From Gondwana to the Yellow Sea, evolutionary diversifications of true toads *Bufo* sp. in the Eastern Palearctic and a revisit of species boundaries for Asian lineages

Siti N. Othman¹, Spartak N. Litvinchuk², Irina Maslova³, Hollis Dahn⁴, Kevin R. Messenger⁵, Desiree Andersen¹, Michael J. Jowers⁶, Yosuke Kojima⁷, Dmitry V. Skorinov², Kiyomi Yasumiba⁸, Ming-Feng Chuang⁹, Yi-Huey Chen¹⁰, Yoonhyuk Bae¹, Jennifer Hoti^{1,11}, Yikweon Jang^{1*}, Amaël Borzée^{12*}

¹ Department of Life Sciences and Division of EcoScience, Ewha Womans University, Seoul, Republic of Korea

² Institute of Cytology, Russian Academy of Sciences, Tikhoretsky pr. 4, St. Petersburg 194064, Russian Federation

³ Federal Scientific Center of the East Asia Terrestrial Biodiversity Far Eastern Branch of Russian Academy of Sciences, Vladivostok, 690022, Russia

⁴ Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, M5S 1A1, Canada

⁵ Herpetology and Applied Conservation Laboratory, College of Biology and the Environment, Nanjing Forestry University, 159 Longpan Rd, Nanjing, Jiangsu, China

⁶ CIBIO/InBIO (Centro de Investigação em Biodiversidade e Recursos Genéticos), Universidade do Porto, Campus Agrário De Vairão, 4485-661, Vairão, Portugal

⁷ Graduate School of Human and Environmental Studies, Kyoto University, Sakyo, Kyoto, Japan

⁸ Tokyo University of Agriculture and Technology, 3-5-8 Saiwai-cho Fuchu, Tokyo, Japan

⁹ Department of Life Sciences and Research Center for Global Change Biology, National Chung Hsing University, Taichung 402, Taiwan

¹⁰ Department of Life Science, Chinese Culture University, Taipei, Taiwan

¹¹ Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy

¹² Laboratory of Animal Behaviour and Conservation, College of Biology and the Environment, Nanjing Forestry University, 159 Longpan Rd, Nanjing, Jiangsu, China

*Corresponding author:

E-mail: amaelborzee@gmail.com (AB)

E-mail: jangy@ewha.ac.kr (YJ)

Impact Statement

A robust framework of integrative phylogeography and advanced taxonomy to resolve the historical biogeography and species boundaries of true toads across the Eastern Palearctic.

Major subject areas, keywords, and research organism

Fossilised birth death calibration, East Palearctic, central Asian desertification, continental collision, ice age refugia, taxonomic resolution, *Bufo* sp.

Abstract

Taxa with vast distribution ranges often display unresolved phylogeographic structures and unclear taxonomic boundaries resulting into hidden diversity. This hypothesis-driven study reveals the evolutionary history of Bufonidae, covering the phylogeographic patterns found in Holarctic bufonids from the West Gondwana to the phylogenetic taxonomy of Asiatic true toads in the Eastern Palearctic. We used an integrative approach relying on fossilised birth-death calibrations, population dynamic, gene-flow, species distribution and species delimitation modelling to resolve the biogeography of the clade and highlight cryptic lineages. We verified the near-simultaneous Miocene radiations within Western and Eastern Palearctic *Bufo*, c. 14.49 – 10.00 Mya, temporally matching with the maximum dust outflows in Central Asian deserts. Contrary to earlier studies, we demonstrated that the combined impacts of long dispersal and ice-age refugia equally contributed to the current genetic structure of *Bufo* in East Asia. Our findings reveal a climate-driven adaptation in septentrional Eastern Asian *Bufo*, explained its range shifts towards northern latitudes. We resolve species boundaries within the Eastern Palearctic *Bufo*, and redefine the taxonomic and conservation units of the northeastern species: *B. sachalinensis* and its subspecies.

Introduction

Evolutionary history is critical in explaining the diversification within and between species, and the dynamics between species and environments in ecological zones (Ricklefs, 2006). Factors such as past geological events, natural dispersion, and anthropogenic changes (Dufresnes et al., 2020; Othman et al., 2020) can drive and induce different evolutionary scenarios such as physiological adaptation, genetic variability and phenotypic divergence (Luquet et al., 2015). However, understanding phylogeography and the processes contributing to genetic divergence can be more challenging in taxonomic groups distributed across vast ranges. This difficulty results from the ecological responses to variability in environmental conditions and geographical features (Zhao and Yu, 2012). Thus, integrating macroevolution of species groups and microevolution within populations is necessary to understand the evolutionary mechanisms of complex study systems (Li et al., 2018).

Amphibians are an excellent model for studying the factors affecting distribution and evolution. Specifically, true toads in the family Bufonidae (bufonids) are well suited to this area of study due to their high species diversity (Rojas et al., 2018) and adaptive response to past climate change that led to an evolutionary-recent global radiation (Van Bocxlaer et al., 2010). Numerous phylogeographic and systematic studies have characterised the evolutionary diversifications of bufonids, notably in the Holarctic (Dufresnes et al., 2019; Garcia-Porta et al., 2012; Pauly et al., 2004; Recuero et al., 2012; Stöck et al., 2006), the Neotropics (Pramuk 2006), and in the Western Ghats (Van Bocxlaer et al., 2009). However, Asian *Bufo* are generally used as an outgroup (Garcia-Porta et al., 2012; Recuero et al., 2012) resulting in only a few studies focusing on the spatiotemporal origin of *Bufo* in the Eastern Palearctic. Consequently, the resolution of evolutionary diversification patterns in Asian *Bufo* lineages is limited, and the

existing phylogenetic studies are generally geared towards regional sampling and mitochondrial markers (Borzée et al., 2017; Chen et al., 2013; Liu et al., 2000; Macey et al., 1998; Yu et al., 2014). Although some recent taxonomic revisions have used multi-locus data (Fong et al., 2020; Zhan and Fu, 2011), the lack of primary fossils calibrations has resulted in molecular dating estimates that mostly depend on paleogeological events and secondary calibrations. Perhaps as a result, contradicting biogeographical hypotheses have been posited regarding the Asian lineage of *Bufo* (Borzée et al., 2017; Fu et al., 2005; Igawa et al., 2006; Macey et al., 1998) and inconsistent phylogenetic species boundaries have been proposed (Macey et al., 1998; Liu et al., 2000; Fu et al., 2005; Pyron and Wiens, 2011; Zhan and Fu, 2011).

To date, three biogeographical hypotheses relevant to the initial emergence of the *Bufo* genus in the Eastern Palearctic have been proposed: (hypothesis 1) the split between a Western and Eastern lineage due to the desertification of Central Asia during the Middle Miocene, c. 12.00 Mya (Garcia-Porta et al., 2012); (hypothesis 2) vicariant speciation at the time of the earliest emergence of a high-altitudes *Bufo* group distributed in Eastern Tibet, followed by a subsequent dispersal to low elevations in the Late Miocene, c. 10.00 – 5.00 Mya (Macey et al., 1998); and (hypothesis 3) the isolation of an insular endemic clade on the Japanese Archipelago following the drift of the archipelago away from the Eurasian continent in the Late Miocene, c. 6.00 Mya (Igawa et al., 2006). Although differing in the details, the hypotheses proposed here agree on the Miocene paleogeological events as key factors in the segregation of the Eastern lineage of *Bufo*.

Comparable to the Western *Bufo*, two or more species of Eastern Palearctic *Bufo* form species complexes with similar morphologies and unclear taxonomic boundaries, such as the *B. gargarizans* and *B. japonicus* complexes (Matsui, 1986; Zhan and Fu, 2011; Arntzen et al.,

2013). The divergence of *B. gargarizans* and *B. japonicus* complexes from the other *Bufo* in East Asia, and the radiations within each species complex coincide with the Plio-Pleistocene climatic oscillations (Fu et al., 2005), sea level fluctuations and ice age glaciations (Borzée et al., 2017), and selection pressure in the Anthropocene (Hase et al., 2013). Most *Bufo* species distributed on the East Asian mainland are restricted to the Qinghai-Tibetan Plateau (QTP) and adjacent high-altitudes areas. The major exception to this is the *B. gargarizans* complex that is widely distributed in northeast Asia. The species colonised its current range from a single source through long dispersal on the Asian mainland (Fu et al., 2005; Zhan and Fu, 2011). Concomitantly, the species expanded and formed a continental lineage of *Bufo* in septentrional East Asia, here defined as the northern-most range of current “*B. gargarizans*”, including clades distributed in the Korean Peninsula and the Russian Far East (Matsui, 1986; Maslova, 2016). In contrast to the Asian mainland clades, *Bufo* septentrional East Asian clades show a marked impact of Pleistocene glaciations on range shifts (Borzée et al., 2017), and possible ice age refugia on the Korean Peninsula, and in the southernmost Amur River Basin (Borzée et al., 2017; Fong et al., 2020). One of the clades is also likely to have been isolated on the Korean Peninsula since the Last Glacial Maximum (LGM), with limited genetic exchange with other clades. This is plausibly because of the Yellow Sea level rise during the LGM (Li et al., 2016), as reflected by the monophyly of the Korean *B. gargarizans* clade (Borzée et al., 2017).

The taxonomic assignment of East Asian toads using their morphometric and geographic distribution included up to 11 distinctive species in the past (see the chronology of taxonomic updates in Figure 1-figure supplement 1). Seven recognised taxa within *Bufo* were assigned to East Asia, ranging from western People’s Republic of China (hereafter China) to continental East Asia and associated islands on the continental shelf, and the Amur River Basin.

129 Additionally, four taxa were specific to the Japanese archipelago (Matsui, 1986). A systematics
130 revision based on mitochondrial markers then simplified the East Asian mainland *Bufo* to five
131 members, with synonymisation of some monophyletic groups and invalidation of paraphyletic
132 taxa (Liu et al., 2000). In a more recent multi-locus study (Pyron and Wiens, 2011), 10 species
133 were considered valid, although some under-sampled taxa remain unresolved (see the type
134 locality distributions and chronology of systematic revision in Supplementary file 1A, Figure 1-
135 figure supplement 1).

136 Clade sorting based on morphological and life-history characteristics also influences the
137 tree topology of Asian *Bufo*. Stream-breeders such as *B. andrewsi*, *B. torrenticola*, *B. stejnegeri*
138 and the former '*Torrentophryne*' all demonstrate adaptation to the lotic ecosystem and high-
139 altitude environment (Liao et al., 2016). For instance, adults lack tympanums (Tsuji and
140 Kawamichi, 1996). Ancestral state reconstruction indicates that the semi-aquatic natural histories
141 of East Asian *Bufo* likely arose independently (Fong et al., 2020). Arguably, life history
142 characteristics may not reflect a phylogenetic relationship, however such adaptive traits may still
143 help offer explanations in the role of morphological and ecological characters on the systematic
144 of Asian *Bufo*, and help revisit previous studies in cases of convergence.

145 Other factors to consider are anthropogenic activities as they have also impacted the
146 current distribution and phylogeny of bufonids, especially in insular regions (Othman et al.,
147 2020). An example is the intergradation between the two *B. j. japonicus* subspecies in Japan that
148 were originally geographically isolated from each other. Historically, *B. j. formosus* occurred in
149 the east and *B. j. japonicus* occurred in the west (Miura, 1995). This was followed by the
150 invasion of northern Japan (i.e. Hokkaido and Sado) by the former after anthropogenic
151 introductions (Hase et al., 2013; Suzuki et al., 2020). The distribution of *B. gargarizans* has

similarly been strongly influenced by human activities, particularly through the traditional medicine trade (Zhan et al., 2020). As a result, contemporary genetic patterns in the species complex show muddled genetic signals (Lee et al., 2021). From this it can be inferred that the past and ongoing trade may lead to the reduction in local diversity within the *B. gargarizans* complex.

The issues related to the taxonomic groupings in the *B. gargarizans* complex are majorly rooted in the inconsistency of the hierarchical systematic ranks used. For instance, the validity of the endemic Taiwan toad, also called Central Formosa toad, *B. bankorensis* as a species was inconsistent across taxonomic revisions despite its position within the *B. gargarizans* clade (Matsui, 1986; Liu et al., 2000; Yu et al., 2014). These inconsistencies confuse not only the interpretations of evolutionary history, but also the ecology. For instance, ecologists have regarded *B. andrewsi* as a species of its own with life history traits related to adaptation to high altitude (Liao et al., 2016), but systematists considered *B. andrewsi* as a junior synonym of *B. gargarizans* (Fu et al., 2005). Clearly, an integrative re-evaluation of East Asian *Bufo* species is required. Species delimitation via coalescent methods is a promising tool for addressing these inconsistencies, and a notable improvement to conventional molecular phylogenetic analysis. This methodology has been used to untangle taxonomic issues linked to phenotypic variations and cryptic lineages in other Asian toads such as *Ingerophrynus celebensis* (Evans et al., 2008). Additionally, the inclusion in our species tree topology of some previously undersampled bufonids from the eastern tip of Tibet to the northern South East Asia such as *B. pageoti*, *B. tuberculatus*, a high elevation-restricted *B. gargarizans* subspecies, and *B. gargarizans minshanicus* (hereafter *B. minshanicus*), may increase the robustness of our analyses and resolve the inconsistent placement of these clades within the nomenclature.

The spatiotemporal origin of the East Asian *Bufo* lineage is uncertain, and characterised by a limited understanding of the evolutionary processes involved. In addition, species delimitations present a serious taxonomic discrepancy within the *B. gargarizans* complex, resulting in repeated calls for taxonomic clarification at a fine scale. This manuscript is structured into two sections, with different taxonomic and geographic scales (Figure 1). The first section focuses on resolving the biogeography of Holarctic bufonid species, following the Bufonidae lineage since the breakdown of the West Gondwana. We use a combination of fossilised birth-death calibrations and multi-locus coalescent-based species tree methods to estimate the most probable time and routes of colonisation of bufonids into the Palearctic (Figure 1). In the second section, we elucidate the biogeographic pattern of the East Palearctic *Bufo* and resolve the taxonomy of the *B. gargarizans* complex. Here, we evaluate the best species tree topology by testing five alternative hypotheses to recover the taxonomic relationship between Asian species of *Bufo*. We rely on an intensive and widespread sampling, and integrated biogeographical analyses, niche differentiation between divergent clades and model-based species delimitation approach to resolve the taxonomic boundaries of the *B. gargarizans* species complex, including the recently expanded septentrional East Asian clade (Figure 1).

Results

Our study addresses biogeographic scenarios explained in two different sections: (1) the biogeography of Holarctic bufonids and, (2) the biogeography of Eastern Palearctic *Bufo* and the taxonomic revision of the species complexes.

Section 1: Biogeography of Holarctic bufonids

The goal was to refine the time estimates of the split between clades of Holarctic bufonids, in coherence with fossil data (Figure 2, Figure 2-figure supplement 1 and Supplementary file 1B).

Fossilised birth death dating

The dated species tree using fossilised birth-death showed that following the split of West Gondwana, bufonids first diverged into American clades and African-Eurasian clades. The subsequent splits of genera of both ancestors followed a contemporary timeline, with the first divergence dating from the Early Oligocene to Middle Miocene (Fig 2).

In Eurasia, radiations between the Oligocene and Early Miocene resulted in the segregation of the clades of North African origin c. 26.08 Mya (95% highest posterior density [HPD 95%]/ Mya for all dating estimates in Table 1; Figure 2). The emergence of the main extant clade was dated from the Early to Middle Miocene. The monophyletic *Bufotes* emerged c. 21.72 Mya (Table 1; Figure 2) with the North African-Eurasian group diverging from the Western Himalayan *Bufotes* c. 19.81 Mya (Table 1; Figure 2). Later, the Iberian *Epidalea* and its sister genus the Eurasian *Strauchbufo* emerged c. 10.88 Mya (Table 1; Figure 2). The Palearctic *Bufo* split into two clades, the Western and the Eastern Palearctic *Bufo* between the Early and the Middle Miocene c. 14.49 Mya (Table 1; Figure 2). The credible interval showed a considerable overlap in the timing of radiations within *Bufo* during the Middle Miocene (10.00 – 15.00 Mya), indicated by the isolation of the European *Bufo* clade c. 11.03 Mya (Table 1; Figure 2), and a subsequent emergence of the Asian *Bufo* clade c. 9.99 Mya (Table 1; Figure 2).

The earliest split among the American bufonids segregated the Neotropical clades predominantly during the Late Miocene c. 24.25 Mya (Table 1; Figure 2). The most basal divergence was *Rhinella* in the southern continent c. 14.16 Mya (Table 1; Figure 2), from which

the younger Central American *Incilius* c. 11.90 Mya (Table 1; Figure 2), and Nearctic *Anaxyrus* may have diverged c. 9.56 Mya (Table 1; Figure 2). The alternative calibrated time tree analysis under a Yule speciation model resulted in a topology and time estimates that generally corresponded to that of the fossilised birth-death analysis (Table 1).

Section 2: Evolutionary diversification of Eastern Palearctic *Bufo*

Here, we focused on reconstructing the historical biogeography of Eastern Palearctic *Bufo* and resolving the taxonomic boundaries of the species complex within the genus. The vast distribution and taxonomic inconsistencies in Eastern Palearctic *Bufo* warrant a careful examination of the hypotheses proposed by previous studies.

Optimum species tree topology for Palearctic *Bufo*

The vast distribution range of the *Bufo* genus in the Eastern Palearctic resulted in cryptic diversity. Hence, we aimed to resolve the topology of Palearctic *Bufo* lineages (N species = 26) following five hypotheses (Figure 3 and Figure 3-figure supplement 1). We derived these hypotheses from ranges, life histories, geological events and the likelihood of single or multiple origins. Here, our focal taxa included the Palearctic *Bufo* species and three group members of the “*Torrentophryne*” genus, a clade paraphyletic to *Bufo*, to elucidate the validity of the genus and determine its relationship with Asian *Bufo* (Figure 3 and Figure 3-figure supplement 1). The nested sampling analyses supported an optimum tree topology for the Eastern Palearctic *Bufo* linked to Miocene geological events with the highest Marginal Likelihood estimation (Model C: Supplementary file 1C, Figure 3), followed in likelihood by the topology of a single origin for the East Asian *Bufo* clade (Model E: Supplementary file 1C, Figure 3). Whereas, the topologies structured by life history and morphological trait recovered the lowest likelihoods, and did not favour the monophyly of “*Torrentophryne*” (Model B: Supplementary file 1C, Figure 3). The

best-fit topology of Model C supported three well-resolved monophyletic clades, a *Bufo* (PP: 1.0; Figure 3), a Western Palearctic *Bufo* clade (PP: 0.90; Figure 3) and an Eastern Palearctic *Bufo* clade (PP: 1.0; Figure 3). Despite weakly recovering the East Asian mainland *Bufo* clade, the species tree strongly supported the distinction of the Japanese *Bufo* subclades (PP: 1.0; Figure 3). The alternative topologies from the four suboptimum models (Model A, B, D and E) were qualified by lower MLE values than that of Model C (details in Supplementary file 1C and Figure 3-figure supplement 1).

Phylogeny and population structure of East Asian *Bufo*

We first inferred the haplotype network of both mtDNA and nuclear data using the median-joining method. The haplotype of the concatenated mtDNA resulted in 98 haplotypes (see genetic diversity and neutrality tests in Supplementary file 1D; Figure 4A). The AMOVA based on the six monophyletic clades (N populations = 8) recovered from the mtDNA phylogenetic analyses supported the population structure and showed that 55.48% of the molecular variance was attributed to differences among clades ($df = 5$). We found 22.08% of the molecular variance among populations within clades ($df = 7$) and 22.44% of the variance was found within individuals ($df = 213$; Supplementary file 1E). The average fixation index over all the loci tested showed that $F_{SC} = 0.50$, $F_{ST} = 0.78$ and $F_{CT} = 0.55$, and that there was a negative correlation between geographical distance and genetic differentiations (N populations = 13; Pearson' $r = -0.059$, Figure 4-figure supplement 1). The analysis of the nuclear data (*POMC-RAG-I-Rho*) from eight populations resulted in a single group of haplotypes ($N = 54$) with a haplotype diversity (Hd) of 0.972 (Figure 4B), an average pairwise difference of 7.667 (± 4.533), and a nucleotide diversity of 0.007 (± 0.005). Additionally, the Mantel Test for nuclear data revealed a low correlation between the geographic distance and genetic variation for the diploid

populations of the East Asian *Bufo* (*N* population; 8; Pearson's $r = 0.072$, Figure 4-figure supplement 1).

To increase the phylogenetic resolution of *Bufo* across East Asia, we enlarged the sampling range to septentrional Eastern Asia where it covered the distributions of *B. sachalinensis sachalinensis* in the Amur River Basin and the *B. sachalinensis cf. sachalinensis* subclade restricted to the Korean peninsula (currently "*B. gargarizans*"). We also conducted two independent molecular phylogenetic analyses using concatenated mtDNA fragments of the control region (*CR*) and NADH dehydrogenase 2 (*ND2*); and concatenated the nuDNA of the gene fragments proopiomelanocortin (*POMC*), recombination activating gene 1 (*RAG-1*) and rhodopsin (*Rho*). The phylogenetic trees derived from the mtDNA (*N* taxa = 221) and nuDNA data (*N* = 44) inferred six monophyletic clades, respectively. The mtDNA and nuDNA trees recovered discordant topologies, and they resulted in different composition of clades (Figures 4C and 4D). We found *B. gargarizans* from the Asian mainland to be segregated into multiple clades scattered across the nuDNA tree (see details in Table 2, Figure 4). We highlighted other discordance found between the monophyletic *B. j. formosus* and *B. j. japonicus* in the mtDNA tree (Clades 1 and 2; Figure 4C), in which both Japanese *B. japonicus* clustered in a clade based on the nuDNA tree (Clade 4; Figure 4D). In addition, nuDNA resolved *B. bankorensis* and the septentrional East Asia clades of *B. sachalinensis* as a monophyletic (Figure 4D). Despite these inconsistencies, we found both mtDNA and nuDNA trees to similarly manifest a segregation between the septentrional East Asian clades of *B. sachalinensis* and *B. gargarizans* (Figure 4).

Additionally, we reconstructed a 16S rRNA only tree to include an individual *B. gargarizans* from Vietnam. The individual did not cluster with East Asian lowlands *B.*

gargarizans, rather it was nested inside a clade of high elevation-restricted species including *B. andrewsi* and *B. tibetanus* (Clade C: Figure 4-figure supplement 2).

Diploid genotype clusters

The Structure analysis based on the 1,030 bp of multilocus *POMC-RAG-1-Rho* supported two clusters ($K = 2$) within East Asian *Bufo*, with Mean (LnProb) equal to -827.838 and mean (similarity score) among 10 runs equal to 0.974 (Figure 4D). The first cluster included the populations of *Bufo* in the western, central, southeastern and northeastern Asian mainland along with the Japanese *Bufo* (Figure 4D). We recorded a negligible amount of admixture between population clusters of *B. gargarizans* in the western mainland, eastern mainland and the Japanese Archipelago (see admixture portions in Supplementary file 1F and Figure 4D). However, we found significant amounts of admixture in central and the southeastern mainland Asia (Supplementary file 1F, Figure 4D). The second cluster was restricted to *B. sachalinensis* of septentrional East Asia with significant admixture, ranging from the Korean Peninsula to the Amur River Basin and Sakhalin Island (Supplementary file 1F, Figure 4D). This cluster also included the population of *B. bankorensis* distributed in Taiwan Island (Figure 4D).

The effect of introgression

The cytonuclear discrepancy (Figure 4) may be the result of introgression or/and incomplete lineage sorting. Thus, we further evaluated the pattern of introgression in our nuclear data (*POMC-RAG1-Rho*, $N = 44$, nucleotide length = 1,030 bp) using ABBA or BABA test. To do so, we employed Patterson's D-statistic to compare the number of allelic ABBA and BABA sites. Here, the D-statistic value we obtained was equal to 1, with the ABBA-BABA pattern calculated among the sites failing to reach equal frequencies (50:50). The probability of specific sites carried the allelic patterns of ABBA or BABA was equivalent to 1, and the number of

segregating sites that fit the pattern of ABBA or BABA in at least one population was equivalent to 0. The unsymmetrical frequencies violate the assumption that only incomplete lineage sorting affects the *Bufo* nuclear tree, and showed the possibility of introgression as a significant factor in shaping the nuclear genetic structure.

Molecular dating and ancestral range

Despite of topological discordance, all nuDNA and mtDNA trees consistently revealed a distinction between the clades of Eastern Central Asia and septentrional East Asia in the *B. gargarizans* complex. Thus, we also provided dating estimates using the nuDNA dataset as an alternative to the mtDNA estimates (Table 3). In the context of dating estimates, we considered mtDNA estimates more informative than nuclear estimates based on two factors; (1) the higher number of taxa in the mtDNA tree than in the nuDNA tree resulting to a clearer phylogeographic structure with higher support values for the clades recovered, and (2) the clades recovered from the mtDNA tree highly matched the best-supported topology of species tree of Eastern Palearctic *Bufo* (Figure 2). Overall dating for the major *B. gargarizans* clades from the nuDNA dataset (especially the 95% HPD ranges) were in agreement with the mitochondrial estimates, with nuDNA ingroup nodes slightly younger than mtDNA (Table 3).

Here, we compared the three hypotheses related to the phylogeography of the *B. gargarizans* complex, explained by: QTP vicariance and dispersal (Macey et al., 1998), dominance of long dispersals (Fu et al., 2005) and ice-age refugia (Borzée et al., 2017). Our dating and ancestral range estimates supported the contribution of vicariance and dispersal to the earliest diversification of the East Asian *Bufo*. We dated the events concern the basal clade of Eastern Asian *Bufo* from the Early to Late Miocene c. 14.20 Mya (see HPD 95% in Table 3; Figure 5). The events were subsequently followed by the emergence and isolation of the

Japanese *Bufo* from the Eastern Asian lineage in Late Miocene (Figure 5-A1), the Japanese *Bufo* group then splitting into two distinct species, *B. j. formosus* and *B. j. japonicus*, c. 8.23 Mya (Table 3, Figure 5). These events were followed by the radiations within the East Asian *Bufo* lineage resulting in the independent divergence of *B. stejnegeri* in the Korean Peninsula, c. 8.90 Mya (Table 3, Figure 5) and a split between the high altitudes East Asian clades: *B. tibetanus*, *B. andrewsi* and *B. gargarizans* c. 11.32 Mya (Table 3, Figure 5). These East Asian clades are present in the areas of the QTP and likely to have dispersed from high elevation areas (Figure 5-A2), resulting in the divergence between *B. tibetanus* and *B. andrewsi*, dated between the Pliocene and the Late Miocene c. 5.24 Mya (Table 3, Figure 5).

We found the combined effects of long dispersal and ice-age refugia to contribute equally to the radiation of the *B. gargarizans* complex in the Plio-Pleistocene, although the trade of the species over the last millennium has muddled the genetic signature (Figure 5-A3). The deepest split within the *B. gargarizans* complex is estimated to have occurred around the same timeframe and at high elevation, delineating the *B. minshanicus* clade (*B. gargarizans* subspecies) in the southwestern to central mainland Asia (i.e. Sichuan and Shaanxi) c. 5.27 Mya (Table 3; Figure 5). Later, *B. gargarizans* may have dispersed to lower latitudes with the splitting from the monophyletic clade of *B. gargarizans gargarizans* ranging from southeastern to northeastern mainland Asia (i.e. Shanghai, Zhejiang, Hubei, Jinan and Shenyang) c. 2.85 Mya (Table 3; Figure 5). A widely dispersed *B. gargarizans* clade further diverged in the central, southeastern and northeastern mainland Asia (i.e. Sichuan, Shaanxi, Hubei and Jiangsu, Dalian) c. 1.85 Mya (Tables 3; Figure 5), nested within the *B. g. popei* clade (Figure 5). The *B. bankorensis* clade is a more recent divergence c. 1.31 Mya (Table 3; Figure 5), sharing a common ancestor with the southeastern *B. g. gargarizans* (Table 3, Supplementary file 1G, Figure 5).

We estimate the emergence of *B. j. formosus* in Japan to have occurred between the Early Pleistocene and the Late Miocene c.3.93 Mya (Table 3, Figure 5), followed by a recent divergence of *B. j. japonicus* between the Pleistocene and the Pliocene c. 1.91 Mya (Table 3, Figure 5). In northeast Asia, the *B. stejnegeri* clade may have independently emerged at the Pliocene-Pleistocene boundary c. 2.77 Mya (Table 3, Figure 5). Finally, the *B. gargarizans* clade dispersed and diverged eastward of the Yellow Sea during the Pleistocene c. 1.88 Mya (Table 3, Figure 5), established a population on the Korean Peninsula c. 1.40 Mya (Table 3, Figure 5) and expanded further to the Amur River Basin c. 0.76 Mya (Table 3, Figure 5).

LGM population expansion

We tested the ice-age refugia hypothesis to infer the impact of past Yellow Sea fluctuations on the East septentrional *Bufo* clades. We demonstrated that the rise in the past Yellow Sea level resulted in population expansions for the Korean and Russian *Bufo* clades, as showed by significantly negative values of Tajima's *D* and Fu's *F_s* (Refer to Clade 6; Supplementary file 1D). The results were consistent with the Bayesian Skyline Plot, which indicated a recent population expansion in the Amur River Basin clade c. 0.48 Mya (effective population size trajectory (mean [HPD]/*N_e*) 1.18 [0.30-2.94]) until present (*N_e* = 3.73 [0.23 – 20.64]) with a mean likelihood: -1930.53 [HPD: -1942.89, -1918.43], Figure 6). In comparison, the populations of the *Bufo* distributed on the southwestern margin of Yellow Sea consistently declined since the Late Pleistocene c. 0.55 Mya (*N_e* = 5.34 [13.46 – 1.124]) until present (*N_e* = 3.43 [0.48-11.84]) with a mean of likelihood: -2954.63 [HPD: -2969.60, -2941.08], Figure 6).

Ecological niche modelling

To clarify the divergence in ecological requirements between *B. gargarizans* and *B. sachalinensis*, we examined the niche overlap between the two clades (Figure 7). The habitat

suitability model for *B. gargarizans* (Figure 7A) was characterised by an AUC of 0.9239 ± 0.0185 and a TSS of 0.6741 ± 0.0408 , while the model for the *B. sachalinensis* (Figure 7B) was characterised by an AUC of 0.9632 ± 0.0043 and a TSS of 0.8723 ± 0.0115 . The “I” niche overlap statistic between the two models was 0.4198, while the “D” statistic was 0.1566. These values were significantly lower than the average values of the null distribution, with the mean of “I” in the null distribution being 0.9788 ($p < 0.0001$) and the mean of “D” in the null distribution being 0.8546 ($p < 0.0001$). Here, our result showed the overlap to be significantly less than expected, and therefore supporting the segregation in environmental requirements between the two East Asian *Bufo* clades.

Population migration in septentrional East Asia

We then examined the impact of gene flows and migration on local adaptation of the focal *B. sachalinensis* clades. Along the latitudinal gradient, our results suggested a weak gene flows outwards from the Amur River Basin and *B. s. sachalinensis*, in comparison with the symmetrical gene flow between the subspecies of *B. sachalinensis* restricted to the Korean Peninsula and *B. gargarizans* (Figure 8). This limited gene flow did not hinder the local adaptation of *B. sachalinensis* clade distributed at northern latitudes.

In particular, the migration trajectory obtained from the unlinked multi-locus data showed a lack of gene flow with a symmetrical pattern of gene exchange between the populations of *Bufo* distributed on the northeastern mainland and populations distributed on the Korean Peninsula (refer to migration rates and theta (Θ) estimates: Supplementary file 1H, Figure 8). We demonstrated a comparable and moderate gene flow when focused on the average population migration rate (M) of haploids ($2Nm$) and diploids ($4Nm$) of the two populations. The gene flow rate from the northeastern mainland towards the Korean Peninsula (M_1 to M_3 ; Figure 8) was

0.106, and in the opposite direction (M_3 to M_1 ; Figure 8) was 0.214 (Supplementary file 1H). Contrary to the symmetrical migration patterns between the northeastern mainland and Korean Peninsula, we estimated an asymmetrical migration pattern toward septentrional Eastern Asian *B. s. sachalinensis* in the Amur River Basin (Supplementary file 1H, Figure 8). We detected a higher rate of gene flow into the Amur River Basin, transferred from the eastern mainland and from the Korean Peninsula (Supplementary file 1H, Figure 8). We however found a lower rate of gene flow from the Amur River Basin towards the northeastern mainland and the Korean Peninsula (Supplementary file 1H, Figure 8).

Additionally, the migration pattern between the two subpopulations of *B. stejnegeri* in Korea was also symmetrical, with a negligibly low rate of gene flow from the northern towards the southern population, and vice-versa (Supplementary file 1H, Figure 8).

Species boundaries and taxonomy updates

Here, we provided support to delimit the cryptic *B. s. sachalinensis* distributed in septentrional Eastern Asia (Amur River Basin). The path sampling and Bayes factor analyses generally supported lineage-splitting as the best speciation pattern as opposed to the lumping of the clades of East Asian *Bufo* (Supplementary file 1I, Figure 9). Out of the eight species delimitation models tested (model A to H; Supplementary file 1I, Figure 9), Bayes factors determined Model A to be the most favourable alternative, receiving the highest support out of all scenarios determined by the path sampling analyses (Supplementary file 1I, Figure 9). Our model supported the split of the East Asian *Bufo* genus into seven independent taxonomic units (Model A; MLE: -502.801, Bayes factor: 3.198; Supplementary file 1I, Figure 9). These correspond to *B. j. japonicus*, *B. j. formosus*, *B. andrewsi*, *B. gargarizans*, *B. stejnegeri*, *B. bankorensis* and the clade we refer to as *B. sachalinensis* (see justification below). The Model A

also supported a species-level boundary between the two subspecies of *B. japonicus*: *B. j. japonicus* and *B. j. formosus*. The species delimitation model suggested the taxonomic merging of *B. sachalinensis* cf. *sachalinensis* clade in the Korean Peninsula with the *B. s. sachalinensis* clade in the Amur River Basin (Supplementary file 1I, Figure 9). Moreover, the model A validated the split between the clades distributed in septentrional East Asia from *B. g. gargarizans* clades in East Asia (Model A; Supplementary file 1I, Figure 9).

Following multiple calls for a taxonomic revision and the various species descriptions and synonymisation (Matsui, 1986; Igawa et al., 2006; Borzée et al., 2017; Lee et al., 2021), we suggest the following updates to the taxonomy. It is to be noted that we do not describe any new species as a long list of valid former names is available (Supplementary file 1I, Figure 1-figure supplement 1). We provided the taxonomy updates on the basis of species delimitation analysis (Figure 9), and the following recommendations were corroborated by the allopatric distribution of the delineated clades (Figure 9) and the differentiation in niche requirements (Figure 8). Thus, we presented the most robust taxonomic framework applied to this group since the description of *B. gargarizans*.

Here we refer to *Bufo gargarizans* Cantor, 1842 (see type locality in Supplementary file 1A and Figure 1-figure supplement 1) as *B. gargarizans gargarizans* (the Zhoushan toad based on the type locality) following the literature (Dubois and Bour, 2010). The subspecies is currently geographically widespread, a consequence of its synonymisation with several clades of taxonomically unstable *Bufo* in the East Asian mainland, including the intergraded subspecies *B. gargarizans popei* Matsui, 1986, two high elevation-restricted *Bufo* species described under the epithets of *B. andrewsi* Schmidt, 1925, and *B. minshanicus* Stejneger, 1926 (see type localities and ranges in Supplementary file 1A, Figure 1-figure supplement 1). Here we provide evidence

for the independent evolution of *B. andrewsi* (Figure 9) and therefore the validity of *B. andrewsi* at the species level, against the previous synonymisation (Fu et al., 2005). The distinction between *B. bankorensis* and *B. gargarizans* (Figure 8) is supported by the polyphyletic structure in their mtDNA lineage (Figures 4B, 5), further reinforced by the independent evolutionary lineage as shown by the multilocus BFD tree (Figure 9). Hence we support the species status of *B. bankorensis* and reject their synonymisation (Liu et al., 2000). The species was described under the epithet *B. bankorensis* Barbour, 1908 (Supplementary file 1A, Figure 1-figure supplement 1).

Most notably, we propose to redefine the *Bufo* clades in septentrional East Asia (Figure 9), which includes the clades referred to as *B. sachalinensis* cf. *sachalinensis* and *B. s. sachalinensis* in the analysis section. We therefore recommended the elevation of the septentrional East Asia clade currently named *B. gargarizans* to the species level, under the taxonomic epithet *Bufo sachalinensis* (Life Science identifier (LSID): urn:lsid:zoobank.org:pub:XXX) and under the common name Sakhalin toad. The species was described under the name *B. vulgaris sachalinensis* Nikolsky, 1905 (Type locality: Sakhalin Island, Russia, syntypes ZISP 1934–1936 and MNKNU 26290; Supplementary file 1A, Figure 1-figure supplement 1). We dated the split within *B. gargarizans* from the Middle Pleistocene (Figure 5), and the two species segregated along a diversity of ecological requirements (Figure 9). We then propose the assignment of clades in the Korean Peninsula and the Amur River Basin to different subspecies under *B. sachalinensis* to reflect evident geographic and genetic distinction. This framework aligns with the lineage-splitting pattern suggested by Model B in the nuDNA species delimitation analysis which provides a model support value only marginally lower than that of Model A (Figure 9), while our mtDNA phylogeny provides robust support for subspecies-level differentiation of

these groups. The population on the Korean Peninsula does not have precedence in the taxonomy and it is not the type locality to any *Bufo* that is not synonymous with *B. stejnegeri* (Supplementary file 1A, Figure 1-figure supplement 1). Hence, we refer to this evolutionarily significant unit as *B. s. cf. sachalinensis*.

Discussion

Our reconstruction of the Holarctic bufonids biogeography using species tree with fossilised birth-death calibration and phylogenetic hypotheses testing verified that the earliest split between Western and Eastern Palearctic *Bufo* occurred during the Middle Miocene (c. 14.46-10.00 Mya; Table 1), temporally matching with the maximum dust outflows in the deserts of Central Asia (Guo et al., 2002). Subsequently, we retraced the single origin Asian *Bufo* in eastern Asia before the split between the continental and Japanese lineages, corroborated by the two Miocene paleogeological events: the rapid uplifts of regions adjacent to the QTP (Hengduan mountains; Xing and Ree, 2017) and the drift of the Japanese Archipelago (Barnes, 2003).

Our revisit of the three phylogeography hypotheses did not favour any specific hypothesis explaining the present geographic distribution of eastern Palearctic *Bufo*, in disagreement with the long colonisation hypothesis (Zhan and Fu, 2011). Instead, our results provided support to the combination of the three elements: vicariance of the western mainland clade, followed by long dispersal and possible refugia in northeastern Asia during the last ice age. We detected a loss of genetic structure within *B. gargarizans* clades, possibly due to introgression resulting from the trade of the species over a millennium (Lee et al., 2021). The recent segregation around the Yellow Sea, as shown by the species distribution models and migration pattern, also provides supports to the delineation of a septentrional Asian *Bufo* clade associated with range shift towards northern latitudes. We therefore resurrect the previously

described *B. sachalinensis* and elevate it to the species level. Our findings resolved the taxonomic boundaries in the *B. gargarizans* complex, and redefine the taxonomic and conservation units: *B. g. gargarizans*, *B. s. sachalinensis* and *B. sachalinensis* cf. *sachalinensis*, with the later waiting for a subspecies description.

Radiation of Holarctic bufonids

Fossil calibration estimates refined the known divergence time for Neotropical and Eurasian Bufonid lineages, and rejected the hypothesis of shared origin (Sanchiz, 1997). Although we could not confirm the African biogeographic origin of the Western Palearctic bufonids (Figure 2), our time estimates show that the emergence of bufonids into Europe was subsequent to the landmass connection of Eurasia and Africa (Frazão et al., 2015). In addition, our ancestral range analysis rejected the hypothesis that bufonids dispersed out of East Asia and into Europe (Sanchiz, 1997; Rage and Roček, 2003). Instead, we confirm the dispersion of bufonids from Europe and into East Asia, in agreement with the Central Asian desertification (Garcia-Porta et al., 2012) as the factor of segregation between the Western and Eastern Palearctic bufonids. The Miocene radiations within the Western Palearctic *Bufo* clade (c. 14.49 – 11.03 Mya; Table 1, Figure 2) are consistent with estimates of the highest dusting outflow in the deserts of Central Asia (c. 15.00 – 13.00 Mya; Guo et al., 2002). The desertification may have triggered an early colonisation of Asian *Bufo* in central Eastern Asia (c. 9.99 Mya; Figure 2), a process that may have taken place prior to the second period of maximum dusting emission in Central Asia (c. 8.00 – 7.00 Mya; Guo et al., 2002).

Mito-nuclear discordance

The phylogenetic trees of East Asian *Bufo* reveal a conspicuous discordance between the topologies of the mtDNA and nuDNA, especially in the placement of regional populations of the

B. gargarizans complex and the inconsistency in positioning *B. bankorensis* and *B. andrewsi* (Figure 4). For instance, the mtDNA phylogeny placed the *B. gargarizans* complex within a nested monophyletic clade, with *B. gargarizans* in the east Asia, the Korean Peninsula and the Amur River Basin as sister groups to each other (mtDNA tree: clades 5 and 6; Figure 4). Conversely, the nuclear tree did not recover a clear geographic structure for the *B. gargarizans* complex (Figure 4). This discordance might reflect the dissimilarity in the evolutionary rates of nuclear and mitochondrial markers as *RAG-1* in amphibians is known to have a slower rate of evolution in comparison to mtDNA (San Mauro et al., 2004).

Introgression and incomplete lineage sorting could have contributed to the discordance between mitochondrial and nuclear trees. Here, the D-statistical analysis provides support to the occurrence of introgression resulting in mitonuclear disequilibrium among populations of *B. gargarizans* in the Central East Asian mainland (Patterson's D Statistic = 1.0). Introgression is a common occurrence in amphibians, and may be naturally and anthropogenically occurring along contact zones (Bell and Irian, 2019). Ancestral polymorphism may have contributed to the incomplete lineage sorting, a point already discussed between *B. gargarizans* and *B. bankorensis* (Yu et al., 2014). We recommend testing the possibility of incomplete lineage sorting in contact zones between the parapatric members of the *B. gargarizans* complex: *B. tibetanus*, *B. andrewsi* and *B. gargarizans* (Figure 7).

Geologically driven divergence in Asia

Past geological events such as the central Asian desertification, the orogeny of the Tibetan plateau and the Eastern Pacific drift had a more important impact than geographical distribution and selection in shaping the genetic structure of the Eastern Palearctic *Bufo* (Supplementary file 1C, Figure 3). Selection is a crucial aspect in evolution, however our model

indicates that selection based on life history strategies had a peripheral role in influencing the evolutionary pathways of Asian bufonids. While the relationship between Asian *Bufo* and the species previously assigned to “*Torrentophryne*” remains uncertain (low support of PP = 0.03; Figure 3), the most accurate species tree topology rejects the monophyly of the lotic breeders “*Torrentophryne*” (Model C vs. Model E; Supplementary file 1C, Figure 3). The paraphyletic nature of “*Torrentophryne*” is also corroborated by previous taxonomy studies (Liu et al., 2000; Pyron and Wiens, 2011). Nevertheless, our topology models validate the convergence among semi-aquatic members of Asian *Bufo* (*B. andrewsi*, *B. stejnegeri* and *B. torrenticola*) and the homoplasy between *Torrentophryne* and *Bufo* (Supplementary file 1C, Figure 3).

Miocene to Pleistocene diversifications in Asia

Subsequent to the early radiation in association with the Central Asian dust events at the end of the Middle Miocene, we confirmed a single origin of Asian *Bufo* from Eurasia through central Asia during the Late Miocene (Supplementary file 1G, Figure 5). The Asian *Bufo* shared an ancestor before the divergence between the East Asian mainland and Japanese Archipelago lineages took place (Figure 5) as a result of the two most important vicariance events that occurred contemporarily during the Miocene (Figures 5-A1 and 5-A2). Our dating estimates, previous hypotheses (Igawa et al., 2006; Macey et al., 1998) and paleogeological events are consistent for two main points: (i) an early radiation of the Asian continental *Bufo* triggered by the QTP vicariance (Miocene; c. 8.36 Mya; Table 3, Supplementary file 1G, Figure 5-A2), and (ii) the isolation of *Bufo* clades on the Japanese archipelago (Late Miocene; c. 7.71 Mya; Table 3, Supplementary file 1G, Figure 5-A1). The early radiation of the Asian continental *Bufo* occurred in high elevation-restricted clades: *B. tibetanus* and *B. andrewsi* (Figure 5) and coincides with the pre-completion of the orogenesis of the Hengduan Mountain (Late Miocene; c

8.0 to 10.0 Mya; Xing and Ree, 2017). The estimated time of isolation of the Japanese *Bufo* from the Asian continental clades (c. 7.71 Mya; Table 3, Supplementary file 1G, Figure 5) is pertinent, as the event may have occurred before the complete separation of the Japanese Archipelago from the Eurasian landmass (c. 5.00 Mya; Barnes, 2003).

The major radiations within the primary clades of East Asian mainland *Bufo* took place predominantly between the Pliocene and Pleistocene (Figure 5-A3), with dispersals principally driving the diversifications of the *B. gargarizans* complex clades in the lowlands (Figure 5). In comparison, divergence in the insular and septentrional East Asia clades was mostly driven by vicariance and dispersals (clades of Japanese *Bufo* and *B. sachalinensis*; Figure 5). In the case of *B. stejnegeri*, although the estimated timing is more recent (Pliocene; c. 3.31 Mya; Table 3, Supplementary file 1G, Figure 5) relative to the estimates in the literature (Pliocene; c. 4.30 Mya; Fong et al., 2020), our results are in agreement on the independent rise of this clade on the Korean Peninsula prior to complete formation of the Baekdu Mountains shields (2.80-1.50 Mya; Kim et al., 2017).

The Yellow Sea as a biogeographic boundary

Previous studies estimated the isolation of *Bufo* populations on the Korean Peninsula to date from the Pleistocene (Borzée et al., 2017; Lee et al., 2021), and our results confirm this pattern. However, additional samples from further east demonstrated the presence of a deeper segregation between the clades around the Yellow Sea, with *B. gargarizans* to the southwest and *B. sachalinensis* in septentrional Eastern Asia (Figure 6). This pattern is visible through both mtDNA and nuDNA. A late colonisation of the septentrional East Asia regions has also been supported by an increase in population size in relation with the LGM, a pattern not visible for the *B. gargarizans* clades (Figure 6). The isolation of *B. sachalinensis* clade on the Korean Peninsula

is in synchrony with the final stage of the Sanduo Event that triggered the widening of the Yellow Sea basin (Neogene-Quaternary; Lei et al., 2018). Most probably, the drift between clades was subsequently induced by variations in water depth, temperature and the tidal circulation of the Yellow Sea (Li et al., 2016). This pattern is also visible in other amphibian clades that became isolated during the quaternary (Borzée et al., 2020).

Delineation of climatic niche selection

In congruence with the best species delimitation model (Figure 9), the niche overlap statistics demonstrated the segregation of the two clades tested: *B. sachalinensis* in septentrional East Asia and *B. gargarizans* in eastern central Asia (or southwestern margin of the Yellow Sea; Figure 7). The two clades have different environmental requirements, with the minimum temperature being one of the key factors segregating the niche of two species (Figure 7-figure supplement 1). The spatial heterogeneity demonstrated by both clades may have led the species to adapt locally, resulting in specific niches, as exemplified in closely related bufonids: *B. bufo* from the Western Palearctic (Luquet et al., 2015). Local adaptation may also derive from the apomorphic trait related to behavioural and phenotypic plasticity, as seen in the case of bufogenin production in *B. gargarizans* along a latitudinal gradient with different climates (Cao et al., 2019).

Ineffective gene flow on the septentrional East Asia

Low gene flow rates among the septentrional East Asian clades (Figure 8) indicate a strong effect of genetic drift, especially in the isolated clade of *B. s. cf. sachalinensis* in the Korean Peninsula. The latest glaciations may not have induced range shifts as the clade was already isolated on the peninsula following the rising level of the Yellow Sea. In addition, low migration rates support the natural selection and gradual adaptation of the *B. s. sachalinensis*

clade to the cold climate of the northern latitudes. Long distance dispersal had a weak impact on *B. s. sachalinensis* population in the Amur River Basin (Figure 8), probably due to the reduced capacity for dispersal to compensate for the energy investment on low-temperature tolerance (Kolbe et al., 2010).

This pattern of unidirectional gene flow may however need to be seen from a different perspective. The lack of gene flow here may also highlight that the *Bufo* trade may not have been as widespread on the Amur River Basin and Korean Peninsula as it was in China, a pattern correlated with the size of human settlements. The presence of two individuals within the range of *B. gargarizans* showing 95% and 87% admixtures with the *B. sachalinensis* clade (Supplementary file 1F, Figure 4D) is likely the result of human-led translocations. *Bufo* toads have been used for traditional Chinese medicine for more than a millennium (Zhan et al., 2020; Lee et al., 2021) and gene flow towards the regions relying on the genus for medicine may actually reveal the directionality of the trade, and escape of individuals at markets, rather than natural dispersal.

Taxonomic recommendations

The distinction among *B. tibetanus*, *B. minshanicus* and *B. andrewsi* cannot be assessed in detail here due to the scarcity of samples. However, *B. tibetanus* and *B. andrewsi* are genetically meaningful as species, each clade being monophyletic and part of a sister clade to *B. gargarizans* (Fu et al., 2005). We however note the possible presence of *B. andrewsi* in Zhongdian, the type locality of *B. tibetanus*. Here, we resolve the long-standing question about the relationship between the *B. gargarizans* complex, *B. andrewsi* and *B. bankorensis* (Chen et al., 2013; Fu et al., 2005). In spite of synonymisation between *B. andrewsi* and *B. gargarizans* through mtDNA phylogeny and allozymes data (Fu et al., 2005), the coalescent perspective on

delimiting the species in the present study supports the taxonomy validity of *B. andrewsi* as a distinct species, a point strongly supported by morphometric data (Liao et al., 2016). Here we demonstrated the impact of the QTP orogenesis on the divergence of *B. andrewsi* from *B. tibetanus* during the Late Miocene (Table 3, Figure 5). Both paleogeography and species delimitation patterns justify the species epithet, and the clade should be referred to as *B. andrewsi* Schmidt, 1925 in later studies. We recommend further research on the relationship with the clade described as *B. tibetanus* in light of the Late Miocene divergence between these two clades (Figure 5).

The improved sampling resolution throughout East Asia, combined with the use of advanced taxonomic tools, helped resolve the long-standing cryptic boundaries between *B. gargarizans* and *B. sachalinensis* (Matsui, 1986). The splitting of *B. sachalinensis* from *B. gargarizans* results in the newly elevated species containing a single evolutionary lineage (Figure 9). This suggests that the species adapted to the local environment after dispersion from a glaciation refugia and colonisation towards northern latitudes during the Late Pleistocene (Figures 5 and 9). A similar pattern of northward gradient colonisation observed in other amphibians in the area such as *Kaloula borealis* (Othman et al., 2021). The presence of two distinct clades, *B. sachalinensis* cf. *sachalinensis* and *B. s. sachalinensis* highlights a strong allopatric structure in the septentrional East Asia clades rather than a case of recent divergence (Figure 4), highlighting the significance of segregated conservation units for these two lineages.

We corroborate the validity of *B. bankorensis* as an independent species through the species delimitation analysis (Figure 9), contemporaneous with the divergence time of other clades within the *B. gargarizans* complex, and an active period of glaciations (Mid to Late-Pleistocene; Figure 5). This clarifies the presence of a common ancestor, followed by a single

evolutionary lineage for each of the clades within the *B. gargarizans* complex, as well as *B. bankorensis* on Taiwan Island. As supporting evidence, the niche segregation between *B. bankorensis* and the other *B. gargarizans* clades is elevation related. *Bufo bankorensis* is distributed in mountainous areas (Lin and Lue, 2004), *B. gargarizans* in the lowlands and *B. minshanicus* at higher elevation on the QTP, sympatric with *B. andrewsi* (Fu et al., 2005).

Despite the current support on the splitting of *B. j. formosus* and *B. j. japonicus* in relation to the Pliocene-Pleistocene divergence (Figures 5 and 9), we do not recommend further taxonomic acts as the taxonomy of the clades is already resolved (Dubois and Bour, 2010; Matsui, 1984), and the type locality of the later subspecies is not restricted to any specific area in Japan. We do however provide evidence about the independent evolution of the two subspecies of *B. japonicus* through historical events and species delimitation, which could be used as the phylogenetic support needed for the subspecies to be elevated to the species level.

It is worth pointing out the widespread risk for most *Bufo* species studied in this work. Anthropogenic introductions have already resulted in threats to insular populations such as *B. j. formosus*, which suffered from introgression with *B. j. japonicus* (Hase et al., 2013). Emphasis should be placed on limiting the trade to regions within the geographic occurrence of the clade traded to avoid such risks, and especially for isolated clade with high genetic homogeneity and small genetic variability, such as *B. s. cf. sachalinensis* ($\pi = 4.76 \pm 2.34$; Supplementary file 1D). A risk reflected through the declining effective population size in Korean population compared to the East Chinese one (Figure 6). In addition, trading *Bufo* species can result in artificial introduction, as seen with the introduction of *B. j. formosus* on Hokkaido (Suzuki et al., 2020) and the spread of pathogens (Borzée et al., 2021).

In summary, ever since the Miocene and throughout the ice ages, past paleogeological events have been the most powerful factors promoting the genetic structures of *Bufo* in the Eastern Palearctic. A segregation between the West and East Palearctic *Bufo* clades in relation with the Central Asian dusting resulted in the vicariances of Asian *Bufo*, mainly under the influences of the QTP orogenesis and the Japanese continental drift. Unlike the simple allopatric speciation, co-distribution and artificial secondary contact have resulted in a blurry genetic structure of the clades inhabiting the central East Asian lowlands: *B. gargarizans gargarizans* and *B. g. popei* both showcase the likely loss of genetic diversity during the Anthropocene. Glaciations impacted the range expansions of *Bufo* to septentrional East Asia, and led to microclimate adaptation along the latitudinal gradient of populations of the *B. gargarizans* complex. As a long-sought resolution to long-standing systematics problems, we elevated *B. sachalinensis* from septentrional East Asia to the species level. This recommendation is supported by the population expansion resulting from the LGM, the gene flow towards northern latitudes, and a significant distinction in niche suitability when compared with *B. gargarizans*. Concurrently, the revision identified two significant conservation units linked to the subspecies *B. s. sachalinensis* in the Amur River Basin and *B. s. cf. sachalinensis* on the Korean Peninsula. Future research on behaviour and call properties could further elaborate on the differentiation highlighted here with genetic tools, and genomic studies may help confirm the genetic patterns demonstrated herein. In addition, this study insists on a strong call for the genetic preservation of Asian *Bufo* clades to prevent future loss of genetic diversity.

Materials and Methods

Taxa sampling and dataset design

We sampled 274 *Bufo* individuals across East Asia. We supplemented our dataset with Genbank sequences to cover the totality of the described range of the species, from the northern distribution in Sakhalin Island to the southern distribution limit in Vietnam. The specimens collected represent four recognised species and two subspecies: *B. andrewsi*, *B. gargarizans*, *B. stejnegeri*, *B. bankorensis*, *B. japonicus japonicus* and *B. j. formosus* (Supplementary file 1J). In addition, we used sequences from Genbank ($N = 158$; Supplementary file 1J), notably adding samples listed as *B. tibetanus* and *B. minshanicus* from western China and a sample of *B. gargarizans* from Vietnam (Supplementary file 1J). We selected five loci from the *CR*, mitochondrial protein-coding *ND2* and nuclear protein-coding markers *POMC*, *RAG-1* and *Rho*. To amplify these targeted gene fragments, we adapted universal primers used in closely related studies (Supplementary file 1K). In addition, we designed primer pairs for several gene fragments (Supplementary file 1K), modified from the most homologous sequences in Genbank. Details on the molecular works such as the isolation of gene fragments and the Polymerase Chain Reaction (PCR) amplifications are explained (Appendix 1) with the final primers used and the PCR conditions (Supplementary file 1K).

From these samples, we designated three datasets varying in number of taxa and geographical scale (Figure 1), with considerable overlap of samples. We described the methodology and the datasets as follows: (1) Estimates of divergence dates in the Holarctic bufonids species tree using a fossilised birth-death model (Heath et al., 2014) and estimating divergences to a time frame coherent with the minimum fossils age. For the species tree reconstruction, we included additional sequence data from two loci, a mitochondrial ribosomal large subunit 16S rDNA (*16S*) and a nuclear protein coding C-X-C chemokine receptor type 4 (*CXCR4*) available in Genbank. Here, the final dataset comprised seven unlinked molecular loci

obtained from gene fragments of *CR-16S-ND2-CXCR4-POMC-RAG-1-Rho* (2,666 bp) and included 313 sequences belonging to 39 taxa distributed across South and Central America, Nearctic and Eurasia (Supplementary file 1J). (2) Evaluation of the best species tree topology for the Palearctic *Bufo* species under a hypothesis testing approach derived from geographical, geological, ecological, and life history variables. In this second dataset, we assembled 2,959 bp of unlinked multi-locus gene fragments (*CR-16S-ND2-CXCR4-POMC-RAG1-Rho*) from 267 individuals of 24 recognised taxa (Supplementary file 1J). (3) Estimation of the divergence dates and reconstruction of the ancestral ranges of the Eastern Palearctic *Bufo* genus from mtDNA and nuDNA sequences data. Here, we refer to the unlinked mtDNA (*CR-ND2*: 894 bp; $N = 221$) as dataset 3a and the unlinked nuDNA data (*POMC-RAG-1-Rho*: 1030 bp; $N = 44$) as dataset 3b. We based the time tree calibrations on primary and secondary calibrations, using a combination of fossil-based estimates, paleogeological events and literature data under various combinations of molecular clocks and tree priors (see detailed calibration points in Appendix 1).

Section 1: Reconstruction of Holarctic bufonids biogeography

We calibrated the species tree under the fossilised birth-death process for 39 recognised species of Holarctic bufonids. To do so, we included species groups with documented fossils or close fossils relatives. To estimate the divergence date, we employed a relaxed molecular clock model with the fossilised birth-death and Yule as tree priors. We then determined the best site model for each unlinked marker (Supplementary file 1K) using jModelTest 2 v.2.1.10 (Darriba et al., 2015). We relied on six calibration points based on the age ranges of the fossil records, and assumed the age as minimum constraint (Donoghue and Yang, 2016).

We retrieved the documented fossil record for all Holarctic bufonids from the database of vertebrates (FosFARbase, accessed 01 June 2020; URL: <http://www.wahre-staerke.com/>) to

determine our calibration points. We then verified the accuracy of each fossil's age range by cross-referencing the literature (Supplementary file 1B). To ensure the reliability of the calibration, we excluded the commonly used Nearctic fossil of "*Bufo*" *praeivius* (20 Mya; Supplementary file 1B) due to known ambiguities (Pauly et al., 2004) and the only conditional attribution of this species to the genus. Instead, we relied on the approximation of the age of their oldest fossil relatives (*Rhinella* of Southern America; Supplementary file 1B). The detailed descriptions for each calibration point are further elaborated in the Appendix 1).

We performed four independent analyses using the Markov Chain Monte-Carlo (MCMC) procedure for 200 million iterations using StarBEAST with the Relaxed Clock (*BEASTRLC) package installed through Beast v.2.6.3 (Bouckaert et al., 2019). We ensured the adequacy of the MCMC samplings by assessing the effective sample size (ESS) values of each parameter (ESS > 200) using Tracer v.1.7 (Rambaut et al., 2018). We assembled all the generated trees in LogCombiner v.2.6.1 (Bouckaert et al., 2019) and summarized the trees with a Maximum Clade Credibility (MCC) tree after discarding 25% of the trees with a posterior probability limit of 0.5 using Tree Annotator v2.6.3 (Bouckaert et al., 2019).

Section 2: Inferring patterns of diversification of Eastern Palearctic *Bufo*

Species tree topology estimation for Palearctic *Bufo*

We estimated the most probable topology of the species tree for the Palearctic *Bufo*. We emphasised the inclusion of all optimum factors that control the topology of the tree, we also included species from the former genus *Torrentophryne*, which is currently synonymised with East Asian *Bufo* in our datasets.

We tested five alternative scenarios for the species tree topology, derived from relevant geological, ecological and life history data associated with the phylogeny of East Asian bufonids.

The factors shaping the tree topology model for each species and detailed descriptions are as such: (1) geographic range and isolation by distance, (2) life history (i.e. terrestrial or semi-aquatic) and adaptive morphological traits (i.e. visible tympanums vs. lacking of external tympanum, presence or absence of adhesive abdomen in tadpoles), (3) a single origin for the eastern Palearctic *Bufo* lineages in relation with past geological events (i.e. QTP Miocene orogenesis and continental plates shifts; Igawa et al., 2006; Macey et al., 1998), (4) alternative topology 1: two independent origins to the Eastern Palearctic *Bufo*: the East Asian mainland *Bufo* vicariance followed by “*Torrentophryne*” radiations around the QTP regions, and the Japanese *Bufo* divergence directly from the Western Palearctic ancestor through colonisation before the Miocene Pacific drift, and (5) alternative topology 2: a single origin of the Eastern Palearctic *Bufo* from the QTP vicariance resulting into the following monophyletic clades: the East Asian mainland *Bufo*, the Japanese *Bufo*, and high elevation “*Torrentophryne*” group in Tibet.

Here, we set *Bufo* as our external calibration point. Four internal calibration dates for ingroup taxa (*Bufo* genus) were adapted from the time frame we determined during the previous step for the Holarctic bufonids dataset using the fossilised birth-death method. We employed a relaxed molecular clock and Yule prior for each designated species tree dataset. We ran two independent analyses with MCMC samplings for 30 million iterations using the *BeastRLC package (Heled and Drummond, 2010) implemented in Beast v.2.6.3 (Bouckaert et al., 2019). The convergence of the MCMC runs of each analysis were assessed by the ESS values obtained for each parameter (ESS > 200) using Tracer v.1.7 (Rambaut et al., 2018). Finally, we evaluated all the supports for species tree topology models with the nested sampling analyses using the nested sampling (NS) package (Maturana et al., 2019) installed through Beast v.2.6.3 (Bouckaert

et al., 2019). We set the sampling parameters as follows: particle count = 20, the length of MCMC sub chains = 20,000, and epsilon = 10^{-6} .

Phylogenetics and divergence dating of Asian *Bufo*

To reconstruct the phylogenetic relationship, we searched for the best-fit evolutionary model and partitioning scheme for each nucleotides dataset (Supplementary file 1M) using PartitionFinder v.2.1.1 (Lanfear et al., 2017). We evaluated six partitions for the concatenated mtDNA dataset, considering three partitions represented a *CR* fragment and two partial introns of *ND2*, and three codon reading frames for exon of *ND2* (Supplementary file 1M). We adjusted the setting to suit “Mr Bayes” mode, using the greedy algorithm. We selected optimal partitioning schemes, and the best-fit substitution models using the Bayesian Information Criterion (BIC). We implemented the best-fit model for each sequence data based on the suggested partitions.

In order to evaluate the recombination pattern in the nuDNA sequences, we conducted a phi test analysis on the concatenated sequences of 1,030 bp of *POMC-RAG-I-Rho* ($N = 44$) using Splitsree v.4.14.6 (Huson, 1998). The phi test did not statistically support the presence of significant recombination ($p = 0.82$) in the nuDNA dataset and the data was therefore reliable for further phylogenetic analyses. We obtained the best-fit evolutionary model for the 11 partitions, evaluated on each partial intron of *POMC* and *Rho*, and three codons position for each exon of *POMC*, *RAGI* and *Rho* (Supplementary file 1M).

For both mtDNA and nuDNA datasets, we analysed the haplotypes with median-joining method (Bandelt et al., 1999) and inferred the network using POPART v.1.7 (Leigh and Bryant, 2015). We reconstructed concatenated genes trees with Bayesian Inference (BI) method for 100 million of MCMC chains with 20% burn-in using Mr. Bayes v.3.2.6 (Huelsenbeck and Ronquist, 2001), until the run reached convergence with a standard deviation values lower than 0.05. We

repeated the reconstruction of the trees following the maximum likelihood (ML) approach using IQ-TREE implemented in the web version of W-IQ-TREE (<http://iqtree.cibiv.univie.ac.at>) (Trifinopoulos et al., 2016). To infer the population-based structure, we ran population genetic statistical analyses such as genetic differentiation, neutrality tests, and hierarchical analysis of molecular variance (AMOVA) on the mtDNA data (Appendix 1). We inferred the potential genetic cluster (K) on the nuclear data, and further examined the introgression patterns following population admixture analysis (Appendix 1).

Then, we estimated the divergence dates of the East Asian *Bufo* clades on the unlinked data of mtDNA (dataset 3a) and nuDNA (dataset 3b) independently. We first examined the most appropriate molecular clock to be used for the datasets (Appendix 1, Supplementary file 1N). As a result, we employed a relaxed molecular clock model for the divergence dates estimates with the mtDNA dataset, and a strict clock model for the nuDNA dataset (Supplementary file 1N). We reconstructed the time tree based on the fixed topology obtained from the phylogenetic relationship through ML and BI trees. For the mtDNA data, the topology of the time tree was also consistent with previous study (Borzée et al., 2017). We set an external calibration point based on the fossil of *Bufo* as outgroup, and seven internal calibration points for the focal clades (ingroup), relying on combined evidences from fossil records, paleogeological events and secondary calibration from the literature. We relied on a log normal distribution with real space mean for the calibrations of fossil records, and we used a normal distribution for the calibration of the paleogeological events and the secondary calibrations from the literature. The enforced internal calibrations are described in more details (Appendix 1).

To reduce bias in prior choice, we tested the mtDNA time tree using the same seven internal calibration points and different combinations of relaxed clocks and tree priors such as

Yule and birth-death. We performed all divergence dating analyses in BEAST v.2.5.2 (Bouckaert et al., 2019). To maintain consistency, each MCMC analysis was run for 30 million generations, with two independent runs. We diagnosed the stationary states of all parameters using Tracer v.1.7 (Rambaut et al., 2018). We then used tree annotator (Rambaut et al., 2018) to sort a maximum clade credibility (MCC) tree to summarise all trees generated with a median height, after 25% burn-in and 0.5 as minimum posterior probability. We further inferred the historical biogeography of the genus to test possible ancestral areas for each clade using the Bayesian dispersal-vicariance (BBM-DIVA) in RASP v.3.1 (Yu et al., 2015).

Past population dynamics

We tested the hypothesis of an ice age refugia (Borzée et al., 2017) against the earlier hypotheses of vicariance during the QTP followed by dispersal (Macey et al., 1998) and the hypothesis of dispersion through long dispersal (Fu et al., 2005). We assessed support for these hypotheses by examining past population dynamics for the two distinct mtDNA *B. gargarizans* clades across a major biogeography barrier, the Yellow Sea. By doing so, we also gained a phylogeographic insight into the impact of the Yellow Sea on the demographic history of *B. gargarizans* during the ice age.

We divided our data into two sets following the clades identified by the mtDNA phylogenetic tree. These clades, hereafter referred to as the southwestern ($N = 75$) and the septentrional East Asia ($N = 123$) clades, were also geographically segregated by the Yellow Sea margins. To estimate the past population dynamics since the Last Glacial Maximum (LGM) period, we constructed for each clade a Bayesian Skyline Plot (Drummond et al., 2005) inferred from the 819 bp of concatenated *CR-ND2* using BEAST v 2.5.2 (Bouckaert et al., 2019). We calibrated the datasets with a combination of strict molecular clock and the Coalescent Bayesian

Skyline as prior, under a lognormal distribution. To calibrate the southwestern clade, we set the mean in the real space to 2.5, standard deviation to 0.5, and offset to 0.0, creating an interval of HPD 5% to 95% that ranged between 0.96 Mya and 5.02 Mya. For the calibration of the septentrional East Asian clade, we set the mean in the real space to 0.8, standard deviation to 0.5, and offset to 0.0, creating an interval of HPD 5%: 0.30 Mya to 95%: 1.84 Mya. We ran two independent analyses on each dataset with 5 million MCMC iterations and a burn-in of 1,000 samples. We looked for the stationary state of the runs and evaluated the ESS values obtained for all the parameters with Tracer v.1.7 (Rambaut et al., 2018), before plotting the Bayesian Skyline Plot with the same program.

Ecological niche modelling

Given the support for two clades within the *B. gargarizans* complex, here named *B. gargarizans* and *B. sachalinensis*, we determined the boundaries of each clade through ecological niche modelling. First, we determined the niche overlap between the two clades, and computed habitat suitability models for each clade using maximum entropy (MaxEnt) modelling (Phillips et al., 2017). To build our model, we combined the geographic information linked to the genetic data and occurrence data downloaded from GBIF (GBIF.org, accessed 27 August 2020; DOI: <https://doi.org/10.15468/dl.ugtsma>). We used only datapoints from the Asian mainland, and manually checked each datapoint. We then assigned each point to either *B. gargarizans* or *B. sachalinensis* based on the geographic distance to the nearest locality with genetic data, resulting in 404 occurrences for *B. gargarizans* (375 GBIF, 29 genetics) and 1076 occurrence for *B. sachalinensis* (1053 GBIF, 23 genetics). Three genetic datapoints belonging to the septentrional East Asian clade were located within range of the western mainland clade. These geographically

discordant datapoints are likely the result of human introduction, only the datapoint assigned to the nominal clade were used for modelling.

To build the models, we used 19 bioclimatic variables (Supplementary file 1O) and three terrain variables extracted from a digital elevation model (DEM; United States Geological Survey) including elevation, slope, and mean slope within a grid cell of 12.27 by 12.27 km. The spatial resolution for all environmental layers was 0.0417 decimal degrees, or approximately 4.1 km. Duplicate presence points (points in the same grid cell in environmental layers) were removed as an option in the MaxEnt model platform to reduce spatial bias, resulting in a final sample size of 275 independent datapoints for *B. gargarizans* and 468 independent datapoints for *B. sachalinensis*. We ran MaxEnt for ten bootstrap replicates with a random test percentage of 20%, and we used the averages of all replicates as final models. Here we decided to include all 22 layers, despite the risk of correlation, as we needed to use the same variables for the two clades for downstream analyses and to avoid the exclusion of relevant variables, based on the previous method (Borée et al., 2020). Models were evaluated using area under the curve (AUC) and true scale statistic (TSS) (Allouche et al., 2006). The niche overlap of the models was evaluated using the niche overlap function in the ‘dismo’ package (Hijmans et al., 2015) in R version 3.5.1 and RStudio version 1.1.442. Both “I” and “D” statistics were calculated.

To determine whether niche overlap of the two *Bufo* clades was significantly different than expected if the two clades had the same environmental requirements, we created a null distribution using MaxEnt modelling. To simulate randomized occurrences, we pooled occurrences from both clades and then randomly assigned individuals to *B. gargarizans* (404 points) or *B. sachalinensis* (1076 points). We then ran MaxEnt using the same parameters as described above for 20 runs and calculated “I” and “D” overlap statistics between corresponding

pairs of suitability outputs. This created a null distribution that we then tested against the niche overlap values calculated from the initial run separating the two clades using a one sample Student's t-Test. Finally, we reconstructed maps in ArcMap 10.6 (ESRI, Redlands, USA).

Estimation of migration patterns

Finally, we estimated the impact of evolutionary forces such as gene flow on the population structure of the septentrional East Asian clades. In order to measure the magnitude and direction of gene flow, we carried out a migration test on a total of 60 *Bufo* individuals representing three sub-populations in the northern latitudes herein defined as: northeastern mainland ($N = 12$), Korean Peninsula ($N = 24$) and Amur River Basin ($N = 24$). Following the assumption that the mutation rate varies among loci and migration is asymmetrical between population, we computed the migration pattern from the following unlinked loci *CR* (894 bp), *ND2* (536 bp), *POMC* (496 bp) and *RAG-1* (308 bp). We set the mutation rate to vary among the loci and standardised the migration rate as $4Nm = 1.0$.

In addition, we estimated the migration pattern of *B. stejnegeri* on the Korean Peninsula, between the northern and southern populations. This analysis was based on 12 individuals sampled north ($N = 9$) and south ($N = 3$) of the Republic of Korea and aimed at testing whether migration is significantly higher in one direction than the other. We conducted this analysis using a full migration matrix model in MIGRATE v4.4.3 (Beerli et al., 2019). We computed two independent analyses for each locus following the MCMC method for 10 million iterations, with four parallel chains and a burn-in of 1,000. We evaluated the ESS value of all parameters sampled for each analysis with Tracers v.1.7 (Rambaut et al., 2018).

Species delimitation modelling

Due to the uncertainty in the taxonomic placement of many East Asian Bufonids, we examined the species boundaries of the Asian *Bufo*, with emphasis of the five putative species within the *B. gargarizans* species complex. We tested different species delimitation scenarios such as sympatric ranges and intergradation zone between subspecies. To account for human-led population displacements, we excluded *B. gargarizans popei* from the species delimitation model itself, and lumped it with *B. gargarizans gargarizans* under *B. gargarizans*. We relied on synonymised subspecies clades to define geographic areas (Pyron and Wiens, 2011), hence the designated models included all clades of the *B. gargarizans* complex in the mainland with unresolved taxonomy: *B. andrewsi*, *B. bankorensis*, *B. gargarizans* from the central East Asian mainland of its range, and septentrional East Asian *B. gargarizans*. We also included the two Japanese *Bufo* subspecies, *B. j. formosus* and *B. j. japonicus*, as indicated by the two distinct clades in our phylogeny, and the Korean water toad, *B. stejnegeri*. We used previously named clades that are now synonymised to test all models possible, and reintroduced synonymised names for the purpose of referring to these clades in our models.

We justified the use of these names following the earliest taxonomically valid binomial name (for which type specimen and type locality have been recorded; Supplementary file 1A, Figure 1-figure supplement 1). The clade of septentrional East Asian *Bufo* distributed on the Korean Peninsula is segregated from *B. gargarizans* (Matsui, 1986), and we deviated from the name “*asiaticus*” originating from the described species “*B. vulgaris* var. *asiaticus*” Steindachner, 1867 as the type locality for the “*asiaticus*” clade is located within the range of *B. gargarizans* (type locality: Shanghai, China; Supplementary file 1A, Figure 1-figure supplement 1). We leave the Korean clade unnamed, hereafter referring to it as “*B. sachalinensis* cf. *sachalinensis*”. The clade distributed in the Amur River Basin (referred to as Far Eastern Russia

in other works) was referred to as *B. sachalinensis sachalinensis*, in reference to “*B. vulgaris* var. *sachalinensis*” Nikolsky, 1905 (type locality: Sakhalin Island, Russia; Supplementary file 1A, Figure 1-figure supplement 1). See the results section for taxonomic resolutions regarding the justification of these names.

We designed eight competing models for species delineation derived from 1,030 bp of concatenated nuDNA (*POMC-RAG-1-Rho*, *N* sequence = 16). We set the species delimitation model under splitting or lumping scenarios. We tested the presence of a single evolutionary lineage for disputed clades within the *B. gargarizans* complex in the mainland, such as the validity of *B. bankorensis* and *B. andrewsi*. We also evaluated the species boundary between the cryptic clades of *B. sachalinensis* cf. *sachalinensis* and *B. s. sachalinensis* on the Korean Peninsula and the Amur River Basin, respectively. Detailed descriptions of the eight alternative species delimitation models are provided in Supplementary file 1I.

Using a coalescent-based approach, we ran the eight species delimitation datasets under the Bayes Factor Delimitation (BFD) method with SNAPP Package (Bryant et al., 2012) implemented in Beast v.2.5.2 (Bouckaert et al., 2019). We calculated the coalescent rates and set the mutation rate to 53.66 (U) and 0.504 (V). We then chose gamma mutation models with log likelihood correction selected. We set the lambda prior based on the number of taxa in the designated model. We ran each dataset separately with two independent analyses before fixing the final parameter in the operator. We ran the analyses with the MCMC procedure to 1 million generation and we sampled every 1,000. We evaluated the convergence state of each model by ensuring that the effective samples size (ESS value) obtained was higher than 200 for each parameter using Tracer v1.7.1 (Rambaut et al., 2018).

To compare the support of each species delimitation model, we then ran a stepping stone/path sampling analysis with Path sampler in the Model Selection Package (Maturana et al., 2019) installed in Beast v.2.5.2 (Bouckaert et al., 2019). For each dataset, we used 0.5 million chains length and eight steps in the paths. We computed the path sampling analyses with the MCMC procedure for 1 million generations until obtaining the marginal likelihood estimation (MLE) value. We determined the rank of support for each species delimitation model using the MLE and Bayes factor values. We calculated Bayes factor with the formula: $\text{Bayes factor} = 2 \times (\text{MLE value of model } X_0) - (\text{MLE value of model } X_1)$. The species delimitation model with a positive Bayes factor value was the most favourable model, relative to models with negative Bayes factor value.

Data availability

Sampling in the Republic of Korea were collected in 2017 under the Ministerial authorisation number 2017–16, and the samples from Jirisan National Park were collected under the Ministerial authorisation number 2019–01. Samples from the People’s Republic of China were collected under the authorisation provided by Nanjing Forestry University. IACUC permit is not required for the in-situ experiment in this study, in accord to the rules of Ewha Woman’s University Institutional Biosafety Committee.

The electronic edition of this article conforms to the requirements of the amended International Code of Zoological Nomenclature, and hence the new name contained herein are available under that Code from the electronic edition of this article. This published work and the nomenclatural acts it contains have been registered in ZooBank, the online registration system for the ICZN. The ZooBank LSIDs (Life Science Identifiers) can be resolved and the associated

information viewed through any standard web browser by appending the LSID to the prefix “<http://zoobank.org/>”. The LSID for this publication is: urn:lsid:zoobank.org:pub:XXX. The electronic edition of this work was published in a journal with an ISSN, and has been archived with accessibility from the following digital repositories: PubMed Central.

All Sequences generated in present study deposited to the Genbank database [<https://www.ncbi.nlm.nih.gov/> 927 genbank/] under the accession number MW081664 – MW081847 (*CR*), MW467646-MW467777 (*ND2*), MW489915-MW489964 (*POMC*), MW489986-MW490035 (*RAG-1*), MW507752-MW507780 (*Rho*). Input files in the form of BEAST XML generated for all molecular dating analyses and species delimitation modelling are available from the Mendeley Data repository <http://dx.doi.org/10.17632/wdtw6kn2t4.1> (Othman et al., 2021).

Acknowledgments

The authors thank Sungsik Kong, Jordy Groffen, Yi Yang, Vasiliy I. Kazakov, Ye Inn Kim and Sera Kwon for their commitment and helps during field work. We also thank Hakyung Kang for the keen assistance in obtaining sampling permits from local authorities. We credit Christophe Dufresnes, Antonio Rodríguez Arduengo (*Rhinella marina*), David Leo Berton (*Bufo verrucosissimus*) and Ben Zerante (*Bufo japonicus*) for the photos of toads used in Figure 1. We acknowledge Professor Yeong-Choy Kam of Tunghai University for providing logistic helps during field work. We also thank Jonathan J. Fong and an anonymous reviewer for their comments during the review.

Funding

This work was supported by a Korean Environmental Industry and Technology Institute through ‘Development of behavioural ecological methods for population regulation of invasive species’ funded by Korea Ministry of Environment (KEITI 2017002270003) and a Rural Development Administration of Korea grant (PJ015071) to YJ, the Foreign Youth Talent Program (QN2021014013L) from the Ministry of Science and Technology of the People’s Republic of China to AB and the Russian Foundation of Basic Research (20-04-00918) to SNL.

Competing interests

The authors declare that they do not have competing interests.

References

- Allouche O, Tsoar A, Kadmon R. 2006. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* **43**:1223–1232. doi:10.1111/j.1365-2664.2006.01214.x
- Arntzen JW, Recuero E, Canestrelli D, Martínez-Solano I. 2013. How complex is the *Bufo bufo* species group? *Molecular Phylogenetics and Evolution* **69**:1203–1208. doi:10.1016/j.ympev.2013.07.012
- Bailon S, Hossini S. 1990. The oldest European Bufonidae (Amphibia, Anura): Miocene species from France. *Annales de Paleontologie* **76**:121–132.
- Bandelt H-J, Forster P, Rohl A. 1999. Median-Joining Networks for Inferring Intraspecific Phylogenies. *Molecular Biology* **16**:37–48. doi:10.1093/oxfordjournals.molbev.a026036
- Barbour T. 1908. Some new reptiles and amphibians. *Bulletin of the Museum of Comparative Zoology* **51**:315–325.
- Barnes GL. 2003. Origins of the Japanese Islands: The New “Big Picture.” *Japan Review* **15**:3–

1043 50. doi:10.2307/25791268

1044 Beerli P, Mashayekhi S, Sadeghi M, Khodaei M, Shaw K. 2019. Population Genetic Inference
 1045 With MIGRATE. *Current Protocols in Bioinformatics* **68**:1–28. doi:10.1002/cpbi.87

1046 Bell RC, Irian CG. 2019. Phenotypic and genetic divergence in reed frogs across a mosaic hybrid
 1047 zone on São Tomé Island. *Biological Journal of the Linnean Society* **128**:672–680.
 1048 doi:10.1093/biolinnean/blz131

1049 Böhme M, Ilg A. 2003. fosFARbase. www.wahre-staerke.com/

1050 Borzée A, Kielgast J, Wren S, Angulo A, Chen S, Magellan K, Messenger KR, Hansen-
 1051 Hendrikx CM, Baker A, Santos MMD, Kusrini M, Jiang J, Maslova I V., Das I, Park D,
 1052 Bickford D, Murphy RW, Che J, Van Do T, Nguyen TQ, Chuang MF, Bishop PJ. 2021.
 1053 Using the 2020 global pandemic as a springboard to highlight the need for amphibian
 1054 conservation in eastern Asia. *Biological Conservation* **255**:108973.
 1055 doi:10.1016/j.biocon.2021.108973

1056 Borzée A, Messenger KR, Chae S, Andersen D, Groffen J, Kim YI, An J, Othman SN, Ri K,
 1057 Nam TY, Bae Y, Ren J-L, Li J-T, Chuang MF, Yi Y, Shin Y, Kwon T, Jang Y, Min M-S.
 1058 2020. Yellow sea mediated segregation between North East Asian *Dryophytes* species.
 1059 *PLoS ONE* **15**:e0234299.

1060 Borzée A, Santos JL, Sánchez-Ramírez S, Bae Y, Heo K, Jang Y, Jowers MJ. 2017.
 1061 Phylogeographic and population insights of the Asian common toad (*Bufo gargarizans*) in
 1062 Korea and China: population isolation and expansions as response to the ice ages. *PeerJ*
 1063 **5**:e4044. doi:10.7717/peerj.4044

1064 Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J,
 1065 Jones G, Kühnert D, De Maio N, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, du

- Plessis L, Poppinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu CH, Xie D, Zhang C, Stadler T, Drummond AJ. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS computational biology* **15**:e1006650. doi:10.1371/journal.pcbi.1006650
- Bouckaert RR. 2010. DensiTree: Making sense of sets of phylogenetic trees. *Bioinformatics* **26**:1372–1373. doi:10.1093/bioinformatics/btq110
- Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, Roychoudhury A. 2012. Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution* **29**:1917–1932. doi:10.1093/molbev/mss086
- Cantor T. 1842. General features of Chusan, with remarks on the flora and fauna of that island. *Annals and Magazine of Natural History Series* **1**:481–493.
- Cao Y, Cui K, Pan H, Wu J, Wang L. 2019. The impact of multiple climatic and geographic factors on the chemical defences of Asian toads (*Bufo gargarizans* Cantor). *Scientific Reports* **9**:17236. doi:10.1038/s41598-019-52641-4
- Chen CC, Li KW, Yu TL, Chen LH, Sheu PY, Tong YW, Huang KJ, Weng CF. 2013. Genetic structure of *Bufo bankorensis* distinguished by amplified restriction fragment length polymorphism of Cytochrome b. *Zoological Studies* **52**:1–9. doi:10.1186/1810-522X-52-48
- Darriba D, Taboada GL, Doallo R, Posada D. 2015. jModelTest 2 : more models , new heuristics and high- performance computing. *Europe PMC Funders Group* **9**:6–9. doi:10.1038/nmeth.2109.jModelTest
- Donoghue PCJ, Yang Z. 2016. The evolution of methods for establishing evolutionary timescales. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**:20160020. doi:10.1098/rstb.2016.0020

1089 Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2005. Bayesian coalescent inference of past
 1090 population dynamics from molecular sequences. *Molecular Biology and Evolution*
 1091 **22**:1185–1192. doi:10.1093/molbev/msi103

1092 Dubois A, Bour R. 2010. The nomenclatural status of the nomina of amphibians and reptiles
 1093 created by by Garsault (1764), with a parsimonious solution to an old nomenclatural
 1094 problem regarding the genus *Bufo* (Amphibia, Anura), comments on the taxonomy of this
 1095 genus, and co. *Zootaxa* **2447**:1–52. doi:10.5281/zenodo.195113

1096 Dufresnes C, Berroneau M, Dubey S, Litvinchuk SN, Perrin N. 2020. The effect of
 1097 phylogeographic history on species boundaries: a comparative framework in *Hyla* tree
 1098 frogs. *Scientific Reports* **10**:5502. doi:10.1038/s41598-020-62382-4

1099 Dufresnes C, Mazepa G, Jablonski D, Oliveira RC, Wenseleers T, Shabanov DA, Auer M, Ernst
 1100 R, Koch C, Ramírez-Chaves HE, Mulder KP, Simonov E, Tiutenko A, Kryvokhyzha D,
 1101 Wennekes PL, Zinenko OI, Korshunov O V., Al-Johany AM, Peregontsev EA, Masroor R,
 1102 Betto-Colliard C, Denoël M, Borkin LJ, Skorinov D V., Pasynkova RA, Mazanaeva LF,
 1103 Rosanov JM, Dubey S, Litvinchuk S. 2019. Fifteen shades of green: The evolution of
 1104 *Bufo* toads revisited. *Molecular Phylogenetics and Evolution* **141**:106615.
 1105 doi:10.1016/j.ympev.2019.106615

1106 Earl DA, VonHoldt BM. 2012. STRUCTURE HARVESTER: A website and program for
 1107 visualizing STRUCTURE output and implementing the Evanno method. *Conservation*
 1108 *Genetics Resources* **4**:359–361. doi:10.1007/s12686-011-9548-7

1109 Estes R, Wassersug RJ. 1963. A Miocene toad from Colombia, South America. *Breviora* **193**:1–
 1110 13.

1111 Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the

1112 software STRUCTURE: A simulation study. *Molecular Ecology* **14**:2611–2620.
 1113 doi:10.1111/j.1365-294X.2005.02553.x

1114 Evans BJ, McGuire JA, Brown RM, Andayani N, Supriatna J. 2008. A coalescent framework for
 1115 comparing alternative models of population structure with genetic data: Evolution of
 1116 Celebes toads. *Biology Letters* **4**:430–433. doi:10.1098/rsbl.2008.0166

1117 Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform
 1118 population genetics analyses under Linux and Windows. *Molecular Ecology Resources*
 1119 **10**:564–567. doi:10.1111/j.1755-0998.2010.02847.x

1120 Fong JJ, Yang BT, Li PP, Waldman B, Min MS. 2020. Phylogenetic systematics of the Water
 1121 Toad (*Bufo stejnegeri*) elucidates the evolution of semi-aquatic toad ecology and
 1122 Pleistocene glacial refugia. *Frontiers in Ecology and Evolution* **7**:523.
 1123 doi:10.3389/fevo.2019.00523

1124 Frazão A, Da Silva HR, De Moraes Russo CA. 2015. The Gondwana breakup and the history of
 1125 the Atlantic and Indian Oceans unveils two new clades for early neobatrachian
 1126 diversification. *PLoS ONE* **10**:e0143926. doi:10.1371/journal.pone.0143926

1127 Fu J, Weadick CJ, Zeng X, Wang Y, Liu Z, Zheng Y, Li C, Hu Y. 2005. Phylogeographic
 1128 analysis of the *Bufo gargarizans* species complex: A revisit. *Molecular Phylogenetics and*
 1129 *Evolution* **37**:202–213. doi:10.1016/j.ympev.2005.03.023

1130 Garcia-Porta J, Litvinchuk SN, Crochet PA, Romano A, Geniez PH, Lo-Valvo M, Lymberakis P,
 1131 Carranza S. 2012. Molecular phylogenetics and historical biogeography of the west-
 1132 palearctic common toads (*Bufo bufo* species complex). *Molecular Phylogenetics and*
 1133 *Evolution* **63**:113–130. doi:10.1016/j.ympev.2011.12.019

1134 Guo ZT, Ruddiman WF, Hao QZ, Wu HB, Qiao YS, Zhu RX, Peng SZ, J.J.Wei, Yuan BY, Liu

1135 TS. 2002. Onset of Asian desertification by 22Myr ago inferred from loess deposits in
1136 China. *Nature* **416**:159–163.

1137 Hase K, Nikoh N, Shimada M. 2013. Population admixture and high larval viability among
1138 urban toads. *Ecology and Evolution* **3**:1677–1691. doi:10.1002/ece3.578

1139 Heath TA, Huelsenbeck JP, Stadler T. 2014. The fossilized birth-death process for coherent
1140 calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences*
1141 *of the United States of America* **111**:E2957–E2966. doi:10.1073/pnas.1319091111

1142 Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multilocus data.
1143 *Molecular Biology and Evolution* **27**:570–580. doi:10.1093/molbev/msp274

1144 Hijmans RJ, Phillips S, Leathwick J, Elith J. 2015. Dismo: Species Distribution Modeling. R
1145 Package Version 1.0-12. <http://cran.r-project.org/package=dismo>

1146 Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees.
1147 *Bioinformatics* **17**:754–755. doi:10.1093/bioinformatics/17.8.754

1148 Huson DH. 1998. SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics*
1149 **14**:68–73. doi:10.1093/bioinformatics/14.1.68

1150 Igawa T, Kurabayashi A, Nishioka M, Sumida M. 2006. Molecular phylogenetic relationship of
1151 toads distributed in the Far East and Europe inferred from the nucleotide sequences of
1152 mitochondrial DNA genes. *Molecular Phylogenetics and Evolution* **38**:250–260.
1153 doi:10.1016/j.ympev.2005.09.003

1154 Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorrell JC, Cullingham CI, Andrew RL. 2017.
1155 The K = 2 conundrum. *Molecular Ecology* **26**:3594–3602. doi:10.1111/mec.14187

1156 Kamata H, Kodama K. 1994. Tectonics of an arc-arc junction: an example from Kyushu Island at
1157 the junction of the Southwest Japan Arc and the Ryukyu Arc. *Tectonophysics* **233**:69–81.

doi:10.1016/0040-1951(94)90220-8

Kim S, Tkalčić H, Rhie J. 2017. Seismic constraints on magma evolution beneath Mount Baekdu (Changbai) volcano from transdimensional Bayesian inversion of ambient noise data. *Journal of Geophysical Research: Solid Earth* **122**:5452–5473. doi:10.1002/2017JB014105

Kolbe JJ, Kearney M, Shine R. 2010. Modeling the consequences of thermal trait variation for the cane toad invasion of Australia. *Ecological Applications* **20**:2273–2285. doi:10.1890/09-1973.1

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resource* **15**:1179–1191. doi:10.1002/cnecr.27633.Percutaneous

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017. Partitionfinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* **34**:772–773. doi:10.1093/molbev/msw260

Lee C, Fong JJ, Jiang J-P, Li P-P, Waldman B, Chong JR, Lee H, Min M-S. 2021. Phylogeographic study of the *Bufo gargarizans* species complex, with emphasis on Northeast Asia. *Animal Cells and Systems* **0**:1–11. doi:10.1080/19768354.2021.2015438

Lei B, Xu M, Chen J, Liang J, Zhang Y. 2018. Structural characteristics and evolution of the South Yellow Sea Basin since Indosinian. *China Geology* **1**:466–476. doi:10.31035/cg2018063

Lei F, Qu Y, Song G. 2014. Species diversification and phylogeographical patterns of birds in response to the uplift of the qinghai-tibet plateau and quaternary glaciations. *Current*

1181 *Zoology* **60**:149–161. doi:10.1093/czoolo/60.2.149

1182 Leigh JW, Bryant D. 2015. POPART: Full-feature software for haplotype network construction.

1183 *Methods in Ecology and Evolution* **6**:1110–1116. doi:10.1111/2041-210X.12410

1184 Li J, Huang JP, Sukumaran J, Knowles LL. 2018. Microevolutionary processes impact

1185 macroevolutionary patterns. *BMC Evolutionary Biology* **18**:123. doi:10.1186/s12862-018-

1186 1236-8

1187 Li XS, Zhao YX, Feng ZB, Liu CG, Xie QH, Zhou QJ. 2016. Quaternary seismic facies of the

1188 South Yellow Sea shelf: depositional processes influenced by sea-level change and tectonic

1189 controls. *Geological Journal* **51**:77–95. doi:10.1002/gj.2811

1190 Liao WB, Luo Y, Lou SL, Lu D, Jehle R. 2016. Geographic variation in life-history traits:

1191 growth season affects age structure, egg size and clutch size in Andrew's toad (*Bufo*

1192 *andrewsi*). *Frontiers in Zoology* **13**:6. doi:10.1186/s12983-016-0138-0

1193 Liedtke HC, Müller H, Rödel MO, Menegon M, Gonwouo LGN, Barej MF, Gvoždík V, Schmitz

1194 A, Channing A, Nagel P, Loader SP. 2016. No ecological opportunity signal on a

1195 continental scale? Diversification and life-history evolution of African true toads (Anura:

1196 Bufonidae). *Evolution* **70**:1717–1733. doi:10.1111/evo.12985

1197 Lin C-F, Lue K-Y. 2004. Altitudinal differences in temporal distribution, spatial preference and

1198 timing of breeding climax of frogs and toads in the Central Taiwan. *Endemic Species*

1199 *Research* **6**:39–50.

1200 Lischer HEL, Excoffier L. 2012. PGDSpider: An automated data conversion tool for connecting

1201 population genetics and genomics programs. *Bioinformatics* **28**:298–299.

1202 doi:10.1093/bioinformatics/btr642

1203 Liu W, Lathrop A, Fu J, Yang D, Murphy RW. 2000. Phylogeny of East Asian bufonids inferred

1204 from mitochondrial DNA sequences (Anura: Amphibia). *Molecular Phylogenetics and*
 1205 *Evolution* **14**:423–435. doi:10.1006/mpev.1999.0716
 1206 Luquet E, Léna JP, Miaud C, Plénet S. 2015. Phenotypic divergence of the common toad (*Bufo*
 1207 *bufo*) along an altitudinal gradient: Evidence for local adaptation. *Heredity* **114**:69–79.
 1208 doi:10.1038/hdy.2014.71
 1209 Macey JR, Schulte JA, Larson A, Fang Z, Wang Y, Tuniyev BS, Papenfuss TJ. 1998.
 1210 Phylogenetic relationships of toads in the *Bufo bufo* species group from the eastern
 1211 escarpment of the Tibetan Plateau: a case of vicariance and dispersal. *Molecular*
 1212 *Phylogenetics and Evolution* **9**:80–87. doi:10.1006/mpev.1997.0440
 1213 Mantel N. 1967. The detection of disease clustering and a generalized regression approach.
 1214 *Cancer Research* **27**:209–220.
 1215 Martin SH, Davey JW, Jiggins CD. 2015. Evaluating the use of ABBA-BABA statistics to locate
 1216 introgressed loci. *Molecular Biology and Evolution* **32**:244–257.
 1217 doi:10.1093/molbev/msu269
 1218 Maslova I V. 2016. The protection of amphibians and reptiles in the Russian far east. *Nature*
 1219 *Conservation Research* **1**:26–35. doi:10.24189/ncr.2016.024
 1220 Matsui M. 1986. Geographic variation in Toads of the *Bufo bufo* complex from the Far East,
 1221 with a description of a new subspecies. *Copeia* 561–579.
 1222 Matsui M. 1984. Morphometric variation analyses and revision of the Japanese toads (genus
 1223 *Bufo*, Bufonidae). *Contributions from the Biological Laboratory, Kyoto University* **26**:209–
 1224 428.
 1225 Maturana P, Brewer BJ, Klaere S, Bouckaert RR. 2019. Model Selection and Parameter
 1226 Inference in Phylogenetics Using Nested Sampling. *Systematic Biology* **68**:219–233.

doi:10.1093/sysbio/syy050

McGee S. 2002. Simplifying likelihood ratios. *Journal of General Internal Medicine* **17**:647–650. doi:10.1046/j.1525-1497.2002.10750.x

Mendelson JR, Mulcahy DG, Williams TS, Sites JW. 2011. A phylogeny and evolutionary natural history of mesoamerican toads (Anura: Bufonidae: *Incilius*) based on morphology, life history, and molecular data. *Zootaxa* **3138**:1–34. doi:10.11646/zootaxa.3138.1.1

Miura I. 1995. Two differentiated groups of the Japanese toad, *Bufo japonicus japonicus*, demonstrated by C-banding analysis of chromosomes. *Caryologia* **48**:123–136. doi:10.1080/00087114.1995.10797322

Nokariya H, Hasegawa Y. 1985. Fossil frogs from Pinza-Abu Cave. Miyako Island, Okinawa, Japan. In reports of excavation of Pinza-Abu Cave. Naha: Okinawa Prefectural Board of Education.

Othman SN, Chen Y-H, Chuang M-F, Andersen D, Jang Y, Borzée A. 2020. Impact of the Mid-Pleistocene Revolution and anthropogenic factors on the dispersion of Asian Black-spined toads (*Duttaphrynus melanostictus*). *Animals* **10**:1157. doi:https://doi.org/10.3390/ani10071157

Othman SN, Putri ET, Messenger KR, Bae Y, Yang Y, Bova T, Reed T, Amin H, Chuang M-F, Jang Y, Borzée A. 2021. Impact of the Miocene orogenesis on *Kaloula* spp. radiation, and implication of local refugia on genetic diversification. *Integrative Zoology* **0**:1–24. doi:https://doi.org/10.1111/1749-4877.12538

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. *Genetics* **192**:1065–1093. doi:10.1534/genetics.112.145037

1250 Pauly GB, Hillis DM, Cannatella DC. 2004. The history of a nearctic colonization: Molecular
 1251 phylogenetics and biogeography of the nearctic toads (*Bufo*). *Evolution* **58**:2517–2535.
 1252 doi:10.1111/j.0014-3820.2004.tb00881.x

1253 Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME. 2017. Opening the black box: an
 1254 open-source release of Maxent. *Ecography* **40**:887–893. doi:10.1111/ecog.03049

1255 Portik DM, Papenfuss TJ. 2015. Historical biogeography resolves the origins of endemic
 1256 Arabian toad lineages (Anura: Bufonidae): Evidence for ancient vicariance and dispersal
 1257 events with the Horn of Africa and South Asia. *BMC Evolutionary Biology* **15**:152.
 1258 doi:10.1186/s12862-015-0417-y

1259 Pramuk JB. 2006. Phylogeny of South American *Bufo* (Anura: Bufonidae) inferred from
 1260 combined evidence. *Zoological Journal of the Linnean Society* **146**:407–452.
 1261 doi:10.1111/j.1096-3642.2006.00212.x

1262 Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus
 1263 genotype data. *Genetics* **155**:945–959. doi:10.1007/s10681-008-9788-0

1264 Pyron RA, Wiens JJ. 2011. A large-scale phylogeny of Amphibia including over 2800 species,
 1265 and a revised classification of extant frogs, salamanders, and caecilians. *Molecular*
 1266 *Phylogenetics and Evolution* **61**:543–583. doi:10.1016/j.ympev.2011.06.012

1267 R Core Team. 2020. R: A language and environment for statistical computing. Vienna, Austria:
 1268 R Foundation for Statistical Computing.

1269 Rage JC, Roček Z. 2003. Evolution of anuran assemblages in the Tertiary and Quaternary of
 1270 Europe, in the context of palaeoclimate and palaeogeography. *Amphibia Reptilia* **24**:133–
 1271 167. doi:10.1163/156853803322390408

1272 Ramasamy RK, Ramasamy S, Bindroo BB, Naik VG. 2014. STRUCTURE PLOT: A program

1273 for drawing elegant STRUCTURE bar plots in user friendly interface. *SpringerPlus* **3**:431.
1274 doi:10.1186/2193-1801-3-431

1275 Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in
1276 Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**:901–904.
1277 doi:10.1093/sysbio/syy032

1278 Ratnikov VY. 2001. Osteology of Russian toads and frogs for paleontological researches. *Acta*
1279 *zoologica cracoviensia* **44**:1–23.

1280 Recuero E, Canestrelli D, Vörös J, Szabó K, Poyarkov NA, Arntzen JW, Crnobrnja-Isailovic J,
1281 Kidov AA, Cogălniceanu D, Caputo FP, Nascetti G, Martínez-Solano I. 2012. Multilocus
1282 species tree analyses resolve the radiation of the widespread *Bufo bufo* species group
1283 (Anura, Bufonidae). *Molecular Phylogenetics and Evolution* **62**:71–86.
1284 doi:10.1016/j.ympev.2011.09.008

1285 Ricklefs RE. 2006. Evolutionary diversification and the origin of the diversity-environment
1286 relationship. *Ecology* **87**:3–13. doi:10.1890/0012-9658(2006)87[3:edatoo]2.0.co;2

1287 Rojas RR, Fouquet A, Ron SR, Hernández-Ruz EJ, Melo-Sampaio PR, Chaparro JC, Vogt RC,
1288 Carvalho VT de, Pinheiro LC, Avila RW, Farias IP, Gordo M, Hrbek T. 2018. A Pan-
1289 Amazonian species delimitation: High species diversity within the genus *Amazophrynella*
1290 (Anura: Bufonidae). *PeerJ* **6**:e4941. doi:10.7717/peerj.4941

1291 San Mauro D, Gower DJ, Oommen O V., Wilkinson M, Zardoya R. 2004. Phylogeny of
1292 caecilian amphibians (Gymnophiona) based on complete mitochondrial genomes and
1293 nuclear RAG1. *Molecular Phylogenetics and Evolution* **33**:413–427.
1294 doi:10.1016/j.ympev.2004.05.014

1295 Sanchiz B. 1998. Salientia In: Wellnhofer P, editor. *Encyclopedia of Paleoherpetology* 4/4.

1296 Munich: Dr. Friedrich Pfeil. pp. 1– 275.

1297 Sanchiz B. 1997. Factores causales de biodiversidad: especiación, extinción y migración en la
1298 formación de la batracofauna europea In: Aguirre E, Morales J, Soria D, editors. Registros
1299 Fósiles e Historia de La Tierra. Madrid, Ed. Complutense. pp. 185–206.

1300 Schmidt KP. 1925. New Chinese amphibians and reptiles. *American Museum Novitates* **175**:1–3.

1301 Stejneger L. 1926. A new toad from China. *J Washington Acad Sci* **16**:445–446.

1302 Stöck M, Moritz C, Hickerson M, Frynta D, Dujsebayaeva T, Eremchenko V, Macey JR,
1303 Papenfuss TJ, Wake DB. 2006. Evolution of mitochondrial relationships and biogeography
1304 of Palearctic green toads (*Bufo viridis* subgroup) with insights in their genomic plasticity.
1305 *Molecular Phylogenetics and Evolution* **41**:663–689. doi:10.1016/j.ympev.2006.05.026

1306 Suzuki D, Kawase T, Hoshina T, Tokuda T. 2020. Origins of nonnative populations of *Bufo*
1307 *japonicus formosus* (Amphibia: Bufonidae) in Hokkaido, Japan, as inferred by a molecular
1308 approach. *Current Herpetology* **39**:47–54.

1309 Syromyatnikova E V. 2015. A new species of *Bufo* (Amphibia, anura) from the Miocene of
1310 Russia. *Russian Journal of Herpetology* **22**:281–288.

1311 Taira A. 2001. Tectonic Evolution of the Japanese Island Arc System. *Annual Review of Earth*
1312 *and Planetary Sciences* **29**:109–134. doi:10.1146/annurev.earth.29.1.109

1313 Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a fast online
1314 phylogenetic tool for maximum likelihood analysis. *Nucleic acids research* **44**:W232–
1315 W235. doi:10.1093/nar/gkw256

1316 Tsuji H, Kawamichi T. 1996. Breeding Activity of a Stream-Breeding Toad, *Bufo torrenticola*.
1317 *Japanese Journal of Herpetology* **16**:117–128. doi:10.5358/hsj1972.16.4_117

1318 Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012.

1319 Primer3-new capabilities and interfaces. *Nucleic Acids Research* **40**:1–12.
1320 doi:10.1093/nar/gks596

1321 Van Bocxlaer I, Biju S, Loader SP, Bossuyt F. 2009. Toad radiation reveals into-India dispersal
1322 as a source of endemism in the Western Ghats-Sri Lanka biodiversity hotspot. *BMC*
1323 *Evolutionary Biology* **9**:131. doi:10.1186/1471-2148-9-131

1324 Van Bocxlaer I, Loader SP, Roelants K, Biju SD, Menegon M, Bossuyt F. 2010. Gradual
1325 adaptation toward a range-expansion phenotype initiated the global radiation of toads.
1326 *Science* **327**:679–682. doi:10.1126/science.1181707

1327 Xing Y, Ree RH. 2017. Uplift-driven diversification in the Hengduan Mountains, a temperate
1328 biodiversity hotspot. *Proceedings of the National Academy of Sciences of the United States*
1329 *of America* **114**:E3444–E3451. doi:10.1073/pnas.1616063114

1330 Yu TL, Lin H Du, Weng CF. 2014. A new phylogeographic pattern of endemic *Bufo bankorensis*
1331 in Taiwan Island is attributed to the genetic variation of populations. *PLoS ONE* **9**:e98029.
1332 doi:10.1371/journal.pone.0098029

1333 Yu Y, Harris AJ, Blair C, He X. 2015. RASP (Reconstruct Ancestral State in Phylogenies): A
1334 tool for historical biogeography. *Molecular Phylogenetics and Evolution* **87**:46–49.
1335 doi:10.1016/j.ympev.2015.03.008

1336 Zhan A, Fu J. 2011. Past and present: Phylogeography of the *Bufo gargarizans* species complex
1337 inferred from multi-loci allele sequence and frequency data. *Molecular Phylogenetics and*
1338 *Evolution* **61**:136–148. doi:10.1016/j.ympev.2011.06.009

1339 Zhan X, Wu Huan, Wu Hong, Wang R, Luo C, Gao B, Chen Z, Li Q. 2020. Metabolites from
1340 *Bufo gargarizans* (Cantor, 1842): A review of traditional uses, pharmacological activity,
1341 toxicity and quality control. *Journal of Ethnopharmacology* **246**:112178.

1342 doi:10.1016/j.jep.2019.112178

1343 Zhang ML, Xiang XG, Xue JJ, Sanderson SC, Fritsch PW. 2016. Himalayan uplift shaped

1344 biomes in Miocene temperate Asia: Evidence from leguminous *Caragana*. *Scientific*

1345 *Reports* **6**:1–7. doi:10.1038/srep36528

1346 Zhao Y, Yu Z. 2012. Vegetation response to Holocene climate change in East Asian monsoon-

1347 margin region. *Earth-Science Reviews* **113**:1–10. doi:10.1016/j.earscirev.2012.03.001

1348

Figure and Figure Supplement Legends

Figure 1 Resolution of the diversifications of bufonids in Holarctic and the Eastern

Paelearctic. (A) Dating of the species tree of Holarctic bufonids, refined using the fossilized birth-death method. Here the box-plots represent the age-range of fossils for each focal species. **(B)** Hypothesised dispersal pathways of genus *Bufo* in the Palearctic based on the Gondwana origin and central Asia desertification hypotheses. The map displays the range of the bufonids genera used in the dating analyses. **(C)** Molecular dating estimates and ancestral range reconstruction addressing three phylogeographic hypotheses derived from Macey et al. (1998), Fu et al. (2005) and Borzée et al. (2017) to elucidate the evolutionary history of East Asian *Bufo*. **(D)** Retracing the hidden diversity in the *Bufo gargarizans* species complex in East Asia and determining taxonomic and conservation units.

Figure 1-figure supplement 1. Chronology of species description in the East Asian *Bufo*

genus and timeline of taxonomy updates. (A) The map indicates the location of the type locality for each described *Bufo* taxa in East Asia since the 18th century. **(B)** The timeline of taxonomic revisions recorded in the literature using various approach on East Asian *Bufo* since 1980.

Figure 2 Fossils used for calibration and molecular dating of Holarctic Bufonidae. (A)

Fossilised birth-death species trees for 39 Holarctic bufonids characterised from the unlinked multi-locus *CR-16S-ND2-CXCR4-POMC-RAG1-Rho* with an abundance of fossils representative of the six genera of Bufonidae used as source of primary calibrations. The captions (i) and (ii) on

the tree are geographically explained in **(B)**, representing the hypothesised dispersal pathways of Holarctic bufonids and *Bufo* spp. in the Palearctic. Similarly, the letter coded branches of the trees are marked on the map. The map displays the range of bufonids genera (IUCN, 2020) with relevant natural features in Central Asia. Black, red and yellow arrows in dated tree and maps indicate the dispersal pathways predicted for neotropic bufonids, *Bufo* and *Bufo* in the Palearctic, respectively.

Figure 2-figure supplement 1. Fossil abundance and distribution of Holarctic bufonids. The record of fossil distribution are retrieved from the open database FosFARbase (Böhme and Ilg, 2003) and the distribution of bufonids fossil embedded with the global paleomap of Early Miocene to Pleistocene modelled by Matthews et al. (2016), Cao et al. (2018) and Ray et al. (2001) for molecular dating using fossilised birth-death method. The species with fossil record in bar chart are represented by different colors and matching with the colour of fossils site in the maps.

Figure 3 Species tree topology estimates based on hypotheses for 26 recognised Palearctic bufonids. **(A)** Geographic range of the Palearctic *Bufo* genus and related “*Torrentophryne*” species included in the reconstruction of the species tree. **(B)** Ranking of the topology models for the species tree examined using nested sampling analyses. **(C)** Best-supported topology on dated species trees reconstructed from the unlinked multilocus data (*CR-16S-ND2-CXCR4-POMC-RAG1-Rho*) under a relaxed clock and Yule prior. Each model tree (model A to E) represents the hypothesis tested for the tree topology with the rank of its likelihood based on the marginal

likelihood estimation (MLE) values. The geology driven factor (model C) was selected to be the most accurate scenario for the species tree topology due to the highest value of MLE.

Figure 3-figure supplement 1. Alternative models of species tree topologies for Palearctic

Bufo. Each model is tested with nested sampling analyses and support is based on the MLE values. **(A)** Rank of support for tree topologies tested on different hypotheses. **(B)** The four suboptimal models here have lower MLE values than the best supported model C in Figure 3.

Figure 4 Discordance in phylogenetic and haplotype relationships of the *Bufo* genus in East Asia inferred from mtDNA and nuclear protein coding. **(A)** Median joining network based on

mtDNA *CR-ND2* (894 bp) obtained from 221 individuals. **(B)** Phylogenetic relationships inferred from the same dataset of concatenated mtDNA *CR-ND2*. **(C)** Haplotype relationship of 44 individuals of East Asian *Bufo* derived from diploid data (nuclear *POMC-RAG-1-Rho*; 1030 bp). The discordance in topology between the mitochondrial and nuclear trees are detailed in Table 2. **(D)** Phylogenetic relationship and population clustering inferred from SNP of the nuclear protein coding analysed in STRUCTURE and CLUMPP (see Supplementary file 1F for details about the populations). The colour code used in the map of East Asia match with the colours coded for each particular clade and haplotype group.

Figure 4-figure supplement 1. The tests of isolation by distance on *Bufo* distributed in the Eastern Asia. **(A)** Mantel test for mitochondrial *CR-ND2* (N individual = 224, N population = 13). **(B)** Mantel test for nuclear *POMC-RAG-1-Rho* (N = 49, n population = 9).

Figure 4-figure supplement 2. The phylogenetic relationship between Bufonidae with emphasis on East Asian *Bufo* lineage inferred on partial 16S fragment. The grey shaded box highlights the paraphyletic of the placement of *Bufo gargarizans* sampled in Vietnam (Clade C) in a different subclade segregated from the low-elevated *B. gargarizans* clades from Eastern Mainland (Clade A) and the septentrional East Asia (Clade B). The name of each taxa in tree represents the names of sampled individuals.

Figure 5 Dated phylogeny and biogeography patterns of *Bufo* in the Eastern Palearctic. (A) Hypothesised dispersal pathways for *Bufo* in the Eastern Palearctic. **(B)** Dated Maximum Clade Credibility (MCC) tree, ancestral ranges reconstruction and colonisation history for Palearctic bufonids inferred from the linked mtDNA *CR-ND2* (*N* individuals = 132). The dispersal pathways illustrated indicate two Miocene vicariance events for the MRCA of East Asian *Bufo* (A1-A2) and the routes hypothesised to have been followed by the *B. gargarizans* complex for dispersion through the Asian mainland (A3). The Miocene tectonic plates models and the Tibetan-Himalayan mapping were both retrieved from established datasets (refer to Supplementary file 1O) and projected to this paleomap.

Figure 6 Ice age refugia hypothesis for two clades of *Bufo* in East Asia. (A) Past population dynamics of *Bufo sachalinensis* show an increase in effective population size (*N_e*). **(B)** Variability in past sea levels of the Yellow Sea since the Last Glacial Maxima (LGM). The maps represent present day distribution of Eastern Asia *Bufo* and the LGM condition in East Asia with the projection of continental shelves during ice ages (refer to Supplementary file 1O). The

variation of the Yellow Sea level during LGM to present day were extracted from literature (Li et al. 2016). All maps were produced using QGIS v10.1 (ESRI, CA, USA).

Figure 7 Niche suitability models for two segregated clades of East Asian *Bufo*. (A) The niche suitability of the southwestern clade of *B. gargarizans*. (B) The divergence in niche suitability of the septentrional East Asian clade of *B. sachalinensis*. The colours represent the climate suitability of the populations with the red area indicating the highest suitability, the grey area represents unsuitable habitats and black lines represent the maximum sensitivity plus specificity threshold. The white dots represent the occurrence based on the phylogenetic relationship data, and black dots represent the occurrence recorded from GBIF.

Figure 7-figure supplement 1. Response curves of Maxent Jackknife method. The response on bioclimatic variables on the two *Bufo* populations, the southwestern (*B. gargarizans*) and the septentrional (*B. sachalinensis*) East Asia.

Figure 8 Migration trajectory estimated using MIGRATE-N among three northern latitudinal populations of *Bufo* characterised by mtDNA and nuDNA data. (A) Predicted migration pattern of *Bufo* in the northeast Asia, the Korean Peninsula and the Amur River Basin. We detected an asymmetric flow from and towards the Amur River Basin, indicates a weak gene flow from *B. s. sachalinensis* and reflecting a potential local adaptation to the climatic condition of the area. (B) The symmetrical pattern of migration between the north and south populations of *B. stejnegeri* in the Republic of Korea. Bold arrows indicate higher rate and thin-dashed arrows indicate lower rate in one asymmetric migration between two regions.

Figure 9 Species delimitation test using coalescent BFD approach inferred from nuDNA data (*RAG-1- POMC-Rho*). (A) Geographic range for Eastern Asia *Bufo* based on best-supported species delimitation model. The range of each valid species is coloured following the species colour codes in the map. The marking for sympatric and intergradation range (the overlapped range between two subspecies in a broad species complex range) are indicated in the legend. (B) Topology of species delimitation Model A, followed by Model B as the best-supported scenarios among the eight alternative models tested. The ranking is supported through the highest Marginal likelihood estimate (MLE) and positive Bayes factor values. The comparison of all alternative models explained in Supplementary file 1I.

Table and Tables legends

Table 1. Timeframe estimate from the coalescent species tree for Holarctic bufonids

following the Gondwanan origin hypothesis. The key nodes represent the speciation events based on six internal calibrations of the minimum age range obtained from 102 fossils records under a log normal distribution as priors. Dating analyses of the bufonids species tree shows the comparable datation estimated under a relaxed clock with the fossilised birth-death and Yule speciation models. The label for each clade is matched with species tree in Figure 2.

Table 2. Comparison of the cladistic characteristic between the mitochondrial and nuclear

trees of East Asian *Bufo*. The comparison between concatenated mtDNA and nuDNA trees demonstrated the discordances between the recovered clades of *Bufo gargarizans* distributed in East Asia and showed an identical patterns of divergence in the septentrional East Asian *Bufo sachalinensis* clades. The number of clade in the description is matching the phylogenetic trees in Figure 4.

Table 3. Molecular dating analyses for the East Asian *Bufo*. This analysis is based on linked

mtDNA genes (*CR-ND2*) and unlinked multiple nuclear protein coding genes (*POMC-RAG-1-Rho*) under a relaxed molecular clock with different tree priors. The node numbers are matching the clades in the dated phylogenetic tree and hypothesised dispersal pathways in Figure 5.

Table 1. Timeframe estimate from the coalescent species tree for Holarctic bufonids following the Gondwanan origin hypothesis. The key nodes represent the speciation events based on six internal calibrations of the minimum age range obtained from 102 fossils records under a log normal distribution as priors. Dating analyses of the bufonids species tree shows the comparable datation estimated under a relaxed clock with the fossilised birth-death and Yule speciation models. The label for each clade is matched with species tree in Figure 2.

Clade	Key nodes	Dating analysis methods	
		Fossilised birth-death (Median [HPD 95%] /Mya)	Yule (Median [HPD 95%] /Mya)
i	North Africa-Eurasia origin	26.08 [22.54 – 32.51]	25.88 [22.55 – 31.81]
ii	Neotropical origin	24.25 [13.21 – 39.07]	23.40 [13.10 – 35.61]
a (1)	Emergence of <i>Bufotes</i> (West Himalaya)	21.72 [18.90 – 25.25]	21.67 [19.63 – 27.68]
a (2)	Emergence of <i>Bufotes</i> (North Africa-Eurasia)	20.42 [18.10 – 23.28]	20.37 [18.79 – 25.02]
c	Emergence of Palearctic <i>Bufo</i>	14.49 [9.76 – 22.70]	14.52 [9.83 – 22.58]
d	Radiation of western Palearctic <i>Bufo</i>	11.03 [9.14 – 15.42]	11.04 [9.12 – 15.37]
e	Emergence and early radiation of eastern Palearctic <i>Bufo</i>	9.99 [4.66 – 16.57]	10.11 [4.58 – 6.34]
b	Emergence of <i>Epidalea</i> and <i>Strauchbufo</i> (Eurasia)	10.88 [7.11 – 17.78]	10.87 [7.03 – 17.51]
h	Emergence of <i>Rhinella</i> (South America)	14.16 [11.16 – 22.24]	14.00 [11.15 – 21.54]
g	Emergence of <i>Incilius</i> (Central America)	11.90 [7.15 – 22.40]	11.80 [7.16 – 21.52]
f	Emergence of <i>Anaxyrus</i> (Nearctic)	9.56 [6.16 – 18.14]	9.53 [6.13 – 17.58]

Table 2. Comparison of the cladistic characteristic between the mitochondrial and nuclear trees of East Asian *Bufo*. The comparison between concatenated mtDNA and nuDNA trees demonstrated the discordances between the recovered clades of *Bufo gargarizans* distributed in East Asia and showed an identical patterns of divergence in the septentrional East Asian *Bufo sachalinensis* clades. The number of clade in the description is matching the phylogenetic trees in Figure 4.

Distribution range	Clade	Description of cladistics	
		Concatenated mtDNA (<i>CR-ND2</i>)	Concatenated nuDNA (<i>POMC-RAG1-Rho</i>)
Japanese Archipelago	<i>Bufo japonicus formosus</i>	Monophyletic (Clade 1)	Grouped together in a clade of Japanese <i>Bufo</i>
Japanese Archipelago	<i>Bufo japonicus japonicus</i>	Monophyletic (Clade 2)	
Korean Peninsula	<i>Bufo stejnegeri</i>	Monophyletic (Clade 3)	<i>B. stejnegeri</i> were grouped with Korean <i>B. gargarizans</i> in a nested clade of southeastern mainland and septentrional East Asian <i>B. gargarizans</i> (Clade 3)
northeastern Mainland Asia	<i>Bufo gargarizans</i>	Monophyletic (clade	monophyletic (Clade 3) Contained of multiple clades of <i>B. gargarizans</i> restricted to the southeastern and northeastern Asian mainland
eastern Mainland Taiwan Island	<i>Bufo gargarizans</i> <i>Bufo bankorensis</i>	Polyphyletic with a clade of <i>B. bankorensis</i> distributed in Taiwan Island (Clade 5)	Formed multiple clades across Eastern Asian <i>Bufo</i> lineages. A clade grouped with northeastern <i>B. gargarizans</i> . Another clade is polyphyletic with of <i>B. bankorensis</i> of Taiwan Island and <i>B. gargarizans</i> distributed in septentrional East Asia (Clade 3).
central Mainland	<i>Bufo gargarizans</i>		Formed two distinctive clades: (i) A clade that restricted to western Mainland (Clade 2). (ii) Belonged to monophyletic groups of southeastern, western and northeastern Mainland <i>B. gargarizans</i>
septentrional East Asia (Korean Peninsula)	<i>Bufo sachalinensis</i> cf. <i>sachalinensis</i>	Monophyletic (Clade 6)	Monophyletic (Clade 3)
septentrional East Asia (Amur River Basin)	<i>Bufo sachalinensis sachalinensis</i>	Monophyletic (Clade 6)	Monophyletic (Clade 3)

1510 **Table 3. Molecular dating analyses for the East Asian *Bufo*.** This analysis is based on linked mtDNA genes (*CR-ND2*) and
1511 unlinked multiple nuclear protein coding genes (*POMC-RAG-1-Rho*) under a relaxed molecular clock with different tree priors. The
1512 node numbers are matching the clades in the dated phylogenetic tree and hypothesised dispersal pathways in Figure 5.

Key events (node number)	mtDNA (<i>CR-ND2</i>)		nuDNA (<i>POMC-RAG-1-Rho</i>)	
	Relaxed molecular clock		Strict molecular clock	
	Yule prior (Median [HPD 95%] /Mya)	Birth-death prior (Median [HPD 95%]/Mya)	Mean (Median [HPD 95%]/Mya)	Birth-death prior (Median [HPD 95%]/Mya)
Root age of East Asian <i>Bufo</i>	10.47 [7.88 – 13.40]	17.93 [11.81 – 26.63]	14.20 [9.46 – 20.02]	10.12 [6.60 – 12.77]
Emergence of Japanese <i>Bufo</i> (A1)	7.75 [5.98 – 9.52]	8.70 [6.97 – 10.44]	8.23 [6.48 – 9.98]	7.64 [4.99 – 8.92]
Crown clade of <i>B. j. formosus</i>	3.71 [1.71 – 5.92]	4.15 [2.38 – 6.13]	3.93 [2.05 – 6.03]	-
Crown clade of <i>B. j. japonicus</i>	2.24 [0.72 – 4.33]	1.58 [0.57 – 2.84]	1.91 [0.65 – 3.59]	-
Crown clade of <i>B. stejnegeri</i>	3.32 [1.54 – 5.24]	2.22 [0.72 – 4.12]	2.77 [1.13 – 4.68]	4.09 [1.46 – 5.43]
MRCA of East Asian mainland <i>Bufo</i> (A2)	8.38 [6.13 – 10.88]	14.25 [9.67 – 19.44]	11.32 [7.90 – 15.16]	-
Crown clade of <i>B. tibetanus</i> - <i>B. andrewsi</i>	5.21 [3.25 – 7.45]	5.26 [3.01 – 7.66]	5.24 [3.13 – 7.55]	-
Stem node of <i>B. gargarizans</i> complex (A3)	6.85 [4.70 – 9.32]	11.25 [7.28 – 16.08]	9.04 [5.99 – 12.70]	-
Stem clade of Chinese mainland of <i>B. gargarizans</i>	5.19 [3.16 – 6.72]	5.34 [3.45 – 7.83]	5.27 [3.31 – 7.28]	-
Segregation between <i>B. gargarizans</i> inhabited high (<i>B. minshanicus</i>) and low elevation (<i>B. g. gargarizans</i>)	4.70 [3.23 – 6.53]	2.68 [1.82 – 3.26]	3.69 [2.53 – 4.90]	3.77 [2.39 – 3.83]

Crown clade of <i>B. andrewsi</i> and <i>B. gargarizans</i> inhabited high elevated range (<i>B. minshanicus</i>)	2.89 [2.18 – 3.54]	2.55 [1.86 – 3.28]	2.72 [2.02 – 3.41]	2.79 [1.18 – 3.10]
Crown clade of <i>B. g. gargarizans</i> in the southeastern Mainland	3.14 [1.92 – 4.49]	2.56 [1.48 – 3.86]	2.85 [1.70 – 4.18]	3.05 [1.04 – 3.45]
Nested clades of <i>B. gargarizans popei</i> and <i>B. g. gargarizans</i> in the Central, southeastern and northeastern Mainland	2.21 [1.37 – 3.20]	1.28 [0.71 – 1.95]	1.85 [1.04 – 2.58]	3.65 [1.18 – 5.28]
Crown clade of <i>B. bankorensis</i>	1.34 [0.74 – 1.94]	1.28 [0.72 – 1.84]	1.31 [0.73 – 1.89]	1.80 [0.50 – 1.80]
Stem clade of septentrional East Asian <i>B. sachalinensis</i>	1.95 [1.55 – 2.34]	1.81 [1.38 – 2.28]	1.88 [1.47 – 2.31]	2.21 [0.93 – 2.16]
Crown clade of Korean <i>B. sachalinensis</i> cf. <i>sachalinensis</i>	1.58 [1.18 – 1.93]	1.22 [0.82 – 1.61]	1.40 [1.00 – 1.77]	-
Crown clade of Russian <i>B. sachalinensis sachalinensis</i>	1.06 [0.62 – 1.53]	0.46 [0.19 – 0.79]	0.76 [0.41 – 1.16]	-

Additional files

Supplementary file 1

Supplementary file 1A: Type localities of the species members of East Asian *Bufo* discussed in this study. The information retrieves from electronic catalogues of Amphibian species of the world (Frost, 2021). See the reference list at the end of the supplementary file for the resources mentioned in this table.

Supplementary file 1B: Cross-references for the oldest age of fossil species used to calibrate divergence times under the fossilised birth-death model. The collection of the fossil abundance and distribution retrieved from the database of vertebrates (FosFARbase).

Supplementary file 1C: Rank of support for topology models to resolve speciation patterns of the Eastern Palearctic *Bufo* species. We conducted nested sampling analyses on species tree models to evaluate the most likely species tree topology. The Miocene geological events (Model C) was the highest supported factor in driving the speciation of Asian *Bufo* based on the lowest MLE and positive value of Bayes factor obtained.

Supplementary file 1D: Population genetic analyses and test of population expansion derived on 13 subpopulations of *Bufo* distributed in East Asia. The statistical analyses inferred from 894 bp of concatenated *CR-ND2* of 214 individuals, consisted of gene diversity, mean of pairwise diversity (π), nucleotide diversity, Test of goodness of fit (Harpending's

raggedness), Neutrality Test; Tajima's D, Li and Fu statistic, and mismatch distribution.

Haplotype data had 214 segregating sites, among which 178 were informative.

Supplementary file 1E: AMOVA statistic on six clades obtained through mitochondrial phylogenetic tree. The results (%) show a differentiation for different populations, different individual and within individual.

Supplementary file 1F: CLUMPP analysis for best cluster ($K = 2$) for STRUCTURE analysis.

Supplementary file 1G: Dating of the origin of the *Bufo* genus in East Asia. Here we determined the divergence time for all relevant nodes related to ancestral dispersion routes and the most probable ancestral regions.

Supplementary file 1H: Migration patterns among the septentrional populations of East Asian mainland *Bufo*. We inferred the migration patterns for three populations of East Asian *Bufo* (northeastern Mainland, Amur River Basin and Korean Peninsula) and two populations of *B. stejnegeri* on the Korean Peninsula. Maximum likelihood estimates the average of effective population size (Θ) and 25-97.5% confidence intervals (in parentheses), with the computed gene flow rates between the regions characterised on the mtDNA data ($2Nm$) and unlinked nuDNA multi-loci ($4Nm$).

Supplementary file 1I: Species delimitation models and their respective scenarios. We tested each model using the coalescent approach and ranked them based on the probability of a single

independent evolutionary lineage determined by the highest to lowest value of the Marginal likelihood estimation (MLE) and positive to negative values of the Bayes Factor.

Supplementary file 1J: Sampling locations of East Asian *Bufo* spp. and the Genbank accession numbers for all sequences used in the species tree and phylogenetic analyses. See the reference list at the end of the supplementary file for the resources mentioned in this table.

Supplementary file 1K: Information of primers used to amplify targeted gene fragments. Some specific primers designed in this study using Primer3 v.0.4.0 (Untergasser et al. 2012).

Supplementary file 1L: Most probable unlinked site models for species tree reconstruction on each multi-locus. The model was selected with jModelTest 2 v.2.1.10 (Darriba et al., 2015).

Supplementary file 1M: Best evolutionary models for sequences evolution model of each mtDNA and nuclear data. The partition recovered with Partition Finder v.2.1.1 (Lanfear et al. 2017).

Supplementary file 1N: Calculated likelihood values for Homogeneity test analysed on ML tree built from mtDNA and nuclear data.

1578 **Supplementary file 1O:** The datasets used in the reconstruction of paleogeographic maps in this
1579 study.

1580

1581 Supplementary file 2

1582 The references and notes for all citations in Supplementary file 1A and Supplementary file 1J.

1583

Appendix 1

This supplementary text includes the supportive information of materials and methods and results used to build our conclusion, such as the extended population genetic analyses, and the details of calibration points used for molecular dating estimation. We include the assessment of model selections for the species tree topology and species delimitation modelling. We also include the detailed information of all datasets used for the reconstruction of the paleogeographic maps for the molecular dating analyses.

Materials and methods:

Molecular analyses

Genomic DNA was extracted from tissues and swab samples with the DNeasy Blood and Tissue Kit following the instruction of the manufacturer (Qiagen Group, Hilden, Germany). We amplified DNA through PCRs with a total volume of 20 µl per tube, containing 35 to 50 ng/µl of template DNA. The final concentration of the other PCR reagents was such as: 0.125 µM for each forward and reverse primer, 1x Ex taq Buffer (Takara; Shiga, Japan), 0.2 mM of dNTPs Mix (Takara; Shiga, Japan), 1.875 mM of Magnesium Chloride (MgCl₂), 0.1 unit/µl of Ex taq (HR001A, Takara; Shiga, Japan), and double distilled water added to make up the final volume. All fragments targeted were amplified with the following PCR thermal profiles: (1) 5 min at 95 °C for pre-denaturation; a repeated 35 cycles of (2) 1 min at 95 °C for denaturation; (3) 30 sec at specific temperature for annealing; (4) 1 min at 72 °C for elongation. The amplification was ended by a final elongation at 72 °C for 5 min. PCR reactions were performed in a SimpliAmp™ Thermal Cycler (Applied Biosystems, USA). Products were visualized on 1.5% agarose gel

loaded with three microliters of PCR products, run on an Agaro-Power™ System (A-7020, Bioneer; Republic of Korea) and visualized with a Nucleic acid Bioimaging Instrument Blue Illuminator (NeoScience; Republic of Korea) using TopGreen Nucleic Acid 6x Loading Dye (GenomicBase, Republic of Korea). Samples were sent for purification and sequencing for both forward and reverse directions by Cosmogenetech (Cosmogenetech Co., Ltd., Seoul, Republic of Korea). These sequence data have been submitted to the GenBank databases (Supplementary file 1J).

Population genetic analyses

We ran the isolation by distance (IBD) analysis on the mtDNA dataset to test for the correlation between genetic and geographical distances using a Mantel test (Mantel, 1967). The statistical significance of the correlation between matrices was assessed with a Mantel randomization test (10,000 permutations). We then determined the genetic differentiation among populations by calculating the F-statistics, and we determined the proportion of genetic variability found among populations (F_{ST}), among populations within groups (F_{SC}) and among groups (F_{CT}). We estimated the fixation index following the random allelic permutation procedures in ARLEQUIN v.3.1 (Excoffier and Lischer, 2010), where we assessed the statistical significance with 10,000 permutations tests. To estimate the genetic population structure, we then grouped the samples according to the six segregated clades inferred from our phylogenetic tree from the mtDNA dataset, and performed an AMOVA as implemented in the Arlequin package.

Population clustering

We converted the 44 sequences of concatenated *POMC-RAG-1-Rho* (1,030 bp) to SNP format using PSGspider v.2.1.1.5 (Lischer and Excoffier, 2012). We then assigned individuals to one of the eight following populations according to the previously determined clades: 1) western

Mainland, 2) central Mainland, 3) southeastern Mainland, 4) northeastern Mainland, 5) Korean Peninsula, 6) Amur River Basin (along with the southern Primorsky Kray and the Sakhalin Island), 7) Taiwan Island, and 8) the Japanese Archipelago. Then, we used an admixture model to determine the most likely clusters in our SNP dataset using sampling location as prior (LocPrior). We computed the statistic under 100,000 MCMC chains with 1,000 burn-in for 15 iterations using STRUCTURE v.2.3.4 (Pritchard et al., 2000). Because of the uncertainties in the real number of subpopulations, we set the range of cluster (K) from K = 1 to K = 10 using the formula of K = 1 to K = n+2 (n = number of populations) (Evanno et al., 2005; Janes et al., 2017). In order to select the best possible number of clusters in the population and interpret the raw results obtained, we further characterised the result obtained using STRUCTURE Harvester (Earl and VonHoldt, 2012) and the Clump algorithm implemented in the CLUMPAK server (Kopelman et al., 2015). We visualised the results using Structure Plot 2.0 (Ramasamy et al., 2014).

Evaluation of introgression pattern

To test the frequencies of discordancy in our multilocus *POMC-RAG-1-Rho* dataset (1,030 bp), we tested for incomplete Lineage Sorting (ILS) or introgression following population admixture of East Asian *Bufo* clades. To do so, we examined the SNP genealogies following the modified gene flow analysis of D- statistic (Patterson et al., 2012). We calculated Patterson's D-statistic under the assumption that the rates of substitution are equal between the unlinked loci. We computed the D-statistics using the 'CalcD' function in the Evobir package implemented in cran R (R Core Team, 2020) with 1,000 replicates per run. In addition, we applied a simple χ^2 variant of the ABBA-BABA test to compare the frequencies of discordant SNP genealogies and trace the pattern of incomplete lineage sorting, based on the rule of equal ratio (50:50) of ABBA-

BABA test (Martin et al., 2015). This test assumes a strict pattern of evolutionary history in a tree where the substitution rate is always equal and the loci are unlinked (represented by a D-statistics equivalent to 0, where the equal 50:50 ratio of ABBA-BABA test is achieved). In opposition, a D-statistic value significantly different from 0 suggests the presence of introgression in the evolutionary tree, where the event distorted the equal substitution rate along the evolutionary history timeframe.

Species tree: Fossil calibrations process

For the species tree reconstruction we compared the topology of our species tree based on the global phylogenetic relationship of Bufonidae proposed by (Portik and Papenfuss, 2015). We calibrated the maximum and minimum age constrained based on a total fossils evidence. Each calibration point set up was described as:

(a) A minimum age of 11.0 Mya for the emergence of the *Rhinella* genus in South America based on the oldest fossil of *R. marina* found in Colombia (Estes and Wassersug, 1963). We enforced this calibration point under a lognormal distribution with mean in real space = 10.0, log standard deviation (SD) = 1.1, and offset = 11.0. This calibration resulted in a credibility interval of 5% to 95% that ranged between 11.9 Mya and 44.3 Mya.

(b) A minimum age of 18.0 Mya for the origin of toads belonging to the *Bufotes* genus based on the oldest fossil of *B. viridis* sensu lato found in France (Bailon and Hossini, 1990). We enforced this calibration point under a lognormal distribution with mean in real space = 8.0, log (SD) = 0.6, and offset = 16.0. This calibration resulted in a credibility interval of 5% to 95% that ranged between 18.5 Mya and 33.9 Mya.

(c) A minimum age of 7.0 Mya for the origin of *Epidalea* and *Strauchbufo* based on the oldest fossil of *E. calamita* found in Spain (Sanchiz, 1998, 1997) and *S. raddei* found in Russia (Syromyatnikova, 2015). We enforced this calibration point under a lognormal distribution with mean in real space = 6.0, log (SD) = 0.8, and offset = 6.0. This calibration resulted in a credibility interval of 5% to 95% that ranged between 7.17 Mya and 22.2 Mya.

(d) A minimum age of 9.2 Mya for the origin of the Western Palearctic *Bufo* group, based on the oldest fossil of *B. bufo* found in the Czech Republic (Rage and Roček, 2003). We enforced this calibration point under a lognormal distribution with mean in real space = 6.0, log (SD) = 0.8, and offset = 8.0. This calibration resulted in a credibility interval of 5% to 95% that ranged between 9.17 Mya and 24.2 Mya.

(e) A minimum age of 1.0 Mya for the origin of the Eastern Palearctic *Bufo* group, based on the oldest recorded fossil of *B. gargarizans* found in Beijing, China (Ratnikov, 2001). We enforced this calibration point under a lognormal distribution with mean in real space = 3.0, log (SD) = 0.8, and offset = 0.5. This calibration resulted in a credibility interval of 5% to 95% that ranged between 1.08 Mya and 8.62 Mya.

We assembled the species trees from all analyses with LogCombiner v.2.6 (Bouckaert et al., 2019). Thereafter the species tree reconstruction, we also assessed and visualised the assembled species trees with DensiTree v.2.0 (Bouckaert, 2010).

Calibrating the time tree of East Asian *Bufo*

For the divergence dating, we first evaluated the most appropriate molecular clock to be used for the datasets. To do so, we tested a strict molecular clock by calculating the likelihood ratio (LR) following the equation: $LR = 2 (-Ln_A - (-Ln_0))$, adapted from (McGee, 2002) using PAUP

v.4.0. The significant Chi-test for the mtDNA dataset (df: 130; $p < 0.0001$; Supplementary file 1N) supports a variation in the rates of substitution in our trees, and indicates the impossibility to reject the global clock hypothesis. As a result, we used of a relaxed molecular clock model for the divergence dates estimates with the mtDNA dataset, in agreement with recent studies on amphibians (Liedtke et al., 2016). In opposition, the homogeneity test accepted the hypothesis stating that our nuDNA data evolved according to a strict molecular clock, supported by a non-significant Chi-test (df: 42; p -value = 0.019; Supplementary file 1N). Hence, we employed a strict clock model for the nuDNA dataset. Here, we calibrated the time tree for the east Asian bufo lineage based on the primary calibration (fossils evidence) and past ecological events, then followed by the secondary calibrations obtained from *CR* mitochondrial estimation in the literatures. The seven internal calibration points are:

(a) A minimum age of 6.0 Mya (Late Miocene; Kamata and Kodama, 1994) to a maximum age of 10.0 Mya (Middle Miocene; Taira, 2001) for the isolation of the Japanese *Bufo*: *B. j. formosus* and *B. j. japonicus*. The time range is based on the age of the drift of the Japanese plates away from Eurasia. We enforced this calibration point under a normal distribution with mean = 1.0, sigma (σ) = 1.0 and offset = 7.0, resulting in a credibility interval of 5% to 95% that ranged between 6.36 Mya and 9.64 Mya.

(b) A minimum age of 5.0 Mya to a maximum age of 10.0 Mya for the emergence TMRCA of Asian mainland *Bufo* (*B. tibetanus*) in western China (Macey et al., 1998), matching with an active uplifting phase of the QTP during the Miocene (Zhang et al., 2016). We enforced this calibration under a normal distribution with mean = 1.0, $\sigma = 1.5$ and offset = 7.0, resulting in a credibility interval of 5% to 95% that ranged between 6.53 Mya and 10.5 Mya.

- (c) A minimum age of 2.7 Mya to a maximum age of 6.2 Mya for the estimated independent origin of *B. stejnegeri* in Northeast Asia (Late Miocene - Early Pliocene: c. 4.3 Mya; Fong et al., 2020). We enforced this secondary calibration under a normal distribution with mean = 1.0, $\sigma = 1.0$ and offset = 3.5, resulting in a credibility interval of 5% to 95% that range between 2.86 Mya and 6.14 Mya.
- (d) A minimum age of 2.5 Mya to a maximum age of 4.0 Mya for the split between *B. andrewsi* and *B. gargarizans* present on highlands and *B. gargarizans* present on lowland of the southwestern Asian mainland (Fu et al., 2005; Zhan and Fu, 2011). This calibration point matches with the Plio-Pleistocene uplift of the QTP (Lei et al., 2014). We enforced this secondary calibration under a normal distribution with mean = 0.3, $\sigma = 0.4$ and offset = 2.5, resulting in a credibility interval of 5% to 95% that ranged between 2.14 Mya and 3.46 Mya.
- (e) A minimum age of 0.65 Mya to a maximum age of 1.35 Mya for the emergence of *B. bankorensis* on the Taiwanese Island, originating from a single colonisation even by *B. gargarizans* from the Asian mainland (Yu et al., 2014). We linked this calibration point with the oldest related fossil: *B. gargarizans miyakonis* in Miyako Island from the Ryukyu archipelago (Late Pleistocene: < 1.0 Mya) (Nokariya and Hasegawa, 1985). We enforced this calibration under a lognormal distribution with mean = 1.0, $\log(\text{SD}) = 0.3$ and offset = 0.05, resulting in a credibility interval of 5% to 95% that ranged between 0.63 Mya and 1.62 Mya.
- (f) A minimum age of 0.7 Mya for the only recorded fossil assigned to “*Bufo*” (presumed *B. gargarizans*) found in Beijing (Ratnikov, 2001). We enforced this calibration on the *B. gargarizans* monophyletic clade occurring on the southeastern and northeastern Asian

mainland under a lognormal distribution with mean = 0.5, log (SD) = 1.0 and offset = 0.5, resulting in a credibility interval of 5% to 95% that ranged between 0.56 Mya and 2.07 Mya.

(g) A minimum age of 0.8 Mya to a maximum age of 2.0 Mya for the Pleistocene refugium and isolation of *B. gargarizans* on the Korean Peninsula, originating from the southwestern mainland and followed by the colonisation to northern latitudes (Pleistocene; 2.0 Mya – 0.8 Mya) (Borzée et al., 2017). We enforced this secondary calibration on the southwestern and septentrional East Asian clades of *B. gargarizans* under a normal distribution with mean = 0.4, $\sigma = 0.3$ and offset = 1.0, resulting in a credibility interval of 5% to 95% that ranged between 0.91 Mya and 1.89 Mya.

Results:

Molecular Phylogeny of East Asian *Bufo*

Among the six clades recovered by the concatenated *CR-ND2*, we found two distinctive clades within the Japanese *Bufo* corresponding to the subspecies *B. j. formosus* (ML: 97.8/100, PP: 0.98; clade 1 on Fig. 4a) and the subspecies *B. j. japonicus* (ML: 99.5/100, PP: 1.0; clade 2 on Fig. 4B). Clade 3 represented *B. stejnegeri* on the Korean Peninsula (ML: 80.1/100, PP: 1.0; Fig. 4B). Clade 4 (ML: 97.6/100, PP: 0.5; Fig. 4B) consisted of *B. tibetanus* (ML: 99.6/100, PP: 0.98; Fig. 4B) and its sister clade, *B. andrewsi* (ML: 100/100, PP: 1.0; Fig. 4B), both distributed in Western Mainland Asia. Clade 5 included the populations of *B. gargarizans* restricted to central of East Asia (ML: 90.8/100, PP: 0.5; Fig. 4B), containing sister clades of *B. gargarizans* in the central Mainland (ML: 93.1/100, PP: 0.7; Fig. 4B), *B. bankorensis* in Taiwan Island, *B. gargarizans* in southeastern Mainland Asia (ML: 87.6/100, PP: 0.92; Fig. 4B) and another *B. gargarizans* in northeastern Mainland Asia (ML: 89.9/100, PP: 0.93; Fig. 4B). Clade 6

1762 recovered *B. gargarizans* that was geographically restricted to septentrional East Asia (Clade 6:
 1763 ML: 96.8/100, PP: 0.99; Fig. 4B): contained a subclade of *Bufo* distributed in Amur River Basin
 1764 (ML: 93.9/100, PP: 1.0; Fig. 4B) and a subclade distributed in Korean Peninsula (ML: 59.2/100,
 1765 PP: 1.0; Fig. 4B).

1766 Whereas, the six clades recovered from nuDNA tree clustered *B. gargarizans* restricted to
 1767 the southeastern and northeastern Asian Mainland (ML: 78.9/100; PP: 0.73; Clade 1; Fig. 4D).
 1768 The second clade contained of *B. gargarizans* distributed in Central Mainland Asia (ML:
 1769 79.1/100; PP: 0.94; Clade 2; Fig. 4D). The third clade nested the two Korean *Bufo* clades,
 1770 supporting the placement of *B. stejnegeri* and the Korean *B. gargarizans* in a clade (ML:
 1771 75.5/100; PP: 0.77; Clade 3; Fig. 4D). The same clade also assigned *B. gargarizans* distributed in
 1772 southeastern Mainland Asia, *B. bankorensis* distributed in Taiwan Island (ML: 93.4/100; PP: 1.0;
 1773 Clade 3; Fig. 4D) and a subspecies of *B. gargarizans* distributed in the Amur River Basin (ML:
 1774 79.6/100; PP: 0.70; Clade 3; Fig. 4D). We also recovered a clade consisting of the Japanese *Bufo*
 1775 clade (ML: 75.5/100; PP: 0.69; Clade 4; Fig. 4D), and a monophyletic clade of *B. gargarizans*
 1776 geographically restricted to the northeastern Mainland (ML: 86/100; PP: 0.70; Clade 5; Fig. 4D).
 1777 Finally, Clade 6 recovered another *B. gargarizans* of the mainland group restrictly distributed
 1778 across western, central, southeastern and northeastern Asia (ML: 84.5/100; PP: 0.75; Fig. 4D).

1779

1780

Key Resources Table				
Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information

Biological sample (<i>Bufo gargarizans</i>)	Buccal DNA swabs and road kills tissues of <i>Bufo</i> individuals across East Asia.	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Biological sample (<i>Bufo andrewsi</i>)	Buccal DNA swabs and road kills tissues of <i>Bufo</i> individuals across East Asia.	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Biological sample (<i>Bufo bankorensis</i>)	Buccal DNA swabs and road kills tissues of <i>Bufo</i> individuals across East Asia.	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Biological sample (<i>Bufo stejnegeri</i>)	Buccal DNA swabs and road kills tissues of <i>Bufo</i> individuals across East Asia	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/

Biological sample (<i>Bufo japonicus formosus</i>)	Buccal DNA swabs and road kills tissues of <i>Bufo</i> individuals across East Asia	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Biological sample (<i>Bufo japonicus japonicus</i>)	Buccal DNA swabs and road kills tissues of <i>Bufo</i> individuals across East Asia	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Biological sample (<i>Bufo sachalinensis sachalinensis</i>)	Buccal DNA swabs and road kills tissues of <i>Bufo</i> individuals across East Asia	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Gene (<i>B. gargarizans</i>)	Multi loci (<i>CR-ND2-POMC-RAG-1-CRCX4-Rho</i>)	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Gene (<i>B. andrewsi</i>)	Multi loci (<i>CR-ND2-POMC-RAG-1-CRCX4-Rho</i>)	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Gene (<i>B. bankorensis</i>)	Multi loci (<i>CR-ND2-POMC-RAG-1-CRCX4-Rho</i>)	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/

Gene (<i>B. stejnegeri</i>)	Multi loci (<i>CR-ND2-POMC-RAG-1-CRCX4-Rho</i>)	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Gene (<i>B. j. formosus</i>)	Multi loci (<i>CR-ND2-POMC-RAG-1-CRCX4-Rho</i>)	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Gene (<i>B. j. japonicus</i>)	Multi loci (<i>CR-ND2-POMC-RAG-1-CRCX4-Rho</i>)	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Gene (<i>B. s. sachalinensis</i>)	Multi loci (<i>CR-ND2-POMC-RAG-1-CRCX4-Rho</i>)	GenBank, this study	Voucher list in Supplementary file 1J	https://www.ncbi.nlm.nih.gov/genbank/
Commercial assay or kit	Qiagen DNeasy Blood and Tissue Kit	QIAGEN Group, Hilden, Germany	Cat. No. / ID: 69504	
Commercial assay or kit	PCR purification CaspaseGlo 3/7	PROMEGA	G8090	
Commercial assay or kit	BigDye™ Terminator v3.1 Cycle Sequencing Kit	Cosmo Genetech Corp. (Seoul, Republic of Korea)	Thermo Fisher Scientific. Gangnam, Republic of Korea	
Sequence- based reagent	Control B-H	Goebler et al. (1999)	PCR primers for control region (<i>CR</i>); Supplementar y file 1K	Forward: GTCCATTGGA GGTTAAGATC TACCA

Sequence-based reagent	CR-BGarF and CR-BGarR	Borzée et al. (2017)	PCR primers for control region (<i>CR</i>); Supplementary file 1K	Forward: TTGGACGATA GCAAGGAACA CTC Reverse: CCTGACTTCT CTGAGGCCGC TTT
Sequence-based reagent	conBG-L and conBG-H	Liu et al. (2000)	PCR primers for control region (<i>CR</i>); Supplementary file 1K	Forward: GCACGATAGC AAGGAACAC Reverse: CCGCTTTAAG GTACGATA
Sequenced-based reagent	ND1-L-int	Fu et al. (2005)	PCR primers for NADH dehydrogenase 1 (<i>ND1</i>) and NADH dehydrogenase 2 (<i>ND2</i>); Supplementary file 1K	Forward: CGAGCATCC TACCCACGAT TTCG
Sequenced-based reagent	ND1-H4980	Macey et al. (1998)	PCR primers for NADH dehydrogenase 1 (<i>ND1</i>) and NADH dehydrogenase 2 (<i>ND2</i>); Supplementary file 1K	Reverse: ATT TTTCGTAGTT GGGTTTGRTT
Sequenced-based reagent	BGND2F and BGND2R	This study	PCR primers for <i>ND2</i> ; Supplementary file 1K	Forward: TCTCATTCCC AATCTCACTT CTACT Reverse: GCC TCACCCTCCG ACAATA

Sequenced-based reagent	POMC DRV F1 and POMC DRV R1	Vieites et al. (2009)	PCR primers for proopiomelanocortin (<i>POMC</i>); Supplementary file 1K	Forward: ATATGTCATG ASCCAYTTYC GCTGGAA Reverse: GGCRTTTYTTG AAWAGAGTC ATTAGWGG
Sequenced-based reagent	snoBGRAG1 F and snoBGRAG1 R	This study	PCR primers for recombination activating gene 1 (<i>RAG-1</i>); Supplementary file 1K	Forward: TGAGAAACGC AGAGAAAGCC C Reverse: GACGGGTGGC ATCACAAAG G
Sequenced-based reagent	BGRho01-F and BGRho01-R	This study	PCR primers for rhodopsin (<i>Rho</i>); Supplementary file 1K	Forward: CGACTACACC CTGAAGCC Reverse: CCAACAGATA AGGAAGAAG ACCAC
Chemical compound, drug	Ethyl alcohol anhydrous 94.5%~99.9%	DAEJUNG	(EP/GR) 500ML /1L; CAS: 64-17-5	
Software, algorithm	Beast v.2.6.1 and STARBEAST	Bouckaert et al., 2019	https://www.beast2.org/	
Software, algorithm	SNP and AFLP Package for Phylogenetic analysis (SNAPP) in BEAST v.2.6.1	Grummer et al., 2014; Bouckaert et al. 2019	https://www.beast2.org/snapp/	

Software, algorithm	QGIS v.10.1.	ESRI, CA, USA		
Software, algorithm	ArcMap v. 10.6	ESRI, Redlands, USA.		
Software, algorithm	MAXENT v.3.1	http://www.cs.princeton.edu/~schapire/maxent/ ; Phillips et al. 2004, 2006		
Software, algorithm	R version 3.5.1 and	R Core Team (2017)		
Software, algorithm	R Studio version 1.1.442	R Studio Team (2020)		
Software, algorithm	MIGRATE v.3.6.11	Beerli et al. 2019	https://peterbeerli.com/migrate-html5/	
Software, algorithm	MrBayes is 3.2.7	Huelsenbeck, and Ronquist. (2001)		

1 Reconstruction of biogeography for Holarctic bufonids

2 Reconstruction of biogeography for Eastern Palearctic *Bufo*

Phylogeography

C

a) Hypothesis of the QTP vicariance then dispersal

b) Hypothesis of long dispersals

c) Hypothesis of ice-age refugia

D

Tracing cryptic lineage & conservation unit

B

Bufonidae

- Anaxyrus*
- Incilius*
- Rhinella*
- Bufotes*
- Epidalea*
- Strauchbufo*
- Bufo*

Hypotheses testing & advanced taxonomy

Genetic conservation

Lower Miocene, c.15 Mya Upper Miocene, c. 10 Mya

Pliocene, c.5.0 Mya

Pleistocene, <2.0 Mya

Fossil distribution of bufonids

A Distribution ranges for examined taxa **B** Ranked species topology models

C Best – supported topology (Model C)

A Species topology model

B Suboptimal models for species tree topologies

Model E –

A single East Asian origin with three segregated groups

Model D –

Two segregated origins: East Asian mainland *Bufo* vs. Japanese *Bufo*

Model A –

Likelihood based on distributional range

Model B –

Likelihood based on life history and adaptive morphology

A mtDNA *CR-ND2*

B Nuclear *POMC-RAG-1-Rho*

Figure 6

Figure 7

Figure 8

A Mapping the range of species boundaries

B Species delimitation models