



UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Ecole de biologie

**MONITORING THE INVASION OF THE MARSH FROG (*PELOPHYLAX
RIDIBUNDUS*) WITHIN THE POPULATIONS OF NATIVE WATER FROGS IN
THE NATURAL RESERVE OF THE GRANDE CARICAIE
USING GENETIC TOOLS**

Travail de Maîtrise universitaire ès Sciences en comportement, évolution et conservation
Master Thesis of Science in Behaviour, Evolution and Conservation

par

Julien LEUENBERGER

Directeur : Prof. Nicolas Perrin
Superviseurs : Antoine Gander & Dr. Benedikt Schmidt
Département d'écologie et évolution (DEE)

Janvier, 2013

Résumé

En suisse, comme dans de nombreux autres pays d'Europe occidentale, les populations sympatriques de grenouilles vertes *Pelophylax lessonae* et *Pelophylax esculentus* (système L-E) ont récemment été perturbées par l'introduction de la grenouille rieuse *Pelophylax ridibundus*. La grenouille verte *P. esculentus*, un hybride naturel entre *P. lessonae* et *P. ridibundus*, présente un mode de reproduction particulier appelé hybridogénèse où le génome *lessonae* est éliminé avant la méiose tandis que le génome *ridibundus* est transmis clonalement. Ce mode de reproduction inhabituel est susceptible de favoriser le potentiel invasif de *P. ridibundus* grâce à un processus génétique de remplacement d'espèces. La récente introduction de *P. ridibundus* au sein du système L-E rend possible la formation de nouveaux hybrides risquant de perturber le fragile équilibre des populations indigènes. Le but de notre étude a été de suivre l'invasion de *P. ridibundus* dans la réserve naturelle de la Grande Cariçaie en Suisse romande. Nous avons démontré le besoin d'outils génétiques (en particulier des marqueurs microsatellites) pour pouvoir déterminer efficacement les trois espèces de grenouilles vertes, étant donné que les critères morphologiques ne permettent pas une identification exacte. L'utilisation de 18 marqueurs microsatellites sur 141 grenouilles vertes échantillonnées nous a permis de définir la composition génétique des trois espèces et de détecter la présence d'hybrides *P. esculentus* récemment formés. L'analyse phylogénétique a révélé que certaines *P. ridibundus* et *P. esculentus* possèdent des génomes mitochondriaux de plusieurs autres espèces non-indigènes de grenouilles du genre *Pelophylax*. Bien que le mécanisme génétique de remplacement des espèces semble menacer la persistance des espèces indigènes de grenouilles vertes, de larges populations de *P. lessonae* et de *P. esculentus* ont été observées dans des zones non favorables à *P. ridibundus*. En effet, le type de milieu semble être un facteur majeur à prendre en compte pour la conservation des espèces indigènes de grenouilles vertes.

Abstract

In Switzerland as in many other Western European countries the populations of the sympatric water frogs *Pelophylax lessonae* and *Pelophylax esculentus* (L-E system) have recently been disturbed by the introduction of the Eurasian marsh frog *Pelophylax ridibundus*. The water frog *P. esculentus*, a natural hybrid between *P. lessonae* and *P. ridibundus*, displays a particular mode of reproduction called hybridogenesis where the *lessonae* genome is eliminated before the meiosis whereas the *ridibundus* genome is clonally inherited. This unusual reproductive system may favour the invasive potential of *P. ridibundus* due to a genetic mechanism of species replacement. The recent introduction of *P. ridibundus* in the L-E system allows the formation of newly hybrids which will disrupt the delicate equilibrium in the established populations. The aim of our study was to monitor the invasion of *P. ridibundus* in the natural reserve of the Grande Cariçaie in Switzerland. We demonstrated the essential need of genetic tools (especially microsatellite markers) to determine efficiently between the three species of the *P. esculentus* complex since the morphological criteria do not allow an accurate identification. Using 18 microsatellite markers over 141 water frog samples, we determined the genetic composition of the three species and detected the presence of newly formed *P. esculentus* hybrids. The phylogenetic analysis revealed that surprisingly some *P. ridibundus* and *P. esculentus* individuals own mitochondrial genomes from several other non-indigenous water frog species. Even if the genetic mechanisms underlying species replacement seemed to threaten the persistence of native water frogs, relatively large *P. lessonae* and *P. esculentus* populations have been found in areas not favourable for *P. ridibundus*. Indeed the type of environment seemed to be a major factor to take into account for the conservation of the native water frogs.

49 **Keywords**

50 Hybridogenesis, invasive species, monitoring, *Pelophylax esculentus* complex, *Pelophylax*
51 *ridibundus*, species identification

Introduction

Introduction of new species made accidentally or by purpose became more frequent with the high rate of international trade (Westphal *et al.* 2008). The spread of these new introduced species could have a strong impact on the biodiversity by threatening strongly indigenous species causing sometimes native species replacement (Wilcove *et al.* 1998; Vorburger & Ribi 1999). To affect strongly and for long term the indigenous populations, the introduced species should evolve into invasive alien species. They have to survive into their new environment, to establish themselves and to increase their density which is possible with a high propagule pressure (Williamson & Fitter 1996). Hybridization with an indigenous species could also enhance the invasive potential and thus be a key process to become invasive (Luquet *et al.* 2011). If the invader can hybridize with a closely related native species, this may lead to long-term introgression, which makes native population less fit. Without introgression, hybridization also affects native species but at a lower rate than with introgression (Rhymer & Simberloff 1996; Huxel 1999).

One famous example of invasive amphibian species is the marsh frog, *Pelophylax ridibundus*, native from Eastern Europe and some regions of Asia such as Turkey and Kazakhstan. As it successfully invades several regions, its distribution range enlarges upon Central and Western Europe (Kuzmin *et al.* 2009). *P. ridibundus* is part of the *P. esculentus* complex, which is composed of two parental species, *P. ridibundus* (genotype RR) and *P. lessonae* (genotype LL), and the hybrid species *P. esculentus* (genotype RL). Due to the presence of the hybrid *P. esculentus*, living in sympatry with *P. lessonae* in Central and Western Europe, the genetic material of *P. ridibundus* was already existent in these regions inside *P. esculentus* before the invasion of *P. ridibundus* (Uzzell & Berger 1975). The natural hybrid *P. esculentus* exhibits a hemiclinal mode of reproduction called hybridogenesis. During hybridogenesis, one parental genome is excluded from the germ line before the

77 meiosis and the other parental genome is clonally inherited. Hybridity is then restored by
 78 mating with the parental species whose genome was excluded. Indeed, the hybrid *P.*
 79 *esculentus* discards the *lessonae* genome (L) from the germ line producing only gametes with
 80 *ridibundus* genome (R). As a consequence of this reproductive system, the hybrid species *P.*
 81 *esculentus* depends on *P. lessonae* to be maintained in a population. Thus, mating between *P.*
 82 *lessonae* and *P. ridibundus* or between *P. lessonae* and *P. esculentus* will produce *P.*
 83 *esculentus* whereas the unique way to obtain *P. lessonae* is a breeding between two *P.*
 84 *lessonae*. In population with only native *P. esculentus* and *P. lessonae*, mating between two
 85 *P. esculentus* produces mainly unviable *P. ridibundus* offspring that die before the
 86 metamorphosis due to deleterious mutations fixed in the R genome. These deleterious
 87 mutations have been accumulated through Muller's ratchet (Muller 1964) in the clonally
 88 inherited R genome where no recombination event occurs (Meyer *et al.* 2009; Holsbeek &
 89 Jooris 2009). However, mating between *P. esculentus* hybrids can produce viable and
 90 competitive *P. ridibundus* offspring if at least one of the hybrids has been recently formed and
 91 displays a new R genome without the deleterious mutation load. These viable offspring could
 92 take an important part in species replacement by the invasive *P. ridibundus* (Luquet *et al.*
 93 2011). This situation is supposed to become more frequent as more recent hybrids will be
 94 formed with the recent arrival of new *P. ridibundus* individuals in Central and Western
 95 Europe. If this is confirmed, the *P. ridibundus* density will increase exponentially in the
 96 population since several crosses will produce *P. ridibundus* (*P. ridibundus* x *P. ridibundus*; *P.*
 97 *esculentus* x *P. esculentus*; *P. ridibundus* x *P. esculentus*). To maintain the indigenous species
 98 *P. lessonae*, mating within *P. lessonae* are needed (Vorburger & Reyer 2003). Moreover, due
 99 to size-related behavioral reasons, mating between small *P. lessonae* male and large *P.*
 100 *ridibundus* female is more likely to occur than between two *P. lessonae* (Plötner *et al.* 2008).

Therefore *P. lessonae* is expected to decline and as a consequence *P. esculentus* will decline as well (Vorburger & Reyer 2003).

In Switzerland, as in many places in Western Europe, only *P. lessonae* and *P. esculentus* coexisted in the so-called L-E system (Holsbeek & Jooris 2009). They colonized Central and Western Europe after the last glaciations 10'000 years ago from Eastern Europe refugees (Vorburger 2001b). The importation of the marsh frog, *P. ridibundus*, for legs consumption and scientific purpose disturbed this L-E system since this new species escaped or was released intentionally by human in the wild (Nöllert & Nöllert 2003; M. Antoniazza personal communication). The first confirmed observation of *P. ridibundus* in Switzerland was made in 1950 around Geneva (Grossenbacher 1988). Afterwards the species has been observed in several places in Switzerland indicating several potential different introductions (Meyer *et al.* 2009). The data collected from the CSCF and the KARCH (CSCF & KARCH 2012) suggest that *P. ridibundus* extended its range from Geneva into two directions: along the Rhone Valley in the Canton of Valais and until the three lakes region (Lake of Neuchâtel, Lake of Bienne and Lake of Morat; Western Switzerland). It appeared also in different areas of Eastern Switzerland as for instance around the airport of Zürich (Vorburger & Reyer 2003; CSCF & KARCH 2012). In several regions where *P. ridibundus* occurs in high density the two native species are close to extinction as observed in the Bois de Finges in the Canton of Valais (Marchesi *et al.* 1999). In this study, we are interested to survey the invasion of *P. ridibundus* within *P. lessonae* and *P. esculentus* populations along the South shore of the Lake of Neuchâtel in the natural reserve of the Grande Cariçaie. The marsh frog *P. ridibundus* was for the first time observed in 1977 in Yvonand, located in the Eastern part of the Grande Cariçaie whereas in Cudrefin, located in the Western part, no observation was made before 1990 (CSCF & KARCH 2012). The frequent observations of *P. lessonae* in the Grande Cariçaie these past years combined with the recent colonization of some areas by *P.*

ridibundus encouraged us to study the problematic caused by the invasive marsh frogs in this region.

Species identification is the first essential step for any conservation studies or any species monitoring. Although the three species of the *P. esculentus* complex display different morphological, morphometric and bio-acoustic characters, the high variation of these characters within species makes the identification based on phenotypic traits difficult (Nöllert & Nöllert 2003; Meyer *et al.* 2009). The morphometric measures allow a reliable identification between the two parental species when both males and females are considered separately, which could be done only on adult species rather during the breeding season (Holsbeek *et al.* 2009a). The main difficulty is to distinguish the hybrid *P. esculentus* from the parental species since its phenotypic characters are often intermediate between the two parental species (Nöllert & Nöllert 2003; Holsbeek *et al.* 2009a). The identification can become more complex as the hybrid *P. esculentus* can also be triploid with either two R genomes (RRL) or two L genomes (RLL) leading to a high overlapping at least of the morphometric criteria (Christiansen 2005). Various genetic tools have been developed to genotype species of the *Pelophylax* genus allowing an efficient differentiation between the different species or hybrids: PCR-RFLP (Restriction Length Polymorphism) (Patrelle *et al.* 2011); size difference of the serum albumin intron (Hauswaldt *et al.* 2012); mitochondrial sequencing (Holsbeek *et al.* 2009b); and microsatellites (Christiansen 2005). These recently developed methods using molecular tools allow a more accurate identification of the species and also the detection of potential multiple origins in the invasive water frogs in Europe (Holsbeek *et al.* 2009b). The use of molecular markers became crucial for conservation genetics and especially for amphibians in our case study, providing essential knowledge on population genetics as population diversity and population size, structure of the population, and providing also a better understanding of hybridization between species (Beebee 2005).

In order to monitor the invasion of *P. ridibundus* within the populations of *P. lessonae* and *P. esculentus* in the Grande Cariçaie, we intend (i) to evaluate the efficiency of morphological and morphometric criteria to distinguish between the three water frog species and to test the efficiency of the molecular tool developed by Hauswaldt *et al.* (2012) using frog samples from Switzerland; (ii) to test the potential of genetic tools to monitor the invasion of *P. ridibundus* within the native species of the *P. esculentus* complex in the natural reserve of the Grande Cariçaie. A large panel of microsatellite markers has been used to study the structure within and between the three species, and mitochondrial DNA sequencing has been used to determine the origin of the invasive species usually identified as *P. ridibundus* in the Grande Cariçaie reserve. Finally, we intend to (iii) see the effect of human maintenance (i.e mowing) on the distribution of the three species among the study area.

Material and methods

Study site and field sampling

Water frog individuals were collected from three areas in the natural reserve of the Grande Cariçaie, located along the South shore of the Lake of Neuchatel in Switzerland during the 2012 breeding season (Mai to August; Table 1). In each area, frogs were sampled in two different environments (zones) that differ by the effect of human maintenance: one is left undisturbed since 2002 (control treatment, TE) and the other is affected by human work such as mowing, tracks left by the machines, creation of new ponds (maintenance treatment, EN). At least three nights of sampling were done per sampling site (for a total of six sampling sites). Individuals were sampled for buccal cells (two sterile buccal swabs per individual; Broquet *et al.* 2007) before being left in a net for a maximum of three nights to avoid

recapture between sampling sessions. Buccal swabs were stored dry at -20°C before analysis. A total of 141 water frogs (adults and juveniles) were captured during the 27 nights of field work (Table 1). Because no frog was found in “Chevroux TE” we limited our analyses to five distinct sampling sites (Table 1).

DNA extraction and Microsatellites amplification

DNA was extracted from buccal swabs using the QIA Amp DNA Mini Kit (Qiagen, Valencia; PA; USA). It was made using the Spin-Column Protocol (manual extraction) and using the BioSprint 96 workstation, both following the manufacturer’s protocols with overnight sample incubation in proteinase K at 56°C.

The DNA samples from the 141 individuals were analyzed with a set of 18 microsatellite markers selected for their applicability in the different genomes of the species under study (Table 2). These markers are known to be either specific to the R, the L or both of the genomes (Zeisset *et al.* 2000; Garner *et al.* 2000; Hotz *et al.* 2001; Arioli 2007; Christiansen 2009; Christiansen & Reyer 2009), as a consequence they allow to identify the three water frog species based on their amplification pattern (Table 2). The initial identification of *P. ridibundus*, *P. lessonae* and *P. esculentus* individuals was made by eye according to the amplification pattern of the 18 microsatellite markers.

Polymerase chain reactions (PCR) were performed in 10 µl containing 3 µl DNA extraction, 3 µl Qiagen Multiplex PCR Master mix and between 2.2 to 3 µl of primers mix (Table 2). The thermal profile of PCR amplification was: an initial denaturation step of 15 min at 95°C followed by 35 cycles of 30 s denaturation at 94 °C, 90 s annealing at T_a (Table 2), 60 s extension at 72 °C, and a final extension of 30 min at 60 °C. PCR products were diluted five times for the solution with primers mix 1 and two times for the solution with

primers mixes 2, 3 and 4. The diluted PCR product (3 µl) was run on an ABI 3130xl Genetic Analyzer (Applied Biosystems, Life Technologies Corporation, Carlsbad, CA) with internal size standard (GeneScan-350 LIZ). The alleles were scored with the Genemapper software (Applied Biosystems 2004).

Species determination with serum albumin intron amplification

The method of species determination developed by Hauswaldt *et al.* 2012 based on the length difference of the serum albumin intron between *P. ridibundus* and *P. lessonae* was used in order to test its reliability. Amplicons of *P. ridibundus* should present one or two bands around 800 bp, *P. lessonae* should present one band smaller than 350 bp whereas *P. esculentus* should present both a large and a small band (Hauswaldt *et al.* 2012). To test the reliability of the method, the results were compared by eye with the species identification resulting from the amplification patterns of the microsatellite markers.

The serum albumin intron region of the DNA was amplified by PCR using the two primers Pel-SA-F1 and Pel-SA-R2, designed by Hauswaldt *et al.* (2012). PCR was performed for the 141 individuals in a total volume of 20 µl, each containing 1µl of 10x buffer, 0.12 µM of each primer, 0.1 mM of dNTPs, 0.4 U of Taq DNA polymerase (Qiagen, Valencia, PA, USA) and 1 µl of extracted DNA. PCR amplification was conducted under the following conditions: an initial denaturation step of 1.30 min at 94°C followed by 35 cycles of 30 s denaturation at 94 °C, 40 s annealing at 59 °C, 100 s extension at 72 °C, and a final extension of 10 min at 72 °C. Species were visually identified on a 1.5% agarose gel where 3 µl of PCR product were run.

Morphology data measurements and analyses

226
227 For each adult captured, six morphological criteria were evaluated and two morphometric
228 measures were taken (Nöllert & Nöllert 2003). The six morphological criteria are listed here
229 below with the expected trait values for *P. ridibundus* and *P. lessonae* respectively in
230 brackets. For *P. esculentus*, the expected trait value was always in between those for *P.*
231 *ridibundus* and *P. lessonae*. Color of the thighs (grey; green and yellow); color of the iris (half
232 black and half yellow; yellow); color of the vocal sacs (grey, white); length of the dorsal line
233 (from the back to the vent; from the back to the eyes); shape of the metatarsal tubercle (flat,
234 round); position of the stub when folding the legs along the body called “stub test” (after the
235 snout, around the eyes). Percentage of correct and false attribution to one of the three species
236 (true positives and false positive) was calculated for each morphological criteria. The two
237 morphometric measures were the snout-to-vent length (SVL) and the tibia length (TL).
238 Measurements were carried out using a caliper with an accuracy of 0.1 mm. The differences
239 of SVL and SVL/TL ratio between the three species were tested respectively with a one way
240 ANOVA and a Kruskal-Wallis χ^2 , using the R software (R Core Team 2012).

241
242 ***Population genetic analyses***
243
244 Clustering of individuals among the whole samples and within each species separately was
245 performed with a Bayesian model-based algorithm using the software STRUCTURE version
246 2.3.3 (Pritchard *et al.* 2000). The most likely number of cluster (K) was calculated with the
247 method based on ΔK developed by Evanno *et al.* (2005) implemented in STRUCTURE
248 HARVESTER (Earl, Dent A. and von Holdt, Bridgett M. 2012). These analyses were carried
249 out with the 18 loci using the admixture model without prior species information. All analyses
250 were conducted using 1'000 burn-in steps and 100'000 iterations.

To visualize the genome differences between *P. ridibundus*, *P. lessonae* and *P. esculentus* a first Principal Component Analysis (PCA) was performed with the software PCAGEN version 2.0 (Goudet 2002). A second PCA was performed to visualize the differences between the R genome of *P. ridibundus*, the L genome of *P. lessonae* and the R* and L* genomes of *P. esculentus* (the * symbol is used to distinguish between the parental species genome and the two genomes present in the hybrid species *P. esculentus*). To perform this analysis we separated the R* from the L* alleles by eye, according to their size which is either specific to the R* or the L* component of the *P. esculentus* genome. The algorithm of PCAGEN calculates the most informative axes based on the F_{ST} values using allelic frequencies of the individuals. PCAGEN is generally used to perform PCA with allelic frequencies of populations. In our analyses each individual was considered as a single population. Unfortunately the statistical significance (p-values) per axis could not be calculated since permutations could not be performed with only one individual per population. The significant factors used for representing graphically the PCA analysis were selected using the scree plot (Cattell 1966). The first PCA analysis was based on two factors although three could be used (Figure S1a) and the second PCA analysis was based on two factors (Figure S1b).

In order to observe if gene diversity differed between the R genome from *P. ridibundus* and the R* genome from *P. esculentus*, and between the L genome from *P. lessonae* and the L* genome from *P. esculentus*, expected heterozygosity was calculated for all individuals using 10 informative loci (the loci specific to one or the other parental species were removed). To test the difference between gene diversity a Mann-Whitney U test was performed using the R software (R Core Team 2012).

For more information on the R* genome in the *P. esculentus* individuals a haplotype network was built using the TCS software version 1.21 (Clement *et al.* 2000). The algorithm

used in TCS, developed by Templeton *et al.* (1992) calculates the estimation of genealogies between DNA sequences. Since TCS use nucleotide sequences as input data, the R* haplotypes were coded manually to nucleotide sequences based on the different alleles size of each loci. The different haplotype sequences were aligned manually using SEAVIEW version 4.2.12 (Gouy *et al.* 2010). The TCS cladogram estimation was performed with a fixed connection limit of 30 steps meaning that hemiclones with more than 30 mutational steps (Single Nucleotide Polymorphism (SNP), insertion and deletion) were not connected with the other hemiclones of the network. Nucleotide gaps were considered as a fifth state for measuring the difference between sequences.

Sequencing of mitochondrial DNA

The extracted DNA from all the 141 individuals was used to amplify the cytochrome-b gene. The region was amplified using primers pair L14850 and H15410 (Tanaka *et al.* 1994) PCR was performed in a total volume of 25 µl, containing 2.5 µl of 10X buffer, 0.2mM of each primer, 0.05mM of dNTPs, 8mM of MgCl₂, 1 U of Taq DNA polymerase (Qiagen, Valencia, PA, USA) and 2 µl of extracted DNA. PCR amplification was performed under the following conditions: an initial denaturation step of 3 min at 94°C followed by 35 cycles of 30 s denaturation at 94 °C, 30 s annealing at 47°C, 60 s extension at 72 °C and a final extension of 7 min at 72°C. To check for amplification, 3 µl of PCR product were run on a 1.5% agarose gel. For the individuals where amplification has been clearly observed on gel, PCR fragments were purified and sequenced at GATC sequencing service (Cologne, D; <http://www.gatc-biotech.com>).

Phylogenetic analyses of cytochrome-b gene

The cytochrome-b DNA sequences were aligned by eye and tests were conducted on the total fragment (466 bp; all codon positions were used). The tree was rooted using a sequence of *Pelophylax saharica*. In addition, published sequences of European and Near East *Pelophylax* species, i.e. *P. bedriagae*, *P. bergeri*, *P. cerigensis*, *P. cretensis*, *P. epeiroticus*, *P. kurtmuelleri*, *P. ridibundus*, *P. shqipericus*, or *P. perezi* (Sumida *et al.* 2000; Lymberakis *et al.* 2007) were induced. The jModelTest 0.1.1 (Guindon & Gascuel 2003; Posada 2008) was used to select the model of DNA substitution for Maximum likelihood (ML) analyses. The HKY+G (Posada 2003) model best fitted the dataset with Akaike's Information Criterion (AIC). Then ML heuristic searches and bootstrap analyses (1000 replicates) were performed using phym (Guindon & Gascuel 2003).

Measure of the effect of the human maintenance on species distribution

The effect of human maintenance on several areas of the Grande Cariçaie could not be tested statistically for two main reasons: design was unbalanced since no individuals were captured in one sampling site in the control environment ("Chevroux TE") and only few *P. ridibundus* individuals were sampled, almost all located in one sampling site.

Results

Determination of the three species based on microsatellites

Genetic determination of the three *Pelophylax* species using microsatellite markers was successful and showed clear results. By eye, three categories of individuals corresponding to

the three species of water frogs were established based on the amplification of the markers: (i) individuals where alleles specific to the L genome were not expressed at several loci, attributed to *P. ridibundus*, (ii) individuals where alleles specific to the R genome were not expressed at several loci, attributed to *P. lessonae*, and (iii) individuals where alleles specific to both the R and the L genomes were expressed at all loci, attributed to *P. esculentus*. Each of the 141 individuals could be distinctly assigned to one of these three categories. Based on the amplification patterns, we identified unambiguously 16 *P. ridibundus*, 38 *P. lessonae* and 87 *P. esculentus*. Depending on the amplification of loci either in the L genome, in the R genome or in both genomes among our samples, we could deduced the genome specificity of the markers; our results diverge from the genome specificity published previously (Table 2). The markers for loci Res22 and RICA5 amplified respectively in the R genome and in the L genome except for one allele. The allele size 99 at locus Res22 was amplified in one *P. lessonae* individual and the allele size 246 at locus RICA5 was amplified in one *P. ridibundus* individual.

Based on the alleles in the *P. esculentus* individuals we could detect the presence of null alleles (i.e. non-amplified alleles) in the R and the L genomes: in the R genome, for the loci RICA2a34 (2.3%), ReGa1a23 (4.6%) and Rrid013A (1.2%) and in the L genome, for the loci Re1Caga10 (26.4%) and Rrid135A (70.1%). The presence of null alleles was not taken into account in our analysis since they do not affect our results.

Efficiency of the determination based on serum albumin intron

Species determination with the amplification of the serum albumin intron showed different results compared with microsatellites: we obtained 39 *P. ridibundus*, 27 *P. lessonae* and 70 *P. esculentus*. Many *P. esculentus* (based on microsatellites amplification) were assigned to *P.*

ridibundus (21 individuals) or to *P. lessonae* (2 individuals). Other misidentifications occurred as two *P. lessonae* were assigned to *P. ridibundus* and nine *P. esculentus* were assigned to *P. lessonae*. Five other individuals could not be assigned to one of the three species. False positives were detected in all species except in *P. ridibundus*.

Efficiency of the determination based on morphological criteria

The six morphological criteria seemed not very informative to determine efficiently between the three species of the *P. esculentus* complex. Rate of true positives was remarkably high (more than 90% of correct attribution) using the color of the thighs and the vocal sacs for distinguishing *P. ridibundus*, and using the color of the vocal sacs for distinguishing *P. lessonae* (Table 3). In other words based on the color of the thighs, 92.3% of the sampled *P. ridibundus* were correctly identified as *P. ridibundus*. Except for these three criteria associated to the mentioned species, the overall percentage of correct attribution was low (going from 34.6% to 84.6%; Table 3). Rate of false positives was remarkably small (less than 5% of false attribution) for *P. lessonae* using the color of the iris, the color of the vocal sacs and the shape of the metatarsal and also for *P. ridibundus* when applying the “stub test” (Table 4). In other words, it means that if we apply the “stub test” to a random individual and the stub goes past to the snout we are able to affirm that it is a *P. ridibundus* species with 1.8% of possible wrong attribution. For *P. ridibundus* (except with the “stub test”) and especially for *P. esculentus* the percentage of false attribution was quite high (going from 5.3 to 46.4; Table 4).

The Snout-Vent Length (SVL) differed significantly between the three species ($F_{2,138} = 30.932$, $p < 0.001$; Figure 1). The *P. ridibundus* being the largest species, *P. lessonae* being the smallest species and the *P. esculentus* species had an intermediate size between the two

parental species. Because of the intra-species variance the SVL of each taxa was overlapping each other. Similar significant results (Kruskal-Wallis $\chi^2 = 34.8067$, $df = 2$, $p < 0.001$) were obtained with the ratio between Snout-Vent Length (SVL) and Tibia Length (TL) (Figure 2) where the hybrid species shows an intermediate position between the two parental species.

Genetic structure of the three species

The method based on the ΔK estimated to three the most likely number of clusters (K) in our sampling, corresponding to the three *Pelophylax* species (Figure 3). All individuals were clustered into one of the three distinct species with a probability higher than 90%, except for six of them: four *P. esculentus* individuals (1_EN_29, 1_EN_40, 2_EN_19 and 3_EN_02) received mixed assignments to the “ridibundus” and “lessonae” clusters, one *P. esculentus* (2_TE_05) was assigned half “esculentus” and half “ridibundus”, and one *P. ridibundus* (3_EN_13) individual was assigned half “ridibundus” and half “esculentus”. No further structure was found when analyzing each species separately.

The PCA analysis of the three species allowed identifying the same clusters as found with STRUCTURE (Figure 4). The first factor (31.22% of total variance) mostly differentiates *P. lessonae* from *P. ridibundus*, while the second factor (9.26% of total variance) differentiates *P. ridibundus* from *P. esculentus*. The six individuals poorly assigned by STRUCTURE also appear as outliers on Figure 5: this includes one *P. ridibundus* individual assigned to an intermediate position between the “ridibundus” and the “esculentus” cluster on factor 2, and five *P. esculentus* individuals assigned to an intermediate position between the “lessonae” and “ridibundus” clusters.

Analysis of the two components of the P. esculentus genome

401

402 The second PCA analysis, performed after discrimination between the lessonae (L*) and

403 ridibundus-like (R*) components of the *P. esculentus* genome also provided clear-cut results

404 (Figure 5). The first factor (37.78% of variance explained) mostly differentiates the two

405 lessonae (L and L*) from the two ridibundus (R and R*) genomes, while the second axis

406 (5.27% of variance explained) discriminates between the R and the R* genomes. The

407 projection of each genome gave clear information of the origin of the two components of the

408 *P. esculentus* genome. The L* genomes clearly stem from the same genetic pool as the L

409 genomes. The R* genomes were more diverse, and mostly distinct from the R genome. The

410 R* genomes of the five *P. esculentus* outliers were closer to the R genomes than to the other

411 R* genomes, and the *P. ridibundus* outlier was intermediate between the R and R* genomes.

412 The analysis of gene diversity also provided contrasted results (Figure 6). Expected

413 heterozygosity was highest in the R genomes of *P. ridibundus*, and lowest in the R* genomes

414 of *P. esculentus* (Wilcoxon rank sum test, $W = 190$, $p < 0.001$). The L genomes of *P. lessonae*

415 and the L* genomes of *P. esculentus* were intermediate, and did not differ significantly from

416 each other (Wilcoxon rank sum test, $W = 101$, $p = 0.908$).

417 With the haplotype network we found a total of twelve different hemiclones among

418 the R* genomes of *P. esculentus* (Figure 7a). A major hemiclone (H1) was present in 2/3 (59

419 individuals) of the *P. esculentus* individuals, one other important hemiclone (H3) was present

420 in 13 individuals, and 7 other hemiclones were genetically close (less than 30 mutational steps

421 from one to another hemiclone). Four hemiclones (H9, H10, H11 and H12) present in one or

422 two individuals were totally isolated from the network due to their high number of mutational

423 steps between them and the others hemiclones. These four isolated hemiclones corresponded

424 to those in the five particular *P. esculentus* individuals mentioned before.

Geographically, no relation between the projection of the R* genomes observed with the second PCA (Figure 5) and their distribution among the sampling sites was observed as we could expect if the invasion of *P. ridibundus* was West to East oriented in Switzerland (data not shown). Also no relation between the twelve hemiclones and their geographical distribution across space was observed (Figure 7b). The two main hemiclones (H1 and H3) were present in the five sampling sites and the four genetically distant hemiclones were spread along the five sampling sites.

Diverse origins of the ridibundus genome

All the *P. ridibundus* individuals except one (3_EN_13) and three *P. esculentus* (1_EN_40, 2_EN_19 and 3_EN_02) individuals presented a good amplification of the cytochrome-b gene. Thus 18 sampled individuals were used for the phylogenetic analysis. The phylogenetic tree highlighted the occurrence of several distinct lineages of mtDNA from invasive water frogs in the study region. The mitochondrial genome of 15 individuals, including two *P. esculentus* (2_EN_19 and 3_EN_02) was closely related to *P. ridibundus* lineages from Eastern Europe, the genome of two individuals, including one *P. esculentus* (1_EN_40) to lineages of *P. cerigensis/bedriagae* from Near East (Turkey) and the genome of one individual to *P. kurtmuelleri* from Southern Greece (Sumida *et al.* 2000; Lymberakis *et al.* 2007; Figure 8). These results were not taken into account for the taxonomy of the species for the following sections.

Distribution of the three species along the south shore of the Lake of Neuchâtel

The distribution of each *P. esculentus* complex species along the sampling area is shown on Figure 9. The *P. lessonae* species and the hybrid species *P. esculentus* have been sampled in all sites and *P. esculentus* was in majority in three of them. The *P. ridibundus* species were present only in two sampling sites located further East and further West.

Distribution of the three species according to the human maintenance of the sites

The proportion of each species per zone either left undisturbed since 2002 (TE) or affected by human work (i.e. mowing) for the sites maintenance (EN) is shown on the Figure 10. All the *P. ridibundus* individuals were captured in the EN zone. Although *P. esculentus* and *P. lessonae* were captured in both the EN and the TE zone, a higher proportion of *P. esculentus* were found in the EN zones whereas a higher proportion of *P. lessonae* were found in the TE zones.

Discussion

Monitoring the water frogs from the *P. esculentus* complex is essential for taking conservation measures to control the invasion of the marsh frog *P. ridibundus* in Central and Western Europe (Holsbeek & Jooris 2009). Here we proposed to monitor this invasion in the natural reserve of the Grande Cariçaie in Western Switzerland. We showed that the use of specific genetic tools is essential to distinguish correctly between *P. ridibundus*, *P. lessonae* and *P. esculentus*. Genetic tools provide also essential information on population structure within and between the three species and can be used to determine the introgression level of *P. ridibundus* within native water frogs populations. The role of human maintenance in species distribution was not proved but the observed trend encourage us for further studies.

These findings are discussed and perspectives for the preservation of the *P. esculentus* complex are given.

Limit of the morphological criteria and efficiency of the genetic tools for species identification in the *P. esculentus* complex

The six morphological criteria used in this study did not allow an accurate identification of the three species of the *P. esculentus* complex. By combining the false positive and the true positive values (Tables 3 and 4), we could deduce the combination of criteria to use to obtain the best accuracy in species identification. The most accurate identification for male *P. ridibundus* consists in combining the color of the vocal sacs (100% true positive) and the “stub test” (1.8% false positive). This allows first to detect all *P. ridibundus* individuals plus several individuals from other species and second to remove 98.2% of the wrong species. Concerning female *P. ridibundus*, the color of the thighs (6.3% false positive) should be used instead of the vocal sacs in males. To distinguish male *P. lessonae*, the color of the vocal sacs (93.8% true positive) and the color of the iris (0% false positive) should be checked. For female *P. lessonae*, identification seems less accurate than in males, however the “stub test” (72.7% true positive) should be applied instead of the vocal sacs in males. A correct identification of *P. esculentus* seemed not possible due to the low percentage of true positives and the high percentage of false positives (Tables 3 and 4). These results corroborate with the difficulty to distinguish between the three species of the *P. esculentus* complex based on phenotypic traits and especially to identify *P. esculentus* as described by Nöllert and Nöllert (2003). Finally mistakes could be done by assigning an individual into a false category within the criteria since they are all sorted into three qualitative categories, which led to misassignments due to the subjectivity of the observer. To improve the use of these six

criteria, we propose for further studies to test for a significant difference and an absence of overlapping between the three species if quantitative variables instead of qualitative ones are used.

The morphometric data measured on the tree species revealed significant differences between species which support the results found by Holsbeek *et al.* (2009a). Both the SVL and the SVL/TL ratio measures are significantly different between the three species taken together and between each pairs of species (Figures 1 and 2). However, at the individual level this difference could be small since the measures overlap within the three species which was also shown by Holsbeek *et al.* (2009a). Concerning *P. esculentus*, species determination can be hazardous since its phenotype is intermediate between its two parental species *P. ridibundus* and *P. lessonae*, which explains the limits of morphological species identification (Lodé & Pagano 2000). As a consequence, the SVL and the SVL/TL ratio cannot be used to determine efficiently the three water frogs and give only information on the average size of each species.

The method developed by Hauswaldt (2012), based on the amplification of serum-albumin intron, seems not to work for water frogs sampled in Western Switzerland since we found many wrong species assignation compared with the identification made with the microsatellite markers. The method of Hauswaldt (2012) has been developed on water frogs from Eastern and Northern Europe and seems not adequate for designing a new reliable method for distinguishing water frogs of the *P. esculentus* complex from other European regions. Thus, the use of the 18 microsatellite markers seemed the most accurate method for species identification since they assigned each individual into one of the three distinct categories corresponding to the three species. Six out of the 18 markers are specific either to the *P. lessonae* genome or to the *P. ridibundus* genome as they do not express respectively the R-specific alleles and the L-specific alleles (Table 2). These loci specificity do not correspond

exactly to those found in the literature (Zeisset *et al.* 2000; Garner *et al.* 2000; Hotz *et al.* 2001; Arioli 2007; Christiansen & Reyer 2009; Table 2). The difference in geographical origin of the sampled frogs could explain these unconcordant results as explained by Holsbeek *et al.* (2009a).

P. ridibundus: species with a limited distribution but with high genetic diversity coming from several origins

The invasive marsh frog *P. ridibundus* seemed not very frequent in the marshland along the Grande Cariçaie since only 16 individuals (11% of the total number of individuals) have been collected during this survey (Figure 9). However, the small sampling size of *P. ridibundus* may not represent their real density in this region, since only three areas were visited, all characterized as inland wetlands. The habitat preference of *P. ridibundus* consists in environment connected to freshwater with high concentration of oxygen such as water channels and open areas connected to the lake (Schmeller *et al.* 2007). Indeed, in our study region, many *P. ridibundus* have been observed in open areas connected with the lake, located close to the sampling sites (M. Antoniazza personal communication).

Genetically *P. ridibundus* seemed to have a high intra-specific variation as shown by the first PCA (Figure 4) and the genetic diversity measured with the expected heterozygosity (Figure 6). The projection of *P. ridibundus* individuals in the biplot graph of the first PCA was more dispersed than the well-clustered projection of *P. lessonae* individuals and the number of alleles at each locus was higher for *P. ridibundus* than for *P. lessonae*. This genetic diversity of *P. ridibundus* compared with *P. lessonae* was also observed by Vorburger and Reyer (2003). This high genetic diversity could be explained by the multiple introductions through time of *P. ridibundus* coming from different location in Eastern Europe (Meyer *et al.*

2009). Therefore the introduced *P. ridibundus* originated from a large genetic pool coming from a large geographic range, and based on the phylogenetic analysis they do not only diverged geographically but also taxonomically. Based on cytochrome-b amplification from supposed *P. ridibundus* and *P. esculentus* individuals we found the presence of genetic lineages of *P. ridibundus*, *P. cerigensis/bedriagae*, and *P. kurtmuelleri* (Figure 8). The detection of multiple mtDNA lineages of invasive water frog species in Switzerland will make the species identification even more difficult. This case is similar to the one observed in Belgium where first two and then three different lineages of invasive water frogs were genetically detected respectively in 2008 and 2009 (Holsbeek *et al.* 2008, 2009b). Even if the mtDNA of supposed *P. ridibundus* come from several water frog species only one nuclear genetic pool seems to exist since all supposed *P. ridibundus* (except one: 3_EN_13) were assigned to one single cluster with the cluster analysis (Figure 3). A large admixture between nuclear genome of all introduced water frog species could explain their assignation to only one genetic cluster, which corresponds to the supposed *P. ridibundus* species.

One *P. ridibundus* individual (3_EN_13) was assigned half to the “ridibundus” and half to the “esculentus” cluster with the cluster analysis (Figure 3). This individual certainly results from a cross between a *P. esculentus* and an introduced *P. ridibundus*. The mtDNA of this individual should be different from the one found in the other *P. ridibundus* since the amplification of the cytochrome-b gene was not successful, leading to the supposition that mtDNA from *P. lessonae* lineage is present in this individual. The *P. lessonae* mtDNA is certainly present in many *P. esculentus* individuals by frequent crosses between males *P. esculentus* and females *P. lessonae* occurring in the L-E system (Spolsky & Uzzel 1986). This hypothesis should be confirmed using universal primers amplifying mtDNA in all the three species of the *P. esculentus* complex.

P. lessonae: threatened species with encouraging results for the nearly future

The Pool frog, *P. lessonae*, is considered as nearly threatened (NT) together with *P. esculentus* in the 2005 red list of threatened species in Switzerland (Schmidt & Silvia 2005). These two species are considered as a whole due to the difficulty to distinguish them morphologically, which reinforces the need of genetic tools for an efficient survey. Based on its red list status and the rare observation of *P. lessonae* in the Bois de Finges, a Swiss natural reserve, since the invasion of *P. ridibundus* (Marchesi *et al.* 1999), we expected to collect only few *P. lessonae* individuals in our study areas. Despite the known presence of *P. ridibundus*, we collected a relatively high number of 38 *P. lessonae* individuals (27% of the total number of individuals), which suggests a high density of *P. lessonae* in this region. They have been collected in each sampling site and are dominant in the marshlands of “Gletterens TE” (Figure 9) which indicates an important settlement of *P. lessonae* populations in this region.

The genetic diversity within *P. lessonae* individuals, shown by their spatial position in the first and second PCA (Figures 5 and 6) and also by their expected heterozygosity (Figure 6) is smaller than within *P. ridibundus* individuals. As discussed before, *P. ridibundus* showed a higher genetic diversity because of its several geographical and taxonomical origins. Even if the expected heterozygosity within *P. lessonae* is lower than in *P. ridibundus*, its mean around 0.5, (Figure 6) is high enough to avoid a decrease of the species due to a small genetic diversity in the population. Another good point for the persistence of *P. lessonae* in the nearly future is the absence of introgression: no alleles from the R genome was found in the genome of *P. lessonae* since the results obtained by the cluster analysis (Figure 3) and the first PCA (Figure 4) showed a clear and distinct classification of all the *P. lessonae* individuals into one taxa.

P. esculentus: species with large populations and a large number of hemiclones partly due to the invasion of P. ridibundus

The hybrid *P. esculentus* is the most represented species along the study area with a total of 87 individuals captured (62% of the total number of individuals; Figure 9). The species was found on each sampling site and dominated largely three of them. The predominance of *P. esculentus* within populations of the L-E system with or without invasion of *P. ridibundus* is frequent as observed in other surveys (Marchesi *et al.* 1999; Hotz *et al.* 2007). One hypothesis of this high density of *P. esculentus* could be due to behavioral reasons during the breeding period. Indeed, males *P. lessonae* favored breeding with *P. ridibundus* females which are larger than *P. lessonae* females (Plötner *et al.* 2008). Thus males *P. lessonae* could preferentially breed with larger *P. esculentus* females for the same reason resulting in *P. esculentus* progeny. This situation together with breeding competition between the two native species could easily tip the scale toward *P. esculentus* predominance.

The R* genome of *P. esculentus* also called hemiclone, is transmitted clonally through hybridogenesis to the progeny since *P. esculentus* lives in sympatry with its parental species *P. lessonae* (L-E system). Due to its clonal transmission, no recombination events occur in the R* genome leading to a much lower genetic diversity than in the R genome of *P. ridibundus* where recombination events occur (Muller 1964; Figure 6). With the introduction of *P. ridibundus* in the L-E system, the genetic diversity and thus the hemiclone diversity became higher (Hotz *et al.* 2007) which can be explained by the presence of recently formed *P. esculentus* through primary hybridization. In this study even if *P. ridibundus* was not observed at high density a high number of twelve different hemiclones was detected (Figure 7a). It is higher than the 10 hemiclones found by Vorburger and Reyer (2003) in more than

170 *P. esculentus* sampled close to the airport of Zürich in Switzerland. Nevertheless another important parameter to take into account when calculating hemiclone diversity is the kind of marker used and the number of loci. With the addition of two more microsatellite loci, into a set of seven allozyme markers, Hotz *et al.* (2001) found twice more hemiclones. In their study, Vorburger and Reyer (2003) have used six allozyme markers which explain why they found a smaller number of hemiclones than us. The number of hemiclones we found is probably underestimated and the use of several other microsatellite markers would allow a better resolution by finding new hemiclones. Among the twelve hemiclones, H1 is the most common overall and is, with H3, found in each sampling site. We supposed at least the H1 hemiclone to be present in *P. esculentus* since a long time, which could correspond to the colonization event following the last glaciations. The other 10 hemiclones were only present in one to four *P. esculentus* individuals spread among the five sampling sites (Figure 7b). Eight of them, including H1 and H3, are relatively genetically close since they are separated from two to 27 mutational steps. The four other hemiclones (G, H, I and J) are genetically more diverged than 30 mutational steps both among each other than with the eight others hemiclones (Figure 7a). The eight closely connected hemiclones may be a result of occasional recombination events between R genomes (Vorburger 2001a, 2001b). Indeed it has been shown that newly formed *P. ridibundus* resulting from mating between *P. esculentus* individuals with different hemiclones followed by meiotic recombination can survive at low rate (Vorburger & Reyer 2003; Holsbeek & Jooris 2009; Luquet *et al.* 2011). These newly formed *P. ridibundus* females (males cannot result from such crosses) may lead to the creation of new *P. esculentus* hemiclones by mating with *P. lessonae* (reviewed by Holsbeek & Jooris 2009).

We have good reasons to consider the four disconnected hemiclones (Figure 7) as newly created, coming from primary hybridization between *P. lessonae* and introduced *P.*

649 *ridibundus*. Effectively, the different analyses performed to define the structure of the *P.*
 650 *esculentus* genome show that the R* genome of five *P. esculentus* individuals (1_EN_29,
 651 1_EN_40, 2_EN_19, 3_EN_02 and 2_TE_05) is closer to the R genome of *P. ridibundus* than
 652 to the other R* genomes. With clustering analysis (Figure 3), four of them (1_EN_29,
 653 1_EN_40, 2_EN_19 and 3_EN_02) were assigned both to the “ridibundus” and the
 654 “lessonae” cluster and the fifth individual (2_TE_05) was assigned half to the “ridibundus”
 655 and half to the “esculentus” cluster. To be assigned partly to the “ridibundus” cluster they
 656 should share similar allelic frequencies with the majority of the *P. ridibundus* individuals
 657 (Pritchard *et al.* 2000), which is probably due to the presence of *P. ridibundus* alleles on their
 658 R* genome. These five *P. esculentus* individuals were assigned to an intermediate position
 659 between the “lessonae” and “ridibundus” cluster on the first PCA (Figure 4) and to an
 660 intermediate position between the R and the R* genomes cluster on the second PCA (Figure
 661 5). The intermediate position shown by the first PCA (Figure 4) suggests that one part of their
 662 genome shares similarities with the “ridibundus” cluster and the other part shares similarities
 663 with the “lessonae” cluster. The second PCA (Figure 5) demonstrates that the L* haplotype
 664 shares exactly the same genetic pool with the L genome of *P. lessonae* individuals whereas
 665 the R* hemiclone seems to have two different origins. The majority of the *P. esculentus*
 666 individuals owns a R* hemiclone different from the R genome of *P. ridibundus*, certainly
 667 inherited clonally since generations by hybridogenesis between *P. lessonae* and *P. esculentus*.
 668 Five other *P. esculentus*, corresponding to the five individuals described ahead, have one of
 669 the four distinct hemiclones clustered close to the R genome of *P. ridibundus*. These four R*
 670 hemiclones certainly share the same or at least a very similar genetic pool with the R genome
 671 of the introduced *P. ridibundus* individuals. The results obtained from these analyses support
 672 the presence of five new *P. esculentus* hybrids resulting from a recent mating between native
 673 *P. lessonae* and invasive *P. ridibundus*. These new hybrids are distributed along the study

area as they have been found in four different sites (Figure 7b). This large distribution can be either due to a high dispersion of *P. esculentus* as shown by Marchesi *et al.* (1999) or by the presence of larger populations of *P. ridibundus* than we actually surveyed which could increase the formation of new hybrids. The detection of recent hybridization between *P. lessonae* and introduced *P. ridibundus* in the Grande Cariçaie corresponds to the beginning of the genetically process of species replacement developed by Vorburger and Reyer (2003): with the increase of new hybridization events, more matings between two *P. esculentus* will arise and lead to *P. ridibundus* offspring. Thus the density of *P. ridibundus* will increase whereas the density of the two native species will decrease.

The mtDNA phylogenetic analysis reveals that one *P. esculentus* out of the three analyzed possessed a mtDNA from *P. cerigensis* whereas the two other possessed a mtDNA from *P. ridibundus* (Figure 8). Although this indicates the occurrence of crosses between *P. lessonae* and *P. cerigensis*, it is impossible to estimate when these crosses occurred. All the other *P. esculentus* individuals did not amplify because they certainly owned a *P. lessonae* mtDNA instead of a *P. ridibundus* mtDNA which seems to be often the case as shown by Spolsky & Uzzel (1986).

Distribution of the three species influenced by a particular habitat type

Species from the *P. esculentus* complex exhibit different ecologies: *P. ridibundus* prefers habitat with relatively high amount of oxygen and low salinity as found in wetlands connected to freshwater sources (Schmeller *et al.* 2007) whereas *P. esculentus* and *P. lessonae* are more adapted to habitat with low oxygen amount (Plenet *et al.* 2001, 2005). Meyer *et al.* (2009) also described that *P. lessonae* is more abundant in marshes or small ponds with high level of vegetation whereas *P. esculentus* and *P. ridibundus* prefer large wetlands like lake shores or

water channels. We also observed a trend in habitat differences: *P. ridibundus* seemed to be favored in areas affected by human maintenance like the EN zones or rather not to be adapted in undisturbed areas since no individuals were found in the undisturbed zones (TE) (Figure 10). The presence of *P. lessonae* and *P. esculentus* in both maintained (EN) and undisturbed (TE) zones seemed to show their adaptation for the two environments (Figure 10). These observations should be considered as hypotheses because the sampling took place only in three areas and only few *P. ridibundus* individuals were sampled with a majority of them collected in “Chevroux EN”. Due to the absence of observation in “Chevroux TE”, located several hundred of meters from “Chevroux EN”, we cannot distinguish between the effect of the maintenance and the effect of the area on species distribution. Furthermore, the absence of observation does not mean a real absence of the three species since it could be due to detectability reasons and be only false absences (Kéry & Schmid 2004). To detect with efficiency an effect of the human maintenance on water frogs distribution we propose to perform other surveys with more numerous sampling sites and samples. Even if we observe only a tendency, we point out the importance of undisturbed areas to preserve *P. lessonae* and *P. esculentus* from the invasion of *P. ridibundus*.

Perspectives and conclusion

Our study provides information to improve the conservation of the two native species *P. lessonae* and *P. esculentus* from the invasion of *P. ridibundus*. We showed the difficulty to distinguish especially *P. esculentus* from its two parental species based on morphological criteria. Thus, for any survey of this complex of species the use of genetic tools and especially microsatellite markers is essential for species identification and to bring further knowledge on the structure and genetic composition of the three species. The detection of several

mitochondrial genomes in either invasive water frogs or *P. esculentus* individuals could affect differently the L-E system than expected with the only presence of *P. ridibundus*. It could affect the hybridogenetic reproductive system since hybrids resulting from Southern European *P. ridibundus* did not induce hybridogenesis and hybrids resulting from *P. bedriagae* may be sterile (Holsbeek *et al.* 2008; Holsbeek & Jooris 2009). Large-scale study should be done to clarify the origin and taxonomy of the exotic water frogs in Switzerland. Even if we did not prove the effect of human maintenance on species distribution, the high proportion of both *P. esculentus* and *P. lessonae* collected in areas where *P. ridibundus* seemed absent or at least in little density suggests habitat differentiation between the species. For the conservation of the two native species these observations are favorable and should be explored. Further studies should be done to obtain clear-cut results on the effect of human maintenance to provide useful information for natural reserve management. This study also provides a good framework for a long-term monitoring of the water frogs both in the study region, the Grande Cariçaie, and in other regions where the L-E system is threatened by the invasion of *P. ridibundus*. It will allow a global knowledge on the level of invasion with the possibility to detect recently formed hybrids.

Acknowledgments

I wish to thank the association of the Grande Cariçaie and especially Michel Antoniazza and Alexandre Ghiraldi for sharing their knowledge respectively on the ecology of the region and on GIS; Sylvain Dubey for its important contribution on the phylogeny analyses; Consolée Aletti, Catherine Berney, Christophe Dufresnes, Romain Savary and Roberto Sernier for their help in the lab; Nicolas Pruvost for general advices; Cécile Benjemia, Sébastien Biollay, Olivier Darbellay, Ludovic Dutoit, Florent Goetschi, Nicolas Hazi, Alexandra Jost, Honorine Lovis, Thomas Martignier, Lea Megali, Guillaume Mottaz, Anna Murakozy, Marion Podolak, Emilie Rapin, Manuel Ribaux, Nicolas Rodrigues, Patrick Roelli, Aude Rogivue, Marina Stojiljkovic and Quentin Theiler for their help during the endless nights of field work; Caroline Betto-Colliard for reviewing this report; finally I especially thank Antoine Gander and Benedikt Schmidt for their support during the study and Nicolas Perrin for its availability and its advice all along this Master project. The samples were collected with the authorisation of the “Canton de Vaud” (N°2012-24).

755 **Literature cited**

- 756 Arioli, M. 2007. Reproductive Patterns and Population Genetics in Pure Hybridogenetic
757 Water Frog Populations of *Rana esculenta*. University of Zurich. Retrieved from
758 www.dissertationen.unizh.ch.
- 759 Beebee, T. J. C. 2005. Conservation genetics of amphibians. *Heredity* **95**:423–427.
- 760 Broquet, T., L. Berset-Braendli, G. Emaresi, and L. Fumagalli. 2007. Buccal swabs allow
761 efficient and reliable microsatellite genotyping in amphibians. *Conservation genetics*
762 **8**:509–511.
- 763 Cattell, R. B. 1966. Scree test for number of factors. *Multivariate behavioral research* **1**:245–
764 276.
- 765 Christiansen, D. G. 2005. A microsatellite-based method for genotyping diploid and triploid
766 water frogs of the *Rana esculenta* hybrid complex. *Molecular ecology notes* **5**:190–193.
- 767 Christiansen, D. G. 2009. Gamete types, sex determination and stable equilibria of all-hybrid
768 populations of diploid and triploid edible frogs (*Pelophylax esculentus*). *BMC*
769 *evolutionary biology* **9**:135.
- 770 Christiansen, D. G., and H.-U. Reyer. 2009. From clonal to sexual hybrids: genetic
771 recombination via triploids in all-hybrid populations of water frogs. *Evolution* **63**:1754–
772 68.
- 773 Clement, M., D. Posada, and K. a Crandall. 2000. TCS: a computer program to estimate gene
774 genealogies. *Molecular ecology* **9**:1657–1659.
- 775 CSCF (Centre Suisse de Cartographie de la Faune), and KARCH (Centre de coordination
776 pour la protection des amphibiens et des reptiles de Suisse). 2012. Distribution des
777 espèces: *Pelophylax ridibundus*. Retrieved January 6, 2013, from
778 [http://lepus.unine.ch/carto/index.php?nuesp=70126&rivieres=on&lacs=on&hillsh=on&](http://lepus.unine.ch/carto/index.php?nuesp=70126&rivieres=on&lacs=on&hillsh=on&year=1990)
779 [year=1990](http://lepus.unine.ch/carto/index.php?nuesp=70126&rivieres=on&lacs=on&hillsh=on&year=1990).
- 780 Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals
781 using the software STRUCTURE: a simulation study. *Molecular ecology* **14**:2611–2620.
- 782 Garner, T. W. J., B. Gautschi, S. Rothlisberger, and H.-U. Reyer. 2000. A set of CA repeat
783 microsatellite markers derived from the pool frog, *Rana lessonae*. *Molecular ecology*
784 **9**:2173–2175.
- 785 Goudet, J. 2002. PCAGEN. Retrieved from
786 <http://www2.unil.ch/popgen/softwares/pcagen.htm>.
- 787 Gouy, M., S. Guindon, and O. Gascuel. 2010. SeaView version 4: A multiplatform graphical
788 user interface for sequence alignment and phylogenetic tree building. *Molecular biology*
789 *and evolution* **27**:221–224.

- 790 Grossenbacher, K. 1988. Verbreitungsatlas der Amphibien der Schweiz. Documenta
791 Faunistica Helvetiae 7:1–207.
- 792 Guindon, S., and O. Gascuel. 2003. A Simple, Fast, and Accurate Algorithm to Estimate
793 Large Phylogenies by Maximum Likelihood. *Systematic Biology* 52:696–704.
- 794 Hauswaldt, J. S., M. Höer, M. Ogielska, D. G. Christiansen, D. Dziewulska-Szwajkowska, E.
795 Czernicka, and M. Vences. 2012. A simplified molecular method for distinguishing
796 among species and ploidy levels in European water frogs (*Pelophylax*). *Molecular*
797 *ecology resources* 12:797–805.
- 798 Holsbeek, G., and R. Jooris. 2009. Potential impact of genome exclusion by alien species in
799 the hybridogenetic water frogs (*Pelophylax esculentus* complex). *Biological Invasions*
800 12:1–13.
- 801 Holsbeek, G., G. E. Maes, L. De Meester, and F. a M. Volckaert. 2009a. Conservation of the
802 introgressed European water frog complex using molecular tools. *Molecular ecology*
803 18:1071–1087.
- 804 Holsbeek, G., J. Mergeay, H. Hotz, J. Plötner, F. a M. Volckaert, and L. DE Meester. 2008. A
805 cryptic invasion within an invasion and widespread introgression in the European water
806 frog complex: consequences of uncontrolled commercial trade and weak international
807 legislation. *Molecular ecology* 17:5023–5035.
- 808 Holsbeek, G., J. Mergeay, F. a. M. Volckaert, and L. Meester. 2009b. Genetic detection of
809 multiple exotic water frog species in Belgium illustrates the need for monitoring and
810 immediate action. *Biological Invasions* 12:1459–1463.
- 811 Hotz, H., G.-D. Guex, P. Beerli, R. D. Semlitsch, and N. B. M. Pruvost. 2007. Hemiclone
812 diversity in the hybridogenetic frog *Rana esculenta* outside the area of clone formation:
813 the view from protein electrophoresis. *Journal of Zoological Systematics and*
814 *Evolutionary Research* 46:56–62.
- 815 Hotz, H., T. Uzzel, G. Guex, D. Alpers, R. D. Semlitschj, and P. Beerli. 2001.
816 Microsatellites : a tool for evolutionary genetic studies of western Palearctic water frogs.
817 *Mitt. Mus. Nat.kd. Berl., Zool.* 77:43–50.
- 818 Huxel, G. R. 1999. Rapid displacement of native species by invasive species : effects of
819 hybridization. *Biological Conservation* 89:143–152.
- 820 Kuzmin, S. et al. 2009. *pelophylax ridibundus* IUCN Red List of Threatened Species.
821 Retrieved January 6, 2013, from <http://www.iucnredlist.org/>.
- 822 Kéry, M., and H. Schmid. 2004. Monitoring programs need to take into account imperfect
823 species detectability. *Basic and Applied Ecology* 5:65–73.
- 824 Lodé, T., and a Pagano. 2000. Variations in call and morphology in male water frogs:
825 taxonomic and evolutionary implications. *Comptes rendus de l'Académie des sciences.*
826 *Série III, Sciences de la vie* 323:995–1001.

- 827 Luquet, E., C. Vorburger, F. Hervant, P. Joly, B. Kaufmann, D. S. Schmeller, J. P. Léna, O.
828 Grolet, L. Konecny, and S. Plénet. 2011. Invasiveness of an introduced species: the role
829 of hybridization and ecological constraints. *Biological Invasions* **13**:1901–1915.
- 830 Lymberakis, P., N. Poulakakis, G. Manthou, C. S. Tsigenopoulos, a Magoulas, and M.
831 Mylonas. 2007. Mitochondrial phylogeography of *Rana* (Pelophylax) populations in the
832 Eastern Mediterranean region. *Molecular phylogenetics and evolution* **44**:115–25.
- 833 Marchesi, P., J. Fournier, and A. Rey. 1999. Etat des populations de “grenouilles vertes” *Rana*
834 *lessonae*, *Rana kl. esculenta* du Bois de Finges (Salquenen, VS). *Bulletin de la*
835 *Murithienne* **117**:13–22.
- 836 Meyer, A., S. Zumbach, B. Schmidt, and J.-C. Monney. 2009. *Amphibiens et reptiles de*
837 *Suisse*. 336 pages. Haupt Verlag. Bern.
- 838 Muller, H. J. 1964. The relation of recombination to mutational advance. *Mutation Research*
839 **1**:2–9.
- 840 Nöllert, A., and C. Nöllert. 2003. *Guide des amphibiens d’Europe*. 383 pages. Delachaux et
841 Niestlé. Paris.
- 842 Patrelle, C., T. Ohst, D. Picard, A. Pagano, S. Sourice, M.-G. Dallay, and J. Plötner. 2011. A
843 new PCR-RFLP-based method for an easier systematic affiliation of European water
844 frogs. *Molecular ecology resources* **11**:200–5.
- 845 Plénet, S., F. Hervant, and P. Joly. 2001. Ecology of the hybridogenetic *Rana esculenta*
846 complex : differential oxygen requirements of tadpoles. *Evolutionary Ecology* **14**:13–
847 23.
- 848 Plénet, S., P. Joly, F. Hervant, E. Fromont, and O. Grolet. 2005. Are hybridogenetic
849 complexes structured by habitat in water frogs? *Journal of evolutionary biology*
850 **18**:1575–86.
- 851 Plötner, J., T. Uzzell, P. Beerli, C. Spolsky, T. Ohst, S. N. Litvinchuk, G.-D. Guex, H.-U.
852 Reyer, and H. Hotz. 2008. Widespread unidirectional transfer of mitochondrial DNA: a
853 case in western Palaearctic water frogs. *Journal of evolutionary biology* **21**:668–681.
- 854 Posada, D. 2008. jModelTest: phylogenetic model averaging. *Molecular biology and*
855 *evolution* **25**:1253–1256.
- 856 Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using
857 multilocus genotype data. *Genetics* **155**:945–959.
- 858 R Core Team. 2012. *R: A Language and Environment for Statistical Computing*. Vienna,
859 Austria. Retrieved from <http://www.r-project.org>.
- 860 Rhymer, J. M., and D. Simberloff. 1996. Extinction by Hybridization and Introgression.
861 *Annual Review of Ecology and Systematics* **27**:83–109.

- 862 Schmeller, D. S., A. Pagano, S. Plénet, and M. Veith. 2007. Introducing water frogs – Is there
863 a risk for indigenous species in France ? **330**:684–690.
- 864 Schmidt, B., and Z. Silvia. 2005. Liste Rouge des espèces menacées en Suisse: Amphibiens.
865 48 pages. OFEFP. Berne.
- 866 Spolsky, C., and T. Uzzel. 1986. Evolutionary History of the Hybridogenetic Hybrid Frog
867 *Rana esculenta* as Deduced from mtDNA Analyses. *Molecular biology and evolution*
868 **3**:44–56.
- 869 Sumida, M., M. Ogata, and M. Nishioka. 2000. Molecular phylogenetic relationships of pond
870 frogs distributed in the Palearctic region inferred from DNA sequences of mitochondrial
871 12S ribosomal RNA and cytochrome b genes. *Molecular phylogenetics and evolution*
872 **16**:278–85.
- 873 Tanaka, T., M. Matsui, and O. Takenaka. 1994. Estimation of phylogenetic relationships
874 among Japanese brown frogs from mitochondrial cytochrome b gene (Amphibia: Anura).
875 *Zoological science* **11**:753–7.
- 876 Templeton, A. R., K. A. Crandall, and F. Sing. 1992. Cladistic Analysis of Phenotypic
877 Associations With Haplotypes Inferred From Restriction Endonuclease Mapping and
878 DNA Sequence Data. 111. Cladogram Estimation. *Genetics* **132**:619–633.
- 879 Uzzell, T., and L. Berger. 1975. Electrophoretic Phenotypes of *Rana ridibunda*, *Rana*
880 *lessonae*, and Their Hybridogenetic Associate, *Rana esculenta*. *Proceedings of the*
881 *Academy of Natural Sciences of Philadelphia* **127**:13–24.
- 882 Vorburger, C. 2001a. Non-hybrid offspring from matings between hemiclinal hybrid
883 waterfrogs suggest occasional recombination between clonal genomes. *Ecology Letters*
884 **4**:628–636.
- 885 Vorburger, C. 2001b. Fixation of deleterious mutations in clonal lineages: evidence from
886 hybridogenetic frogs. *Evolution* **55**:2319–2332.
- 887 Vorburger, C., and H. Reyer. 2003. A genetic mechanism of species replacement in European
888 waterfrogs ? *Conservation genetics* **4**:141–155.
- 889 Vorburger, C., and G. Ribi. 1999. Aggression and competition for shelter between a native
890 and an introduced crayfish in Europe. *Freshwater Biology* **42**:111–119.
- 891 Westphal, M. I., M. Browne, K. MacKinnon, and I. Noble. 2008. The link between
892 international trade and the global distribution of invasive alien species. *Biological*
893 *Invasions* **10**:391–398.
- 894 Wilcove, D. S., D. Rothstein, J. Dubow, A. Phillips, and E. Losos. 1998. Threats to Imperiled
895 Quantifying Species in the United States. *BioSciences* **48**:607–615.
- 896 Williamson, M., and A. Fitter. 1966. The Varying Success of Invaders. *Ecology* **77**:1661–
897 1666.

898 Zeisset, I., G. Rowe, and T. J. C. Beebee. 2000. Polymerase chain reaction primers for
899 microsatellite loci in the north European water frogs *Rana ridibunda* and *R. lessonae*.
900 *Molecular ecology* **9**:1173–1174.

901

Table 1: Sampling sites of *Pelophylax* spp. along the South shore of the Lake of Neuchâtel (Western Switzerland) with the environment (left undisturbed (TE) or affected by human maintenance (EN)) and the geographical coordinates. The total number of frogs collected per location (N) is divided per corresponding sexes (male, female or juvenile).

Location	Environment	Coordinates	N	Males	Females	Juveniles
Chevroux	TE	06°55'03.9" E ; 46°53'59.8" N	0	0	0	0
Chevroux	EN	06°54'26.4" E ; 46°53'45.3" N	26	4	17	5
Gletterens	TE	06°56'20.4" E ; 46°54'44.1" N	26	15	11	0
Gletterens	EN	06°56'12.9" E ; 46°54'32.2" N	28	20	6	2
Motte	TE	06°58'50.6" E ; 46°56'23.6" N	35	16	10	9
Motte	EN	06°59'06.3" E ; 46°56'30.8" N	26	8	18	0

Table 2: The microsatellite markers used in this study distributed in four primers mixes. For each marker both the genome specificity deduced from our samples and previously published, the label of the forward primers, the volume of primers used in our protocols and the annealing temperature are indicated.

Primers mix	Marker	Genome specificity from our samples	Genome specificity as published	Sequences and genome specificity published by	Forward primers labelled with	Volume (µl) of each primers [10mM]	Annealing temperature (°C)
Primers mix 1	Ga1a19 redesigned	L + R	L + R	Arioli 2007; Christiansen 2009	FAM	0.4	58°C
	RICA1b5	L + R	L + R	Garner 2000	ATTO550	0.2	58°C
	RICA5	L + R	L	Garner 2000	HEX	0.3	58°C
	RICA1b6	L + R	L + R	Arioli 2007	FAM	0.2	58°C
	Rrid064A	R	R	Christiansen & Reyer 2009	Dyomics630	0.4	58°C
Primers mix 2	Re2Caga3	R	R	Arioli 2007	ATTO550	0.1	55°C
	Res16	L + R	L + R	Zeisset 2001	FAM	0.2	55°C
	Res20	L	L	Zeisset 2001	HEX	0.5	55°C
	RICA2a34	L + R	L + R	Christiansen & Reyer 2009	Dyomics630	0.5	55°C
Primers mix 3	ReGA1a23	L + R	L	Christiansen & Reyer 2009	ATTO550	0.2	58°C
	Rrid169A	R	R	Christiansen & Reyer 2009	HEX	0.4	58°C
	Res22	L + R	R	Zeisset 2001	FAM	0.2	58°C
	Rrid013A	L + R	L + R	Hotz et al. 2001	FAM	0.1	58°C
	Rrid059A	L + R	L + R	Hotz et al. 2001	ATTO550	0.2	58°C
Primers mix 4	Re1Caga10	L + R	L + R	Arioli 2007	HEX	0.4	58°C
	RICA1a27	L	L	Christiansen & Reyer 2009	Dyomics630	0.4	58°C
	Rrid135A	L + R	R	Christiansen & Reyer 2009	ATTO550	0.3	58°C
	RICA18	L	L	Garner 2000	FAM	0.2	58°C

Table 3: Percentage of correct attribution (true positives) to either *P. ridibundus*, *P. lessonae* or *P. esculentus* based on the corresponding morphological criteria.

	Thighs color	Iris color	Vocal sacs color	Dorsal line length	Metatarsal tubercle shape	“Stub test”
<i>P. ridibundus</i>	92.3	84.6	100	83.3	61.5	46.2
<i>P. lessonae</i>	70.6	44.1	93.8	44.8	63.6	72.7
<i>P. esculentus</i>	83.3	34.6	42.1	46.2	79.5	67.5

Table 4: Percentage of false attribution (false positives) to either *P. ridibundus*, *P. lessonae* or *P. esculentus* based on the corresponding morphological criteria.

	Thighs color	Iris color	Vocal sacs color	Dorsal line length	Metatarsal tubercle shape	“Stub test”
<i>P. ridibundus</i>	6.3	46.4	38.9	33.6	11.7	1.8
<i>P. lessonae</i>	6.6	0.0	2.4	9.5	4.4	25.6
<i>P. esculentus</i>	23.4	42.6	5.3	42.9	34.8	34.8

Figure 1: Snout-to-vent length (SVL) for each species of the sampled water frog, males and females combined. The boxes represent the quartiles (50% of the data) with the median (black line), dotted lines (whiskers) extend to the most extreme measures that are no more than 1.5 times the interquartile range from the box; measures outside of this range are shown as points.

Figure 2: Snout-to-vent length (SVL) by tibia length (TL) ratio for each species of the sampled water frog, males and females combined. The boxes represent the quartiles (50% of the data) with the median (black line) and whiskers extend to the most extreme measures that are no more than 1.5 times the interquartile range from the box; measures outside of this range are shown as points.

Figure 3: Individual Bayesian clustering run with the most likely number of clusters $K = 3$. No prior information was used to distinguish between the three species of the *P. esculentus* complex. Each species is separated by a black line. Each bar represents one individual. The “ridibundus” cluster is shown in blue, the “lessonae” cluster in red and the “esculentus” cluster in green.

Figure 4: Biplot of factor scores obtained from the principal component analysis performed on the alleles frequencies of the three species of the *P. esculentus* complex: *P. ridibundus* in blue, *P. esculentus* in green and *P. lessonae* in red. Factor 1, in the X, axis explains 31.22% of the variance and the factor 2, in the Y axis explains 9.26% of the variance.

Figure 5: Biplot of factor scores obtained from the principal component analysis performed on the alleles frequencies of the R (ridibundus) and L (lessonae) genome from the two parental species, respectively *P. ridibundus* and *P. lessonae* and the R* (ridibundus) and L*

(*lessonae*) components of the *P. esculentus* genomes. The R genome is shown in blue, the L genome in red, the R* genome in black and the L* genome in yellow. Factor 1, in the X, axis explains 37.78 % of the variance and the factor 2, in the Y axis explains 5.27% of the variance.

Figure 6: Genetic diversity measured by the expected heterozygosity of the R and L genome from the two parental species, respectively *P. ridibundus* and *P. lessonae* and the R* and L* components of the *P. esculentus* genomes. The boxes represent the quartiles (50% of the data) with the median (black line) and the whiskers extend to the most extreme measures that are no more than 1.5 times the interquartile range from the box; the measures outside of this range are shown as points.

Figure 7: Parsimony network of haplotypes based on the alleles of the 12 *ridibundus*-like (R*) hemiclones of the *P. esculentus* individuals and the geographical distribution of these hemiclones. Each hemiclone is represented by a unique and different color. (a) Each point corresponds to one mutational step (SNP, insertion or deletion) between hemiclones. Each hemiclone is shown by an oval with its size corresponding to the frequency of the hemiclone in the *P. esculentus* populations. (b) The frequency and the geographical distribution of the hemiclones are represented by a pie chart located on each sampling site. The size of the pie chart corresponds to the frequency of *P. esculentus*.

Figure 8: Phylogeny of the cytochrome-b fragment from mtDNA of European and Near East *Pelophylax* spp. using a maximum likelihood analysis. In bold are the individuals from our sampling sites, in the Grande Cariçaie. The *P. ridibundus* indication referred to the individuals defined as *P. ridibundus* with the microsatellites amplification. Values of

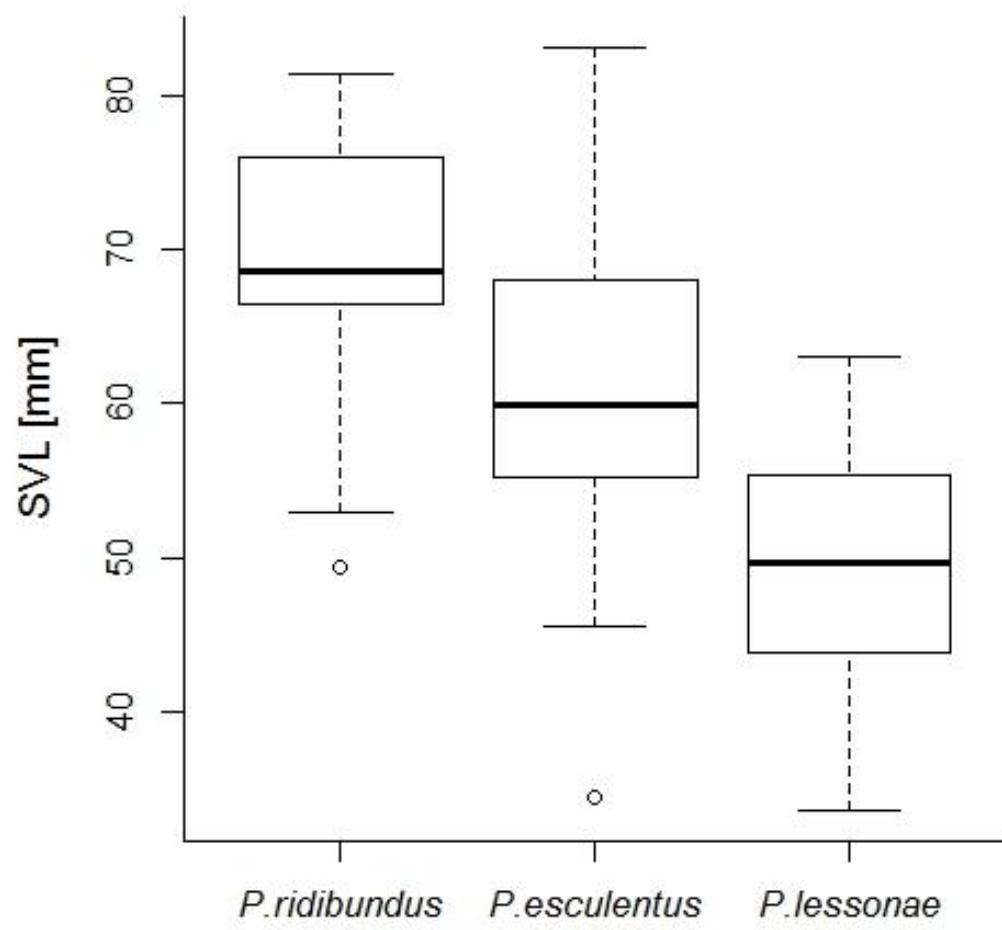
976 bootstrap support for the major branches (percentage of 1000 replications) are indicated in
977 branches.

978

979 **Figure 9:** Distribution of the three species among the study area. The proportion of *P.*
980 *ridibundus* (blue), *P. lessonae* (red) and *P. esculentus* (green) per sampling site is represented
981 by a pie charts located on each sampling site.

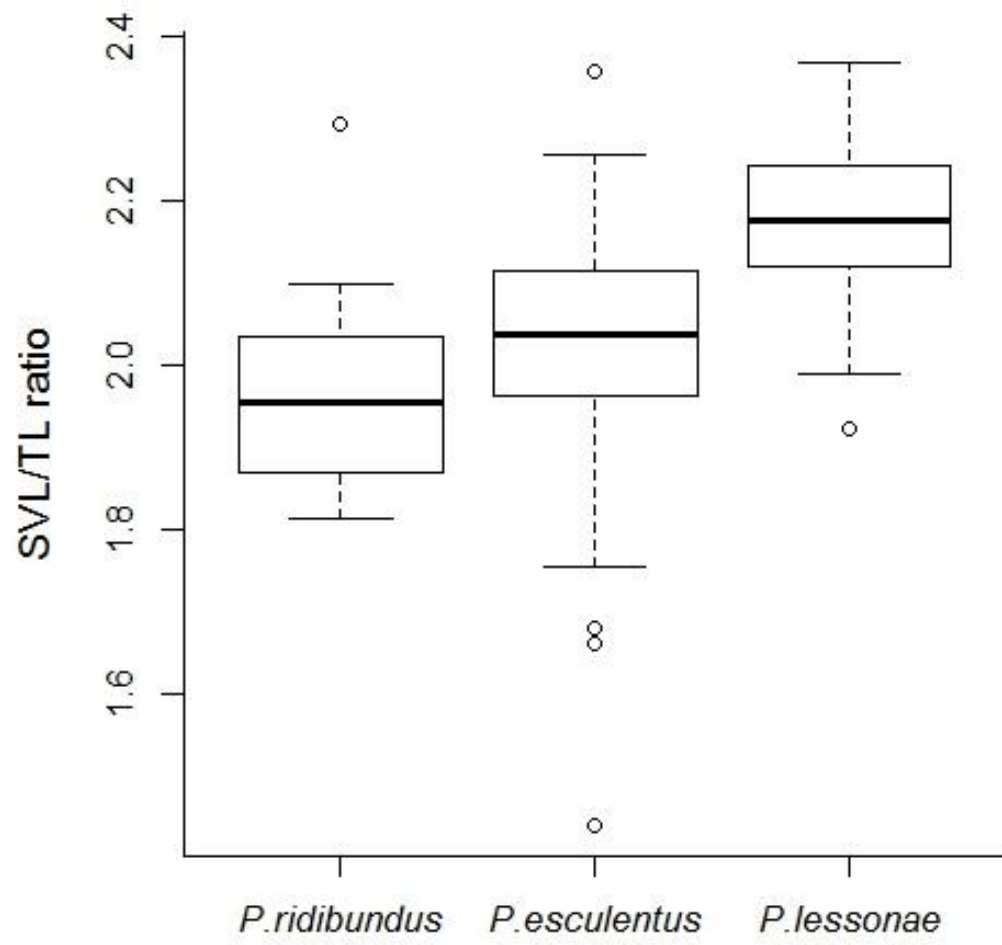
982

983 **Figure 10:** Relative frequency of *P. ridibundus*, *P. esculentus* and *P. lessonae* found on the
984 two environments: affected by human maintenance, called EN zones (grey) or undisturbed,
985 called TE zones (white).



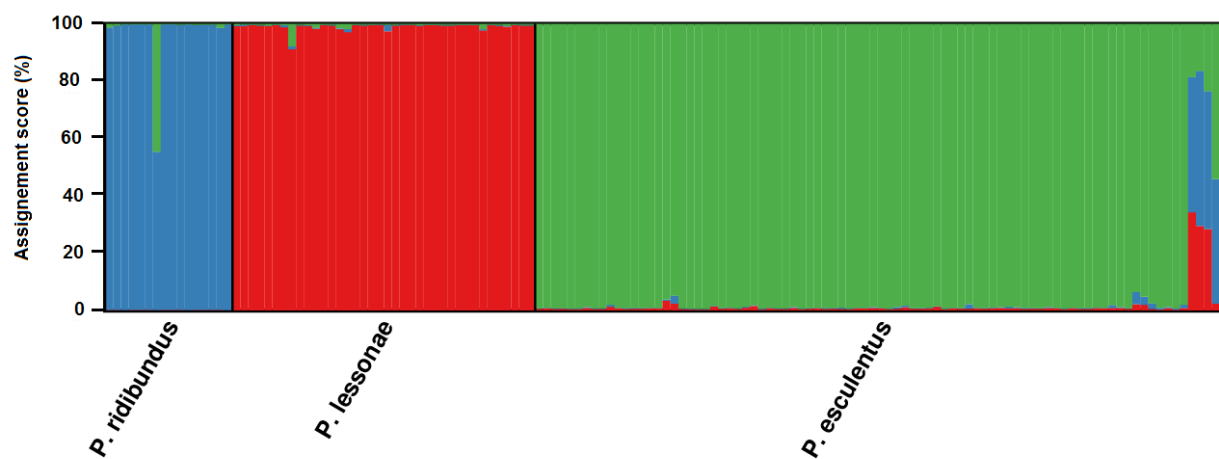
986

987 **Figure 1**



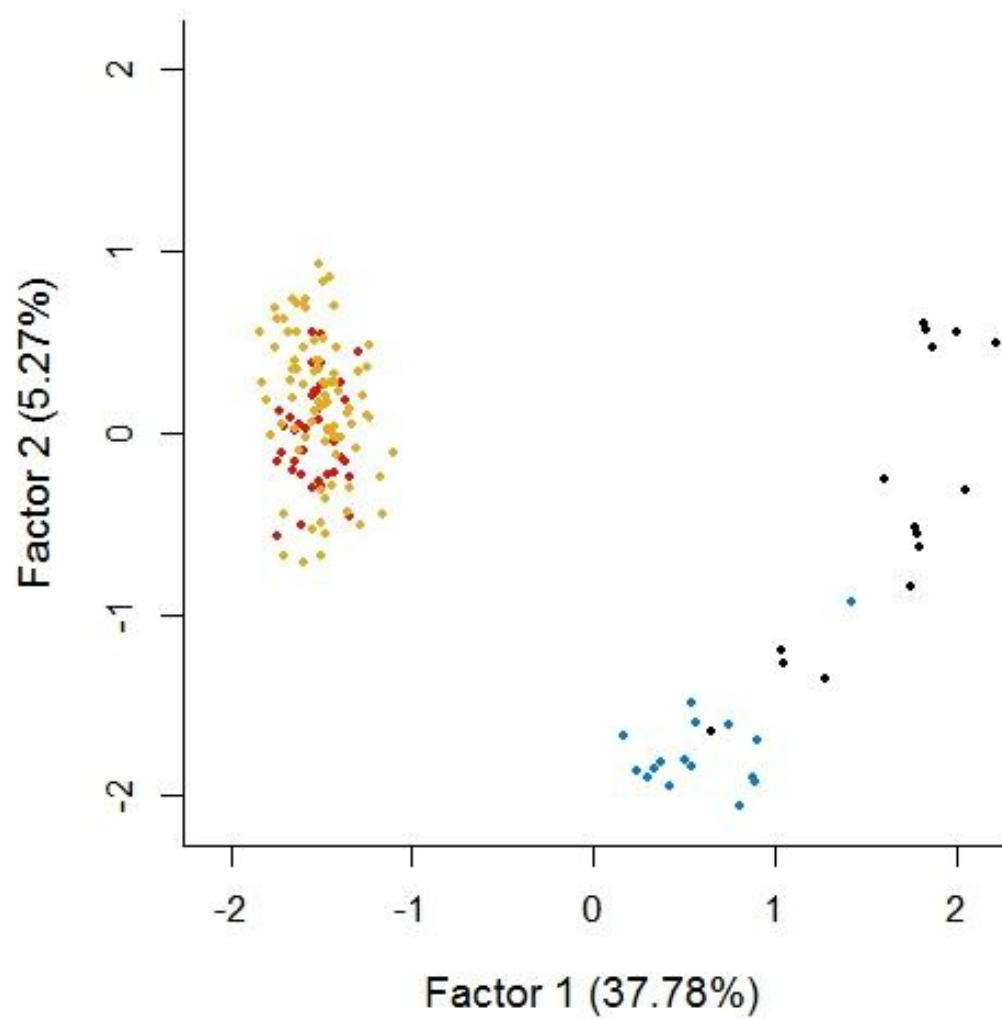
988

989 **Figure 2**



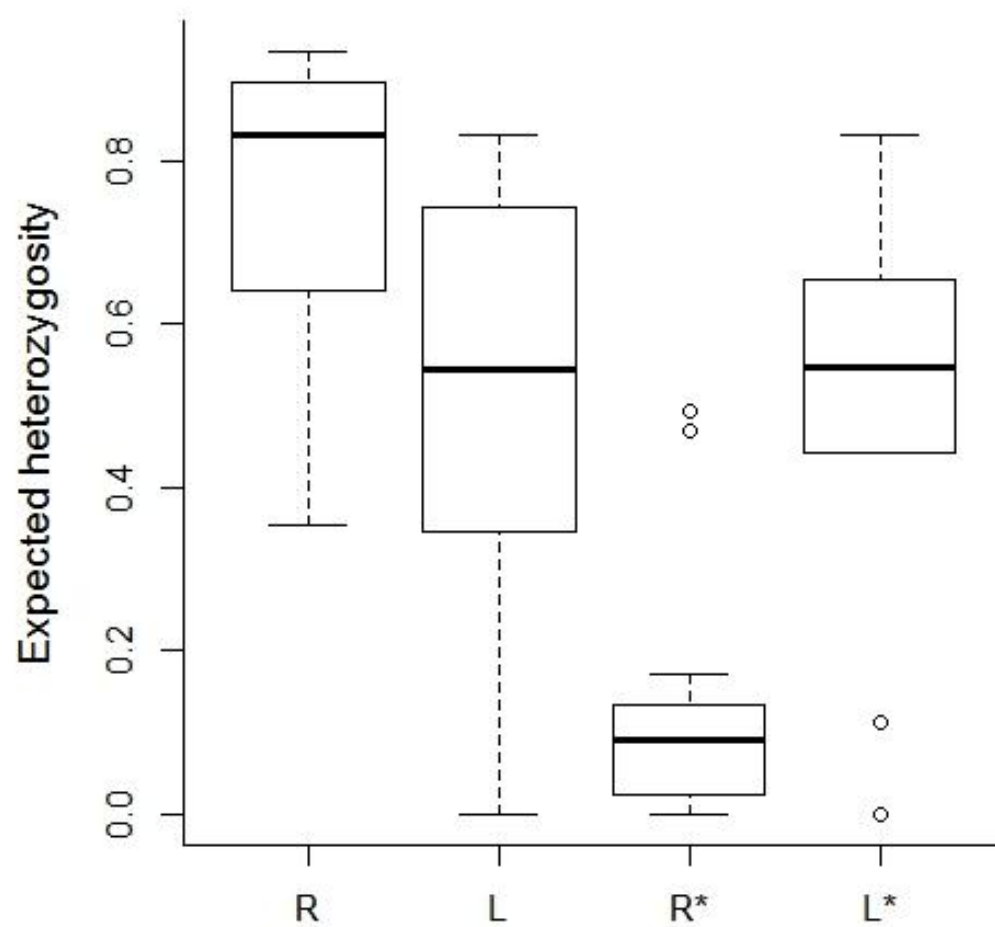
990

991 **Figure 3**



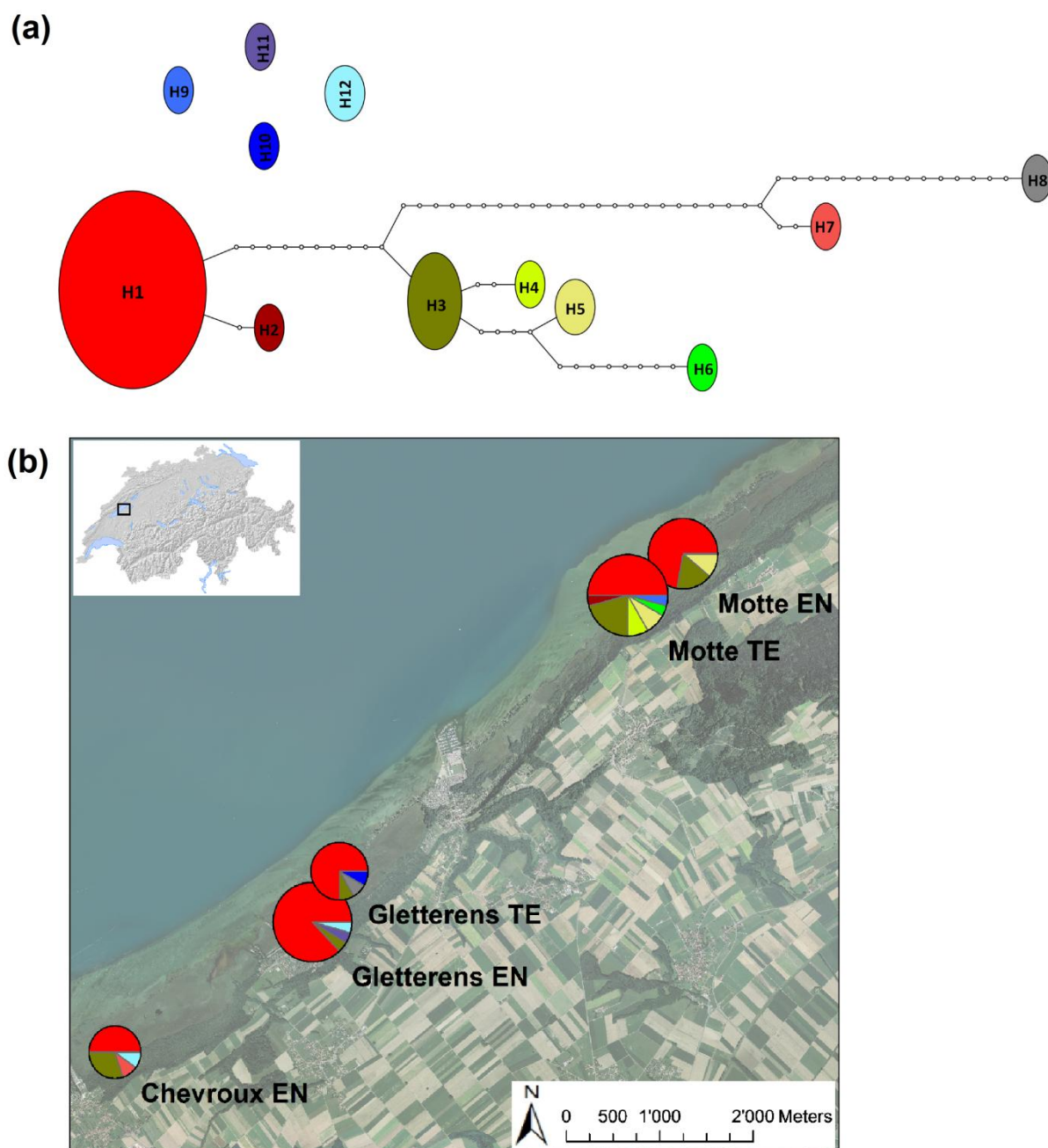
994

995 **Figure 5**



996

997 **Figure 6**



998

999 **Figure 7**

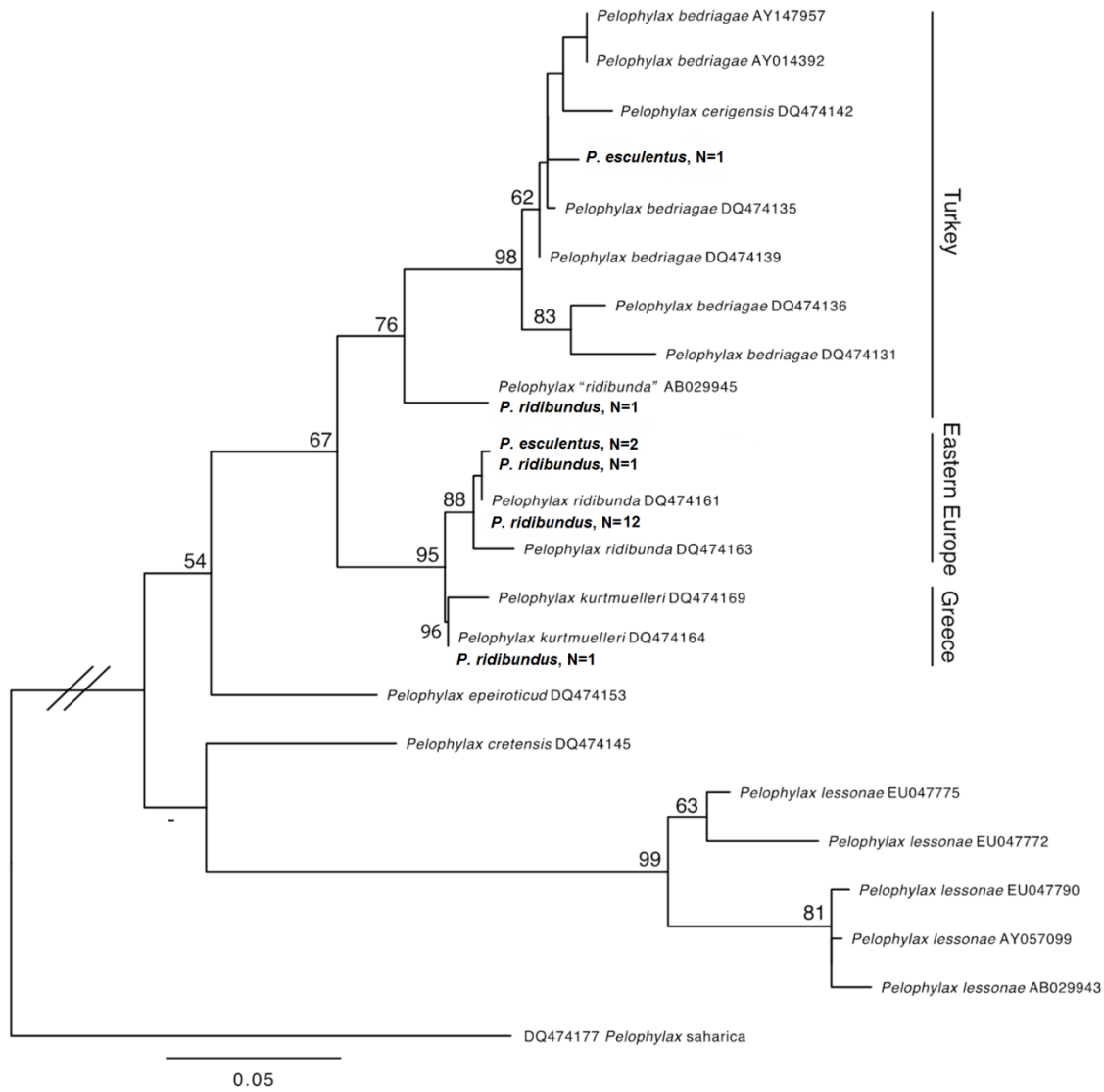
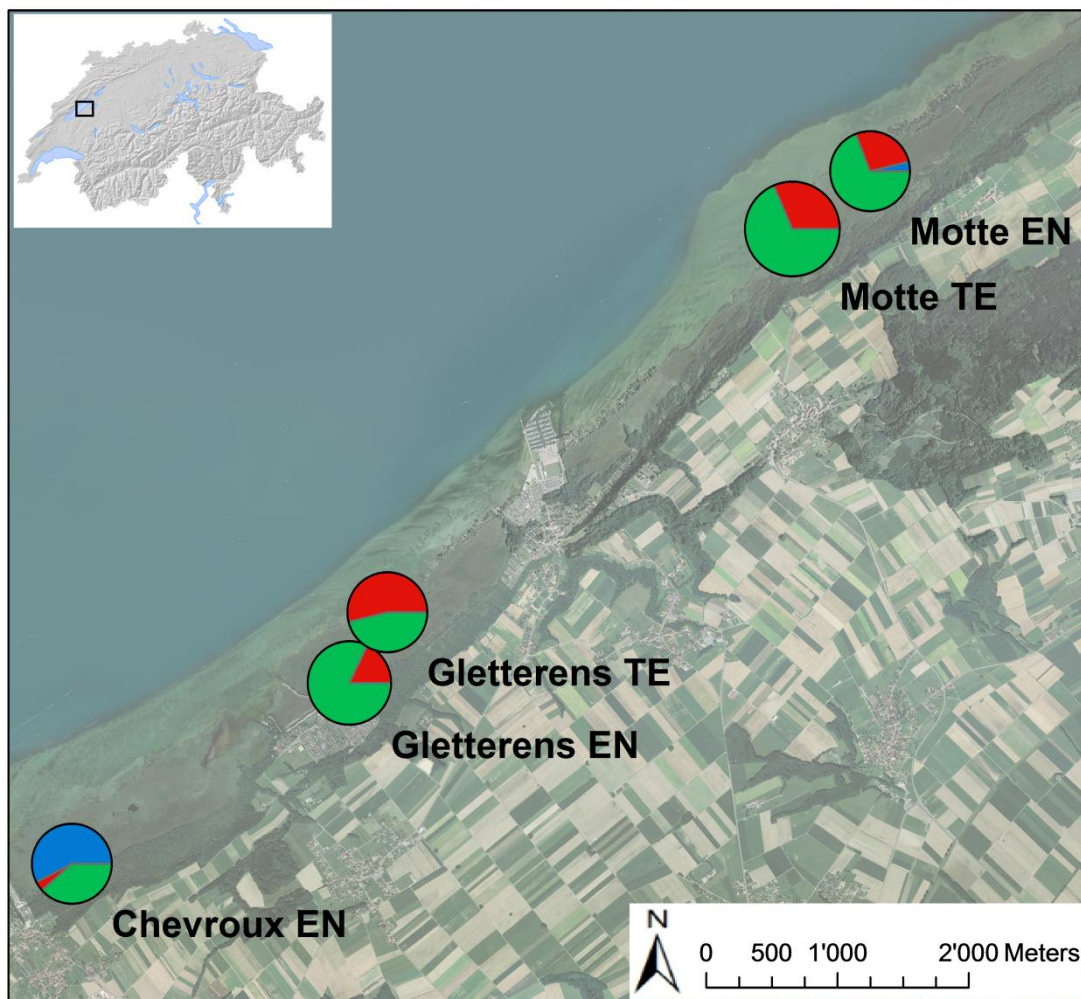
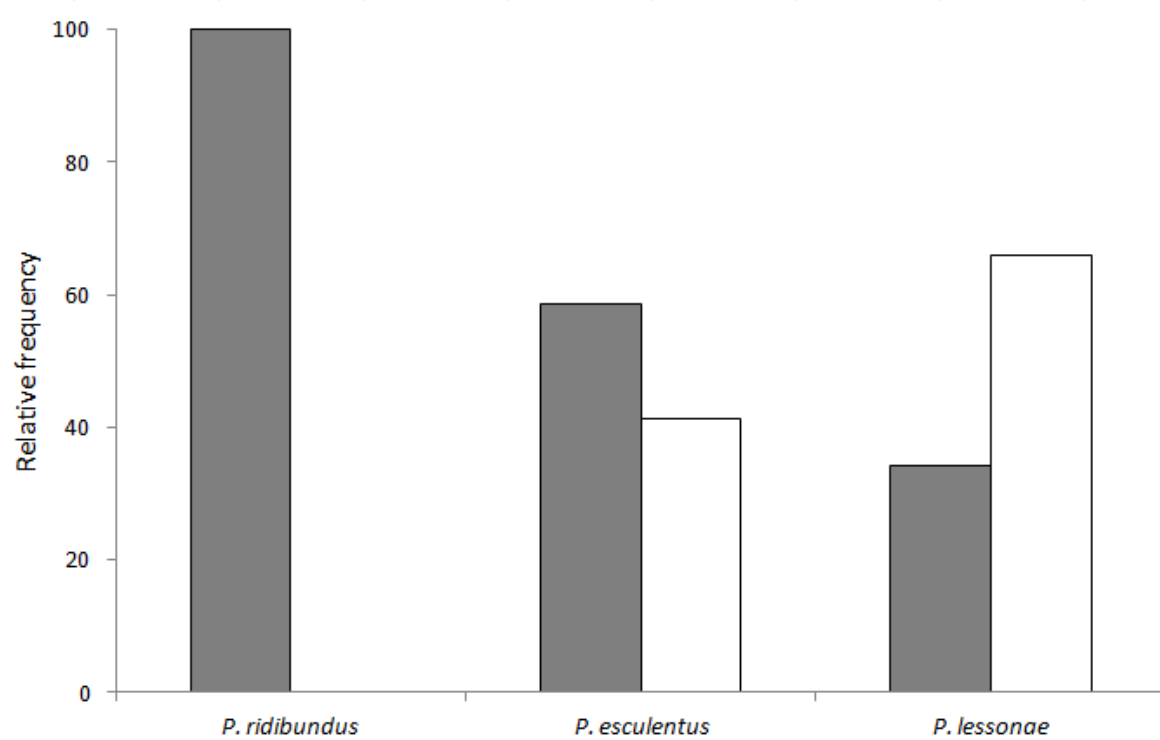


Figure 8



1002

1003 **Figure 9**

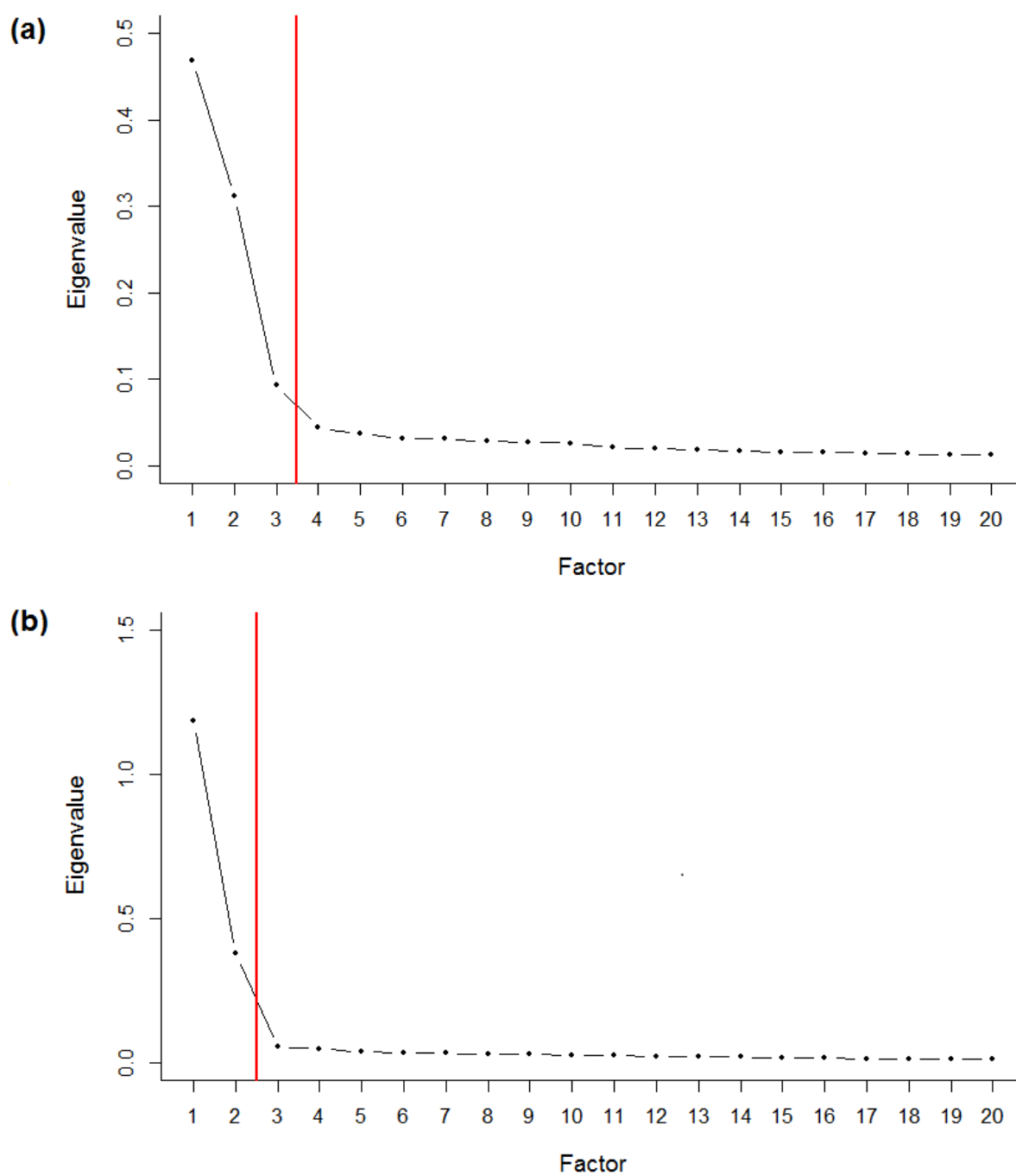


1004

1005 **Figure 10**

Supporting information

Figure S1: Scree plot of the eigenvalues from the 20 first factors of the two performed principal component analysis (PCA). The position of the red line referred to the Cattell's scree test which is used to select the significant factors of a PCA: the factors located after the elbow (indicated by the red line) cannot be used since they do not explain enough variance. Panel (a) represent the scree plot for the first PCA performed on the three species and the panel (b) represent the scree plot for the second PCA performed on the genome of the two parental species (R and L) and the R* and L* components of *P. esculentus*.



1015

1016 **Figure S1**