Impacts of metallic trace elements on an earthworm community in an urban wasteland: Emphasis on the bioaccumulation and genetic characteristics in Lumbricus castaneus

Helene Audusseau, Franck Vandenbulcke, Cassandre Dume, Valentin Deschins, Maxime Pauwels, Agnès Gigon, Matthieu Bagard, Lise Dupont

Impacts of metallic trace elements on an earthworm community in an urban wasteland: emphasis on the bioaccumulation and genetic characteristics in *Lumbricus castaneus*

Hélène Audusseau [1,2,3*], Franck Vandenbulcke [4], Cassandre Dume [1,4], Valentin Deschins [1], Maxime Pauwels [5], Agnès Gigon [1], Matthieu Bagard [1] and Lise Dupont [1]

1. Univ. Paris Est Creteil, Sorbonne Université, CNRS, INRA, IRD, Institut d'écologie et des sciences de l'environnement de Paris, 94010 Créteil, France
2. Department of Zoology, Stockholm University, Stockholm, Sweden
3. UK Centre for Ecology & Hydrology Maclean Building, Benson Lane, Wallingford, Oxon, OX10 8BB, UK.
4. Université de Lille, EA 4515-LGCgE - Laboratoire Génie Civil et géo-Environnement, Cité scientifique, SN3, F-59655 Villeneuve d'Ascq, France
5. Université de Lille, CNRS, UMR 8198 – Unité Evolution-Ecologie-Paléontologie, F-59000 Lille, France

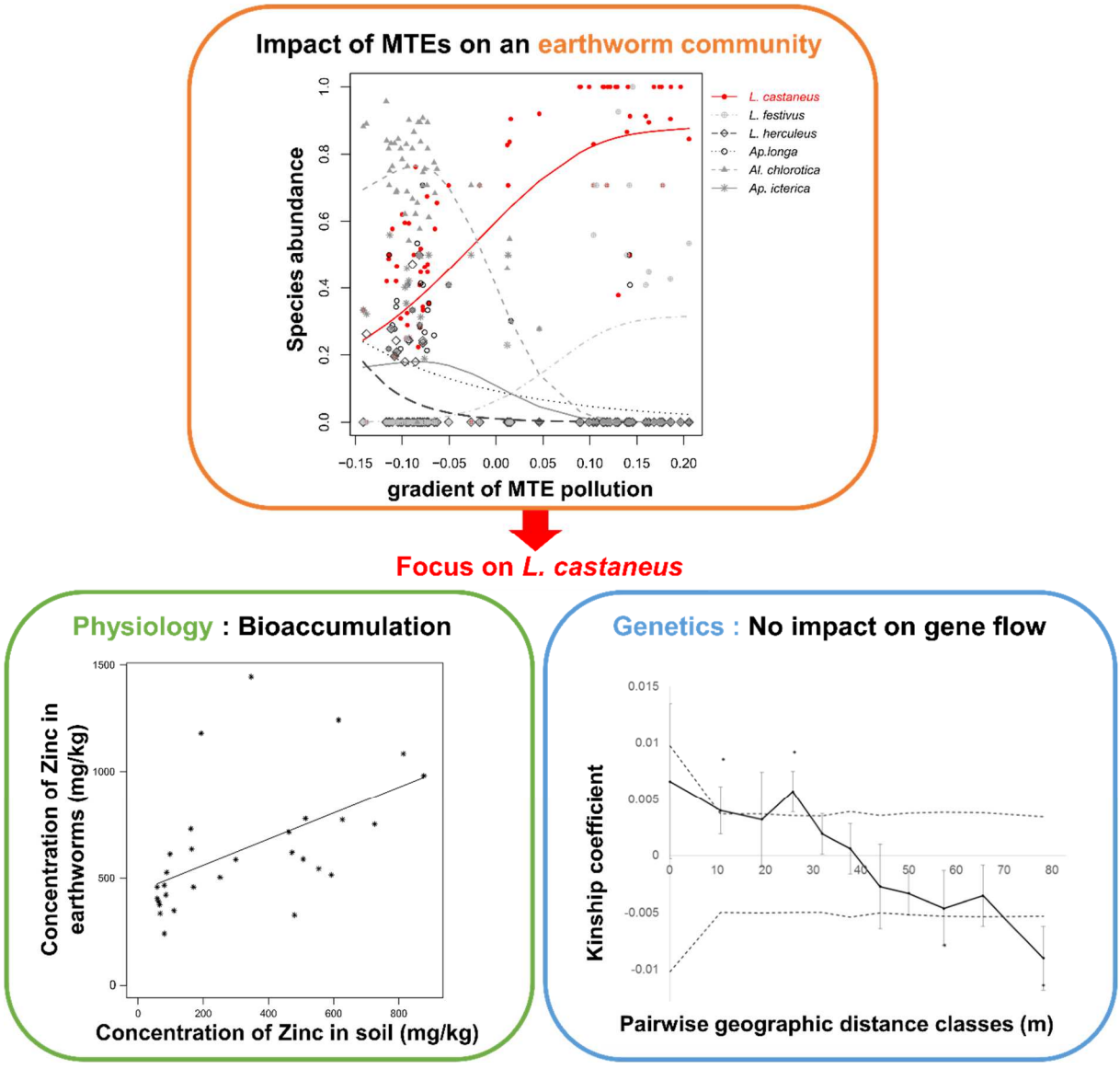
* Corresponding author; e-mail address: helene.audusseau@zoologi.su.se

Highlights

- Impact of soil pollution on earthworms at different levels of biological organisation.
- Differential effects of MTEs on earthworms were investigated in an urban wasteland.
- Community structure and *L. castaneus* physiology and genetics were studied.
- MTEs affected earthworm community and bioaccumulation, but not population genetic.
- *L. castaneus* is a promising model to study the molecular basis of MTE tolerance.

25 **Graphical abstract**

26



27

Abstract

Metallic trace elements (MTEs) soil pollution has become a worldwide concern, particularly regarding its impact on earthworms. Earthworms, which constitute the dominant taxon of soil macrofauna in temperate regions and are crucial ecosystem engineers, are in direct contact with MTEs. The impacts of MTE exposure on earthworms, however, vary by species, with some able to cope with high levels of contamination. We combined different approaches to study the effects of MTEs at different levels of biological organisation of an earthworm community, in a contaminated urban wasteland. Our work is based on field collection of soil and earthworm samples, with a total of 891 adult earthworms from 8 species collected, over 87 quadrats across the study plot. We found that MTE concentrations are highly structured at the plot scale and that some elements, such as Pb, Zn, and Cu, are highly correlated. Comparing species assemblage to MTE concentrations, we found that the juvenile and adult abundances, and community composition, were significantly affected by pollution. Along the pollution gradient, as species richness decreased, *Lumbricus castaneus* became more dominant. We thus investigated the physiological response of this species to a set of specific elements (Pb, Zn, Cu, and Cd) and studied the impacts of MTE concentrations at the plot scale on its population genetic. These analyses revealed that *L. castaneus* is able to bioaccumulate high quantities of Cd and Zn, but not of Cu and Pb. The population genetic analysis, based on the genotyping of 175 individuals using 8 microsatellite markers, provided no evidence of the role of the heterogeneity in MTE concentrations as a barrier to gene flow. The multidisciplinary approach we used enabled us to reveal the comparatively high tolerance of *L. castaneus* to MTE concentrations, suggesting that this is a promising model to study the molecular bases of MTE tolerance.

51 **Keywords**

52 Soil contamination, MTEs, community structure, population genetics.

1. Introduction

Metallic trace elements (MTEs) occur naturally in the earth's crust, but increasing quantities of metals are being released into the environment by human activities. Soil pollution by MTEs has become a global concern (Hou et al., 2017; Weissmannova & Pavlovsky 2017; Rodriguez-Eugenio et al., 2018), particularly in urban areas, as urban soil appears to be more contaminated than agricultural and natural soils (Ajmone-Marsan & Biasioli 2010). In urban soils, anthropogenic sources of MTEs include traffic emissions, industrial discharges and municipal wastes (McLlwaine et al., 2017; Weissmannova & Pavlovsky 2017; Jia et al., 2018). While the most common hazardous MTEs in soils (namely arsenic As, cadmium Cd, chromium Cr, copper Cu, mercury Hg, nickel Ni, lead Pb and zinc Zn) can be non-degradable, persistent, and bioaccumulate and biomagnify in food chains, the specific impact and toxicity of the different MTEs are dictated by their chemical forms (Knox et al., 2000; Weissmannova & Pavlovsky 2017). MTEs tend to accumulate in soil, and sometimes in food webs, representing, beyond certain concentrations and durations of exposure, a significant risk to the health of living organisms, including humans (Tyler et al., 1989).

Earthworms constitute a dominant taxon of soil macrofauna, and their activities of recycling organic matter and modifying soil structure are crucial to the functioning of the soil ecosystem (Blouin et al., 2013). These keystone species are affected by MTEs present in the soil as they are in direct contact with the bioavailable contaminants through the soil porewater (van Gestel et al., 2009). At the community level, fieldworks showed that earthworm biomass and species richness are inversely correlated to metal concentrations in the soils, and especially to Pb, Zn, and Cd, contents (Terhivuo et al., 1994; Spurgeon & Hopkin 1996, 1999; Nahmani et al., 2003; Leveque et al., 2015; Wang et al., 2018b). At the species level, studies have aimed at evaluating metal toxicity, either focusing on their effect on fitness or assessing

MTE lethal and sub lethal concentrations (reviewed in Nahmani et al., 2007). However, there are great differences among species in their sensitivity to MTEs. Some species of earthworms, such as *Lumbricus rubellus*, *L. castaneus*, and *L. terrestris*, were shown to persist even in highly contaminated sites (Spurgeon & Hopkin 1996) and this persistence can often be linked with differences in physiological abilities. For instance, these species have the ability to protect themselves from the toxic effects of metals by sequestering, detoxifying, and storing excess metal (Spurgeon & Hopkin 1996; Vijver et al., 2004; Iordache & Borza 2012; Grumiaux et al., 2015). This protection involves the induction of a gene coding a metal sequestering metallothionein (Sturzenbaum et al., 1998; Brulle et al., 2006), allowing species to tolerate high concentration of MTEs through bioaccumulation processes. In particular, cadmium and zinc are MTEs that have been shown to bioaccumulate in numerous earthworms (van Straalen et al., 2001; Tischer 2009), among which *L. castaneus* (Tischer 2009). *L. castaneus* also accumulated higher quantities of Pb and at higher pace than *L. rubellus*, *Aporrectodea caliginosa* and *A. rosea,* in a laboratory rearing experiment (Terhivuo et al., 1994).

At the population level, MTEs are likely to induce microevolutionary processes through (i) mutations and increased allelic diversity, (ii) emphasis of the effects of genetic drift and bottlenecks and (iii) natural selection, leading to the disappearance of the most sensitive genotypes (Ribeiro & Lopes 2013). In that respect, population genetic approaches using neutral molecular markers (‘neutral’ refers to a locus that has no effect on fitness, Holderegger et al., 2006) are commonly used to infer microevolutionary processes such as mutation, genetic drift and gene flow (Kirk & Freeland 2011).

Since the responses of individual species and communities are highly dependent on the soil physicochemical properties and the MTE cocktail present locally, the results found in the literature are not always consistent (see the contrasted responses of earthworm communities

to Pb concentration from van Gestel et al., 2009 and Leveque et al., 2015). Here, we aimed at providing a comprehensive study of the effects of MTEs on different levels of the biological organisation of a community of earthworm species in a polluted site. Our approach was first exploratory and aimed at studying in their entirety the links between the concentrations of MTEs found on the study site and the earthworm community (abundance and richness). Second, we focused on the specific effect of Pb, Zn, Cd, and Cu, on *L. castaneus.* Pb, Zn, Cd, and Cu, are known to be the most widespread anthropogenic contaminant elements in urban soils (Argyraki et al., 2018), with Pb being one of the major concern in many cities (Ajmone-Marsan and Biasioli 2010). The species, *L. castaneus,* was selected as it was found throughout the site, which makes it possible to study variations in the response of this species to a range of MTE concentrations. Specifically, we hypothesized that *L. castaneus* would show increased levels of bioaccumulation with increased concentrations of these MTEs. We also investigated the neutral genetic variation of this species using 8 microsatellite markers in order to determine how mutation, genetic drift, and gene flow, affect the genetic characteristics of earthworm populations in a contaminated urban soil. To our knowledge, this is the first field study that combines different approaches to assess the responses of earthworms to metallic trace elements in the field.

120 ## 2. Material and methods

121 ### 2.1 Study area

122 The study was carried out in an urban wasteland of 8 ha situated east of the city of
123 Villeneuve-Le-Roi, in the region of Paris (2°26'16.7"E 48°44'26.8"N, France). The area is
124 located inside a loop of the Seine River and near an industrial zone. In 2013, a quantitative
125 evaluation of sanitary risks was carried out at the request of the city, which wanted to convert
126 the site and dedicate it to urban agriculture. The report concluded that an area of 3 ha north of
127 the site presented high levels of 5 MTEs (Cd, Cu, Hg, Pb, and Zn); this polluted zone was
128 then separated from the unpolluted zone by a fence (Guittard 2013). Further, the agro-
129 pedological report by Sol Paysage (2013) defined the soil from the polluted zone as mainly
130 composed of coarse sand with a C/N ratio of 14.5, a cation exchange capacity (CEC, 0-25 cm
131 depth) of 18.7 meq/100g, and a pH of 8.2. The soil from the unpolluted zone is defined as
132 loamy with a C/N ratio of 9.6, a CEC (0-25 cm depth) of 10.1 meq/100g, and a pH of 7.7 (Sol
133 Paysage 2013).

134 The study plot covered a surface of 50 m x 60 m that straddles in approximately equal
135 proportions the zones initially identified as polluted and unpolluted. Based on previous work
136 on related species, this surface area is appropriate to study the fine-scale population genetic
137 structure of earthworms in order to infer microevolutionary processes at the intra-population
138 scale (Novo et al., 2010; Dupont et al., 2015).

139 ### 2.2 Earthworms collection

140 Earthworms were collected over two years. The first sampling aimed to collect
141 specimens for the studies of community structure and population genetics. It was done over
142 three consecutive days in 2016 (March 29th, 30th, and April 1st). Earthworms were collected

143 from 87 quadrats (50 cm x 50 cm), chosen following a stratified sampling protocol across the
144 study plot (Fig. 1a). 52 quadrats were sampled in the polluted zone and 35 in the unpolluted
145 zone. We sampled more densely in the polluted zone, in order to better capture the variations
146 in soil pollution, than in the unpolluted zone, where the soil is assumed to be homogeneous.
147 Note that the resolution of our sampling is likely to have captured the heterogeneity of soil
148 pollution, which is at a smaller spatial scale than the 50 cm x 50 cm quadrat of our sampling.

149 Over each quadrat, 10L of AITC (allyl isothyocyanate) and isopropanol diluted in
150 water (1:100:10000L) were poured in two stages. All earthworms were collected as they came
151 out of the soil until no more individuals came out (we waited up to 15 minutes for each
152 quadrat). They were first transferred to a solution of 10% dilution alcohol and water, and then
153 stored in 100% alcohol before taxonomic identification in the laboratory and analyses. The
154 number of adults of each species as well as the number of juveniles were counted per quadrat.
155 Earthworms were identified using the taxonomic keys of Bouché (1972) and Sims and Gerard
156 (1999). Further, the identification of a subset of 55 individuals, for which either the taxonomic
157 identification was uncertain or belonging to species rarely occurring within our sampling, was
158 confronted with the results of the barcode identification, carried out using a fragment the
159 cytochrome c oxidase subunit I (COI) mitochondrial gene (Genbank accession number
160 MN519732 - MN519786, Hebert et al., 2003; Dupont et al., 2019).

161 For the second sampling, carried out the 6th of April 2017, only individuals from *L.*
162 *castaneus* were collected in order to study, in this species, the bioaccumulation of a subset of
163 MTEs. 136 *L. castaneus* individuals were sampled from 30 out of the 87 quadrats. These
164 quadrats were selected to reflect a lead (Pb) concentration gradient (estimated on the basis of
165 2016 soil data, Fig. 1a).

166 **2.3 Soil collection and measure of MTE pollution**

167 Soil samples were taken simultaneously with the first earthworm sampling (in 2016) to
168 measure pH and metal pollution of soils. For each quadrat, three soil samples were collected
169 on three of the four sides of the quadrat, and pooled prior to be analyzed. In the laboratory,
170 soil samples were dried at 40°C for 24 hours, grounded, and sieved to 2 mm. Soil pH was
171 measured in 1 M KCl and in distilled H_2O according to the ISO 10390:2004 standard. The
172 concentrations of the different elements in the soils, including the MTEs, were measured by
173 X-ray Fluorescence (XRF) using the Epsilon 3XL panalytical and were analyzed with the
174 software Omnian. As soil moisture is known to significantly affect XRF-measurements
175 (Parsons et al., 2013), 10 randomly chosen soil subsamples were dried for 24 hours at 104°C
176 in order to measure their total humidity. The average soil moisture of these 10 soils was of 1.4
177 ± 0.5% (mean ± sd), which testify of the accuracy of XRF-measurements (Parsons et al.,
178 2013). However, a pilot investigation on a subset of 10 soil samples showed relatively high
179 variance of the concentrations of the different elements between 3 repeated measurements of
180 each soil sample. Therefore, for the rest of the soil samples, the measurements were
181 performed twice per sample in order to increase the reliability of measured concentrations.
182 The average concentration of each element was used in the analyses described below.
183 Detailed data on the pH and the concentrations of each element, and their variation across the
184 soil samples, are available in Table A1.

185 The variation in soil chemical composition was first explored using a PCA on zero-
186 centred and normed data of pH and concentrations of the different elements. We only
187 included in the PCA analysis the elements that were detected in a minimum of 80 quadrats.
188 The missing values were replaced by the average concentration of the element on all the
189 quadrats, so as not to influence the centroid of the PCA. Then, we tested for each element
190 their spatial autocorrelation at the scale of the sampling with a Mantel test based on 9999

permutations (Table A1). All the above-mentioned statistical analyses were done in R 3.6.1 (R core Team 2019) using the ade4 library (Dray & Dray & Dufour 2007).

In order to investigate bioaccumulation processes in Cd, Pb, Zn, and Cu, in *L. castaneus* (see below), the composition of the soil samples of the 30 quadrats where *L. castaneus* were collected for the bioaccumulation study was also quantified using inductively coupled plasma-optical emission spectrometry (ICP-OES). Soil samples were first lyophilized and grinded with a mortar and a pestle. The mineralisation consisted in the digestion of 300 mg of sample in 7 mL of concentrated HNO3, using a Berghof microwave digestion system (speed wave MWS-2-Microwave pressure digestion). The soil samples were analyzed by ICP-OES (ICPOES IRIS Interpid II XSP Thermo, Thermo Scientific, Whatman, MA, USA). We used commercial mussel tissue (ERM®-CE278) as the certified reference material, here and in the bioaccumulation study in *L. castaneus* described below. Measured concentrations in the reference material never differed by more 10% from the certified concentrations.

2.4 Earthworm community and soil pollution

The link between earthworm assemblage and the changes in the set of chemical elements detected in a minimum of 80 quadrats was performed using a canonical redundancy analysis (RDA) with the Vegan package (Oksanen et al., 2018). We followed the recommendation of Legendre and Gallagher (2001) and Hellinger–transformed the species data prior to analysis to tune down dominant species. The statistical significance of the RDA, the canonical axes, and the different elements were tested by the mean of permutation tests (n = 999 permutations). For the species whose abundances showed to be significantly structured by the first axis of the RDA, the shapes of these relationships were investigated. In order to model changes in Hellinger-transformed abundances of species along RDA1, we used

generalized additive models with a quasi-Poisson family and specified the number of knots to 3.

Moreover, MTEs do not only affect adult earthworm populations. MTEs are also known to interfere with species reproduction and, thereby, influence the population dynamics of these species. The effects on the abundance of juveniles of the soil score on the axes of the PCA (PCA1 and PCA2), while correcting for the effect of the abundance of adults present in each quadrat, were tested using a generalized linear model with a Poisson distribution.

2.5 Bioaccumulation of metals in whole *L. castaneus* bodies

The level of MTEs in *L. castaneus* specimens were measured on individuals that were starved for 48 hours to empty the intestinal content, frozen for at least 48 hours, and finally lyophilized for about 60 hours. Then, the specimens were reduced to powder using liquid nitrogen before mineralisation. The mineralisation is a digestion in acid medium (using HNO_3, H_2SO_4 and HCl) at high temperature (for details on the method, see Bernard et al., 2010). The obtained solution was analyzed by ICP-OES (Varian 720-ES, USA) and Cd, Cu, Pb, and Zn, were quantified in the samples. Bioaccumulation factors (BAFs) were calculated according to the following equation: $BAF_{Me} = \frac{c_{Me\ earthworm}}{c_{Me\ soil}}$, where $c_{Me\ earthworm}$ is the total metal concentration in the body of the earthworm ($mg.kg^{-1}$) and $c_{Me\ soil}$ is the total concentration of the same metal in the soil ($mg.kg^{-1}$). Based on preliminary visual investigations of the data, the relationships between the concentration of Cd, Cu, Pb, and Zn, in the specimens and in the soil were investigated using a linear model for Zn and Cd, and included a quadratic term for Pb and Cu.

2.6 Analysis of neutral genetic variation: microsatellite genotyping

Neutral fine-scale genetic structure of *L. castaneus* populations was investigated to determine if the population in the contaminated zone has undergone a strong demographic bottleneck that might be linked to high mortality and to investigate if there was any limits to gene flow between both plots. Total genomic DNA of 175 *L. castaneus* sampled across 63 of the 87 study quadrats (Fig. 1a) was extracted using the NucleoSpin® 96 Tissue kit (Macherey-Nagel). Individuals were genotyped at the eight microsatellite loci described in Dupont et al. (2019). Loci were amplified by polymerase chain reaction (PCR) following the protocol detailed in Dupont et al. (2019). The migration of PCR products was carried out on an ABI 3130 xl Genetic Analyzer using the LIZ500 size standard (Applied Biosystems); alleles were scored using GeneMapper 5 software (Applied Biosystems). All PCR results were repeated and individuals missing three or more loci (e.g. failed PCR, poor-quality DNA extract) were excluded from our data set.

The genetic diversity of the *L. castaneus* population was analyzed by computing allele frequencies, number of alleles (N_{all}), and expected heterozygosity (H_e) using Genetix V 4.05 (Belkhir et al., 2004). The null independence between loci was tested from statistical genotypic disequilibrium analysis using Genepop V4.4 (Rousset 2008). Null allele frequencies were estimated using the software Microchecker (Van Oosterhout et al., 2004; van Oosterhout et al., 2006). Departure from Hardy-Weinberg expectation was quantified by calculating the Weir and Cockerham's (1984) estimator of the fixation index, F_{is}, and conformity to Hardy-Weinberg equilibrium was assessed with exact tests implemented in Genepop V4.4.

Moreover, we investigated the occurrence of a cryptic population structure using the Bayesian model implemented in Geneland V 4.0.3 (Guillot et al., 2005) that simultaneously analyses the spatial and genetic data. The analyses were conducted using both the

uncorrelated and correlated allele frequency models. The correlated frequency model is more powerful at detecting subtle differentiation, but it is also more sensitive to departures from model assumptions (e.g. presence of isolation-by-distance), and more prone to algorithm instabilities, than the uncorrelated frequency model (Guillot et al., 2005). The putative presence of null allele(s) was taken into account in the model (Guillot et al., 2005). The Markov Chain Monte Carlo (MCMC) was run 5 times to check for convergence allowing K to vary from one to three clusters and using 10^6 MCMC iterations.

To further examine the spatial genetic structure of *L. castaneus* at the individual scale, a spatial autocorrelation analysis was conducted using Spagedi 1.2 (Hardy & Vekemans 1999; Hardy & Vekemans 2002). Such an analysis provides a measure of the genetic relatedness between pairs of individuals as a function of their Euclidean distances. Kinship coefficients between individuals (F_{ij}) were estimated as described in Loiselle et al. (1995). We identified 10 classes of spatial distance in order to reach approximately 1400 pairs of individuals per spatial distance class, apart from the 0 distance class with 301 pairs of individuals coming from the same quadrat and having the same spatial coordinates. The average multilocus relationship coefficients per distance class were estimated and their significance per class was tested with 10000 permutations of multilocus genotypes. To visualize the spatial genetic structure, we plotted the kinship coefficient against geographical distance.

3. Results

3.1 Soil pollution heterogeneity

XRF-measurements allowed the relative quantification of 38 elements over all soils sampled. First, it was shown that the concentrations of elements greatly varied among samples. Al, As, Ca, Cr, Fe, K, Mg, Mn, Pb, Rb, S, Si, Sr, Ti, V, Zn, Zr were found in each of

283 the 87 soils sampled (Table A1). Ba, Eu, Ga, Hg, Ir, Mo, Nb, Ni, Os, Re, Sb, Ta, Te, Th, Yb
284 were found in less than 64 of the sampled soils (Table A1). The accuracy of the XRF-
285 measurements greatly varied between measurements and between elements. For example, the
286 coefficient of variation between replicate measurements for As, S, and Cr were of 52.5, 26.1,
287 and 15.8%, respectively. Therefore, even though these elements were also found to
288 significantly vary among samples, suggesting that they are likely to structure the environment,
289 their potential effect should be interpreted with caution as the reliability of their quantification
290 is uncertain. However, for Pb, Cu, and Zn, which are the MTEs of interest, the coefficients of
291 variation among XRF-measures were quite low (from 6.5 to 11.2, Table A1). The XRF-
292 measures for these elements were also highly correlated with the measurements obtained by
293 ICP-OES (Pb: $r = 95.3$, $n = 30$, $p < 0.001$; Cu: $r = 78.0$, $n = 28$, $p < 0.001$; Zn: $r = 99.4$, $n =$
294 30, $p < 0.001$), suggesting that their quantification by XRF is reliable. Cd, which is also a
295 MTE of interest, was always below the level of detection by XRF.

296 The two first axes of the PCA capture 66.6% of the variation in chemical composition
297 found between soil samples (Fig. 1b). The first axis accounts for most of the variation, 52.9%,
298 found among soil samples. This axis is characterized by soils mainly rich in Ti, K, Al, Rb, and
299 Si, at one end, and in Sr, Pb, Zn, Ca, and Cu, at the other end. Soil contamination in Pb, Zn,
300 and Cu, is highly correlated (Fig. 1b). Besides, Mantel's tests have shown that most of the
301 elements that contribute the most to PCA1 also exhibit strong spatial autocorrelation, as is the
302 case for Pb, Zn, and Cu (Mantel test: Pb = 0.44, $p < 0.001$; Zn = 0.40, $p < 0.001$; Cu = 0.25,
303 $p < 0.001$, see Table A1). Such a spatial autocorrelation suggests that the pollution is
304 structured at the scale of our sampling. The visual inspection of the interpolation plot for
305 PCA1 shows that the pollution is indeed structured with a source pollution at the northwest of
306 the site (Fig. 1a, Appendix B).

3.2 Community composition

A total of 891 adults, 21 subadults, and 1129 juveniles, were collected over the 87 samples. On average, we found 10.24 ± 0.92 adults, 12.98 ± 1.33 juveniles, and a juvenile to adult ratio of 1.54 ± 0.15 in each quadrat (mean ± se). No earthworm was collected in four out of the 87 quadrats. Eight species were found: *Al. chlorotica* (L1, n = 418), *L. castaneus* (n = 279), *Ap. rosea* (L4, n = 49), *Ap. icterica* (n = 47), *Ap. longa* (n = 38), *L. festivus* (n = 25), *Ap. giardi* (n = 15), *L. herculeus* (n = 14). We were also able to identify subadults of *Ap. longa* (n = 7), *L. festivus* (n = 6), *Ap. giardi* (n = 4), *Al. chlorotica* (L1, n = 2), and of *Ap. rosea* (L4, n = 2).

The canonical RDA showed that soil chemical composition accounted for 43.1% of the variation in species abundance and the permutation test confirmed the significance of this model ($F = 3.88$, $p = 0.001$, Fig. 2a). Only the first axis of the RDA (RDA1) was significant ($F = 81.84$, $p = 0.001$) and explained 35.4% of the total variance. pH H_2O, Mg, and Fe, significantly explained the variation captured by the RDA ($F(\text{pH } H_2O) = 32.75$, $p = 0.001$; $F(\text{Mg}) = 26.73$, $p = 0.001$; $F(\text{Fe}) = 4.09$, $p = 0.009$). Although not found to be significant, the first axis aligns closely with an increase in the concentration of soils in Cu, Zn, and Pb (Fig. 2a). Thus, in what follows, we interpret the changes in community composition along the first axis of the RDA as a response to an increase of MTE pollution.

The generalized additive models built to investigate for each species changes in the Hellinger-transformed abundance data along RDA1 were significant for 6 of the 8 recorded species and highlighted species differences of sensitivity to soil pollution by MTEs (Fig. 2b, Table A2). Note that while the models were not significant for *Ap. rosea* and *Ap. giardi,* these species were not observed in soils with high scores on the RDA1. Still, in the following our interpretations are limited to the changes in these 6 species, which significantly vary along RDA1. For high RDA1 values, species richness is restricted to *L. festivus* and *L. castaneus*,

the latter being the dominant species. For low RDA1 values, the communities are much more diverse, with the all 6 species being observed in soils with low scores on RDA1. The relative dominance of *Al. chlorotica* in the community sharply decreased with increasing values along the RDA1 axis, as for *Ap. icterica, Ap. longa,* and *L. herculeus,* but the amplitudes of their variation were weaker (Fig. 2b). *L. festivus* was not observed for soils with a low score on RDA1. The explained deviations of the models were high, up to 90.3% for *Al. chlorotica,* suggesting that changes in species abundance and of community assemblage is to a large extent determined by changes in soil chemical composition (Table A2). More specifically, these changes in species abundance along RDA1 are likely to reflect, in part, the increase in concentrations of MTEs in soils, and in particular of Zn, Pb, and Cu (Fig. 2a).

Last, the number of juveniles is positively correlated to the number of adults (estimate = 0.051, z-value = 16.74, $p < 0.001$, Table A3). In addition, the number of juveniles is negatively correlated to PCA1 (estimate = -0.090, z-value = -6.57, $p<0.001$) but positively correlated to PCA2 (estimate = 0.13, z-value = 6.47, $p<0.001$).

3.3 Bioaccumulation in *L. castaneus*

L. castaneus showed high BAFs for Zn [0.68 – 7.81] and Cd [5.41 – 56.97]. The models testing the relationship between the concentration of MTEs in earthworms and the soils showed that bioaccumulation of Zn and Cd by *L. castaneus* was linear (Fig. 3). The models explained 25.9 and 22.3 % of the variance in the data for Zn and Cd, respectively. In contrast, *L. castaneus* showed low BAFs for Cu [0.0027 – 0.26] and Pb [0.005 – 0.23], indicating that at the site, these MTEs do not bioaccumulate in *L. castaneus*. Yet, the concentrations of Cu and Pb in earthworms and the soils are correlated according to a quadratic relationship. The models captured 33.5% and 39.7% of the variance in the data for Cu and Pb, respectively (Fig. 3).

3.4 Population genetic structure of *L. castaneus*

A high neutral genetic diversity was observed within the 175 analysed genotypes (N_{all} = 16.25; H_e = 0.670). Two pairs of loci departed significantly from linkage equilibrium (LC18 – LC33 and LC10-LC36). Since Dupont et al. (2019) showed no physical linkage between these loci, this result could be explained by inbreeding (Nordborg 2000). A second dataset composed of only one genotype per quadrat (i.e. 63 genotypes) was created. All loci were unlinked in this second dataset and a lower proportion of them displayed heterozygote deficiency (i.e. significant F_{is} estimate, Table 1). Geneland Bayesian analysis requiring linkage equilibrium was carried out with this second dataset and identified only one group of individuals in the study plot, regardless of the model of allelic frequency chosen.

Spatial autocorrelation analysis revealed, however, local genetic structure at the scale of the study plot. We found a significant negative relationship between the kinship coefficient and the geographic distance between pair of individuals (b ± se = -0.211 ± 0.079, $p < 0.001$). In particular, positive values of kinship coefficient are measured between individuals collected in close quadrats (mean distance of 10 m and 25 m), which means that neighbouring individuals have a higher genetic relatedness than random pairs of individuals (Fig. 4). Conversely, negative kinship coefficient values, are observed between individuals collected in more distant quadrats (mean distance of 60m and 80m) and indicate isolation by distance (Fig. 4).

4. Discussion

The study plot displays a high heterogeneity in soil chemical composition. In particular, the concentrations of MTEs are highly structured and reflect the division of the plot into the polluted and unpolluted zones. The levels of pollution in Pb, Cu, Cd, and Zn, overlap

with the range of variation found elsewhere in lawn and forest soils of the Paris region (Foti et al., 2017) and the median value of European urban soils and world soils for Pb, Cu, Zn, Ni, Cr, Cd, and As (Baize 1997; Adriano 2001; Desaules 2012; Luo et al., 2012, summarized in Foti et al., 2017). The concentrations of MTEs were, for part, highly correlated, which makes it difficult to disentangle their respective impact on earthworms, suggesting potential cocktail effects. Based on the existing knowledge and literature on the impact of MTEs on earthworm species, in the following we relate the variation of soil chemical composition along RDA1 to a gradient of pollution, since this axis is associated with increasing concentrations of Pb, Cu, and Zn (Fig. 2a).

4.1 Impact of MTEs on earthworm community

Along the gradient of pollution, we observed marked changes in community composition and abundances. *Al. chlorotica* and *L. castaneus* were the dominant species in our sampling. *Al. chlorotica* dominates low-polluted quadrats and is absent from the most polluted part of the plot, while the relative abundance of *L. castaneus* increases along the pollution gradient. Previous studies have already shown that epigeic species, such as *L. castaneus*, are more resistant than endogeic species, such as *Al. chlorotica*, to the effects of MTE pollution (e.g. Spurgeon & Hopkin 1999; Mirmonsef et al., 2017). An explanation for this difference is that endogeic species that live and feed in the mineral soil layers are probably more exposed to the bioavailable fraction of metals, per comparison with epigeic species which are active in the superficial soil and litter layers (Mirmonsef et al., 2017).

Overall, species richness decreased along the pollution gradient, as did the abundance of juveniles compared to that of adults. The latter result is consistent with the literature as Cd, Pb, Cu, and Zn, are known to reduce reproduction rates in several species. For instance, Spurgeon et al. (1994) found a significant negative effect of high concentrations of these four

MTEs, particularly of Cd and Cu, on the cocoon production of the epigeic earthworm *Eisenia fetida* in an artificial soil. Significant decreases in cocoon production was also observed for *L. rubellus* in Cu-amended sandy soil and sandy loam (Ma 1984). Conversely, Reinecke et al. (2001) showed that in three different species, *Eudrilus eugeniae*, *Perionyx excavates*, and *E. fetida*, cocoon viability, but not production, was detrimentally affected by Pb concentrations.

Although MTEs are known to have an impact on earthworms, we cannot exclude the hypothesis that the observed changes of community composition across the plot may be explained by differences in soil properties between zones. The unpolluted zone is characterized by a sandy soil while a loamy soil was found in the polluted zone, and these differences also have certainly an impact on the earthworm community composition. Indeed, spatial variations in the abundance of earthworms are commonly observed and can be partly explained by variations in soil properties (e.g. Nuutinen et al., 1998). It is however noteworthy that the most abundant earthworm species in the polluted zone is *L. castaneus*, an epigeic species known to resist to and bioaccumulate MTEs (present study, Terhivuo et al., 1994; Spurgeon & Hopkin 1996; Tischer 2009).

4.2 Impact of MTEs on the bioaccumulation of *L. castaneus*

We showed that *L. castaneus* bioaccumulate Zn and Cd, and that the bioaccumulation of these two MTEs was highly correlated ($R^2 = 0.89$, $t = 10.5$, $p < 0.001$).

Zn is an essential element necessary for earthworm growth, maturation, and reproduction, and might therefore be required to a number of metabolic processes (Nannoni et al., 2014; Wang et al., 2018a). High concentrations of Zn in earthworm tissues have been recorded elsewhere (Wang et al., 2018a) but it has also been shown that some earthworm species, e.g. *Eisenia fetida*, are able to regulate their uptake of Zn and, thus, do not accumulate this metal (Bernard et al., 2010; Brulle et al., 2011). In our sampling, *L. castaneus*

427 was found to accumulate Zn up to 7.8 fold the concentration found in the soil. The
428 relationship between the concentration of Zn in *L. castaneus* and in soil is linear, suggesting
429 that the bioaccumulation of Zn has not reached a threshold.

430 Although the concentration of Cd was relatively low in the soil, ranging from 0.15 to
431 1.06 mg/kg, this metal accumulated in *L. castaneus* up to 18.06 mg/kg, a result highlighting
432 its strong bioavailability. *L. castaneus* wasshown to bioaccumulate up to 56.9 fold the
433 concentration of Cd found in the surrounding soil. This strong ability to accumulate Cd has
434 been reported for other epigeic and endogeic earthworm species (Bernard et al., 2010; Latif et
435 al., 2013).

436 Conversely, the bioaccumulation factors of Pb and Cu in *L. castaneus* were low. Other
437 studies have shown that Cu and Pb are bioaccumulated by earthworms when their
438 concentrations in soils are particularly high ([Pb] > 900 mg/kg Bernard et al., 2010). The soil
439 concentrations of Cu and Pb at our study site are, comparatively, lower, which might explain
440 why these metals were not found to bioaccumulate in *L. castaneus.* Alternatively, Mirmonsef
441 et al. (2017) proposed that the bioaccumulation of Cu or other heavy metals in earthworm
442 populations can occur in populations that have been exposed for many generations to these
443 metals, as natural selection and genetic adaptation in these populations would have resulted in
444 an increase in their efficiency to sequester and detoxify these MTEs. The low
445 bioaccumulation of Cu and Pb by the population of *L. castaneus* at our study site could then
446 reflect the yet limited duration of their exposure to pollution.

447 **4.3 Impact of MTEs on the genetic characteristics of *L. castaneus***

448 The genetic variability of a population exposed to MTEs may be altered in different
449 ways. Genetic changes in the population may either result from genotoxic exposure (i.e. direct
450 effect) or from microevolutionary processes (i.e. indirect effect). Direct effects are related to

DNA or chromosome alterations which, when they are exerted on gametes and are passed on to the next generation, can significantly impact exposed populations (Medina et al., 2007). In this work, we focused on indirect effects of MTEs, which are population-mediated processes. We assumed that soil pollution by MTEs may alter the diversity of neutral genetic markers, such as microsatellites, through random genetic drift associated to drastic reduction in population size. However, the indices of genetic diversity computed in the population were high, similar to those of other *L. castaneus* populations genotyped with the same markers (Dupont et al., 2019), providing no support for the hypothesis that this population would have undergone a significant reduction of genetic diversity through genetic drift. Alternatively, the genetic structure estimated from neutral markers could be shaped by natural selection of resistant phenotypes. In such instance, we would expect to observe at least two differentiated genetic clusters, corresponding to the polluted and unpolluted zones. Yet, no genetic clustering was revealed at the scale of the study plot. Last, the spatial autocorrelation analysis revealed a pattern of isolation by distance that does not support the role of the heterogeneity of soil pollution by MTEs as a barrier to gene flow in this species. This result, added to the fact that the abundance of *L. castaneus* was found to be high in the polluted quadrats, suggest that direct genotoxic effects might be negligible for this species.

5. Conclusion

It is particularly difficult to study the consequences of MTE pollution on soil biodiversity in the field, mainly because of confounding and cocktail effects (e.g. Ye et al., 2017). Although laboratory experiments are valuable in testing theory and in providing quantitative estimates of survival and reproduction rates of species under controlled conditions (e.g. level of contamination), they can be difficult to implement when the pollution is multifactorial and heterogeneous, as generally observed in urban areas. Moreover, these

475 experiments frequently use the laboratory earthworm models, *E. fetida* and *E. andrei,* while
476 they are often rare in the field (e.g. Coelho et al., 2018).

477 Here, we used a multidisciplinary approach to study in the field the response to MTE
478 pollution of an earthworm community in an urban area and to further our understanding of the
479 bioaccumulation capacities, population genetic structure and gene expression of a MTE
480 tolerant species in response to pollution. *L. castaneus* was identified as the most tolerant
481 species to MTEs of the study site. In sites contaminated by MTEs, the maintenance of
482 earthworm populations and their associated functions in the ecosystem (Pauwels et al., 2013)
483 rely on the evolution of molecular mechanisms of metal tolerance, which, however, remain
484 poorly understood. As mentioned elsewhere (Stapley et al., 2010; Vandegehuchte & Janssen
485 2014; Evans 2015), the study of gene expression profile in populations under different
486 selection pressure should provide new insights into the molecular mechanisms of metal
487 tolerance in earthworms, and help identify candidate functional genes that may be under
488 selection. Although still expensive, the Next-generation sequencing (NGS), now permits
489 direct transcriptome sequencing, and can provide such qualitative and quantitative
490 information on the expression of genes.

Acknowledgements

This work was funded by the Ile de France region through the "Partenariats institutions-citoyens pour la recherche et l'innovation" (Picri) calls for projects (ReFUJ project). H. Audusseau acknowledges support from the Swedish Research Council (2016-06737). This work benefited from discussions with J. Mathieu (Sorbonne – University) and T. Lerch (UPEC).

References

Adriano DC (2001) Trace metals in terrestrial environments: biogeochemistry, bioavailability and risks of metals, p. 866. Springer-Verlag, New York.

Ajmone-Marsan F & Biasioli M (2010) Trace elements in soils of urban areas. Water Air and Soil Pollution, 213, 121-143.

Argyraki A, Kelepertzis E, Botsou F, Paraskevopoulou V, Katsikis I & Trigoni M (2018) Environmental availability of trace elements (Pb, Cd, Zn, Cu) in soil from urban, suburban, rural and mining areas of Attica, Hellas. Journal of Geochemical Exploration, 187, 201-213.

Baize D (1997) Teneurs totales en éléments traces métalliques dans les sols (France). Références et stratégies d'interprétation., p. 410. INRA Editions, Paris.

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).

Bernard F, Brulle F, Douay F, Lemiere S, Demuynck S & Vandenbulcke F (2010) Metallic trace element body burdens and gene expression analysis of biomarker candidates in *Eisenia fetida*, using an "exposure/depuration" experimental scheme with field soils. Ecotoxicology and Environmental Safety, 73, 1034-1045.

Blouin M, Hodson ME, Delgado EA, Baker G, Brussaard L, Butt KR, Dai J, Dendooven L, Peres G, Tondoh JE, Cluzeau D & Brun JJ (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.

Brulle F, Lemiere S, Waterlot C, Douay F & Vandenbulcke F (2011) Gene expression analysis of 4 biomarker candidates in *Eisenia fetida* exposed to an environmental metallic trace elements gradient: A microcosm study. Science of the Total Environment, 409, 5470-5482.

Brulle F, Mitta G, Cocquerelle C, Vieau D, Lemiere S, Lepretre A & vandenBulcke F (2006) Cloning and real-time PCR testing of 14 potential biomarkers in *Eisenia fetida* following cadmium exposure. Environmental Science & Technology, 40, 2844-2850.

Coelho C, Foret C, Bazin C, Leduc L, Hammada M, Inacio M & Bedell JP (2018) Bioavailability and bioaccumulation of heavy metals of several soils and sediments (from industrialized urban areas) for *Eisenia fetida*. Science of the Total Environment, 635, 1317-1330.

Desaules A (2012) Critical evaluation of soil contamination assessment methods for trace metals. Science of the Total Environment, 426, 120-131.

Dray S & Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software, 22, 1-20.

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, Pauwels M, Dume C, Deschins V, Audusseau H, Gigon A, Dubs F & Vandenbulcke F (2019) Genetic variation of the epigeic earthworm *Lumbricus*

539 *castaneus* populations in urban soils of the Paris region (France) revealed using eight
540 newly developed microsatellite markers. Applied Soil Ecology, 135, 33-37.

541 Evans TG (2015) Considerations for the use of transcriptomics in identifying the 'genes that
542 matter' for environmental adaptation. Journal of Experimental Biology, 218, 1925-
543 1935.

544 Foti L, Dubs F, Gignoux J, Lata J-C, Lerch TZ, Mathieu J, Nold F, Nunan N, Raynaud X,
545 Abbadie L & Barot S (2017) Trace element concentrations along a gradient of urban
546 pressure in forest and lawn soils of the Paris region (France). Science of the Total
547 Environment, 598, 938-948.

548 Grumiaux F, Demuynck S, Pernin C & Lepretre A (2015) Earthworm populations of highly
549 metal-contaminated soils restored by fly ash-aided phytostabilisation. Ecotoxicology
550 and Environmental Safety, 113, 183-190.

551 Guillot G, Mortier F & Estoup A (2005) Geneland : A program for landscape genetics.
552 Molecular Ecology Notes, 5, 712-715.

553 Guittard A (2013) Evaluation quantitative des risques sanitaires (EQRS) du site de Pierre Frite
554 à Villeneuve-Le-Roi (94) (ed. CG94). BG Ingénieurs Conseils.

555 Hardy OJ & Vekemans X (1999) Isolation by distance in a continuous population:
556 reconciliation between spatial autocorrelation analysis and population genetics
557 models. Heredity, 83, 145-154.

558 Hardy OJ & Vekemans X (2002) SPAGEDi: a versatile computer program to analyse spatial
559 genetic structure at the individual or population levels. Molecular Ecology Notes, 2,
560 618-620.

561 Hebert PDN, Cywinska A, Ball SL & Dewaard JR (2003) Biological identifications through
562 DNA barcodes. Proceedings of the Royal Society of London Series B: Biological
563 Sciences, 270, 313-321.

564 Holderegger R, Kamm U & Gugerli F (2006) Adaptive vs. neutral genetic diversity:
565 implications for landscape genetics. Landscape Ecology, 21, 797-807.

566 Hou DY, O'Connor D, Nathanail P, Tian L & Ma Y (2017) Integrated GIS and multivariate
567 statistical analysis for regional scale assessment of heavy metal soil contamination: A
568 critical review. Environmental Pollution, 231, 1188-1200.

569 Iordache M & Borza I (2012) The bioremediation potential of earthworms (Oligochaeta:
570 Lumbricidae) in a soil polluted with heavy metals. Journal of Food Agriculture &
571 Environment, 10, 1183-1186.

572 Jia Z, Li S & Wang L (2018) Assessment of soil heavy metals for eco-environment and
573 human health in a rapidly urbanization area of the upper Yangtze Basin. Scientific
574 Reports, 8, 3256.

575 Kirk H & Freeland JR (2011) Applications and implications of neutral versus non-neutral
576 markers in Molecular Ecology. International Journal of Molecular Sciences, 12, 3966-
577 3988.

578 Knox A, Seamans JC, Mench MJ & Vangronseveld J (2000) Remediation of metal-and
579 radionuclides- contaminated soils by in situ stabilization techniques. In:
580 Environmental restoration of metals-contaminated soils (ed. Islandar, I.K.). Lewis
581 Publishers, Boca Raton.

582 Latif R, Malek M & Mirmonsef H (2013) Cadmium and lead accumulation in three endogeic
583 earthworm species. Bulletin of Environmental Contamination and Toxicology, 90,
584 456-459.

585 Legendre P & Gallagher ED (2001) Ecologically meaningful transformations for ordination of
586 species data. Oecologia, 129, 271-280.

587 Leveque T, Capowiez Y, Schreck E, Mombo S, Mazzia C, Foucault Y & Foucault G (2015)
588 Effects of historic metal(loid) pollution on earthworm communities. Science of the
589 Total Environment, 511, 738-746.

590 Loiselle BA, Sork VL, Nason J & Graham C (1995) Spatial genetic structure of a tropical
591 understory shrub, *Psychotria officinalis* (Rubiaceae). American Journal of Botany, 82,
592 1420-1425.

593 Luo XS, Yu S, Zhu YG & Li XD (2012) Trace metal contamination in urban soils of China.
594 Science of the Total Environment, 421, 17-30.

595 Ma W (1984) Sublethal toxic effects of copper on growth, reproduction and litter breakdown
596 activity in the earthworm *Lumbricus rubellus*, with observations on the influence of
597 temperature and soil pH. . Environmental Pollution Series a-Ecological and
598 Biological, 33, 207-219.

599 McLlwaine R, Doherty R, Cox SF & Cave M (2017) The relationship between historical
600 development and potentially toxic element concentrations in urban soils.
601 Environmental Pollution, 220, 1036-1049.

602 Medina MH, Correa JA & Barata C (2007) Micro-evolution due to pollution: possible
603 consequences for ecosystem responses to toxic stress. Chemosphere, 67, 2105-2114.

604 Mirmonsef H, Hornum HD, Jensen J & Holmstrup M (2017) Effects of an aged copper
605 contamination on distribution of earthworms, reproduction and cocoon hatchability.
606 Ecotoxicology and Environmental Safety, 135, 267-275.

607 Nahmani J, Hodson ME & Black S (2007) A review of studies performed to assess metal
608 uptake by earthworms. Environmental Pollution, 145, 402-424.

609 Nahmani J, Lavelle P, Lapied E & van Oort F (2003) Effects of heavy metal soil pollution on
610 earthworm communities in the north of France. Pedobiologia, 47, 663-669.

611 Nannoni F, Rossi S & Protano G (2014) Soil properties and metal accumulation by
612 earthworms in the Siena urban area (Italy). Applied Soil Ecology, 77, 9-17.

613 Nordborg M (2000) Linkage disequilibrium, gene trees and selfing: An ancestral
614 recombination graph with partial self-fertilization. Genetics, 154, 923-929.

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

617 Nuutinen V, Pitkanen J, Kuusela E, Widbom T & Lohilahti H (1998) Spatial variation of an
618 earthworm community related to soil properties and yield in a grass-clover field.
619 Applied Soil Ecology, 8, 85-94.

620 Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara
621 RB, Simpson GL, Solymos P, Stevens MH, Szoecs E & Wagner HH (2018) vegan:
622 Community Ecology Package. R package version 2.5-3.

623 Parsons C, Grabulosa EM, Pili E, Floor GH, Roman-Ross G & Charlet L (2013)
624 Quantification of trace arsenic in soils by field-portable X-ray fluorescence

spectrometry: Considerations for sample preparation and measurement conditions. Journal of Hazardous Materials, 262, 1213-1222.

Pauwels M, Frerot H, Souleman D & Vandenbulcke F (2013) Using biomarkers in an evolutionary context: Lessons from the analysis of biological responses of oligochaete annelids to metal exposure. Environmental Pollution, 179, 343-350.

Reinecke AJ, Reinecke SA & Maboeta MS (2001) Cocoon production and viability as endpoints in toxicity testing of heavy metals with three earthworm species. Pedobiologia, 45, 61-68.

Ribeiro R & Lopes I (2013) Contaminant driven genetic erosion and associated hypotheses on alleles loss, reduced population growth rate and increased susceptibility to future stressors: an essay. Ecotoxicology, 22, 889-899.

Rodriguez-Eugenio N, McLaughlin M & Pennock D (2018) Soil pollution : a hidden reality, p. 142. FAO.

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

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.

Snape JR, Maund SJ, Pickford DB & Hutchinson TH (2004) Ecotoxicogenomics: the challenge of integrating genomics into aquatic and terrestrial ecotoxicology. Aquatic Toxicology, 67, 143-154.

Sol Paysage (2013) Villeneuve-Le-Roi Projet d'aménagement d'une friche - Pierrefrite. Rapport d'études agro-pédologiques. Conseil Général 94, Orsay.

Spurgeon DJ & Hopkin SP (1996) The effects of metal contamination on earthworm populations around a smelting works: Quantifying species effects. Applied Soil Ecology, 4, 147-160.

Spurgeon DJ & Hopkin SP (1999) Seasonal variation in the abundance, biomass and biodiversity of earthworms in soils contaminated with metal emissions from a primary smelting works. Journal of Applied Ecology, 36, 173-183.

Spurgeon DJ, Hopkin SP & Jones DT (1994) Effects of cadmium, copper, lead and zinc on growth, reproduction and survival of the earthworm *Eisenia fetida* (Savigny) - Assessing the environmental impact of point-source metal contamination in terrestrial ecosystems. Environmental Pollution, 84, 123-130.

Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, Ekblom R, Bennison C, Ball AD, Beckerman AP & Slate J (2010) Adaptation genomics: the next generation. Trends in Ecology & Evolution, 25, 705-712.

Sturzenbaum SR, Kille P & Morgan AJ (1998) Heavy metal-induced molecular responses in the earthworm, *Lumbricus rubellus* genetic fingerprinting by directed differential display. Applied Soil Ecology, 9, 495-500.

Terhivuo J, Pankakoski E, Hyvarinen H & Koivisto I (1994) Pb uptake by ecologically dissimilar earthworm (lumbricidae) species near a lead smelter in South Finland. Environmental Pollution, 85, 87-96.

Tischer S (2009) Earthworms (Lumbricidae) as bioindicators: the relationship between in-soil and in-tissue heavy metal content. Polish Journal of Ecology, 57, 513-523.

669 Tyler G, Pahlsson AMB, Bengtsson G, Baath E & Tranvik L (1989) Heavy-metal ecology of
670 terrestrial plants, microorganisms and invertebrates - a review. Water Air and Soil
671 Pollution, 47, 189-215.

672 van Gestel CAM, Koolhaas JE, Hamers T, van Hoppe M, van Roovert M, Korsman C &
673 Reineckec SA (2009) Effects of metal pollution on earthworm communities in a
674 contaminated floodplain area: Linking biomarker, community and functional
675 responses. Environmental Pollution, 157, 895-903.

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

679 van Oosterhout C, Weetman D & Hutchinson WF (2006) Estimation and adjustment of
680 microsatellite null alleles in nonequilibrium populations. Molecular Ecology Notes, 6,
681 255-256.

682 van Straalen NM, Butovsky RO, Pokarzhevskii AD, Zaitsev AS & Verhoef SC (2001) Metal
683 concentrations in soil and invertebrates in the vicinity of a metallurgical factory near
684 Tula (Russia). Pedobiologia, 45, 451-466.

685 Vandegehuchte MB & Janssen CR (2014) Epigenetics in an ecotoxicological context.
686 Mutation Research-Genetic Toxicology and Environmental Mutagenesis, 764, 36-45.

687 Vijver MG, Van Gestel CAM, Lanno RP, Van Straalen NM & Peijnenburg W (2004) Internal
688 metal sequestration and its ecotoxicological relevance: A review. Environmental
689 Science & Technology, 38, 4705-4712.

690 Wang K, Qiao Y, Zhang H, Yue S, Li H, Ji X & Liu L (2018a) Bioaccumulation of heavy
691 metals in earthworms from field contaminated soil in a subtropical area of China.
692 Ecotoxicology and Environmental Safety, 148, 876-883.

693 Wang K, Qiao YH, Zhang HQ, Yue SZ, Li HF, Ji XH & Crowley D (2018b) Influence of
694 cadmium-contaminated soil on earthworm communities in a subtropical area of China.
695 Applied Soil Ecology, 127, 64-73.

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

698 Weissmannova HD & Pavlovsky J (2017) Indices of soil contamination by heavy metals -
699 methodology of calculation for pollution assessment (minireview). Environmental
700 Monitoring and Assessment, 189, 616.

701 Ye SJ, Zeng GM, Wu HP, Zhang C, Liang J, Dai J, Liu ZF, Xiong WP, Wan J, Xu PA &
702 Cheng M (2017) Co-occurrence and interactions of pollutants, and their impacts on
703 soil remediation-A review. Critical Reviews in Environmental Science and
704 Technology, 47, 1528-1553.

Figure legends

Figure 1 (a) Interpolation maps of soil score on PCA1. We interpolated the data using ordinary kriging and validated our interpolation plot by cross validation of the residuals (see Appendix B for further explanation on the interpolation plot). The pie charts indicate the location of the quadrats and the purpose which the collected earthworms were used for. (b) Biplot showing the contribution to the two PCA axes of pH and chemical elements, as well as the position of soil samples (quadrats) in this two-dimensional space.

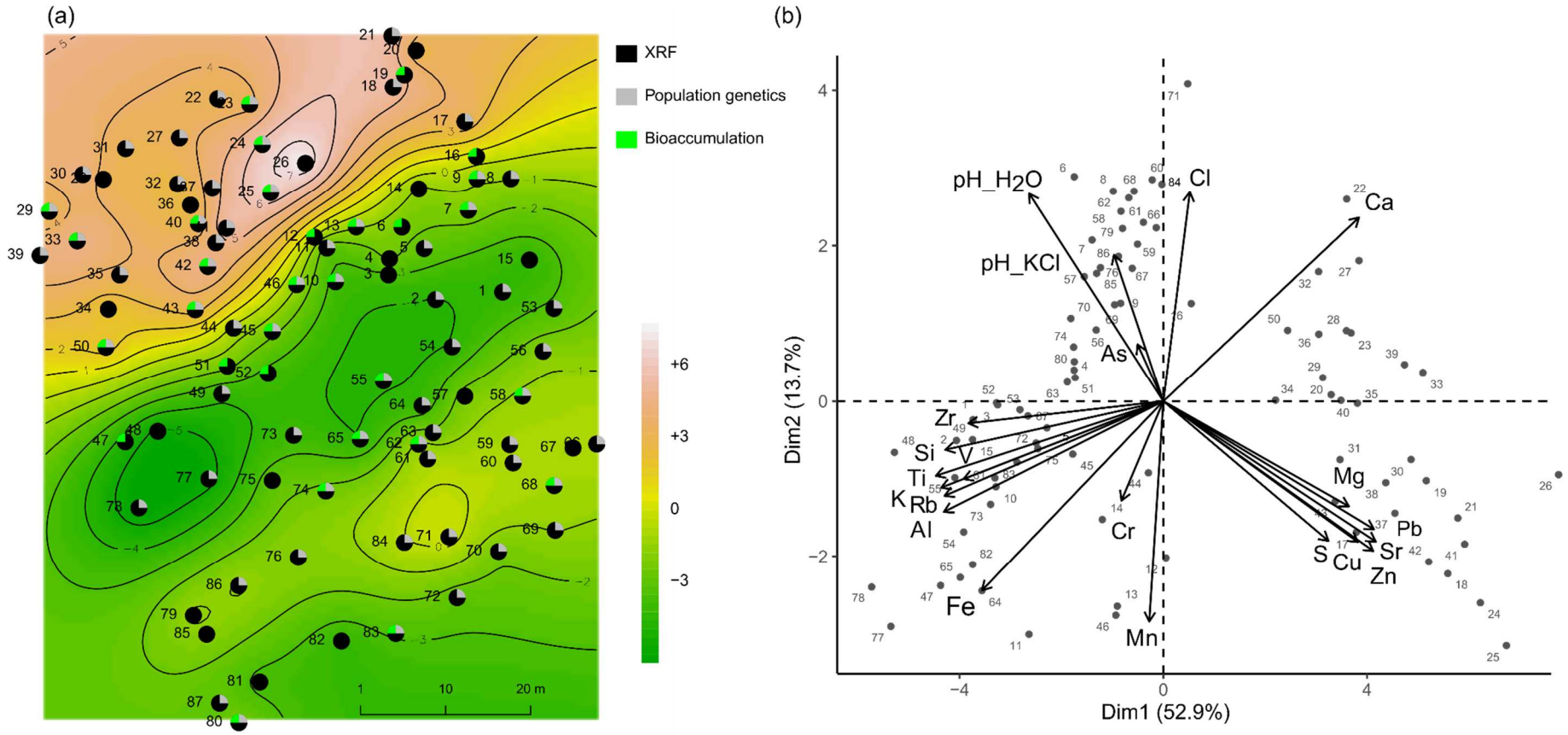
Figure 2 (a) RDA triplots of the Hellinger-transformed earthworm abundance data constrained by the elements measured in a minimum of 80 quadrats. (b) Changes in Hellinger-transformed abundance data along RDA1 for *L. castaneus, L. festivus, L. herculeus, Ap.longa, Al. chlorotica,* and *Ap. icterica*. The fitted GAM is depicted by a line for each species and the dots correspond to the raw data.

Figure 3 Relationship between the concentrations of MTEs measured by ICP in earthworms and soils for Cu and Pb (left panel), Zn (central panel), and Cd (right panel).

Figure 4 Average kinship coefficients, F_{ij}, between pairs of *L. castaneus* individuals plotted against the geographical distance. Dashed lines represent 95% confidence intervals for F_{ij} under the null hypothesis that genotypes are randomly distributed. Significant values: *P < 0.05.

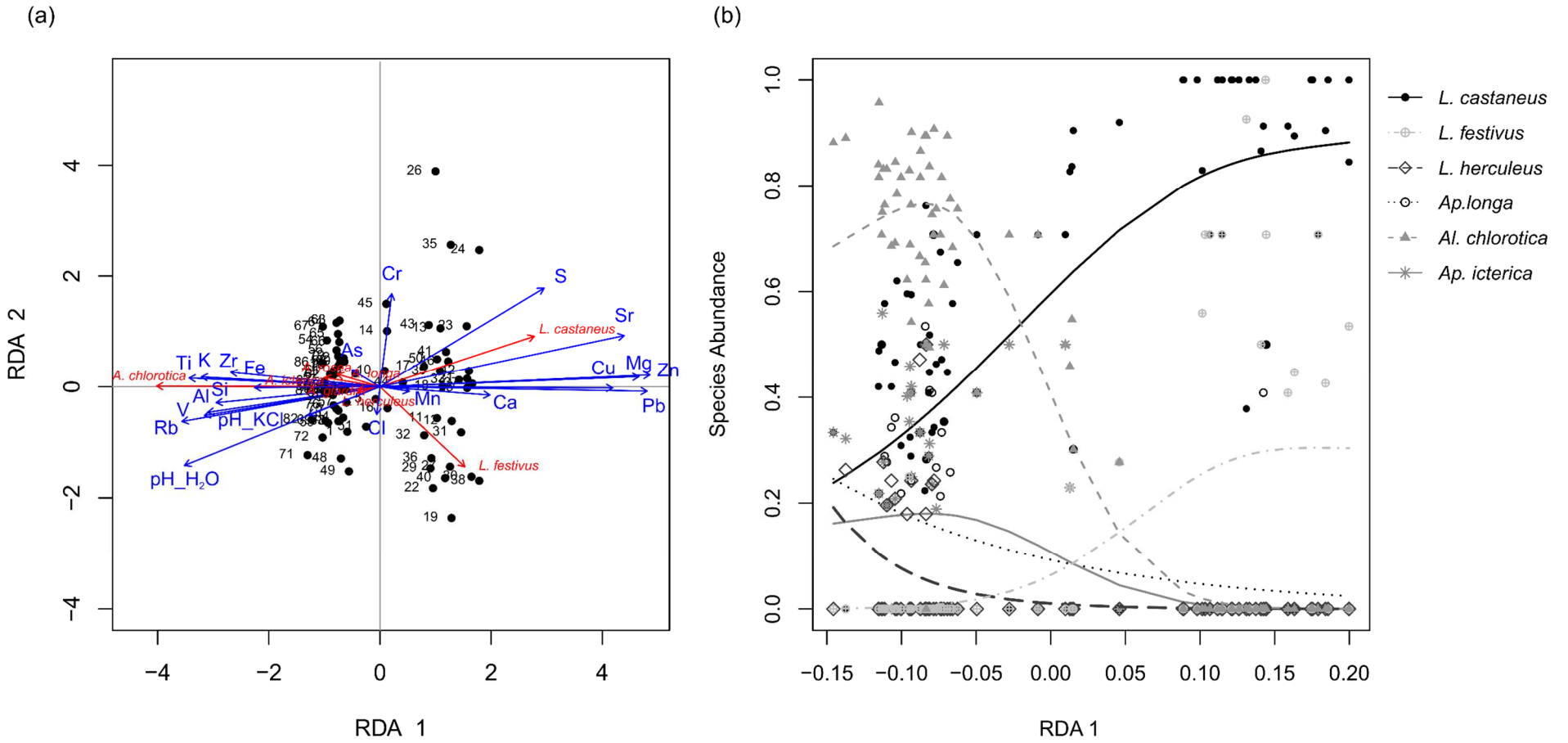
726 **Figure 1**

727



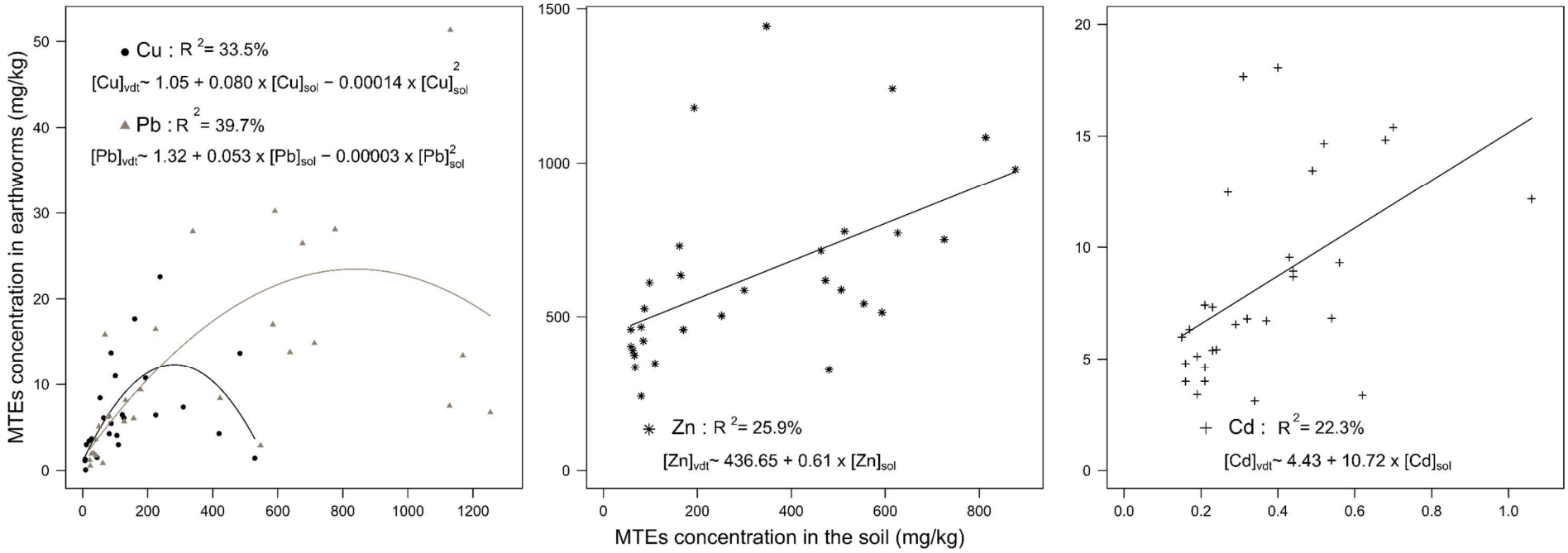
728
729

730 **Figure 2**



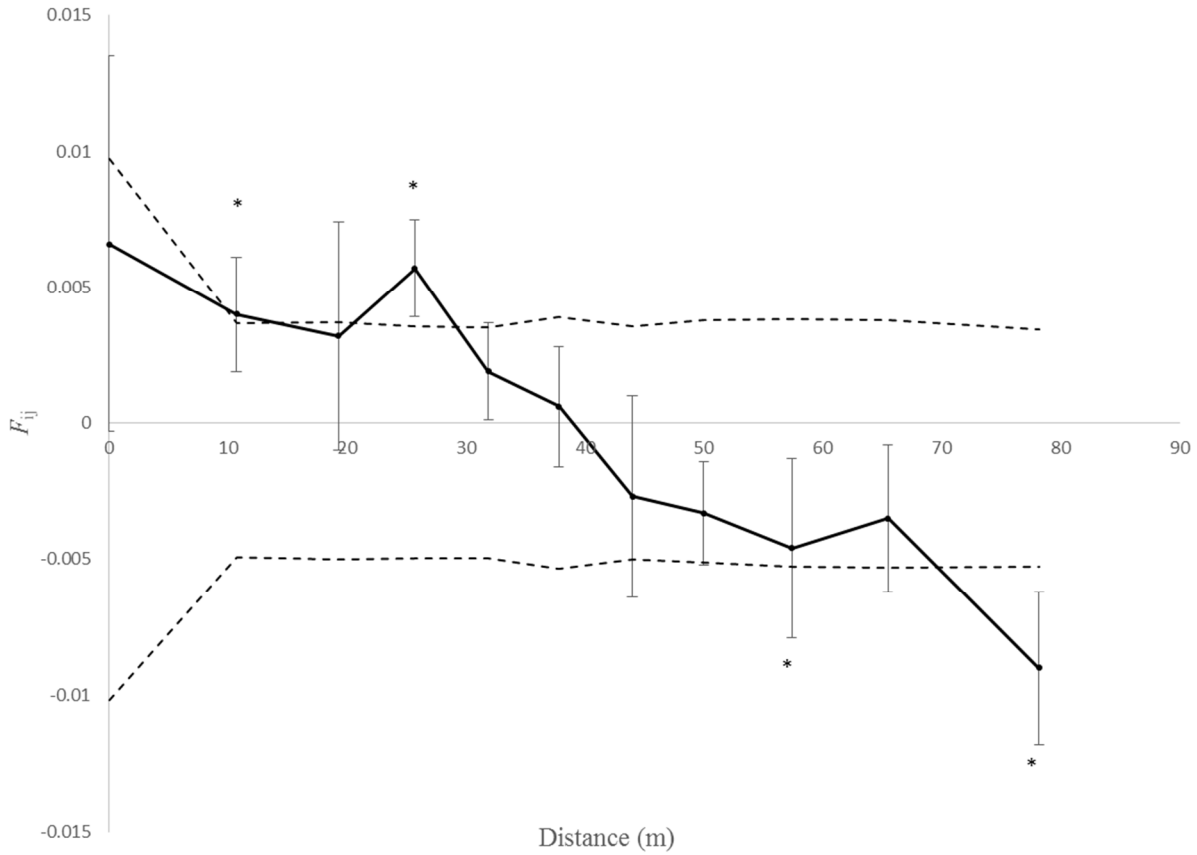
731
732

733 **Figure 3**



734
735

736 **Figure 4**



737

738 **Table 1 Characteristics of microsatellite data in the global dataset (175 genotypes) and in a subsample of one genotype per quadrat (63**
739 **genotypes). Are indicated: the number of alleles for each locus (Na), the estimator of the fixation index (F_{is}) with significant value in**
740 **bold, and the null-allele frequency (Null).**

Locus		**LC02**	**LC05**	**LC10**	**LC16**	**LC18**	**LC27**	**LC33**	**LC36**
175	N_a	5	10	18	36	22	10	10	19
genotypes	F_{is}	**0.184**	-0.081	**0.117**	**0.331**	**0.175**	-0.010	**0.239**	**0.334**
	Null	**0.096**	0.000	**0.057**	**0.162**	**0.082**	0.000	**0.118**	**0.164**
63	N_a	4	7	14	28	17	9	9	17
genotypes	F_{is}	0.118	-0.077	**0.166**	**0.362**	0.093	-0.098	**0.316**	**0.301**
	Null	0.000	0.000	**0.076**	**0.177**	0.000	0.000	**0.159**	**0.155**

741

Appendix A

Table A1 Summary measures of the elemental composition and pH of the soil sampled. N corresponds to the number of quadrats in which the element were found and quantified by XRF. The variance between repeated XRF-measures for a soil sample is given by the minimum and maximum standard error ("min se" or "max se") for all measured soils. It corresponds to range of standard error found when at least two replicated XRF measures report the element in a sample. Mean CV corresponds to the average coefficient of variation found across measurements. Spatial autocorrelation across the study site was measure for each element using the Mantel test. Note that Mantel tests for Ir, Nb, Th, Mo, Sb, Te, Os, Ta were not performed as the sample size for those element was too small (< 8).

element	pH H_2O	pH KCl	S	Pb	K	Zn	Cr	As	Mg	Sr	V	Al	Zr	Ti	Rb	Ca	Si	Fe	Mn
min value (in pH unit and mass concentration in mg/kg)	6.41	6.42	57.0	14.0	1741.0	25.5	13.0	3.0	372.0	91.0	8.5	2826.5	46.0	581.5	21.0	43682.0	16030.0	7962.0	178.0
max (in pH unit and mass concentration in mg/kg)	7.87	7.45	7262.5	940.0	3534.0	638.5	85.0	18.5	2205.0	441.0	33.0	8192.0	129.0	1463.0	45.5	62140.0	31495.5	13324.0	246.7
mean (in pH unit and mass concentration in mg/kg)	7.28	7.24	651.3	214.5	2629.4	182.4	26.3	9.0	821.9	162.8	19.6	5202.2	81.4	1076.8	31.2	54107.9	24380.6	9932.1	206.5
sd (in pH unit and mass concentration in mg/kg)	0.27	0.18	1143.4	243.7	486.0	177.9	8.1	2.7	466.8	87.8	5.4	1103.5	16.5	209.4	5.6	4826.3	3403.6	1231.3	16.1
N	87	87	87	87	87	87	87	87	87	87	87	87	87	87	87	87	87	87	87
max/min ratio	1.23	1.16	127.4	67.1	2.0	25.0	6.5	6.2	5.9	4.8	3.9	2.9	2.8	2.5	2.2	1.4	2.0	1.7	1.4
min se	-	-	0.9	0.0	10.0	0.0	0.0	0.0	0.0	0.0	0.0	10.5	0.5	0.0	0.0	5.5	6.0	0.5	0.0
max se	-	-	307.5	76.5	352.7	116.5	30.0	8.5	134.0	98.0	9.0	886.0	15.0	530.5	4.0	1559.3	1847.0	854.0	40.0
mean CV (%)	-	-	26.1	6.5	6.3	7.9	15.8	52.5	3.2	5.7	16.4	11.3	8.9	6.7	5.3	0.9	3.7	2.5	7.6
Mantel test, p-value	0.03, p=0.23	0.005, p=0.43	0.06, p=0.15	0.44, p<0.001	0.19, p<0.001	0.40, p<0.001	0.015, p=0.37	0.04, p=0.14	0.44, p<0.001	0.26, p<0.001	0.18, p<0.001	0.16, p<0.001	0.14, p<0.001	0.25, p<0.001	0.27, p<0.001	0.04, p=0.12	0.16, p<0.001	0.11, p=0.004	0.06, p=0.052

749

element	Cl	Cu	Sn	P	Br	Y	Eu	Re	Hg	Ni	Ba	Yb	Ga	Ir	Nb	Th	Mo	Sb	Te	Os	Ta
min value (in pH unit and mass concentration in mg/kg)	1.5	9.0	27.0	6.0	2.0	5.0	1.0	2.0	1.0	6.0	56.0	0.0	0.5	0.5	4.0	1.0	1.0	27.0	27.0	8	3
max (in pH unit and mass concentration in mg/kg)	44.0	608.0	98.0	256.7	5.5	13.5	130.0	10.0	5.0	19.0	203.3	101.0	4.0	3.0	6.0	4.0	1.0	36.0	27.0	8	3
mean (in pH unit and mass concentration in mg/kg)	21.3	70.6	39.4	94.3	3.9	9.2	8.5	4.8	2.7	10.7	127.1	7.2	2.1	1.8	4.7	3.0	1.0	31.5	-	-	-
sd (in pH unit and mass concentration in mg/kg)	8.7	86.0	13.1	73.9	0.8	1.6	15.7	1.5	1.1	3.5	34.7	16.5	1.0	0.9	0.9	1.4	0.0	4.5	-	-	-
N	84	79	75	71	68	66	62	50	44	43	43	34	31	8	3	3	2	2	1	1	1
max/min ratio	29.3	67.6	3.6	42.8	2.8	2.7	130.0	5.0	5.0	3.2	3.6	-	8.0	6.0	1.5	4.0	1.0	1.3	-	-	-
min se	0.5	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.5	0.5	0.0	0.0	0.0	-	-	-	-	-	-	-
max se	23.0	437.0	26.5	29.5	1.0	1.0	126.0	3.5	1.5	6.5	21.0	99.0	1.0	0.6	-	-	-	-	-	-	-
mean CV (%)	42.6	11.2	9.6	22.9	8.6	5.1	27.3	39.3	38.1	31.9	6.8	46.1	37.5	63.8	-	-	-	-	-	-	-
Mantel test, p-value	0.06, p=0.054	0.25, p<0.001	0.26, p<0.001	0.40, p<0.001	0.15, p=0.007	0.18, p=0.001	-0.01, p=0.41	0.007, p=0.41	0.26, p<0.001	0.14, p=0.01	0.14, p=0.019	-0.035, p=0.47	0.31, p=0.001	-	-	-	-	-	-	-	-

750

Table A2 Summary table of the GAM univariate models built for each species that showed a significant relationship between Hellinger-transformed abundance data and RDA1.

Model	Edf	Ref.df	F	p-value	Deviance explained (%)	Adj.R2
L. cataneus	1.75	1.94	35.01	<0.001	36.7	48.1
Ap. icterica	1.84	1.97	3.66	0.036	34.8	19.1
L. festivus	1.87	1.98	7.32	<0.001	45.5	28.2
L. herculeus	1.00	1.00	4.74	0.030	29.2	9.9
A. chlorotica	1.98	2.00	52.75	<0.001	90.3	86.0
A. longa	1.00	1.00	8.67	0.004	16.0	11.0

754 **Table A3. Analysis of deviance table (type II LR Chi-square tests) showing the effect of the soils score on PCA1 and PCA2 and of the**

755 **number of adults on juvenile abundances.**

756

# juveniles	df	Chisq	*p*
# adults	1	254.97	<0.001
PCA1	1	46.64	<0.001
PCA2	1	42.23	<0.001

757

Appendix B

The values of each quadrat on the first axis of the PCA (PCA1) were interpolated into a pollution map with a mesh size of 50cm (similar to the sampling size). To create this map of pollution we used the kriging interpolation method to account for spatial autocorrelation among points and used a Matern distribution to model variogram. The reliability of the interpolation map was checked through cross-validation of the residuals using 60 randomly selected points for the modelling set and 27 for the validation set. Note that we did not build such a map for the individual elements of particular interest in this study as they are all very much correlated to each other and to PCA1. Pearson correlation tests between Pb, Cu, and Zn, followed by Bonferroni correction for multiple testing showed that Pb and Cu were significantly correlated to Zn at 0.94 and 0.92, respectively. Cd was excluded as it was found in none of the soils analysed by XRF. All the above mentioned statistical analyses were done in R 3.6.1 (R core Team 2019) using gstat (Pebesma 2004, Gräler et al. 2016) and raster (Hijmans 2019) libraries.

References:

Pebesma EJ (2004) Multivariable geostatistics in S: the gstat package. Computers & Geosciences, 30: 683-691.

Gräler B, Pebesma E & Heuvelink G (2016) Spatio-Temporal Interpolation using gstat. The R Journal 8(1), 204-218

Hijmans RJ (2019). raster: Geographic Data Analysis and Modeling. R package version 2.9-5. https://CRAN.R-project.org