



HAL
open science

Genetic evidence of multiple introductions and mixed reproductive strategy in the peregrine earthworm *Pontoscolex corethrurus*

S. Taheri, T. Decaëns, L. Cunha, G. G Brown, E. da Silva, M. L C Bartz, D. Baretta, L. Dupont

► **To cite this version:**

S. Taheri, T. Decaëns, L. Cunha, G. G Brown, E. da Silva, et al.. Genetic evidence of multiple introductions and mixed reproductive strategy in the peregrine earthworm *Pontoscolex corethrurus*. *Biological Invasions*, Springer Verlag, 2020, 22 (8), pp.2545-2557. 10.1007/s10530-020-02270-0 . hal-03002216

HAL Id: hal-03002216

<https://hal.sorbonne-universite.fr/hal-03002216>

Submitted on 12 Nov 2020

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.

1 **Genetic evidence of multiple introductions and mixed reproductive strategy in the**
2 **peregrine earthworm *Pontoscolex corethrurus***

3
4 S. Taheri^{1,2,3}, T. Decaëns⁴, L. Cunha^{5,6}, G. G. Brown⁶, E. Da Silva⁶, M. L. C. Bartz⁷, D.
5 Baretta⁸, L. Dupont^{1,2,3*}

6
7 ¹ *Univ Paris-Est Creteil, CNRS, INRAE, IRD, IEES-Paris, F-94010 Creteil France*

8 ² *Sorbonne Université, IEES-Paris, F-75005 Paris, France*

9 ³ *Université de Paris, IEES-Paris, F-75013 Paris, France*

10 ⁴ *Centre d'Ecologie Fonctionnelle et Evolutive (CEFE UMR 5175, CNRS–Université de*
11 *Montpellier–Université Paul-Valéry Montpellier–EPHE), 1919 Route de Mende, F-34293*
12 *Montpellier, FRANCE*

13 ⁵ *Cardiff school of Biosciences, Cardiff University, Museum Avenue, Cardiff, UK CF10 3AX*

14 ⁶ *Embrapa Florestas, Estrada da Ribeira km. 111, Colombo, PR, 83411-000, Brazil*

15 ⁷ *Universidade Positivo, Rua Prof Pedro Viriato Parigot de Souza, 5300, CEP 81280-330,*
16 *Campo Comprido, Curitiba, Parana, Brazil*

17 ⁸ *Universidade do Estado de Santa Catarina, Centro de Educação Superior do Oeste, Chapeco,*
18 *Brazil*

19
20 *Corresponding author: lise.dupont@u-pec.fr

21

22 **Keywords** : founder effect, genetic diversity, parthenogenesis, recombination, soil
23 macrofauna, tropical invasive species.

24

25

26 **Abstract**

27 *Pontoscolex corethrurus* is a well-known invasive earthworm in tropical zone which is
28 believed to have originated from the Guayana Shield in South America and was described as
29 parthenogenetic. A recent phylogenetic study revealed four cryptic species in the *P. corethrurus*
30 complex (L1, L2, L3 and L4), among them L1 was particularly widespread and was proposed
31 as *P. corethrurus sensu stricto*. Here, our aims were to investigate the genetic variation of *P.*
32 *corethrurus* L1 in its presumed native and introduced ranges and to examine its reproductive
33 strategy. An extensive dataset of 478 cytochrome oxidase I gene (COI) sequences, obtained in
34 specimens sampled all around the world, revealed a weak COI haplotype diversity with one
35 major haplotype (H1) present in 76% of the specimens. Analyses of the genetic variation of 12
36 L1 populations were done using both nuclear (226 AFLP profiles) and mitochondrial (269 COI
37 sequences) genetic information. The high AFLP genotype diversity at the worldwide scale and
38 the fact that no genotype was shared among populations, allowed to reject the ‘super-clone’
39 invasion hypothesis. Moreover, a similar level of mean genetic diversity indices were observed
40 between the introduced and native ranges, a pattern explained by a history of multiple
41 introductions of specimens from different parts of the world. At last, occurrence of identical
42 AFLPs genotypes (i.e. clones) in several population confirmed asexual reproduction, but
43 recombination was also revealed by gametic equilibrium analysis in some populations
44 suggesting that *P. corethrurus* L1 may have a mixed reproductive strategy.

45

46

47 **Introduction**

48 Many species have been reported as having extremely widespread distributions. Among
49 these cosmopolitan species, earthworms are poorly represented except for a few taxa. Because
50 of the limited dispersal ability of earthworms it is even striking that out of approximately 5358
51 described species according to the online Drilobase database, about 120 species are considered
52 peregrine i.e. widely ranging, often owing to human actions (Blakemore 2009; Csuzdi 2012;
53 Hendrix et al. 2008). Over the past centuries, global commerce has favored transport of
54 temperate more than tropical earthworm species but Hendrix et al. (2008) alerted that the
55 emergence of new tropical invasive species is likely. Insights into the invasiveness of
56 earthworms in tropical regions may be gained by examining the specific attributes responsible
57 for the success of the most well-known tropical peregrine earthworm, *Pontoscolex corethrurus*
58 (Müller 1857).

59 This pan-tropical earthworm is believed to have originated from the Guayana Shield in
60 South America (Righi 1984) which includes Guyana, Suriname, French Guiana, most of
61 Southern Venezuela, as well as parts of Colombia and Brazil. A recent phylogenetic study using
62 two mitochondrial (cytochrome c oxidase I and 16S-rDNA genes) and two nuclear (internal
63 transcribed spacer 2 and 28S-rDNA) markers, in addition to a large-scale multilocus sequence,
64 revealed the high complexity of the genus *Pontoscolex* (Taheri et al. 2018a). The commonly
65 treated as a unique entity *P. corethrurus* was shown to correspond to a complex of four distinct
66 species (i.e., defined as L1, L2, L3 and L4). Two highly divergent genetic lineages (L1 and L4)
67 were already described by Cunha et al. (2014), using sequences of two mitochondrial genes
68 and Amplified Fragment Length Polymorphism (AFLP) markers, in three populations of the
69 São Miguel Island in the Azores archipelago. Taheri et al. (2018a) revealed that *P. corethrurus*
70 L1 was the most widespread lineage in the world. It was present in the location where this
71 species was first described by Fritz Müller in 1857 (Müller 1857), suggesting that L1

72 corresponds to *P. corethrurus sensu stricto*. Most published studies in the past referencing “*P.*
73 *corethrurus*” have probably dealt with the *P. corethrurus* L1 species, although the risk to have
74 investigated another lineage is not null. *P. corethrurus* was described as an endogeic
75 earthworm, tolerant to a wide range of biotic and abiotic environmental conditions (review in
76 Taheri et al. 2018b), and reproducing parthenogenetically, although the possibility of sexual
77 reproduction could not be ruled out (Dupont et al. 2012; Gates 1973). The mode of
78 parthenogenesis in *P. corethrurus* is not well known. It is generally accepted that
79 parthenogenetic earthworms are automictic (i.e. mode of parthenogenesis that retains meiosis)
80 with a premeiotic doubling of the chromosome number, followed by regular meiosis which
81 restore diploidy in the following egg (Diaz Cosin et al 2011). Because replicated chromosomes
82 pair prior to meiosis I, the offspring are genetically identical to the mother (such as in apomictic
83 parthenogenesis with suppression of meiosis), and heterozygosity is maintained (Simon et al.
84 2003, Lutes 2010). In other types of automictic parthenogenesis, diploidy is restored by
85 duplication or fusion of the female gamete and it rapidly leads to complete homozygosity
86 (Simon et al. 2003).

87 Several peregrine earthworm species are known to have a parthenogenetic reproduction
88 often accompanied by polyploidy, in addition to some other advantageous traits such as high
89 fecundity, small size, resistant cocoons, wide and rapid dispersal, feeding diet plasticity
90 (reviewed in Fernandez et al. 2011a) and resistance to drought (Bartz, Brown, Decaëns &
91 Lapiéd personal observations) and to waterlogged soils (Bartz personal observation). Two main
92 features allow parthenogenetic species to invade and establish in new habitats faster than
93 sexually reproducing species; (i) the potential increase of the population in numbers per
94 generation is double that of an amphimictic species and (ii) the ability of a single individual to
95 establish a new colony (Cuellar 1977). However, the lack of recombination makes asexual
96 lineages prone to accumulate deleterious mutations that limit their evolutionary potential.

97 Mixed reproductive systems can combine the advantages of both sexual and asexual
98 reproduction in terms of genetic load reduction, fixation of beneficial mutations, and adaptation
99 to fluctuating environments (Bazin et al. 2014). In earthworms, mixed reproductive systems,
100 combining parthenogenesis and amphimixis, is only known in the peregrine species
101 *Aporrectodea trapezoides*. Although the most abundant clone of *A. trapezoides* was showed to
102 be purely parthenogenetic, two pseudogametic (i.e. parthenogenetic reproduction requiring
103 sperm to stimulate embryogenesis without incorporating the parental DNA) populations and
104 two populations where amphimictic and parthenogenetic forms coexist were discovered (De
105 Sosa et al. 2017; Fernandez et al. 2011b).

106 Population genetics tools can contribute significantly to the understanding of the
107 reproductive strategies of peregrine species (e.g. Dupont et al. 2007). The elucidation of genetic
108 structure patterns and variation within and among populations in distant locations has also the
109 potential to provide insights into the geographic origin, migration pathways and colonization
110 history (e.g. single vs multiple introductions) of peregrine species (Sakai et al. 2001). For
111 instance, molecular markers have proved to be powerful tools for inferring the evolutionary
112 history of *A. trapezoides* (Fernandez et al. 2011a; Fernandez et al. 2016) which has Palearctic
113 origin but a current worldwide distribution. Analyzing two mitochondrial and two nuclear
114 markers, Fernandez et al (2011a) revealed a relatively high clonal diversity in this species
115 whereas one clone was shared by almost one third of the sampled individuals and was widely
116 distributed. Such “super-clones” are thought to be “general-purpose genotypes” (Lynch 1984)
117 which show a broad ecological tolerance resulting from interclonal selection in temporally
118 variable environments (Vorburger et al. 2003).

119 To date, only two studies focused on *P. corethrurus* population genetics and were of
120 limited geographical scale. Dupont et al (2012) investigated the genetic variability of *P.*
121 *corethrurus* in its presumed native range using AFLPs. Among the six populations from French

122 Guiana, a higher level of genetic diversity was found in populations originating from the most
123 disturbed sites while a higher number of clones was observed in the other populations. In the
124 São Miguel Island in the Azores archipelago, Cunha et al. (2014) found a higher intra-lineage
125 genetic diversity in the population of *P. corethrurus* L1 living in pineapple greenhouses, most
126 probably because of repeated introductions. Here, our aim was to achieve a large-scale
127 population genetics study of *P. corethrurus* L1, investigating samples coming from its
128 presumed native and introduced ranges, and using cytochrome oxidase I (COI) mitochondrial
129 gene, in addition to AFLPs markers. The genetic variability and the genetic differentiation of
130 the different *P. corethrurus* L1 populations were studied in order to test the following
131 hypothesis: (i) a “super-clone” has invaded most of the introduced areas, (ii) the clonal diversity
132 is lower in the introduced range than in the native one because mutations did not have time to
133 accumulate, and (iii) this species has a mixed reproductive strategy.

134 2. Methods

135

136 2.1. Mitochondrial DNA analysis

137 We first assessed genetic diversity and its distribution among *P. corethrurus* L1 by re-
138 analyzing the extensive global dataset of cytochrome oxidase I (COI) sequences published in
139 Taheri et al (2018a). Of the 662 COI sequences in Taheri et al (2018a) study, 478 belonged to
140 *P. corethrurus* L1 specimens coming from a total of 49 locations belonging to five
141 biogeographical realms (Genbank accession and information on the samples in supplementary
142 data Table S1). In order to investigate the relationships among the different L1 haplotypes, a
143 haplotype network was constructed based on derived haplotypes in Network v. 5.0.0.1, using
144 median joining calculations (Bandelt et al. 1999).

145 For the analysis of genetic diversity at the population level, we selected 12 populations
146 with a minimum sampling size of 9 individuals (Table 1, 269 COI sequences); 5 populations
147 came from the presumed native range of the species (i.e. the Guayana Shield) and more
148 precisely from French Guiana [Cayenne (CAY), Mitaraka (MIT), Pararé A (PARA), Pararé B
149 (PARB) and Pararé C (PARC)] and 7 other populations came from the presumed introduced
150 range represented by north and south of Brazil [Caxiuana (CAX)], Belem (BEL), Joinville
151 (JNL), Orleans (ORL)], Mexico [Tlalcotlen (TLC)], Gabon [La Lopé (LOP)] and Thailand
152 [Chachoengsao (CHG)]. For these populations, number of polymorphic sites, number of
153 haplotypes, haplotype (gene) diversity and nucleotide diversity (P_i ; Nei 1987) were calculated,
154 using DnaSP v. 5 software (Librado and Rozas 2009). The historical stability of populations
155 was assessed with a test of neutrality (Tajima's D , Tajima 1989) that also served as indicator
156 of population expansion in DnaSP. We quantified the amount of genetic differentiation of
157 population groups using a hierarchical Analysis of Molecular Variance (AMOVA)
158 implemented in Arlequin V3.5 (Excoffier & Lischer, 2010). In this analysis, the COI data set
159 was partitioned at three levels: groups of populations from native *versus* introduced range,

160 among-populations within groups and among all populations. One thousand random
161 permutations were used to infer the significance of the variance components.

162

163 **2.2. Nuclear DNA analysis using AFLPs**

164 *2.2.1 AFLP procedure*

165 The genetic variation of the 12 previously described populations was also investigated
166 using AFLP markers. DNA extraction was performed as described in Taheri et al (2018a).
167 AFLP analysis of 226 specimens was carried out according to Vos et al. (1995) with a few
168 modifications: approximately 50 ng/ μ l of purified genomic DNA of each specimen was
169 digested with two digestive enzymes. The first digestion was done in 10 μ l with Taq1 (20U,
170 New England BioLabs, Ipswich, MA, USA (NEB)), Buffer Taq1 (10X, NEB), BSA (1mg.ml-
171 1, NEB) and incubated for 1h30 at 65°C by thermal cycler (T100TM, BIO-RAD Laboratories
172 Inc., Foster city, CA, USA). The second digestion of the DNA from the last step was done in
173 15 μ l with a solution of: EcoRI (20U, NEB), Buffer EcoRI (10X, NEB), BSA (1mg.ml-1, NEB)
174 and incubated for 1h30 at 37°C. We verified DNA digestion quality by gel electrophoresis and
175 negative controls were run at each time. Adapters were ligated using 50pmoles μ l⁻¹ of double-
176 stranded Taq1 adapter (Taq top 5'-GACGATGAGTCCTGAC and Taq bottom 5'-
177 CGGTCAGGACTCAT, Eurofins Genomics, Germany) and 5pmoles μ l-1 of double-stranded
178 EcoRI adapter (Eco top 5'-CTCGTAGACTGCGTACC-3' and Eco bottom 5'-
179 AATTGGTACGCAGTCTAC-3', Eurofins Genomics, Germany), T4 DNA ligase buffer (10X,
180 Promega, Madison, WI, USA), ATP (10mM, NEB), BSA (1mg.ml-1, NEB), T4 DNA ligase
181 (3U, Promega). Samples with a total volume of 50 μ l (prepared solution + digested DNA from
182 previous step + adjusted water) was then incubated for 3h at 37°C. Digestion-ligation
183 production of 20 μ l was diluted with 40 μ l of AE buffer (QIAGEN Sciences, Maryland, USA).
184 Then each dilution was divided in 3 separate samples of 15 μ l each. Pre-selective PCRs

185 contained 5pmoles of E01 primer (GACTGCGTACCAATTCA) and 5pmoles of T01 primer
186 (GATGAGTCCTGACCGAA), MgCl₂ (25mM, Promega), dNTPs (10mM, Invitrogen™, Life
187 technologies, Carlsbad, CA, USA), GoTaq buffer (5X, Promega), DNA Taq polymerase (5U,
188 Promega), adjusted with water to have a total volume of 50µl in each sample (added to diluted
189 ligation product). The PCR pre-selective was done by an initial denaturation step at 94°C for
190 2min, followed by annealing setup of 30 cycles containing; 94°C for 30s, 56°C for 1 min and
191 72°C for 60s, and finally elongation setup of 72°C for 10 minutes. After this setup, the three
192 sub samples were reassembled in one, and then 1µl of each sample was diluted in 19µl of AE
193 buffer (QIAGEN). Selective PCR reactions contained 5µl of pre-amplified DNA with 15µl of
194 a solution containing MgCl₂ (25mM, Promega), dNTPs (10mM, Invitrogen), GoTaq buffer (5X,
195 Promega), 5pmoles.µl⁻¹ T32 primer (5'-GATGAGTCCTGACCGAAAC-3') or 5pmoles.µl⁻¹
196 T38 primer (5'-GATGAGTCCTGACCGAAACT-3'), DNA Taq polymerase (5U, Promega)
197 and one primer combination E32-FAM was used (5'-GACTGCGTACCAATTCAA-3'). A
198 touchdown thermal cycling (PTC-100) started with denaturing setup of 94°C for 2 min,
199 following by 9 cycles; 94°C for 30s, 65°C for 30s with 1°C diminution per cycle, 72°C for 60s,
200 following by 26 cycles containing 94°C for 30s, 56°C and 72°C during 1min and finally 72°C
201 for 10 minutes. After each setup DNA solutions were centrifuged by ROTANTA 460R (Hettich
202 Lab Technology, Tuttlingen, Germany). Amplified products were mixed with formamide
203 (Hi_Di™, Applied Biosystems, Foster city, CA, USA (AB) and a Genescan™-500LIZ™
204 (AB) size standard (9.5µl of formamide and 0.5 µl of Genescan™-500LIZ™ for 2µl of
205 amplified product). Fragments were separated on an ABI PRISM™ 3130 Genetic Analyzer
206 (platform INSERM, Henri Mondor hospital, Créteil). Raw data were visualized, and the
207 fragments manually scored using Genemapper V5 (Applied Biosystem) software. Processed
208 data were exported as presence/absence matrix.

209 Genetic markers, including AFLPs, can be prone to genotyping errors with various
210 potential sources (Bonin et al. 2004). Therefore, we estimated the genotyping error rate of our
211 dataset by re-genotyping and blind scoring of 54 randomly chosen individuals. The error rate
212 was calculated by mismatch error rate based on the formula proposed by Bonin et al. (2004),
213 multiplied by the total number of markers.

214

215 *2.2.2 Genetic diversity analysis*

216 Genetic diversity statistics, including number of different genotypes by considering the
217 ‘error rate’ calculated from replicates (i.e. 1.16), genotype diversity, gene diversity, proportion
218 of variable markers and the frequency down-weighted marker value (DW, Schonswetter et al.
219 2005) that indicates the rarity of each marker within the dataset were calculated using AFLPdat
220 program (Ehrich 2006). Genotype and gene diversity were calculated based on Nei’s formula
221 (Nei 1987). Similar genotypes within locations were removed and further analyses on *P.*
222 *corethrurus* L1 were carried out on a dataset without clones.

223

224 *2.2.3 Multilocus gametic disequilibrium*

225 In order to evaluate the evidence for recombination, measures of multilocus gametic
226 disequilibrium were calculated and tested for significance with 500 randomizations in
227 Multilocus software (<http://www.bio.ic.ac.uk/evolve/software/multilocus/>), using 3 different
228 methods. First, we performed a character compatibility analysis. In a data set with a bi-allelic
229 marker with presence or absence (1 or 0) at each locus such as AFLPs, the presence of all four
230 possible combinations of alleles at two different loci (e.g. 0/0, 1/0, 0/1, 1/1) is an indirect signal
231 of recombination. In contrast, a hierarchical data structure (e.g. 0/0, 0/0, 1/1, 1/1) fits to a model
232 of mutations that are vertically transmitted and accumulated within a clonal lineage. The
233 proportion of compatible pairs of loci was thus computed to probe the predominant mating

234 system in the populations. Moreover, the index of association (I_A) and an alternative measure
235 of index of association that is less sensitive to the number of loci (\bar{r}_d , Agapow and Burt 2001)
236 were computed. We further tested for gametic disequilibrium based on the distribution of allelic
237 mismatches between pairs of genotypes over all loci using an exact test implemented in
238 Arlequin. To adjust for multiple comparisons, the SGoF method (Carvajal-Rodriguez et al.
239 2009) as implemented in the software Myriad
240 (<http://myriads.webs.uvigo.es/MyriadsReadme.htm>) was applied.

241

242 *2.2.4 Genetic structure among populations*

243 Analysis of molecular variance (AMOVA) was conducted in Arlequin V3.5 in the same
244 way as for the COI dataset. In addition, an unbiased estimate of differentiation among
245 populations, $\theta^{(II)}$ was obtained using the Bayesian method proposed by Holsinger et al. (2002)
246 and implemented in the software Hickory v1.1. The data were run with the default parameters
247 using the f-free model.

248 To illustrate the relationships among populations, a split network was constructed using
249 the software Splitstree version 4.1.4.6 (Huson and Bryant 2006) on AFLP profiles. We used the
250 distance-based Neighbor-Net (N-net) method for construction of networks. The N-net provides
251 good visualization of the data when it presents complex evolutionary steps or reticulate
252 relationship among genotypes (Huson and Bryant 2006). The networks were constructed based
253 on Nei's distance (GD) matrix between populations calculated with a Bayesian method using
254 AFLP-Surv version 1.0 (Vekemans et al. 2002) with non-uniform distribution by assuming
255 deviation from Hardy-Weinberg equilibrium; F_{is} values were estimated by Hickory software
256 using the full model ($f = 0.033$, Holsinger et al. 2002). Analyses were done with 1000
257 permutations and 1000 bootstrap values.

258

259 3. Results

260

261 3.1 COI genetic variation

262 An extensive global dataset of COI sequences obtained in L1 specimens was used to
263 assess the distribution of mitochondrial genetic diversity among populations. Thirteen COI
264 haplotypes were inferred from 478 sequences within 49 locations. COI haplotype network is
265 shown in Figure 1. Haplotype one was shared by most individuals i.e., 364 out of 479, followed
266 by haplotype two with 64 individuals. Five haplotypes (i.e., H3, H7, H8, H9, and H10) were
267 exclusively found in the presumed native range while six haplotypes were found exclusively in
268 the presumed introduced range (H4, H5, H6, H11, H12 and H13). In 5 populations (PARA,
269 PARB, PARC, ORL and LOP), no polymorphic site was observed. In the native range, the
270 highest values of COI genetic diversity indices were obtained in the MIT population (Table 1)
271 with 10 polymorphic sites and a haplotype diversity of 0.693. In the introduced range, the
272 highest COI genetic diversity was observed in CAX (11 polymorphic sites, haplotype diversity
273 of 0.561). No difference of mean haplotypic diversity was observed between the native (mean
274 gene diversity = 0.24, SD = 0.46) and the introduced (mean gene diversity = 0.26, SD = 0.32)
275 ranges (2-tailed test, $t = -0.07$, $p = 0.943$).

276 The null hypothesis of neutrality was rejected due to significant negative value in the
277 BEL population indicating a recent population expansion in this location. Positive and
278 significant Tajima's D tests were obtained in the CAY and MIT populations indicating
279 population sub-structure in these sites. AMOVA analysis showed that 64 % of the genetic
280 variation was partitioned among all populations, 25 % was attributed to differences among
281 populations within groups, and 11 % of the variation was due to differences between native and
282 introduced groups (Table 2). The genetic differentiation among the native and introduced

283 groups (Φ_{CT}) was not significant while the genetic differentiation among all populations was
284 important ($\Phi_{ST} = 0.361$, $P < 0.001$).

285

286 **3.2 AFLPs genetic variation**

287 A total of 394 AFLPs loci were scored and among them 318 were polymorphic. The
288 difference between the number of individuals analyzed using AFLPs and the number of
289 different AFLP genotypes in the dataset indicates the number of clones in each population. High
290 number of clones (85%) were found in LOP location (Gabon) with 17 similar genotypes out of
291 20 individuals (Table 1). Important proportion of clones was also observed in CHG (Thailand)
292 with 10 similar genotypes out of 35 individuals (29%) and in CAX (Brazil) with 4 similar
293 genotypes out of 18 individuals (22%). In five populations i.e., CAY, PARC, BEL, JNL, and
294 TLC, no clones were found. The highest rates of rare alleles were observed in CAY, MIT, CAX
295 and LOP with DW values of 1804, 830.7, 708.9 and 692.7, respectively. The smallest DW
296 values were found in JNL and PARA with 354.1 and 362.2 values respectively.

297 Further analyses were done on data set without clones, reducing the dataset from 226
298 profiles to 185. The highest gene diversity was found in JNL, CAY and PARC (0.168, 0.165
299 and 0.163 respectively) while the lowest values were obtained in LOP, PARA, PARB and ORL
300 (0.012, 0.052, 0.083 and 0.087 respectively, Table 1). No difference of mean gene diversity
301 was observed between the native (mean gene diversity = 0.12, SD = 0.01) and the introduced
302 (mean gene diversity = 0.11, SD = 0.01) ranges (2-tailed test, $t = 0.269$, $p = 0.793$).

303 All indices of multilocus gametic disequilibrium were congruent (Table 1). All
304 populations showed significant gametic disequilibrium and character compatibility (except
305 LOP but the analysis couldn't be correctly carried out in this population because an extremely
306 low polymorphism) as expected in an asexually reproducing parthenogenetic species. However,
307 the values of these indices were highly different among populations. Thus, the TLC population

308 showed a particularly high value of index of association ($I_A = 42.52$) while the lowest significant
309 value was obtained in JNL ($I_A = 2.23$).

310 The genetic differentiation Φ_{ST} measured by AMOVA analysis was 0.364; 64 % of the
311 variation was partitioned within populations, 31 % was attributed to differences among
312 populations within groups, and 5 % of the variation was due to differences between the native
313 and introduced groups. All three hierarchical levels were significant with $P < 0.05$ (Table 2) but
314 the genetic differentiation among groups was low ($\Phi_{CT} = 0.052$, $P = 0.001$). A similar level of
315 genetic differentiation among populations was obtained using the f-free model with Hickory
316 software: $\theta^{(II)} = 0.342$. Relationships between populations are illustrated by the Neighbor-Net
317 results obtained from Nei's genetic distances (Fig. 2, supplementary data Table S2): LOP and
318 CAY were the most differentiated to each other and to the other locations while the sample
319 from Asia (CHG) was genetically close to the populations from South America. Moreover, all
320 populations from the native range were grouped together.

321

322 **Discussion**

323 *The “super-clone” hypothesis*

324 Despite the very extensive geographical range of *P. corethrurus* in the tropical and sub-
325 tropical zone, and the limited potential for natural dispersal of earthworms (Costa et al., 2013),
326 our sampling of Africa, Asia, South America and Central America and various islands revealed
327 a low global mitochondrial haplotypic diversity with a total of 13 COI haplotypes and, among
328 them, three major haplotypes (H1, H2 and H3). These haplotypes were not very divergent: H1
329 was only separated from H2 by four mutations and by six mutations from H3. The haplotype
330 H1 was observed in all continents and was in high frequency in the introduced populations. All
331 populations from the presumed native range showed the H1 haplotype and it was the sole
332 haplotype in three of them (PARA, PARB and PARC). The haplotype H2 was also frequent in
333 both native and introduced ranges but was absent from Africa while the haplotype H3 was only
334 observed in two populations (MIT and CAX). These results could suggest that H1 and H2
335 correspond to “super-clones” such as those observed in aphids. For instance, population studies
336 of the parthenogenetic aphid *Myzus persicae* revealed that some genotypes are widespread both
337 in time and space (Vorburger et al. 2003). It was proposed that these predominant genotypes
338 possess general-purpose genotypes, i.e. genotypes with broad ecological tolerances as a result
339 of interclonal selection in temporally variable environments (Lynch 1984). In the case of *P.*
340 *corethrurus*, a successful invasive earthworm species that has a pan-tropical distribution, this
341 postulate was appealing and was tested through our population genetic study using AFLP
342 markers. Although clones were observed in several populations and sometimes in high
343 frequency, no AFLP profile was shared among populations. For instance, the LOP population
344 that had the highest number of clones, had a high DW value, meaning that the AFLP genotypes
345 within this population were highly different from the other genotypes of the total dataset.
346 Overall, AFLP results allow to reject the hypothesis that a “super-clone” was responsible for

347 the success of *P. corethrurus* invasion. The most likely factors that may be suggested to explain
348 the lack of COI haplotype diversity in *P. corethrurus* populations are a slow evolutionary rate
349 of the COI gene (i.e. constraint on the ability of the mitochondrial genome to accumulate
350 mutations) or, more likely, genetic drift associated with historical events, such as recent
351 population bottlenecks or founder effects (e.g. López-Legentil & Turon, 2007). The fourfold
352 higher effective population size for the nuclear genome compared to mitochondrial loci, which
353 results in genetic drift being four times weaker for nuclear loci may explain the discrepancy
354 between the mitochondrial genetic depletion and the relatively high genetic variability of the
355 nuclear loci.

356 ***Genetic variation in introduced versus native range***

357 The number of COI haplotypes was slightly higher in the presumed introduced range
358 per comparison with the presumed native range (9 and 7 respectively), and comparison of mean
359 haplotypic diversity values between populations from the native and introduced ranges showed
360 no significant difference. This is explained by the fact that the COI haplotypic diversity was
361 highly variable among populations. Within the native range, several populations did not show
362 any COI polymorphism while other populations displayed some of the highest values of COI
363 diversity of the dataset (Table 1, Fig 1). Similar patterns of discordant genetic diversity among
364 populations from the native range was observed with the nuclear markers; AFLPs gene diversity
365 was two to three times higher in CAY and PARC populations per comparison with PARA and
366 PARB. The congruence between AFLPs and mitochondrial data suggests that the low COI
367 polymorphism of this species is not due to a marker-specific effect but most probably due to
368 population and species history.

369 Strong genetic drift effects are expected when populations are founded by a small
370 number of colonists (Sakai et al. 2001). Our data for *P. corethrurus* fit this theoretical
371 expectation in the LOP population where an extremely low AFLP genetic diversity was

372 observed, congruent with the lack of COI polymorphism (Table 1). Except the LOP population,
373 which was probably very recently founded, levels of genetic variation measured from AFLP
374 data in introduced populations were comparable to those found in populations from the
375 presumed native range (Table 1). For instance, the JNL population (Brazil) showed high gene
376 diversity (0.168) comparable to those of CAY (0.165). Moreover, no enhanced genetic
377 differentiation, except for the specific case of LOP, was observed in the introduced range
378 although repeated founder effects are likely to increase genetic divergence among populations
379 (Slatkin 1977). The absence of reduced genetic diversity in most of the introduced range is also
380 certainly a result of *P. corethrurus* rapid population expansion (e.g. Zenger et al. 2003). The
381 signature of population expansion is, for instance, still visible in the population of Belem where
382 a negative and significant Tajima's D value was obtained. Moreover, it is noteworthy that the
383 native range of the species is not known with certainty. In particular, the case of Northern
384 Brazilian sites (CAX and BEL) is tricky. We hypothesized that they correspond to the
385 introduced range because they do not belong to the Guayana Shield but they might also
386 correspond to the native area.

387 Our results suggest that successive introduction from multiple sources has enhanced
388 genetic diversity in not only some introduced localities but also in the presumed native range.
389 Significant sub-structures were indeed revealed by positive and significant Tajima's D value in
390 two populations of the native range (i.e. CAY and MIT). A previous study has already showed
391 that the suburban environment of Cayenne, where the CAY population was sampled, is prone
392 to constant introduction of individuals (Dupont et al. 2012). In the locality of Mitaraka (MIT
393 population) situated in the Amazonian primary forest where human activity is extremely
394 limited, the explanation could be completely different. In this locality, specimens were collected
395 in five different sub-localities (supplementary data Table S1) separated by a maximum of 3 km.
396 This distance is therefore probably sufficient to detect sub-structure using AFLP markers,

397 highlighting the low active dispersal of the species. At the opposite, in more disturbed sites,
398 passive dispersal linked to human activities is probably the predominant mode of spread
399 (Dupont et al. 2012). In South America, successive introduction and rapid expansion through
400 passive dispersal could be occurring since humans arrived and dispersed in Amazonia (Cunha
401 et al. 2016).

402

403 ***The mixed reproductive strategy of Pontoscolex corethrurus sensu stricto***

404 *P. corethrurus* has several life history traits that predispose it to invasiveness (Taheri et
405 al. 2018b). Among them, the ability of a single individual to establish a population is an
406 important characteristic of several invasive species. (Dybdahl and Drown 2011). The important
407 number of clones revealed in some populations by the AFLP analysis highlights the prominence
408 of parthenogenetic reproduction in this species. For instance, in LOP only 15% of the AFLP
409 genotypes were different. The importance of asexual reproduction is also confirmed by the high
410 levels of gametic disequilibrium (Table 1) that were observed in several populations such as
411 TLC ($I_A = 42.52$; $\bar{r}_d = 0.39$), CAX ($I_A = 22.34$; $\bar{r}_d = 0.12$), PARB ($I_A = 18.51$; $\bar{r}_d = 0.21$) and
412 ORL ($I_A = 13.37$; $\bar{r}_d = 0.16$).

413 Although asexual reproduction provides an immediate demographic advantage, it is
414 expected that in the long term, asexual populations deteriorate in fitness due to mutation
415 accumulation or the inability to cope with changing environment (Lynch et al. 1993; Muller
416 1932). Thus, asexual lineages depend on occasional sexual recombination to add variation to
417 their genomes, purge deleterious mutations, and adapt to changing environments. A mixed
418 strategy of asexual reproduction with low level of sex combines the best of both strategies
419 (Combosch and Vollmer 2013; Hurst and Peck 1996). It has been previously suggested that *P.*
420 *corethrurus* was capable of such a mixed reproductive strategy (Dupont et al. 2012; Gates 1973;

421 Taheri et al. 2018b) and this hypothesis was tested here using the AFLP genotypes. We showed
422 that low values of gametic disequilibrium estimators in some populations (e.g. $I_A = 2.23$; $\bar{r}_d =$
423 0.01 and $I_A = 2.54$; $\bar{r}_d = 0.01$ in JNL and PARC populations respectively, Table 1) associated
424 to an absence of clones revealed the occurrence of recombination within these populations.
425 Meiotic recombination occurs during amphimictic reproduction and automictic
426 parthenogenesis. Because recombination occurs between sister chromosomes in the premeiotic
427 doubling of chromosome happening before the regular meiosis of the automictic
428 parthenogenesis (i.e. the type of parthenogenesis described in earthworms), gametic
429 disequilibrium is not expected in this case. Although we cannot exclude that another type of
430 automictic parthenogenesis occurs in *P. corethrurus*, our results seem to support the hypothesis
431 previously proposed of a plastic reproduction strategy in *P. corethrurus*.

432

433 **Conclusion**

434 Successive founder effects and demographic bottlenecks are most probably responsible
435 of the low COI genetic diversity revealed in our worldwide scale study. Although mtDNA
436 contain historic genetic mutations and is suitable for resolving taxonomic uncertainties, AFLP
437 markers are more appropriate for inferring contemporaneous genetic patterns. Comparing
438 results obtained using both kind of markers revealed that the founder event in LOP was
439 probably more recent than in the other populations where COI polymorphism was also null
440 (PARA, PARB, PARC and ORL). A major question in invasion biology is how newly founded
441 and subsequently isolated populations overcome the detrimental effects of low genetic
442 diversity, the so-called ‘genetic paradox’ (Allendorf and Lundquist 2003; Roman and Darling
443 2007). The establishment success of newly founded *P. corethrurus* populations may be due to
444 recurrent introductions from multiple sources, accumulation of new mutations and high level
445 of intrinsic phenotypic plasticity (i.e. Taheri et al. 2018b), in addition to a mixed reproductive

446 strategy. Asexual reproduction improves *P. corethrurus* ability to reach high abundance (e.g.
447 Taheri et al. 2018b) and is a particularly important trait in the first stage of the invasion, when
448 sexual partners are rare. During establishment and expansion of the population, events of sexual
449 reproduction will then promote the genetic diversity. Such a mixed reproductive strategy is a
450 way to purge the genome of deleterious mutations (i.e. no clonal population senescence after
451 many generations), elegantly defining a mode of escape to Muller's ratchet (Muller 1932).

452

453 **Acknowledgements**

454 S. Taheri acknowledge the SIE Doctoral School of the University Paris Est Créteil (France).
455 Sampling was supported by the CNRS Nouragues call and labex CEBA through the strategic
456 project DIADEMA, the Réserve de la Trinité and Our Planet Reviewed (in French Guyana),
457 CNPq, NERC, EU, Fundação Araucária, the Newton Fund and FAPESC (in Brazil), National
458 Science Foundation, USA (in Taiwan), the ECOTROP consortium and the Agence Nationale
459 des Parcs Nationaux (in Gabon), the International Joint Laboratory LUSES (in Thailand). We
460 would like to thank the Fritz Müller Museum of Blumenau for allowing collecting on their
461 property, and I. Barois, C.-H. Chang, M. Coulis, S. James, E. Lapied, and P. Lavelle, for
462 sampling specimens in the different countries covered by the COI study (dataset already
463 published in a previous paper and re-used here).

464

465 **References**

- 466 Agapow PM, Burt A (2001) Indices of multilocus linkage disequilibrium. *Mol Ecol Notes*
467 1:101-102. <https://doi.org/10.1046/j.1471-8278.2000.00014.x>
- 468 Allendorf FW, Lundquist LL (2003) Introduction: Population biology, evolution, and control
469 of invasive species. *Conserv Biol* 17:24-30. <https://doi.org/10.1046/j.1523-1739.2003.02365.x>
- 470 Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific
471 phylogenies. *Mol Biol Evol* 16:37-48. <https://doi.org/10.1093/oxfordjournals.molbev.a026036>
- 472 Bazin E, Mathe-Hubert H, Facon B, Carlier J, Ravigne V (2014) The effect of mating system
473 on invasiveness: some genetic load may be advantageous when invading new environments.
474 *Biol Invasions* 16:875-886. <https://doi.org/10.1007/s10530-013-0544-6>
- 475 Blakemore RJ (2009) Cosmopolitan earthworms—a global and historical perspective. In: Shain
476 DH (ed) *Annelids in modern biology*. Wiley-Blackwell, Hoboken, pp 257-283
- 477 Bonin A, Bellemain E, Eidesen PB, Pompanon F, Brochmann C, Taberlet P (2004) How to
478 track and assess genotyping errors in population genetics studies. *Mol Ecol* 13:3261-3273
- 479 Carvajal-Rodriguez A, de Una-Alvarez J, Rolan-Alvarez E (2009) A new multitest correction
480 (SGoF) that increases its statistical power when increasing the number of tests. *Bmc*
481 *Bioinformatics* 10. <https://doi.org/20910.1186/1471-2105-10-209>
- 482 Combosch DJ, Vollmer SV (2013) Mixed asexual and sexual reproduction in the Indo-Pacific
483 reef coral *Pocillopora damicornis*. *Ecol Evol* 3:3379-3387. <https://doi.org/10.1002/ece3.721>
- 484 Costa D, Timmermans MJTN, Sousa JP, Ribeiro R, Roelofs D, Van Straalen NM (2013)
485 Genetic structure of soil invertebrate populations: collembolans, earthworms and isopods. *Appl*
486 *Soil Ecol* 68: 61–66.
- 487 Csuzdi C (2012) Earthworm species, a searchable database *Opuscula Zoologica Budapest*
488 43:97-99

489 Cuellar O (1977) Animal parthenogenesis. Science 197:837-843.
490 <https://doi.org/10.1126/science.887925>

491 Cunha L et al. (2016) Soil animals and pedogenesis: the role of earthworms in anthropogenic
492 soils. Soil Science 181:110-125. <https://doi.org/10.1097/ss.0000000000000144>

493 Cunha L, Montiel R, Novo M, Orozco-terWengel P, Rodrigues A, Morgan AJ, Kille P (2014)
494 Living on a volcano's edge: genetic isolation of an extremophile terrestrial metazoan. Heredity
495 112:132-142. <https://doi.org/10.1038/hdy.2013.84>

496 De Sosa I, Marchan DF, Novo M, Diaz Cosin DJ, Giribet G, Fernandez R (2017) Insights into
497 the origin of parthenogenesis in oligochaetes: Strong genetic structure in a cosmopolitan
498 earthworm is not related to reproductive mode. Eur J Soil Biol 81:31-38.
499 <https://doi.org/10.1016/j.ejsobi.2017.06.003>

500 Diaz Cosin DJ, Novo M, Fernandez R (2011.) Reproduction of earthworms : sexual selection
501 and parthenogenesis. In: Karaca A (ed) Biology of earthworms. Springer-Verlag Berlin
502 Heidelberg, pp 69–86

503 Dominguez J, Velando A, Aira M, Monroy F (2003) Uniparental reproduction of *Eisenia fetida*
504 and *E andrei* (Oligochaeta : Lumbricidae): evidence of self-insemination. Pedobiologia 47:530-
505 534

506 Dupont L et al. (2012) Genetic signature of accidental transfer of the peregrine earthworm
507 *Pontoscolex corethrurus* (Clitellata, Glossoscolecidae) in French Guiana. Eur J Soil Biol 53:70-
508 75

509 Dupont L, Viard F, David P, Bishop JDD (2007) Combined effects of bottlenecks and selfing
510 in populations of *Corella eumyota*, a recently introduced sea squirt in the English Channel.
511 Divers Distrib 13:808-817

512 Dybdahl MF, Drown DM (2011) The absence of genotypic diversity in a successful
513 parthenogenetic invader. *Biol Invasions* 13:1663-1672. <https://doi.org/10.1007/s10530-010->
514 9923-4

515 Ehrich D (2006) AFLPDAT: a collection of R functions for convenient handling of AFLP data.
516 *Mol Ecol Notes* 6:603-604. <https://doi.org/10.1111/j.1471-8286.2006.01380.x>

517 Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform
518 population genetics analyses under Linux and Windows. *Mol Ecol Res* 10: 564-567.
519 <https://doi.org/10.1111/j.1755-0998.2010.02847.x>

520 Fernandez R, Almodovar A, Novo M, Gutierrez M, Cosin DJD (2011a) A vagrant clone in a
521 peregrine species: Phylogeography, high clonal diversity and geographical distribution in the
522 earthworm *Aporrectodea trapezoides* (Duges, 1828). *Soil Biol Biochem* 43:2085-2093

523 Fernandez R, Bergmann P, Almodovar A, Cosin DJD, Heethoff M (2011b) Ultrastructural and
524 molecular insights into three populations of *Aporrectodea trapezoides* (Duges, 1828)
525 (Oligochaeta, Lumbricidae) with different reproductive modes. *Pedobiologia* 54:281-290.
526 <https://doi.org/10.1016/j.pedobi.2011.04.003>

527 Fernandez R, Novo M, Marchan DF, Cosin DJD (2016) Diversification patterns in
528 cosmopolitan earthworms: similar mode but different tempo. *Mol Phylogenet Evol* 94:701-708.
529 <https://doi.org/10.1016/j.ympev.2015.07.017>

530 Gates GE (1973) Contributions to a revision of the earthworm family Glossoscolecidae. I.
531 *Pontoscolex corethrurus* (Müller, 1857). *Bulletin of Tall Timbers research station* 14:1-12

532 Hendrix PF, Callahan MA, Drake JM, Huang CY, James SW, Snyder BA, Zhang WX (2008)
533 Pandora's Box contained bait: the global problem of introduced earthworms. *Annu Rev Ecol*
534 *Evol S* 39:593-613

535 Holsinger KE, Lewis PO, Dey DK (2002) A Bayesian approach to inferring population
536 structure from dominant markers. *Mol Ecol* 11:1157-1164

537 Hurst LD, Peck JR (1996) Recent advances in understanding of the evolution and maintenance
538 of sex. *Trends Ecol Evol* 11:A46-A52

539 Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol*
540 *Biol Evol* 23:254-267. <https://doi.org/10.1093/molbev/msj030>

541 Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA
542 polymorphism data. *Bioinformatics* 25:1451-1452

543 López-Legentil S, Turon X (2007) Lack of genetic variation in mtDNA sequences over the
544 amphiatlantic distribution range of the ascidian *Ecteinascidia turbinata*. *Mol Phylogenet Evol*
545 45:405-408. <https://doi.org/10.1016/j.ympev.2007.06.003>

546 Lutes AA, Neaves WB, Baumann DP, Wiegraebe W, Baumann P (2010) Sister chromosome
547 pairing maintains heterozygosity in parthenogenetic lizards. *Nature* 464:283-286.
548 <https://doi.org/10.1038/nature08818>

549 Lynch M (1984) Destabilizing hybridization, general-purpose genotypes and geographic
550 parthenogenesis. *Quarterly Review of Biology* 59:257-290. <https://doi.org/10.1086/413902>

551 Lynch M, Burger R, Butcher D, Gabriel W (1993) The mutational meltdown in asexual
552 populations. *J Hered* 84:339-344. <https://doi.org/10.1093/oxfordjournals.jhered.a111354>

553 Müller F (1857) Description of a new earthworm (*Lumbricus corethrurus*). *Annals and*
554 *Magazine of Natural History* 20:13-15

555 Muller HJ (1932) Some genetic aspects of sex. *Am Nat* 66:118-138. doi:10.1086/280418

556 Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York

557 Righi G (1984) *Pontoscolex* (Oligochaeta : Glossoscolecidae), a new evaluation. *Studies on*
558 *Neotropical Fauna and Environment* 19:159-177

559 Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions.
560 *Trends Ecol Evol* 22:454-464. <https://doi.org/10.1016/j.tree.2007.07.002>

561 Sakai AK et al. (2001) The population biology of invasive species. *Annu Rev Ecol Syst* 32:305-
562 332

563 Schonswetter P, Stehlik I, Holderegger R, Tribsch A (2005) Molecular evidence for glacial
564 refugia of mountain plants in the European Alps. *Mol Ecol* 14:3547-3555.
565 <https://doi.org/10.1111/j.1365-294X.2005.02683.x>

566 Simon JC, Delmotte F, Rispe C, Crease T (2003) Phylogenetic relationships between
567 parthenogens and their sexual relatives: the possible routes to parthenogenesis in animals. *Biol*
568 *J Linn Soc* 79:151–163. <https://doi.org/10.1046/j.1095-8312.2003.00175.x>

569 Slatkin M (1977) Gene flow and genetic drift in a species subject to frequent local extinctions.
570 *Theor Popul Biol* 12:253-262

571 Taheri S et al. (2018a) Complex taxonomy of the 'brush tail' peregrine earthworm *Pontoscolex*
572 *corethrurus*. *Mol Phylogenet Evol* 124:60-70. <https://doi.org/10.1016/j.ympev.2018.02.021>

573 Taheri S, Pelosi C, Dupont L (2018b) Harmful or useful? A case study of the exotic peregrine
574 earthworm morphospecies *Pontoscolex corethrurus*. *Soil Biol Biochem* 116:277-289.
575 <https://doi.org/10.1016/j.soilbio.2017.10.030>

576 Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA
577 polymorphism. *Genetics* 123:585-595.

578 Vekemans X, Beauwens T, Lemaire M, Roldan-Ruiz I (2002) Data from amplified fragment
579 length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship
580 between degree of homoplasy and fragment size. *Mol Ecol* 11:139-151

581 Vorburger C, Lancaster M, Sunnucks P (2003) Environmentally related patterns of reproductive
582 modes in the aphid *Myzus persicae* and the predominance of two 'superclones' in Victoria,
583 Australia. *Mol Ecol* 12:3493-3504. <https://doi.org/10.1046/j.1365-294X.2003.01998.x>

584 Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J,
585 Kuiper M et al (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*
586 23:4407-14

587 Zenger KR, Richardson BJ, Vachot-Griffin AM (2003) A rapid population expansion retains
588 genetic diversity within European rabbits in Australia. *Mol Ecol* 12:789-794.
589 <https://doi.org/10.1046/j.1365-294X.2003.01759.x>

590

591

592 **Figure legends**

593

594 **Fig. 1:** Locations of the *P. corethrurus* L1 specimens used in the study and COI haplotype
595 network based on 478 sequences in 49 locations; with the name of the countries where samples
596 were collected in bold (for details about each site, see Table S1). Each colour represents a
597 different haplotype and pie charts in the map display the relative frequencies for a given locality
598 of each of the haplotypes. The white small circles indicate the presence of islands that could
599 not be visualized otherwise. The location of the 12 populations used in the population genetics
600 analysis are indicated by their code in the map. Codes of populations from the native range (i.e.
601 Guayana Shield) are underlined. In the network, connecting lines show mutational pathways
602 among haplotypes. Only one mutation separates the haplotypes, unless the number of mutation
603 is indicated.

604

605 **Fig. 2:** Neighbor-Net networks done on AFLP profiles with Splitstree based on Nei's genetic
606 distance matrix for 12 populations of *P. corethrurus* L1. Populations from the presumed native
607 range are in black while populations from the introduced range are in blue. The country of
608 origin of each population is indicated by the following codes : GUY – French Guiana, BRA –
609 Brazil, MEX – Mexico, GAB – Gabon, THA – Thailand.

610

611

Table 1: Overall genetic variability in 12 *P. corethrurus* L1 populations, based on 318 polymorphic AFLP markers and 633 bp fragment of the COI gene.

		Native range					Introduced range						
Sites		CAY	MIT	PARA	PARB	PARC	CAX	BEL	JNL	ORL	TLC	LOP	CHG
AFLP	N _{AFLP} ^a	17	20	19	8	13	18	21	22	24	9	20	35
	VM% ^b	52.284	31.472	21.574	22.335	46.701	48.223	41.878	51.523	22.589	28.173	2.030	34.518
	DW (from means) ^b	1804	830.7	362.2	406.1	516.6	708.9	570.7	354.1	426.1	467.2	692.7	471.1
	NoH ^c	17	17	16	7	13	14	21	22	21	9	3	25
	Genotype diversity ^d	1	0.979	0.965	0.964	1	0.935	1	1	0.978	1	0.416	0.973
	Gene diversity ^d	0.165	0.112	0.052	0.087	0.163	0.138	0.132	0.168	0.083	0.113	0.012	0.110
	PCLP ^d	0.91	0.97	0.99	1.00	0.94	0.96	0.93	0.88	0.98	0.99	1.00	0.96
	I _A ^d	6.78	6.82	11.91	18.51	2.54	22.34	12.77	2.23	13.37	42.52	1.29	11.98
	\bar{r}_d ^d	0.03	0.06	0.15	0.21	0.01	0.12	0.08	0.01	0.16	0.39	0.21	0.09
LD (%) ^d	8.66	14.78	8.32	9.67	8.49	16.04	19.84	4.06	34.17	19.64	0.00	20.12	
COI	N _{COI} ^a	12	64	16	16	15	12	15	17	13	9	20	60
	NPS ^e	6	10	0	0	0	11	6	5	0	6	0	1
	Number of haplotype	2	4	1	1	1	4	2	2	1	2	1	2
	Haplotype diversity	0.530	0.693	0	0	0	0.561	0.133	0.221	0	0.500	0	0.381
	P _i (per site) ^f	0.00503	0.00671	0	0	0	0.00426	0.00126	0.00174	0	0.00474	0	0.00060
	Tajima's D ^g	2.294*	2.762**	-	-	-	-1.080	-1.983*	-0.744	-	1.566	-	0.999

^aThe number of individuals that have been genotyped using AFLPs or sequenced for COI are given by N_{AFLP} and N_{COI} respectively. ^bPercentage of variable markers (VM%) and Frequency-down-weighted marker values (DW) were computed using the AFLP dataset with clones. ^cThe number of different AFLP genotypes in the population is given by NoH. ^dFor the computation of genotype diversity, gene diversity, proportion of compatible locus pairs (PCPL), indexes of association (I_A and \bar{r}_d) and gametic disequilibrium (LD), clones were removed from the AFLP dataset; significant values (P<0,002) are shown in bold. ^eNPS correspond to the number of COI polymorphic sites. ^fP_i correspond to the COI nucleotide diversity. ^gFor Tajima's D, significance is given by: * P<0.05, **P<0.01.

Table 2: Analysis of molecular variance (AMOVA) of COI and AFLP (in italics) genotypes for 5 populations from the presumed native range and 7 populations from the introduced range of *P. corethrurus*.

	Variance components	% of total variation	P-value	Φ statistics
Among groups (native <i>versus</i> introduced)	0.142	11.22	0.194	$\Phi_{CT} = 0.112$
	<i>1.49</i>	<i>5.16</i>	<i>0.001</i>	<i>$\Phi_{CT} = 0.052$</i>
Among populations within groups	0.316	24.91	0.000	$\Phi_{SC} = 0.281$
	<i>9.08</i>	<i>31.26</i>	<i>0.000</i>	<i>$\Phi_{SC} = 0.330$</i>
Among all populations	0.809	63.87	0.000	$\Phi_{ST} = 0.361$
	<i>18.47</i>	<i>63.58</i>	<i>0.000</i>	<i>$\Phi_{ST} = 0.364$</i>

Figure1:

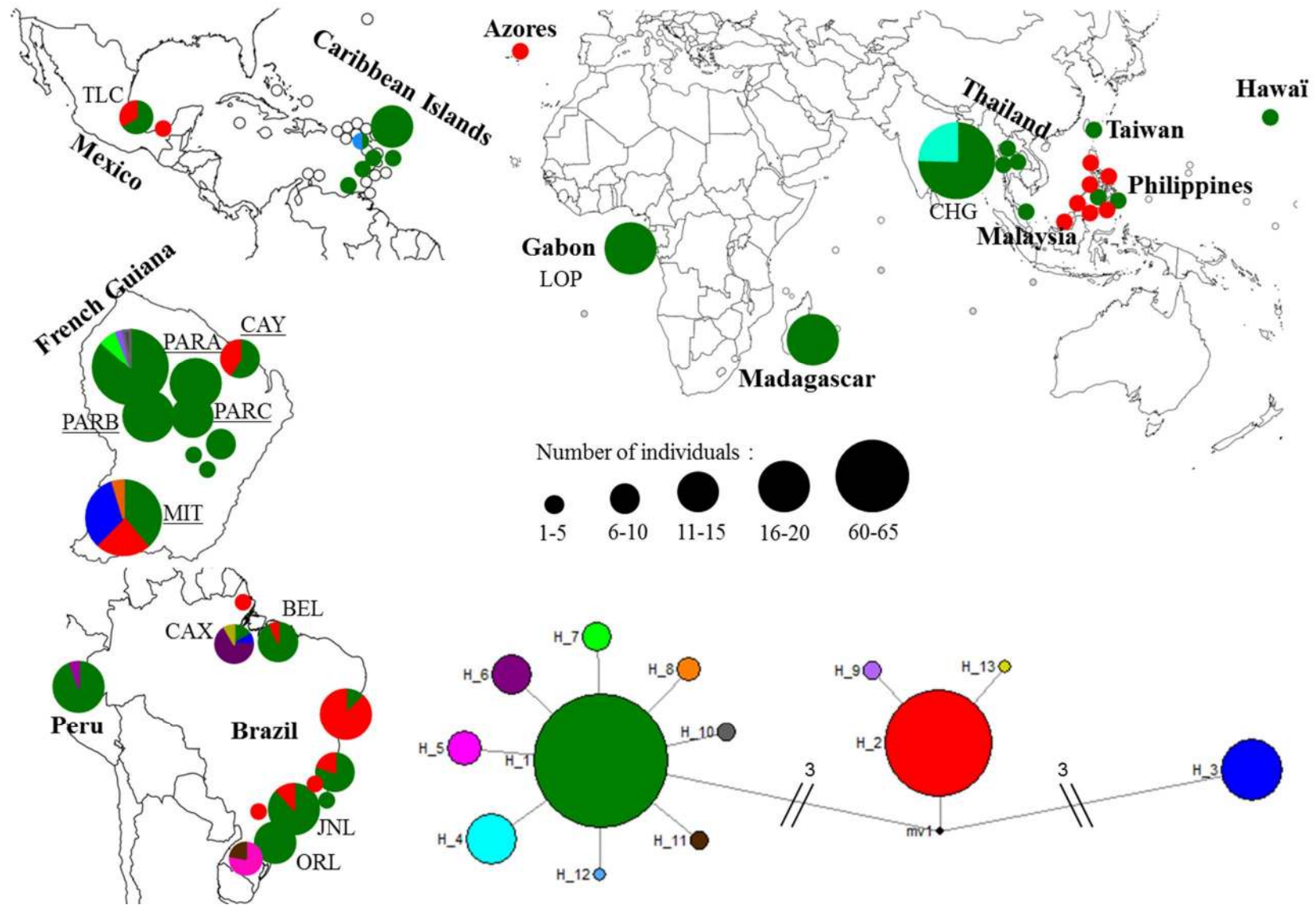


Figure 2

