

**Phylogeography, evolutionary history, and effects of glaciations in a species**  
**(*Zootoca vivipara*) inhabiting multiple biogeographic regions**

**Short running title:** Biogeography of *Zootoca vivipara*

Jose L. Horreo<sup>1</sup>, Maria L. Pelaez<sup>1,2</sup>, Teresa Suárez<sup>3</sup>, Merel C. Breedveld<sup>1,3</sup>, Heulin B.<sup>4</sup>,  
Yann Surget-Groba<sup>5</sup>, Tuula A. Oksanen<sup>6</sup>, Patrick S. Fitze P. S.<sup>1\*</sup>

\*Corresponding author: patrick.fitze@mncn.csic.es

<sup>1</sup>Department of Biodiversity and Evolutionary Biology, Museo Nacional de Ciencias  
Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006 Madrid, Spain.

<sup>2</sup>Department of Cellular and Molecular Physiopathology, Centro de Investigaciones  
Biologicas (CSIC), Calle Ramiro de Maetzu 9, Madrid, E-28040, Spain.

<sup>3</sup>Animal Ecology, University of Potsdam, Maulbeerallee 1, 14469 Potsdam, Germany.

<sup>4</sup>Station Biologique, CNRS UMR 6553, Paimpont 35380, France.

<sup>5</sup>Institut des Sciences de la Forêt tempérée, Université du Québec en Outaouais, 58, rue  
principale Ripon, Québec J0V 1V0, Canada.

<sup>6</sup>University of Jyväskylä, Department of Biological and Environmental Science, P.O.  
Box 35, FI-40014 University of Jyväskylä, Finland.

**ACKNOWLEDGEMENTS**

We thank M. Urieta for help during sampling of *Z. vivipara*. J. L. Horreo was supported  
by two Spanish MINECO postdoc grants (FPDI-2013-16116 and IJCI-2015-23618). W.  
Mayer and M. L. A. Ceirans provided samples. Project funds were provided by the  
Swiss National Science Foundation (PPOOP3\_128375, PP00P3\_152929/1 to P.S.F.),  
the Spanish Ministry of Education and Science (CGL2008-01522, CGL2012-32459,

CGL2016-76918 to P.S.F.), and the Academy of Finland (grant no. 7108955 to T. Oksanen).

## **ABSTRACT**

**Aim.** During glaciations, the distribution of temperate species inhabiting the northern hemisphere generally contracts into southern refugia; and in boreo-alpine species of the northern hemisphere, expansion from northern refugia is the general rule. Little is known about the drivers explaining vast distributions of species inhabiting multiple biogeographic regions (major biogeographic regions defined by the European Environmental Agency). Here we investigate the fine-scale phylogeography and evolutionary history of the Eurasian common lizard (*Zootoca vivipara*), the terrestrial reptile with the world's widest and highest latitudinal distribution, that inhabits multiple biogeographic regions.

**Location** Eurasia

**Methods** We generated the largest molecular dataset to date of *Z. vivipara*, ran phylogenetic analyses, reconstructed its evolutionary history, determined the location of glacial refuges, and reconstructed ancestral biogeographic regions.

**Results** The phylogenetic analyses revealed a complex evolutionary history, driven by expansions and contractions of the distribution due to glacials and inter-glacials, and the colonization of new biogeographic regions by all lineages of *Z. vivipara*. Many glacial refugia were detected, most were located close to the southern limit of the last glacial maximum. Two subclades recolonized large areas covered by permafrost during the last glaciation: namely, Western and Northern Europe, and North-Eastern Europe and Asia.

**Main conclusions** In *Z. vivipara*, most of the glacial refugia were located in the South of their current distribution. Previous studies suggested the existence of northern

refuges, but the species' inability to overwinter on permafrost and the lack of genetic support suggest that the presence of a refugia in the north of the Alps is unlikely. This species currently inhabits boreo-alpine climates and retracted during previous glaciations into southern refugia, as temperate species. The fast expansion of two clades started from two distinct glacial refugia. These phylogeographic patterns were highly congruent with those of *Vipera berus*. Together they suggest that glacial retraction, the location of the refugia, and absence of competition may have promoted colonization.

**Keywords** ancestral area reconstruction, ancestral biogeographic region reconstruction, biogeography, glacial refuges, last glacial maxima, molecular diversity, phylogeny, post-glacial recolonization.

## INTRODUCTION

Current biogeographic patterns of many organisms are strongly determined by the past climatic, geographic and geologic history (e.g. Mittelbach et al., 2007). In many regions of the world, glaciations played a major role in shaping diversity (e.g. Fernández-Palacios et al., 2016). In the Northern hemisphere, glaciations typically led to south-northward contraction-expansion patterns (Hewitt, 1996), either due to the reduction of suitable habitat and the formation of cold stage refugia in the south during glaciation with subsequent expansion during interglacials, or vice versa (warm-stage refugia; Stewart et al., 2010). Species inhabiting temperate conditions frequently exhibit contraction during glaciation and subsequent recolonization during interglacials (Hewitt, 1996). Instead, boreo-alpine and steppe species are suggested to exhibit contraction during inter-glacials and, at least the former, recolonization during glacial periods (Kajtoch et al., 2016). Consequently, species distributions will contract into Southern, Northern or Montane, and extrazonal refugia, depending on the inhabited biogeographic region (Stewart et al., 2010; Kajtoch et al., 2016). However, not all species inhabit a single biogeographic region and thus predictions about the contraction-expansion patterns and the location of refugia may not be straight-forward. For example, the distribution of the common European viper (*Vipera berus*, Linnaeus, 1758), the world's most widely distributed terrestrial snake species, ranges from Scotland in the west to Pacific Russia in the Far East, and from Albania and Greece in the south to north of the Arctic Circle (Ursenbacher et al., 2006). It thus inhabits temperate, boreal, alpine, Atlantic, and continental climates (Ursenbacher et al., 2006). Potential refugia have been reconstructed in South-Central France (West of the Alps; Western subclade), in the lowland plains in the Balkans, on the Northern Italian Peninsula, in the northeast and east of the Carpathian Mountains (Carpathian subclade),

86 and in the North of the Alps (Ursenbacher et al., 2006). This points to refugia South and  
87 West of the glaciers (South-Central France, Balkans, Northern Italy), but as well in the  
88 North of the Alps and North of the Carpathians. The former refugia are in locations  
89 predicted for temperate zone species, while the latter are rather locations predicted for  
90 boreo-alpine species. Similarly, the Eurasian common lizard *Zootoca vivipara*  
91 (Lichtenstein, 1823), the terrestrial reptile with the world's widest distribution and the  
92 highest latitudinal distribution in the Northern hemisphere (Hikida, 2002), inhabits  
93 temperate, boreal, alpine, Atlantic, and continental climates (Surget-Groba et al., 2001).  
94 Potential refuges have been reconstructed in Southern France (North of Pyrenees),  
95 North-Western and Northern Spain (Surget-Groba et al., 2001; Mila et al., 2013), Italy  
96 (Surget-Groba et al. 2002) and the Carnian Alps (Heulin et al., 2000). This points to  
97 refugia South and West of the glaciers, but also North of the Pyrenees. The findings for  
98 both species suggest that the distribution of species can be limited by both glacial and  
99 interglacial periods and that responses of a single species to glacial and interglacial  
100 periods may be similar to those of species inhabiting temperate and boreo-alpine  
101 regions (Hewitt, 1996; Kajtoch et al., 2016). These different responses may exist if  
102 different lineages adapted to different climates. For example, if some lineages behave as  
103 boreo-alpine species (with expansions during glaciations and contractions during post-  
104 glacials), and others as temperate climate species (with expansion during post-glacials  
105 and contractions during glacials). However, for both *Z. vivipara* and *V. berus*, the  
106 sampling used in previous studies was limited (Ursenbacher *et al.*, 2006; Provan &  
107 Bennet, 2008). For instance, in *V. berus* the sampling was not sufficient to localize more  
108 precisely the putative Northern refugia (Ursenbacher et al., 2006). Consequently, fine-  
109 scale phylogenetic and phylogeographic knowledge from widely distributed species is  
110 required to understand whether species exhibit at the same time biogeographic patterns

of temperate zone species and of boreo-alpine species.

Here we study the phylogeography and location of refugia of the Eurasian common lizard (*Z. vivipara*), and investigate whether different lineages exhibit different responses to glacials and interglacials. Earlier studies detected six major genetic clades (Surget-Groba et al., 2001; Surget-Groba et al., 2002; Surget-Groba et al., 2006), two with oviparous (A and B) and four with viviparous (C, D, E and F) reproduction. The viviparous populations are widely distributed along the Palearctic area (from western Europe up to Sakhalin and Hokkaido Islands; Heulin & Guillaume, 1989) where their northern limit is closely associated with the southern limit of the permafrost (Berman et al., 2016). The oviparous populations are present only in two areas of its southern distribution: first in northern Spain and the Pyrenees (Heulin & Guillaume, 1989), and second, in northern Italy and the Carnian Alps (Heulin et al., 2000, Surget-Groba et al. 2002). Two subclades (B1 and B2) have been described by Surget-Groba et al. (2001) and more recently, regional studies with more intense sampling have provided evidence for additional well-supported subclades (Mila et al., 2013). This suggests not only that independent species may exist (Cornetti et al., 2015), but the current phylogeny and phylogeography are more complex than believed.

The current biogeographic hypotheses for *Z. vivipara* suggest: (i) the existence of latitudinal contractions/expansions in the Western (French and Spanish) oviparous populations (clade B) during Pleistocene glaciations with refugia in Southern France (North of the Pyrenees), in the south of the Pyrenees, and Northern Spain (i.e., South-Western Europe; Surget-Groba et al., 2001), (ii) an Italian refuge of the oviparous clade A (Surget-Groba et al. 2002), and that (iii) for clades C to F the location of potential refuges is unknown and (iv) clades D and E, respectively, may have exhibited post-

glacial northwards expansion to Northern Europe and post-glacial north-east expansion to North-Eastern Europe and Asia (Surget-Groba et al., 2001). Taken together this information indicates (i) that *Z. vivipara* exhibits biogeographic patterns typical for temperate-species, namely, Southern refuges during glaciations, and (ii), its occurrence in high northern latitudes (at least up to 66° N; Surget-Groba et al. 2006; Berman et al. 2016) and the refuge North of the Pyrenees suggests that some lineages exhibit biogeographic patterns typical for boreo-alpine species. However, sample size, genetic methods and markers used in previous studies do not allow to draw robust conclusions (Surget-Groba et al., 2001) and the reason why some, but not all clades exhibit enormous geographic distributions is unclear.

Here we generated the largest molecular dataset of *Z. vivipara* known to date. It covers almost the entire natural distribution and all known lineages, including fine scale sampling of areas with limited sampling and of areas that have not been sampled. Genetic and biogeographic analyses are used to test the following hypotheses. First, are the locations of *Z. vivipara*'s refugia in areas commensurate with those predicted by temperate and/or boreo-alpine species? Second, are contraction/expansion patterns associated with glacials/inter-glacials? Third, do different clades/lineages inhabit different biogeographic regions? Fourth, do different clades/lineages inhabit areas with different climates? Fifth, are the results congruent with the idea that a single species exhibits at the same time biogeographic patterns typical for temperate and typical for boreo-alpine species?

## **MATERIALS AND METHODS**

### **Used samples**

The 231 samples used in this study (Fig. 1) cover all described clades (clades A-F; Surget-Groba et al., 2001, 2006) of *Zootoca vivipara* and almost all the species distribution across Eurasia (specimen details, see Appendix S1 in Supporting Information). The number of samples per clade ranged between 24 and 63, and two outgroups were employed for the rooting of the phylogenetic tree, namely the common wall lizard (*Podarcis muralis*; Laurenti, 1768) and the western green lizard (*Lacerta bilineata*; Daudin, 1802). These outgroups were chosen because the genera are closely related to *Zootoca* and because both have been used in previous studies (Surget-Groba et al., 2001; Surget-Groba et al., 2002; Surget-Groba et al., 2006). DNA was extracted from lizard tissues using a commercial ChargeSwitch® gDNA Micro Tissue Kit (Invitrogen™, Carlsbad, USA) that provides better DNA-extraction quality compared to other methods (Horreo et al., 2015). Tail, skin sheds, or other tissues were obtained from collaborators (see Surget-Groba *et al.* 2006 and acknowledgments) and from own sampling under licences from Instituto Aragonés de Gestión Ambiental LCE/Mp 24/2012/426 and 500201/24/2013/12572; Service Nature Languedoc-Roussillon arrêtés N° 2014304-004; DREAL Aquitaine; Préfecture Pyrénées Atlantiques: Arrêtés N° 02/2006, 47/2007, 03/2008, 06/2009, 08/2010, 07/2012; Parque Nacional des Pyrénées 2012-55 & 2016\_8; Xunta de Galicia 071/2014; Generalitat de Catalunya; Asturias; Parque Nacional Picos de Europa; Bizkaiko Foru Aldundia Zk/N° 1079; University of Jyväskylä 49/29.11.2004 and 21/22.5.2006; Finnish Ministry of the Environment 3/5714/2003 and 8/5714/2006; County Administrative Board of Norrbotten 522-1341-05; Finnish Environment Institute KSU-2004-L-535/254).

## **Genetic markers**



In all samples, three nuclear (nZV1, nZV2 and nZV3; 1266 bp; Horreo et al., 2018) and three mitochondrial (ND2, CytB, and 16S rRNA; 1355 bp) gene fragments were amplified. Polymerase chain reaction (PCR) for gene amplification was done with 100 ng of template DNA in a total reaction volume of 25 µl employing a 5PRIME MasterMix Kit. PCRs cycles: initial denaturation (5 min) at 94°C; 35 cycles (30 s) at 94°C; annealing (30 s) at 59°C for the nuclear genes and at 53°C, 53°C and 50°C for the ND2, CytB and 16S genes; extension (90 s) at 72°C; and a final extension at 72°C for 5 min. Five of the six genes amplified in all 231 *Z. vivipara* and the two outgroups, and ND2 amplified in 201 *Z. vivipara* and the outgroups. To amplify the remaining 30 *Z. vivipara* samples, a new reverse primer was designed (ND2-H; 5'-3' sequence: GATGGAAGCCCACTGGTTAG) and amplified at 55°C together with the Met F6 primer of Macey *et al.*, (1992). Heterozygote sites were coded with IUPAC codes. Fragments were concatenated following a total evidence approach (Kluge, 1998), leading to a dataset of 2621 basepairs (bp) of 231 *Z. vivipara* and two outgroups. Alignments will be made available upon request.

### **Phylogenetic and molecular clock analyses**

For nuclear DNA, haplotypes of heterozygotes were inferred employing PHASE 2.1 (Stephens *et al.*, 2001), as well as SeqPHASE for creating input and output files (Flot, 2010). Recombination was tested on them with DNAsp (Librado & Rozas, 2009) under the four-gamete test methodology (Hudson & Kaplan, 1985). This software was also employed to estimate the genetic variability of the different clades, measured as the number of variable sites, the haplotype diversity and the nucleotide diversity, in both nuclear (*n*) and mitochondrial (*mt*) DNA (in the case of nuclear DNA, each gene was analysed separately).

King & Lee (2015) showed that Bayesian analyses need to account for rate heterogeneity, because not accounting for variable evolutionary rates can cause highly anomalous and even incorrect results. Consequently, Bayesian inference with a Random Local Clock (RLC), implemented in BEAST v.2.3.1 (Bouckaert et al., 2014) was used. Birth-Death model evolution and 100 million MCMC chains were run. jModelTest v.2 (Posada, 2008) with the Akaike inference criteria (AIC) was used to evaluate different models of nucleotide substitution. The GTR model fitted best in *mt* DNA, *n* DNA, and in the concatenated dataset. Consequently, GTR with no partitions was used for all analyses. Phylogenetic Maximum Likelihood analyses were done with MEGA v.7 (Kumar et al., 2016), using the GTR substitution model with 100 bootstrap replications. Molecular calibration of the tree was done using the dating of the node between the oviparous clade A of *Z. vivipara* and the other clades (4.5 Mya ago; 95% confidence interval (CI) = 6.1-2.6 Mya; Cornetti et al., 2014), since the estimate of this node was based on precise geological calibration (Cornetti et al., 2014), by means of the divergence between the Peloponnese wall lizard (*Podarcis peloponnesiaca*; Bibron & Bory, 1833) and the Cretan wall lizard (*P. cretensis*; Wettstein, 1952) when Crete became isolated from Peloponnese during the Messinian (5.2+/-0.1 Mya; Beerli et al., 1996).

#### **Phylogeographic and Biogeographic inferences**

Ancestral area reconstruction was done with RASP v.3.2 software (Yu et al., 2015), which implements the two most widely and accepted phylogeographic reconstruction methods. Both methods were used: a) the Lagrange Dispersal-Extinction-Cladogenesis method (DEC; Ree & Smith, 2008) and b) the Statistical-Dispersal Vicariance Analysis with options allow extinctions and allow reconstruction (Maximum = 100) checked (S-

DIVA; Yu et al., 2010), and individual and combined statistics were derived. Four major phylogeographic areas were defined representing the current phylogeographic distribution of *Z. vivipara*'s clades: Italy (including the areas covered by clade A), Northern and Central Europe and southern Balkan (covered by the clades C, E, and F; hereafter referred to as European phylogeographic area), Asia (clade D), and South-Western Europe (clade B), and nodes were allowed to be attributed to a maximum of two phylogeographic areas. The same approach was used to reconstruct ancestral biogeographic regions. The biogeographic region (EEA, 2012) in which a specimen was captured was determined and nodes were allowed to be attributed to a maximum of two biogeographic regions. Refugia during last glacial maximum (LGM: 21,000 years ago) were inferred based on the current clade distributions and locations where the species could not survive during the last glaciation, due to continental ice sheet cover and permafrost.

## RESULTS

### Genetic variability

No recombination was found within nuclear genes (minimum number of recombination events,  $R_m$ , was zero in all cases). Haplotype diversity ranged between 0.056 (clade B, nZV1) and 0.913 (clade C, *mt* DNA), and nucleotide diversity between 0.001 (clade B, nZV1 and nZV2, clade C, nZV3) and 0.053 (clade D, *mt* DNA; Table 1). The TCS haplotype network (Fig. 2) showed a clear separation of all extant clades (A, B, C, D, E, F) by several mutation steps, and star-like patterns in approximately half of the haplotype groups (*i.e.*, subclades).

### Phylogenetic analyses

Most ancestral nodes of the *Z. vivipara* clades (A, B, D, E and F) had a posterior probability (pp) of  $\geq 0.99$  (Fig. 3; Appendix S2, S3). Clade B, D, E, and F contained statistically well-supported and geographically separated subclades (Table 2) and clade A contained two subclades with moderate statistical support (A1, A2, Fig. 1). Clade B consisted of three subclades with pp = 1 and one sample with node support of pp = 0.85 (OF25, Fig. 3). Individuals belonging to the NC Spain and the NE Spain subclade (Mila et al., 2013) grouped together in the same subclade (B3+B4), exhibited monophyly, and were not statistically supported (pp = 0.74, Appendix S2). Subclades B2 and B1 included individuals belonging to the Southern France (Mila et al., 2013; B2 in Surget-Groba et al. 2006) and NW Spain subclade (Mila et al., 2013; B1 in Surget-Groba et al. 2006; note: all except one individual of B1, belong to the NW Spain subclade). Clade D consisted of three statistically well-supported subclades (pp = 1): D1, D2, and D3. Two samples (VRO1B and VRO5D) belonging to clade D were isolated from the rest (pp = 0.89 and 1) by > 10 mutational steps (Fig. 2), and both stem from two distinct geographic locations (Western Romanian Carpathians and Southern Carpathians) represented by only one sample (Fig. 3), suggesting that additional subclades may exist. Clade E consisted of two subclades with pp = 1: E1 and E2. Clade F consisted of two subclades with pp = 1: F1 and F2 (Fig. 1). Within clade C, no evidence for subclades existed. The molecular dating is shown in Fig. 3, and the age of ancestral nodes in Appendix S3. Maximum Likelihood phylogenetic tree (Appendix S4) exhibited the same clade topology as the Bayesian trees.

## Phylogeography

Over large areas the distribution of the subclades is allopatric. A1 inhabits NW Italy and A2 NE Italy (Fig. 1). B1 inhabits NW Spain, B2, Southern France (North of the

283 Pyrenees), B3 and B4 NC Spain and the NE Spain, respectively (Mila *et al.*, 2013). D1  
284 inhabits the North and East of Eurasia (Northern Sweden, Finland, Latvia, Lithuania,  
285 Belarus, Ukraine, Russia including Altai and Tuva Regions, and Mongolia, Table 2),  
286 D2 the Central East-Carpathian Mountains (Romania); and D3 the NW of and the  
287 central Inner East-Carpathian Mountains (Theiss region in Hungary, Romania,  
288 Slovakia, Ukraine; Table 2). VRO1B and VRO5D, belonging to clade D, originated  
289 from the Southern and Western Romanian Carpathians (Romania), respectively. E1  
290 inhabited Western, Central and Northern Europe, and the southern Balkans, and E2 the  
291 North of the East-Carpathian Mountains (frontier between Slovakia, Ukraine and  
292 Hungary), Eastern Poland, and Kaliningrad Oblast (Table 2). F1 is mainly distributed in  
293 SW Austria and F2 mainly in SE Austria (Fig. 1), while clade C inhabits an area south  
294 of clade E and north of clade F in central Austria.

295       Geographic overlap among clades existed in two areas (Fig. 1): in the Eastern  
296 Alps among clades A, C, E, and F (in Austria, NE Italia, and Northern Slovenia), and in  
297 the Central East-Carpathian Mountains between clades D and E (in Northern Romania,  
298 North-Eastern Hungary, and Eastern Slovakia). In the Central European Contact zone  
299 geographic overlap existed among clades A and E (in NE Italy, and Kärnten), among A  
300 and F (in Northern Slovenia, Kärnten, and SW Steiermark), among E and F (in East  
301 Tirol and Kärnten), and among C and F (in Steiermark). In subclades, evidence for  
302 geographic overlap existed between subclades F1 and F2 (in Southern Austria, Northern  
303 Slovenia and Hungary), between subclades B2 and B3+B4, and to a lesser extent  
304 between the subclades A1 and A2 (in Northern Italy and in Southern-central Austria).

305       Two subclades exhibit a very large geographic distribution: subclade E1 prevails  
306 from the Alps to Scandinavia and from Ireland to Poland and it also inhabits the  
307 Southern Balkans (Fig. 1), and subclade D1 prevails from Southern Ukraine to Northern

Scandinavia and from Lithuania/Latvia and Ukraine to Mongolia and Japan. The distribution of the other clades and subclades was much smaller and some occupy areas smaller than 20,000 km<sup>2</sup> (clade C and subclades B1, B2, B3+B4, D3, F1 and probably D2).

### **Biogeography**

The ancestral area reconstruction suggested that the ancestor of *Z. vivipara* lived in the Italian/European phylogeographic area. Clade A remained in this area, while the ancestor of clades B to F was reconstructed in the European phylogeographic area. The ancestor of clades E and D colonized the Asian phylogeographic area, where clade D diversified. Clade E remained in the European and Asian phylogeographic area until today. The ancestor of clades B, C and F was reconstructed with high probability in the European phylogeographic area, from where the ancestor of clade B and C colonized the Southwestern European area (Fig. 3). Later on the ancestor of clade C remained only in the European and the ancestor of clade B diversified in the Southwestern European phylogeographic area.

The reconstruction of the ancestral biogeographic regions shows that the ancestor of *Zootoca* inhabited the alpine biogeographic region (Fig. 3). Clade A and the ancestors of the other clades remained in the alpine region with the exception of the ancestor of clade D that colonized the continental biogeographic region (Fig. 3). Within clades, the ancestor of B1 and B3+B4 colonized the Atlantic, the ancestor of F1 the pannonian, the ancestor of D1 the boreal, and the ancestor of D3 the pannonian biogeographic region. Seven subclades recently colonized one (B3+B4, B2, C, F2, E2, D1) to three (E1) new biogeographic regions. Only two subclades changed

biogeographic region (B1, D1), all others still inhabit the alpine region, and D1 and E1 inhabit 3 and 4 biogeographic regions, respectively.

## **DISCUSSION**

### **Genetic structure**

The phylogenetic analyses unravelled six major clades (Fig. 3). These clades coincided with the major groups of the TCS haplotype network (Fig. 2). The positions of the clades in the phylogenetic tree (Fig. 3) were identical to those of an earlier phylogenetic maximum parsimony tree based on mitochondrial DNA (Surget-Groba et al., 2006), and they differed from the positions of another tree based on mitochondrial DNA and a tree based on nuclear DNA (Cornetti et al., 2014). Our phylogenetic Bayesian tree included more taxa and more genetic information than the previous ones, and all but one of the nodes grouping different clades had high statistical support ( $pp \geq 0.99$ , Fig. 3). Several subclades (Fig. 3, Table 2) were detected and they correspond to separated groups in the TCS haplotype network (Fig. 2; exception subclades of clade A). Our analyses provide robust evidence for 9 and moderate evidence for 2 subclades (Fig. 3, Table 2), and they suggest that at least another 2 subclades (Western Romanian Carpathians, and Southern Carpathians) may exist. However, their validity needs confirmation with more extensive sampling.

### **Current geographic distribution**

Clades and subclades exhibit current allopatry of over large areas (Fig. 1), with strong geographic overlap of clades only occurring in the Central European Contact Zone located in the East of the Eastern Alps (mainly in Kärnten, Steiermark, Northern Slovenia, and North-Eastern Italy; Fig. 1). Moderate overlap of 3 subclades existed in

the Pyrenees and in the Carpathian Mountains and no overlap existed among subclades of D and E, and between B1 and the other subclades of B. However, given the close geographic proximity of D2, D3 and E2, more precise fine-scale sampling is required to clarify whether geographic overlap may exist at a smaller geographic scale. Secondary contacts among clades or subclades have been described in the Pyrenees (subclades B2 and B3+B4; Mila et al., 2013) and in the Central European contact zone (clades A and D; Lindtke et al., 2010). In this study, populations of the two contact locations were not included and our study provided no evidence for clade mixing outside these two contact locations, since in the TCS network (Fig. 2) clades were clearly separated by many mutational steps and no intermediate haplotypes existed. The strong genetic structure and the described mixing in narrow contact zones (Lindtke et al., 2010; Mila et al., 2013) suggest that the observed contacts are rather recent, what points to future clade mixing due to current climate change. In contrast to these contacts, a study investigating a contact zone between clade A and E located in Northern Italy (ca. 270 km east of the contact zone described by Lindtke et al., 2010) detected no introgression and hypothesized that reproductive isolation may exist (Cornetti et al. 2014). The absence of clade mixing over most of the distribution and the fact that in the Central European contact zone the colonization of the different locations is difficult to explain without contacts among clades, is in agreement with reproductive isolation. However, detailed fine-scale sampling and experimental studies are required to provide robust evidence for or against these hypotheses.

### **Evolutionary history and refugia**

Phylogenetic analyses and molecular dating indicate that the oviparous clade A and the other *Z. vivipara* clades split during the Pliocene 4.4 (4.2-4.6; 95% CI, Appendix S3)



Mya, and the ancestral area reconstruction (Fig. 3) suggested that the origin of *Z. vivipara* was in the Italian or the European phylogeographic area (Fig. 4A). Given that earlier diverging genera (*Phoenicolacerta*, *Atlantolacerta*, *Gallotia*; Pyron et al., 2013) inhabit the middle east and Cyprus (*Phoenicolacerta*), Northern Africa (*Atlantolacerta*), and the Canary Islands (*Gallotia*), the most likely scenario is that the ancestor of *Z. vivipara* inhabited the Italian or the south of the European phylogeographic area (Fig. 4A) and that the ancestor of the clades C, F, and B moved Northwards (most likely through the Vienna Bassin). The ancestral area reconstruction and the high haplotype diversity in the Carpathians (Figs. 2 and 3), suggest that the common ancestor of clade D and E colonized the Carpathians, where diversification into clade D and E happened (Fig. 4A), and from where clade D (i.e. subclade D1) and E (i.e. subclade E1) colonized Asia and Northern to Western Europe (see below), respectively. The ancestor of clade B colonized South Western Europe, where differentiation into B2 and the other subclades happened 1.6 (1.2-2.0; 95% CI) Mya (Fig. 3; Appendix S3).

The divergence of clades B to E occurred between 2.7 (2.2-3.1; 95% CI) and 2.0 (1.6-2.4; 95% CI) Mya (Fig. 3), at the Pliocene / Pleistocene boundary, when major climatic changes existed that led to habitat fragmentation (Bennet, 1990), segregation into refugia and diversification of several ectothermic species in the Mediterranean region and in the Carpathian Mountains (e.g. Ursenbacher et al., 2006). In contrast, the evolution of the subclades is younger and happened between 1.6 Mya (1.2-2.0; 95% CI) and 0.8 Mya (0.5-1.0; 95% CI) (see Appendix S3), most likely as a consequence of glaciations and diversification due to genetic drift in different refugia (Fig. 4B).

Star-like haplotype networks, typical for fast population expansion (Hewitt, 1996), existed in clade A and in subclades B2, B3+B4, F2, D1 and E1. Moreover, two subclades (D1 and E1) exhibit vast current distributions in areas covered by ice during

the last glacial maximum (LGM: 21,000 years ago, Fig. 4). These patterns are congruent with glacial refugia and postglacial geographic expansion in Southern France (B2; Surget-Groba et al., 2001), North-East and North-Central Spain (B3+B4), and Northern Italy/Slovenia (Surget-Groba et al., 2002). This points to the typical geographic south-northward contraction-expansion due to glaciations and inter-glacials (Hewitt, 1996) in B and A, and to a refugia of clade B2 in the North of the glacial shield covering the Pyrenees, and subsequent southward expansion (Fig. 4B). E1 and D1 exhibited longitudinal expansion as well as northward expansion (Fig. 4). The former colonizing Western and Northern Europe, as well as Southern Europe (Bulgaria, Montenegro, Serbia) and the later colonizing Eastern Europe and Asia as well as North-Eastern Europe and Northern Asia. The ancestral area reconstruction, haplotype network, and current distribution (Fig. 1, 2 and 3) suggest that subclade D1 had a refugia in the Asian region, most likely close to the black sea given the extent of the glaciers and the permafrost during the last glacials (see below; Fig. 3B; Vandenberghe et al., 2014). The location of the refugia of subclade E1 was reconstructed in the European or Asian area (Fig. 3), and the genetic results do not allow to draw more precise conclusions about its location.

Subclade E1 (Fig. 2) inhabits two isolated geographic areas (Northwestern Europe and the Southern half of the Balkans). The phylogenetic analyses unravelled that two haplotypes prevailed exclusively on the Balkans (one in the Pirin Mountains, and the other one in Montenegro, Serbia and NW Bulgaria; Fig. 2). Moreover, no evidence for the existence of subclades or a monophyletic origin of the Balkan specimens existed and the TCS network showed that the genetically closest haplotypes located in Northern Europe differed by one base pair (Fig. 2). This points to the existence of historic connections between the two geographic areas (Heulin et al., 2011), allowing the

colonization of the Balkans from the north and subsequent northward migration (Figs. 2 and 4B).

While several subclades and clades suggest fast colonization patterns, others exhibit high diversity, no star-like patterns, and many mutational steps among haplotypes (Fig. 2, clade C, subclades D3, E2, F1), suggesting that they diverged in different subrefugia, i.e. refugia within refugia (e.g., Hewitt, 1996), from where no important expansion happened. Clade F consisted of two statistically supported subclades (F1 and F2), and in clade C no clusters existed, suggesting two potential refugia of clade F and one of clade C, all of them most likely located in the Pannonian/Vienna basin (Fig. 4B).

### **Biogeographic implications**

The results of the phylogeographic analyses point to refuges located in Western Europe (North, South, and West to the Pyrenees), Central Europe (most likely in the Pannonian/Vienna Basin), Eastern Europe (East, West, and probably as well South of the Carpathian Mountains), North to the black Sea, and in Southern Europe (Northern Italy). No explicit evidence existed for refuges North of the Alps, but the precise location of the refuge of the ancestor of subclade E1 could not be reconstructed and the ancestral area reconstruction points to the European or Asian area and thus to a refuge north, south, or east of the Alps.

The existence of permafrost even during Würm (115,000 – 10,000 years ago) down to the 46° N latitude (Fig. 4B), is incongruent with northern refuges, since *Z. vivipara* cannot endure temperatures in hibernacula (5 - 20 cm below ground) below -10°C and since permafrost temperatures at this depth are frequently below -15°C (Berman et al., 2016). More likely and more congruent with the TCS network is a

refuge South-East or in the East of the Alps, colonization of the Balkan and subsequent colonization of Northern and Western Europe (Figs. 2 and 4). The fast and recent geographic expansion of subclades E1 and D1 is in line with the northward-shift of the southern permafrost limit and the rapid colonization of newly available habitat. Southern refuges of clade C and F in the Pannonian/Vienna Basin, of clade A in Northern Italy, and of D1 in the north of the black sea, is in line with permafrost forcing *Z. vivipara* into areas south to the southern permafrost limit during the last glaciations (Fig. 4B). The split of clade B from the rest of *Z. vivipara*, its location in the Pyrenees, and the absence in the rest of Europe of individuals belonging to this clade, could be explained by early Pleistocene glaciations during which populations of *Z. vivipara* located North of the Alps may have gone extinct, thus disconnecting SW European populations from the rest of the distribution of *Z. vivipara*. Finally, the haplotype network (Fig. 2) and the current geographic distributions (Fig. 1) also suggest the existence of current admixture among subclades of clades F and A, most likely as a consequence of postglacial colonization of areas previously covered by permafrost, in line with interbreeding observed between clades A and E in the central European contact zone (Lindtke et al. 2010), and between subclades of the clade B in the Pyrenean suture zone (Mila et al. 2013).

All detected refugia were south of the southern permafrost limit of the last glaciation (Fig. 4B), and thus *Z. vivipara*'s refuges are in locations predicted for temperate zone species (Hewitt, 1996). This suggests that species inhabiting the boreo-alpine region do not necessarily exhibit Northern refuges. Interestingly, the reconstruction of ancestral biogeographic region shows that *Z. vivipara* originally inhabited the alpine region and that all 13 subclades colonized new biogeographic regions in the last 0.9 Mya (Fig. 3), suggesting adaptation to different climates. The two

most widely distributed subclades (D1, E1) colonized the continental and boreal region (Fig. 3) after glacial retraction and E1 also colonized the Atlantic region, *i.e.* the biogeographic regions covering the biggest part of their current distribution. Since in the other subclades the colonization of new biogeographic regions happened approximately at the same time (Fig. 3) and since none of the biogeographic regions mainly inhabited by D1 or E1 is exclusively inhabited by one subclade (the continental biogeographic region is inhabited by: A1, A2, C, D1, D2, E1; boreal by D1, E1; Atlantic, by B1, B2, B3+B4, E1; Table 2), adaptation of D1 and E1 to different climates cannot explain their success.

The reconstructed refugia of D1 and E1 were in the East and North of the refugia of the other clades/subclades (Fig. 4B) and thus at the forefront where new habitat became available due to glacial retraction. This suggests that the refugias' location may have allowed for colonization with low or inexistent competition with other clades/subclades, potentially explaining the large and recent expansion of the two clades. This idea is supported by the lack of large geographic expansion of the clades/subclades that retracted into refuges more south- or more west-wards (A, B, C, F, D2, D3). These clades/subclades colonized the same biogeographic regions approximately at the same time as D1 and E1 (Fig. 3) and in contrast to D1 and E1, their refugia were in close geographic proximity. This suggests that they may have been exposed to intense competition with other clades (see also Central European contact zone, Fig. 1), what may have hindered the colonization of newly available habitat.

## **General conclusions**

The European common lizard (*Z. vivipara*) exhibits a complex evolutionary history that was mainly driven by glacial expansion and retraction patterns. The biogeographic

patterns are congruent with those observed in pure temperate zone species and the current distribution is very similar to the distribution of *V. berus*, a terrestrial snake species that preys on *Z. vivipara* (e.g. Sebela, 1989) and inhabits the same biogeographic regions. Detected refugia were in similar locations as in *Z. vivipara* (Ursenbacher et al., 2006) and a northern refuge of a subclade of *V. berus* has been proposed. However, *V. berus* has been classified as non-freeze tolerant (Andersson & Johansson, 2001), suggesting that in both species the subclade with potential northern refugia may have had a refugia south of the southern permafrost limit. In both species, the clade for which a northern refuge has been suggested colonized Sweden and Norway, and similarly, in both species the same clade that colonized Asia, also colonized Finland and was present in the Carpathian Mountains. In both species, the glacial refuges of the two clades with rapid geographic expansion have been reconstructed in the same geographic locations. This points to similar evolutionary patterns, potentially due to the colonization of areas previously covered by permafrost. Colonization of areas previously covered by permafrost also resulted in similar colonization patterns outside non-avian sauropsids, for example in the bank vole *Myodes glareolus* (Deffontaine et al., 2005). *Zootoca vivipara* and *V. berus* inhabit quite similar ecological niches, and the ecological niches of all other European terrestrial Sauropsids include either more thermophile habitats (the largest amount of species), or are restricted to narrow alpine areas (e.g. *Iberolacerta*, and *Vipera walser*), suggesting that only species with relatively good cold tolerance and inhabiting an ample thermal niche may benefit from glacial retraction and colonize vast areas. In both species, the refuge in the East of the Carpathian Mountains is the most eastward, and the refuge of the other fast expanding subclade (E1, Central European subclade; Ursenbacher et al., 2006) the most northward glacial refuge. This suggests that glacial

retraction, the location of the refugia, and absence of competition, may have been important for fast geographic expansion and thus for colonization success of both species. The results further show that although inhabiting the boreo-alpine region, both species and all their subclades/lineages exhibit responses to glaciations as do temperate zone species, suggesting that adaptations leading to real boreo-alpine behaviour (southward expansion from Northern refugia) may take many million years to evolve.

## REFERENCES

- EEA (2012). *European Environmental Agency Report No5 Biogeographic regions in Europe*. ISSN 1725-9177.
- Andersson, S. & Johansson, L. (2001). Cold hardiness in the boreal adder, *Vipera berus*. *Cryo Letters*, 22, 151-156.
- Beerli, P., Hotz, H., & Uzzell, T. (1996). Geologically dated sea barriers calibrate an average protein clock in water frogs of the Aegean region. *Evolution*, 50, 1676-1687.
- Bennet, K.D. (1990). Milankovitch cycles and their effects on species in ecological and evolutionary time. *Palaeobiology*, 16, 11-21.
- Berman, D.L., Bulakhova, N.A., Alfimov, A.V., & Meshcheryakova, E.N. (2016). How the most northern lizard, *Zootoca vivipara*, overwinters in Siberia. *Polar Biology*, 39, 2411-2425.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., & Drummond, A.J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, 10, e1003537.

555 Clary, D.O., & Wolstenholme, D.R. (1985). The mitochondrial DNA molecule of  
556 *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code.  
557 *Journal of Molecular Evolution*, 22, 252-271.

558 Cornetti, L., Ficetola, G.F., Hoban, S., & Vernesi, C. (2015). Genetic and ecological  
559 data reveal species boundaries between viviparous and oviparous lizard lineages.  
560 *Heredity*, 115, 517-526.

561 Cornetti, L., Menegon, M., Giovine, G., Heulin, B., & Vernesi, C. (2014).  
562 Mitochondrial and nuclear DNA survey of *Zootoca vivipara* across the eastern  
563 italian Alps: evolutionary relationships, historical demography and conservation  
564 implications. *PLOS ONE*, 9, e85912.

565 Deffontaine, V., Libois, R., Kotlik, P., Sommer, R., Nieberding, C., Paradis, E., Searle,  
566 J.B., & Michaux, J.R. (2005). Beyond the Mediterranean peninsulas: evidence  
567 of central European glacial refugia for a temperate forest mammal species, the  
568 bank vole (*Clethrionomys glareolus*). *Molecular Ecology*, 14, 1727-1739.

569 Fernández-Palacios, J.M., Rijdsdijk, K.F., Norder, S.J., Otto, R., Nascimiento, L.  
570 Fernández-Lugo, S.F., Tjorve, E., & Whittaker, R.J. (2016). Towards a glacial-  
571 sensitive model of island biogeography. *Global Ecology and Biogeography*, 25,  
572 817-830.

573 Flot, J.F. (2010). SeqPHASE: a web tool for interconverting PHASE input/output files  
574 and FASTA sequence alignments. *Molecular Ecology Ressources*, 10, 162-166.

575 Heled, J., & Drummond, A.J. (2008). Bayesian inference of population size history  
576 from multiple loci. *BMC Evolutionary Biology*, 8, 289.

577 Heulin, B., & Guillaume, C.P. (1989). Extension géographique des populations ovipares  
578 de *Lacerta vivipara*. *Revue d'Ecologie*, 44, 39-45.



579 Heulin, B., Guillaume, C.P., Vogrin, N. ,& Surget-Groba, Y. (2000). Further evidence  
 580 of the existence of oviparous populations of *Lacerta (Zootoca) vivipara* in the  
 581 NW of the Balkan Peninsula. *C. R. Acad. Sci. Paris Ser. III*, 323, 461-468.  
 582 Heulin, B., Ghiemi, S., Vogrin, N., Surget-Groba, Y., & Guillaume, C.P. (2002).  
 583 Variation in eggshell characteristics and intrauterine egg retention between two  
 584 oviparous clades of the lizard *Lacerta vivipara*: insight into the oviparity-  
 585 viviparity continuum in squamates. *Journal of Morphology*, 252, 255-262.  
 586 Heulin, B., Surget-Groba, Y., Sinervo, B., Miles, D., & Guiller, A. (2011). Dynamics of  
 587 haplogroup frequencies and survival rates in a contact zone of two mtDNA  
 588 lineages of the lizard *Ecography*, 34, 436-447.  
 589 Hewitt, G.M. (1996). Some genetic consequences of ice ages, and their role in  
 590 divergence and speciation. *Biological Journal of the Linnean Society*, 58, 247-  
 591 276.  
 592 Hikida, H. (2002). *Natural history of the reptiles*, Tokyo.  
 593 Horreo, J.L., Peláez, M.L. & Fitze, P.S. (2015). Skin sheds as a useful DNA source for  
 594 lizard conservation. *Phyllomedusa*, 14, 73-77.  
 595 Horreo, J.L., Peláez, M.L., Suárez, T., & Fitze, P.S. (2018). Development and  
 596 characterization of 79 nuclear markers amplifying in viviparous and oviparous  
 597 clades of the European common lizard. *Genetica*, 146, 115-121.  
 598 Hudson, R.R., & Kaplan, N.L. (1985). Statistical properties of the number of  
 599 recombination events in the history of a sample of DNA sequences. *Genetics*,  
 600 111, 147-164.  
 601 Kajtoch, L., Cieslak, E., Varga, Z., Paul, W., Mazur, M.A., Sramkó, G., & Kubisz, D.  
 602 (2016). Phylogeographic patterns of steppe species in Eastern Central Europe: a

603 review and the implications for conservation. *Biodiversity and Conservation*, 25,  
604 2309-2339.

605 King, B., & Lee, M.S.Y. (2015). Ancestral state reconstruction, rate heterogeneity, and  
606 the evolution of reptile viviparity. *Systematic Biology*, 64, 632-544.

607 Kluge, A.G. (1998). Total evidence or taxonomix congruence: cladistics or consensu  
608 classification. *Cladistics*, 14, 151-158.

609 Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary  
610 Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and*  
611 *Evolution*,

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

614 Lindtke, D., Mayer, W., & Böhme, W. (2010). Identification of a contac zone between  
615 oviparous and viviparous common lizards (*Zootoca vivipara*) in central Europe:  
616 reproductive strategies and natural hybridization. *Salamandra*, 46, 73-82.

617 Macey, J.R., Larson, A., Ananjeva, N.B., Fang, Z., & Papenfuss, T.J. (1992). Two  
618 novel gene Orders and the role of light-strand replication in rearrangement of the  
619 vertebrate mitochondrial genome. *Molecular Biology and Evolution*, 14, 91-104.

620 Mila, B., Surget-Groba, Y., Heulin, B., Gosá, A., & Fitze, P.S. (2013). Multilocus  
621 phylogeography of the common lizard *Zootoca vivipara* at the Ibero-Pyrenean  
622 suture zone reveals lowland barriers and high-elevation introgression. *BMC*  
623 *Evolutionary Biology*, 13, 192.

624 Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Borwn, J.M., Bush,  
625 M.B., Harrison, S.P., Hurlbert, A.H., Knowlton, N., Lessios, H.A., McCain,  
626 C.M., McCune, A.R., McDade, L.A., McPeck, M.A., Near, T.J., Price, T.D.,  
627 Ricklefs, R.E., Roy, K., Sax, D.F., Schluter, D., Sobel, J.M., & Turelli, M.

628 (2007). Evolution and the latitudinal diversity gradient: speciation, extinction  
 629 and biogeography. *Ecology Letters*, 10: 315-331.  
 630 Posada, D. (2008). jModelTest: phylogenetic model averaging. *Molecular Biology and*  
 631 *Evolution*, 25, 1253-1256.  
 632 Provan, J., & Bennet, K.D. (2008). Phylogeographic insights into cryptic glacial  
 633 refugia. *Trends in Ecology & Evolution*, 23, 564-571.  
 634 Pyron, R.A., Burbrink, T.B., & Wiens, J.J. (2013). A phylogeny and revised  
 635 classification of Squamata, including 4161 species of lizards and snakes. *BMC*  
 636 *Evolutionary Biology*, 13, 93.  
 637 Ree, R.H., & Smith, S.A. (2008). Maximum likelihood inference of geographic range  
 638 evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*,  
 639 57, 4-14.  
 640 Sebel, M. (1989). The diet of the common viper (*Vipera berus*) in the area of Mt.  
 641 Králícký Sněžník. *Folia Zoologica*, 29, 117-123.  
 642 Smith, M.F., & Patton, J.L. (1991). Variation in mitochondrial Cytochrome b sequence  
 643 in natural populations of south American akodontine rocents (Muridae:  
 644 Sigmodontinae). *Molecular Biology and Evolution*, 8, 85-103.  
 645 Stephens, M., Smith, N.J., & Donnelly, P. (2001). A new statistical method for  
 646 haplotype reconstruction from population data. *American Journal of Human*  
 647 *Genetics*, 68, 978-989.  
 648 Stewart, J.R., Lister, A.M., Barnes, I., & Dalen, L. (2010). Refugia revisited:  
 649 individualistic responses of species in space and time. *Proceedings of the Royal*  
 650 *Society B-Biological Sciences*, 277, 661-671.

651 Surget-Groba, Y., Heulin, B., Ghielmi, S., Guillaume, C.P., & Vogrin, N. (2002).  
652 Phylogeography and conservation of the populations of *Zootoca vivipara*  
653 carniolica. *Biological Conservation*, 106, 365-372.

654 Surget-Groba, Y., Heulin, B., Guillaume, C.P., Puky, M., Semenov, D., Orlova, V.,  
655 Kupriyanova, L., Ghuira, I. ,& Smajda, B. (2006). Multiple origins of viviparity,  
656 or reversal from viviparity to oviparity? The European Common Lizard  
657 (*Zootoca vivipara*, Lacertidae) and the evolution of parity. *Biological Journal of*  
658 *the Linnean Society*, 87, 1-11.

659 Surget-Groba, Y., Heulin, B., Guillaume, C.P., Thorpe, R.S., Kupriyanova, L., Bogrin,  
660 N., Maslak, R., Mazzotti, S., Venczel, M., Ghira, I., Odierna, G., Leontyeva, O.,  
661 Monney, J.C., & Smith, N. (2001). Intraspecific phylogeography of *Lacerta*  
662 *vivipara* and the evolution of viviparity. *Molecular Phylogenetics and*  
663 *Evolution*, 18, 449-459.

664 Ursenbacher, S., Carlsson, M., Helfer, V., Tegelström, H., & Fumagalli, L. (2006).  
665 Phylogeography and Pleistocene refugia of the adder (*Vipera berus*) as inferred  
666 from mitochondrial DNA sequence data. *Molecular Ecology*, 15, 3425-3437.

667 Vandenberghe, J., French, H.M., Gorbunov, A., Marchenko, S., Velichko, A.A., Jin, H.,  
668 Cui, Z., Zhang, T., & Wan, X. (2014). The Last Permafrost Maximum (LPM)  
669 map of the Northern Hemisphere: permafrost extent and mean annual air  
670 temperatures, 25–17 ka BP. *BOREAS*, 43, 652-666.

671 Yu, Y., Harris, A.J., & He, X.J. (2010). S-DIVA (statistical dispersal-vicariance  
672 analysis): a tool for inferring biogeographic histories. *Molecular Phylogenetics*  
673 *and Evolution*, 56, 848-850.

674 Yu, Y., Harris, A.J., Blair, C., & He, X.J. (2015). RASP (Reconstruct Ancestral State in  
675 Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and*  
676 *Evolution*, 87, 46-49.  
677  
678

679   **BIOSKETCH**

680

681   JLH and PSF are interested in evolutionary biology, ecology, and population genetics,  
682   in order to understand short- and long-term species dynamics. They are particularly  
683   interested in how ecology drives genetic evolution.

684           JLH and PSF designed the study. JLH, MCB, BH, and PSF sampled Spanish  
685   populations and populations in Southern France. BH, YSG, and TAO provided tissue  
686   samples. JLH, MLP, TS and PSF developed genetic markers, amplified the gene  
687   fragments, and prepared the alignments. JLH and PSF analysed the data and wrote the  
688   article, and all authors commented the last version of the manuscript.

689 **Tables**

690 **Table 1.-** Genetic variability of the mitochondrial and nuclear genes per *Z. vivipara* clade (A, B, C, D, E and F). Given are the number of  
 691 individuals (*N*), haplotype and nucleotide diversity (Hd and Nd, respectively).

692

	<b>Clade</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>
<b>DNA type</b>	<b>N</b>	38	63	24	30	52	24
<b>Nuclear</b>	<b>Hd - nZV1</b>	0.555	0.056	0.513	0.645	0.262	0.653
	<b>Hd - nZV2</b>	0.764	0.347	0.809	0.658	0.401	0.595
	<b>Hd - nZV3</b>	0.642	0.252	0.198	0.778	0.110	0.411
	<b>Nd - nZV1</b>	0.006	0.001	0.004	0.007	0.002	0.007
	<b>Nd - nZV2</b>	0.003	0.001	0.007	0.005	0.004	0.004
	<b>Nd - nZV3</b>	0.008	0.002	0.001	0.003	0.002	0.005
<b>Mitochondrial</b>	<b>Hd - <i>mt</i></b>	0.859	0.859	0.913	0.685	0.854	0.866
	<b>Nd - <i>mt</i></b>	0.006	0.011	0.003	0.053	0.005	0.007

693

694

695

696

697



698

699 **Table 2.-** Clades and subclades of *Z. vivipara* and their distribution. Clade and subclade abbreviations, inhabited biogeographic region (EEA,  
700 2012), and number and abbreviation of countries in which their occurrence has been confirmed are given.

701

Clade	Subclade	Distribution range	Biogeographic region	Countries ( <i>N</i> / abbreviations*)	
A	A1	North-West Italy	alpine, continental	2	A, I
A	A2	North-East Italy	alpine, continental	3	A, I, SLO
B	B4	North-East Spain	alpine	1	E
B	B3	North-Central Spain	alpine, atlantic	1	E
B	B1	North-West Spain	atlantic	1	E
B	B2	Southern France	alpine, atlantic	3	AND, E, F
C	none	Austria	alpine, continental	1	A
D	D1	North and East Eurasia	boreal, continental, steppic	9	BY, FIN, J, LT, LV, RUS, S, UA
D	D2	Central East-Carpathians	alpine, continental	1	RO

D	D3	North-West and Central Inner East-Carpathians	alpine, pannonian	4	H, RO, SK, UA
E	E1	Western Europe and Southern Balkan	alpine, atlantic, boreal, continental	15	A, B, BG, CH, D, F, GB, I, IRL, L, MNE, N, NL, S, SRB
E	E2	Eastern Europe	alpine, pannonian	5	H, PL, RUS, SK, UA
F	F1	South-West Austria	alpine, pannonian	3	A, H, SLO
F	F2	South-East Austria	alpine, pannonian	2	A, H

702

703 \*Acronyms: Austria (A), Andorra (AND), Belgium (B), Bulgaria (BG), Belarus (BY), Switzerland (CH), Germany (D), Spain (E), France (F), Finland (FIN), Great Britain

704 (GB), Hungary (H), Italy (I), Ireland (IRL), Japan (J), Luxemburg (L), Lithuania (LT), Latvia (LV), Montenegro (MNE), Norway (N), Netherlands (NL), Romania (RO),

705 Russia (RUS), Sweden (S), Slovakia (SK), Slovenia (SLO), Serbia (SRB), Ukraine (UA).

## Figure legends

**Fig. 1.-** Map showing sampled populations (dots) belonging to all described *Z. vivipara* clades across Eurasia (Sakhalin is not shown). Dot and shade colour represent different clades and subclades unravelled by the phylogenetic analyses (Fig. 3) and shaded areas the approximate area occupied by them. The zoomed map represents the central European contact zone (in Austria, Italy, and Slovenia) and the red population a confirmed secondary contact between clade E and A (Lindtke et al. 2010). The map in the lower right corner shows the species' range following IUCN Red List (<http://maps.iucnredlist.org/>).

**Fig. 2.-** Haplotype network (TCS) of all described extant *Z. vivipara* clades (Surget-Groba et al., 2001; see within figure legend). Each colour circle represents a haplotype and its size is proportional to the haplotype's frequency. Black dots represent mutation steps and connecting line length is proportional to their number. Dots within shaded areas correspond to the same previously described clade (A, B, C, D, E, F; Surget-Groba et al., 2001; colour code see within-figure legend), circled haplotypes correspond to subclades with statistical support on the phylogeny (Fig. 3). Asterisk in haplotypes belonging to E1 indicate haplotypes located on the Balkans. Grey haplotypes correspond to isolated haplotypes (in the phylogeny; Fig. 3).

**Fig. 3.-** Bayesian phylogenetic tree of the 231 *Z. vivipara* samples and the outgroups included in this study (see Appendix 2 for a tree showing all 231 specimens). Node numbers are posterior probabilities. Clade width is proportional to the number of comprised individuals. Different colours represent different clades and different

tonalities different subclades. Subclade, clade and currently inhabited biogeographic regions are given in the middle of the two trees. The map shows the major phylogeographic areas and their nomenclature. Reconstructions of ancestral area (left tree, above the lines) and ancestral biogeographic region (right tree) are given for each node. Numbers beside the reconstructed area / biogeographic region correspond to the attribution probability (in %), with upper numbers corresponding to the combined statistic and lower numbers to the reconstruction method with the highest support. Acronyms correspond to phylogeographic areas: Italy (IT), Northern Europe and Balkan (EU), Asia (AS), and South-Western Europe (SWE) and biogeographic regions (EEA, 2012): alpine (alp), atlantic (atl), continental (cont), boreal (bor), pannonian (pan), and steppic (step).

**Fig. 4.-** Map showing the reconstructed evolutionary history of *Z. vivipara* in Eurasia. A) suggested colonization routes during Pliocene and starting Pleistocene of the different clades, subclades and their ancestors, and B) glacial refugia during Pleistocene glaciations and post-glacial expansion patterns. In A), the species' reconstructed initial area (around 4.4 Mya) is delimited by a yellow ellipse and the most likely colonization routes are indicated with arrows. In B) potential location of glacial refugia of the different clades and subclades are delimited with ellipsis and directions of postglacial expansion with arrows; the colours correspond to those of the clades/subclades in Figs. 2 and 3. Potential, yet to be confirmed refugia are indicated in grey. The limit of the southern boundary of the permafrost limit during the LGM (around 21,000 years ago) is indicated in red.

755 **Supporting Information**

756

757 **Appendix S1.** Sample name, source population, and clade affiliation of the 231 *Z.*

758 *vivipara* and 2 outgroups employed in this work. Coordinates correspond to

759 geographical latitude and longitude.

760 **Appendix S2.** Full Bayesian phylogenetic tree showing all 231 *Z. vivipara* samples and

761 the outgroups. All node supports  $PP \geq 0.90$  are given on the upper-left of the node.

762 **Appendix S3.** Bayesian phylogenetic tree with molecular dating of all nodes. Given are

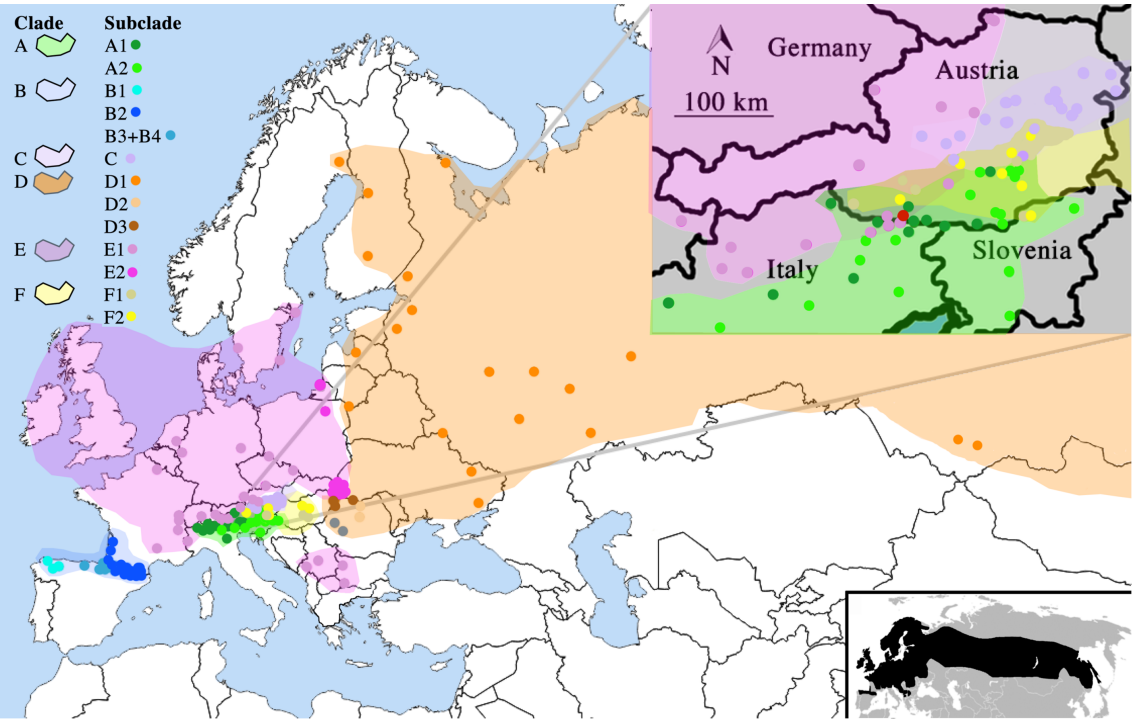
763 ages (in million years ago) and 95% CI (in square brackets).

764 **Appendix S4.** Maximum Likelihood phylogenetic tree of the *Z. vivipara* clades and the

765 outgroups. Node numbers are bootstrap values.

766

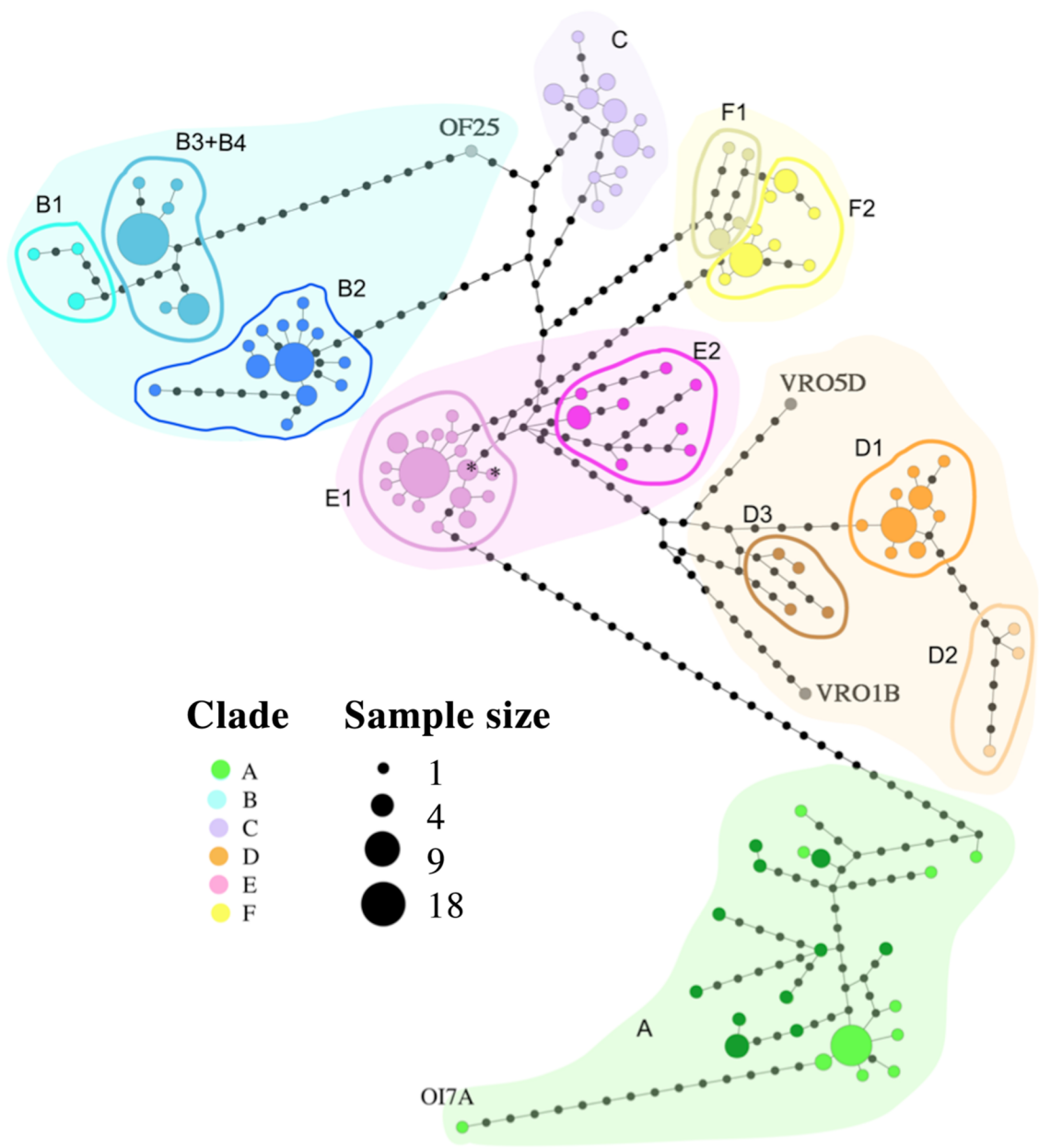
767 **Figure 1**



768

769

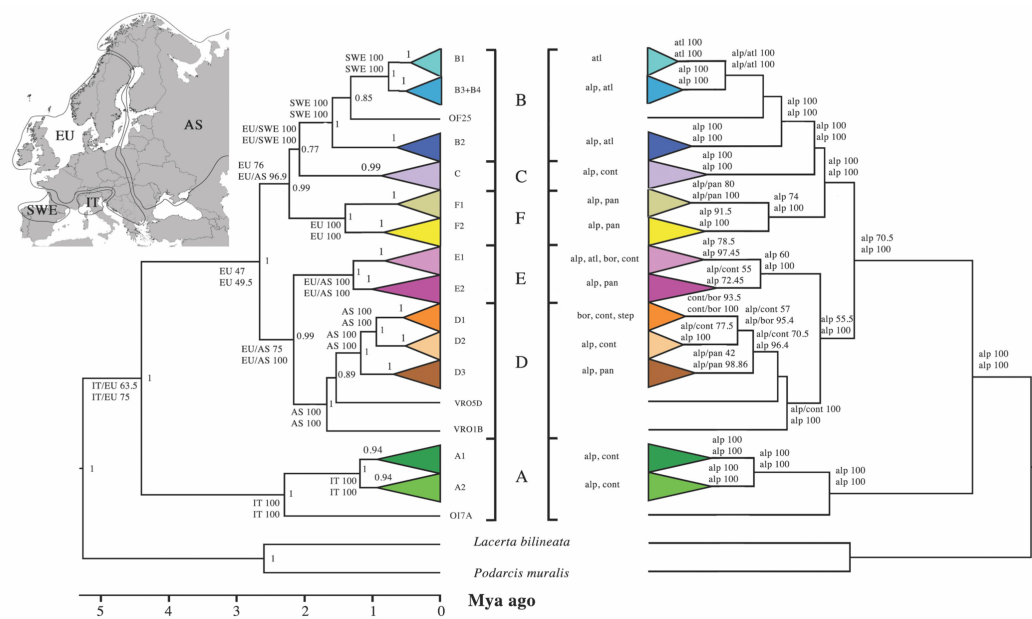
770 **Figure 2**



771

772

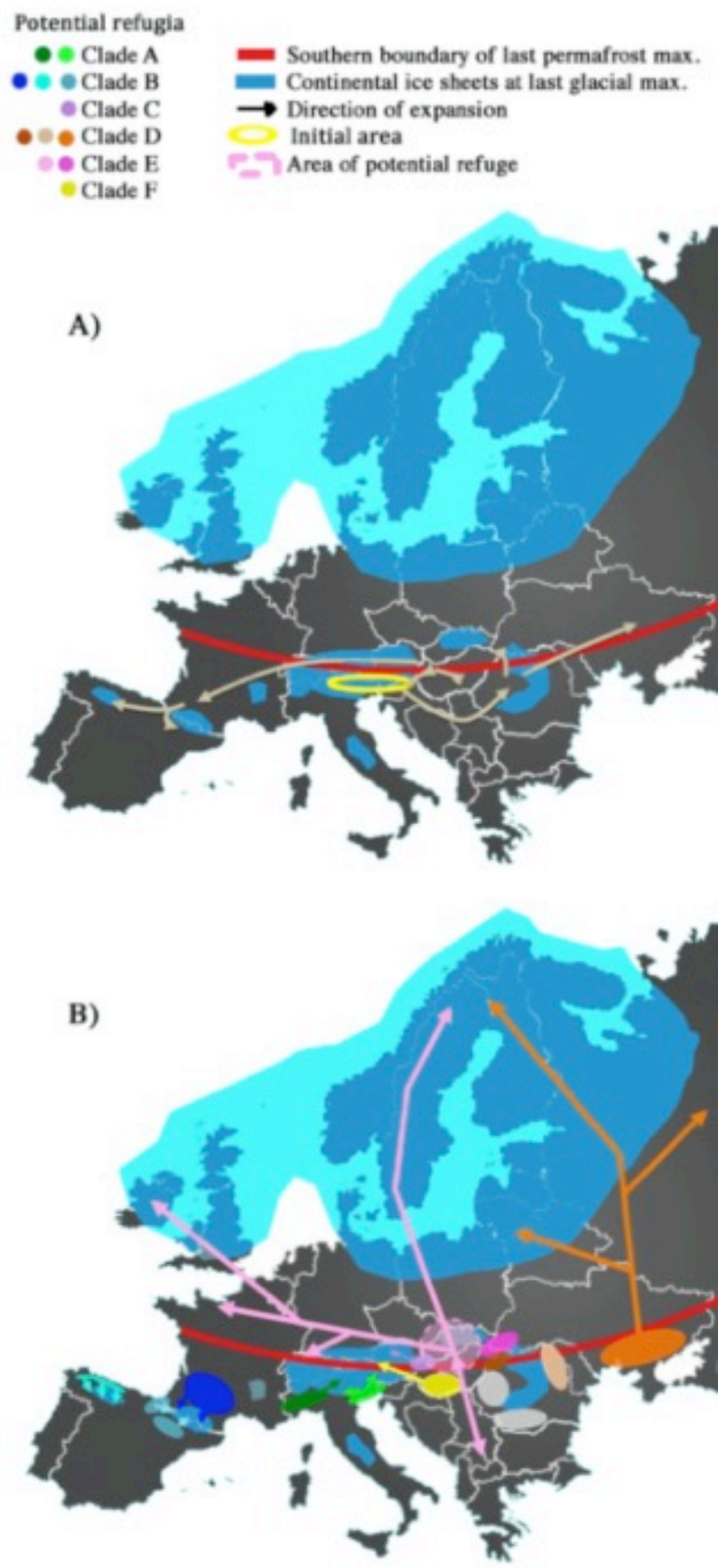
773 **Figure 3**



774

775





778 **Appendix S1.-** Sample name, source population, and clade affiliation of the 231 *Z.*  
779 *vivipara* and 2 outgroups employed in this work. Coordinates correspond to  
780 geographical latitude and longitude.  
781

Sample	Clade	Population	Longitude	Latitude
OI25 C	A	Stampoden	13.31	46.55
OI13 A	A	Alba	13.23	46.46
OI17A	A	Avoltri	12.75	46.58
ZE7	A	Bacher	16.40	46.42
ZZ5	A	Bodental	14.22	46.47
OI18 A	A	Bollone	10.59	45.74
OI8 B	A	Busatello	11.09	45.10
OI14 A	A	Cansiglio	12.40	46.08
OI7 A	A	Cavazzo	13.03	46.38
ZF50	A	Claut	12.50	46.25
OI12 A	A	Deffar	13.48	46.55
ZF51	A	Forni	12.62	46.40
ZZ1-2	A	Gallizien	14.53	46.53
OI34 A	A	Monte Grappa	11.79	45.87
OI33 A	A	Hastallone	8.11	45.92
HK1	A	Hebalm	14.97	46.93
TV5-2	A	Hörfeld	14.52	47.02
OSL7 A	A	Ig	14.54	45.96
OI32 A	A	Inghiaie	11.31	45.99
HK3	A	Kamp	14.92	46.88
VS9 A	A	Morobia	9.12	46.15
Zu9	A	Obdacher	14.43	47.02
OSL6 A	A	Pavlicevo	14.59	46.42
ZZ3	A	Petzen	14.77	46.53
ZZ4-2	A	Petzen	14.77	46.53
OSL4 A	A	Pohorje-Kot	15.61	46.49
ZU2-2	A	Preblau	14.80	46.93
OI39 A	A	Rasun	12.12	46.78
ZF29	A	Rivignano	13.00	45.92
OI40 A	A	Sanguigno	9.87	45.98
ZE2	A	Snežnik	14.42	45.58
ZF43	A	Tramonti	12.78	46.32
ZU5	A	Urban	14.18	46.75
OI1 A	A	Varese	8.70	45.78
OI27 A	A	Valtorta	9.42	45.90
OAUI A	A	Waisach	13.25	46.70
ZZ2	A	Zauchen	14.57	46.52
OSL5 A	A	Zelenci	13.76	46.49
691	B	Archilondo	-1.13	43.04
692	B	Archilondo	-1.13	43.04

OF45A	B	Bouesou	-0.67	43.00
BUG1	B	Bugangue	-0.64	43.14
09-585	B	Burgos	-3.70	43.12
CS_1200	B	Candanchu	-0.55	42.78
CS_1201	B	Candanchu	-0.55	42.78
CS_1202	B	Candanchu	-0.55	42.78
CS_1203	B	Candanchu	-0.55	42.78
CS_1204	B	Candanchu	-0.55	42.78
PF_14_158	B	Culibillas	-0.40	42.78
ETH2	B	Ethenee	-1.02	43.24
CS_527	B	Formigal	-0.40	42.79
CS_526	B	Formigal	-0.40	42.79
CS_531	B	Formigal	-0.40	42.79
OF33C	B	Gabas	-0.42	42.90
PF136	B	Ibañeta	-1.32	43.02
OF25	B	Iraty	-1.07	43.04
AG09201	B	Irun	-1.80	43.31
ISA 3	B	Issarbe	-0.80	43.03
OF39 A	B	Lescun	-0.65	42.90
Lizaso_014	B	Lizaso	-1.67	42.96
OF41A	B	Peyrenere	-0.55	42.80
OF31 B	B	Pourtalet	-0.40	42.79
RC21	B	Roncesvalles	-1.31	43.02
RHU1	B	La Rhune	-1.63	43.31
Lizaso_016	B	Lizaso	-1.67	42.96
Lizaso_017	B	Lizaso	-1.67	42.96
Lizaso_018	B	Lizaso	-1.67	42.96
OE8 B	B	Ancares	-6.91	42.87
09-503	B	Xistral	-7.54	43.47
09-532	B	Somiedo	-6.23	43.00
09-505	B	Xistral	-7.54	43.47
PF-15-109	B	Vielha	0.76	42.63
PF-15-134	B	Urdiceta	0.28	42.68
OF36A	B	Aubisque	-0.34	42.97
BLA3	B	Berguery	-0.55	43.12
PF-15-098	B	Bonaigua	0.99	42.65
PF-15-100	B	Bonaigua	0.99	42.65
CAP1	B	Capitole	-0.89	44.22
CET1	B	Cette	-0.58	42.93
PF-15-125	B	Aspin	0.33	42.94
PF-15-126	B	Aspin	0.33	42.94
PF_14_159	B	Culivillas	-0.40	42.78
PF-15-001	B	Els Estanyols	2.08	42.51
PF-15-003	B	Els Estanyols	2.08	42.51
EST 1	B	Estaing	-0.23	42.91

HOU 1	B	Hourcq	-0.43	42.91
PF-15-123	B	Mourtis	0.77	42.90
Lizaso_015	B	Lizaso	-1.67	42.96
OF29 A	B	Moura de Montrol	-1.29	43.51
OF28A	B	Palomieres	0.16	43.05
OF30A	B	Pinet	1.98	42.86
PF_14_234	B	Pla de Beret	0.95	42.72
PF_14_236	B	Pla de Beret	0.95	42.72
PF_14_228	B	Pla de Beret	0.95	42.72
OF27A	B	St. Raphael	-0.73	44.92
PF-15-075	B	Soldeu	1.67	42.60
SS1 1	B	Sud Sagette	-0.40	42.89
PF-15-110	B	Vielha	0.76	42.63
PF-15-135	B	Urdiceta	0.28	42.68
PF-15-124	B	Mourtis	0.77	42.90
PF-15-078	B	Soldeu	1.67	42.60
ZL7-2	C	Aspang	16.08	47.53
VAU7 A	C	Breitenstein	15.81	47.66
ZX1	C	Haindlkar	14.62	47.57
ZN1	C	Illmitz	16.80	47.70
NZ10	C	Lassingtal	15.03	47.73
NZ11	C	Lassingtal	15.03	47.73
NZ14	C	Lunz	15.07	47.08
ZL2-2	C	Mönichkirchen	16.03	47.50
ZL3-2	C	Mönichkirchen	16.03	47.50
ZM2	C	Moosbrunn	16.45	48.02
VD9-2	C	Nikolai	14.05	47.15
NZ17	C	Oistal	15.15	47.77
NZ18	C	Oistal	15.15	47.77
VD3-2	C	Planner	14.20	47.40
ZX5-2	C	Präbichl	14.95	47.52
ZX4	C	Präbichl	14.95	47.52
ZT4	C	Ramsau	13.58	47.45
ZL6-2	C	Rax	15.72	47.72
ZL10	C	Schneeberg	15.83	47.78
VAU6 A	C	Semmering	15.85	47.62
VD8	C	Sölktal	13.93	47.32
ZL4	C	Straßegg	15.37	47.38
ZL5-2	C	Straßegg	15.37	47.38
VAU2 B	C	Weiden am See	16.87	47.92
ACLV14-26	D	Adazi	24.48	57.11
VR21A	D	Altai	87.25	51.80
VRO1 B	D	Craiului	22.33	46.20
VRO4 A	D	Rodnei	24.95	47.47
VRO4 B	D	Rodnei	24.95	47.47

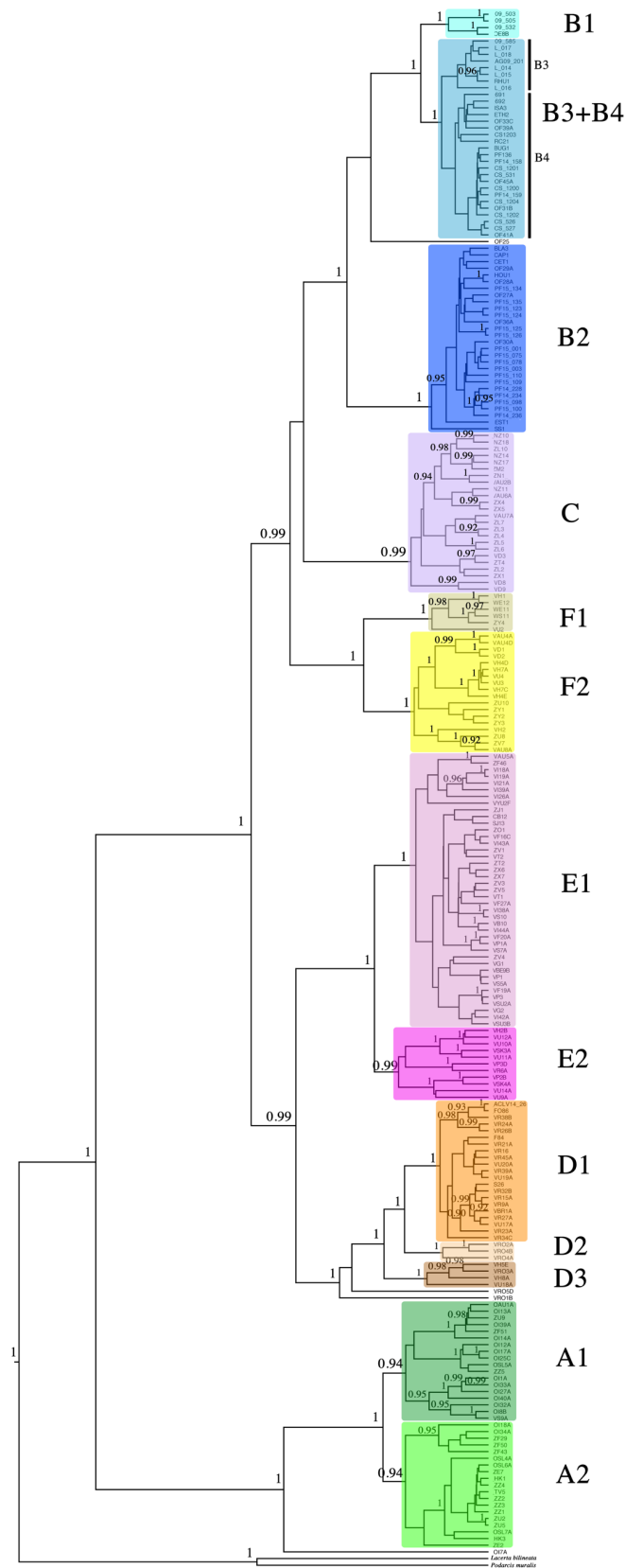
VH5 E	D	Batorliget	22.27	47.77
VR32 B	D	Chernogolovka	38.39	56.01
VU17 A	D	Chervonnyy	33.50	52.33
VU18 A	D	Delovoe	24.23	47.90
VRO2 A	D	Eremitu	24.92	46.66
VU19 A	D	Eskhar	36.58	49.80
VH8 A	D	Fehergyarmat	22.51	48.01
VBR1 A	D	Grodno	23.91	53.91
VRO5 D	D	Hunedoara	23.20	45.66
VR23 A	D	Kousznetchnoie	29.84	61.15
F84	D	Jyväskylä	25.75	62.18
VR38 B	D	Kara-Khol	89.26	51.47
VR9A	D	Krasnitsy	30.35	59.45
VRO3 A	D	Marghita	22.35	47.31
VU20 A	D	Novoandrejevka	37.35	47.64
FO86	D	Oulu	25.82	65.03
S26	D	Överkalix	22.84	66.38
VR45 A	D	Raievka	48.93	52.28
VR26 B	D	Srednii	33.82	66.34
VR34 C	D	Shahovskoe	43.03	56.04
VR16 A	D	Sakhalin	142.65	51.44
VR39 A	D	Tambov	41.46	53.19
VR15 A	D	Tschuvachia	46.68	55.02
VR27 A	D	Uzmino	28.77	58.43
VR24 A	D	Ydmyrtia	53.09	56.94
VI26 A	E	Pramollo	13.29	46.55
ZT2	E	Almsee	13.95	47.77
VP3	E	Ältasjön	18.20	59.27
VF27 A	E	Autreppes	3.86	49.91
CB12	E	Bjelasica	19.41	43.01
VF20 A	E	Bonnevaux	6.17	46.84
VSK3 A	E	Botany	22.09	48.45
VF19 A	E	Chambery	6.12	45.69
VG2	E	Duisburg	6.55	51.05
VS10 A	E	Fuela	9.93	46.78
ZJ1	E	Franking	12.90	48.08
VU10 A	E	Gertsovtsi	22.65	48.58
VSU3 B	E	Gothenburg	12.31	57.37
VAU5 A	E	Hausalm	12.95	46.62
ZV1-2	E	Heidenreichstein	15.15	48.85
ZV3	E	Heidenreichstein	15.15	48.85
ZX7	E	Hengstpass	14.47	47.68
ZX6	E	Hengstpass	14.47	47.68
VS7 A	E	Hochainplangen	8.48	46.73
VBE9 B	E	Turracher Hohe	13.88	46.91

VU12 A	E	Irskava	23.23	48.43
VU11 A	E	Klimets	23.24	48.84
VSK4 A	E	Kolonica	22.27	48.96
VYU2 F	E	Kopaonik	20.69	43.46
VP3 D	E	Krutyn	21.42	53.71
VI19 A	E	Lanza	13.15	46.58
VF16 C	E	Mas de la Barque	3.88	44.39
VG1	E	Overasseltse	5.83	51.76
VU9 A	E	Perechin	22.76	48.62
VU14 A	E	Perechin	22.76	48.62
VB10	E	Petrohan	23.13	43.12
SJ13	E	Pirin	23.40	41.76
VI18 A	E	Pramollo	13.28	46.56
VT1	E	Radun	17.99	49.88
VT2 A	E	Radun	17.99	49.88
VI42 A	E	Ribba	7.04	44.88
VS5 A	E	Roch	6.18	46.53
VSU2 A	E	Runsten	16.69	56.70
VR6 A	E	Rybachii	20.85	55.17
ZF46	E	Sauris	12.70	46.48
ZO1	E	Stubachtal	12.60	47.22
VI39 A	E	Sutrio	12.96	46.51
VP1 A	E	Szklarska	15.54	50.83
VH2 B	E	Tarpa	22.53	48.13
ZV4	E	Thaures	15.07	48.87
ZV5-2	E	Thaures	15.07	48.87
VI38 A	E	Tonale	10.59	46.26
VI44 A	E	Tovel	10.95	46.22
VP2 B	E	Ustrzyki	22.65	49.10
VI43 A	E	Martello	10.68	46.48
VI21 A	E	San Nicolo	11.77	46.42
VP1	E	Wurzen	12.63	51.32
WE-11	F	Christebaueralm	13.38	46.87
WS11	F	Defreggen	12.47	46.82
VAU4 D	F	Emberfer Alm	13.15	46.77
VAU4 A	F	Emberfer Alm	13.15	46.77
VH7 A	F	Felsobesnyo	19.27	47.23
VH7 C	F	Felsobesnyo	19.27	47.23
ZY1-2	F	Gaberl	14.92	47.10
ZY2	F	Gaberl	14.92	47.10
ZY4	F	Gleinalm	15.00	47.02
VAU8 A	F	Godingberg	14.93	46.78
VU3	F	Izsak	19.35	46.80
VU-2	F	Izsak	19.35	46.80
VU4-2	F	Izsak	19.35	46.80

ZV7	F	Klippitztörl	14.73	46.93
ZU8	F	Klippitztörl	14.73	46.93
VH1	F	Koralm	14.95	46.48
VH2	F	Koralm	14.95	46.48
WE12	F	Mallnitz	13.25	46.98
VD1	F	Murau	14.15	47.10
VD2	F	Murau	14.15	47.10
VH4 D	F	Osca	19.23	47.27
VH4 E	F	Osca	19.23	47.27
ZU10	F	Prebl	14.78	46.95
ZY3	F	Wieser Alm	15.28	47.33

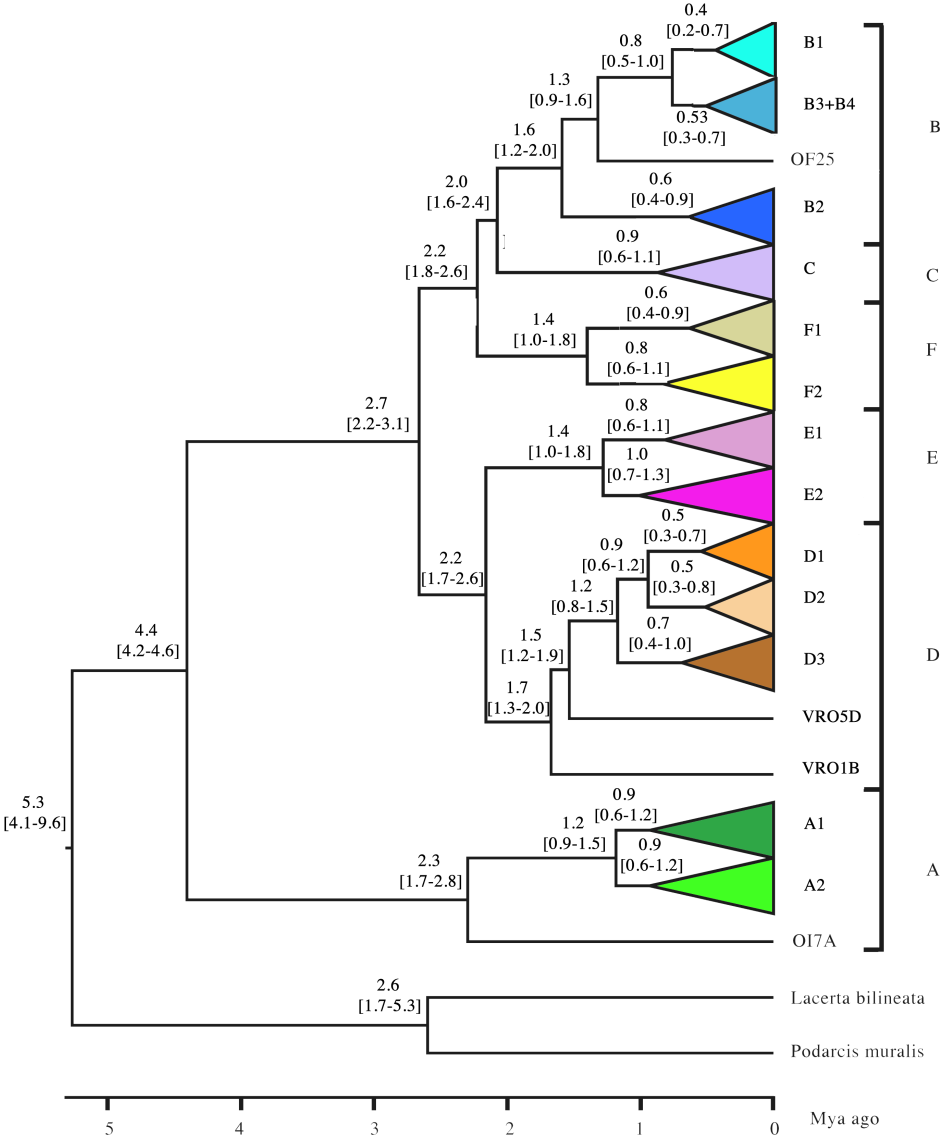
782  
783  
784  
785  
786  
787

788 **Appendix S2.** Full Bayesian phylogenetic tree showing all 231 *Z. vivipara* samples and  
 789 the outgroups. All node supports PP  $\geq 0.90$  are given on the upper-left of the node.





**Appendix S3.** Bayesian phylogenetic tree with molecular dating of all nodes. Given are ages (in million years before present) and 95% CI (in square brackets).



**Appendix S4.** Maximum Likelihood phylogenetic tree of the *Z. vivipara* clades and the outgroups. Node numbers are bootstrap values.

