Molecular Ecology (2015) 24, 4371-4391

Genetic diversity and distribution patterns of diploid and polyploid hybrid water frog populations (*Pelophylax esculentus* complex) across Europe

ALEXANDRA HOFFMANN,* JÖRG PLÖTNER,† NICOLAS B. M. PRUVOST,* DITTE G. CHRISTIANSEN,* SANDRA RÖTHLISBERGER,* LUKÁŠ CHOLEVA,‡§ PETER MIKULÍČEK,¶ DAN COGĂLNICEANU,** ISTVÁN SAS-KOVÁCS,†† DMITRY SHABANOV,‡‡ SVYATOSLAV MOROZOV-LEONOV§§ and HEINZ-ULRICH REYER*

*Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, Zurich CH-8057, Switzerland, †Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Invalidenstraße 43, Berlin 10115, Germany, ‡Laboratory of Fish Genetics, Institute of Animal Physiology and Genetics, AS CR, v.v.i., Rumburská 89, Liběchov 277 21, Czech Republic, §Department of Biology and Ecology, Faculty of Science, University of Ostrava, Chittussiho 10, Ostrava 710 00, Czech Republic, ¶Department of Zoology, Faculty of Natural Sciences, Comenius University in Bratislava, Mlynska dolina B-1, Bratislava SK-84215, Slovak Republic, **Faculty of Natural Sciences and Agricultural Sciences, University Ovidius Constanţa, Al. Universităţii 1, corp B, Constanţa, Romania, ††Department of Biology, Faculty of Science, University of Oradea, Universitatii str. 1, Oradea 410087, Romania, ‡‡Karazin Kharkiv National University, Svobody sq 4, Kharkiv 62077, Ukraine, §§Department of Evolutionary Genetic Basics of Systematics, Schmalhausen Institute of Zoology, B. Khmelnitskogo st., 15 Kyiv-30, Kyiv MSP UA-01601, Ukraine

Abstract

Polyploidization is a rare yet sometimes successful way for animals to rapidly create geno- and phenotypes that may colonize new habitats and quickly adapt to environmental changes. In this study, we use water frogs of the Pelophylax esculentus complex, comprising two species (Pelophylax lessonae, genotype LL; Pelophylax ridibundus, RR) and various diploid (LR) and triploid (LLR, LRR) hybrid forms, summarized as P. esculentus, as a model for studying recent hybridization and polyploidization in the context of speciation. Specifically, we compared the geographic distribution and genetic diversity of diploid and triploid hybrids across Europe to understand their origin, maintenance and potential role in hybrid speciation. We found that different hybrid and parental genotypes are not evenly distributed across Europe. Rather, their genetic diversity is structured by latitude and longitude and the presence/absence of parental species but not of triploids. Highest genetic diversity was observed in central and eastern Europe, the lowest in the northwestern parts of Europe. This gradient can be explained by the decrease in genetic diversity during postglacial expansion from southeastern glacial refuge areas. Genealogical relationships calculated on the basis of microsatellite data clearly indicate that hybrids are of multiple origin and include a huge variety of parental genomes. Water frogs in mixed-ploidy populations without any parental species (i.e. all-hybrid populations) can be viewed as evolutionary units that may be on their way towards hybrid speciation. Maintenance of such all-hybrid populations requires a continuous exchange of genomes between diploids and triploids, but scenarios for alternative evolutionary trajectories are discussed.

Keywords: all-hybrid populations, founder effects, geographic distribution, hybrid speciation, microsatellites, mtDNA, *Pelophylax*

Received 12 November 2014; revision received 9 July 2015; accepted 10 July 2015

Correspondence: Heinz-Ulrich Reyer, Fax: +41 44 635 57 11; E-mail: uli.reyer@ieu.uzh.ch

Introduction

Genetic diversity, caused by mutational changes and recombination of genes, is an essential precondition for the evolution and the origin of species. In concert with selection, it enables populations to adapt to ever changing environments. Hybridization between genetically distinct forms is one of the several mechanisms that provide genomic variability; but when heterospecific genomes are combined, genetic incompatibilities may result in a variety of genomic disorders expressed as embryonic mortality, low fertility or even hybrid sterility (reviewed by Coyne & Orr 2004). Interspecific hybridization is thus often maladaptive. Therefore, many authors see hybrids as 'evolutionary dead ends', whereas others consider that hybridization contributes to adaptive divergence and to the origin of new hybrid species (reviewed by Seehausen 2004; Abbott et al. 2013). There are an increasing number of examples that interspecific hybridization promotes speciation and species divergence. Most come from plants (reviewed by Otto & Whitton 2000; Soltis & Soltis 2009), but several also from animals (e.g. Arnold 1997; Dowling & Secor 1997; Mallet 2007, 2008). Key to this success is the hybrids' ability to circumvent meiotic disturbances during gametogenesis of heterospecific (non-coadapted) genomes. Some hybrid taxa achieve that by clonal reproduction, for example parthenogenetic, gynogenetic and hybridogenetic organisms (Dawley 1989), whereas others produce mainly diploid gametes, leading to allopolyploidy, that is offspring carrying at least three haploid sets of chromosomes from two or more parental species (Arnold 1997).

With the increasing application of molecular tools, evidence is accumulating that both clonal reproduction and allopolyploidy are more frequent than previously believed (Dawley 1989; Gregory & Mable 2005; Kearney et al. 2009; Lamatsch & Stöck 2009; Mable et al. 2011). But to what extent do the polyploids contribute to speciation? While it is well established that clonality can lead to speciation (Birky & Barraclough 2009; Abbott et al. 2013), the importance of allopolyploidy is still controversially disputed (Mason & Pires 2015). Theoretically, tetraploidization can instantly restore normal meiosis, as well as reproductive isolation (Coyne & Orr 2004). In plants, allopolyploidization results in higher reproductive isolation from diploid progenitors than homoploid hybridization (Rieseberg & Willis 2007). Yet, polyploidization seems to represent a rare path to speciation when compared to the more common genetic mechanisms such as mutational changes, recombination and genetic drift (reviewed by Choleva & Janko 2013). Moreover, theoretical models and experimental results suggest that induction of polyploidy is not required for instantaneous reproductive isolation and hybrid speciation (Seehausen 2004). Against this background of discrepancies, which may in part be due to differences among taxa, additional comparative data from various groups of organisms are needed to evaluate the role of allopolyploid populations in hybrid speciation.

The Pelophylax esculentus complex

European water frogs (genus Pelophylax) represent a suitable model for studying speciation in the context of hybridization and polyploidization. They comprise at least seven evolutionary species and several hybrid forms, among them are three hybridogenetic taxa (reviewed by Plötner 2005). In this study, we focus on a complex of water frog taxa, which consists of two species, Pelophylax ridibundus (Pallas, 1771) (the marsh or lake frog, genotype RR) and Pelophylax lessonae (Camerano, 1882) (the pool frog, genotype LL), and their hybrid Pelophylax esculentus (Linnaeus, 1758) (the edible frog). Pelophylax esculentus comprises diploid individuals with one lessonae (L) and one ridibundus (R) genome (genotype LR) and triploid individuals that possess either two L genomes and one R genome (genotype LLR) or two R genomes and one L genome (genotype LRR).

Unlike its parental species and several hybrid forms with Mendelian reproduction, almost all *P. esculentus* reproduce hybridogenetically; that is, they exclude one of their parental genomes during gametogenesis (either the R or the L genome) and pass the remaining one clonally to their gametes. This reproductive mode, first described in the fish genus *Poeciliopsis* (Schultz 1969), enables diploid hybrids to reproduce via back-crossing with the parental species that provides the genome excluded from the hybrid germline. Because the genome delivered by the syntopic parental species undergoes Mendelian inheritance (e.g. Schmeller *et al.* 2001b), *P. esculentus* reproduces hemiclonally.

Hemiclonal individuals are usually unable to successfully procreate by mating with other hybrids (Dawley 1989) because of the irreversible accumulation of deleterious mutations in the clonally transmitted genome (Vorburger 2001; Guex *et al.* 2002; Vorburger *et al.* 2009). Hence, in populations with only diploid hybrids, *P. esculentus* is usually reproductively dependent on its syntopic parental species and therefore is considered a sexual parasite (Graf & Polls Pelaz 1989) that uses either *P. lessonae* or *P. ridibundus* as sexual hosts. Under these conditions, it is not surprising that the distribution area of *P. esculentus* is largely congruent with the ranges of the parental species (e.g. Plötner 2005).

In most parts of northwestern Europe, however, both P. lessonae and P. ridibundus are extremely rare or absent, while the hybrid occupies a variety of habitats up to 56° latitude (Ebendal 1979; Berger & Berger 1994; Arioli et al. 2010; Jakob et al. 2010). In these areas, 'allhybrid' (or 'pure') P. esculentus populations exist that reproduce and persist independently of the parental species. Such populations consist of both diploid (LR) and triploid individuals (LLR and/or LRR). Triploid individuals usually produce haploid gametes, while diploid females produce both haploid and diploid eggs (Uzzell et al. 1975). LR individuals originate from the fusion of haploid gametes while LLR and LRR frogs result from combinations between diploid and haploid gametes (Günther et al. 1979; Christiansen 2009; Arioli et al. 2010; Jakob et al. 2010). When two diploid gametes fuse, viable tetraploid individuals (LLRR) can occur, but in nature they are very rare (Borkin et al. 2004, 2006; Christiansen 2009; Arioli et al. 2010; Jakob et al. 2010). Most triploid frogs transmit the genome that is present in two copies; that is, the L genome is passed on by LLR triploids and the R genome by LRR triploids. Diploid (LR) females pass on the R genome, less frequently the L genome, and/or produce diploid LR eggs, while LR males form sperms with an R or an L genome (Günther et al. 1979; Christiansen 2009; Pruvost et al. 2013a). In a few areas, however, triploid hybrids, exclusively LLR males, produce diploid LL sperms (Tunner & Heppich-Tunner 1992; Brychta & Tunner 1994; Mikulíček & Kotlík 2001; Pruvost et al. 2013a; Mikulíček et al. 2015).

In addition to 'diploid populations' with just diploid hybrids plus one or both parental species and 'all-hybrid populations' with only diploid and triploid hybrids, there are also populations that combine features of the two types; that is, they comprise diploid and triploid hybrids plus one or both of the parental species. Such populations can be very diverse and complex with up to five different genotypes: LL, LLR, LR, LRR and RR in various combinations. In this study, we use the term 'mixed-ploidy populations' for populations with various combinations of di- and triploids, including all-hybrid populations.

The above summary is based on several more or less detailed studies of frogs from usually only a few populations in geographically fairly restricted areas. These studies indicate that the system is highly complex with several different regional patterns. Moreover, some of the studies show that the produced gamete types, in concert with environmental and geographic factors, have a significant influence on population structure (Christiansen 2009; Arioli *et al.* 2010; Jakob *et al.* 2010; Christiansen & Reyer 2011; Pruvost *et al.* 2015). Such locally restricted results can only deliver pieces of the puzzle concerning the evolutionary history of population systems and genomic composition of water frog populations and breeding systems. We therefore investigated the genetic structure and variability within frog populations over a large geographic scale, ranging from 43° to 60°N latitude and 8° to 36°E longitude in Europe. Specifically, we pursued the following three goals:

- 1 Characterization of population compositions across Europe, with respect to proportions of parental species and diploid and polyploid hybrids. We expected that all populations with diploid hybrids should also contain a parental species and/or triploid hybrids because, based on the presently known reproductive modes (see above), these are the indispensable donors of the premeiotically eliminated genome in diploids. We also expected all populations with triploids to contain diploids because these are usually the only providers of diploid gametes. Deviation from these expected composition patterns would indicate a previously undescribed breeding pattern, thus adding to our knowledge about this complex frog system.
- 2 Investigation of genetic diversity and genetic differences of parental species and hybrid forms in terms of nuclear DNA and mitochondrial DNA in relation to geography and population composition. We expected latitudinal (and perhaps also longitudinal) gradients in genetic variation, with lower diversity in the north and perhaps also the west. Water frogs must have survived glaciations in southern (and perhaps also eastern) glacial refugia and moved north- and westward during postglacial periods. Such range expansion often goes along with a loss of alleles and, hence, decreasing diversity towards the edge of the distribution range.
- **3** Assessment of how, where and how often diploid and polyploid hybrid populations have originated and how they are propagated. We do that by comparing the genetic similarity within and between diploid and triploid hybrids from populations ranging from north of the Baltic Sea to Bulgaria. Very close genetic relatedness of parental genomes in triploids from different populations would indicate a single origin of triploidy. Alternatively, if parental genomes in triploids than to each other, this would indicate multiple local origins of triploidy.

Methods

Sampling and population types

Genetic samples of water frogs from 72 and 102 populations were collected for microsatellite and mtDNA analysis, respectively. Sample size per population ranged from 5 to 238 individuals, with an average of 32 frogs per locality and >15 frogs sampled at 68% of the localities (see Tables S1 and S2, Data accessibility details). Adult water frogs were captured by hand or with a net. Tissue samples were taken from toe clips and stored in 80% ethanol until processed. Genotypes and ploidy of individuals were determined via microsatellite analysis and erythrocyte size measurement (methods are explained in more detail below). Based on this determination, populations were classified as 'diploid' when only diploid hybrids were found in the sample and no information indicated that polyploids were present in this area. We classified a population as of 'mixed ploidy' when, in addition to diploid hybrids, at least one polyploid individual was found. Population types were further subdivided based on cumulative information on their structure, that is the existence of LL, RR, LR, LLR and LRR individuals (Table S1, Data accessibility details). We distinguished between all LL (only LL, no hybrids), all RR (only RR, no hybrids), diploid L-E (LL and LR), diploid R-E (RR and LR), diploid L-E-R (LL, RR and LR) and mixedploidy populations (any combination with polyploid genotypes) including all-hybrid (E-E) populations.

Microsatellite marker selection and genotype determination

Microsatellite analysis was performed on samples from 72 localities (Table S1, Data accessibility details). DNA extraction, PCR and electrophoresis followed the protocols of Christiansen & Reyer (2009). To obtain useful data for subsequent population genetic analyses, it was important to select a set of microsatellites that work on the broad geographic range. We therefore started with 18 microsatellites known to be polymorphic and speciesspecific in water frogs from western, northern and central European populations, but allele distribution and variability in other European parts are not yet entirely known (for details, see Table 1). The 18 primer pairs were combined in two multiplex mixes of 9 pairs each. Sometimes these multiplex mixes were split into four mixes because allele overlap was observed during processing of samples from increasingly distant localities. PCRs with single primer pairs were routinely run to check the results from both individual primers and split-up primer mixes. We ran PCR products on an ABI 3730 Avant capillary sequencer (Applied Biosystems, Life Technologies Corporation, Carlsbad, CA, USA) with an internal size standard (GeneScan-500 LIZ; Applied Biosystems). Alleles were scored and peak heights were measured using the program GENEMAPPER 3.7 (Applied Biosystems 2004).

Multilocus genotypes were established from the allele data in a stepwise procedure. First, alleles were prescored by one of us (SR) without knowledge of locality or genomic composition (LL, LLR, LR, LLRR, LRR and RR). Scoring was then repeated independently by a person who had been in the field (AH, NBMP). Samples for which results did not match phenotypic and molecular determination were then run and scored again. As most primers amplified in only one of the two genomes (L or R, Table 1), it was immediately obvious from the microsatellite profile whether an individual had a parental or hybrid genotype. With prior expectations of genome specificity from previous studies (Christiansen 2005, 2009; Christiansen & Reyer 2009, 2011), supported by field notes describing the supposed taxon of the individual based on morphological characters, L specificity and R specificity could be inferred for the majority of loci. The hybrid types were also clearly distinguishable from parental genotypes by their amplification patterns. Moreover, LR, LLR and LRR genotypes could be distinguished because several L-specific microsatellites exhibited two distinct alleles in LLR individuals and several R-specific markers exhibited two alleles in LRR individuals. Based on these amplification patterns, consensus genotypes were determined for all individuals; that is, each individual was assigned to the genotype that was unambiguously revealed by all loci.

The ploidy of the consensus genotypes was verified by analyses of dosage effects at four loci (Res16, Ga1a19, RICA1b5 and RICA1b6), following the method established by Christiansen (2005). At these not very polymorphic loci, often just one L and one R allele are amplified in hybrids. In LLR frogs, the L peak is clearly higher than the R peak; in LR frogs, they are of similar height; and in LRR frogs, the L peak is lower than the R peak. Using this principle of dosage effect, plots of log10(peak height1/peak height2) were drawn for all pairwise combinations of alleles at these four loci in the entire data set. These plots were visually examined for groups of individuals corresponding to 2:1, 1:1 and 1:2 allele ratios. Depending on the genome specificity of the alleles, these ratios could be translated into LL, LLR, LR, LRR and RR types (LLL, LLLR, LRRR and RRR were not found). LLRR tetraploids might appear as LR at some dosage effect loci, because for both genotypes, the L:R allele ratio is 1:1. The chance of mistaking LLRR for LR, however, is very low because like triploids, tetraploids should be revealed by amplification of two distinct L or two R alleles at one or another locus of the polymorphic loci analysed. To verify the diploid state of individuals with an 'LR' allelic profile, we measured the average size of ten erythrocytes from blood smears from all individuals sampled in the field, except for frogs from Bulgaria, Romania and Ukraine, for which

Marker	Repeat motif	References	GenBank ID	Allele numbers diversity			and	genetic
				L	R	L + R	HeL	HeR
RICA1b6	(TG) _n	Arioli et al. (2010)	EF121548	3	18	2	0.486	0.730
RICA1b5	(CA) _n	Garner <i>et al.</i> (2000)	AF286388	7	8	0	0.092	0.275
Ga1a19 redesigned*	(CT) _n	Christiansen (2009); Arioli <i>et al.</i> (2010)	EF121547	4	29	1	0.038	0.652
Res16	(CA) _n	Zeisset et al. (2000)	AF195843	2	9	4	0.086	0.468
Rrid064A	$(GT)_n \dots \dots \dots (TA)_n (GT)_n$	Christiansen & Reyer (2009)	EU445524		19			0.733
Re2Caga3	(CAGA) ₆ TTA(GATA) ₁ (GACA) ₁ (GATA) ₂₀ (GACA) ₁	Arioli et al. (2010)	EF121550		41			0.900
Res22	(CA) _n	Zeisset et al. (2000)	AF195846		27			0.644
Rrid013A	(ACT) _n	Hotz et al. (2001)	FJ024047		6			0.249
Rrid059A redesigned	(CA) _n	Hotz et al. (2001); Christiansen & Reyer (2009)	FJ024048		29			0.645
Rrid135A	$(CA)_n(TA)_n$	Christiansen & Reyer (2009)	EU445526		27			0.718
RICA2a34	(GT) _n	Christiansen & Reyer (2009)	EU445521	18			0.734	
Ga1a23	(GA) _n	Christiansen & Reyer (2009)	EU445523	22			0.860	
Ca1A27	(CA) _n	Christiansen & Reyer (2009)	EU445522	15			0.827	
RICA18	(CA) _n	Garner et al. (2000)	AF286386	22			0.687	

Table 1 Microsatellite markers tested and selected in this study. Allele numbers and genetic diversity for L genomes (HeL) and R genomes (HeR) are given pooled across all samples

The marker indicated with * was not used in genealogical analyses of the L genome of mixed-ploidy populations (for explanation, see text).

blood smears were not available. This method recognizes that erythrocytes of polyploid frogs are larger than those of diploid frogs; in case of triploid individuals, the size difference is approximately one-third (Uzzell & Berger 1975; Günther 1977; Schmeller *et al.* 2001a). The results of erythrocyte measurements confirmed the ploidy level deduced from the consensus genotype, and we did not recognize any case where a tetraploid individual was mistaken as a diploid LR.

Based on these methods, genome specificity could be unambiguously assigned to the majority of previously unknown alleles (Table S4). Occasionally, however, we encountered single loci or alleles that were in conflict with the consensus genotype in certain geographic areas. For example, a locus may have yielded two L alleles when LR was expected by consensus genotype, or alleles considered as L specific occurred in the R genome. In these cases, the samples were examined again or rerun in PCR and fragment analysis. If loci were still incongruent with the consensus genotype after this extra round of evaluation, they were treated in one of the three ways:

1 When only 1–3 frogs from the same locality (populations with n < 15: 1–2 individuals; populations with n > 15: up to three individuals) showed the same kind of incongruence at the same locus, this locus was coded as missing data in the 1–3 frogs

concerned. We did this, because from that few individuals, it could often not be decided whether the incongruence was caused by allele unspecificity (see point 2 below) or the occurrence of null alleles (see point 3 below). With only one locus in 1–3 frogs in the whole data set, the problem was considered as quantitatively unimportant.

- 2 When more frogs were concerned, the conflicts could often be attributed to allele unspecificity; that is, the allele could not be assigned to either L or R because it repeatedly occurred in both genomes. In such cases, single alleles were re-assigned to either the L or R genome to fit the consensus genotype. In a data set used for calculation of genetic distances and Bayesian tests, allelic unspecificity pertained to only one allele (127) in locus Res16. This allele reached frequencies of 0.60 and 0.03 in *P. ridibundus* and *P. lessonae*, respectively. Another *P. lessonae*-specific allele at this locus was 121 with a frequency of 0.97. When a hybrid with a Res16 genotype 121/127 was found, allele 121 was attributed to its L and 127 to its R genome.
- **3** In the few remaining cases, the affected samples were deleted from the analysed data set, as conflicts could clearly be attributed to null alleles, because missing values occurred for some loci in certain geographic areas, despite multiple reruns. This indicated the occurrence of null alleles caused by mismatches

between the primers and their templates. In particular, populations from Ukraine and Romania were affected by the failure of amplification in the R genome. Because P. ridibundus is genetically very heterogeneous in southern and eastern Europe (Hotz et al. 1985; Plötner et al. 2008; Akin et al. 2010), nonamplification of some markers in individuals from eastern European populations was probably caused by the occurrence of null alleles. Null alleles can bias the estimates of population genetic parameters when they occur in high frequencies. In our study, however, we did not often observe persistent nonamplification after repeated runs of samples. We thus assume that null alleles occurred at low frequencies only (<1.0%), which was also found in two earlier studies on water frogs (Christiansen 2009; Pruvost et al. 2015).

After these exclusions, 14 markers were eventually used for statistical analyses: four for the L genome, six for the R genome and four that amplified in both genomes (Table 1). For genealogical analyses, we used only 13 markers, four for the L genome, six for the R genome and three for both parental genomes. Details are explained in the next section.

Microsatellite data analysis

For population genetic analyses, the L and R genomes were analysed separately because they almost never recombine. As a measurement for genetic diversity, we used He, the expected heterozygosity according to Nei (1978), calculated by the program SPAGEDI 1.3 (Hardy & Vekemans 2002). SPAGEDI can handle a mix of haploid and diploid data; it was also used to calculate Nei's (1978) $D_{\rm s}$, $F_{\rm st}$ and geographic distance matrices among populations. To test the influence of geographic distribution and population parameters on genetic diversity based on microsatellite data, we performed generalized linear type II models (GLMs) using SYSTAT 11 (SYSTAT Software Inc. 2004). Mantel's tests between genetic distance and geographic distance matrices were calculated with the program ZT (Bonnet & Van de Peer 2002). Pairwise geographic distances were calculated from GPS coordinates using the online software geographic distance matrix generator 1.2.3 (Ersts 2012).

To test whether R and L genomes present in different types of hybrids and parental species are related to each other, and to find geographic structuring of both genomes, Bayesian assignment tests implemented in STRUC-TURE 2.3.3 (Pritchard *et al.* 2000; Falush *et al.* 2007) and distance-based methods implemented in POPTREEW (Takezaki *et al.* 2014) were applied. L and R parental genomes were analysed separately. For POPTREEW analyses, D_A distances (Nei *et al.* 1983) were calculated on the basis of seven and ten polymorphic markers, respectively (Table 1). One marker (Ga1a19 redesigned), which was used in analyses of genetic diversity, yielded only two alleles in the L genome across all mixedploidy populations and was therefore omitted in the genealogical analysis of the L genome (indicated by an asterisk in Table 1). For tree constructions, the neighbour-joining (NJ) method was applied. Tree robustness was evaluated by bootstrapping (Felsenstein 1985) with 1000 replicates.

Bayesian algorithms implemented in STRUCTURE use an iterative approach to assign genotypes into K clusters without a priori knowledge of the population membership of individuals, assuming Hardy-Weinberg (H-W) and linkage equilibrium within the inferred clusters (Pritchard et al. 2000). These assumptions are unlikely to be met in populations of hybrid and clonal organisms because of fixed heterozygosity and linkage of multilocus genotypes. Nevertheless, several studies demonstrated that the Bayesian models implemented in STRUCTURE are robust to deviations from these assumptions and provide biologically meaningful structuring supported by other independent analyses (e.g. Halkett et al. 2005; Schmidt et al. 2011). Because L and R genomes were analysed separately in STRUCTURE, the excluded genome in hybrids was coded as missing data (see Pritchard et al. 2010). For instance, analysing the L genome of LR and LRR hybrids, R-specific alleles were excluded and coded as missing. All individuals were then considered to be diploid (the STRUCTURE software does not enable simultaneous analysis of different ploidy levels). For assignment tests, admixture and uncorrelated allele models were applied. The most likely number of clusters (K) was assessed using the ΔK statistics (Evanno et al. 2005) implemented in the online program STRUCTURE HARVESTER (Earl & vonHoldt 2012, available at http://taylor0.biology.ucla.edu/structureHarvester/#), and assuming prior values of Kbetween one and ten. The analyses were based on runs of 10⁶ iterations, following a burn-in period of 100 000 iterations. A series of 10 independent runs for each K was made with the same parameters to test the accuracy of the results. Graphical representations of the STRUCTURE results were carried out using an online application CLUMPAK (Kopelman et al. 2015, available at http://clumpak.tau.ac.il).

Mitochondrial DNA sequence analysis

To estimate haplotype diversity, we analysed mtDNA sequences (the complete ND2 and ND3 gene; in total, 1378 bp) of 1175 samples from 105 localities (Tables S2 and S3). Because haplotype diversity is rather low

within a single population (Plötner *et al.* 2008), we sequenced mtDNA of usually 5–20 and in a few case up to 50 individuals per population. We also included mtDNA sequences from an earlier study (Arioli 2007; chapter 5) as well as additional populations for which we had too few samples to be used for microsatellite analysis. DNA extraction, PCR and sequencing were conducted using the protocols described by Plötner *et al.* (2008). Both genes were sequenced in the sense and antisense directions.

Mitochondrial DNA sequences were initially aligned using the algorithm CLUSTALW as implemented in the program MEGA 6.0 (Tamura *et al.* 2013). Subsequently, the alignments were improved manually. The model that best describes the substitution patterns of concatenated ND2 + ND3 sequences (1378 bp) was selected on the basis of the Bayesian information criterion (Schwarz 1978) as implemented in MEGA. Model selection was based on maximum likelihood (ML). The ML algorithm was also applied to estimate a haplotype genealogy using all sites for gaps/missing data treatment and Nearest-Neighbor-Interchange as the heuristic search method. Nodal support was evaluated by bootstrapping (Felsenstein 1985) with 1000 replicates.

Results

Population types and genotype distribution

Based on microsatellite profiles, we genotyped and analysed a total of 2062 frogs from 72 localities. The minimum distance between localities was 2.63 km, and the maximum was 1863.5 km. The most numerous taxon was *Pelophylax esculentus* with 63% of all individuals, followed by *P. ridibundus* (25.5%) and *P. lessonae* (11.5%). *P. esculentus* occurred at 50 localities (69%), *P. ridibundus* at 40 (56%) and *P. lessonae* at 27 (38%). All sample localities are listed in Table S1 (Data accessibility details) and referred to in square brackets throughout this study; for example, [1] refers to the population in Uppsala.

We found 20 all-*P. ridibundus* populations (26% of localities), almost exclusively south of 48° latitude and east of 16° longitude, especially in the proximity of the numerous tributaries to the Danube River and the Black Sea (Fig. 1a). One all-*P. ridibundus* population [7], however, was situated quite remotely from the rest in the Baltic area. We found only two all-*P. lessonae* populations, one in Sweden, the other in Estonia ([1], [5]; 2.6% of all samples); both constituted the northernmost populations sampled.

The remaining 50 populations included the hybrid *P. esculentus*. Of these, 26 (36% of total) were classified as diploid and 24 (33%) as mixed-ploidy populations



Fig. 1 (a) Sample locations of individuals used for microsatellite analyses. Symbols and colours indicate populations of the following types: \bigcirc = all *Pelophylax lessonae*, \bigcirc = all *Pelophylax ridibundus*, \bigcirc = mixed ploidy, green = diploid hybrids with one or both parental species (\diamondsuit = L-E, \blacktriangle = R-E, \blacksquare = L-E-R populations; L = *P. lessonae*, R = *P. ridibundus*, E = *Pelophylax esculentus*). In areas where several sampling sites are close together, the symbols have been slightly displaced from the real location for better clarity. (b) Geographic distribution of the L-genotype clusters from Fig. 5. In both graphs, letters denote relevant mountain ranges. WA = Western Alps, EA = Eastern Alps, WC = Western Carpathians, EC = Eastern Carpathians, SC = Southern Carpathians, DM = Dinaric Mountains, BM = Balkan Mountains.

that contained triploid LLR and/or LRR individuals. These two population types show different geographic distributions which overlap only in some areas (Figs 1 to 3). Mixed-ploidy populations are fairly evenly distributed between 48° and 56°N latitude. In contrast, 20 of the 26 diploid populations (77%) occur at lower latitudes (42°–50°N) and four at higher latitudes (57°–61°N), two with hybrids (L-E system) and two without (pure LL populations). In terms of longitude, distributions of diploid and mixed-ploidy populations differed, too. Nineteen of the 24 mixed-ploidy populations (79%)

were sampled between 10° and 20°E longitude, two around 25° and three at >32°E longitude. In contrast, diploid populations were most numerous in the range 15° -26°E longitude and sampled only sporadically at lower or higher longitudes.

In agreement with our expectations (see goal 1 in the Introduction), most diploid hybrids occurred in populations that also contained parental species and/or triploids. Of the four populations where we caught diploid hybrids only, three were of low sample sizes (N = 5, 7 and 9), but in the Polish population [11], we caught 43 frogs. As expected too, all triploid individuals co-occurred with diploid LR hybrids. In 50% of the mixed-ploidy populations, polyploid hybrids additionally co-occurred with either P. ridibundus or P. lessonae. Only in two populations ([11] and [54]), diploid and polyploid hybrids lived in sympatry with both parental species. Fourteen localities (19.4%) included both types of triploid hybrids, LLR and LRR. In another six populations, LLR was the only type of triploid hybrids, whereas two populations included only LRR. Only four tetraploid LLRR were detected in three populations ([4], [11] and [24]), that is in 0.2% of all samples.

Effects of geographic and population parameters on genetic diversity

In the L genome, 100 microsatellite alleles were found (range: 5–22 per locus) across the entire sample of 1506 individuals from 50 populations (Table 1). From the R genome, 220 microsatellite alleles (range: 6–41 per locus) were obtained from 1807 individuals from 66 populations (Table 1). Accordingly, genetic diversity was generally lower in the L genome than in the R genome (mean HeL = 0.32 for populations with LL and all types of hybrids; mean HeR = 0.44 for populations with RR and all types of hybrids). This is evident by a comparison of Figs 2a and 3a including the intercept of the regression lines. All four markers that amplified in both genomes showed lower genetic diversity in the L than in the R genome (Table 1).

We investigated whether geographic locality and composition of the population influences genetic diversity (*He*) in the R and/or L genome. For *He*L and *He*R, we performed separate GLMs in which we incorporated latitude, longitude, population type (diploid vs. mixed ploidy), the proportions of parental species (% LL, % RR) and polyploid hybrids (% LLR, % LRR) as independent variables and tested for interactions between geographic parameters and population type (poptype × latitude, poptype × longitude) (Table 2). Both *He*L and *He*R showed a negative correlation with increasing latitude (Table 2, visualized in Figs 2 and 3), which means that genetic diversity in both genomes



Fig. 2 Genetic diversity of the L genome (*HeL*) plotted against (a) latitude and (b) longitude for all-*Pelophylax lessonae* populations (yellow), three types of diploid populations (green) and mixed-ploidy populations (black). The regression with equation and the explained variance (R^2) in (a) is based on all population types. For details, see Tables 2 and S1 (Data accessibility details).

decreases from south to north. *HeR* also increased from west to east. None of these geographic effects differ between diploid and mixed-ploidy populations, as indicated by the lack of significant interactions with population type (Table 2).

As expected (see goal 2 in the Introduction), *HeL* significantly increased with the percentage of *P. lessonae* (% LL) and *HeR* with the percentage of *P. ridibundus* (% RR; Table 2). However, against our expectations, the proportions of the two polyploid hybrid types had no significant effect on either *HeL* or *HeR*, irrespective of whether % LLR and % LRR were entered separately into the model or pooled into % polyploids. Figure 3 also shows that among the diploid populations, *HeR* is higher in pure RR populations than in those including hybrids. This effect of the parental species is confirmed by our result that higher proportions of the parental species correlate with higher genetic diversity in the respective genome.



Fig. 3 Genetic diversity of the R genome (*HeR*) plotted against (a) latitude and (b) longitude for all-*Pelophylax ridibundus* populations (blue), three types of diploid populations (green) and mixed-ploidy populations (black). The regression lines with equations and explained variances (R^2) in (a) and (b) are based on all population types. For details see, Tables 2 and S1 (Data accessibility details).

Isolation by distance

For the entire sample, calculation of global F_{st} values yielded 0.349 for the L genome and 0.294 for the R genome, thus attributing 35% of variation in the L genome and 29% in the R genome to interpopulation differences. When we tested for isolation by distance across all populations, we found genetic distance (given as Nei's D_s) to increase strongly with geographic distance between populations in both genomes (one-tailed Mantel's tests: L: r = 0.63, P < 0.001; R: r = 0.65, P < 0.001; Fig. 4). Mantel's tests on F_{st} values between populations yielded similar results, yet the effect was smaller (one-tailed Mantel's tests for L: r = 0.22, P < 0.001; R: r = 0.35, P < 0.001).

Isolation by distance was supported for the 24 mixedploidy populations in our data set (one-tailed Mantel's tests for L: r = 0.54, P = 0.0003; R: r = 0.43, P = 0.0035), showing that genetic relationships among these polyploid populations are distance dependent.

Genetic structuring of water frog populations based on microsatellites

According to the ΔK statistics, two and four distinct Bayesian clusters were recognized in the R and L genome, respectively (Fig. 5a,b, Table S5). Within the R genome, water frogs from eastern Ukraine formed a distinct cluster (orange in Fig. 5a), supported also in the NJ tree (Fig. 6, Table S6). All other frogs were assigned to the second STRUCTURE cluster. Within the L genome, genotypes from eastern Ukraine formed a distinct cluster in the STRUCTURE (purple in Fig. 5b), as well as in the NJ tree (Fig. 7, Table S8). The second cluster comprised genotypes from populations situated in Germany, Czech Republic, Poland and Sweden (orange in Fig. 5b, blue in Figs 1b and 7). The third cluster was composed mainly of genotypes from the Czech, Slovak, western Ukrainian and partly southern Polish populations (green in Fig. 5b). The fourth cluster (blue in Fig. 5b, red in Figs 1b and 7) included mainly genotypes from populations of northeastern Germany west of the Oder River- and Bornholm Island.

L and R genomes of triploid hybrids (LLR, LRR) were scattered throughout the NJ trees (Figs 6 and 7), as well as were assigned to different STRUCTURE clusters (Fig. 5a, b). Both distance-based and Bayesian methods thus reflect the fact that triploid forms originated independently in different geographic regions. In general, triploid hybrids were more closely related to syntopic diploid hybrids and the parental species than to triploids from other populations (Figs 5 to 7). Exceptions to this rule were mainly populations [31-Borovec] concerning the L genome (Figs 5b and 7) and populations [14-Rügen, 20-Altenhausen, 23-Döbern] concerning the R genome, where triploid hybrids were assigned to different NJ clusters than syntopic diploid hybrids (Fig. 6). This may indicate different geographic origins of L or R genomes of diploids and triploids.

Mitochondrial haplotype diversity and structure

We sequenced 1175 samples from 105 populations, which yielded a total of 75 haplotypes (Tables S8 and S9). Based on the maximum-likelihood (ML) tree, four mitochondrial clusters were distinguished: *P. lessonae*, *P. bergeri* (a species distributed on the Apennine Peninsula, Sicily and Corsica), *P. cf. bedriagae* (a so far unnamed taxon from Anatolia) and *P. ridibundus* (Fig. 8). The *P. lessonae*-specific ND2 + ND3 sequences exhibited 30 variable sites (25 in ND2 and 5 in ND3), which resulted in 40 haplotypes (Table S6). Nucleotide

	HeL			HeR	HeR			
Variables	d.f.	F	Р	d.f.	F	Р		
Population type	1	0.001	0.974	1	0.719	0.400		
% LLR	1	0.122	0.729	1	3.803	0.056		
% LRR	1	0.085	0.772	1	2.222	0.142		
% LL*	1	14.861	< 0.001					
% RR [†]				1	74.631	< 0.001		
Latitude	1	5.638	0.022	1	5.124	0.027		
Longitude	1	2.670	0.108	1	6.962	0.012		
Poptype \times latitude	1	0.154	0.697	1	0.609	0.438		
Poptype \times longitude	1	3.645	0.060	1	1.562	0.217		
Error	41			57				

Table 2 Results from two stepwise GLM analyses of genetic diversity in the L genome (*He*L) and the R genome (*He*R) vs. geographic and population parameters

Significant results are shown in bold.

*Only available in the HeL data set.

[†]Only available in the *He*R data set.



Fig. 4 Isolation by distance in (a) the L genome and (b) the R genome. Pairwise distances $D_{\rm S}$ (according to Nei 1978) between all 72 microsatellite sampling sites are shown. Geographic distance was ln-transformed for better illustration.

diversity (π) among *P. lessonae* haplotypes ranged from 0.0007% to 0.82%, with a mean of 0.0031% (\pm 0.0007 SD). *P. ridibundus*-specific ND2 + ND3 sequences exhibited 35 variable sites (29 in ND2 and 6 in ND3), which

resulted in 32 haplotypes (Table S7). Among these types, π ranged from 0.0007% to 1.57%, with a mean of 0.0068% (± 0.0012 SD). We found 806 samples (68.6%) that carried haplotypes that were P. lessonae specific; 343 (30%) had P. ridibundus-specific haplotypes, and 15 (1.3%) possessed haplotypes characteristic of P. bergeri. Only one haplotype originated from Anatolian water frogs, *P.* cf. *bedriagae*. *P. lessonae*-specific haplotypes were found in P. lessonae, diploid and triploid hybrids of both types (LLR, LRR), and also in P. ridibundus from central Europe (Table S2). In contrast, P. ridibundus-specific haplotypes were never found in P. lessonae but only in P. ridibundus and in hybrids—in the latter, however, less often than the P. lessonae-type mtDNA (mostly in LR, very scarcely in LRR, never in LLR).

The most common *P. lessonae*-specific haplotype (L1) was found in all genotypes and all population types except all-P. ridibundus populations; the most common P. ridibundus-specific haplotype (R1) was detected only in LR and RR genotypes and in L-E-R, R-E, mixedploidy and all-P. ridibundus populations. The best fit model of sequence evolution for the concatenated ND2 and ND3 sequences was the Tamura-Nei (TN93) model (Tamura & Nei 1993) with gamma-shaped rate variation (G = 0.19). The ML analysis yielded significant differentiation between P. lessonae-, P. ridibundus-, P. bergeriand the Anatolian P. cf. bedriagae haplotypes (Fig. 8, Table S9). Within the P. lessonae group, two clusters (LES-1 and LES-2) can be differentiated; polyploid individuals (LLR, LRR) were found in both clusters with a total of 16 haplotypes. Eight P. lessonae-specific haplotypes (from both clusters) were only found in P. lessonae individuals, and seven haplotypes from cluster LES-1 were also found in P. ridibundus, in addition to hybrids



Fig. 5 Structuring in the R (a) and the L (b) genome according to STRUCTURE. Two and four distinct clusters were recognized in the R and L genomes, respectively, according to the ΔK statistics.

and *P. lessonae*. The *P. ridibundus* clade showed no significant differentiation (Fig. 8). Haplotype R1 is most widespread; it was found in both *P. esculentus* and *P. ridibundus*. Haplotype (R6) was detected for the first time in triploid individuals, specifically in 5 LRR individuals from population [52] in Ukraine.

Discussion

Our analysis of more than 2000 frogs from 72 localities with 18 microsatellite markers and two mtDNA genes showed that hybrid and parental genotypes are not evenly distributed across Europe and that genomic composition of populations, as well as latitude and longitude of their locations, affects genetic diversity. Genealogical relationships clearly indicate that hybrids are of multiple origin, with all-hybrid populations as a possible initial basis for hybrid speciation. Below, we discuss the findings in relation to the three goals specified in the Introduction and to results obtained from other species.

Population composition patterns across Europe (goal 1)

The distribution areas of population types overlap considerably (Fig. 1a). Populations with one or both of the parental species and only diploid hybrids are common in central Europe, whereas pure *Pelophylax ridibundus* populations dominate south of the eastern Alps and the Carpathians; pure populations of *P. lessonae* were found only at two remote northern localities in Sweden and Estonia [1; 5]. The apparent concentration of diploid L-E populations close to mountain ranges like the Alps or the Carpathians, and of pure *P. ridibundus* populations close to large rivers and coastal areas of the Black Sea, may be a reflection of both climate-mediated ecological constraints and a possible sample bias.

Mixed-ploidy populations dominate in northwestern Europe and seem to decrease in relative frequency towards central Europe. They do not occur further south than the rims of the Carpathian Mountains, which are thought to have played a significant role as a dispersal barrier for amphibians (Hofman *et al.* 2007;



Fig. 6 R genome-based genealogical relationships of genotypes (RR, LR, LLR, LRR) from 23 mixed-ploidy water frog populations based on Nei's genetic distances (D_A) calculated from polymorphic microsatellite markers. One population [54] could not be included in the analysis because of several null alleles. Labels of terminal units: population number, locality name, genotype and country. Country abbreviations: CZ, Czech Republic; DE, Germany; DK, Denmark; PL, Poland; SE, Sweden; SK, Slovakia; UA, Ukraine. Bootstrap values below 50 are not shown.

Zielinski *et al.* 2014). Previous studies have shown that polyploidy (LLR and LRR, very rarely tetraploidy) in *P. esculentus* occurs frequently and across different population systems (including mixed systems with *P. lessonae* and *P. ridibundus*, or both) in Ukraine and Russia (Borkin *et al.* 2004, 2006), especially along the Donets River, a large fluvial area in eastern Ukraine. We could not confirm a high frequency of triploidy for this area, probably because of a limited sample size. The localities in the Czech Republic [31] and Slovakia ([33], [36]), where we found triploid LLR individuals, basically cover the area where triploids had already been discovered in previous years (Mikulíček & Kotlík 2001; Pruvost *et al.* 2013a, 2015). In Hungary, however,



Fig. 7 L genome-based genealogical relationships of genotypes (LL, LR, LLR, LRR) from 24 mixed-ploidy water frog populations based on Nei's genetic distances (D_A) calculated from seven polymorphic microsatellite markers. The red marked clade comprises genotypes from populations of northeastern Germany west of the Oder River, while the blue marked clade consists of genotypes east and south of the Oder River including populations from Bornholm Island and Sweden; among the remaining populations (shown in black), the pattern is less clear, although Czech and Slovak populations (31, 33, 36) may represent another cluster than eastern Ukrainian ones (51, 52). Labels of terminal units: population number, locality name, genotype and country. Country abbreviations: CZ, Czech Republic; DE, Germany; DK, Denmark; PL, Poland; SE, Sweden; SK, Slovakia; UA, Ukraine. Bootstrap values below 50 are not shown. For the geographic distribution of the clusters, see Fig. 1b.

we failed to find triploid water frogs, although we sampled at two localities ([46], [47]) close to the area where LLR individuals were observed more than two decades ago (Tunner & Heppich-Tunner 1992). Instead, here we found exclusively *P. ridibundus*. Possible explanations for this could be that we either did not sample the exact same population, or that environmental conditions in this area have changed in a way so that *P. ridibundus* has gained a selective advantage over *P. esculentus*. The latter can lead to a (local) decline of polyploid hybrids (cf. Christiansen 2009). The low density of mixed-ploidy populations just slightly north of the region between the Alps and the Carpathian range probably reflects a true rarity of this population type in this area.

Genetic diversity and genetic differences in parental species and hybrid forms (goal 2)

Nuclear DNA. Our nuclear genetic markers suggest that diploid and triploid water frogs exhibit similar genetic variability and that genetic diversity of diploid hybrids increases, as expected, with the percentage of



Fig. 8 Phylogenetic relationships among 75 mtDNA haplotypes, as inferred from maximum-likelihood analysis. The units on the scale bar are expected mutations per site. Filled symbols indicate haplotypes found in polyploids. Bootstrap values are given at nodes. Values higher than 95% are considered statistically significant; values below 80% are not shown.

parental species genotypes, but not with triploids. The genetic diversity of the R genome was higher compared to the L genome; this finding is mainly attributable to the excess of all-P. ridibundus populations compared to all-P. lessonae populations (20 vs. 2) in our sample and, hence, a higher effective population size (Ne) of P. ridibundus. The population size effect is further supported by the fact that the highest genetic diversity values for the R genome (HeR) were found in southeastern Europe, where pure populations of *P. ridi*bundus are very common. All-P. lessonae populations, in contrast, were rarely found (cf. Plötner 2005) and seem to occur only in the marginal northernmost parts of the water frog distribution area, where they exhibit low genetic variation (Sjögren 1991; Sjögren-Gulve & Berg 1999).

In both the L and R genomes, genetic diversity decreased with higher latitude, that is from southern to northern Europe. These clines emerged even after controlling for differences in sympatry between hybrids and parental species and in relative frequencies of parental species and diploid and triploid hybrids by including the percentages of LLR, LRR, LL and RR, respectively (Table 2). In the R genome, we also found an east-west decline in diversity. The clines are probably a result of rapid postglacial colonization and founder events caused by a few successful long-distance dispersers, which after arrival in a new locality may inhibit the establishment of later arriving genotypes through high-density blocking (Waters et al. 2013). This is consistent with the results from several continental species that used southern and eastern glacial refugia, and lost diversity through repeated founder effects when later spreading north- and westward (Schmitt 2009). Such a north-westward expansion route has probably also been taken by water frogs and other European amphibians (Hewitt 1996, 1999; Taberlet et al. 1998; Zeisset & Beebee 2001; Stöck et al. 2012). The effectiveness of long-distance dispersal and high-density blocking in producing geographic partitioning of recolonizing genotypes can be enhanced by ecological superiority of the founder individuals (Waters et al. 2013). This may have contributed to the success of P. esculentus in northern regions. As hybrid larvae develop faster and perform better than those of the parental species at cold temperatures (Negovetic et al. 2001; Pruvost et al. 2013b), *P. esculentus* may have had (and still have) a competitive advantage under the shorter and cooler summers in the north.

Genetic diversity of diploid hybrids (LR) did, as expected, increase with the percentage of parental species genotypes (LL or RR), but not with triploids (LLR and/or LRR). This was a surprise because LL and LLR frogs have the same number of L genomes and thus the same capacity to contribute genetic diversity to the population (and likewise for LRR and RR). Maybe the explanation lies in the history of different population types. It is possible that localities with many individuals of the one or the other of the parental species represent old core areas of that species with high genetic diversity, whereas areas with few parental individuals tend to be relatively newly colonized localities affected by founder effect.

Genetic differentiation of R and L genomes from eastern Ukraine (Bayesian clustering) and their basal positions in the NJ trees indicate that they may represent independent evolutionary lineages. Another distinct lineage in the L genome was detected in western Slovakia, western Ukraine and northeastern Czech Republic and partly in southern Poland. Genetic differentiation of these populations is reinforced by the fact that triploid LLR males from the Czech and Slovak populations ([31], [33] and [36]) produce diploid LL sperm and, thus, differ in their mode of gametogenesis from other triploid forms (Mikulíček & Kotlík 2001; Pruvost *et al.* 2013a; Mikulíček *et al.* 2015).

mtDNA. In areas where P. esculentus occurs, many P. ridibundus carry P. lessonae-specific mtDNA, whereas the reverse pattern, that is a transfer of P. ridibundusspecific mtDNA into the P. lessonae gene pool, was never observed. In eastern and southeastern Europe, however, P. ridibundus carries P. ridibundus-specific mtDNA exclusively, irrespective of the co-occurrence of P. esculentus or P. lessonae. Hybrids reflect this pattern to a large extent. While P. esculentus possesses P. lessonae-specific mtDNA in areas where presently primary hybridization does not (or only occasionally) occur because of the absence or rarity of P. ridibundus, diploid hybrids may carry both types of mtDNA in areas with both parental species. Moreover, all triploids carried P. lessonae-specific mtDNA except five LRR individuals from population [52] in the Ukraine, which possessed a P. ridibundus-specific mitogenome. These findings generally confirm the common patterns of mitochondrial inheritance in European water frogs that have been described elsewhere (Spolsky & Uzzell 1986; Plötner et al. 2008; Mikulíček et al. 2014).

Our results also provide evidence that *P. esculentus* can incorporate genetic material from related taxa liv-

ing close to the distribution boundaries of its parental species. In this study, this holds for the proximity of the Swiss and German populations to the contact zone between P. lessonae and its sister species, the Italian pool frog P. bergeri, as well as to the transition zones between P. ridibundus and Anatolian water frogs in eastern Greece and west of the Caspian Sea. Pelophylax bergeri-specific mtDNA was found in the three westernmost populations from central Germany and Switzerland ([22], [24] and [40]), thus confirming earlier reports for Switzerland and Southern Germany (Hotz et al. 1992; Plötner 2005; Ohst 2008; Plötner et al. 2008). In the two diploid populations [22, 40], P. bergeri-specific mtDNA was carried both by P. lessonae and by LR hybrids. In population [24], we detected P. bergeri-specific mtDNA in diploid hybrids and-for the first time-in one triploid LRR female. Compared to other central European populations, population [24] is very diverse: it contains LR hybrids (61.7%), LRR triploids (25.0%), some RR individuals (11.4%) and a small percentage (2.3%) of tetraploids (LLRR). Besides P. bergeri-specific mtDNA, P. lessonae- and P. ridibundus-specific haplotypes were detected in this population. While the P. lessonae haplotype is typical for water frog populations of north and central Europe, the P. ridibundus haplotypes are rare in central Europe but much more common further east. As indicated by our microsatellite data, this population also contains L genomes that are obviously distinct from L genomes of other central European populations. Whether these genomes are autochthonous for this area or have originated from introduced individuals remains an open question. They may, for instance, have come with fish fry ([24] is a > 200-year-old fish farming pond) or been released from frogs traded for food or as laboratory animals.

Concerning the transition zones between *P. ridibundus* and other water frog species, we have evidence for horizontal transfer of mtDNA from P. cf. bedriagae (cf. meaning similar to yet genetically distinct from P. bedriagae) into a diploid P. esculentus-like hybrid in southeastern Ukraine (population [52]). Frequent introgression of P. cf. bedriagae-specific genes into the gene pool of P. ridibundus and P. esculentus is known from western Europe and adjacent areas where during the last decades, Anatolian frogs have been released intentionally or accidentally into natural habitats. Here, they successfully reproduce, inter alia, by backcrossing with indigenous forms (Holsbeek et al. 2008; Ohst 2008). If, for example, a P. esculentus male mates with a P. cf. bedriagae female or with a P. cf. bedriagae \times P. ridibundus hybrid female carrying P. cf. bedriagae-specific mtDNA, P. esculentus-like genotypes with P. cf. bedriagae-specific mtDNA may originate, provided that the hybrid male inherits the L genome. Besides southwestern Germany, Belgium and Switzerland, hybrids tracing back to crosses *P*. cf. *bedriagae* \times *P*. *ridibundus* were also detected in natural populations in eastern Greece (Hotz *et al.* 2013) and west of the Caspian Sea (J. Plötner & S. N. Litvinchuk, unpublished).

Origin of triploids and their putative role for hybrid speciation (goal 3)

There are various pathways that can lead to polyploidy in animals (Mable et al. 2011; Choleva & Janko 2013). For triploid water frogs, the 'genome addition hypothesis' seems to be the prevalent way of triploidization; that is, an unreduced egg from a diploid hybrid female is fertilized by a haploid sperm. This is supported by both artificial crossing experiments and the correlation between the relative frequencies of diploid gamete types and triploid hybrids in natural populations (e.g. Günther 1970; Uzzell et al. 1975; Berger & Roguski 1978; Christiansen 2009; Pruvost et al. 2013a). The mechanisms, however, that originally have given rise to unreduced gametes (mostly eggs) during gametogenesis are not known. As diploid gamete production varies substantially among species, sexes, populations and individuals, it must be controlled by both genetic and environmental factors. Genetic factors are directly linked with hybridity: because of segregation problems during meiosis, interspecific hybrids are more likely to produce 2n gametes than nonhybrids, not only in plants but also in animals including water frogs (Vrijenhoek et al. 1989; Levin 2002). The fact that approximately two-thirds of polyploid animals have abandoned recombination between the parental genomes and reproduce clonally, testifies to the importance of avoiding meiotic disturbances (reviewed by Vrijenhoek et al. 1989; Beukeboom & Vrijenhoek 1998; Otto & Whitton 2000). Environmental factors that can contribute to such disturbance of proper meiotic segregation, and thus enhance diploid gamete production, include extreme temperatures, poisonous substances, nutrient deficiencies, parasites and other stress-causing factors (Kawamura 1984; Levin 2002; Kondo & Kashiwagi 2004; Mable 2004; David & Pandian 2006). Hence, cold climate during glacial or postglacial periods could have initiated the formation of unreduced gametes and, thus, offers one explanation for the relatively high proportion of polyploid plants and animals found under the harsh conditions at high latitudes and altitudes (Mable 2004; Mason & Pires 2015). Temperature effects on diploid gamete production have been documented for fishes, newts and anurans, but not yet for water frogs (Fankhauser & Griffiths 1939; Kawamura 1984; Kondo & Kashiwagi 2004; David & Pandian 2006).

Ploidy elevation in association with genome multiplication is considered an important mechanism for increasing genotypic and phenotypic diversity which natural selection can act upon. Compared to diploids, triploids with two different copies of one parental genome and one copy of the other parental genome (e.g. L₁L₂R) can potentially exhibit more allele combinations and also recombine the parental genomes they have twice (Günther et al. 1979; Christiansen & Reyer 2009). Moreover, the duplication of homologous genes may result in subfunctionalization, that is differential expression of these homologues in response to an array of stressful conditions (Force et al. 1999; Madlung 2013). For triploid water frogs, evidence for differential gene expression, rather than a gene dosage effect, comes from proteins (J. Plötner, unpublished results), mating calls (Hoffmann & Reyer 2013) and morphometric characters (Plötner et al. 1994; Tunner 2000).

Our microsatellite data confirm that P. esculentus represents a genetically very diverse hybrid form that comprises different genotypes, ploidy levels and hemiclonal lineages, thus allowing for high ecological plasticity. As a result, the hybrid forms inhabit a wider range of habitats than the parental species (e.g. Pagano et al. 2008). They even occur in regions where both parental species are absent (see Fig. 1a and review by Plötner 2005). Here the hybrids are reproductively independent from their parental species and, thus, are evolutionary significant units that may represent a preliminary stage of hybrid speciation. From a global perspective, they are not as independent, however, as some other taxa with triploid forms, including sexually reproducing green toads of the Bufo (Bufotes) viridis complex (Stöck et al. 2002, 2005), parthenogenetic reptiles (Kearney et al. 2009) or gynogenetic fishes (Lamatsch & Stöck 2009). In contrast to these organisms, maintenance of triploidy in water frogs requires the continuous reproductive interaction between diploid hybrids as donors of diploid gametes (usually females producing LR eggs, occasionally LLR males producing LL sperm), and triploid hybrids as donors of haploid gametes (Günther et al. 1979; Som & Reyer 2006).

Thus, similar to systems in which sexual and asexual lineages interact (e.g. Janko *et al.* 2012), the stability and dynamics of these all-hybrid *P. esculentus* populations is based on a mutual reproductive dependence of diploid and triploid hybrids (Günther 1975; Som & Reyer 2006). Whether, at some point, they will also evolve into independent all-triploid populations remains an open question. Together with previous investigations on hemiclone diversity (e.g. Hotz *et al.* 2008; Pagano *et al.* 2008), the genetic diversity presented in this study and the close relationships between diploid and triploid

hybrids from the same populations (Figs 5 and 6) are evidence for past and present multiple primary hybridization events. They also indicate that hybrids are not the result of one unique event, but have repeatedly originated (and continue to originate) independently. Such recurrent origin of polyploids is also known from plants (Soltis & Soltis 1999) and several animals, including ostracods (Little & Hebert 1997), clonal fishes (Alves et al. 2004; Choleva et al. 2012; Janko et al. 2012), reptiles (Moritz et al. 1989) and some amphibian species (Ptacek et al. 1994; Stöck et al. 2005; Holloway et al. 2006; Vrijenhoek 2006). Where the genetically distinct polyploid lineages get in contact with each other, hybridize and recombine their genomes, they will provide an important source of additional genetic variation (Soltis & Soltis 1999).

Based on the high genetic and demographic diversity in water frog populations, we can imagine various scenarios for their future development. These include continuance of the present diploid-triploid assemblages, development into pure triploid populations as in the examples mentioned above, or transition to tetraploidy with restored normal meiosis. The last scenario could be achieved in two ways, which both have been demonstrated for cyprinid fishes of the Squalius alburnoides complex (Alves et al. 1999, 2001, 2004). First, triploids can act as a stepping stone towards tetraploidy, if they produce triploid gametes which upon fertilization with haploid gametes result in tetraploid individuals ('triploid bridge'; Ramsey & Schemske 1998; Mable 2004; Cunha et al. 2008). In water frogs, however, there is no indication for the production of triploid gametes. Second, if 2n eggs and 2n sperm are produced in the same population, tetraploid offspring can result. In water frogs, however, high proportions of 2n eggs and 2n sperm have been found in separate geographic areas: in northwestern European populations, the dominating gamete production pattern is diploid eggs and haploid sperm, whereas in some central European populations, it is haploid eggs and diploid sperm (Christiansen 2009; Pruvost et al. 2013a, 2015; Mikulíček et al. 2015). Thus, the very low occurrence of tetraploid water frogs observed this far can be explained by the low likelihood of 2n egg and 2n sperm encounters. Moreover, even tetraploidy does not guarantee successful chromosome segregation and gamete production, neither in asymmetric tetraploids (e.g. AAAB) nor in symmetrical ones (AABB). The only tetraploid LLRR hybrid male for which we, so far, know the gamete types produced haploid R sperm and a few diploid gametes of unknown genotypic composition (Pruvost et al. 2015). Thus, with high genetic diversity and great variability in origin, composition and gamete production patterns, P. esculentus populations across Europe hold the potential for various evolutionary trajectories. Which ones will be realized, only time can tell. The outcome may not only differ among geographic regions; it will also be highly dependent on how much room their human cohabitants decide to grant these remarkable animals.

Acknowledgements

We are greatly thankful to several people who contributed to this study in one way or another. We especially thank Maria Ogielska and her group, Mariusz Rybacki, Jacek Szymura, Dragica Salamon, Maja Cipot, Miklós Puky, Torsten Ohst, S.-D. Covaciu-Marcov and Jon Loman for their assistance with obtaining catching permits from their respective countries, their valuable help with collecting frogs, fruitful scientific discussions and their outstanding hospitality. Additional samples from Swedish and Danish populations were kindly provided by Martina Arioli and Christian Jakob. We also thank Erik Postma for sharing his expertise in genetic and statistical analvsis and Anna Lemkul who helped with data processing for the genealogical analysis. Robert Schreiber did the sequencing for one part of the mtDNA samples in Berlin. Torsten Ohst and Marie Doležálková provided helpful advice and the occasional troubleshooting during sequencing another part of the mtDNA samples at the University of Zurich. Three anonymous reviewers provided helpful comments and suggestions on earlier versions of the manuscript. Great thanks also go to our field assistants Irene Völlmy, Daniel Hollinger, Ursina Tobler, Julian Wild and Matej Kautman, who most enthusiastically helped with catching frogs and provided helpful and very pleasant company in all sorts of field situations. Lastly, we are grateful to pond owners, rangers, policemen or passers-by who approached us froggers mostly with kind acceptance and curiosity. For frog catching and sampling, we had permits from the respective countries. Tissue samples and/or DNA samples of the studied frogs are deposited in the 'Water Frog Tissue and DNA Collection' of the Zoological Museum of the University of Zürich. This work was supported from SNF grant No. 3100A0-120225/1 to Heinz-Ulrich Reyer. L. Choleva received support No. RVO 67985904 from the Academy of Sciences of the Czech Republic and No. 15-19947Y from the Czech Science Foundation (GACR).

References

- Abbott R, Albach D, Ansell S *et al.* (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–246.
- Akin C, Bilgin CC, Beerli P, Westaway R et al. (2010) Phylogeographic patterns of genetic diversity in eastern Mediterranean water frogs have been determined by geological processes and climate change in the Late Cenozoic. *Journal of Biogeography*, **37**, 2111–2124.
- Alves MJ, Coelho MM, Prospero MI, Collares-Pereira MJ (1999) Production of fertile unreduced sperm by hybrid males of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae): an alternative route to genome tetraploidization in unisexuals. *Genetics*, 151, 277–283.
- Alves MJ, Coelho MM, Collares-Pereira MJ (2001) Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica*, **111**, 375–385.

- Alves MJ, Gromicho M, Collares-Pereira MJ, Coelho MM (2004) Simultaneous production of triploid and haploid eggs by triploid Squalius alburnoides (Teleostei: Cyprinidae). Journal of Experimental Zoology, 301A, 552–558.
- Applied Biosystems (2004) *Genemapper*. Available at: http:// www.lifetechnologies.com/ch/en/home/brands/applied-bio systems.html
- Arioli M (2007) Reproductive patterns and population genetics in pure hybridogenetic water frog populations of Rana esculenta. PhD thesis, University of Zurich, Zurich, Switzerland.
- Arioli M, Jakob C, Reyer HU (2010) Genetic diversity in water frog hybrids (*Pelophylax esculentus*) varies with population structure and geographic location. *Molecular Ecology*, 19, 1814–1828.
- Arnold ML (1997) Natural Hybridization and Evolution. Oxford University Press, New York.
- Berger L, Berger WA (1994) Persistence of all-hybrid water frog populations (*Rana* kl. esculenta) in northern Germany. Genetica Polonica, 35, 73–80.
- Berger L, Roguski H (1978) Ploidy of progeny from different egg size classes of *Rana esculenta* L. *Folia Biologica (Krakow)*, 26, 231–248.
- Beukeboom LW, Vrijenhoek RC (1998) Evolutionary genetics and ecology of sperm-dependent parthenogenesis. *Journal of Evolutionary Biology*, **11**, 755–782.
- Birky CW Jr, Barraclough TG (2009) Asexual speciation. In: Lost Sex: The Evolutionary Biology of Parthenogenesis (eds Schön I, Martens K, van Dijk P. 201–216. Springer, Heidelberg.
- Bonnet E, Van de Peer Y (2002) zt: a software tool for simple and partial Mantel tests. *Journal of Statistical Software*, 7, 1–12.
- Borkin LJ, Korshunov AV, Lada GA et al. (2004) Mass occurrence of polyploid green frogs (Rana esculenta complex) in Eastern Ukraine. Russian Journal of Herpetology, 11, 194–213.
- Borkin LJ, Lada GA, Litvinchuk SN, Melnikov D, Rozanov YM (2006) The first record of mass triploidy in hybridogenetic green frog *Rana esculenta* in Russia (Rostov Oblasť). *Russian Journal of Herpetology*, **13**, 77–82.
- Brychta BH, Tunner HG (1994) Flow cytometric analysis of spermatogenesis in triploid *Rana esculenta*. Zoologica Poloniae, 39, 507.
- Choleva L, Janko K (2013) Rise and persistence of animal polyploidy: evolutionary constraints and potential. *Cytogenetic* and Genome Research, **140**, 151–170.
- Choleva L, Karel J, De Gelas K *et al.* (2012) Synthesis of clonality and polyploidy in vertebrate animals by hybridization between two sexual species. *Evolution*, **66**, 2191–2201.
- Christiansen DG (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 DG (2009) Gamete types, sex determination and stable equilibria of all-hybrid populations of diploid and triploid edible frogs (*Pelophylax esculentus*). *BMC Evolutionary Biology*, **9**, 135.
- Christiansen DG, Reyer H-U (2009) From clonal to sexual hybrids: genetic recombination via triploids in all-hybrid populations of water frogs. *Evolution*, **63**, 1754–1768.
- Christiansen DG, Reyer H-U (2011) Effects of geographic distance, sea barriers and habitat on the genetic structure and diversity of all-hybrid water frog populations. *Heredity*, **106**, 25–36.

- Coyne JA, Orr HA (2004) *Speciation*. Sinauer, Sunderland, Massachusetts.
- Cunha C, Doadrio I, Coelho MM (2008) Speciation towards tetraploidization after intermediate processes of non-sexual reproduction. *Philosophical Transaction of the Royal Society of London B: Biological Sciences*, **363**, 2921–2929.
- David CJ, Pandian TJ (2006) Maternal and paternal hybrid triploids of tetras. *Journal of Fish Biology*, **69**, 1102–1119.
- Dawley RM (1989) An introduction to unisexual vertebrates. In: Evolution and Ecology of Unisexual Vertebrates (eds Dawley RM, Bogart JP), pp. 1–18. New York State Museum, Albany, New York.
- Dowling TE, Secor CL (1997) The role of hybridization and introgression in the diversification of animals. *Annual Review* of Ecology and Systematics, **28**, 593–619.
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.
- Ebendal T (1979) Distribution, morphology and taxonomy of the Swedish green frogs (*Rana esculenta* complex). *Mitteilungen aus dem Zoologischen Museum in Berlin*, **55**, 143–152.
- Ersts PJ (2012) Geographic Distance Matrix Generator (version 1.2.3). Available at: http://biodiversityinformatics.amnh.org/ open_source/gdmg. American Museum of Natural History, Center for Biodiversity and Conservation.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.
- Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, 7, 574–578.
- Fankhauser G, Griffiths RB (1939) Induction of triploidy and haploidy in the newt, *Triturus viridescens*, by cold treatment of unsegmented eggs. *Proceedings of the National Academy of Sciences*, USA, 25, 233–238.
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783–791.
- Force A, Lynch M, Pickett FB (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, 151, 1531–1545.
- Garner TWJ, Gautschi B, Röthlisberger S, Reyer H-U (2000) A set of CA repeat microsatellite markers derived from the pool frog, *Rana lessonae*. *Molecular Ecology*, **9**, 2173–2175.
- Graf J-D, Polls Pelaz M (1989) Evolutionary genetics of the Rana esculenta complex. In: Evolution and Ecology of Unisexual Vertebrates (eds Dawley R, Bogart JP), pp. 289–301. New York State Museum, Albany, New York.
- Gregory TR, Mable BK (2005) Polyploidy in Animals. In: *Evolution of the Genome* (ed. Gregory TR), pp. 427–517. Elsevier, Amsterdam.
- Guex GD, Hotz H, Semlitsch RD (2002) Deleterious alleles and differential viability in progeny of natural hemiclonal frogs. *Evolution*, 56, 1036–1044.
- Günther R (1970) Der Karyotyp von *Rana ridibunda* Pall. und das Vorkommen von Triploidie bei *Rana esculenta* L. (Anura, Amphibia). *Biologisches Zentralblatt*, **89**, 327–342.
- Günther R (1975) Zum natürlichen Vorkommen und zur Morphologie triploider Teichfrösche "*Rana esculenta*" L., in der DDR (Anura, Ranidae). *Mitteilungen aus dem Zoologischen Museum in Berlin*, **51**, 146–158.

- Günther R (1977) Die Erythrozytengröße als Kriterium zur Unterscheidung diploider und triploider Teichfrösche, *Rana "esculenta"* L. (Anura). *Biologisches Zentralblatt*, **96**, 457–466.
- Günther R, Uzzell T, Berger L (1979) Inheritance patterns in triploid *Rana "esculenta"* (Amphibia, Salientia). *Mitteilungen aus dem Zoologischen Museum in Berlin*, **55**, 35–57.
- Halkett F, Plantegenest M, Prunier-Leterme N, Mieuzet L, Delmotte F, Simon JC (2005) Admixed sexual and facultatively asexual aphid lineages at mating sites. *Molecular Ecology*, **14**, 325–336.
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.
- Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.
- Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.
- Hoffmann A, Reyer H-U (2013) Genomic effects on advertisement call structure in diploid and triploid hybrid waterfrogs (Anura, *Pelophylax esculentus*). BMC Ecology, **13**, 47.
- Hofman S, Spolsky C, Uzzell T, Cogălniceanu D, Babik W, Szymura JM (2007) Phylogeography of the fire-bellied toads *Bombina*: independent Pleistocene histories inferred from mitochondrial genomes. *Molecular Ecology*, **16**, 2301–2316.
- Holloway AK, Cannatella DC, Gerhardt HC, Hillis DM (2006) Polyploids with different origins and ancestors form a single sexual polyploid species. *American Naturalist*, **167**, E88–E101.
- Holsbeek G, Mergeay J, Hotz H, Plötner J, Volckaert FA, De Meester L (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.
- Hotz H, Mancino G, Bucci-Innocenti S, Ragghianti M, Berger L, Uzzell T (1985) *Rana ridibunda* varies geographically in inducing clonal gametogenesis in interspecies hybrids. *Jour*nal of Experimental Biology, 236, 199–200.
- Hotz H, Beerli P, Spolsky C (1992) Mitochondrial DNA reveals formation of nonhybrid frogs by natural matings between hemiclonal hybrids. *Molecular Biology and Evolution*, **9**, 610– 620.
- Hotz H, Uzzell T, Guex G-D, Alpers D, Semlitsch RD, Beerli P (2001) Microsatellites: a tool for evolutionary genetic studies of western Palearctic water frogs. *Mitteilungen aus dem Museum für Naturkunde in Berlin, Zoologische Reihe*, 77, 43–50.
- Hotz H, Guex G-D, Beerli P, Semlitsch RD, Pruvost NBM (2008) Hemiclone diversity in the hybridogenetic frog *Rana* esculenta outside the area of clone formation: the view from protein electrophoresis. *Journal of Zoological Systematics and Evolutionary Research*, **46**, 56–62.
- Hotz H, Beerli P, Uzzell T *et al.* (2013) Balancing a cline by influx of migrants: a genetic transition in water frogs of eastern Greece. *Journal of Heredity*, **104**, 57–71.
- Jakob C, Arioli M, Reyer HU (2010) Ploidy composition in allhybrid frog populations in relation to ecological conditions. *Evolutionary Ecology Research*, **12**, 633–652.
- Janko K, Kotusz J, De Gelas K *et al.* (2012) Dynamic formation of asexual diploid and polyploid lineages: multilocus analysis of *Cobitis* reveals the mechanisms maintaining the diversity of clones. *PLoS One*, **7**, e45384.

- Kawamura T (1984) Polyploidy in amphibians. Zoological Science, 1, 1–15.
- Kearney M, Fujita MK, Ridenour J (2009) Lost sex in reptiles. In: *Lost Sex: The Evolutionary Biology of Parthenogenesis* (eds Schön I *et al.*), pp. 447–474. Springer, Heidelberg.
- Kondo Y, Kashiwagi A (2004) Experimentally induced autotetraploidy and allotetraploidy in two Japanese pond frogs. *Journal of Herpetology*, **38**, 381–392.
- Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, doi:10.1111/1755-0998.12387.
- Lamatsch DK, Stöck M (2009) Sperm dependent parthenogenesis and hybridogenesis in teleost fishes. In: Lost Sex: The Evolutionary Biology of Parthenogenesis (eds Schön I et al.), pp. 399–432. Springer, Heidelberg.
- Levin DA (2002) The Role of Chromosomal Change in Plant Evolution. Oxford University Press, Oxford.
- Little TJ, Hebert PDN (1997) Clonal diversity in high arctic ostracodes. *Journal of Evolutionary Biology*, **10**, 233–252.
- Mable BK (2004) Why polyploidy is rarer in animals than in plants: myths and mechanisms. *Biological Journal of the Linnean Society*, **82**, 453–466.
- Mable BK, Alexandrou MA, Taylor MI (2011) Genome duplication in amphibians and fish: an extended synthesis. *Journal of Zoology*, 284, 151–182.
- Madlung A (2013) Polyploidy and its effect on evolutionary success: old questions revisited with new tools. *Heredity*, **110**, 99–104.
- Mallet J (2007) Hybrid speciation. Nature, 446, 279-283.
- Mallet J (2008) Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **363**, 2971–2986.
- Mason AS, Pires JC (2015) Unreduced gametes: meiotic mishap or evolutionary mechanism? *Trends in Genetics*, **31**, 5–10.
- Mikulíček P, Kotlík P (2001) Two water frog populations from western Slovakia consisting of diploid females and diploid and triploid males of the hybridogenetic hybrid *Rana esculenta* (Anura, Ranidae). *Mitteilungen aus dem Museum für Naturkunde in Berlin, Zoologische Reihe,* **77**, 59–64.
- Mikulíček P, Kautman M, Demovič B, Demovic B (2014) When a clonal genome finds its way back to a sexual species: evidence from ongoing but rare introgression in the hybridogenetic water frog complex. *Journal of Evolutionary Biology*, **27**, 628–642.
- Mikulíček P, Kautman M, Kautmann J, Pruvost NBM (2015) Mode of hybridogenesis and habitat preferences influence population composition of water frogs (*Pelophylax esculentus* complex, Anura: Ranidae) in a region of sympatric occurrence (western Slovakia). *Journal of Zoological Systematics and Evolutionary Research*, **53**, 124–132.
- Moritz C, Brown WM, Densmore LD et al. (1989) Genetic diversity and the dynamics of hybrid parthenogenesis in *Cnemidophorus* (Teiidae) and *Heteronotia* (Gekkonidae). In: *Evolution and Ecology of Unisexual Vertebrates* (eds Dawley RM, Bogart JP), pp. 87–112. New York State Museum, Albany, New York.
- Negovetic S, Anholt BR, Semlitsch RD, Reyer HU (2001) Specific responses of sexual and hybridogenetic European waterfrog tadpoles to temperature. *Ecology*, 82, 766–774.

- Nei M (1978) Estimation of average heterozygosity and genetic distance for small numbers of individuals. *Genetics*, 89, 583– 590.
- Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*, **19**, 153–170.
- Ohst T (2008) Genetische Einflüsse allochthoner Wasserfrösche auf endemische Wasserfroschpopulationen (R. kl. esculenta Komplex). PhD thesis, Humboldt-Universität zu Berlin, Berlin.
- Otto SP, Whitton J (2000) Polyploid incidence and evolution. Annual Review of Genetics, 34, 401–437.
- Pagano A, Lesbarreres D, O'Hara R et al. (2008) Geographical and ecological distributions of frog hemiclones suggest occurrence of both 'General-Purpose Genotype' and 'Frozen Niche Variation' clones. Journal of Zoological Systematics and Evolutionary Research, 46, 162–168.
- Plötner J (2005) Die westpaläarktischen Wasserfrösche. Laurenti-Verlag, Bielefeld.
- Plötner J, Becker C, Plötner K (1994) Morphometric and DNA investigations into European water frogs (*Rana* kl. esculenta Synklepton (Anura, Ranidae)) from different population systems. Zeitschrift für Zoologische Systematik und Evolutionsforschung, **32**, 193–210.
- Plötner J, Uzzell T, Beerli P et al. (2008) Widespread unidirectional transfer of mitochondrial DNA: a case in western Palaearctic water frogs. Journal of Evolutionary Biology, 21, 668–681.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Pritchard JK, Wen X, Falush D (2010) Documentation for structure software: version 2.3. Available at: http://pritchardlab.stanford.edu/structure.html
- Pruvost NBM, Hoffmann A, Reyer H-U (2013a) Gamete production patterns, ploidy, and population genetics reveal evolutionary significant units in hybrid water frogs (*Pelophylax esculentus*). *Ecology and Evolution*, **3**, 2933–2946.
- Pruvost NBM, Hollinger D, Reyer H-U (2013b) Genotype-temperature interactions on larval performance shape population structure in hybridogenetic water frogs (*Pelophylax esculentus* complex). *Functional Ecology*, 27, 459–471.
- Pruvost NBM, Mikulíček P, Choleva L, Reyer HU (2015) Contrasting reproductive strategies of triploid hybrid males in vertebrate mating systems. *Journal of Evolutionary Biology*, 28, 189–204.
- Ptacek MB, Gerhardt HC, Sage RD (1994) Speciation by polyploidy in tree frogs: multiple origins of the tetraploid *Hyla* versicolor. Evolution, **31**, 721–736.
- Ramsey J, Schemske DW (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics*, **29**, 467–501.
- Rieseberg LH, Willis JH (2007) Plant speciation. Science, 317, 910–914.
- Schmeller D, Crivelli A, Veith M (2001a) Is triploidy indisputably determinable in hybridogenetic hybrids by planimetric analyses of erythrocytes? *Mitteilungen aus dem Museum für Naturkunde in Berlin, Zoologische Reihe*, 77, 71–77.
- Schmeller D, Seitz A, Crivelli A, Pagano A, Veith M (2001b) Inheritance in the water frog *Rana ridibunda* Pallas, 1771 – is it Mendelian or hemiclonal? *Mitteilungen aus dem Museum für Naturkunde in Berlin, Zoologische Reihe*, 77, 39–42.

- Schmidt DJ, Bond NR, Adams M, Hughes JM (2011) Cytonuclear evidence for hybridogenetic reproduction in natural populations of the Australian carp gudgeon (*Hypseleotris*: Eleotridae). *Molecular Ecology*, **20**, 3367–3380.
- Schmitt T (2009) Biogeographical and evolutionary importance of the European high mountain systems. *Frontiers in Zoology*, 6, 9.
- Schultz RJ (1969) Hybridization, unisexuality and polyploidy in the teleost *Poeciliopsis* (Poeciliidae) and other vertebrates. *The American Naturalist*, **103**, 605–619.
- Schwarz G (1978) Estimating the dimension of a model. *The Annals of Statistics*, **6**, 461–464.
- Seehausen O (2004) Hybridization and adaptive radiation. Trends in Ecology and Evolution, 19, 198–207.
- Sjögren P (1991) Genetic variation in relation to demography of peripheral pool frog populations (*Rana lessonae*). *Evolutionary Ecology*, 5, 248–271.
- Sjögren-Gulve P, Berg LM (1999) Allozyme variation as a demographic predictor at high latitudes: the moor frog and the pool frog at 60 degrees N. *Hereditas*, **130**, 317–323.
- Soltis D, Soltis P (1999) Polyploidy: origins of species and genome evolution. *Trends in Ecology and Evolution*, **14**, 348–352.
- Soltis P, Soltis D (2009) The role of hybridization in plant speciation. *Annual Review of Plant Biology*, **60**, 561–588.
- Som C, Reyer H-U (2006) Demography and evolution of pure hybridogenetic frog (*Rana esculenta*) populations. *Evolutionary Ecology Research*, 8, 1235–1248.
- Spolsky C, Uzzell T (1986) Evolutionary history of the hybridogenetic hybrid frog *Rana esculenta* as deduced from mtDNA analyses. *Molecular Biology and Evolution*, 3, 44–56.
- Stöck M, Lamatsch DK, Steinlein C et al. (2002) A bisexually reproduce all-triploid vertebrate. Nature Genetics, 30, 325– 328.
- Stöck M, Steinlein C, Lamatsch DK, Schartl M, Schmid M (2005) Multiple origins of tetraploid taxa in the Eurasian Bufo viridis subgroup. Genetica, 124, 255–272.
- Stöck M, Dufresnes C, Litvinchuk SN *et al.* (2012) Cryptic diversity among Western Palearctic tree frogs: postglacial range expansion, range limits, and secondary contacts of three European tree frog lineages (*Hyla arborea* group). *Molecular Phylogenetics and Evolution*, **65**, 1–9.
- SYSTAT[®] Software 11 (2004). Systat Inc., Richmond, California.
- Taberlet P, Fumagalli L, Wust-Saucy AG *et al.* (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.
- Takezaki N, Nei M, Tamura K (2014) POPTREEW: web version of POPTREE for constructing population trees from allele frequency data and computing other population statistics. *Molecular Biology and Evolution*, **31**, 1622–1624.
- Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution, 30, 2725–2729.
- Tunner HG (2000) Evidence for genomic imprinting in unisexual triploid hybrid frogs. *Amphibia-Reptilia*, **21**, 135–141.
- Tunner HG, Heppich-Tunner S (1992)A new population system of water frogs discovered in Hungary. In: *Proceedings of the*

6th Ordinary General Meeting of the Societas Europaea Herpetologica, 19–23 August 1991, pp. 453–460, Budapest, Hungary.

- Uzzell T, Berger L (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, Berger L, Günther R (1975) Diploid and triploid progeny from a diploid female of *Rana esculenta* (Amphibia Salientia). *Proceedings of the Academy of Natural Sciences of Philadelphia*, **127**, 81–91.
- Vorburger C (2001) Non-hybrid offspring from matings between hemiclonal hybrid waterfrogs suggest occasional recombination between clonal genomes. *Ecology Letters*, 4, 628–636.
- Vorburger C, Schmeller DS, Hotz H, Guex G-D, Reyer H-U (2009) Masked Damage: mutational load in hemiclonal water frogs. In: *Lost Sex: The Evolutionary Biology of Parthenogenesis* (eds Schön I *et al.*), pp. 433–446. Springer, Heidelberg.
- Vrijenhoek RC (2006) Polyploid Hybrids: multiple origins of a treefrog species. Current Biology, 16, R245–R247.
- Vrijenhoek RC, Dawley CJ, Cole CJ, Bogart JP (1989) A list of known unisexual vertebrates. In: *Evolution and Ecology of Unisexual Vertebrates* (eds Dawley RM, Bogart JP), pp. 19–24. New York State Museum, Albany, New York.
- Waters JM, Fraser CI, Hewitt GM (2013) Founder takes all: density-dependent processes structure biodiversity. *Trends in Ecology and Evolution*, 28, 78–85.
- Zeisset I, Beebee TJ (2001) Determination of biogeographical range: an application of molecular phylogeography to the European pool frog *Rana lessonae*. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **268**, 933–938.
- Zeisset I, Rowe G, Beebee TJC (2000) Polymerase chain reaction primers for microsatellite loci in the north European water frogs *Rana ridibunda* and *R. lessonae. Molecular Ecology*, **9**, 1173–1174.

Zielinski P, Dudek K, Stuglik MT, Liana M, Wiesław Babik W (2014) Single nucleotide polymorphism reveal genetic structuring of the Carpathian newt and provide evidence of interspecific gene flow in the nuclear genome. *PLoS One*, 9, e97431.

A.H. collected samples from several localities, analysed the data, interpreted the results and wrote the first version of the manuscript. J.P. contributed to mtDNAsequencing, statistical analysis and manuscript writing. N.B.M.P. helped in sampling, database management and interpretation of results. D.G.C. and P.M. contributed to the analyses of microsatellite markers and the revision of the manuscript. S.R. did most of the genetic analyses in the laboratory. L.C., P.M., I.S.-K., D.C., D.S. and S.M.-L. provided information about populations in their respective countries, obtained catching permits, helped with sampling and/or sent their own samples (some of them already partly analysed), and contributed to the revision of the manuscript. H.-U.R. designed and coordinated the study, obtained the funding and helped in collecting samples, analysing data and manuscript writing. All authors have read draft versions of the manuscript.

Data accessibility

The following data files are archived in the Dryad Digital Repository: doi:10.5061/dryad.nd1hg.