Population genomics of wall lizards reflects the dynamic history of the Mediterranean Basin

Weizhao Yang[1*], Nathalie Feiner[1], Daniele Salvi[2], Hanna Laakkonen[1], Daniel Jablonski[3], Catarina Pinho[4], Miguel A. Carretero[4,5], Roberto Sacchi[6], Marco A. L. Zuffi[7], Stefano Scali[8], Konstantinos Plavos[1], Panayiotis Pafilis[9], Nikos Poulakakis[10,11], Petros Lymberakis[10], David Jandzik[3], Ulrich Schulte[12], Fabien Aubret[13,14], Arnaud Badiane[15], Guillem Perez i de Lanuza[16], Javier Abalos[16], Geoffrey M. While[17], Tobias Uller[1*]

1. Department of Biology, Lund University, Lund 223 62, Sweden
2. Department of Health, Life and Environmental Sciences, University of L'Aquila, Coppito 67100, L'Aquila, Italy
3. Department of Zoology, Comenius University in Bratislava, Ilkovičova 6, 842 15 Bratislava, Slovakia
4. CIBIO/InBIO Research Centre in Biodiversity and Genetic Resources, Universidade do Porto, Campus de Vairão, 4485-661 Vairão, Portugal
5. Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, R. Campo Alegre, s/n, 4169 - 007, Porto, Portugal
6. Department of Earth and Environmental Sciences, University of Pavia, Pavia 27100, Italy
7. Museum Natural History, University of Pisa, Pisa 56011, Italy
8. Museum of Natural History of Milan, Milano 20121, Italy
9. National & Kapodistrian University of Athens, School of Science, Faculty of Biology, Panepistimiopolis 15701, Athens, Greece
10. Natural History Museum of Crete, School of Sciences and Engineering, University of Crete, Knossos Avenue, Hrakleio 71409, Greece
11. Biology Department, School of Sciences and Engineering, University of Crete, Voutes University Campus, Heraklion 70013, Greece
12. Büro für Faunistische Gutachten - Dr. Ulrich Schulte, Kaiserstraße 2, 33829 Borgholzhausen, Germany
13. Station d'Ecologie Théorique et Expérimentale, CNRS 09200 Moulis, France
14. School of Molecular and Life Sciences, Curtin University, WA 6102, Australia

15. IMBE, Aix-Marseille Université, Avignon Université, CNRS, IRD, Marseille, France

16. Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, APT. 22085, 46071, Valencia, Spain

17. School of Natural Sciences, University of Tasmania, Sandy Bay, Tasmania 7005, Australia

*Email: weizhao.yang@biol.lu.se; tobias.uller@biol.lu.se

Abstract

The Mediterranean Basin has experienced extensive change in geology and climate over the past six million years. Yet, the relative importance of key geological events for the distribution and genetic structure of the Mediterranean fauna remains poorly understood. Here, we use population genomic and phylogenomic analyses to establish the evolutionary history and genetic structure of common wall lizards (*Podarcis muralis*). This species is particularly informative because, in contrast to other Mediterranean lizards, it is widespread across the Iberian, Italian, and Balkan peninsulas, and in extra-Mediterranean regions. We found strong support for six major lineages within *P. muralis*, which were largely discordant with the phylogenetic relationship of mitochondrial DNA. The most recent common ancestor of extant *P. muralis* was likely distributed in the Italian Peninsula, and experienced an "Out-of-Italy" expansion following the Messinian salinity crisis (~5 Mya), resulting in the differentiation into the extant lineages on the Iberian, Italian and Balkan peninsulas. Introgression analysis revealed that both inter- and intraspecific gene flow have been pervasive throughout the evolutionary history of *P. muralis*. For example, the Southern Italy lineage has a hybrid origin, formed through admixture between the Central Italy lineage and an ancient lineage that was the sister to all other *P. muralis*. More recent genetic differentiation is associated with the onset of the Quaternary glaciations, which influenced population dynamics and genetic diversity of contemporary lineages. These results demonstrate the pervasive role of Mediterranean geology and climate for the evolutionary history and population genetic structure of extant species.

Keywords: phylogeography, phylogenomics; introgression; refugia; glaciation; Messinian salinity crisis

Introduction

The reconstruction of evolutionary history is essential if we are to understand the factors and processes that explain contemporary patterns of biodiversity (Avise 2000). The Mediterranean Basin is considered a biodiversity hotspot because of its species richness and high degree of endemism (Médail and Quezel 1997; Myers et al. 2000). However, the processes responsible for Mediterranean diversity and biogeography have proven difficult to resolve because of the complex geological and climatic history of this region (Cavazza and Wezel 2003; Médail and Diadema 2009; Thompson 2005; Hewitt, 2011a). Defining events include the Messinian salinity crisis (MSC; 5.96-5.33 million years ago, Mya; Krijgsman et al. 1999, 2010; Duggen et al. 2003), a period of progressive aridity that led to the extinction of subtropical Tertiary lineages and the diversification of arid-adapted lineages as well as extensive faunal exchange within the Mediterranean and with Africa (Jiménez-Moreno et al. 2010; Fiz-Palacios and Valcárcel 2013). The refilling of the Mediterranean (García-Castellanos et al. 2009) resulted in the isolation of biota on islands and peninsulas, and the onset of the Mediterranean climate of today (3.4-2.8 Mya) significantly changed ecological communities (Tzedakis 2007; Postigo et al. 2009). The later Quaternary climatic oscillations (starting ca. 2.5 Mya), characterized by the alternation of colder (glacial) and warmer (interglacial) periods, further affected the distribution of many species due to recurrent range shifts with associated cycles of demographic expansion-contraction (Hewitt 1996, 2000; Provan and Bennet 2008; Taberlet et al. 1998). During glacial periods, the Mediterranean Basin provided refugial areas, where the long-term persistence of isolated populations frequently resulted in the formation of new allopatric lineages (Hewitt 1996, 2004; Wiens 2004; Gentili et al. 2015; Mairal et al. 2017).

While all these events contributed to the evolution of the Mediterranean fauna, their respective roles for explaining the genetic structure and geographic distribution of extant taxa remain poorly understood (Hewitt 2011a). One reason for this is that reconstruction of evolutionary history can be challenging, especially when lineages have been subject to repeated range expansions and contractions over the Quaternary climatic cycles (e.g., Buckley 2009; Hewitt 2011b; Carstens et al. 2013). Introgression imposes further challenges for reconstructing historical relationships among taxa, as well as estimates of taxonomic diversity (e.g., Naciri and Linder 2015; Mallet et al. 2016). Approaches based on representative sampling of genome-wide genetic variability allow powerful and refined phylogeographical inference overcoming many of the limitations of more traditional molecular markers (Delsuc et al. 2005; Cutter 2013; McCormack et al. 2013). Such phylogenomic approaches can be particularly useful in resolving evolutionary affinities in situations where clades are separated with short internal branches (Pollard et al. 2006) or where hybridization has occurred (Cui et al. 2013; McCluskey and Postlethwait 2015; Malinsky et al. 2018; Chen et al. 2019).

Wall lizards of the genus *Podarcis* are currently represented by 24 to 25 species (Speybroeck et al. 2020) and are a characteristic fauna of the Mediterranean Basin and its islands. Similar to many other Mediterranean animals (reviewed in Hewitt 2011a), wall lizards commonly exhibit high regional endemism, with species typically being restricted to one of the Balkan, Iberian, and Italian peninsulas, or one or several Mediterranean islands (Poulakakis et al. 2005; Psonis et al. 2021; Salvi et al. 2021; Yang et al. 2021). A notable exception is the common wall lizard (*P. muralis*). This species is not only widespread, being distributed from Iberia to Asia Minor, but it is also native to extra-Mediterranean regions in Western, Central and Eastern Europe (Schulte et al. 2008; fig. 1A). Previous studies based on DNA sequence data suggest that such widespread geographic distribution has been accompanied by regional differentiation into more than 20 genetic lineages that purportedly diverged during the Pleistocene glaciations (Salvi et al. 2013). These lineages were defined by divergence in mitochondrial DNA, several of them separated by low genetic divergence (i.e. short internal branches), and their phylogenetic relationships therefore remain largely unresolved (Salvi et al., 2013). Moreover, recent analyses of single nucleotide variants (SNV) data have demonstrated extensive gene flow even between distantly related mtDNA lineages of *P. muralis* (Yang et al. 2018; 2020). This suggests a much more complex scenario than what can be revealed by traditional phylogenetic studies.

In this study, we implemented a phylogenomic approach based on both restriction site-associated DNA sequencing (RAD-Seq) and whole genome sequencing (WGS) to identify the processes that have shaped genetic differentiation and biogeography of *P. muralis.* We had three specific aims. First, to identify lineages of *P. muralis* from both nuclear and mitochondrial genomic data and establish their phylogenetic relationships. Second, to establish evidence for hybridization and introgression between each of the major *P. muralis* lineages and between *P. muralis* and other *Podarcis* species that currently overlap in their distribution. Third, to reconstruct the biogeographic and demographic history of *P. muralis* in Europe, and assess whether these reflect the paleogeographic and climatic history of the Mediterranean Basin.

Results

We collected samples from 55 locations for RAD-Seq covering the previously suggested genetic lineages within *P. muralis* (fig. 1A; supplementary table S1; Gassert et al. 2013; Salvi et al. 2013; Jablonski et al. 2019). In total, 28,039 SNVs were obtained with a mean coverage of 17.10 per site and an average genotyping rate of 0.95. In addition, whole genomes of 16 individuals were sequenced (supplementary fig. S1; supplementary table S2), obtaining 9,699,080 nuclear SNVs with a mean coverage of 11 and genotyping rate above 0.99. We also

sequenced the full mitochondrial genomes of those individuals (13,884 bp). One individual of *P. bocagei*, *P. siculus*, and *P. tiliguerta* served as outgroups for all analyses, and *Archaeolacerta bedriagae* was further included as outgroup for WGS data (see Yang et al. 2021 for a phylogenomic analysis of all *Podarcis* species). In addition, we also included WGS data of 13 additional *Podarcis* lineages (Yang et al. 2021; supplementary table S3) to assess gene flow between *P. muralis* and other *Podarcis* species.

Population structure analysis

We inferred the genetic relationship between the *P. muralis* individuals based on genotypes from RAD-Seq data using a principal component analysis (PCA; Chang et al. 2015) and ADMIXTURE clustering (Alexander et al. 2009). Results of both PCA and ADMIXTURE supported distinct *P. muralis* lineages with a strong geographic structure. In the PCA, the first, second and third principal components (variances explained: 17.80%, 16.79% and 13.15%, resp.) largely separated populations into five different lineages: Western Europe (WE), Balkan, Southern Italy (SI), Southern Alps (SA) and Central Italy (CI) (fig. 1B). In the ADMIXTURE clustering, the best supported number of presumed ancestral populations (K = 5) further divided the Balkan populations into Southern Balkan (SB) and Northern Balkan (NB). The lineage in Southern Italy, SI, showed an admixed pattern of the Central and Northern Italian lineages, CI and SA, at K = 5, but was recovered as an independent cluster (SI) at K = 6 (fig. 1C). Several other admixed populations were also identified by this population structure analysis, including admixture between three lineages, CI, SA, and SB, on the Balkan peninsula (locations CCR and KOP), and between SA and CI in Northern Italy (locations BV, CE, and MG) (fig. 1C).

Phylogenetic relationship between lineages

Following the population genetic structure, we inferred the phylogenetic relationship between the six major lineages (CI, NB, SA, SB, SI, and WE). Concatenated maximum likelihood (ML) analysis yielded a highly resolved phylogeny based on RAD-Seq data: 80% of the branches exhibited bootstrap values > 80%, and the branches leading to the six major lineages showed bootstrap values equal to 100% (fig. 1D). In this phylogeny, SA formed the sister clade to all other lineages, CI formed the sister clade to the WE and the Balkan clade, where each of the latter two were divided into northern and southern sub-clades (note that the subdivision of WE into northern and southern sub-clades was not detected by the ADMIXTURE clustering). The only poorly supported node connecting the major lineages (bootstrap of 85%) was associated with the phylogenetic position of SI as a sister clade to all non-SA lineages. Considering this uncertainty and the admixed pattern of SI in the

ADMIXTURE analysis, we also reconstructed the phylogeny by excluding all SI individuals, in which the phylogenetic relationship between all other lineages remained the same (supplementary fig. S2).

Next, we inferred the ML phylogeny based on WGS data for 16 individuals representing all major distribution ranges (Supplementary fig. S1). This WGS phylogeny was supported by bootstrap values of 100% for all nodes and its topology showed the same relationships among the six major lineages as the RAD-Seq phylogeny (Supplementary fig. S3). One individual from the admixed population KOP (fig. 1C), clustered with the Balkan lineages and was excluded from the following analyses (fig. 2). We also validated the phylogeny using a multispecies coalescent approach in ASTRAL-III (Zhang et al. 2018) based on local ML trees of 200 kb windows across the genome, and again obtained the same topology with bootstrap values of 100% for all branches (Supplementary fig. S4). These results congruently indicated that the obtained phylogeny was highly robust.

We obtained an ML tree based on mitochondrial genome sequences with well supported clades (average bootstrap value > 97.57%; fig. 2). The mtDNA phylogeny was extensively discordant with the phylogeny derived from nuclear data (nuDNA). For example, the SA lineage was grouped with WE and SB in the mtDNA tree (which in turn grouped with the SI mtDNA lineage), while the other major mtDNA clade was formed by only CI and NB. Surprisingly, the mitochondrial genome of the individual from San Remo (SR) was not nested within the SA lineage, but was basal to all other lineages. Other mtDNA-nuDNA discordances were found for populations from Elba and in the contact zone between the SA and SB lineages on the Balkan Peninsula (fig. 2). We confirmed that the mitochondrial genomes of all the *P. muralis* lineages clustered as a monophyletic clade by including mitochondrial genomes of all the13 *Podarcis* species in the phylogenetic analysis (Supplementary fig. S5).

Gene flow analysis

We first tested for interspecific gene flow between *P. muralis* lineages and the 13 *Podarcis* species and lineages (see Yang et al. 2021), using D statistics (supplementary table S3; Patterson et al. 2012). A total of 273 tests were performed, in which 127 (46.52%) significantly deviated from neutrality (Z-score > 3.3). The top 50, and more than 2/3 of all significant tests, involved the WE lineage, which excessively shared alleles with the common ancestor of the Iberian species group of *Podarcis* (here represented by *P. bocagei* and *P. hispanicus*) and the Ibiza wall lizard (*P. pityusensis*). Other signatures of interspecific gene flow were found between *P. siculus* and the CI and SA lineages, but the D and Z-scores were substantially lower than for WE and the Iberian *Podarcis* species (fig. 3A).

Next, we investigated introgression within *P. muralis*. The results of D statistics revealed that 33 out of 35 tests (94.29%) significantly deviated from neutrality (Z-score > 3.3), suggesting substantial genetic exchange between the major *P. muralis* lineages (fig. 3B; detailed information in supplementary table S4). This pattern was supported by further analyses using phyloNet (supplementary fig. S6; Wen et al. 2008) and qpGraph (fig. 3C; Patterson et al. 2012). The phylogenetic network indicated a complex evolutionary history for the WE lineage, and demonstrated that the introgression from Iberian species (i.e., *P. hispanicus* complex) likely happened well before the WE lineage split into a northern and southern clade. This analysis also revealed that the SA lineage experienced introgression from the common ancestor of the two Balkan lineages. A complex history was inferred for the SI lineage, which received one part of its genome (phyloNet: 40%; qpGraph: 32%) from the CI lineage, and the other part (phyloNet: 60%; qpGraph: 68%) from a sister clade to all other *P. muralis* lineages (fig. 3C, supplementary fig. S6). This early diverging and extinct clade also contributed alleles (phyloNet: 44%; qpGraph: 10%) to the common ancestor of all extant lineages, except SA.

The evolutionary history of the Southern Italy lineage

Gene flow analysis suggested that the current SI lineage resulted from the fusion of populations belonging to the CI lineage and an ancient SI lineage that was sister to all other lineages. Thus, the placement of the SI lineage in our inferred phylogeny was not supported by the phylogenetic network with gene flow (compare fig. 1D, fig. 2 and fig. 3C). To clarify the evolutionary history of the SI lineage, we estimated the distribution of different tree topologies across the genome based on 200 kb windows. A total of 5,576 high-quality local trees were inferred with average bootstrap value > 60%. Among these trees, 1,816 trees (32.57%) supported a monophyletic clade formed by the CI and SI lineages (Topology 3, 4 and 6 in fig. 4A&B). The windows supporting this relationship are referred to as the "CI-ancestry genome" of the SI lineage.

To test if introgression contributed to the discordance of lineage phylogeny and local trees besides incomplete lineage sorting, we further calculated the absolute genetic distance (Dxy) between CI and SI, and the fd-statistics (Martin et al. 2015) on the topology (WE/Balkan, CI, SI, Outgroup) for the CI-ancestry part and the rest of the genome. The Dxy of the CI-ancestry genome (0.1127) was significantly lower than the Dxy of the rest of genome (0.1198; p-value < 0.001; 1,000 iterations of permutation test; fig. 4C). Conversely, the fd of the CI-ancestry genome (0.0778) was significantly greater than that of the rest of genome (0.0265; p-value < 0.001; 1,000 iterations of permutation test; fig. 4D). These results

strongly suggest that the CI-ancestry part of the SI genome was derived from an introgression between the CI and an ancient SI lineage.

Following these results, we next inferred the phylogeny using the same multispecies coalescent method but keeping the CI and the ancient SI lineage parts of the genome separated. The CI-ancestry genome of the SI lineage indeed formed a sister taxon with CI close to Western Europe / Balkan clade in the phylogeny of *P. muralis*. However, the ancient SI lineage formed an independent clade that branched off before any of the other *P. muralis* lineages (fig. 4E). This was consistent with the topology inferred by introgression analysis (fig. 3C), which indicated that about 68% of alleles in the genome of the SI lineage introgressed from this early diverged ancient SI lineage, and 32% of alleles introgressed from the CI lineage.

Divergence time estimation

Divergence time estimation between *P. muralis* lineages was performed based on the 200-kb genomic windows (N = 346) for which local trees supported the topology derived from the ancient SI part of the genome (see fig. 4E). Two secondary calibrations were used from a fossil-calibrated Lacertini phylogeny of Garcia-Porta et al. (2019) - the root node (37.55 Mya) and the crown node of *Podarcis* (18.60 Mya). The time-calibrated tree (fig. 5A) revealed an early split of the ancient SI lineage estimated at ca. 6.24 Mya, followed by the separation of the SA lineage and the MRCA of all other lineages during the Messinian salinity crisis at ca. 5.76 Mya. The CI lineage separated shortly afterwards (ca. 4.90 Mya). The divergence between WE and the Balkan lineages was estimated to be ca. 4.05 Mya, and the divergences within the Western Europe and Balkan lineages were almost coinciding (ca. 2.54 Mya and ca. 2.58 Mya). We also estimated the divergence times based on the phylogeny without the ancient SI lineage, using the same methods, which generated consistent results (supplementary fig. S7).

Estimation of biogeographic and demographic history

To infer the possible ancestral distribution range and the biogeographical events leading to the extant distribution of *P. muralis* lineages, we used BioGeoBEARS (Matzke 2013) with a total of three models - dispersal extinction cladogenesis (DEC), dispersal vicariance analysis (DIVALIKE), Bayesian inference of historical biogeography for discrete areas (BAYAREALIKE). We defined seven biogeographic areas: Iberian Peninsula, Western Europe, Southern Alps, Central Italy, Southern Italy, Northern Balkan and Southern Balkan. The results showed that the DIVALIKE was the best-fitting model (AICc = 50.69; supplementary table S5). All three models suggested very similar patterns, in which the

ancestral range of *P. muralis* was located on the Italian Peninsula. The three major lineages found there today (the SI, SA and CI lineages) appear to have separated within the Italian Peninsula, followed by an out-of-Italy dispersal towards the Balkan and Iberian Peninsulas, and into north-western and central Europe (fig. 5A,B; supplementary fig. S8).

We further reconstructed the detailed demographic history of each lineage using the pairwise sequential Markovian coalescence (PSMC; Li and Durbin 2011; see fig. 5). The SI lineage was excluded in this analysis due to its highly heterogeneous genome. With respect to the demographic history, the PSMC model suggested an initial population expansion for all lineages until the Lower Pleistocene (ca. 2 Mya; fig 5). Subsequently, all lineages except CI experienced a population decline at the beginning of the Quaternary climate oscillations (Gelassian and Calabrian). The CI, NB, SB and WE lineages then experienced another population expansion from ca. 700 thousand years ago (Kya) during the Günz complex. Then, all lineages appear to have experienced a decline around the beginning of Upper Pleistocene (200 - 100 Kya), coinciding with a long glacial period (fig. 5C).

Discussion

The population- and phylogenomic analyses of common wall lizards provide insights into how geological and climatic change in the Mediterranean Basin has shaped the evolutionary history of the Mediterranean fauna. We found strong support for a series of diversification events, range shifts, and extensive inter- and intraspecific gene flow taking place over the past six million years, connected to key events in the Mediterranean history (e.g. Cavazza and Wezel 2003; Hewitt 2011a).

Biogeographic and demographic analyses suggest that the MRCA of *P. muralis* stemmed from the Italian Peninsula. While molecular dating is fraught with difficulty, the timing of first lineage divergence corresponds well with the onset of the Messinian salinity crisis (Krijgsman et al. 1999, 2010; Duggen et al. 2003). This suggests that the divergence may have been triggered by declining sea levels and increased land connection, although aridification might have restricted suitable habitats (Jiménez-Moreno et al. 2010; Fiz-Palacios and Valcárcel 2013). The climatic changes that are associated with the end of the Messinian salinity crisis may have promoted allopatry and genetic differentiation that eventually resulted in the three distinct lineages on the Italian peninsula (CI, SI and SA).

The descendants of this ancient Italian assemblage appear to have followed an “Out-of-Italy” dispersal route to the Balkan and Iberian Peninsulas after the Messinian salinity crisis, a pattern that is consistent with dating of *Podarcis* fossil remains from Germany (Böttcher 2007). Both the RAD-Seq and WGS data supported six major genetic lineages for *P. muralis,* each with a distinct geographic distribution, which demonstrates the influence of the Iberian,

316 Italian, and Balkan peninsulas on diversification within the Mediterranean Basin (e.g.,
317 Schmitt 2007; Hewitt 2011a). These lineages were largely discordant with the phylogenetic
318 relationship based on mitochondrial genome data (Gassert et al. 2013; Salvi et al. 2013;
319 Jablonski et al. 2019). Populations from the Iberian Peninsula, France and western Germany
320 belong to the WE lineage (that can be further separated into north and south), and populations
321 from the Balkan Peninsula can be assigned to two closely related lineages in the Southern
322 (SB) and Northern (NB) Balkan. The contemporary distributions of the other three major
323 lineages, Southern Alps (SA), Central Italy (CI), and Southern Italy (SI) are essentially
324 restricted to Italy.

325 These lineages diverged between 6.24 and 2.5 Mya, but gene flow between lineages
326 appears to have been extensive. This explains inconsistencies in estimates of relationships
327 between *P. muralis* and other *Podarcis* species as well as between *P. muralis* lineages
328 (Andrade et al. 2019; Salvi et al. 2021; Yang et al. 2021). These levels of ancient gene flow
329 are consistent with studies of contemporary hybridization within *P. muralis*, which
330 demonstrate that significant parts of a genome can introgress under positive selection (While
331 et al. 2015; Yang et al. 2018; see also Schulte et al. 2012; Beninde et al. 2018). It is also
332 consistent with estimates of ancient gene flow between extant lineages of *Podarcis* (Caeiro-
333 Dias et al. 2020; Yang et al. 2021), which suggests that reproductive isolation evolves slowly
334 in wall lizards. Nevertheless, the hybrid zones between different *P. muralis* lineages appear to
335 be narrow and steep (e.g., Yang et al. 2020), and experimental studies suggest some degree of
336 pre-mating reproductive isolation (Heathcote et al. 2016; MacGregor et al. 2017). Partial
337 reproductive isolation is not unexpected given that some of the extant lineages have been
338 separated for about five million years. The many hybrid zones, involving lineages of different
339 age of divergence, makes *P. muralis* a useful system to study the evolution of pre- and post-
340 copulatory mechanisms of speciation (e.g., Heathcote et al. 2016; Yang et al. 2020).

341 The most extensive introgression between *P. muralis* lineages involved the Southern Italy
342 lineage, and explains the ambiguity with respect to its placement in the phylogenetic analysis.
343 The SI lineage genome turned out to be a mosaic, comprising of roughly one third of alleles
344 from the CI lineage in central Italy and two thirds of alleles from an early-diverged ancient
345 lineage. We further revealed that this ancient lineage has contributed a substantial amount of
346 genetic material to the MRCA of several extant lineages of *P. muralis*. These results
347 demonstrate the value in complementing phylogenetic trees with introgression analyses of
348 genomic data for detecting cryptic events in evolutionary histories (Than and Nakhleh 2009;
349 Ottenburghs et al. 2016, 2017).

350 The WGS data also revealed that some lineages of *P. muralis* experienced ancient
351 introgression from other *Podarcis* species. The WE lineage has received a substantial part of

its genome from the MRCA of the Iberian species group (i.e., the *P. hispanicus* complex; see also Yang et al. 2021). There was also evidence for introgression from *P. siculus* into the SA and CI lineages on the Italian peninsula, although the signal was weak. More extensive genomic data is needed to establish if hybridization between *P. muralis* and local sympatric species has been more widespread and the extent to which is it ongoing.

Overall, the genetic structure revealed by genomic data is strongly discordant with the sub-species division of traditional taxonomy based on morphology (Gruschwitz and Böhme 1986; Biaggini et al. 2011) and molecular phylogenies based on mtDNA (Bellati et al. 2011; Gassert et al. 2013; Salvi et al. 2013; this study). Mito-nuclear discordance is common in nature (e.g. Zink and Barrowclough 2008; Toews and Brelsford 2012; Ivanov et al. 2018), and can result from several different processes, including incomplete lineage sorting (Firneno et al. 2020), introgression (Phuong et al. 2017; Ivanov et al. 2018), and sex-biased dispersal (Dai et al. 2013).

Some instances of mito-nuclear discordance in *P. muralis* are likely the result of introgression. This is particularly well illustrated by the NB, SA and WE lineages, where genomic data support exchange of mtDNA through introgression events between the three lineages, as well as with the MRCA of extant Iberian *Podarcis* species (Yang et al. 2021). Other instances of mito-nuclear discordance are better explained by incomplete lineage sorting. For example, the individual from San Remo, situated on the Italian coast close to the border to France, belonged to the SA lineage according to nuclear data but formed the sister clade to all *P. muralis* in the mitochondrial phylogeny. Since there was no signal of mtDNA or nuclear genomic introgression from closely related species, or evidence of introgression from a "ghost lineage", this discordance appears to be a remnant of an ancient mtDNA that persisted throughout the evolutionary history of the SA lineage. Judging from the evidence for mito-nuclear discordance in other animals (e.g., Firneno et al. 2020), situations like these are probably not unusual but alternative hypotheses can be difficult to rule out (Toews & Brelsford 2012). Indeed, the genetic structure of *P. muralis* in this geographic region (south-western arc of the Alps) is poorly studied, and it is possible that more extensive sampling will reveal nuclear genomic signatures of an extinct lineage. Our data also supported a divergent mtDNA clade on the island of Elba (Bellati et al. 2011), but there was no evidence that this reflects a deep genetic divergence with mainland CI populations or an introgression event. Thus, incomplete lineage sorting of mtDNA haplotypes during or following isolation on Elba may best explain this discordance.

In the Mediterranean Basin, the Pliocene-Pleistocene climatic oscillations have been considered to play a key role for the current patterns of biodiversity and biogeography of animal species (reviewed in Hewitt 2000; Hewitt 2004). In particular, the genetic structure of

388 many species has been explained by cycles of glacial contraction to refugia in southern
389 peninsulas and inter-glacial expansion to northern regions (Hewitt 1996, 2000, 2004; Provan
390 & Bennet 2008; Taberlet et al., 1998). Our results are consistent with this hypothesis, but
391 suggest that all the major extant lineages were already present at the onset of the Quaternary
392 climatic oscillations. Glacial cycles therefore appear to have played a less important role in
393 initiating lineage divergence than might be assumed (reviewed in Hewitt 2004, 2011a), but
394 nevertheless have been crucial for dictating the subsequent dynamics of these lineages.
395 Indeed, our demographic simulations indicated that all lineages were affected by glacial
396 cycles, including a population expansion (ca. 0.7 Mya) during the Günz complex and a severe
397 decline (ca. 0.2 - 0.1 Mya) during the Riss glaciation. However, the inferred population
398 dynamics differed somewhat between lineages. This suggests that lineages may have largely
399 persisted in distinct regions during glacial and inter-glacial periods, including in northern
400 refugia in France and eastern Europe, thereby promoting further genetic differentiation and
401 reproductive isolation. The presence of several refugia within each of these regions might
402 explain the high number of low-divergent mitochondrial sub-lineages found in Mediterranean
403 and extra-Mediterranean regions (Salvi et al., 2013; Jablonski et al. 2019; see also fig. 1D). A
404 similar pattern of mtDNA lineages has been found for other Mediterranean taxa, for example
405 butterflies (Dincă et al. 2019; Hinojosa et al. 2019) and slow worms (Gvoždík et al. 2013).
406 Allopatric isolation during glaciation may have also promoted reproductive isolation, and
407 therefore contributed to the persistence of lineages following secondary contact, and explain
408 the location and apparent stability of contemporary hybrid zones (e.g., between the SA and CI
409 lineages; Yang et al., 2018, 2020).

410 In summary, the range-wide genomic approach employed in our study allowed a
411 disentangling of the evolutionary history of a broadly distributed Mediterranean lizard species
412 in unprecedented detail. We reveal an "Out-of-Italy" origin of the species about five Mya,
413 roughly coinciding with the end of the Messinian salinity crisis. The species diversified into
414 distinct lineages soon after its expansion from the Italian peninsula. The genomes of these
415 lineages carries the signature of several major introgression events and of extinct lineages as
416 well as the demographic imprints left by range expansions and contractions associated with
417 glacial cycles.

418

419 **Materials and methods**

420 **Samples and data collection**

421 We collected samples from a total of 55 locations, covering the previously suggested genetic
422 structure within *P. muralis* (fig. 1A; Gassert et al. 2013; Salvi et al. 2013). Detailed
423 information for these samples is provided in supplementary table S1, and the collection

permits are given in supplementary table S6. For each sample, we extracted total genomic DNA using the DNeasy blood and tissue kit (Qiagen, USA). We genotyped samples from all sites using RAD-Seq (deposited in NCBI Short Reads Archive [SRA] with accession number PRJNA486080). The RAD-Seq libraries were prepared following the protocol in Peterson et al. (2012) with modifications described in Yang et al. (2018). In addition, we conducted whole genome sequencing (WGS) for 16 individuals (supplementary table S2; supplementary fig. S1) with insert size of 300-500 bp on the Illumina HiSeq X platform by NOVOGENE Ltd. (Hong Kong). One individual of *P. bocagei*, *P. siculus*, and *P. tiliguerta* served as outgroups for all analyses and, in addition, *Archaeolacerta bedriagae* was included as outgroup for analyses involving WGS data. We also included WGS of 13 additional *Podarcis* species/lineages (accessible in NCBI under the accession number PRJNA715201; supplementary table S3), representing all species groups in *Podarcis* according to Yang et al. (2021), to investigate the gene flow between *P. muralis* and other *Podarcis* species.

Data processing for sequencing data

We used STACKS (version 2.4; Catchen et al. 2011; Rochette et al. 2019) to process RAD-Seq reads and infer single nucleotide variants (SNVs) for each individual. At first, the "process_radtags" module was used to remove reads with low-quality scores (Phred score < 30), ambiguous base calls, or incomplete barcode or restriction site. Clean reads were mapped to the genome of *P. muralis* (Andrade et al. 2019) using BWA (Li and Durbin 2009). We used sorted bam files as input for the reference-based STACKS pipeline that contains modules "gstacks" and "populations" to estimate SNVs using a Marukilow model (Maruki and Lynch 2017). We also aligned WGS reads to the *P. muralis* genome using BWA (Li and Durbin 2009). We called SNVs and short indel variants using the GATK best practice workflow (DePristo et al. 2011). Only SNVs from autosomes were used in the following phylogeographic analyses.

Mitochondrial genome

We assembled the mitochondrial genomes from WGS reads using NOVOPasty (Dierckxsens et al. 2017). The mitochondrial genome of *P. muralis* (accession FJ460597 from MitoZoa; D'Onorio et al. 2012) was set as a starting reference. A total of 6 Gb sequence reads from each sample were randomly extracted for the baiting and iterative mapping with default parameters. We aligned mitochondrial DNA (mtDNA) sequences using MUSCLE v3.8.31 (Edgar 2004). We excluded all ambiguous regions from the analyses to avoid false hypotheses of orthology.

Population structure analysis

We inferred the genetic relationship between the samples based on genotypes from RAD-Seq data. A principal component analysis (PCA) was conducted in Plink (version 1.9; Chang et al. 2015) based on pairwise genetic distance. In addition, the population structure was inferred assuming different numbers of clusters (K) from 1 to 15 in ADMIXTURE (version 1.3.0; Alexander et al. 2009). We used 10-fold cross-validation (CV) to compare different numbers of clusters, in which the lowest CV value indicates the most likely number of clusters.

Phylogenetic analysis

We reconstructed phylogenetic trees with maximum likelihood (ML) inference using IQ-TREE (Nguyen et al. 2015). We concatenated all SNVs generated from RAD-Seq dataset and inferred the phylogenies under a GTR+ASC model and 1,000 iterations of bootstrap replicates. We excluded 12 samples that showed admixed patterns in ADMIXTURE analysis (admixed ancestry > 10%). We also performed the same concatenation approach on SNVs for the 16 individuals with WGS data. An individual from KOP was excluded in the downstream analyses due to an admixed pattern. In addition, we used the "multispecies coalescent" approach implemented in ASTRAL-III (Zhang et al. 2018) to infer the phylogenetic relationships based on local trees of 200 kb fixed windows across the whole genome.

We inferred phylogenetic trees based on mitogenomic data implementing the highest-ranked model with 1,000 bootstrap replicates using IQ-TREE (Nguyen et al. 2015). We performed model selection for 1/2, and 3 codon positions for protein coding genes, and tRNA/rRNA genes with the setting "-m mf".

Gene flow analysis

We used D statistics (ABBA-BABA test; Patterson et al. 2012) to estimate the gene flow between both inter- and intraspecific lineages using "qpDstat" in AdmixTools (Patterson et al. 2012), using *A. bedriagae* as outgroup. We tested the significance level of D statistics through a block-jackknifing approach as implemented in AdmixTools, in which the z-score > 3.3 is considered significant (Patterson et al. 2012). First, we tested for interspecific introgression, between each of the major *P. muralis* lineages and geographically overlapping *Podarcis* species. We used the WGS data for 13 *Podarcis* species and lineages as "target_taxon" in the test (*muralis*_A, *muralis*_B, target_taxon, outgroup). Second, we also estimated the intraspecific gene flow between the major *P. muralis* lineages (i.e., muralis_A, muralis_B, muralis_C, outgroup).

We further conducted phylogenetic network analysis using phyloNet (Wen et al. 2008) to infer reticulation events between the *P. muralis* lineages (i.e., intraspecific introgression). *P. bocagei* was also included in this analysis to represent the Iberian *Podarcis* species, since we identified a strong signal of introgression in interspecific D statistics and in a previous study (Yang et al. 2021). We made use of high-quality local trees derived from 200 kb windows with mean bootstrap > 80, and extracted 1,000 random trees per run with 10^6 chain-length and 50% burn-in length in MCMC_gt module. We performed 100 independent runs, extracted all output networks with more than 50% posterior probability, and summarized the results by generating a correlation matrix of those networks based on Luay Nakhleh's metric of reduced phylogenetic network similarity (Edelman et al. 2019).

Based on the phylogenetic network, we used the program "qpGraph" from ADMIXTOOLS (Patterson et al. 2012) to fit the evolutionary history for all *P. muralis* lineages while accounting for introgression. qpGraph optimizes the fit of a proposed admixture graph in which each node can be descended either from a mixture of two other nodes, or from a single ancestral node. We calculated the proportion of introgressed alleles ted by f_4-ratio tests (Patterson et al. 2012). To identify the genomic regions with signatures of introgression, we calculated the absolute genetic divergence (Dxy) between lineage pairs, and the f_d statistics based on 200 kb windows across the genome. A significantly low Dxy and high f_d identify an introgressed genomic region tested by 1,000 permutations.

Divergence time estimation

We performed divergence time analysis between *P. muralis* lineages based on the WGS phylogeny using the MCMCtree program in the PAML package (Yang 2007). According to the fossil-calibrated Lacertini phylogeny of Garcia-Porta et al. (2019), the divergence time between *Podarcis* and other closely related clades was estimated at 37.55 million years ago (Mya), and the crown node of *Podarcis* species was at 18.60 Mya. We specified these calibration constraints with soft boundaries by using 0.025 tail probabilities above and below the limit in the built-in function of MCMCtree. To exclude the confounding effect of introgression events on topology and divergence time estimates, we only retained those genomic regions (346 windows with length of 200 kb) whose local trees were consistent with the consensus phylogeny. The independent rate model (clock = 2) was used to specify the rate priors for internal nodes. The MCMC run was first executed for 10^7 generations as burn-in and then sampled every 150 generations until a total of 100,000 samples were collected. We compared two MCMC runs using random seeds for convergence, which yielded similar results.

Biogeographic and demographic analysis

We used BioGeoBEARS (Matzke 2013) to infer the possible ancestral range of *P. muralis* and the number and type of biogeographical events dispersal leading to the distribution of extant lineages. We defined seven biogeographic areas covering the current distribution of this species: Iberian Peninsula, Western Europe, Southern Alps, Central Italy, Southern Italy, Northern Balkans and Southern Balkans. The time-calibrated phylogeny for *P. muralis* lineages was applied under three models - dispersal extinction cladogenesis (DEC), dispersal vicariance analysis (DIVALIKE), and Bayesian inference of historical biogeography for discrete areas (BAYAREALIKE). We selected the best-fitting model for comparisons among models based on AICc.

To reconstruct the detailed demographic history of each lineage, we applied the pairwise sequential Markovian coalescence (PSMC) model (Li and Durbin 2011) with the following parameters "-N25 -t15 -r5 -p4+25*2+4+6". We selected two to three sequenced individuals from each extant lineage. We excluded the SI lineage in this analysis due to its strikingly heterogeneous genome. On the basis of the time-calibrated phylogeny, we made use of a mutation rate of 1.98×10^{-9} mutations per site per year. The generation time for *P. muralis* was set to 2 years (Barbault and Mou 1988).

Data Availability

All sequence data generated in this study have been deposited in NCBI Sequence Reads Archive (SRA) with accession number PRJNA486080 (RAD-Seq data) and PRJNA715201 (whole genome data).

Acknowledgements

Computations were performed on resources provided by SNIC through the center for scientific and technical computing at Lund University under Project SNIC 2017/4-39, and Uppsala Multidisciplinary Center for Advanced Computational Science under Project SNIC 2017/5-8 and Uppstore 2018/2-18. This study was funded by the Swedish Research Council (E0446501), the Crafoord Foundation (20160911), the National Geographic Society, the British Ecological Society, the Royal Society of London, and a Wallenberg Academy Fellowship to TU, starting grants from the Swedish Research Council (2020-03650) and the European Research Council (948126) to NF, a LabEx TULIP grant (ANR-10-LABX-41) to FA and the Slovak Research and Development Agency (APVV-19-0076) to DJab.

565 **References**

566 Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in
567 unrelated individuals. *Genome Res*, 19:1655-64.

568 Andrade PC et al. 2019. Regulatory changes in pterin and carotenoid genes underlie balanced
569 color polymorphisms in the wall lizard. *Proc. Nat. Acad. Sci. USA*, 116: 5633-5642.

570 Avise JC. 2000. Phylogeography: the history and formation of species. *Harvard university*
571 *press*.

572 Barbault R, Mou YP. 1988. Population dynamics of the common wall lizard, Podarcis
573 muralis, in suthwestern France. *Herpetologica*, 44:38–47.

574 Bellati A, Pellitteri-Rosa D, Sacchi R, Nistri A, Galimberti A, Casiraghi M, Fasola M,
575 Galeotti P. 2011. Molecular survey of morphological subspecies reveals new
576 mitochondrial lineages in *Podarcis muralis* (Squamata: Lacertidae) from the Tuscan
577 Archipelago (Italy). *J Zool Syst Evol Res*, 49:240–250.

578 Beninde J, Feldmeier S, Veith M, Hochkirch A. 2018. Admixture of hybrid swarms of native
579 and introduced lizards in cities is determined by the cityscape structure and invasion
580 history. *Proc R Soc Lond B*, 285:20180143.

581 Biaggini M, Bazzoffi P, Gentile R, Corti C. 2011. Effectiveness of the GAEC cross
582 compliance standards Rational management of set aside, Grass strips to control soil
583 erosion and Vegetation buffers along watercourses on surface animal diversity and
584 biological quality of soil. *Ital J Agronom*, 12:e14.

585 Böttcher W. 2007. Fossile Amphibien und Reptilien in Baden-Württemberg. *Die Amphibien*
586 *und Reptilien Baden-Württembergs* (ed. by H. Laufer, K. Fritz and P. Sowig), pp. 62–76.
587 Ulmer-Verlag, Stuttgart.

588 Bryant et al. 2012. Inferring species trees directly from biallelic genetic markers: bypassing
589 gene trees in a full coalescent analysis. *Mol Biol Evol*, 29:1917-1932.

590 Buckley TR, Marske KA, Attanayake D. 2009. Identifying glacial refugia in a geographic
591 parthenogen using palaeoclimate modelling and phylogeography: the New Zealand stick
592 insect Argosarchus horridus (White). *Mol Ecol*, 18:4650-63.

593 Caeiro-Dias, G., Brelsford, A., Kaliontzopoulou, A., Meneses-Ribeiro, M., Crochet, P. A., &
594 Pinho, C. (2020). Variable levels of introgression between the endangered *Podarcis*
595 *carbonelli* and highly divergent congeneric species. *Heredity*, 126:463-476.

596 Carstens BC, Pelletier TA, Reid NM, Satler JD. 2013. How to fail at species delimitation. *Mol*
597 *Ecol*, 22:4369-83.

598 Catchen JM, Amores A, Hohenlohe P, Cresko W, and Postlethwait JH. 2011. Stacks: building
599 and genotyping loci de novo from short-read sequences. *G3-Genes Genom Genet*, 1:
600 171-182.

601 Cavazza W, Wezel FC. 2003. The Mediterranean region - a geological primer. *Episodes*,
602 26:160-168.

603 Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation
604 PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, 4:s13742-015.

605 Chen L, Qiu Q, Jiang Y, Wang K, Lin Z, Li Z, Bibi F, Yang Y, Wang J, Nie W, Su W, et al.
606 2019. Large-scale ruminant genome sequencing provides insights into their evolution and
607 distinct traits. *Science*, 364:eaav6202.

608 Cui R, Schumer M, Kruesi K, Walter R, Andolfatto P, Rosenthal GG. 2013. Phylogenomics
609 reveals extensive reticulate evolution in *Xiphophorus* fishes. *Evolution*, 67:2166-79.

610 Cutter AD, Payseur BA. 2013. Genomic signatures of selection at linked sites: unifying the
611 disparity among species. *Nat Rev Genet*, 14:262-74.

612 Dai C, Wang W, Lei F. 2013. Multilocus phylogeography (mitochondrial, autosomal and Z-
613 chromosomal loci) and genetic consequence of long-distance male dispersal in Black-
614 throated tits (*Aegithalos concinnus*). *Heredity*, 110:457-65.

615 Delsuc F, Brinkmann H, Philippe H. 2005. Phylogenomics and the reconstruction of the tree
616 of life. *Nat Rev Genet*, 6:361-75.

617 DePristo MA, et al. 2011. A framework for variation discovery and genotyping using next-
618 generation DNA sequencing data. *Nat Genet,* 43:491-498.

619 Dierckxsens N., et al. 2017. NOVOPlasty: De novo assembly of organelle genomes from
620 whole genome data. *Nucleic Acid Res,* 45:e18.

621 Dincă V, Lee KM, Vila R, Mutanen M. 2019. The conundrum of species delimitation: a
622 genomic perspective on a mitogenetically super-variable butterfly. *Proc R Soc Lond B*,
623 286: 20191311.

624 D'Onorio M. P., et al. 2012. MitoZoa 2.0: a database resource and search tools for
625 comparative and evolutionary analyses of mitochondrial genomes in Metazoa. *Nucleic*
626 *Acids Res*, 40: D1168-1172.

627 Duggen S, Hoernie K, van den Bogaard P, Rupke L, Morgan JP. 2003. Deep roots of the
628 Messian salinity crisis. *Nature,* 422:602-605.

629 Edelman NB, Frandsen PB, Miyagi M, Clavijo B, Davey J, Dikow RB, García-Accinelli G,
630 Van Belleghem SM, Patterson N, Neafsey DE, Challis R et al. 2019. Genomic
631 architecture and introgression shape a butterfly radiation. *Science*, 366:594-9.

632 Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high
633 throughput. *Nucleic Acids Res*, 32(5):1792-7.

634 Felsenstein J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the
635 author. Department of Genome Sciences, University of Washington, Seattle.

636 Firneno Jr TJ, O'Neill JR, Portik DM, Emery AH, Townsend JH, Fujita MK. 2020. Finding
637 complexity in complexes: Assessing the causes of mitonuclear discordance in a
638 problematic species complex of Mesoamerican toads. *Mol Ecol*, 29:3543-59.

639 Fiz-Palacios O, Valcárcel V. 2013. From Messinian crisis to Mediterranean climate: a
640 temporal gap of diversification recovered from multiple plant phylogenies. *Perspect*
641 *Plant Ecol, Evol Syst*, 15:130-7.

642 García-Castellanos D, Estrada F, Jiménez-Munt I, Gorini C, Fernàndez M, Vergés J, De
643 Vicente R. 2009. Catastrophic flood of the Mediterranean after the Messinian salinity
644 crisis. *Nature*, 462: 778-782.

645 Garcia-Porta J, Irisarri I, Kirchner M, Rodríguez A, Kirchhof S, Brown JL, MacLeod A,
646 Turner AP, Ahmadzadeh F, Albaladejo G, Crnobrnja-Isailovic J. 2019. Environmental
647 temperatures shape thermal physiology as well as diversification and genome-wide
648 substitution rates in lizards. *Nat Commun*, 10:1-2.

649 Gassert F, Schulte U, Husemann M, Ulrich W, Rödder D, Hochkirch A, Engel E, Meyer J,
650 Habel JC. 2013. From southern refugia to the northern range margin: genetic population
651 structure of the common wall lizard, *Podarcis muralis*. *J Biogeogr*, 40:1475-89.

652 Gentili R, Baroni C, Caccianiga M, Armiraglio S, Ghiani A, Citterio S. 2015. Potential warm-
653 stage microrefugia for alpine plants: Feedback between geomorphological and biological
654 processes. *Ecol Complex*, 21:87-99.

655 Giovannotti M, Nisi-Cerioni P, Caputo V. 2010. Mitochondrial DNA sequence analysis
656 reveals multiple Pleistocene glacial refugia for *Podarcis muralis* (Laurenti, 1768) in the
657 Italian Peninsula. *Ital J Zool*, 77:277–288.

658 Gruschwitz M, Böhme W. 1986. *Podarcis muralis* (Laurenti, 1768)—Mauereidechse. In:
659 Handbuch der Amphibien und Reptilien Europas. Bandll/2, Echsen (Sauria) III
660 (Lacertidae III; *Podarcis*). Aula-Verlag, Wiesbaden, pp 155–208.

661 Gvoždík V, Benkovský N, Crottini A, Bellati A, Moravec J, Romano A, Sacchi R, Jandzik D.
662 2013. An ancient lineage of slow worms, genus Anguis (Squamata: Anguidae), survived
663 in the Italian Peninsula. *Mol Phylonet Evol*, 69(3):1077-92.

664 Harris DJ, Arnold EN. 1999. Relationships of Wall Lizards, *Podarcis* (Reptilia: Lacertidae)
665 based on mitochondrial DNA sequences. *Copeia*, 3:749–754.

666 Heathcote RJ, While GM, MacGregor HE, Sciberras J, Leroy C, D'Ettorre P, Uller T. 2016.
667 Male behaviour drives assortative reproduction during the initial stage of secondary
668 contact. *J Evol Biol*, 29:1003-15.

669 Hewitt GM. 1996. Some genetic consequences of ice ages and their role in divergence and
670 speciation. *Biol J Linn Soc*, 58, 247–276.

671 Hewitt GM. 2000. The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907–913.

672 Hewitt GM. 2001. Speciation, hybrid zones and phylogeography – or seeing genes in space
673 and time. *Mol Ecol*, 10:537-549.

674 Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Phil.*
675 *Trans. R. Soc. Lond. B*, 359, 183–195.

676 Hewitt GM. 2011a. Mediterranean peninsulas: the evolution of hotspots. In Biodiversity
677 hotspots. *Distribution and protection of conservation priority areas*. Eds. FE Zachos, JC
678 Habel, Springer, Berlin, Heidelberg, pp. 123-147.

679 Hewitt GM. 2011b. Quaternary phylogeography: the roots of hybrid zones. *Genetica,*
680 139:617-638.

681 Hinojosa JC, Koubínová D, Szenteczki MA, Pitteloud C, Dincă V, Alvarez N, Vila R. 2019. A
682 mirage of cryptic species: Genomics uncover striking mitonuclear discordance in the
683 butterfly *Thymelicus sylvestris*. *Mol Ecol,* 28: 3857-3868.

684 Ivanov V, Lee KM, Mutanen M. 2018. Mitonuclear discordance in wolf spiders: Genomic
685 evidence for species integrity and introgression. *Mol Ecol*, 27:1681-95.

686 Jablonski D, Gvoždík V, Choleva L, Jandzik D, Moravec J, Mačát Z, Veselý M. 2019. Tracing
687 the maternal origin of the common wall lizard (*Podarcis muralis*) on the northern range
688 margin in Central Europe. *Mitochondrion*, 46:149-57.

689 Jiménez-Moreno G, Fauquette S, Suc JP. 2010. Miocene to Pliocene vegetation reconstruction
690 and climate estimates in the Iberian Peninsula from pollen data. *Rev Palaeobot Palynol*,
691 162:403-15.

692 Kalyaanamoorthy B. Q., et al. 2017. ModelFinder: Fast model selection for accurate
693 phylogenetic estimates. *Nat Method*, 14:587-589.

694 Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS. 1999. Chronology, causes and
695 progression of the Messinian salinity crisis. *Nature*, 400:652-655.

696 Krijgsman W, Stoica M, Vasiliev I, Popov VV. 2010. Rise and fall of the Paratethys Sea
697 during the Messinian Salinity Crisis. *Earth Planet Sci Lett,* 290:183-191.

698 Li H., Durbin R. 2011. Inference of human population history from whole genome sequence
699 of a single individual. *Nature,* 475:493-496.

MacGregor HE, While GM, Barrett J, Pérez i de Lanuza G, Carazo P, Michaelides S, Uller T. 2017. Experimental contact zones reveal causes and targets of sexual selection in hybridizing lizards. *Funct Ecol*, 31:742-52.

Mairal M, Sanmartín I, Herrero A, Pokorny L, Vargas P, Aldasoro JJ, Alarcón M. 2017. Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the Eastern Afromontane biodiversity hotspot. *Sci Rep*, 7:1-3.

Malinsky M, Svardal H, Tyers AM, Miska EA, Genner MJ, Turner GF, Durbin R. 2018. Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nat Ecol Evol*, 2:1940-55.

Mallet J, Besansky N, Hahn MW. 2016. How reticulated are species? *BioEssays*, 38:140-9.

Martin SH, Davey JW and Jiggins CD. 2015. Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Mol Biol Evol,* 32:244-257.

Maruki T, Lynch M. 2017. Genotype calling from population-genomic sequencing data. *G3: Genes Genome Genet*, 7:1393-404.

Matzke N. 2013. BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts. University of California, Berkeley, Berkeley, CA.

McCluskey BM, Postlethwait JH. 2015. Phylogeny of zebrafish, a "model species," within *Danio*, a "model genus". *Mol Biol Evol*, 32:635-52.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Mol Phylogenet Evol*, 66:526-38.

Médail F, Diadema K. 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *J Biogeogr*, 36:1333-45.

Médail F, Quezel P. 1997. Hot-spots analysis for conservation of plant biodiversity in the Mediterranean Basin. *Ann Missouri Botanic Garden*, 1:112-27.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature,* 403:853-8.

Naciri Y, Linder HP. 2015. Species delimitation and relationships: the dance of the seven veils. *Taxon*, 64:3-16.

Nguyen L. T., et al. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol Biol Evol*, 32:268-274.

Ottenburghs J, Megens HJ, Kraus RH, Madsen O, van Hooft P, van Wieren SE, Crooijmans RP, Ydenberg RC, Groenen MA, Prins HH. 2016. A tree of geese: A phylogenomic perspective on the evolutionary history of True Geese. *Mol Phylogenet Evol*, 101:303-13.

734 Ottenburghs J, Megens HJ, Kraus RH, Van Hooft P, van Wieren SE, Crooijmans RP,
735 Ydenberg RC, Groenen MA, Prins HH. 2017. A history of hybrids? Genomic patterns of
736 introgression in the True Geese. *BMC Evol Biol*, 17:1-4.

737 Patterson N., et al. 2012. Ancient admixture in human history. *Genetics,* 192:1065-1093.

738 Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an
739 inexpensive method for de novo SNP discovery and genotyping in model and non-model
740 species. *PLoS One*, 7:e37135.

741 Phuong MA, Bi K, Moritz C. 2017. Range instability leads to cytonuclear discordance in a
742 morphologically cryptic ground squirrel species complex. *Mol Ecol*, 26:4743-55.

743 Pollard DA, Iyer VN, Moses AM, Eisen MB. 2006. Widespread discordance of gene trees
744 with species tree in Drosophila: evidence for incomplete lineage sorting. *PLoS Genet*,
745 2:e173.

746 Postigo Mijarra JM, Barrón E, Gómez Manzaneque F, Morla C. 2009. Floristic changes in the
747 Iberian Peninsula and Balearic Islands (south-west Europe) during the Cenozoic. *J*
748 *Biogeogr*, 36:2025-43.

749 Poulakakis N, Lymberakis P, Valakos E, Zouros E, Mylonas M. 2005. Phylogenetic
750 relationships and biogeography of *Podarcis* species from the Balkan Peninsula, by
751 bayesian and maximum likelihood analyses of mitochondrial DNA sequences. *Mol*
752 *Phylogenet Evol*, 37:845-857.

753 Provan J, Bennet KD. 2008. Phylogeographic insights into cryptic glacial refugia. *Trend Ecol*
754 *Evol*, 23:564–571.

755 Psonis N, Antoniou A, Karameta E, Darriba D, Stamatakis A, Lymberakis P, Poulakakis N.
756 2021. The wall lizards of the Balkan peninsula: Tackling questions at the interface of
757 phylogenomics and population genomics. *Mol Phylogenet Evol*, 159:107121.

758 Ree RH, Sanmartin I. 2018. Conceptual and statistical problems with the DEC+J model of
759 founder-event speciation and it comparison with DEC via model selection. *J Biogeogr*,
760 45:741-749.

761 Rochette NC, Rivera-Colon AG, and Catchen JM. 2019. Stacks 2: Analytical methods for
762 paired-end sequencing improve RADseq-based population genomics. *Mol Ecol,* 28:4737-
763 4754.

764 Runemark A, Trier CN, Eroukhmanoff F, Hermansen JS, Matschiner M, Ravinet M, Elgvin
765 TO, Sætre GP. 2018. Variation and constraints in hybrid genome formation. *Nat Ecol*
766 *Evol*, 2:549-556.

767 Salvi D, Harris DJ, Kaliontzopoulou A, Carretero MA, Pinho C. 2013. Persistence across
768 Pleistocene ice ages in Mediterranean and extra-Mediterranean refugia: phylogeographic
769 insights from the common wall lizard. *BMC Evol Biol*, 13:1-8.

770 Salvi D, Pinho C, Mendes J, Harris DJ, 2021. Fossil-calibrated time tree of Podarcis wall
771 lizards provides limited support for biogeographic calibration models. *Mol Phylogenet*
772 *Evol*, 161:107169.

773 Stadler T. 2011. Mammalian phylogeny reveals recent diversification rate shifts. *Proc Natl*
774 *Acad Sci USA*, 108:6187-6192.

775 Schmitt T. 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial
776 trends. *Front Zool*, 4:1-13.

777 Schulte U, Thiesmeier BU, Mayer WE, Schweiger SI. 2018. Allochthone Vorkommen der
778 Mauereidechse (*Podarcis muralis*) in Deutschland. *Z Feldherpetol*, 15:138-56.

779 Schulte U, Veith M, Hochkirch A. 2012. Rapid genetic assimilation of native wall lizard
780 populations (Podarcis muralis) through extensive hybridization with introduced lineages.
781 *Mol Ecol*, 21:4313-26.

782 Speybroeck J, Beukema W, Dufresnes C, Fritz U, Jablonski D, Lymberakis P, Martínez-
783 Solano I, Razzetti E, Vamberger M, Vences M, et al. 2020. Species list of the European
784 herpetofauna – 2020 update by the Taxonomic Committee of the Societas Europaea
785 Herpetologica. *Amphibia-Reptilia,* 41:139-189.

786 Taberlet P, Fumagalli L, Wust-Saucy A-G et al. 1998. Comparative phylogeography and
787 postglacial colonization routes in Europe. *Mol Ecol*, 7:453–464.

788 Than C, Nakhleh L. 2009. Species tree inference by minimizing deep coalescences. *PLoS*
789 *Comput Biol*, 5:e1000501.

790 Firneno Jr TJ, O'Neill JR, Portik DM, Emery AH, Townsend JH, Fujita MK. 2020. Finding
791 complexity in complexes: Assessing the causes of mitonuclear discordance in a
792 problematic species complex of Mesoamerican toads. *Mol Ecol*, 29(18): 3543-3559.

793 Thompson JD. 2005. Plant evolution in the Mediterranean. *Oxford University Press on*
794 *Demand*.

795 Toews DP, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in
796 animals. *Mol Ecol*, 21:3907-30.

797 Tzedakis PC. 2007. Seven ambiguities in the Mediterranean palaeoenvironmental narrative.
798 *Quat Sci Rev*, 26:2042-66.

799 While GM, Michaelides S, Heathcote RJ, MacGregor HE, Zajac N, Beninde J, Carazo P,
800 Pérez i de Lanuza G, Sacchi R, Zuffi MA, Horváthová T. 2015. Sexual selection drives
801 asymmetric introgression in wall lizards. *Ecol Lett*, 18:1366-75.

802 Wiens JJ. 2004. Speciation and ecology revisited: phylogenetic niche conservatism and the
803 origin of species. *Evolution*, 58:193-7.

804 Wen D, et al. 2008. Inferring phylogenetic networks using PhyloNet. *Syst Biol,* 67:735-740.

805 Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*,
806 24:1586-91.

807 Yang W, While GM, Laakkonen H, Sacchi R, Zuffi MA, Scali S, Salvi D, Uller T. 2018.
808 Genomic evidence for asymmetric introgression by sexual selection in the common wall
809 lizard. *Mol Ecol*, 27:4213-24.

810 Yang W, Feiner N, Laakkonen H, Sacchi R, Zuffi MA, Scali S, While GM, Uller T. 2020.
811 Spatial variation in gene flow across a hybrid zone reveals causes of reproductive
812 isolation and asymmetric introgression in wall lizards. *Evolution*, 74:1289-300.

813 Yang W, Feiner N, Pinho C, Kaliontzopoulou A, While GM, Harris DJ, Salvi D, and Uller T.
814 2021. Extensive introgression and mosaic genomes of endemic Mediterranean lizards.
815 *Nat Commun,* 12(2762): 1-8.

816 Zhang C., et al. 2018. ASTRAL-III: polynomial time species tree reconstruction from
817 partially resolved gene trees. *BMC Bioinformatics,* 19:153.

818 Zink RM, Barrowclough GF. 2008. Mitochondrial DNA under siege in avian phylogeography.
819 *Mol Ecol*, 17:2107-21.

Figure legends

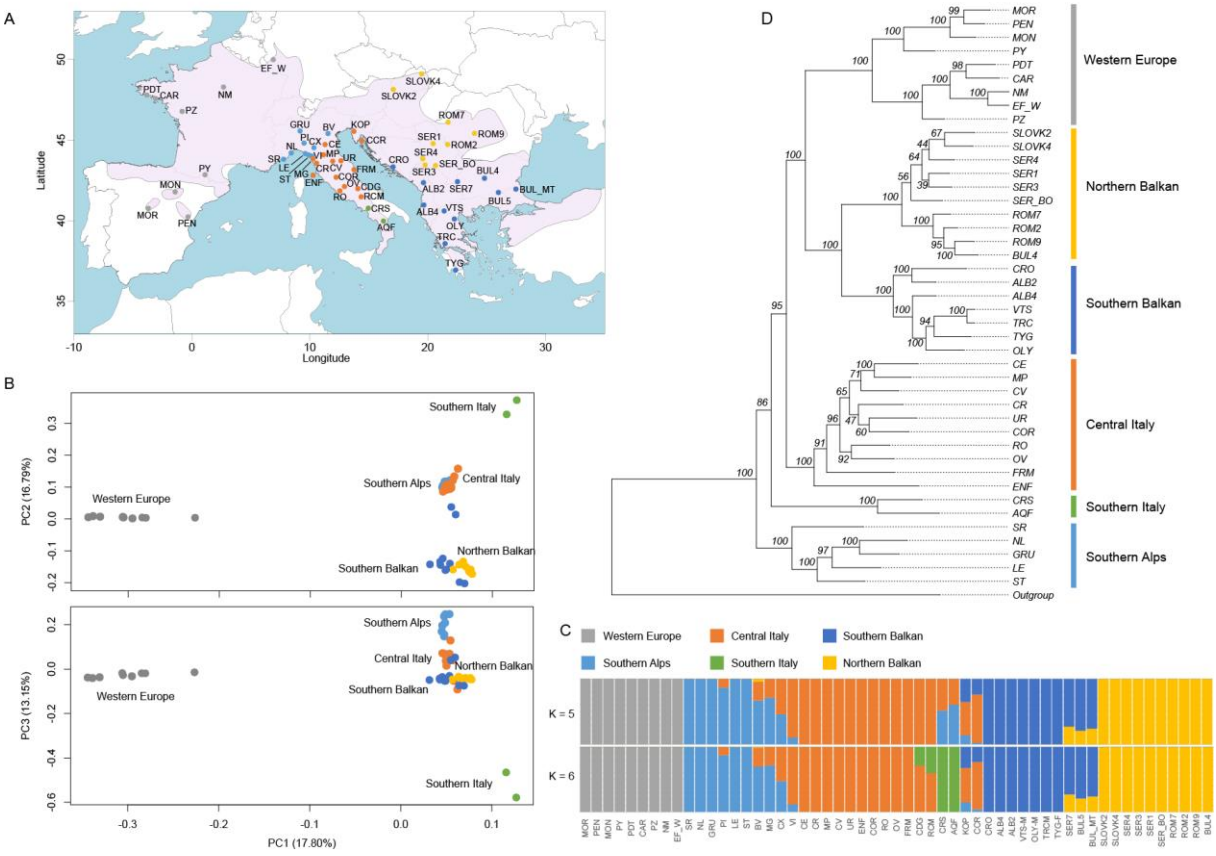
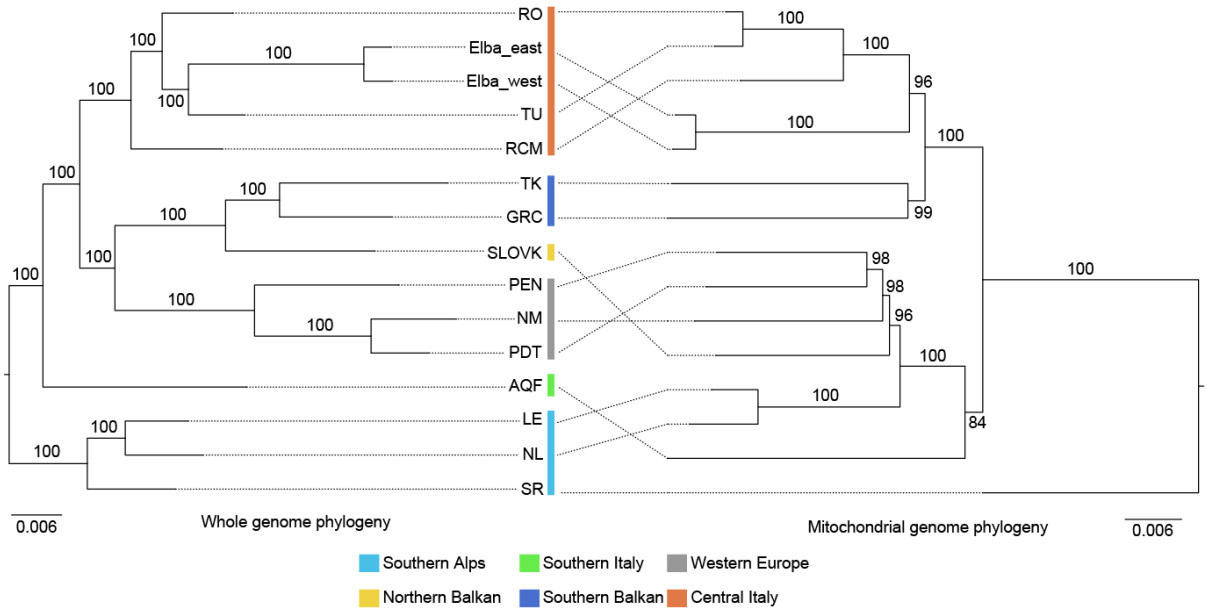


Fig. 1 Population genetic and phylogenetic analyses of *P. muralis*. (A) The sampling localities for the RAD-Seq part of this study. The pink area represents the distribution range of *P. muralis.* For sampling locations of WGS samples, see supplementary fig. S1. (B) PCA plots of genetic distance for all 55 individuals based on RAD-Seq data. (C) Admixture clustering of individuals into five and six groups (K). The proportion of each individual's genome assigned to each cluster is shown by the length of the colored segments. (D) Maximum likelihood phylogeny inferred based on RAD-Seq data. The numbers above branches indicate the bootstrap values for each node. For all panels, genetic lineages are color-coded consistently.



833

834

835 **Fig. 2 Mito-nuclear discordance for *P. muralis* lineages**. The discrepancies between
836 phylogenies based on whole-genome sequencing data (left) and mitochondrial genome data
837 (right). Bootstrap values are provided for all nodes. Colors refer to genetic lineages.

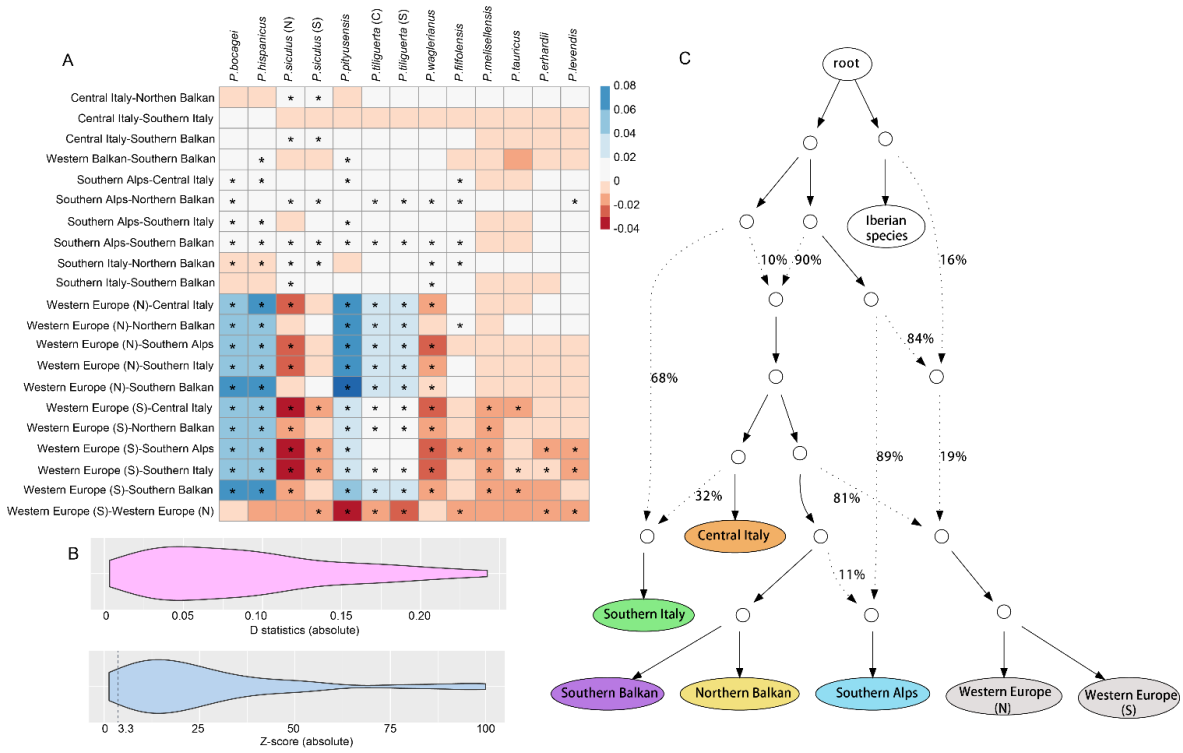


Fig. 3 Analyses of gene flow for *P. muralis* lineages. (A) Interspecific introgression analyses between *P. muralis* and other *Podarcis* species using four-taxon D statistics with the test (muralis_A, muralis_B, target_taxon, outgroup). The lineage pairs of *P. muralis* are listed on the Y axis, and the targeted non-*muralis* species are listed on the X axis. The coloration of each square represents the D statistics, and the asterisk (*) indicates significant deviations from neutrality based on z-scores. (B) Distribution of D statistics and z-scores of intraspecific introgression analysis between *P. muralis* lineages. The result showed that most tests significantly deviated from neutrality. We list detailed information in supplementary table S4. (C) Admixture graph of *P. muralis* generated by qpGraph. Solid lines with arrows indicate tree-like evolution, whereas dash lines with arrows indicate admixture events. The numbers next to branches represent the proportion of alleles from a parental nodes.

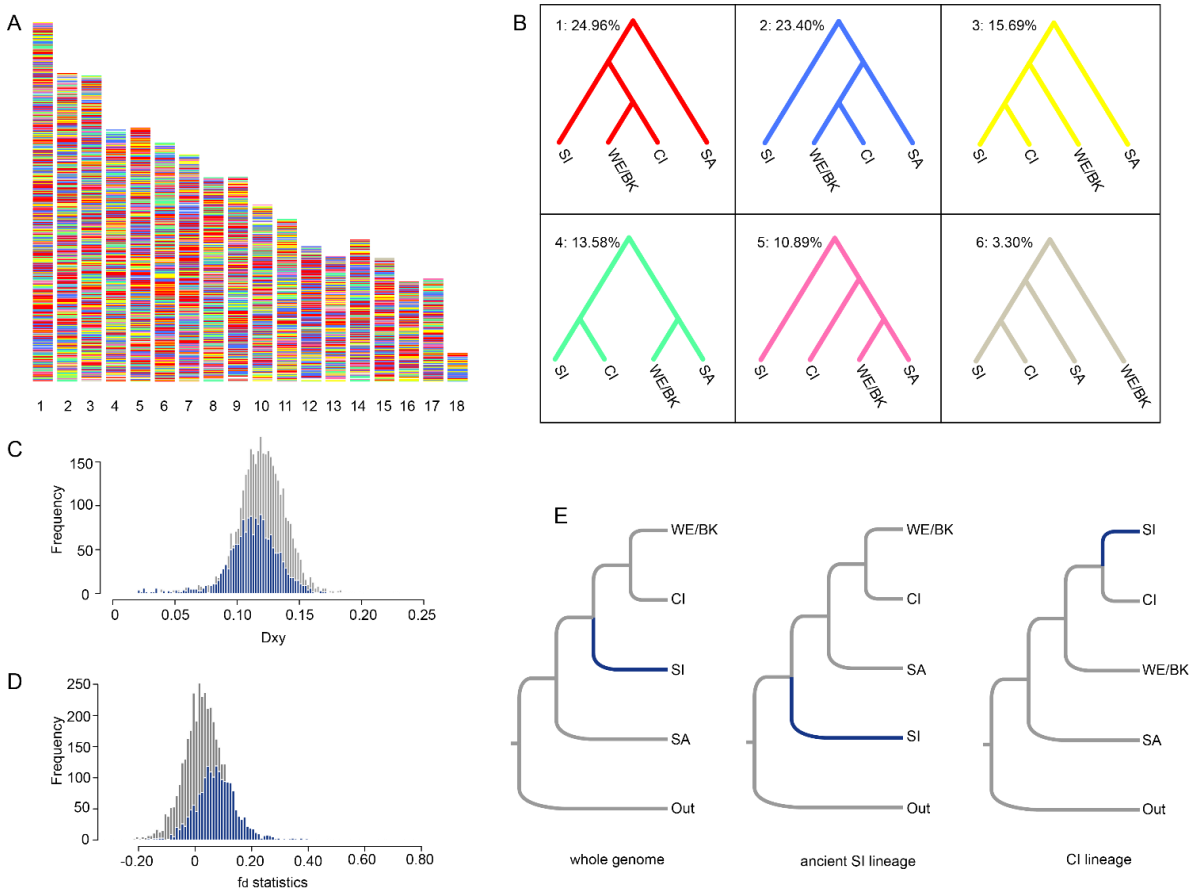
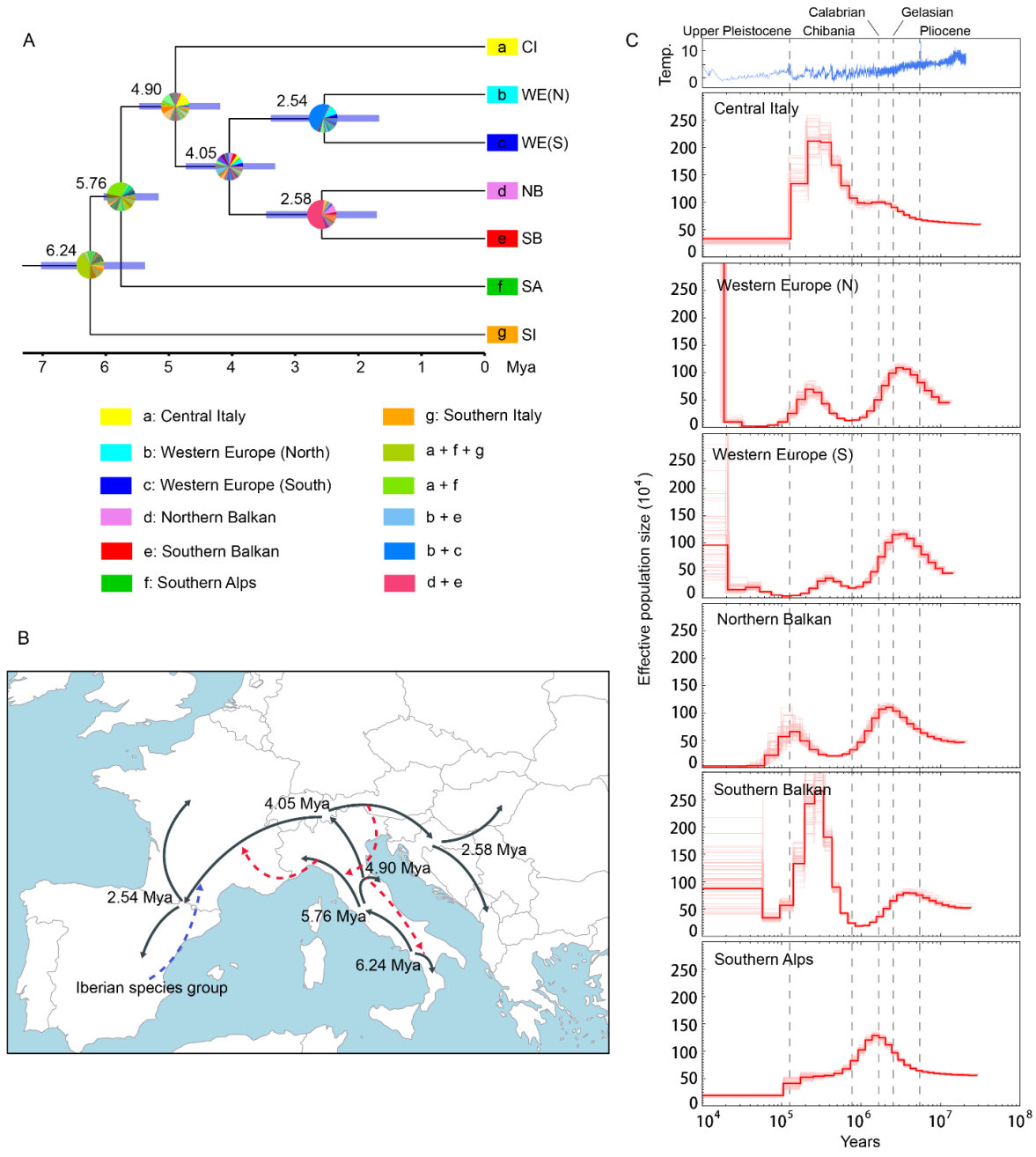


Fig. 4 The ancestry and phylogenetic position of the Southern Italy lineage. (A) The distribution of local tree topologies across the genome. (B) The six most common topologies with the corresponding coloration as in panel (A). The value on the top left is the percentage of all 200 kb windows supporting the specific topology. (C) The distribution of Dxy between the two lineages for the CI-ancestry part of the genome (blue) and the rest of the genome (grey). (D) The distribution of f_d statistics between the two lineages for the CI-ancestry genome (blue) and the rest of the genome (grey). (E) Multispecies coalescent tree topologies for distinct portions of the genome: whole genome, the ancient SI lineage part and the CI-ancestry part of the genome.



862

863 **Fig. 5 Biogeography and demographic dynamics for *P. muralis***. (A) Time-calibrated
864 phylogeny for *P. muralis* lineages with ancestral area reconstructions. The numbers indicate
865 the estimated ages for each node. The blue bars represent the confidence intervals of
866 divergence times. The colored squares at tip nodes represent the current distribution ranges of
867 lineages in seven biogeographic regions, and the pie charts indicate the proportional posterior
868 probability of ancestral ranges inferred by the best-fitting model DIVALIKE. (B) Illustration
869 of the inferred biogeographic history of *P. muralis* on a map of contemporary Europe. The
870 solid lines with arrows indicate the tree-like divergence and dispersal. Dashed lines with
871 arrows indicate interspecific introgression (blue) or intraspecific introgression (red). (C)
872 Demographic history of five major *P. muralis* lineages (note that both Northern and Southern

873 WE are shown in separate panels). Effective population size over time was estimated using
874 the pairwise sequential Markovian coalescent model. The red lines represent the population
875 dynamics, and the light red lines represent the results of 100 bootstrap replicates. The dash
876 lines indicate the start of geological periods. The top panel shows the dynamics of global
877 temperature.