SALAMANDRA **58**(3): 218–230
15 August 2022 | ISSN 0036–3375

SALAMANDRA
German Journal of Herpetology

A new PCR-RFLP method for the identification of parental and hybridogenetic western European Water Frogs, including the *Pelophylax perezi-grafi* system

Angélica Cuevas[1,2,3,*], Cécile Patrelle[4,5,*], François Ciavatti[1,3], Thomas Gendre[3], Patricia Sourrouille[1], Philippe Geniez[6], Paul Doniol-Valcroze[1,3,**] & Pierre-André Crochet[1,**]

[1)] CEFE, CNRS, Univ Montpellier, EPHE, IRD, Montpellier, France

[2)] CEES – Centre for Ecological and Evolutionary Synthesis, Department of Bioscience, University of Oslo, Oslo, Norway

[3)] Conservatoire d'espaces naturels Occitanie, Bâtiment le Thèbes, 26 allée de Mycènes, 34000 Montpellier, France

[4)] EA4688 "Vecpar", UFR de Pharmacie, Université de Reims Champagne-Ardenne, 51 rue Cognacq-Jay, 51096 Reims Cedex, France

[5)] NaturAgora Développement, Bureau d'Etudes des usagers de la nature, 1 chemin du pont de la Planche, 02000 Barenton-Bugny, France

[6)] CEFE, EPHE-PSL, CNRS, Univ Montpellier, IRD, Biogéographie et Ecologie des Vertébrés, Montpellier, France

[*] These authors each contributed equally to this work

[**] These authors each contributed equally to this work

Corresponding author: Paul Doniol-Valcroze, ORCID: 0000-0002-2523-5979, e-mail: pauldoniol-valcroze@orange.fr

Manuscript received: 27 January 2022
Accepted: 17 May 2022 by Stefan Lötters

Abstract. Western European Water Frog species can be morphologically very similar, especially *Pelophylax ridibundus* and *P. perezi*. As these species, together with *P. lessonae*, produce hybridogenetic lineages with highly variable phenotypes, the morphological identification of taxa in this genus remains challenging. This is especially problematic in the current context of the spread of the invasive *P. ridibundus* complex in native populations across Western Europe. In this study, we present a novel molecular method for the identification of western European Water Frogs (genus *Pelophylax*) based on the PCR-RFLP genotyping of nuclear species-specific loci. Using a combination of restriction digestions of species-specific genetic markers, this method facilitates reliable specific identification in the *Pelophylax perezi-grafi* system (PG system) and the *Pelophylax lessonae-esculentus* system (LE system). Our method provides several advantages compared to other genetic techniques previously used. It relies on PCR amplification allowing non-invasive sampling followed by RFLP to assess diagnostic SNPs in the nuclear genome, rendering obsolete the need for sequencing. Lastly, it is easier to replicate than methods based on microsatellites. Importantly, this novel protocol allows the identification of all five taxa of Water Frogs that occur in Western Europe.

Key words. Amphibia, Anura, Ranidae, species identification, hybridogenetic species complex, conservation, invasive species.

Introduction

The western Palearctic Water Frogs (genus *Pelophylax*) constitute one of the best-known examples of hybridogenetic species complexes, giving rise to several hybridogenetic systems across Europe (Günther 1990). These complexes are composed of a fixed bisexual hybrid, originating from hybridization events between *Pelophylax ridibundus* and another species, usually coexisting with at least one of its parental species. The hybrid taxa are maintained by a hemiclonal reproductive system where, during gametogenesis, the hybrid taxon systematically excludes the genome of one of the parental species, frequently the genome from the parent species that lives in sympatry with it. Pre-meiotic DNA synthesis is followed by a compensatory duplication of the remaining genetic material, which generates gametes that contain the genome of just one of the parental species (Vinogradov & Chubinishvili 1999). Thus, recombination usually does not take place except in triploids (Christiansen & Reyer 2009) and the genome that is included in the gametes of the hybrid taxon (typically but not always a *P. ridibundus* hemigenome) is transmitted clonally. Consequently, only the somatic cells of the hybrid contain genetic material of both parents, while gametes are clones of a sin-

gle parent's genetic material. Given this particular reproductive system, hybrid maintenance relies on backcrossing to the parental species whose genome is eliminated: crossings between the hybrid taxa and the co-occurring parental species produce hybrid taxa, while crossings between individuals of the hybrid taxa will typically fail, as the clonally transmitted genome accumulates too many deleterious mutations to function in a diploid state (Graf & Polls Pelaz 1989, Dufresnes & Mazepa 2020). The hybrid taxa are thus the avatars of hemiclones that reproduce by "parasitizing" a sexual species; these hemiclones function well when combined with the sexually reproducing genome of another species, but are usually unable to persist on their own (see Dufresnes & Mazepa 2020 for a review).

The most common and most studied system is the *Pelophylax lessonae-esculentus* system (the LE system, Berger 1973, Heppich 1978), found throughout Europe from western France to the Volga-Urals region of Russia (Graf & Polls Pelaz 1989, Gasc et al. 2004). This species complex, composed of *P.* kl. *esculentus*, a fixed bisexual hybrid resulting from hybridization events between *P. ridibundus* and *P. lessonae*, and the parental species *P. lessonae,* upon which the hybrid taxon depends for its reproduction (Berger 1973, Heppich 1978, Graf & Mueller 1979, Uzzell et al. 1980), has allowed the characterization of the particular reproductive system of the Water Frog complexes. Several other hybridogenetic systems involving these three taxa are distributed across Europe (Graf & Polls Pelaz 1989, Dufresnes & Mazepa 2020). They differ in population composition, especially with regard to the species coexisting with the hybrid, and in the genome transmitted by the hybridogenetic taxon (*ridibundus* in most systems, including the LE system, but rarely *lessonae* in the *ridibundus-esculentus* RE system, see details in Dufresnes & Mazepa 2020). They also differ in the ploidy of the hybrid *P.* kl. *esculentus* (mainly diploid individuals with one *P. lessonae* and one *P. ridibundus* genome in the LE system or including triploids with two genomes of either *P. lessonae* or *P. ridibundus* in other systems) and of the gametes produced by *esculentus* (mainly haploid in the LE system, often diploid in some other systems). Lastly, some pure populations of the *P.* kl. *esculentus* hybrid taxon are known to persist through the coexistence of LR diploids with LLR or LRR triploids (the *esculentus-esculentus* EE system). In some of these systems recombination and sexual reproduction have even been restored, giving rise to sexually reproducing populations of the hybrid taxon *P.* kl. *esculentus* (see Christiansen & Reyer 2009).

Another taxon of hybrid origin is *Pelophylax* kl. *grafi,* a fixed bisexual hybrid between *Pelophylax perezi* and *P. ridibundus*. Coexisting with *P. perezi* to form the *perezi-grafi* system (PG system), it is commonly encountered in the Mediterranean and Atlantic regions of southern and southwestern France and in northeastern Spain (Graf et al. 1977, Uzzell & Tunner 1983, Graf & Polls-Pelaz 1989, Hotz et al. 1994, Crochet et al. 1995, Pagano et al. 2001a, b, Daf et al. 2006). The current consensus is that natural hybridogenetic populations were originally constituted exclusively of individuals of *P. perezi* and *P.* kl. *grafi*, except in the restricted area of overlap with the LE system (see Daf et al. 2006). The current widespread occurrence of *P. ridibundus* within the distribution of the PG system in France is assumed to be the result of expansions from spots of recent human-mediated translocations that triggered the invasion of this region by *P. ridibundus* (Pagano et al. 2001a, b, 2003, Holsbeek & Jooris 2010, Holsbeek et al. 2010, Dufresnes et al. 2017a). As in the LE system, the hybrid taxon is thought to persist by backcrossing with *P. perezi* and by transmitting only the *P. ridibundus* genome.

The origin of *P.* kl. *grafi* remains speculative and two main hypotheses are still under discussion (Arano et al. 1995). One hypothesis proposes direct hybridization between *P. perezi* and *P. ridibundus* during past periods of range overlap. Another hypothesis involves hybridization between *P. perezi* and *P.* kl. *esculentus* (with the latter acting as the donor of the *P. ridibundus* hemigenome) where the distributions of the LE and PG systems currently overlap in western France (Pagano et al. 2001a, b, Daf et al. 2006). Alternatively, *P. ridibundus* could be native (and not introduced as is widely believed) to southeastern France, where it is currently the only *Pelophylax* in most of Provence and Alps regions (Lescure & de Massary 2012, own unpublished data) and comes into contact with the PG system along the lower Rhône valley. However, Hotz et al. (1994) argued that the low clonal diversity of *P.* kl. *grafi* compared with *P.* kl. *esculentus* and the fact that the system was restricted to the eastern limits of the geographic distribution of *P. perezi* supported the idea that *P.* kl. *esculentus* is the source of the *P. ridibundus* genome in the PG system.

One of the major issues precluding a better understanding of the ecology and evolution of the European Water Frogs is the difficulty of reliable species identification (Günther 1990, Pagano & Joly 1999). In the systems involving *P.* kl. *esculentus*, morphology and bioacoustics differ between the parental species and the hybrid (Günther 1990, Plötner 2005), but gene dosage effects in triploid hybrids that contain a double genetic makeup, either from *P. ridibundus* or from *P. lessonae*, blur the validity of morphological and acoustic delimitations (Plötner 2010, Hoffmann & Reyer 2013, Mayer et al. 2013). In the case of *P.* kl. *grafi*, phenotypic differentiation is even more difficult, since *P. perezi* and *P. ridibundus* are extremely similar (Günther 1990) and bioacoustic differences have not been well established yet. Therefore, we still do not know how to confidently identify *P.* kl. *grafi* based on morphological or acoustic characters.

Accurate identification of interspecific hybrid taxa and parental species is not only crucial for understanding the complex reproductive system of hybridogenetic species, but it also represents an important issue in conservation biology (e.g., Congiu et al. 2001). Several studies suggest that the presence of *P. ridibundus* in western European countries (e.g., in the south of France, Spain, Switzerland) is the result of multiple human-mediated translocations (Arano et al. 1995, Pagano et al. 1997, Pagano et al. 2001a, 2003, Vorburger & Reyer 2003, Zeisset & Beebee 2003,

Holsbeek & Jooris 2010, Holsbeek et al. 2008, 2010). *Pelophylax ridibundus* is currently responsible for the decrease of native Water Frog species in many areas and it is considered a threat to other amphibian species in the region (Arano et al. 1995, Pagano et al. 2003, Vorburger & Reyer 2003, Schmeller et al. 2007, Holsbeek & Jooris 2010, Hauswaldt et al. 2012, Leuenberger et al. 2014). Therefore, molecular identification methods of Water Frog complexes are essential tools for monitoring the spread of the invasive *P. ridibundus*, for establishing conservation measures, as well as for ecological and evolutionary studies.

In European Water Frogs, interspecific mitochondrial gene transfer (Spolsky & Uzzell 1984, 1986, Plötner et al. 2008) and the presence of hybridogenetic taxa make mitochondrial markers unreliable, because they are maternally inherited and thus cannot identify hybrid taxa. Such problems do not occur with nuclear markers, and the molecular identification of Water Frogs has long been based on allozymes (Uzzell & Berger 1975, Uzzell & Hotz 1979, Hotz & Uzzell 1982, Günther & Plötner 1994, Hotz et al. 1994, Buckley et al. 1994, Beerli et al. 1996). However, this method comes with technical challenges such as the need for large amounts of tissue and the necessity to preserve samples at very low temperatures upon collection; it also suffers from poor repeatability, as alleles are defined by their relative migration speed, rendering identification of specimens without reference samples difficult. Nuclear DNA-based methods that alleviate some of these difficulties have been proposed, using microsatellites (Garner et al. 2000, Zeisset et al. 2000, Hotz et al. 2001, Christiansen 2005, Sánchez-Montes et al. 2016, Sagonas et al. 2020), Southern blot analysis of enzyme-digested genomic DNA (Tognarelli et al. 2014), or random amplified polymorphic DNA markers (RAPD, Zeisset & Beebee 1998) as alternatives. Microsatellites, however, suffer from a lack of diagnostic loci and usually require assignment analyses in large samples to allow species identification. RAPD approaches suffer from low repeatability, while Southern blot methods, which do not rely on DNA amplification, require large amounts of tissue. None of these methods is thus suitable for non-invasive specific identification of single specimens or are easily transferable from one lab to another.

Alternative methods based on PCR-amplification of diagnostic nuclear loci have thus been proposed. A PCR-RFLP method by Patrelle et al. (2011) and a PCR-migration method by Hauswaldt et al. (2012) have been developed to identify the Water Frogs involved in the LE complex (*P. ridibundus*, *P. lessonae* and *P.* kl. *esculentus*), with the latter method having been successfully adapted to separate *P. ridibundus* from *P. shqipericus* (Vucić et al. 2018). However, these methods have not been tested in the *Pelophylax perezi-grafi* (PG) system and are thus for the moment unsuitable in large areas of Western Europe. For the PG system, the only DNA-based method currently available uses a set of 16 microsatellite loci (Sánchez-Montes et al. 2016); like allozymes, microsatellites can suffer from poor transferability between labs, and identification of specimens without reference samples is difficult. Moreover, given that in some regions the LE system and the PG system overlap, a single reliable method for the identification of the three parental species and the two hybrid taxa involved in the PG and LE systems is necessary. A technique based on nuclear markers with biparental inheritance and where genotyping is possible without the need for reference samples will provide a crucial tool for the reliable identification of the different Water Frog taxa in these two species complexes.

In this study we present a fast, easy and non-invasive PCR-RFLP method that allows the separation of the five widespread European Water Frog species: *Pelophylax ridibundus*, *P. perezi*, *P. lessonae* (including *bergeri*, which is sometimes treated as a valid species) and the hybridogenetic *P.* kl. *esculentus* and *P.* kl. *grafi*. Because recombination in the hybridogenetic lineages cannot be totally excluded (Christiansen & Reyer 2009), we aimed at providing at least two independent markers to diagnose each taxon. We first selected nine nuclear loci found to be informative at interspecific level by previous studies or already published for *Pelophylax* species identification, including the two loci used by Patrelle et al. (2011) and Hauswaldt et al. (2012, Table 1). For those where amplification was successful, sequences were generated for reference samples of the three parental species (*P. ridibundus*, *P. perezi* and *P. lessonae*, see Table 2). This allowed us to identify five nuclear genes with potentially diagnostic substitutions between at least two of the parental species (Table 3). After screening potential enzymes with restriction sites around the diagnostic positions, we tested enzymatic digestions on large series of samples representing the five target taxa to validate the diagnostic character of these substitutions. Because most of the samples used in this last step had not been identified previously by other methods, we employed at least two genes to identify each sample and assessed the concordance between the identification provided by each locus to validate the identification. This concordance between markers also provides a test for the reliability of the diagnostic sites and allows detecting rare recombination events (Christiansen & Reyer 2009).

Material and methods

Sampling

We analyzed a total of 1,111 individuals belonging to all five Western European Water Frog species, *Pelophylax ridibundus*, *P. lessonae*, *P.* kl. *esculentus*, *P. perezi* and *P*. kl. *grafi* (Supplementary File 1). Water Frog sampling in France had been authorized by the following permits from competent regional administrations: AP 02/2007 & 03/2007 by the Préfecture de la Vendée; AP 07-2549 by the Préfecture de la Sarthe; AP by the Préfecture des Ardennes in 2007; AP-2013-06; AP-2013 274-0002; AP-2013-325.0010; AP-2013-189.0001; AP 2019-s-07 by the Préfecture de l'Hérault et du Gard, and samples from other countries were retrieved from the BEV tissue collection.

Table 1. List of the nine nuclear genes evaluated in this study with their corresponding primers' names, sequences and sources. Annealing temperature is only given when amplification was successful. The "Result" column identify which genes were selected. For the selected genes, the last column reports the restriction enzymes that were retained after our tests. * Details for the digestion of NTF3 gene by the XapI enzyme are given in Supplementary Files 2 (PCR-RFLP method) and 3 (alignment).

Gene	Primers	Sequence 5' – 3'	Source	Annealing temperature	Result	Restriction enzyme
CXCRA (Chemokine receptor type 4)	CXCR4_Rana.F CXCR4_Rana.R	5'-TTCACCCTTCCATTCTGGTC-3' 5'-GCCACGGCTTCTGTGATAG-3'	Newman et al. (2016)	–	Not selected(No amplification)	–
SIA (Seven in absentia)	SIA1 (F)SIA2 (R)	5'-TCGAGTGCCCCGTGTGYTTYGAYTA-3' 5'-GAAGTGGAAGCCGAAGCAGSWYTGCATCAT-3'	Frost et al. (2006)	–	Not selected(No amplification)	–
ITS2 (Internal transcribed spacer 2)	ITS2-hinITS2-rück	5'-GGATCACTCGGCTCGTGCGTCGATGAAG-3' 5'-CGGGGATTCGGCGCTGGGCTCTTCCC-3'	Patrelle et al. (2011)	–	Not selected(No amplification in *perezi* or *grafi*)	–
SAI-1 (Serum albumin intron 1)	PEL-SA-F1PEL-SA-R2	5'-TCCATACAAATGTGCTAAGTAGGTT-3' 5'-GACGGTAAGGGGACATAATTCA-3'	Hauswaldt et al. (2012)	60°C	Not selected (Length differences not entirely diagnostic)	–
NTF3 (Neurotrophin 3)	NTF3-FNTF3-R	5'-TCTTCCTTATCTTTGTGGCATCCACGCTA-3' 5'-ACATTGRGAATTCCAGTGTTTGTCGTCA-3'	Newman et al. (2016)	61°C	Not selected* (not fully diagnostic site)	XapI(RzAATTY)
RAG1 (Recombination activating protein 1)	MartFL1AMPR-1rana	5'- AGCTGCAGYCAGTACCACAAAATG-3' 5'-AATTCAGCTGCATTTCCAATGTC-3'	Newman et al. (2016)	59°C	Selected	Dra II(RGzGNCCY)BSM I (GAATGCNz)
TYR1 (Tyrosinase 1)	TYR1bRana (F) TYR1gBufo(R)	5'-AGGTCCTCTTRAGCAAGGAATG-3' 5'-TGCTGGGCATCTCTCCAGTCCCA-3'	Newman et al. (2016)	54°C	Selected	Bal I(TGGzCCA)
POMC (Pro-opiomelanocortin A)	POMC_SH_Rana_F1POMC_SH_Rana_R	5'-ATAYGTCATGAGCCACTTCCGCTGGAA-3' 5'-GTCTTTGGGTGGWCCTTCCCATCG-3'	Susanne Hauswaldt (pers. com.)	61°C	Selected	Hinc II(GTYzRAC)
CMYC2 (Cellular myelocytomatosis intron 2)	CMYC2 MYEL F1 CMYC2 MYEL R3	5'-CAGTGAATGACAGCATTTCCAG-3' 5'-GTCAAAGCCTTCAAAGACCATTG-3'	Dubey & Dufresnes (2017)	60°C	Selected	SspI(AATzATT)

DNA sampling and extraction

Two types of samples were used in this study: 1) mostly unidentified tissue samples stored in 95% ethanol at -20°C in the tissue collection of the Biogéographie et Ecologie des Vertébrés (BEV) team from the UMR5175 – Centre d'Ecologie Fonctionnelle et Evolutive (Montpellier, France); these samples include both toe clippings or buccal swabs from live frogs (not preserved after sampling) or tissue samples from vouchers previously preserved in 95% ethanol in the BEV voucher collection; and 2) a series of 83 samples from the five species originating from the Ardennes (northeastern France), southeastern and southwestern France, and northeastern Spain, collected between 1996 and 2008, stored in ethanol in the Pagano and Patrelle tissue collection by the team "EA 4688 USC ANSES – VECPAR" (Université Reims Champagne-Ardenne, France). These 83 samples had been previously identified using specific allozymic markers or a PCR-RFLP-based method on ITS2 (Pagano et al. 1997, 2001a, b, Daf et al. 2006, Patrelle 2010, Patrelle et al. 2011).

In a first step, we selected a series of *P. ridibundus*, *P. lessonae* and *P. perezi* samples from the BEV sample collection, originating from areas where no other species are known to occur. These reference samples and samples from the Pagano and Patrelle collection, previously identified with other methods, were used to screen nuclear loci for candidate diagnostic single-nucleotide polymorphisms (SNPs). In a second step, a large series of DNA samples from the BEV and the Pagano and Patrelle tissue collections were used to assess the reliability of the method and to provide genetic identification to local naturalists and managers of protected areas (see details in Results and in Supplementary File 1).

We used three different DNA extraction protocols. 1) A modification of the NaCl extraction protocol from Brufold et al. (1992), using an extraction buffer containing Tris (1M, pH 8), NaCl (5M), EDTA (0.5M, pH 8) and dH_2O.

Samples were incubated in a mix of 410 µl of extraction buffer, 80 µl of SDS (10%), and 10 µl of Proteinase K at 65°C for 30 min before centrifugation at 13,000 rpm for 5 min. A volume of 180 µl of NaCl (5M) was added to the supernatant and centrifuged again at 13,000 rpm for 5 min. Finally, 420 µl of cold isopropanol was added to the resulting supernatant and a final centrifugation of 13,000 rpm for 5 min was performed. We washed the DNA pellet with 250 µl of EtOH 80% before elution in 100 µl of water. 2) Secondly, we used the Extract-N-Amp™ Plant PCR Kit (Sigma-Aldrich, Saint-Louis, USA). Samples were mixed with 50 µl of E-Buffer, and after incubation at 95°C for 10 min, 50 µl of D-Buffer were added. 3) Lastly, some samples were extracted with the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. All three extraction protocols showed good results for buccal swabs. However, for tissue samples, the NaCl extraction protocol was chosen given its better performance.

Nuclear marker selection

To identify loci allowing species identification, nine potential nuclear markers were selected based on previous studies, and primers for the amplification of these markers were retrieved from the literature (see list in Table 1). Polymerase Chain Reaction (PCR) amplifications were optimized by performing gradient PCRs in search of the optimal annealing temperature. PCR reactions were conducted in 20 µL volumes with 2 µL of DNA, 10 µL of 1X Taq Polymerase (Extract-N-Amp™ PCR ReadyMix, Sigma-Aldrich, Saint Louis, USA), 0.5 µL of each primer (10 µM), and 7 µL of sterile water. The PCR amplification protocol was as follows: a denaturation step of 5 minutes at 94°C; 41 cycles of denaturation at 94°C for 1 minute, annealing at specific annealing temperature (Table 1) for 1.5 minutes, elongation at 72°C for 1 minute; a final elongation of 10 minutes at 72°C.

Reference samples of the parental species were PCR-amplified for each of the selected markers in Montpellier (BEV samples) or in Reims (Pagano and Patrelle collection) to identify potentially diagnostic SNPs (details in Table 2 and Supplementary File 1). Sequencing was done in both directions with the PCR primers; samples from the BEV collection or from the Pagano and Patrelle collection were sequenced by Eurofins Genomics (Ebersberg, Germany) or Genoscreen (Lille, France), respectively. Sequences were aligned using CodonCode Aligner V.4.2.3 sequence analysis software (CodonCode Corporation, Centerville, USA) or ClustalW (Thompson et al. 1994) included in MEGA V.5 (Tamura et al. 2011). For the CMYC2 marker, we retrieved sequences from GenBank for several individuals of *P. ridibundus*, *P. perezi* and *P. lessonae* (Table 2). Representative samples of newly generated sequences from each marker and parental species have been deposited in GenBank (see Table 2 and Supplementary File 1 for details).

Sequences of each marker were inspected visually across all species using the alignment viewer in MEGA V.5 (Tamura et al. 2011) to identify potentially diagnostic SNPs.

Table 2. Number of sequenced individuals for the four selected markers (or individuals retrieved from Genbank for CMYC2). Some individuals have been sequenced to check ambiguous restriction patterns and have not been uploaded on GenBank.

	RAG1	TYR1	POMC	CMYC2 (GenBank)
P. ridibundus	32	32	5	13
P. perezi	14	9	5	1
P. lessonae	6	6	6	8

Among the six successfully sequenced markers (Table 1), five provided SNPs with alternative alleles in at least one of the three parental species: Recombination-activation protein 1 (RAG1), Tyrosinase (TYR1), Pro-opiomelanocortin (POMC), Neurotrophin-3 (NTF3), and Cellular myelocytomatosis intron 2 (CMYC2).

In addition, when the initial RFLP tests generated ambiguous patterns or discordance between the RAG1 and TYR1 markers for some samples, we sequenced some samples for these markers to inspect the diagnostic SNPs directly from the sequences. In each case, sequencing revealed the expected base for the species and removed the discordance between markers. After confirmation to that effect had been found in these first sequencing trials, subsequent conflicting or ambiguous results were resolved by repeating the RFLP steps until results were compatible across markers or the conflict between markers were confirmed (see below). Most of these sequences were checked for their diagnostic positions, however, they were not corrected in their entirety, have not been deposited in GenBank, and the respective individuals are identified as "sequencing" in Supplementary File 1 instead of having GenBank accession numbers.

PCR-RFLP based method: digestion by restriction enzymes

Seeking to provide a quick and affordable method to identify Water Frog specimens, we developed restriction fragment length polymorphism protocols (RFLP) to genotype samples at the candidate diagnostic SNPs identified in the previous step. We used the RestrictionMapper software (available from www.restrictionmapper.org) and the CLC sequence Viewer 6 software (Giagen, Aarhus, Denmark) to find restricting endonucleases (restriction enzymes) whose recognition sequence include the candidate SNPs. When several enzymes were available for the same SNP, we selected the ones that were cheaper and easier to obtain. We retained the following enzymes for further tests: DraII (= EcoO109I) and BsmI (= Mva1269I) for digesting the RAG1 marker, BalI (= MscI) for the TYR1 marker, SspI for CMYC2, XapI (= Apo I) for NTF3, and HincII for POMC. All enzymes were ordered from the Thermo Scientific FastDigest Restriction Enzymes line (Thermo Fisher Scientific, Waltham, USA).

Table 3. Restriction patterns of the selected enzymes. For each marker (DNA fragment) and restriction enzyme, we provide the size of the fragment(s) resulting from the restriction digestion for each species. * The BsmI restriction site on RAG1 is not fully diagnostic if Italian samples of the LE system are included (see Supplementary File 2 and 3).

DNA fragment	TYR1 (646bp)	RAG1 (983bp)		POMC (476bp)	CMYC2 (298bp)
Restriction enzyme	BalI	DraII	BsmI*	HmcII	SspI
P. ridibundus	646	394/589	983	128/348	298
P. perezi	305/341	983	983	476	-
P. kl. *grafi*	646/305/341	983/394/589	983	476/128/348	-
P. lessonae	646	983	462/521	128/348	~ 150/150
P. kl. *esculentus*	646	983/394/589	983/462/521	128/348	298/~150

To test the reliability of the SNPs, amplicons of the five selected markers were digested with their corresponding restriction enzymes according to manufacturer's instructions. We report here the conditions that we used, but these might need to be changed with the source of the enzymes, as different enzyme concentrations might require different digestion times. All digestions were performed at 37°C, and digestion solutions of TYR1, RAG1 and CMYC2 were incubated for 6 hours, while POMC and NTF3 digestions were incubated for 10 minutes. Digested fragments were then separated in 1% agarose electrophoresis gel for 30 min at 100 V and 80 mA, and visualized with UV light. Fragment lengths were estimated by comparison with a 100-bp DNA Step Ladder (Promega).

As explained above, we always aimed at genotyping each specimen for at least two diagnostic markers. When the RFLP patterns were ambiguous or discordant between markers (each marker suggesting a different identification), we repeated the digestion, and the amplification when necessary, until we obtained two unambiguous and concordant results. Samples that failed to produce unambiguous RFLP patterns for at least two loci were discarded (almost invariably poorly preserved samples that failed to amplify reliably, see below). We only found two truly discordant cases where clear RFLP results were consistently repeated on the same markers twice and produced RFLP patterns suggesting different species (see details below).

Results and discussion

Enzymatic restriction patterns on the markers selected

In the results presented here, we report specimens whose identification was confirmed by previous allozyme genotyping, by concordance between RFLP genotypings, or sequencing of at least two diagnostic loci. Specimens that were genotyped for a single locus have been excluded and their identification has been treated as tentative only.

RAG1 with DraII

The fragment amplified from the RAG1 marker (amplicon: 983 bp, gene fragment of 936 bp after priming sequences are removed) contains several species-specific loci. Five diagnostic SNPs differentiate *P. ridibundus* from *P. perezi* (two of which are shared between *P. perezi* and *P. lessonae*) and one SNP allele is specific of *P. lessonae*. We selected the DraII enzyme, with a target site of RGzGNCCY, that cuts between positions 565 and 566 for *P. ridibundus* alleles only (alignment in Supplementary File 4), generating two fragments of 589 bp and 394 bp for *P. ridibundus* but a single uncut fragment (983 bp) for *P. perezi* and *P. lessonae* (Table 3, Fig 1A). The digestions for *P.* kl. *grafi* and *P.* kl. *esculentus* produce three fragments, viz. one uncut band of 983 bp corresponding to the complete *P. perezi* or *P. lessonae* allele, and two smaller fragments of 589 bp and 394 bp, resulting from cutting the *P. ridibundus* allele (Table 3, Fig. 1A, see Supplementary File 5 for *P. lessonae* and *P.* kl. *esculentus* patterns).

This RAG1/DraII restriction site was genotyped by RFLP or sequencing in a total of 903 specimens (Supplementary File 1): 524 *P. ridibundus* (including three *P.* cf. *ridibundus* / *P. bedriagae* from Azerbaijan), 86 *P. perezi*, 22 *P. lessonae*, 251 *P.* kl. *grafi*, and 20 *P.* kl. *esculentus*; all yielded RFLP patterns or genotypes for the restriction site as was expected for their taxon (Table 3).

TYR1 with BalI

The amplified TYR1 fragment is a 646-bp amplicon (generating a 601-bp fragment after removal of primer sequences) where the enzyme BalI (target site: TGGzCCA) cuts between positions 318 and 319 in *P. perezi*, generating two fragments of almost the same length (305 and 341 bp), while in *P. ridibundus* and *P. lessonae* this restriction site does not exist (alignment in Supplementary File 6). As a result, in the electrophoresis gel, a short fragment of around 320 bp is visible for *P. perezi,* while a full-length fragment of 646 bp is found for *P. ridibundus*, *P. lessonae* and *P.* kl. *esculentus* (Table 3, Fig 1B, see Supplementary File 5 for *P. lessonae* and *P.* kl. *esculentus* patterns). For *P.* kl. *grafi*, two bands of different length are visible, the shortest measuring around 320 bp, corresponding to the two fragments from the *P. perezi* allele, and the longest 646 bp, corresponding to the complete *P. ridibundus* allele (Table 3, Fig 1B). Like in *P. ridibundus*, *P. lessonae* does not have a restriction site for the BalI enzyme, therefore digestion of *P. lessonae* and *P.* kl. *esculentus* samples results in a single band of 646 bp

(Table 3, see Supplementary File 5 for *P. lessonae* and *P.* kl. *esculentus* patterns).

This TYR1/BalI restriction site was genotyped by RFLP or sequencing in a total of 886 specimens (Supplementary File 1): 523 *P. ridibundus* (including seven *P.* cf. *ridibundus* / *P. bedriagae* from Azerbaijan and Turkey), 82 *P. perezi*, 13 *P. lessonae*, 251 *P.* kl. *grafi* and 16 *P.* kl. *esculentus*; all yielded RFLP patterns or genotypes for the restriction site as was expected for their taxon (Table 3).

CMYC2 with SspI

We did not manage to amplify the CMYC2 fragment in the PG system with the primers used. In the LE system, the amplicon is 298 bp, and the 253 bp fragment (after primers were removed) exhibits a unique species-specific SNP for *P. lessonae*. The SspI enzyme, with a target site of AATzATT, only cuts the *P. lessonae* allele, between positions 133 and 134 (alignment in Supplementary File 7), producing a single 298-bp fragment for *P. ridibundus* and two undistinguishable fragments of approx. 150 bp for *P. lessonae* (Table 3, Fig 1C). In *P.* kl. *esculentus*, two bands of 298 bp and ± 150 bp are visible on gel (Table 3, Fig 1C). Even though we did not successfully amplify *P. perezi*, we retrieved the sequence of this species from GenBank, and as can be seen in the alignment in Supplementary File 7, *P. perezi* does not present the SspI restriction site. This implies that after digestion, the amplicons of *P. perezi* or *P.* kl. *grafi* would produce a pattern identical to *P. ridibundus* and that the CMYC2 patterns described here would be truly diagnostic for *P. lessonae* and *P.* kl. *esculentus*.

This CMYC2/SspI restriction site was genotyped by RFLP or sequencing in a total of 169 specimens (Supplementary File 1): 23 *P. ridibundus*, 78 *P. lessonae* and 67 *P.* kl. *esculentus*; all yielded RFLP patterns or genotypes for the restriction site as was expected for their taxon (Table 3).

POMC with HincII

The fragment amplified from POMC (amplicon: 476 bp, 415 bp of sequence after removal of primer) does not include fixed differences between *P. ridibundus* and *P. lessonae;* however, it contains four SNPs that are only present in *P. perezi*. The HincII enzyme, with target site GTYzRAC, digests the POMC fragment between positions 99 and 100 (alignment in Supplementary File 8) in *P. ridibundus* and *P. lessonae*, generating fragments of 128 bp and 348 bp for these species and a full-length fragment (476 bp) for *P. perezi* (Table 3, Fig 2). Among the hybridogenetic species, *P.* kl. *esculentus* has the same restriction profile as its par-

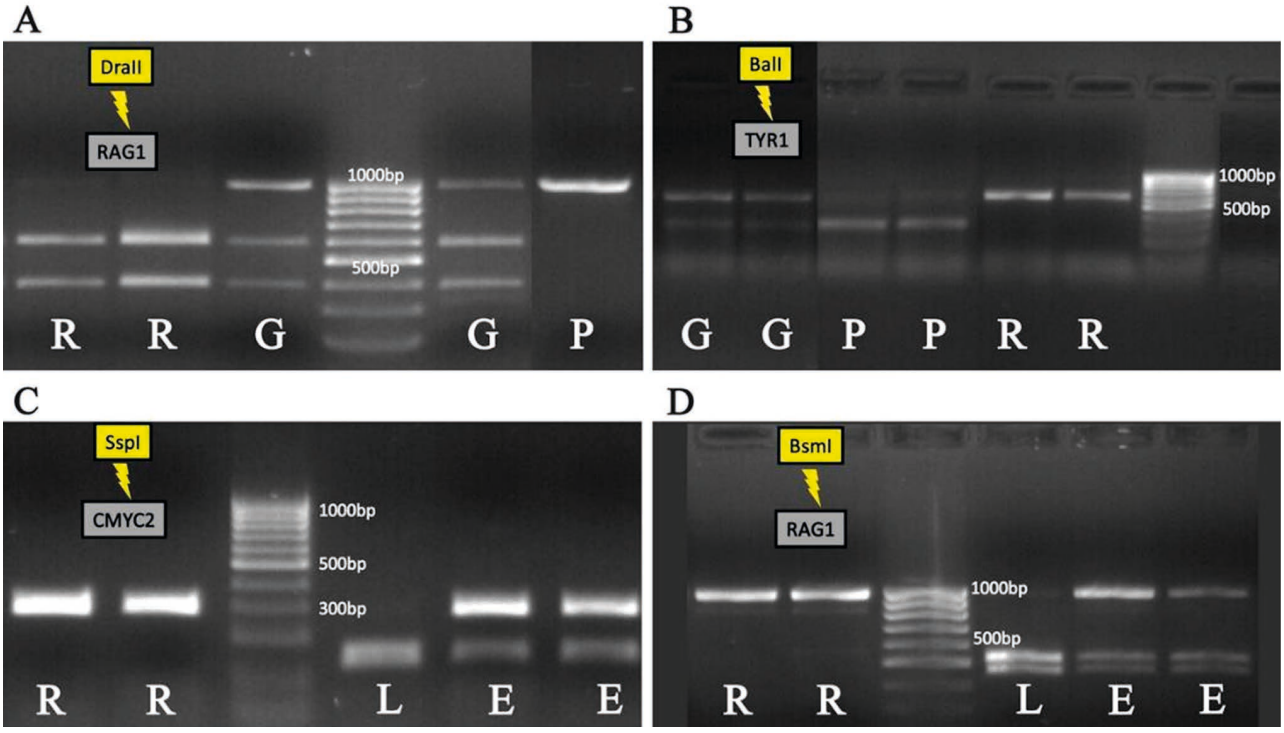

Figure 1. Enzymatic digestion patterns of the species diagnostic markers revealed by electrophoresis gel. Species names are labelled as follow. P: *Pelophylax perezi*; G: *P.* kl. *grafi*; R: *P. ridibundus*; L: *P. lessonae*; E: *P.* kl. *esculentus*. (A) Digestion of RAG1 by Dra II on species of the PG system (Digestion on the LE system is shown on Supplementary File 5). (B) Digestion of TYR1 by BalI on species of the PG system (Digestion on the LE system is shown on Supplementary File 5). (C) Digestion of CMYC2 by SspI on species of the LE system (Digestion on the PG system was not attempted). (D) Digestion of RAG1 by Bsm I on species of the LE system (Digestion on the PG system is shown on Supplementary File 5).

ents *P. ridibundus* and *P. lessonae* (two fragments of 128 and 348 bp in length, Fig 2). *Pelophylax* kl. *grafi* presents three fragments, one of 476 bp, corresponding to the complete *P. perezi* allele, and two smaller fragments of 128 and 348 bp, corresponding to the allele of *P. ridibundus* (Table 3, Fig 2).

This POMC/HincII restriction site was genotyped by RFLP or sequencing in a total of 83 specimens (Supplementary File 1): 18 *P. ridibundus* (including seven *P.* cf. *ridibundus* / *P. bedriagae* from Azerbaijan and Turkey), 19 *P. perezi*, 20 *P. lessonae*, 13 *P.* kl. *grafi* and 13 *P.* kl. *esculentus*; all yielded RFLP patterns or genotypes for the restriction site as was expected for their taxon (Table 3).

RAG1 with BsmI

Initial examinations of RAG1 sequences suggested an additional diagnostic position at the enzyme BsmI restriction site CTTACzGN where this enzyme digests the *P. lessonae* RAG1 allele between positions 437 and 438 (alignment in Supplementary File 4), generating two bands of similar length (462 and 521 bp) but usually visibly different (Table 3, Fig. 1D). For *P.* kl. *esculentus*, the digestion produces three fragments, one of 983 bp, corresponding to the complete *P. ridibundus* allele, and two bands of 462 and 521 bp from the *P. lessonae* allele (Table 3, Fig. 1D). For *P. ridibundus*, *P. perezi* and *P.* kl. *grafi*, the digestion results in a single band of 983 bp (Table 3, see Supplementary File 5 for *P. perezi* and *P.* kl. *grafi* patterns).

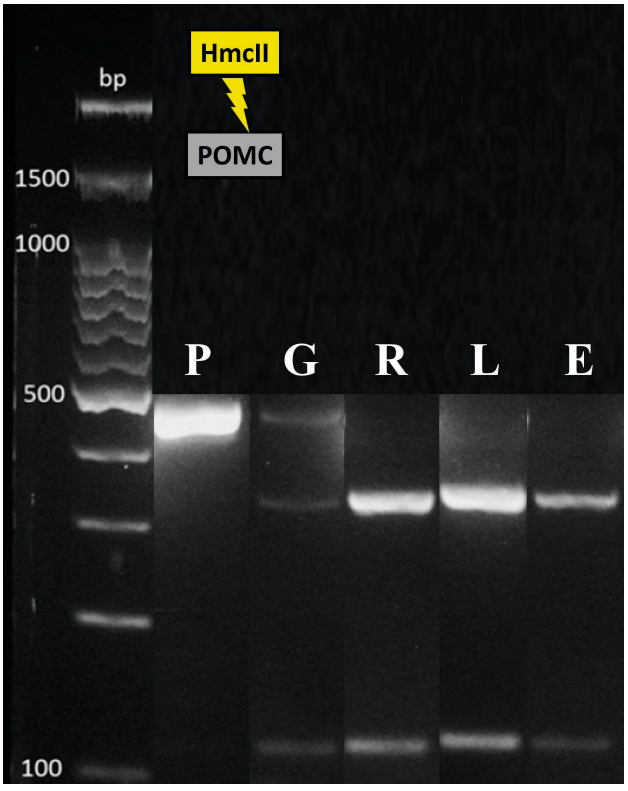

Figure 2. Enzymatic digestion patterns of POMC by HmcII. DNA fragment size is determined by comparison with a 100-bp DNA Step Ladder (Promega). Species names are labeled as follow. P: *Pelophylax perezi*; G: *P.* kl. *grafi*; R: *P. ridibundus*; L: *P. lessonae*; E: *P.* kl. *esculentus*.

As a consequence, a series of 146 specimens were genotyped by RFLP for this RAG1/BsmI restriction site (Supplementary File 1): 11 *P. ridibundus*, 1 *P. perezi*, 68 *P. lessonae*, 1 *P.* kl. *grafi* and 64 *P.* kl. *esculentus*; all yielded RFLP patterns or genotypes for the restriction site as was expected for their taxon (Table 3). However, when adding more sequence data, we realized that this restriction site is not diagnostic for *P. lessonae bergeri* from Italy, as two individuals out of five sequenced Italian *P. lessonae* carried the *P. ridibundus* allele at a heterozygous (BEV.11361) or homozygous (BEV.12379) state (GenBank numbers in Supplementary File 1). These individuals would thus exhibit digestion patterns that do not correspond with the pattern expected for *P. lessonae*. Since we aimed at providing a method that would reliably identify the five widespread European species over their distribution range, we discarded the RAG1/BsmI for our final identification method, but this probably remains a valid option for molecular diagnosis in most of Europe. Note however that Dufresnes et al. (2017a) reported an extensive occurrence of Italian alleles in most of the French populations of *P. lessonae* for two nuclear loci so that the occurrence of Italian RAG1 alleles in French populations of *P. lessonae*, even if we did not detect them, is probably not unexpected.

Repeatability of RFLP genotyping

A variable but sometimes significant proportion of samples failed to amplify or yielded weak bands that were hard to resolve on gel after digestion. This problem was mostly caused by poor PCR amplification due to low-quality DNA. Our samples originated from various sources (Supplementary File 1), including buccal swabs collected by non-professional naturalists and preserved either dried or in 95% ethanol, and old voucher specimens, generating samples of unequal quality. We especially noted that buccal swabs stored dry at room temperature were not as reliably amplified as those stored in ethanol. These issues are inherent to any PCR-based approach and not particular to our RFLP method.

In addition, a small number of samples (45 out of 1,111, i.e., less than 5% of the genotyped samples) were successfully amplified but yielded contradictory results between markers at a first attempt (genotyping results marked in yellow in Supplementary File 1). For most of these samples one or both markers were reamplified and/or the enzymatic digestion was repeated, resolving the contradictions in almost all cases. We suspect that most of these issues were caused by incomplete DNA digestion during the RFLP step, generating incorrect patterns. They were easily resolved by repeating the process (except for two samples, see below), however, this shows that the digestion step can sometimes generate flawed results that are impossible to detect if the identification is not confirmed by another marker.

Conflicting results between markers

Two samples were consistently discordant between two markers even after repeated genotyping (orange results in Supplementary File 1): one sample from a mixed PG system and *P. ridibundus* population (T9844 in Supplementary File 1) produced twice a *P.* kl. *grafi* pattern for DraII on RAG1 (371-, 565- and 936-bp bands) and a *P. ridibundus* pattern for BalI on TYR1 (601-bp band); one sample from a LE population (T12991 in Supplementary File 1) produced twice a *P.* kl. *esculentus* pattern for Bsm I on RAG1 (150- and 300-bp bands) and a *P. ridibundus* pattern for SspI on CMYC2 (300-bp band). The sample from the LE population presents *P.* kl. *esculentus* morphological features and it was identified as such at sampling, but since morphological separation of *P.* kl. *grafi* from *P. perezi* and *P. ridibundus* is difficult, we are unsure of the identification of the second individual. In these two cases the genetic identification was left unresolved, and we suspect that these contradictory results stem from rare recombination events in the genomes of the hybridogenetic taxa (Christiansen & Reyer 2009). These rare cases of contradiction between markers, either due to recombination in hybrid taxa or allele-sharing at very low frequency in the parental species, are another reason that led us to design a method where identifications are always supported by a minimum of two independent markers.

Re-examination of SAI-1 length polymorphism

Hauswaldt et al. (2012) suggested that *P. lessonae* differed from all other Western European Water Frogs by a ~530-bp deletion in SAI-1, allowing easy identification of *P. lessonae* and *P.* kl. *esculentus* by the simple migration of the PCR products on agarose gel. We sequenced the SAI-1 fragment in two *P. perezi* individuals (GenBank OL702782, the other individual yielding a short partial sequence only). Together with the two specimens sequenced by Hauswaldt et al. (2012, FN432377-78), these four sequences confirmed that *P. perezi* has a long SAI-1 allele, identical to *P. ridibundus* alleles after standard migration on agarose gels. As previously suggested, SAI-1 length polymorphism should thus be an easy way of distinguishing *P. lessonae* and *P.* kl. *esculentus* from the other Western European Water Frogs.

However, when examining all (n = 19) *P. lessonae* SAI-1 sequences stored in GenBank, we found that three individuals (JQ965512, MF094367-68) carry long SAI-1 haplotypes that cannot be separated from the alleles of the other species by migration on agarose gel. A quick examination of these sequences revealed that, while JQ965512 is identical to a widespread *P. ridibundus* allele and is possibly a mislabelled sample, the other two sequences carry mutations that are private and group with other *P. lessonae* alleles in a neighbour-joining tree (result not shown). There is thus little doubt that *P. lessonae* can carry long SAI-1 alleles, casting doubt on the length polymorphism of this marker as a safe species-diagnostic tool in European water frogs. As a consequence, we have not retained this marker for the identification method that we propose here.

A novel PCR-RFLP protocol for the genetic identification of Western European water frogs

We here propose a novel method to identify Western European Water Frog taxa by PCR-RFLP based on the restriction patterns of selected markers. As we have detailed above, 1) the digestion step sometimes produces misleading RFLP patterns, and 2) there are very rare cases of discordance between identification suggested by different markers, therefore we developed a protocol that ensures that every identification result will be supported by a combination of two independent markers in all possible situations.

Our method relies on three different combinations of markers, depending on the geographical situation. When only three species are present (*P. ridibundus* and species from either the LE system or the PG system), a combination of PCR-RFLP tests on two markers will be enough to obtain a result. When all five species are potentially present, three markers are needed to ensure that all pairwise diagnoses are supported by two markers. Our protocol is detailed below and illustrated in Figure 3:

1) in Spain and large areas of southern France, the only candidate species are *P. ridibundus*, *P. perezi* and *P.* kl. *grafi*, and a combination of RAG1/DraII & TYR1/BalI facilitates reliable identification (Fig. 3A);

2) in most of Europe, separation of *P. ridibundus*, *P. lessonae* and *P.* kl. *esculentus* can be achieved by a combination of RAG1/DraII with CMYC2/SspI (Fig. 3B);

3) in the rare situations where all five species are potentially present (especially in western France), a first step with a combination of POMC/HmcII with RAG1/DraII provides patterns that are unique to each species; a second step with TYR1/BalI for the PG system or CMYC2/SspI for the LE system allows to verify these identification results (Fig. 3C).

With this protocol, all three pairwise identification results between parental species are based on at least two markers: *P. perezi* and *P. lessonae* are separated by POMC and TYR1, *P. lessonae* and *P. ridibundus* by RAG1 and CMYC2, and *P. perezi* and *P. ridibundus* by RAG1, POMC and TYR1. We also made sure that no final identification result would be based on a combination of non-digested RFLP patterns to prevent misidentification due to complete digestion failure: when a species-specific RFLP pattern is "one band" (no digestion) for a given marker we always combine it with a marker that "cuts" for the same species.

Applicability to other, exotic Water Frog taxa

Although our protocol has been developed mainly on the basis of French samples, it has also been tested on a few samples from the LE system in Italy, where several parental and hybridogenetic taxa are currently treated as conspe-

cific with *P. lessonae* and *P.* kl. *esculentus* (Speybroeck et al. 2020). When and where possible, we also checked on GenBank sequences that the substitutions used to design our tests were also present in Italian samples. Although increasing the sample size of Italian specimens would be ideal, we believe that our method is also applicable to these populations. Within the *P. ridibundus* complex, our data include specimens sampled in French populations carrying either the Central European mtDNA lineage or the Balkans (*P. ridibundus kurtmuelleri*) mtDNA lineage (based on data in Dufresnes et al. 2017b and on our own data). As expected from their low divergence in nuclear DNA, these two lineages conform to the *P. ridibundus* patterns described above. We have not been able to assess the performance of our method in the *P. ridibundus-bedriagae* complex from Anatolia and the Middle East, but the samples that we have sequenced (Supplementary File 1, three from Azerbaijan and five from Turkey, belonging to several mtDNA lineages) share RAG1 and TYR1 restriction sites with *P. ridibundus* from Europe, as has been assessed from our sequence alignments.

Lastly, we wish to emphasize that our PCR-RFLP method does not provide a "positive" identification result for any of the species: it distinguishes between these five species when the samples actually belong to one of these species. Samples of other species of the genus *Pelophylax* such as *P. epeiroticus*, *P. sqhipericus*, *P. cretensis* or *P. saharicus* will most likely amplify for the markers we used and could exhibit RFLP patterns identical to one of the five Western European species. For example, samples of the recently discovered introduced French population of *P. saharicus* (Doniol-Valcroze et al. 2021) were assayed for the RAG1/

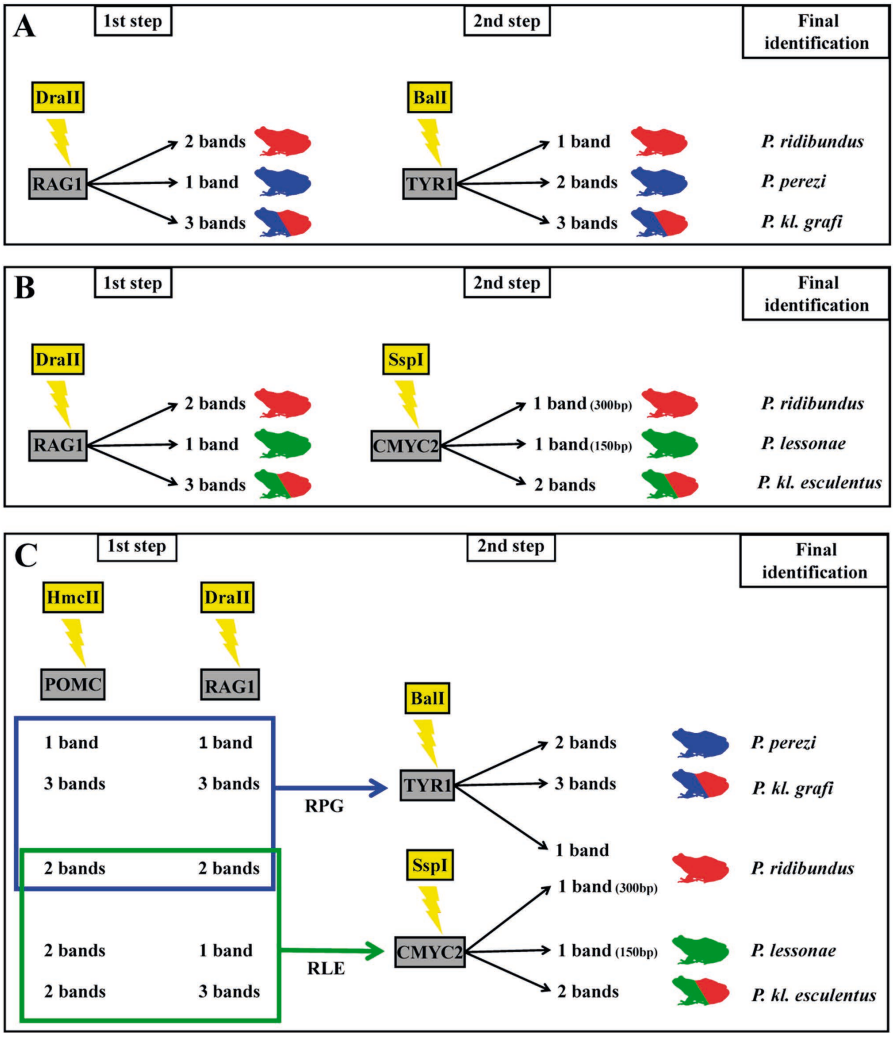

Figure 3. A novel PCR-RFLP protocol for the identification of Western European water frogs. (A) In the PG system, a combination of RAG 1/DraII and TYR1/BalI ensures that all identifications are supported by two independent loci. (B) In the LE system, a combination of RAG I/DraII and CMYC2/SspI has the same property. (C) When all five species are potentially present, the first step requires double digestion of POMC by HmcII and of RAG 1 by DraII, which in combination provides a species-specific pattern and reveals from which system (LE or PG) the samples are originating. This identification can be confirmed by a second step using TYR1/BalI, if the first step suggests the PG system, or a CMYC2/SspI step, if the first step suggest LE.

DraII and TYR1/BalI combinations before their true identity was revealed and they exhibited *P. perezi* RFLP patterns for both markers (Supplementary File 9). Whenever our protocol returns an unexpected identification, either a species outside its normal distribution range (e.g., PG individuals in Northern or Eastern Europe or LE individuals in Iberia) or an identification that does not match the morphology of the specimen, we urge that such identification be tested with other methods (e.g., DNA barcoding) to rule out the possibility of an introduced population of a species not included in our protocol.

Conclusion

Given the difficulties to identify European Water Frog species with morphological or acoustic characters and the conservation concerns raised by the current spread of *Pelophylax ridibundus* at the expense of the native taxa of the PG and LE systems, DNA-based methods have become crucial for monitoring their status. The PCR-RFLP method we propose here generates restriction banding patterns specific for *P. ridibundus, P. perezi*, *P. lessonae* and the two hybrid species *P.* kl. *grafi* and *P.* kl. *esculentus*, facilitating the reliable identification of Water Frogs in most parts of Europe. In comparison to other molecular methods, our PCR-RFLP protocol is fast, relatively inexpensive, easily replicable, and allows non-invasive sampling. Because it systematically relies on at least two independent nuclear markers to reach an identification result, we also believe it is less error-prone than other PCR-based methods currently available.

Acknowledgements

This study was supported by a REPERE grant for the MEDINA project (Conservation de la biodiversité en région Méditerranéenne: une interface recherche – gestion pour identifier les enjeux et alimenter les processus de decision) awarded to the Conservatoire des espaces naturels Languedoc-Roussillon (CEN-LR) and the UMR5175 CEFE and coordinated by John Thompson (CNRS, CEFE). Financial support for this study was also provided by the team EA 4688 "VECPAR", University of Reims Champagne-Ardenne. The authors would like to thank Alain Pagano who provided some specimens, Alban Gerard, Marc Mangeat and Thomas Sire for helping with sample collection and all the students and naturalists named in Supplementary File 1 who collected samples. Cassandre Treyvaud, Blaise Raymond and Adrien Pineau helped genotyping some of the samples used in this study. Part of the samples used in this study were also collected under projects funded by Association Bufo (Strasbourg), URCPIE Midi Pyrénées (L'Isle de Noé) and Nature en Occitanie. Lastly, we are grateful to all the pond owners who allowed us to collect samples on their properties.

References

Arano, B., G. Llorente, M. Garcia-Paris & P. Herrero (1995): Species translocation menaces Iberian waterfrogs. – Conservation Biology, **9**: 196–198.

Beerli, P., H. Hotz & T. Uzzell (1996): Geographical dated sea barriers calibrate a protein clock for Aegean water frogs. – Evolution, **50**: 1676–1687.

Berger, L. (1973): Some characteristics of backcrosses within forms of *Rana esculenta* complex. – Genetica Polonica, **14**: 413–440.

Bruford, M. W., O. Hanotte, J. F. Y. Brookfield & T. Burke (1992): Single-locus and multilocus DNA fingerprinting. – pp. 225–269 in: Hoelzel, A. R. (ed.): Molecular genetic analysis of populations: a practical approach. – RL Press, Oxford.

Buckley, D., B. Arano, P. Herrero, G. A. Llorente & M. Esteban (1994): Moroccan water frogs vs. *R. perezi*: allozyme studies show up their differences. – Zoologica Poloniae, **39**: 377–386.

Christiansen, D. G. (2005): A microsatellite-based method for genotyping diploid and triploid water frogs of the *Rana esculenta* hybrid complex. – Molecular Ecology Notes, **5**: 190–193.

Christiansen, D. G. & H.-U. Reyer (2009): From clonal to sexual hybrids: genetic recombination via triploids in all-hybrid populations of water frogs. – Evolution, **63**: 1754–1768.

Congiu, L., I. Dupanloup, T. Patarnello, F. Fontana, R. Rossi, G. Arlati & L. Zane (2001): Identification of interspecific hybrids by amplified fragment length polymorphism: the case of sturgeon. – Molecular Ecology, **10**: 2355–2359.

Crochet, P.-A., A. Dubois, A. Ohler & H. Tunner (1995): *Rana (Pelophylax) ridibunda* Pallas, 1771, *Rana (Pelophylax) perezi* Seoane 1885 and their associated klepton (Amphibia, Anura): morphological diagnoses and description of a new taxon. – Bulletin du Muséum national d'Histoire naturelle de Paris, **17**: 11–30.

Daf, O. S., A. Pagano & T. Lodé (2006): Taxonomic diversity and sympatry among water frogs from Southern France: evidence for new assemblages. – Amphibia-Reptilia, **27**: 295–299.

Doniol-Valcroze, P., G. Mazepa, F. Grimal, P. Sourrouille, N. Perrin, S. N. Litvinchuk & P.-A. Crochet (2021): Discovery of a *Pelophylax saharicus* (Anura, Ranidae) population in Southern France: a new potentially invasive species of water frogs in Europe. – Amphibia-Reptilia, **42**: 427–442.

Dotti do Prado, F., D. T. Hashimoto, F. F. Mendonça, J. A. Senhorini, F. Foresti & F. Porto-Foresti (2011): Molecular identification of hybrids between Neotropical catfish species *Pseudoplatystoma corruscans* and *Pseudoplatystoma reticulatum*. – Aquaculture Research, **42**: 1890–1894.

Dubey, S. & C. Dufresnes (2017): An extinct vertebrate preserved by its living hybridogenetic descendant. – Scientific Reports, **7**: 1–6.

Dufresnes, C., L. di Santo, J. Leuenberger, J. Schuerch, G. Mazepa, N. Grandjean, D. Canestrelli, N. Perrin & S. Dubey (2017a): Cryptic invasion of Italian pool frogs (*Pelophylax bergeri*) across Western Europe unraveled by multilocus phylogeography. – Biological Invasions, **19**: 1407–1420.

Dufresnes, C., M. Denoël, L. di Santo & S. Dubey (2017b): Multiple uprising invasions of *Pelophylax* water frogs, potentially inducing a new hybridogenetic complex. – Scientific Reports, **7**: 1–10.

Frost, D. R., T. Grant, J. Faivovich, R. H. Bain, A. Haas, C. F. B. Haddad, R. O. De Sa, A. Channing, M. Wilkinson, S. C. Donnellan, C. Raxworthy, J. Christopher, J. A. Campbell, B. L. Blotto, P. Moler, R. C Drewes, R. A. Nussbaum, J. D. Lynch, D. M. Green & W. C Wheeler (2006): The amphibian tree of life. – Bulletin of the American Museum of Natural History, **297**: 1–370.

Garner, T. W. J., B. Gautschi, S. Rothlisberger & H. U. Reyer (2000): A set of CA repeat microsatellite markers derived from the pool frog, *Rana lessonae*. – Molecular Ecology, **9**: 2173–2175.

Gasc, J.-P., A. Cabela, J. Crnobrnja-Isailovic, D. Dolmen, K. Grossenbacher, P. Haffner, J. Lescure, H. Martens, J. P. Martínez Rica, H. Maurin, M. E. Oliveira, T. S. Sofianidou, M. Veith & A. Zuiderwijk (2004): Atlas of Amphibians and Reptiles in Europe, Second edition. – Paris, Societas Europaea Herpetologica & Muséum national d'Histoire naturelle.

Gottschling, M., R. Knop, J. Plötner, M. Kirsh, H. Willems & H. Keupp (2005): A molecular phylogeny of *Scrippsiella* sensu lato (Calciodinellaceae, Dinophyta) with interpretations on morphology and distribution. – European Journal of Phycology, **40**: 207–220.

Graf, J.-D., F. Karch & M. C. Moreillon (1977): Biochemical variation in the *Rana esculenta* complex: a new hybrid form related to *Rana perezi* and *Rana ridibunda*. – Experientia, **33**: 1582–1584.

Graf, J.-D. & W. P. Mueller (1979): Experimental gynogenesis provides evidence of hybridogenetic reproduction in the *Rana esculenta* complex. – Experientia, **35**: 1574–1576.

Graf, J. D. & M. Polls Pelaz (1989): Evolutionary genetics of *Rana esculenta* complex. – pp. 289–302 in: Dawley R. M. & J. P. Bogart (eds): Evolution and Ecology of Unisexual Vertebrates. – Springer, New York.

Günther, R. (1990): Die Wasserfrösche Europas (Anura – Froschlurche). – Ziemsen, Wittenberg-Lutherstadt.

Günther, R. & J. Plötner (1994): Morphometric, enzymological and bioacoustic studies in Italian water frogs (Amphibia, Ranidae). – Zoologica Poloniae, **39**: 387–415.

Hauswaldt, J. S., M. Höer, M. Ogielska, D. G. Christiansen, D. Dziewulska-Szwajkowska, E. Czernicka & M. Vences (2012): A simple molecular method for distinguishing among species and ploidy levels in European water frogs (*Pelophylax*). – Molecular Ecology Resources, **12**: 797–805.

Heppich, S. (1978). Hybridogenesis in *Rana esculenta*: C-band karyotypes of *Rana ridibunda*, *Rana lessonae* and *Rana esculenta*. – Journal of Zoological Systematics and Evolutionary Research, **16**: 27–39.

Hoffmann, A. & H. U. Reyer (2013): Genomic effects on advertisement call structure in diploid and triploid hybrid waterfrogs (Anura, *Pelophylax esculentus*). – BMC Ecology, **13**: 1–17.

Holsbeek, G., J. Mergeay, H. Hotz, J. Plötner, F. A. M. Volckaert & L. De Meester (2008): A cryptic invasion within an invasion and widespread introgression in the European water frog complex: consequences of uncontrolled commercial trade and weak international legislation. – Molecular Ecology, **17**: 5023–5035.

Holsbeek, G. & R. Jooris (2010): Potential impact of genome exclusion by alien species in the hybridogenetic water frogs (*Pelophylax esculentus* complex). – Biological Invasions, **12**: 1–13.

Holsbeek, G., J. Mergeay, F. A. M. Volckaert & L. De Meester (2010): Genetic detection of multiple exotic water frog species in Belgium illustrates the need for monitoring and immediate action. – Biological Invasions, **12**: 1459–1463.

Hotz, H. & T. Uzzell (1982): Biochemically detected sympatry of two water frog species: two different cases in the Adriatic Balkans (Amphibians, Ranidae). – Proceedings of the Academy of Natural Sciences of Philadelphia, **134**: 50–79.

Hotz, H., P. Beerli & C. Spolsky (1992): Mitochondrial DNA reveals formation of nonhybrid frogs by natural matings between hemiclonal hybrids. – Molecular Biology and Evolution, **9**: 610–620.

Hotz, H., T. Uzzell & L. Berger (1994): Hemiclonal hybrid water frogs associated with the sexual host species *Rana perezi*. – Zoologica Poloniae, **39**: 243–266.

Hotz, H. & R. D. Semlitsch. (2000): Differential performance among LDH-B genotypes in *Rana lessonae* tadpoles. – Evolution, **54**: 1750–1759.

Hotz, H., T. Uzzell, G.-D. Guex, D. Alpers & R. D. Semlitsch (2001): Microsatellites: a tool for evolutionary genetic studies of western Palearctic water frogs. – Zoosystematics and evolution, **77**: 43–50.

Joseph, N., E. Krauskopf, M. I. Vera & B. Michot (1999): Ribosomal internal transcribed spacer 2 (ITS2) exhibits a common core of secondary structure in vertebrates and yeast. – Nucleic Acids Research, **27**: 4533–4540.

Lescure, J. & J. C. de Massary (eds) (2012): Atlas des amphibiens et reptiles de France. – Biotope Editions and Publications scientifiques du Muséum, Mèze and Paris.

Leuenberger, J., A. Gander, B. R. Schmidt & N. Perrin (2014): Are invasive marsh frogs (*Pelophylax ridibundus*) replacing the native *P. lessonae*/*P. esculentus* hybridogenetic complex in Western Europe? Genetic evidence from a field study. – Conservation Genetics, **15**: 869–878.

Lodé, T. & A. Pagano (2000): Variations in call and morphology in male water frogs: taxonomic and evolutionary implications. – Comptes Rendus de l'Académie des sciences de Paris, Sciences de la vie/Life Sciences, **323**: 995–1001.

Lymberakis, P., N. Poulakakis, G. Manthalou, C. S. Tsigenopoulos, M. Magoulas & M. Mylonas (2007): Mitochondrial phylogeography of *Rana* (*Pelophylax*) populations in the Eastern Mediterranean region. – Molecular Phylogenetics and Evolution, **44**: 115–125.

Malécot, V., N. Macquaire-Le Pocreau, G. Auvray, V. Kapusta (2009): Polymorphic ITS as a tool to identify hybrids and their parents in cultivated *Genisteae* (Fabaceae). – Acta Horticulturae, **836**: 91–96.

Mayer, M., O. Hawlitschek, A. Zahn & F. Glaw (2013): Composition of twenty green frog populations (*Pelophylax*) across Bavaria, Germany. – Salamandra, **49**: 31–44.

Newman, C. E., J. A. Feinberg, L. J. Rissler, J. Burger, H. B. Shaffer (2012): A new species of leopard frog (Anura: Ranidae) from the urban northeastern US. – Molecular Phylogenetics and Evolution, **63**: 445–455.

Nikoloudakis, N., G. Skaracis & A. Katsiotis (2008): Evolutionary insights inferred by molecular analysis of the ITS1-5.8S-ITS2 and IGS *Avena* sp. sequences. – Molecular Phylogenetics and Evolution, **46**: 102–115.

Pagano, A., P. Joly & H. Hotz (1997): Taxon composition and genetic variation of water frogs in the Mid-Rhône floodplain. – Comptes Rendus de l'Académie des Sciences de Paris, Sciences de la vie/Life Sciences, **320**: 759–766.

Pagano, A. & P. Joly (1999): Limits of the morphometric method for field identification of water frogs. – Alytes, **16**: 130–138.

Pagano, A., T. Lodé & P.-A. Crochet (2001a): New contact zone and assemblages among water frogs of Southern France. – Journal of Zoological Systematics and Evolutionary Research, **39**: 63–67.

Pagano, A., P.-A. Crochet, J.-D. Graf, P. Joly & T. Lodé (2001b): Distribution and habitat use of water frog hybrid complexes in France. – Global Ecology and Biogeography, **10**: 433–441.

Pagano, A., A. Dubois, D. Lesbarrères & T. Lodé (2003): Frog alien species: a way for genetic invasion? – Comptes Rendus Biologies, **326**: 85–92.

Patrelle, C. (2010): Les complexes d'hybridation chez les grenouilles vertes: identification taxonomique, exigences écologiques, et capacité d'acclimatation. – Unpubl. thesis, University of Angers.

Patrelle, C., T. Host, D. Picard, A. Pagano, S. Sourice, M.-G. Dallay & J. Plötner (2011): A new PCR-RFLP-based method for an easier systematic affiliation of European water frogs. – Molecular Ecology Resources, **11**: 200–205.

Plötner, J. (2005): Die westpaläarktischen Wasserfrösche. Von Märtyrern der Wissenschaft zur biologischen Sensation. – Laurenti, Bielefeld.

Plötner, J., T. Uzzell, P. Beerli, C. Spolsky, T. Ohst, S. N. Litvinchuk, G. D. Guex, H. U. Reyer & H. Hotz (2008): Widespread unidirectional transfer of mitochondrial DNA: a case in western Palaearctic water frogs. – Journal of Evolutionary Biology, **21**: 668–681.

Plötner, J. (2010): Möglichkeiten und Grenzen morphologischer Methoden zur Artbestimmung bei europäischen Wasserfröschen (*Pelophylax esculentus*-Komplex). – Zeitschrift für Feldherpetologie, **17**: 129–146.

Sagonas, K., E. Karameta, P. Kotsakiozi & N. Poulakakis (2020): Cross-species testing of nuclear markers in *Pelophylax* water frogs in Greece and examination of their power to detect genetic admixture. – Amphibia-Reptilia, **41**: 253–259.

Sánchez-Montes, G., E. Recuero, J. Gutiérrez-Rodríguez, I. Gomez-Mestre & I. Martínez-Solano (2016): Species assignment in the *Pelophylax ridibundus* x *P. perezi* hybridogenetic complex based on 16 newly characterised microsatellite markers. – The Herpetological Journal, **26**: 99–108.

Schmeller, D. S., A. Pagano, S. Plénet & M. Veith (2007): Introducing water frogs – is there a risk for indigenous species in France? – Comptes Rendus Biologies, **330**: 684–690.

Spolsky, C. & T. Uzzell (1984): Natural interspecies transfer of mitochondrial DNA in amphibians. – Proceedings of the National Academy of Sciences of the USA, **81**: 5802–5805.

Spolsky, C. M. & T. Uzzell (1986): Evolutionary history of the hybridogenetic hybrid frog *Rana esculenta* as deduced from mtDNA analyses. – Molecular Biology and Evolution, **3**: 44–56.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei & S. Kumar (2011): MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. – Molecular Biology and Evolution, **28**: 2731–2739.

Thompson, J. D., D. G. Higgins & T. J. Gibson (1994): CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. – Nucleic Acids Research, **22**: 4673–4680.

Tognarelli, G., M. A. Zuffi, S. Marracci & M. Ragghianti (2014): Surveys on populations of green frogs (*Pelophylax*) of Western Tuscany sites with molecular and morphometric methods. – Amphibia-Reptilia, **35**: 99–105.

Uzzell, T., H. Hotz & L. Berger (1980): Genome exclusion in gametogenesis by an interspecific *Rana* hybrid: evidence from electrophoresis of individual oocytes. – Journal of Experimental Zoology, **214**: 251–259.

Uzzell, T. & H. G. Tunner (1983): An immunological analysis of Spanish and French water frogs. – Journal of Herpetology, **17**: 320–326.

Uzzell, T. & L. Berger (1975): Electrophoretic phenotypes of *Rana ridibunda*, *Rana lessonae*, and their hybridogenetic associate, *Rana esculenta*. – Proceedings of the Academy of Natural Sciences of Philadelphia, **127**: 13–24.

Uzzell, T. & H. Hotz (1979): Electrophoretic and morphological evidence for two forms of green frogs (*Rana esculenta* complex) in peninsular Italy (Amphibia, Salientia). – Mitteilungen aus dem Museum für Naturkunde in Berlin, **55**: 13–27.

Van Doninck, K., I. Schön, L. De Bruyn & K. Martens (2002): A general-purpose genotype in an ancient asexual. – Oecologia, **132**: 205–212.

Vinogradov, A. E. & A. T. Chubinishvili (1999): Genome reduction in a hemiclonal frog *Rana esculenta* from radioactively contaminated areas. – Genetics, **151**: 1123–1125.

Vorburger, C. & H. U. Reyer (2003): A genetic mechanism of species replacement in European waterfrogs? – Conservation Genetics, **4**: 141–155.

Vucić, M., D. Jelić, G. I. Klobučar, B. Prkljačić & M. Jelić (2018): Molecular identification of species and hybrids of water frogs (genus *Pelophylax*) from Lake Skadar, Southeast Adriatic drainages (Amphibia: Ranidae). – Salamandra, **54**: 147–157.

Zeisset, I. & T. J. C. Beebee (1998): RAPD identification of north European water frogs. – Amphibia-Reptilia, **19**: 163–170.

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

Zeisset, I. & T. J. C. Beebee (2003): Population genetics of a successful invader: the marsh frog *Rana ridibunda* in Britain. – Molecular Ecology, **12**: 639–646.

Supplementary data

The following data are available online:

Supplementary File 1. List of all *Pelophylax* specimens with definite identification results.

Supplementary File 2. Digestion protocol and pattern of NTF3 marker using the XapI restriction enzyme.

Supplementary File 3. NTF3 nuclear marker alignment with an indication of the XapI restriction site.

Supplementary File 4. RAG1 nuclear marker alignment with indications of the BsmI and DraII restriction sites.

Supplementary File 5. Digestion patterns of RAG1 using DraII for *Pelophylax lessonae* and *P.* kl. *esculentus*, TYR1 using BalI for *P. lessonae* and *P.* kl. *esculentus*, and RAG1 using BsmI for *P. perezi* and *P.* kl. *grafi*.

Supplementary File 6. TYR1 nuclear marker alignment with an indication of the BalI restriction site.

Supplementary File 7. CMYC2 nuclear marker alignment with an indication of the SspI restriction site.

Supplementary File 8. POMC nuclear marker alignment with an indication of the HincII restriction site.

Supplementary File 9. Digestion patterns of RAG1 using DraII, and TYR1 using BalI, for *P. saharicus* samples originating from southern France.