



HAL
open science

Genetic variation of the epigeic earthworm [*Lumbricus castaneus*] populations in urban soils of the Paris region (France) revealed using eight newly developed microsatellite markers

Lise Dupont, Maxime Pauwels, Cassandre Dume, Valentin Deschins, Helene Audusseau, Agnès Gigon, Florence Dubs, Franck Vandebulcke

► To cite this version:

Lise Dupont, Maxime Pauwels, Cassandre Dume, Valentin Deschins, Helene Audusseau, et al.. Genetic variation of the epigeic earthworm [*Lumbricus castaneus*] populations in urban soils of the Paris region (France) revealed using eight newly developed microsatellite markers. *Applied Soil Ecology*, 2019, 135, pp.33-37. 10.1016/j.apsoil.2018.11.004 . hal-02624014

HAL Id: hal-02624014

<https://hal.inrae.fr/hal-02624014>

Submitted on 19 Oct 2021

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

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



Distributed under a Creative Commons Attribution - NonCommercial| 4.0 International License

1 **Short communication**

2

3 **Genetic variation of the epigeic earthworm *Lumbricus castaneus* populations in urban**
4 **soils of the Paris region (France) revealed using eight newly developed microsatellite**
5 **markers**

6

7 Lise Dupont ¹, Maxime Pauwels ², Cassandre Dume ^{1,3}, Valentin Deschins ¹, Hélène
8 Audusseau ¹, Agnès Gigon ¹, Florence Dubs ¹ and Franck Vandebulcke ³.

9

10 1. Université Paris Est Créteil (UPEC), Sorbonne universités, Paris 7, CNRS, INRA, IRD,
11 Institut d'écologie et des sciences de l'environnement de Paris, 94010 Créteil Cedex, France

12 2. Université de Lille, CNRS, UMR 8198 – Unité Evolution-Ecologie-Paléontologie, F-59000
13 Lille, France

14 3. Université de Lille, EA 4515-LGCgE - Laboratoire Génie Civil et géo-Environnement, Cité
15 scientifique, SN3, F-59655 Villeneuve d'Ascq, France

16

17

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

19

20 **Keywords:** earthworm populations, genetic diversity, inbreeding, urban area

21 **Highlights**

- 22 - Eight new microsatellite loci were developed for the earthworm *Lumbricus castaneus*
- 23 - These markers revealed an important genetic diversity in 6 populations of the Paris region
- 24 - Geographically close populations were genetically significantly different.

25

26 **Abstract**

27 Urban soils are subject to intense environmental pressures, e.g. physical disturbance,
28 surface transformation and pollution, which greatly impact the activity of soil organisms. The
29 epigeic earthworm *Lumbricus castaneus* (Savigny, 1826) is known to be tolerant to trace
30 elements contamination of soils and was found to be abundant in the urban context. In order
31 to investigate how urban environmental stressors shape the population genetic variation of
32 this species, we developed 8 microsatellite loci from a microsatellite-enriched genomic
33 library. Polymorphism was explored in 6 populations (153 individuals) from the Paris region.
34 The number of alleles per locus varied from 7 to 29. The average expected heterozygosity
35 within populations ranged from 0.596 to 0.705. These new microsatellite markers revealed a
36 significant genetic structure at fine spatial scale, highlighting that they could be particularly
37 useful for genetic studies of populations in urban soils.

38

39 **1. Introduction**

40 Levels of genetic variation within and among populations are determined by the joint
41 action of mutation, selection, genetic drift, and gene flow (Hartl and Clark, 2007), which in
42 turn operate within the historical, ecological and biological context of each species (Loveless
43 and Hamrick, 1984). Thus, ecological preferences and life history traits can play a major role
44 in shaping the population genetic structure of a species (Gonzalez and Zardoya, 2007).
45 Understanding the evolution of genetic variation is a priority for biodiversity conservation and
46 environmental biomonitoring programs, in particular because levels of genetic variation is a
47 major determinant of the long-term maintenance of species. This is particularly true in urban
48 environments where environmental disturbances of anthropogenic origin such as pollution,
49 geographic isolation and habitat fragmentation can profoundly affect population genetic
50 parameters (Johnson and Munshi-South, 2017). Despite their challenging environment, some
51 species are able to exploit urban habitats and to maintain their populations (Takami et al.,
52 2004). It is important for biodiversity management programs to determine the genetic
53 background of these species evolving in urban ecosystems.

54 Soils are the foundation for many ecological processes and provide key ecosystem
55 services even in urban areas where they are polluted, physically disturbed and frequently
56 artificial (review in Pickett et al., 2011). Elevated trace element concentrations are universally
57 reported in urban soils and sources have been associated with roadside environments, interior
58 and exterior paint, refuse incinerators, industrial stack emissions, management and industrial
59 waste (Pickett et al., 2011). Such contamination of soils by trace elements is known to affect
60 exposed species, such as earthworms (Mussali-Galante et al., 2014). In particular, metal
61 pollution is expected to profoundly alter the genetic variation within populations, in four
62 different ways : (i) by increasing mutation rates, (ii) by directional selection on tolerant
63 genotypes, (iii) by causing bottleneck events, and (iv) by altering migration (i.e. reducing or

64 increasing gene flow in relation with the level of tolerance of the individuals) (Mussali-
65 Galante et al., 2014; van Straalen and Timmermans, 2002).

66 Microsatellites are commonly used to infer microevolutionary processes such as
67 mutation, genetic drift and gene flow and provide invaluable insights into the overall levels of
68 neutral genetic variation (i.e. neutral means that gene variants detected do not have any direct
69 effect on fitness) within populations (Kirk and Freeland, 2011). Although microsatellite
70 markers have been developed for eight earthworm species to date: *Lumbricus rubellus*
71 (Harper et al., 2006), *L. terrestris* (Souleman et al., 2016; Velavan et al., 2007), *Hormogaster*
72 *elisae* (Novo et al., 2008), *Allolobophora chlorotica* (Dupont et al., 2011), *Aporrectodea*
73 *icterica* (Torres-Leguizamon et al., 2012), *Aporrectodea longa* (Strunk et al., 2012), *Eisenia*
74 *fetida* (Somers et al., 2011) and *Amyntas cortices* (Cunha et al., 2017), they have not been
75 used to investigate genetic changes in population exposed to trace elements yet.

76 Here, our main aim was to develop and characterize microsatellite loci for *L.*
77 *castaneus* (Savigny, 1826), a diploid and obligatory amphimictic earthworm species. This
78 epigeic earthworm is known to be tolerant to trace element contamination of soils (Grumiaux
79 et al., 2015; Spurgeon and Hopkin, 1996). For instance, Spurgeon and Hopkin (1996) found
80 that zinc was the most toxic trace element for six different earthworm species living close to
81 smelting works at Avonmouth (southwest England) but *L. castaneus* was the most tolerant
82 and survived in soil containing up to 3627 mg kg⁻¹ of Zn. This tolerance to elevated trace
83 element concentrations makes *L. castaneus* a good candidate for genetic studies in urban
84 conditions, such as in the Paris region where mean concentration of Zn varies approximately
85 between 107 and 174 mg kg⁻¹ in soils of lawn and woods respectively (Foti et al., 2017).
86 Metal-tolerant populations of *L. castaneus* situated in metal contaminated sites surrounded by
87 non-contaminated areas provide the opportunity to investigate the initial steps in the
88 establishment of differentiated populations under severe selection pressures and strong

89 demographic bottlenecks (e.g. (Mengoni et al., 2000). These populations growing in
90 contaminated sites are thus valuable models for the study of microevolutionary processes
91 (Macnair, 1987). In that context, our main aims were (i) to develop eight polymorphic
92 microsatellite loci from a microsatellite-enriched genomic library and (ii) to assess their
93 usefulness for genetic studies of *L. castaneus* populations in urban environment by examining
94 their genetic variability in six populations of the Paris region.

95

96 **2. Material and Methods**

97 All collected specimens of *Lumbricus castaneus* were identified based on external
98 morphology using the key of Bouché (1972). Earthworms were sampled manually in top 10
99 cm of soil using a solution of Allyl isothiocyanate (AITC) diluted with isopropanol (propan-2-
100 ol, RPE grade, Carlo-Erba) to obtain a 5 g L⁻¹ solution (Zaborski, 2003). Earthworms were
101 rinsed and cleaned under tap water and stored in ethanol 96% until use.

102 For microsatellite-enriched genomic library construction, we maximized the genomic
103 diversity by using eight individuals that were collected in two sites of northern Nord-Pas-de-
104 Calais region, approximately 150 km apart (3 specimens from
105 50°51'46.119"N/1°50'31.765"E and 5 specimens from 50°26'56.465"N/3°2'56.558"E), and
106 one individual that was collected in southern Poland (50°01'38.0"N/20°15'02.9"E). Total
107 genomic DNA was isolated using the NucleoSpin® 96 Tissue kit (Macherey Nagel, France).
108 A stoichiometric mixture of the 9 DNAs was performed to produce a microsatellite enriched
109 genomic library which was constructed by the Genoscreen Company (Lille) by coupling
110 multiplex microsatellite enrichment isolation techniques with the 454 GS-FLX Titanium
111 pyrosequencing (Malausa et al., 2011). Enrichment was performed using probes containing
112 the following microsatellite motifs: TG, TC, AAC, AAG, AGG, ACG, ACAT and ACTC.
113 Overall, 11875 sequences containing a single microsatellite motif were found (Accession no.

114 in the NCBI Short Read Archives SRP154315). This allows the identification of 664
115 independent microsatellite loci candidates for which amplification primer pairs were designed
116 in silico. Among them, 40 microsatellite loci were chosen in order to diversify the repeated
117 motif (di – tri- and tetra-nucleotides motifs), to maximize the number of repeats of the motif,
118 and to allow multiplexing (i.e. different expected amplicon sizes).

119 Amplification trials, tests of polymorphism and subsequent analyses of genetic
120 structure were performed on six populations from the Paris region. Four of these populations,
121 namely the Buttes-Chaumont park (BCH), the Butte du chapeau rouge square (BCR), the
122 Belleville park (BEL), and the Compans square (COM) are geographically close to each other
123 (less than 2 kilometers) and situated in the 19th and 20th districts northeast of Paris. The
124 remaining two populations are from the park of Pierre-Fitte in Villeneuve-le Roi city (VLR 1
125 and VLR2), approximately 16 km away from Paris (Table 1). The population from Buttes-
126 Chaumont Park (BCH) was sampled over two consecutive years (2015 and 2016). Genomic
127 DNA was extracted using the same methodology as previously described. Tests of
128 amplification patterns were carried out on a panel of 8 individuals from these six populations.
129 After basic PCR amplifications, PCR products were checked using gel electrophoresis in 2%
130 agarose gels. Of the 40 primer pairs, 28 had ambiguous banding patterns. The 12 markers that
131 proved to be monolocus were analysed separately using labelled forward primers on an ABI
132 3130 xl Genetic Analyzer (Applied Biosystems, genomic platform of IMRB, Mondor
133 Institute, Créteil). Fragment lengths were manually checked on chromatograms to detect
134 inconsistencies and genotypes were scored against the GeneScan-500 Liz Size Standard
135 (Applied Biosystems) using GeneMapper 5 software (Applied Biosystems).

136 Of the 12 primer pairs, 4 were discarded because of ambiguous banding patterns or
137 monomorphism. The eight microsatellite markers (Table 2 and Table S1 of the supplementary
138 data) showing both consistent amplification results and genetic variation among tested

139 individuals were used to develop one multiplex set to be carried out in 12.5 μ l reactions using
140 10 ng of DNA and the Qiagen $\text{\textcircled{R}}$ Multiplex Kit according to the manufacturer's protocol.

141 In order to quantify genotyping errors, 23 individuals of the VLR2 population were
142 genotyped twice (genotypes were produced from two different DNA extracts from the same
143 individuals). Error rates were quantified by computing the mean error rate per allele (e_a) and
144 the mean error rate per locus (e_l , Pompanon et al., 2005). Moreover, null allele frequencies
145 were estimated using the software FreeNa (Chapuis and Estoup, 2007) with a number of
146 replicates fixed to 1000.

147 From the entire data set (153 individuals), the number of alleles, the observed and
148 expected heterozygosities and the estimator of the inbreeding coefficient (F_{is}) were calculated
149 using the program Genetix V. 4.05 (Belkhir et al., 2004). The allelic richness standardized for
150 sample size (A_r , $N = 14$) was computed using the program Fstat V. 2.9.3.2 (Goudet, 2000).
151 Exact test for genotypic disequilibrium and deviations from Hardy-Weinberg equilibrium
152 were computed using Genepop V. 4.4 (Rousset, 2008). In order to investigate the importance
153 of genetic exchanges among populations, exact tests of allelic differentiation were carried out
154 using Genepop V. 4.4. To adjust for multiple comparisons, sequential Bonferroni correction
155 was used. Moreover, Weir and Cockerham's (1984) estimator of the fixation index F_{st} was
156 calculated with Genepop V. 4.4. The program BAPS V. 6 (Corander and Marttinen, 2006;
157 Corander et al., 2008) was used to carry out an analysis of spatial clustering of groups. We ran
158 10 replicates for $k = 6$ where k is the maximum number of genetically divergent groups
159 (populations).

160 The 658 bp fragment of the COI gene proposed as a standard DNA barcode for
161 animals (Hebert et al., 2003) was sequenced using LCO1490F and HCO2192R primers
162 (Folmer et al., 1994) for six individuals used for the genomic library (DNA of 3 individuals
163 could not be amplified) and for a sub-sample of each study population (54 individuals, details

164 in Table 1, GenBank Accession Numbers MH684623 – MH684682). DNA sequencing was
165 performed in both directions by Eurofins Genomics company and sequences were manually
166 aligned using the BioEdit program (Hall, 1999). Taxonomy of the samples was checked using
167 the identification engine of BOLD (Barcode of Life Data Systems –
168 <https://www.boldsystems.org/>). Using the software MEGA6, we constructed a phylogenetic
169 tree by neighbour joining with p-distance and 1000 bootstraps. To root the tree, we used a
170 sequence of *Lumbricus herculeus* (GenBank Accession Number MH638308). Moreover, as
171 Porco et al. (2018) revealed the existence of three lineages within the morphospecies *L.*
172 *castaneus*, one sequence per lineage was added to the dataset (GenBank Accession Numbers
173 FJ937284, MF121706, GU206163). The neighbor-joining phylogenetic tree revealed that all
174 the study specimens belong to the L1 lineage (Fig S1, supplementary data). Cross-
175 amplification of the new *L. castaneus* microsatellites across mitochondrial lineages could thus
176 not be tested.

177

178 **3. Results and Discussion**

179 In total eight polymorphic microsatellite markers were developed for *L. castaneus*.
180 None of these loci showed significant linkage disequilibrium. Over the whole data set (N =
181 153), allele number per locus ranged from 7 to 29 (Table 2). Null alleles are likely to occur at
182 all loci but not at the same frequency according to the population. Null allele frequencies
183 ranged from 0 to 0.293 (Table 2). The presence of null alleles in population genetics studies
184 using microsatellite markers have been frequently reported for a wide range of taxa (Dakin
185 and Avise, 2004), and in earthworms in particular (Dupont et al., 2011; Harper et al., 2006;
186 Novo et al., 2008; Souleman et al., 2016; Velavan et al., 2007). Several statistical corrections
187 for null allele are available in population genetics software to overcome this problem (e.g.
188 Chapuis and Estoup, 2007). Apart from the risk of null alleles, the reliability of the new

189 microsatellite markers was good. Indeed, the risk of genotyping error was null for the loci
190 LC02, LC05, LC10, LC27 and LC36. The higher values of mean error rate per allele and
191 mean error rate per locus were obtained for the locus LC33 (0.05 and 0.10 respectively).

192 We further evaluated the suitability of the use of these microsatellites for genetic
193 studies of populations exposed to environmental stressors by studying the genetic diversity
194 and the genetic structure of the six populations we sampled in the Paris region. Summary
195 statistics describing the genetic variation within populations are given Table 1. Significant
196 departure from Hardy-Weinberg equilibrium caused by a deficiency in heterozygotes was
197 observed in all sites with F_{is} estimator ranging from 0.132 to 0.366 (Table 1). This
198 heterozygote deficit could be due to null alleles, but also to a Wahlund effect (i.e. cryptic
199 population structure within site) or to inbreeding. Additional population genetic studies within
200 each site are necessary to further interpret these results. The allelic richness and the expected
201 heterozygosity were found to vary slightly depending on the population (from 6.06 to 7.70
202 and 0.596 to 0.705, respectively). The highest values of the genetic diversity estimators were
203 observed in both populations of Villeneuve-le-Roi (VLR1 and VLR2) which were sampled in
204 an urban wasteland of 8 ha. The sample from the Compans square (COM) was the population
205 with the lowest expected heterozygosity and the second lowest allelic richness (Table 1).
206 Because of the spatial isolation and the restricted size of the Compans square (2444 m²), we
207 suggest that the low level of genetic diversity observed in COM may be attributed to
208 important genetic drift effects due to small effective population size.

209 Among populations, a low but significant genetic structure was revealed at the level of
210 the whole study ($F_{st} = 0.033$, $P < 0.001$). As expected, the BCH15 and BCH16 samples,
211 collected at exactly the same locality of the Buttes-Chaumont Park but over two consecutive
212 years, were not genetically differentiated and, thus, are indeed from the same population
213 (Table 3). The highest pairwise F_{st} values were obtained between the COM population and all

214 the other populations (Table 3). This important genetic differentiation of the COM population
215 confirms the suspected founder effect indicated by the low level of genetic diversity observed,
216 and is probably the result of genetic drift. This singularity was confirmed by the spatial
217 clustering of groups that revealed 3 clearly differentiated genetic clusters with one of the
218 cluster corresponding only to COM (Fig. 1). The populations of the Buttes-Chaumont and
219 Butte-du-Chapeau-Rouge Parks, both situated in the 19th district northeast of Paris, clustered
220 as expected with respect to their geographical proximity (Fig. 1). For the same reason, the
221 cluster found for both populations from Villeneuve-le-Roi (VLR1 and VLR2) was predictable
222 but the fact that the Belleville population (BEL) also belongs to this third cluster was
223 surprising. This unexpected pattern of genetic structure highlights the probable role of human
224 transfer of earthworms between sites as had been already shown for instance for the endogeic
225 species *Allolobophora chlorotica* and *Aporrectodea icterica* (Dupont et al., 2015; Dupont et
226 al., 2017; Torres-Leguizamon et al., 2014).

227 To conclude, the eight newly developed microsatellite markers for *L. castaneus* were
228 highly polymorphic and allowed to differentiate populations even at a fine scale (< 2 km) in
229 an urban context. Therefore, we believe that they will be of great utility to investigate the
230 consequences of environmental stressors, such as trace elements contamination, on neutral
231 genetic variation of *L. castaneus* populations.

232 **Acknowledgements**

233 This work was funded by the French National Research Agency (ANR) as a part of the
234 project Ecoville No. ANR-14-CE22-0021, by the town council of Paris through the Paris2030
235 calls for projects (Soil biodiversity in green spaces of Paris project) and by the Ile de France
236 region through the “Partenariats institutions-citoyens pour la recherche et l’innovation” (Picri)
237 calls for projects (ReFUJ project). The development of the microsatellite-enriched genomic
238 library was funded by the “Institut de Recherches Pluridisciplinaires en Sciences de
239 l’Environnement” (IREPSE, Lille 1 University).

240

241 **References**

- 242 Belkhir, K., Borsa, P., Goudet, J., Chikhi, L., Bonhomme, F., 2004. GENETIX 4.05, logiciel sous Windows
243 pour la génétique des populations. Laboratoire Génome, Population, Interactions, CNRS UMR 5000,
244 Université Montpellier II, Montpellier (France).
- 245 Bouché, M.B., 1972. Lombriciens de France. Ecologie et systématique. INRA, Paris.
- 246 Chapuis, M.P., Estoup, A., 2007. Microsatellite null alleles and estimation of population
247 differentiation. *Mol. Biol. Evol.* 24, 621-631.
- 248 Corander, J., Marttinen, P., 2006. Bayesian identification of admixture events using multilocus
249 molecular markers. *Mol. Ecol.* 15, 2833-2843.
- 250 Corander, J., Sirén, J., Arjas, E., 2008. Bayesian spatial modelling of genetic population structure.
251 *CompStat* 23, 111-129.
- 252 Cunha, L., Thornber, A., Kille, P., Morgan, A.J., Novo, M., 2017. A large set of microsatellites for the
253 highly invasive earthworm *Amyntas corticis* predicted from low coverage genomes. *Appl. Soil Ecol.*
254 119, 152-155.
- 255 Dakin, E.E., Avise, J.C., 2004. Microsatellite null alleles in parentage analysis. *Heredity* 93, 504-509.
- 256 Dupont, L., Grésille, Y., Richard, B., Decaëns, T., Mathieu, J., 2015. Fine-scale spatial genetic structure
257 and dispersal constraints in two earthworm species. *Biol. J. Linn. Soc.* 114, 335-347.
- 258 Dupont, L., Lazrek, F., Porco, D., King, R.A., Rougerie, R., Symondson, W.O.C., Livet, A., Richard, B.,
259 Decaens, T., Butt, K.R., Mathieu, J., 2011. New insight into the genetic structure of the *Allolobophora*
260 *chlorotica* aggregate in Europe using microsatellite and mitochondrial data. *Pedobiologia* 54, 217-
261 224.
- 262 Dupont, L., Torres-Leguizamon, M., Rene-Corail, P., Mathieu, J., 2017. Landscape features impact
263 connectivity between soil populations: a comparative study of gene flow in earthworms. *Mol. Ecol.*
264 26, 3128-3140.
- 265 Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of
266 mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Bio. &*
267 *Biotech.* 3, 294-299.
- 268 Foti, L., Dubs, F., Gignoux, J., Lata, J.-C., Lerch, T.Z., Mathieu, J., Nold, F., Nunan, N., Raynaud, X.,
269 Abbadie, L., Barot, S., 2017. Trace element concentrations along a gradient of urban pressure in
270 forest and lawn soils of the Paris region (France). *Sci. Total Environ.* 598, 938-948.
- 271 Gonzalez, E.G., Zardoya, R., 2007. Relative role of life-history traits and historical factors in shaping
272 genetic population structure of sardines (*Sardina pilchardus*). *BMC Evol. Biol.* 7, 197.
- 273 Goudet, J., 2000. FSTAT, a program to estimate and test gene diversities and fixation indices, version
274 2.9.1 ed. Available from <https://www2.unil.ch/popgen/softwares/fstat.htm>. (last access date
275 06/11/18)
- 276 Grumiaux, F., Demuyck, S., Pernin, C., Lepretre, A., 2015. Earthworm populations of highly metal-
277 contaminated soils restored by fly ash-aided phytostabilisation. *Ecotox. Environ. Safe.* 113, 183-190.
- 278 Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program
279 for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41, 95-98.
- 280 Harper, G.L., Cesarini, S., Casey, S.P., Morgan, A.J., Kille, P., Bruford, M.W., 2006. Microsatellite
281 markers for the earthworm *Lumbricus rubellus*. *Mol. Ecol. Notes* 6, 325-327.
- 282 Hartl, D.L., Clark, A.G., 2007. Principles of population genetics, 4 ed. Sinauer, Sunderland, MA.
- 283 Hebert, P.D.N., Cywinska, A., Ball, S.L., Dewaard, J.R., 2003. Biological identifications through DNA
284 barcodes. *P. Roy. Soc. B. - Biol. Sci.* 270, 313-321.
- 285 Johnson, M.T.J., Munshi-South, J., 2017. Evolution of life in urban environments. *Science* 358.
- 286 Kirk, H., Freeland, J.R., 2011. Applications and Implications of Neutral versus Non-neutral Markers in
287 Molecular Ecology. *Int. J. Mol. Sci.* 12, 3966-3988.
- 288 Loveless, M.D., Hamrick, J.L., 1984. Ecological determinants of genetic structure in plant populations
289 *Annu. Rev. Ecol. Syst.* 15, 65-95.
- 290 Macnair, M.R., 1987. Heavy metal tolerance in plants: a model evolutionary system. *Trends Ecol.*
291 *Evol.* 2, 354-359.

292 Malausa, T., Gilles, A., Meglecz, E., Blanquart, H., Duthoy, S., Costedoat, C., Dubut, V., Pech, N.,
293 Castagnone-Sereno, P., Délye, C., Feau, N., Frey, P., Gauthier, P., Guillemau, T., Hazard, L., Le Corre,
294 V., Lung-Escarmant, B., Malé, P.-J.G., Ferreira, S., Martin, J.F., 2011. High-throughput microsatellite
295 isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries. *Mol. Ecol. Resour.*
296 *11*, 638-644.

297 Mengoni, A., Gonnelli, C., Galardi, F., Gabbrielli, R., Bazzicalupo, M., 2000. Genetic diversity and
298 heavy metal tolerance in populations of *Silene paradoxa* L. (Caryophyllaceae): a random amplified
299 polymorphic DNA analysis. *Mol. Ecol.* *9*, 1319-1324.

300 Mussali-Galante, P., Tovar-Sanchez, E., Valverde, M., Rojas, E., 2014. Genetic structure and diversity
301 of animal populations exposed to metal pollution, in: Whitacre, D. (Ed.), *Reviews of environmental*
302 *contamination and toxicology*. Springer International Publishing Switzerland, pp. 79-106.

303 Novo, M., Velavan, T.P., Almodovar, A., Schulenburg, H., Diaz-Cosin, D.J., Michiels, N.K., 2008.
304 Microsatellite markers for the drought-resistant earthworm *Hormogaster elisae*. *Mol. Ecol. Resour.* *8*,
305 901-903.

306 Pickett, S.T.A., Cadenasso, M.L., Grove, J.M., Boone, C.G., Groffman, P.M., Irwin, E., Kaushal, S.S.,
307 Marshall, V., McGrath, B.P., Nilon, C.H., Pouyat, R.V., Szlavecz, K., Troy, A., Warren, P., 2011. Urban
308 ecological systems: Scientific foundations and a decade of progress. *J. Environ. Manage.* *92*, 331-362.

309 Pompanon, F., Bonin, A., Bellemain, E., Taberlet, P., 2005. Genotyping errors: Causes, consequences
310 and solutions. *Nat. Rev. Genet.* *6*, 847-859.

311 Porco, D., Chang, C.-H., Dupont, L., James, S., Richard, B., Decaëns, T., 2018. A reference library of
312 DNA barcodes for the earthworms from Upper Normandy: Biodiversity assessment, new records,
313 potential cases of cryptic diversity and ongoing speciation. *Appl. Soil. Ecol.* *124*, 362-371.

314 Rousset, F., 2008. GENEPOP ' 007: a complete re-implementation of the GENEPOP software for
315 Windows and Linux. *Mol. Ecol. Resour.* *8*, 103-106.

316 Somers, C.M., Neudorf, K., Jones, K.L., Lance, S.L., 2011. Novel microsatellite loci for the compost
317 earthworm *Eisenia fetida*: A genetic comparison of three North American vermiculture stocks.
318 *Pedobiologia* *54*, 111-117.

319 Souleman, D., Grumiaux, F., Frerot, H., Vandenbulcke, F., Pauwels, M., 2016. Isolation and
320 characterization of eight polymorphic microsatellites markers for the earthworm *Lumbricus*
321 *terrestris*. *Eur. J. Soil Biol.* *74*, 76-80.

322 Spurgeon, D.J., Hopkin, S.P., 1996. The effects of metal contamination on earthworm populations
323 around a smelting works: Quantifying species effects. *Appl. Soil. Ecol.* *4*, 147-160.

324 Strunk, H., Hochkirch, A., Veith, M., Hankeln, T., Emmerling, C., 2012. Isolation and characterization
325 of eleven polymorphic microsatellite markers for the earthworm *Aporrectodea longa* (Ude). *Eur. J.*
326 *Soil Biol.* *48*, 56-58.

327 Takami, Y., Koshio, C., Ishii, M., Fujii, H., Hidaka, T., Shimizu, I., 2004. Genetic diversity and structure
328 of urban populations of Pieris butterflies assessed using amplified fragment length polymorphism.
329 *Mol. Ecol.* *13*, 245-258.

330 Torres-Leguizamon, M., Mathieu, J., Decaëns, T., Dupont, L., 2014. Genetic structure of earthworm
331 populations at a regional scale: inferences from mitochondrial and microsatellite molecular markers
332 in *Aporrectodea icterica* (Savigny 1826). *Plos One* *9*, e101597.

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

335 van Straalen, N.M., Timmermans, M.J.T.N., 2002. Genetic variation in toxicant-stressed populations:
336 An evaluation of the "genetic erosion" hypothesis. *Hum Ecol Risk Assess* *8*, 983-1002.

337 Velavan, T.P., Schulenburg, H., Michiels, N., 2007. Development and characterization of novel
338 microsatellite markers for the common earthworm (*Lumbricus terrestris* L.). *Mol. Ecol. Notes* *7*, 1060-
339 1062.

340 Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure.
341 *Evolution* *38*, 1358-1370.

342 Zaborski, E.R., 2003. Allyl isothiocyanate: an alternative chemical expellant for sampling earthworms.
343 *Appl. Soil Ecol.* *22*, 87-95.

344 **Figure legends**

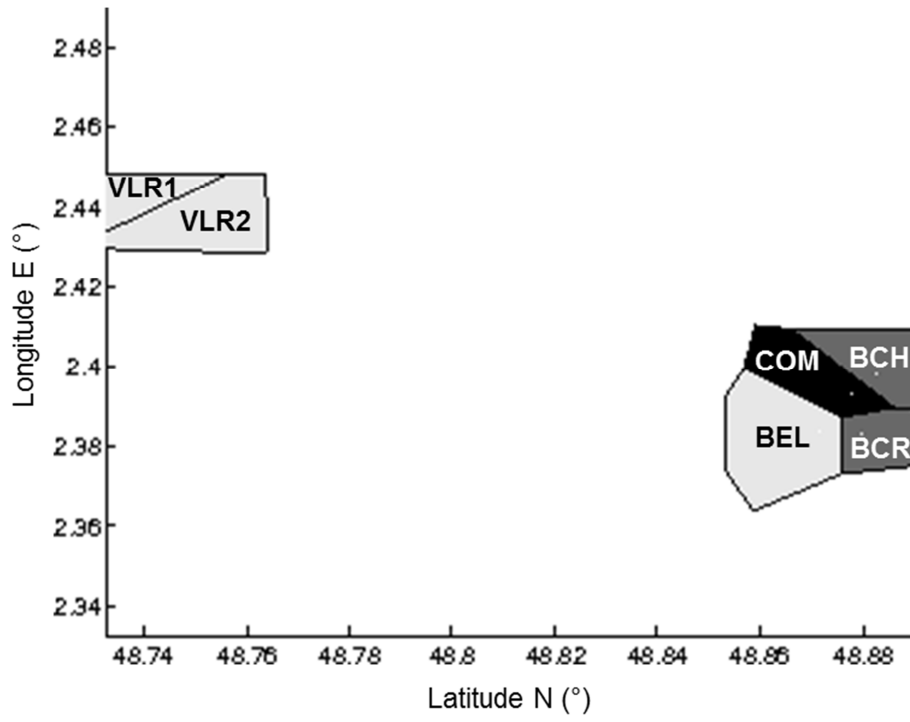
345 **Figure 1**

346 Spatial clustering analysis using a Voronoi tessellation in BAPS. The optimal partition value

347 was for $K = 3$. Different greyscales represent the three genetic clusters. Each cell of the

348 tessellation corresponds to the physical neighbourhood of an observed data point.

349



350

351

352 **Table 1:** Characteristics of the six study samples, with geographical location (locality, latitude and longitude), year of collection (date), sample
353 size (N), number of COI sequences (N_{seq}), allelic richness after rarefaction to a population size of 14 (A_r), observed heterozygosity (H_o), expected
354 heterozygosity (H_e), and fixation index F_{is} which was always associated to a highly significant probability of the exact test for deviations from
355 Hardy Weinberg expectations (*).

Locality	Latitude N (°)	Longitude E (°)	Sample	Date	N	N_{seq}	A_r	H_o	H_e	F_{is}
Buttes-Chaumont park- Paris	48.879978	2.383333	BCH15	2015	19	19	6.06	0.526	0.603	0.132*
			BCH16	2016	25	0	6.64	0.487	0.651	0.257*
Belleville park- Paris	48.871901	2.383527	BEL	2016	26	2	7.26	0.463	0.643	0.271*
Butte du chapeau rouge square - Paris	48.882608	2.398003	BCR	2016	24	2	7.05	0.466	0.646	0.284*
Compans square - Paris	48.877923	2.393220	COM	2016	18	2	6.13	0.424	0.596	0.298*
Site 1 – Pierre-Fitte –Villeneuve - le - Roi	48.7399292	2.4391225	VLR1	2015	18	18	7.56	0.450	0.705	0.366*
Site 2 – Pierre-Fitte –Villeneuve - le - Roi	48.7406666	2.4379152	VLR2	2016	23	11	7.70	0.570	0.676	0.160*

356

357

358 **Table 2:** Characteristics of eight microsatellite loci isolated from *Lumbricus castaneus*, with locus name, repeat array, primer sequences (F:
359 forward primer, R: reverse primer), allele size range (ASR), total number of allele over the whole data set (N_{tall}), range of the estimation of null
360 allele frequency in the study samples (Null), mean error rate per allele (e_a) and mean error rate per locus (e_l).

Locus ID	Repeat array	Primer sequence (5' – 3')	ASR (bp)	N_{tall}	Null	e_a	e_l
LC02	(TC) ₁₀	F : AT565-GTTATCGGTGCCTTCCATGT R : GTGACCATCCTTTGATTGCC	093 - 115	7	0.000 – 0.195	0.00	0.00
LC05	(GT) ₈	F : FAM-ATACGTGTCTGGAAGGGTGG R : CTCCGTTGTACCGCTGTGTA	120 - 170	7	0.000 – 0.135	0.00	0.00
LC10	(AATC) ₁₀	F : AT550-GAAGCCATACTGCCACTGGT R : GCTTTACTTGTTGCCATTCTGTT	141 - 208	19	0.029 – 0.208	0.00	0.00
LC16	(GACA) ₈	F : AT565-AACAACCGAAAACCTGCAAGTC R : CCTCAAGGCAAGCTCAGGTA	225 - 344	29	0.053 – 0.293	0.01	0.05
LC18	(TCAA) ₁₁	F : FAM-GAGGCCAAAACCCATCACTA R : TGTCTTTCAGGGCAGAAGTG	223 - 286	22	0.021 – 0.253	0.01	0.05
LC27	(CT) ₅	F : AT550-GCCATTTTGTCAATTCGGTCT R : AACATAACGCAACACCCACA	237 - 266	13	0.000 – 0.139	0.00	0.00
LC33	(TTA) ₇	F : YAKIMAYELLOW-CTCAGAATTGACGTTGGCAG R : ATCATGCATGGACAGCGTT	110 - 132	10	0.063 – 0.290	0.05	0.10
LC36	(GAAT) ₆	F : YAKIMAYELLOW-GACGTAACGCAATGTGATGG R : TCCAGGCAGAGTTATTTCCGG	151 - 215	19	0.000 – 0.264	0.00	0.00

361

362 **Table 3** Pairwise multi-loci F_{ST} estimates between different populations. Significant exact
 363 tests of allelic differentiation after sequential Bonferroni correction are indicated in bold.

BCH16	0.004					
BEL	0.054	0.037				
BCR	0.025	0.018	0.014			
COM	0.069	0.069	0.050	0.040		
VLR1	0.030	0.021	0.016	0.007	0.039	
VLR2	0.048	0.036	0.031	0.032	0.072	0.004
	BCH15	BCH16	BEL	BCR	COM	VLR1

364

365