and cultivated soil environments. *Applied Soil Ecology*, **50**, 1-13.

Frankham R (1995) Conservation genetics. *Annual Review of Genetics*, **29**, 305-327.

Goudet J (2000) FSTAT, a program to estimate and test gene diversities and fixation indices. Available from http://www.unil.ch/izea/softwares/fstat.html.

Hall LA Beissinger SR (2014) A practical toolbox for design and analysis of landscape genetics studies. *Landscape Ecology*, **29**, 1487-1504.

Holderegger R Wagner HH (2008) Landscape genetics. *Bioscience*, **58**, 199-207.

Jouquet P, Dauber J, Lagerlof J, Lavelle P Lepage M (2006) Soil invertebrates as ecosystem engineers: Intended and accidental effects on soil and feedback loops. *Applied Soil Ecology*, **32**, 153-164.

Keyghobadi N (2007) The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, **85**, 1049-1064.

Lavelle P, Bignell D, Lepage M*, et al.* (1997) Soil function in a changing world: the role of invertebrate ecosystem engineers. *European Journal of Soil Biology*, **33**, 159-193.

Lavelle P, Decaens T, Aubert M*, et al.* (2006) Soil invertebrates and ecosystem services. *European Journal of Soil Biology*, **42**, S3-S15.

Manel S Holderegger R (2013) Ten years of landscape genetics. *Trends in Ecology & Evolution*, **28**, 614-621.

Manel S, Schwartz MK, Luikart G Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189-197.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209-220.

Mathieu J, Barot S, Blouin M*, et al.* (2010) Habitat quality, conspecific density, and habitat pre-use affect the dispersal behaviour of two earthworm species, *Aporrectodea icterica* and *Dendrobaena veneta*, in a mesocosm experiment. *Soil Biology and Biochemistry*, **42**, 203-209.

Matthysen E (2012) Multicausality of dispersal : a review. In: Dispersal Ecology and Evolution (eds. Clobert, J., Baguette, M., Benton, T.G., Bullock, J.M.), pp. 3-18. Oxford University Press, Oxford.

McGarigal K, Cushman SA Ene E (2012) FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps., p. Available at the following web site: http://www.umass.edu/landeco/research/fragstats/fragstats.html. produced by the authors University of Massachusetts, Amherst.

McInerney CE, Allcock AL, Johnson MP, Bailie DA Prodohl PA (2011) Comparative genomic analysis reveals species-dependent complexities that explain difficulties with microsatellite marker development in molluscs. *Heredity*, **106**, 78-87.

McRae BH (2006) Isolation by resistance. *Evolution*, **60**, 1551-1561.

McRae BH Nürnberger B (2006) Isolation by resistance. *Evolution*, **60**, 1551-1561.

Montgomery DC Peck EA (1992) Introduction to Linear Regression Analysis. Wiley, New York.

Novo M, Almodovar A, Fernandez R, Gutierrez M Cosin DJD (2010) Mate choice of an endogeic earthworm revealed by microsatellite markers. *Pedobiologia*, **53**, 375-379.

Palm J, van Schaik NLMB Schroeder B (2013) Modelling distribution patterns of anecic, epigeic and endogeic earthworms at catchment-scale in agro-ecosystems. *Pedobiologia*, **56**, 23-31.

Piry S, Luikart G Cornuet JM (1999) BOTTLENECK: A Computer Program for Detecting Recent Reductions in the Effective Population Size Using Allele Frequency Data. *Journal of Heredity*, **90**, 502-503.

Pompanon F, Bonin A, Bellemain E Taberlet P (2005) Genotyping errors: Causes, consequences and solutions. *Nature Reviews Genetics*, **6**, 847-859.

R Development Core Team (2012) R : A language and environment for statistical computing. In: *R Foundation for Statistical Computing*, Vienna, Austria.

Richard B, Legras M, Margerie P*, et al.* (2012) Spatial organization of earthworm assemblages in pastures of northwestern France. *European Journal of Soil Biology*, **53**, 62-69.

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

Ryman N Palm S (2006) POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. *Molecular Ecology Notes*, **6**, 600-602.

Sims RW Gerard BM (1999) Synopsis of the British Fauna (31)- earthworms, pp. 1-169. The Linnean Society of London and the Estuarine and Brackish-water Sciences association, London.

Souleman D, Grumiaux F, Frerot H, Vandenbulcke F Pauwels M (2016) Isolation and characterization of eight polymorphic microsatellites markers for the earthworm Lumbricus terrestris. *European Journal of Soil Biology*, **74**, 76-80.

Spear SF, Balkenhol N, Fortin MJ, McRae BH Scribner K (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular Ecology*, **19**, 3576-3591.

Storfer A, Murphy MA, Evans JS*, et al.* (2007) Putting the 'landscape' in landscape genetics. *Heredity*, **98**, 128-142.

Storfer A, Murphy MA, Spear SF, Holderegger R Waits LP (2010) Landscape genetics: where are we now? *Molecular Ecology*, **19**, 3496-3514.

Sullivan TJ, Dreyer AP Peterson JW (2009) Genetic variation in a subterranean arthropod (Folsomia candida) as a method to identify low-permeability barriers in an aquifer. *Pedobiologia*, **53**, 99-105.

Taylor P, Fahrig L With KA (2006) Landscape connectivity: back to the basics. In: Connectivity Conservation (eds. Crooks, K., Sanjayan, M.A.), pp. 29–43. Cambridge University Press, Cambridge.

Taylor PD, Fahrig L, Henein K Merriam G (1993) Connectivity Is a Vital Element of Landscape Structure. *Oikos*, **68**, 571-573.

Torres-Leguizamon M, Mathieu J, Decaëns T Dupont L (2014) 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.

Torres-Leguizamon M, Mathieu J, Livet A Dupont L (2012) Isolation of polymorphic microsatellite markers in *Aporrectodea icterica* (Savigny 1826). *Soil Biology & Biochemistry*, **51**, 16-19.

Tyser RW Worley CA (1992) Alien flora in grasslands adjacent to road and trail corridors in glacier national-park, Montana (USA) *Conservation Biology*, **6**, 253-262.

Van Oosterhout C, Hutchinson WF, Wills DPM Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535-538.

Van Strien MJ, Keller D Holderegger R (2012) A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Molecular Ecology*, **21**, 4010-4023.

Vanbergen AJ, Watt AD, Mitchell R*, et al.* (2007) Scale-specific correlations between habitat heterogeneity and soil fauna diversity along a landscape structure gradient. *Oecologia*, **153**, 713-725.

Velavan TP, Schulenburg H Michiels N (2007) Development and characterization of novel microsatellite markers for the common earthworm (*Lumbricus terrestris* L). *Mol. Ecol. Notes*, **7**, 1060-1062.

Velavan TP, Weller S, Schulenburg H Michiels NK (2009) High genetic diversity and heterogeneous parasite load in the earthworm *Lumbricus terrestris* on a German meadow. *Soil Biology & Biochemistry*, **41**, 1591-1595.

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

Zeller K, McGarigal K Whiteley A (2012) Estimating landscape resistance to movement: a review. *Landscape Ecology*, **27**, 777-797.

Zuur AF, Ieno EN Elphick CS (2010) A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, **1**, 3-14.

Data Accessibility

The data sets with the microsatellite genotypes of *A. chlorotica* and *A. icterica* are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.sc6bb

Author Contributions

J.M initiated the project. L.D. and J.M. designed the study. L.D. drafted the manuscript. L.D., M.T.L., P. R.-C. and J.M. collected the data. L.D and P. R.-C. performed molecular analyses. J.M. built the GIS database and conducted the landscape genetic analyses (Buffer and connectivity). L.D. and J.M. performed statistical analyses and wrote the manuscript.

Figure legends

Figure 1

Illustration of the different landscape genetic approaches used in this work. a) Local genetic structure is compared to local landscape features within a buffer around each population; b) genetic difference is linked to geographical distance between populations (IBD); c) genetic difference is linked to the functional distance between populations, calculated as the least cost path between populations (LCP), based on a resistance map; d) genetic difference is linked to the functional distance between populations, calculated as the sum of the cost weighted paths between populations (cwd), based on a resistance map.

Figure 2

Geographical distribution of earthworm sampling sites near Yvetot in Normandy (France). "No species" means that neither *A. icterica* nor *A. chlorotica* were found.

Figure 3

Best correlation found between genetic distance (D_c distance) and resistance distance scenario (cwd) for *A. chlorotica* (A) and *A. icterica* (B).

Table 1. Geographical location*, sample size (N), genetic diversity (A_r: standardized allelic richness; H_e: expected heterozygosity), fixation index F_{is} (significant values are in bold) and Wilcoxon test P value for heterozygote excess compared to expectations at mutation-drift equilibrium (p_{wil}) for each study plot and each species. *Latitude/longitude range in Lambert II étendu.

Plot	**Latitude**	**Longitude**	***Allolobophora chlorotica***					***Aporrectodea icterica***				
			N	A_r	H_e	F_{is}	p_{wil}	**N**	A_r	H_e	F_{is}	p_{wil}
P_A	484578.509	2513691.360	42	8.61	0.777	**0.128**	0.727	22	2.73	0.523	**0.352**	0.055
P_B	484100.668	2513845.836	30	8.44	0.794	0.091	0.527	30	2.51	0.481	**0.094**	**0.016**
I02	484907.071	2511340.636	-	-	-	-	-	12	2.27	0.397	**0.274**	0.281
I03	484824.797	2511382.199	-	-	-	-	-	28	2.32	0.395	0.115	**0.015**
I04	483636.894	2511737.970	-	-	-	-	-	12	2.55	0.438	**0.103**	0.344
I07	483713.045	2511076.294	28	8.60	0.775	**0.185**	0.371	24	2.95	0.494	**0.380**	0.281
I08	483823.122	2511110.760	-	-	-	-	-	31	2.71	0.484	**0.209**	0.406
I10	483411.237	2511337.635	32	7.75	0.781	**0.233**	0.422	14	2.61	0.495	0.167	**0.008**
I11	485746.167	2507374.639	27	7.45	0.765	**0.195**	0.527	29	2.81	0.449	**0.172**	0.406
I15	481854.533	2514843.110	27	9.28	0.777	**0.141**	0.727	30	2.54	0.461	**0.065**	**0.008**
I18	479823.687	2512611.239	27	8.08	0.742	**0.205**	0.973	30	3.69	0.575	**0.277**	0.344
I19	492929.229	2508850.328	-	-	-	-	-	43	3.00	0.493	**0.335**	0.711
I20	477908.835	2515839.136	29	7.06	0.725	**0.293**	0.808	33	2.15	0.373	**0.184**	0.312
I25	480555.744	2508635.999	29	6.30	0.720	**0.229**	0.527	31	2.84	0.506	**0.263**	0.078

I27	489902.819	2513148.683	30	8.81	0.788	**0.156**	0.770	34	2.62	0.485	**0.317**	**0.023**
I31	482053.828	2513328.538	-	-	-	-	-	27	1.73	0.226	0.115	0.594
I32	483297.736	2513724.785	29	8.16	0.786	**0.179**	0.273	30	2.70	0.422	**0.305**	0.656
I33	481742.786	2511459.498	32	8.61	0.761	**0.182**	0.902	29	2.52	0.454	**0.179**	0.148
I34	482430.931	2510886.051	32	7.95	0.747	**0.237**	0.875	30	2.33	0.366	**0.403**	0.078
I36	487833.362	2509076.597	46	6.67	0.777	0.067	**0.010**	-	-	-	-	-

Table 2 Role and cost of landscape elements in the different landscape scenarios tested in this study. The cost of grasslands, crops and roads changed among scenarios. Costs of permanent water bodies (P.W.B.), temporary water bodies (T.W.B.), deciduous forest (D. Forest), coniferous forest (C. Forest), hedges and urban area (Urban A.) were fixed in all scenarios.

	Role of elements			**Cost of elements**								
Scenario	**Grasslands**	**Crops**	**Roads**	**Grasslands**	**Crops**	**Roads**	**P. W. B.**	**T. W. B.**	**D. Forest**	**C. Forest**	**Hedges**	**Urban A.**
1	Barrier	Barrier	Barrier	50	50	50	20	1	20	20	20	50
2	Barrier	Barrier	Neutral	50	50	20	20	1	20	20	20	50
3	Barrier	Barrier	Corridor	50	50	1	20	1	20	20	20	50
4	Barrier	Corridor	Barrier	50	1	50	20	1	20	20	20	50
5	Barrier	Corridor	Neutral	50	1	20	20	1	20	20	20	50
6	Barrier	Corridor	Corridor	50	1	1	20	1	20	20	20	50
7	Barrier	Neutral	Barrier	50	20	50	20	1	20	20	20	50
8	Barrier	Neutral	Neutral	50	20	20	20	1	20	20	20	50
9	Barrier	Neutral	Corridor	50	20	1	20	1	20	20	20	50
10	Corridor	Barrier	Barrier	1	50	50	20	1	20	20	20	50
11	Corridor	Barrier	Neutral	1	50	20	20	1	20	20	20	50
12	Corridor	Barrier	Corridor	1	50	1	20	1	20	20	20	50
13	Corridor	Corridor	Barrier	1	1	50	20	1	20	20	20	50
14	Corridor	Corridor	Neutral	1	1	20	20	1	20	20	20	50

15	Corridor	Corridor	Corridor	1	1	1	20	1	20	20	20	50
16	Corridor	Neutral	Barrier	1	20	50	20	1	20	20	20	50
17	Corridor	Neutral	Neutral	1	20	20	20	1	20	20	20	50
18	Corridor	Neutral	Corridor	1	20	1	20	1	20	20	20	50
19	Neutral	Barrier	Barrier	20	50	50	20	1	20	20	20	50
20	Neutral	Barrier	Neutral	20	50	20	20	1	20	20	20	50
21	Neutral	Barrier	Corridor	20	50	1	20	1	20	20	20	50
22	Neutral	Corridor	Barrier	20	1	50	20	1	20	20	20	50
23	Neutral	Corridor	Neutral	20	1	20	20	1	20	20	20	50
24	Neutral	Corridor	Corridor	20	1	1	20	1	20	20	20	50
25	Neutral	Neutral	Barrier	20	20	50	20	1	20	20	20	50
26	Neutral	Neutral	Neutral	20	20	20	20	1	20	20	20	50
27	Neutral	Neutral	Corridor	20	20	1	20	1	20	20	20	50

Table 3 Summary of the forward multiple regression between landscape features (predictors) and genetic diversity (A_r: Allele Richness, H_e: expected heterozygosity) of the populations of the two earthworm species *A. icterica* and *A. chlorotica* in the region of Yvetot, Normandy, France. A patch represents an element in the landscape. Predictors are: Edge Density (length of patch edge/surface of the buffer), Patch Richness (number of different types of patches in the buffer), Patch Diversity (Shannon Diversity of the different Land Use types), Crop surface, Grassland surface and Total length of roads.

	Aporrectodea icterica		***Allolobophora chlorotica***	
	A_r	H_e	A_r	H_e
Anova table				
F	3.10	4.85	1.62	1.41
p - value	**<0.05**	**<0.01**	ns	ns
adjusted r2	0.41	0.56	0.22	0.16
Predictors				
Edge Density	ns	ns	ns	ns
Patch Richness	ns	-0.10	ns	ns
Patch Diversity	ns	0.74	ns	ns
Crop surface	ns	ns	ns	ns
Grassland surface	ns	ns	ns	ns
Total length of roads	ns	ns	ns	ns

Table 4 Statistical power for detecting various true levels of population differentiation (F_{st}) by means of Fisher's exact test when using both sets of microsatellite loci, allele frequencies, and sample sizes. The power is expressed as the proportion of simulations that provide statistical significance at the 0.05 level.

True F_{st}	*A. chlorotica*	*A. icterica*
0.000	0.064	0.070
0.001	0.417	0.196
0.002	0.852	0.432
0.005	1.000	0.964
0.010	1.000	1.000

Table 5 Summary of the network regression analyses of landscape connectivity comparing **A**. Euclidian distance and genetic differentiation (pairwise F_{st} , pairwise F_{st} estimates corrected for null alleles, D_c genetic distance and D_c genetic distance corrected for null alleles) between populations (Isolation by distance, IBD).and **B**. cost weighted distances (cwd) and genetic differentiation between populations. The most likely landscape connectivity scenarios are indicated. The roles of the landscape elements in the most likely scenarios are specified (b = barrier, c = corridor and n = neutral); when applicable the most frequent role is in bold. In **C**. the best congruent scenarios are presented. NA = Not Applicable

A. **IBD**

	Genetic distance	**Geographic distance**	**scenario**	**Grasslands**	**Crops**	**Roads**	**r2**	**pvalue**
A. chlorotica	F_{st}	euclidian	NA	-	-	-	0.29	$<10^{-5}$
	$F_{st\text{-}COR}$	euclidian	NA	-	-	-	0.28	$<10^{-5}$
	D_c	euclidian	NA	-	-	-	0.4	$<10^{-5}$
	$D_{c\text{-}COR}$	euclidian	NA	-	-	-	0.41	$<10^{-5}$
A. icterica	F_{st}	euclidian	NA	-	-	-	0.01	0.12
	$F_{st\text{-}COR}$	euclidian	NA	-	-	-	0.02	0.09
	D_c	euclidian	NA	-	-	-	0.03	$<10^{-5}$
	$D_{c\text{-}COR}$	euclidian	NA	-	-	-	0.02	0.06

B. **Best species specific scenario**

	Genetic distance	**Geographic distance**	**scenario**	**Grasslands**	**Crops**	**Roads**	**r2**	**pvalue**
A. chlorotica	F_{st}	CWD	10	**c**	**b**	**b**	0.44	$<10^{-5}$
	$F_{st\text{-}COR}$	CWD	10	**c**	**b**	**b**	0.43	$<10^{-5}$

	D_c	CWD	8, 9, 10	b, **c**	**b**, n	**b**, n, c	0.54	$<10^{-5}$
	$D_{c\text{-}COR}$	CWD	8, 9, 14	b, **c**	n, c	n, c	0.54	$<10^{-5}$
A. icterica	F_{st}	CWD	1, 2, 3, 11, 12, 21	b, n, c	**b**, c	b, n, c	0.04	$<10^{-4}$
	$F_{st\text{-}COR}$	CWD	1, 2, 3, 11, 12, 21	b, n, c	**b**, c	b, n, c	0.05	$<10^{-4}$
	D_c	CWD	1, 2, 3, 11, 12, 21	b, n, c	**b**, c	b, n, c	0.08	$<10^{-5}$
	$D_{c\text{-}COR}$	CWD	1, 11, 12, 21, 25, 26, 27	b, n, c	**b**, n	b, n, c	0.06	$<10^{-5}$

C. **Best congruent scenario**

	Genetic distance	**Geographic distance**	**scenario**	**Grasslands**	**Crops**	**Roads**	**r2**	**pvalue**
A.chlorotica	F_{st}	CWD	10	**c**	**b**	**b**	0.44	$<10^{-5}$
	$F_{st\text{-}COR}$	CWD	10	**c**	**b**	**b**	0.43	$<10^{-5}$
	D_c	CWD	9, 10	b, **c**	n, **b**	**b**, c	0.54	$<10^{-5}$
	$D_{c\text{-}COR}$	CWD	9, 10, 11	b, **c**	n, **b**	**b**, n, c	0.54	$<10^{-5}$
A.icterica	F_{st}	CWD	10	**c**	**b**	**b**	0.04	$<10^{-4}$
	$F_{st\text{-}COR}$	CWD	10	**c**	**b**	**b**	0.04	$<10^{-4}$
	D_c	CWD	9, 10	b, **c**	n, **b**	**b**, c	0.07	$<10^{-5}$
	$D_{c\text{-}COR}$	CWD	9, 10, 11	b, **c**	n, **b**	**b**, n, c	0.06	$<10^{-4}$

Figure 1

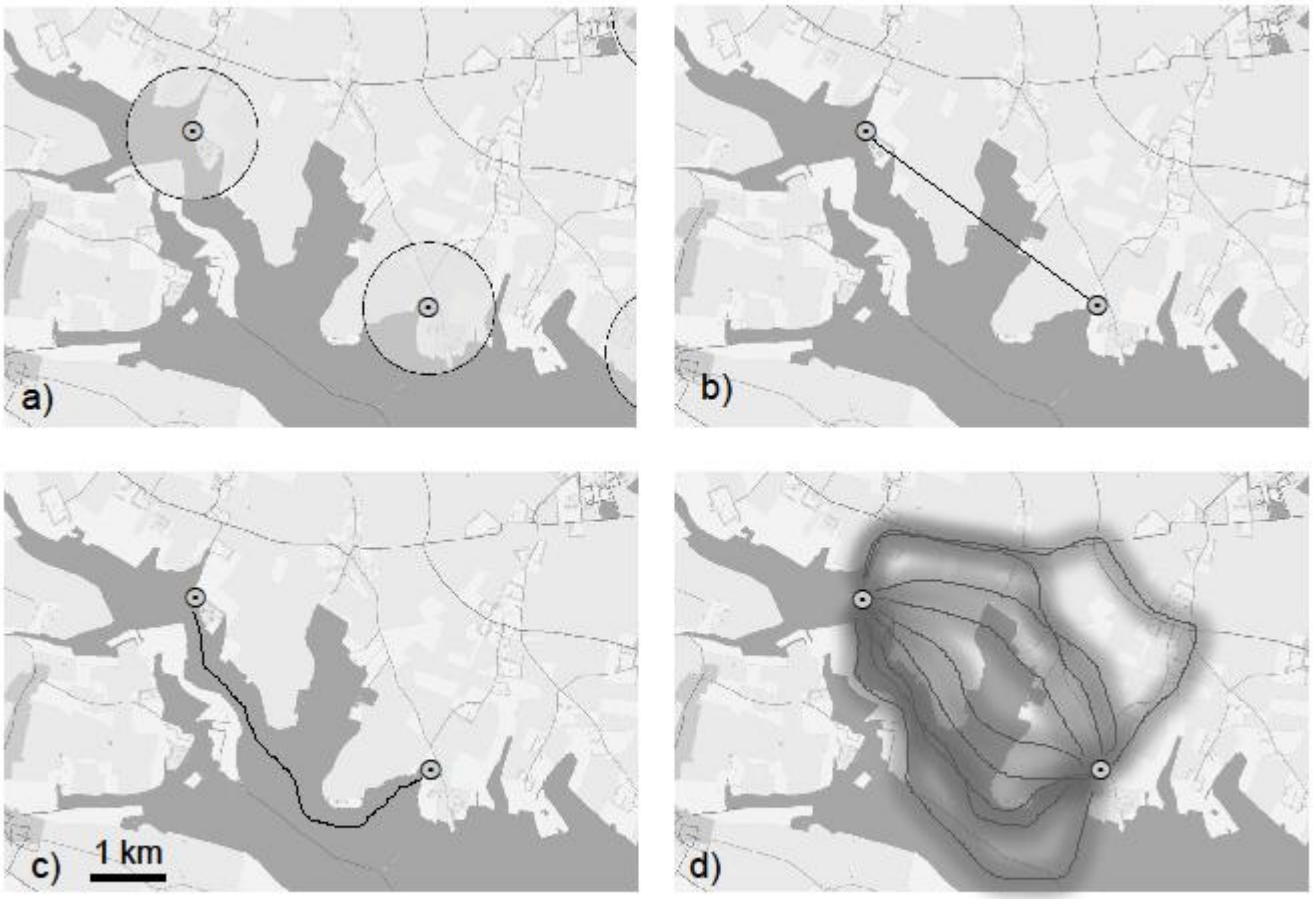


Figure 2

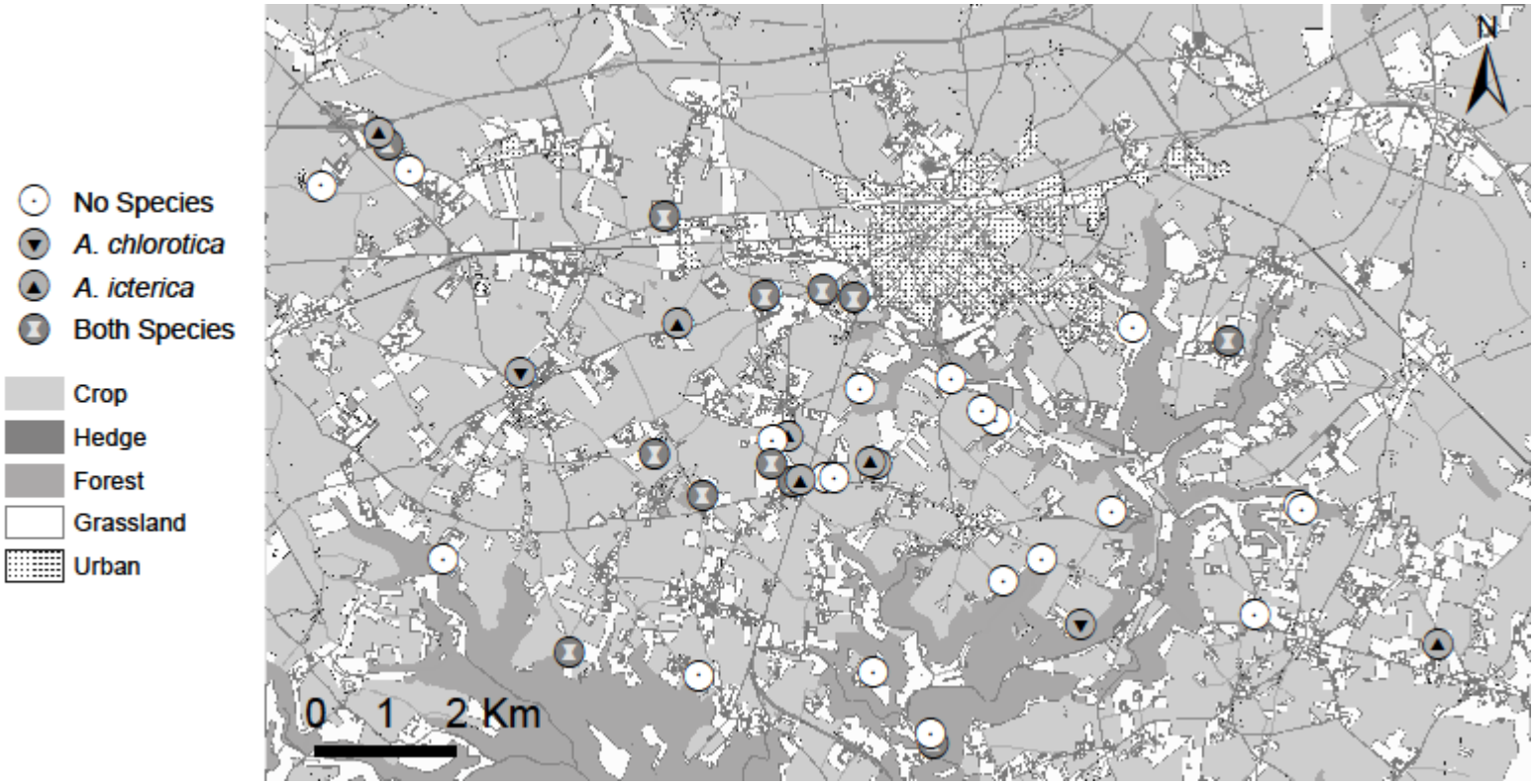
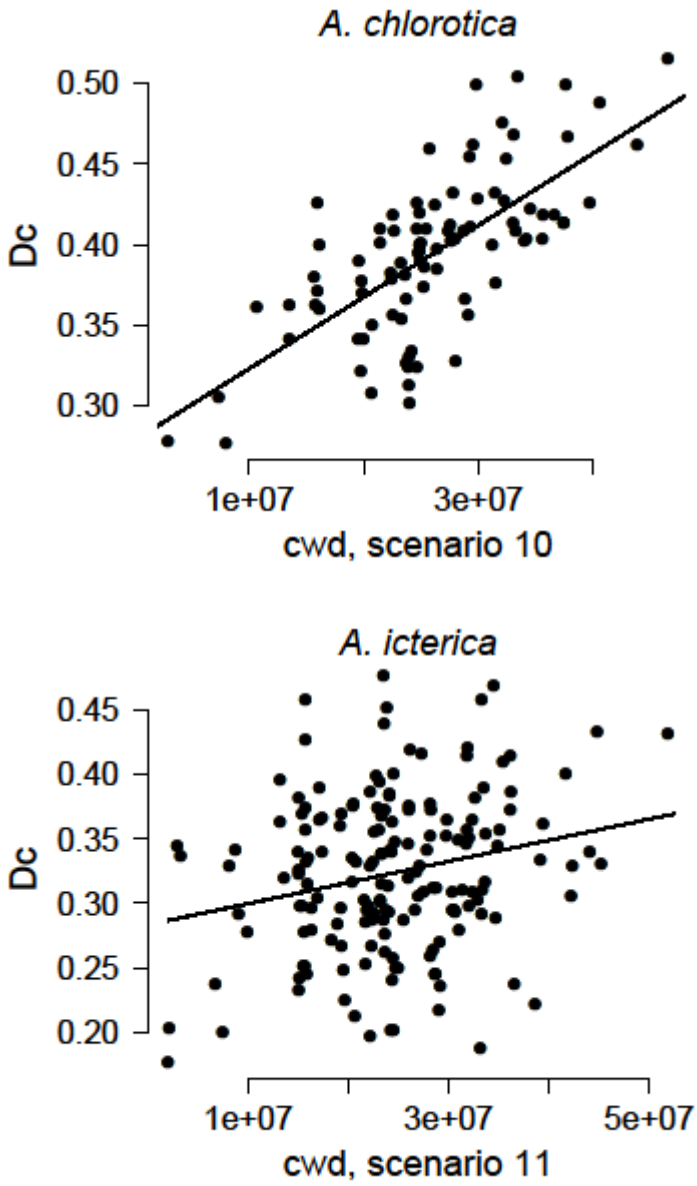

No Species
A. chlorotica
A. icterica
Both Species
Crop
Hedge
Forest
Grassland
Urban
N
0 1 2 Km

Figure 3


A. chlorotica
0.50
0.45
0.40
0.35
0.30
Dc
1e+07
3e+07
cwd, scenario 10
A. icterica
0.45
0.40
0.35
0.30
0.25
0.20
Dc
1e+07
3e+07
5e+07
cwd, scenario 11