The effect of phylogeographic history on species boundaries: a comparative framework in *Hyla* tree frogs

Christophe Dufresnes¹,²*, Matthieu Berroneau³, Sylvain Dubey²,⁴,⁵, Spartak N. Litvinchuk⁶,⁷ & Nicolas Perrin²

Because it is indicative of reproductive isolation, the amount of genetic introgression across secondary contact zones is increasingly considered in species delimitation. However, patterns of admixture at range margins can be skewed by the regional dynamics of hybrid zones. In this context, we posit an important role for phylogeographic history: hybrid zones located within glacial refugia (putatively formed during the Late-Pleistocene) should be better defined than those located in post-glacial or introduced ranges (putatively formed during the Holocene and the Anthropocene). We test this hypothesis in a speciation continuum of tree frogs from the Western Palearctic (*Hyla*), featuring ten identified contacts between species spanning Plio-Pleistocene to Miocene divergences. We review the rich phylogeographic literature of this group and examine the overlooked transition between *H. arborea* and *H. molleri* in Western France using a multilocus dataset. Our comparative analysis supports a trend that contacts zones resulting from post-glacial expansions and human translocations feature more extensive introgression than those established within refugial areas. Integrating the biogeographic history of incipient species, i.e. their age since first contact together with their genetic divergence, thus appears timely to draw sound evolutionary and taxonomic inferences from patterns of introgression across hybrid zones.

How and when the continuous process of speciation leads to discrete entities is a central question in evolutionary biology and systematics¹. Reproductive isolation evolves as diverging lineages accumulate genetic changes, ultimately reducing the probability of mating (pre-zygotic barriers) and the fitness of interspecific hybrids (post-zygotic barriers)². Hybridizing taxa thus lay at the boundaries of the biological species concept³ and offer opportunities to understand how much divergence is needed to maintain incipient species apart, despite occasional gene flow⁴.

Since it reflects the balance between dispersal and selection against hybrids, the amount of admixture at phylogeographic transitions serves as an ad hoc measure of reproductive isolation under natural conditions. Although scarce, comparative hybrid zone analyses support a gradual buildup of reproductive isolation with the time diverging lineages have spent in allopatry, where a remarkably short temporal window separates widely admixing and genetically isolated taxa⁵–⁷. Characterizing this “grey zone of speciation” is attractive for species delimitation. Hence, assessing the amount of admixture at phylogeographic transitions is increasingly viewed as a key step in integrative taxonomy, notably for controversial species⁶.

However, patterns of introgression at hybrid zones can also be influenced by an array of extrinsic factors that confound with reproductive isolation, e.g. local barriers to gene flow due to landscape features, demographic events of extinction and recolonization, or changes in ecological conditions shifting species transitions⁶–¹². We posit that the time since contact should also play a major role, in the light of the biogeographic history of species. In the Northern hemisphere, the Quaternary climatic oscillations frequently remodeled lineage distributions...
and set up the contact zones observed nowadays. Species transitions located in glacial refugia may have existed since or even prior to the last glaciation, and may thus be as old as several hundred thousand of years. In contrast, contact zones formed by post-glacial recolonization post-date the Last Glacial Maximum (LGM, ~21,000 years ago), and are no older than a few thousands of years. Post-glacial expansions should lead to wide, mosaic parapatric distributions as species colonized new, empty ranges, across the climatically instable northern regions, giving rise to multiple local contacts that may not have reached their ecological or genetic equilibrium yet (e.g.12). As a result, these transitions would exhibit admixture over large distances, even if the interacting species are partially reproductively isolated. Reciprocally, contact zones that have persisted in glacial refugia had more time to purge hybrid genotypes and should feature more stable transitions. These processes could also affect opportunities for reinforcement, i.e. the evolution of assortative mating as a response to selection against unfit or maladapted hybrids14,25. Species could start establishing pre-mating barriers earlier in refugial than in post-glacial ranges, where they are thus expected to experience less gene flow.

Given these contrasting outcomes, did the biogeographic history of secondary contacts significantly affect their amount of admixture, in turn blurring species delimitation assessments? This question is fundamental for both speciation and systematic biology, but because the answer requires replicate transitions between the same species pairs, or at least between pairs of similar divergence, it has never been formally tested. Several unrelated studies yet support remarkable latitudinal differences in gene flow. The rate of hybridization between the collared and pied flycatchers (Ficedula) is higher in the Baltic Isles than in Central Europe26, where both species have been in sympathy for a longer time and evolved character displacement on male plumage color via reinforcement13.

Time since contact was also invoked to explain differences in hybrid zone structures among Quercus oaks. In house spiders (Tegenaria), two sister species co-exist in England and widely hybridize in northern, but not in southern parts of the country. These scattered observations match our expectations of wider hybrid zones in recently colonized ranges, for the same species pair.

The Western-Palearctic radiation of tree frogs (Hyla) provides a unique framework to compare phylogeographic transitions in respect to their late-Quaternary history. The phylogeny and phylogeography of this group have been extensively characterized over the last decade, for all species individually, but also at range margins. W-Palearctic tree frogs diversified from the Miocene (>10 Ma) to the Plio-Pleistocene (3–4 Ma) and now include ten species forming replicate transitions across refugial, post-glacial, or even introduced ranges. Here we revisit their patterns of admixture in the light of the history of contacts. We first detail the overlooked hybrid zone between H. arborea and H. molleri preliminary located in W-France11–13, using a population genetic approach. This pair is of high comparative interest since it diverged almost simultaneously (5–6 Ma) with several other species pairs (H. arborea/orientalis, H. arborea/perrini, H. savignyi/felixarabica). Second, we assess how admixture across phylogeographic transitions varies depending on species divergence and relative time since first contact, by combining independently published data from ten contact zones.

Methods

Population genetics of the Hyla arborea/molleri hybrid zone. A total of 251 adult tree frogs (n = 231 H. arborea/molleri, as well as 20 syntopic H. meridionalis) were captured from 25 localities (loc. 4–29) throughout Western France (File S1). DNA was collected with non-invasive buccal swabs, stored dry at −20 °C, and extracted with the BioSprint robotic workstation (Qiagen). The experiments conducted in this research (DNA sampling) were performed in accordance with the regional authorities (Prefectures of Ille-et-Vilaine, Charente and Gironde) under collecting permits, and relevant guidelines were followed. No individuals were harmed or killed, and all were released at their place of capture.

Twenty-six microsatellite loci cross-amplifying in H. arborea and H. molleri were genotyped in all samples: Ha-T3, Ha-T52, Ha-M2, WHA5–22, Ha-T45, Ha-T11, Ha-T49, Ha-T51, Ha-T56, Ha-T64, Ha-T41, Ha-T69, Ha-T54, Ha-T55, Ha-T50, Ha-T58, Ha-T53, Ha-T62, Ha-T66, Ha-T60, Ha-T61, Ha-T68, Ha-T63, Ha-T67, WHA1–25, WHA1–67, Ha-A1124,25 (see methods therein). Eight loci did not amplify in H. meridionalis, in line with the older divergence time. In addition, we re-analyzed microsatellite reads from reference populations of H. arborea (from Switzerland, the Netherlands, and France, loc. 1–3, n = 80), H. molleri (from Spain, loc. 30, n = 16) and H. meridionalis (from southern France, loc. 31, n = 16) from our previous work24–26. The new and published genotypes were generated by the same genetic analyzer (an ABI3130) and (re)scored together in GenMapper 4.0 with a single set of bins and panels, to ensure the correspondence of alleles.

The total microsatellite dataset (Western France + reference populations) included 363 individuals. We first performed Bayesian clustering with STRUCTURE27 by running 10 chains of 100’000 iterations, after a burnin of 10’000, for each value of K from 1 to 11. The most likely number of clusters (K) was estimated with STRUCTURE HARVESTER14, concatenated over replicates with CLUMPP29, and graphically displayed with Distruct8. Second, we performed a Principal Component Analysis (PCA) of individual genotypes with the R packages ade4 and adegenet14. Third, we computed pairwise genetic distance (Fst) for populations with n ≥ 5 using the R package hierfstat21.

Mitochondrial DNA from 230 hybrid zone samples was profiled by enzyme restriction with MseI, which differentially cuts the three mitotypes (methods in19). In addition, we included the published mtDNA barcoding data from two additional populations located close to our loc. 15–16 (n = 9 H. arborea/molleri and 3 H. meridionalis, see File S1).

Finally, we fitted sigmoid clines to mitochondrial frequency and nuclear admixture coefficients (STRUCTURE’s Q) along a north-south transect running along the Atlantic coast (loc. 4–12, 15–16, 21–24, 26–29), using the R package hzar21. We performed model selection between two- (width and center) to eight-parameter clines (width, center, parental frequencies, position and steepness of exponential tails), and kept the model with the best AIC score.
**Phylogeography of W-Palearctic Hyla.** In order to provide an up-to-date and comprehensive overview of species distributions, range limits, and the extent of the ten contacts documented so far in W-Palearctic *Hyla*, we combined our new *H. arborea/molleri* population genetic data with those from 24 previously-published studies on the phylogeography and hybrid zones of this group, totaling records from 857 different localities. These included species-specific phylogenetic/phylogeographic accounts, intra-specific genetic work on *H. meridionalis* and *H. cartaginensis*, *H. sarda* and *H. intermedia* and *H. perrini*. We built species distribution models (SDM) to predict the ecological niches and the LGM ranges for all ten taxa. Such analyses have been independently conducted in previous work on some taxa (20, 46–48), but here we aimed to provide comparable outcomes across all of them, obtained from a single methodology. To this end, analyses were performed with MaxEnt 3.4.1 (53), implementing occurrence filtering, tests for distinct candidate sets of environmental variables, and multiple combinations of model parameters (feature classes, regularization multipliers, and sets of variables), using accessible areas for model calibration and multiple statistical criteria (partial ROC, omission rates, and AICc) for model selection.

A total of 6,363 localities were used, comprising our own records and previously published data (File S3), filtered to avoid spatial autocorrelation and duplication by NicheToolBox (48). We selected localities of each species to obtain the maximum number of localities that were at least 10 km (0.093°) apart, which resulted in 3,685 presence-only locations for analyses. Altitude and 19 bioclimatic layers (30 arc seconds and 2.5 arc minutes spatial resolutions) representative of the climatic data over ~1950–2000 were extracted from the WorldClim 1.4 database (http://www.worldclim.org). An additional three layers were considered: the aridity index (Global Aridity and Potential Evapo-Transpiration; http://www.cgiar-csi.org/data/global-aridity-and-pot-database), spatial homogeneity of global habitat (EarthEnv; http://www.earthenv.org|textured.html), and the global percent of tree coverage (https://github.com/globalmaps/gm_ve_v1). To consider topography in the model, four landscape layers were calculated with QGIS: aspect, exposition, slope, and terrain roughness index. For the LGM predictions, we used the same 19 bioclimatic layers (2.5 arc minutes spatial resolution) and applied two general atmospheric circulation models: the Community Climate System Model (CCSM; http://www2.cesm.ucar.edu/) and the Model for Interdisciplinary Research on Climate (MIROC (54)). Analyses were conducted with species-specific masks covering the area of occurrence and adjacent regions.

To eliminate predictor collinearity before generating the model, we calculated Pearson’s correlation coefficients for all pairs of bioclimatic variables using ENMTools (48). For correlated pairs (\(|r| > 0.75\)), we excluded the variable that appeared less biologically important for *Hyla* tree frogs. Analyses were performed separately for each species, and we ran MaxEnt models with 10 replicate bootstrap resamplings. Model calibration consisted in creating models developed with distinct regularization multipliers (0.5 to 6 at intervals of 0.5), feature classes (resulted from all combinations of linear, quadratic, product, threshold, and hinge response types), and from two different sets of layers. The first set included altitude, the aridity index, spatial homogeneity, of global habitat (EarthEnv; http://www.earthenv.org|textured.html), and the global percent of tree coverage (https://github.com/globalmaps/gm_ve_v1). To consider topography in the model, four landscape layers were calculated with QGIS: aspect, exposition, slope, and terrain roughness index. For the LGM predictions, we used the same 19 bioclimatic layers (2.5 arc minutes spatial resolution) and applied two general atmospheric circulation models: the Community Climate System Model (CCSM; http://www2.cesm.ucar.edu/) and the Model for Interdisciplinary Research on Climate (MIROC (54)). Analyses were conducted with species-specific masks covering the area of occurrence and adjacent regions.

To eliminate predictor collinearity before generating the model, we calculated Pearson’s correlation coefficients for all pairs of bioclimatic variables using ENMTools (48). For correlated pairs (\(|r| > 0.75\)), we excluded the variable that appeared less biologically important for *Hyla* tree frogs. Analyses were performed separately for each species, and we ran MaxEnt models with 10 replicate bootstrap resamplings. Model calibration consisted in creating models developed with distinct regularization multipliers (0.5 to 6 at intervals of 0.5), feature classes (resulted from all combinations of linear, quadratic, product, threshold, and hinge response types), and from two different sets of layers. The first set included altitude, the aridity index, spatial homogeneity, of global habitat (EarthEnv; http://www.earthenv.org|textured.html), and the global percent of tree coverage (https://github.com/globalmaps/gm_ve_v1). To consider topography in the model, four landscape layers were calculated with QGIS: aspect, exposition, slope, and terrain roughness index. For the LGM predictions, we used the same 19 bioclimatic layers (2.5 arc minutes spatial resolution) and applied two general atmospheric circulation models: the Community Climate System Model (CCSM; http://www2.cesm.ucar.edu/) and the Model for Interdisciplinary Research on Climate (MIROC (54)). Analyses were conducted with species-specific masks covering the area of occurrence and adjacent regions.

To eliminate predictor collinearity before generating the model, we calculated Pearson’s correlation coefficients for all pairs of bioclimatic variables using ENMTools (48). For correlated pairs (\(|r| > 0.75\)), we excluded the variable that appeared less biologically important for *Hyla* tree frogs. Analyses were performed separately for each species, and we ran MaxEnt models with 10 replicate bootstrap resamplings. Model calibration consisted in creating models developed with distinct regularization multipliers (0.5 to 6 at intervals of 0.5), feature classes (resulted from all combinations of linear, quadratic, product, threshold, and hinge response types), and from two different sets of layers. The first set included altitude, the aridity index, spatial homogeneity, of global habitat (EarthEnv; http://www.earthenv.org|textured.html), and the global percent of tree coverage (https://github.com/globalmaps/gm_ve_v1). To consider topography in the model, four landscape layers were calculated with QGIS: aspect, exposition, slope, and terrain roughness index. For the LGM predictions, we used the same 19 bioclimatic layers (2.5 arc minutes spatial resolution) and applied two general atmospheric circulation models: the Community Climate System Model (CCSM; http://www2.cesm.ucar.edu/) and the Model for Interdisciplinary Research on Climate (MIROC (54)). Analyses were conducted with species-specific masks covering the area of occurrence and adjacent regions.

**Results**

**Hyla arborea/molleri hybrid zone.** Our data supported widespread nuclear admixture between *H. arborea* and *H. molleri*, spanning from the Garonne basin (loc. 21) to the Massif Central (east of loc. 14) and Normandy (loc. 4) (Fig. 1). The mitochondrial picture was broadly concordant, with mitotypes shared in several localities north of the Garonne (loc. 15–20). Clines were wide (width \(w = 98\) km for nuclear loci; \(w = 43\) km for mtDNA), remarkably coincident (center \(c = 407\) km for nuclear loci; \(c = 406\) km for mtDNA), and both bear a long introgression tail on the *H. arborea* side (the selected cline model under the AIC criterion, Fig. 2). There was no trace of mitochondrial or nuclear introgression with the sympatric and syntopic *H. meridionalis* (Fig. 1). The PCA (File S5) and pairwise F\(_{st}\) (File S6) disentangled the three gene pools and confirmed that individuals collected in Western France represent the entire *H. arborea–molleri* spectrum.
Species transitions. The distribution of W-Palearctic *Hyla* taxa is shown in Fig. 3, based on genetic barcoding from 25 studies (including the present paper). A summary of species interactions and the corresponding literature is provided in Table 1 and File S2, from which we categorized the following. (1) One case of sympatry without gene flow despite occasional hybridization: *H. meridionalis/arborea-molleri*; (2) Five steep parapatric transitions, with introgression (if any) restricted at range margins, and narrowcline width (<50 km): *H. savignyi/felixarabica* in N-Israel; *H. arborea/perrini* in NE-Italy; *H. meridionalis/carthagineensis* in NE-Algeria; *H. arborea/orientalis* in the Balkans; *H. savignyi/orientalis* in S-Turkey, where evidence for hybridization/admixture is
presently lacking. (3) Three wide continuous transitions with large nuclear clines (>50 km) and admixture over more than a hundred kilometers: *H. arborea*/*orientalis* in Poland; *H. arborea*/*molleri* in SW-France; *H. intermedia*/*perrini* in Central Italy; the *H. arborea*/*perrini* hybrid swarm from W-Switzerland was also assigned to this category. Considering the ten contacts altogether, the relationship between the amount of admixture and the divergence time was significant (ordinal logistic regression: $\chi^2 = 7.72$, $n = 10$, df = 1, $P = 0.005$).

**LGM distributions and relative ages of the species transitions.** All species distribution models (SDM) received very good fits, indicative of robust predictions (File S7). Projected distributions for each species are shown in Figs. 4 and 5. In Europe, LGM conditions appeared unsuitable across the northern parapatric ranges of *H. arborea* and *H. orientalis*, but were more favorable in the Balkans (Fig. 4). For *H. molleri*, the models identified hospitable glacial ranges along the coasts (Fig. 4). In Italy, the probabilities of occurrence remained high for

---

**Table 1.** Summary of ten secondary contact zones among W-Palearctic tree frogs, in respect to the divergence time between species pairs (Div. in Ma; taken from20; nodes are labelled in Fig. 6), the putative epoch of first contact based on previous studies and past species distribution modelling (see File S2), and the type of transition (see File S2, Fig. 3 and the listed references), with the corresponding category in our meta-analysis (‘sympatry without gene flow’, ‘steep transition’, ‘wide transition/hybrid swarm’); w: hybrid zone width, as measured from cline analyses of transects.

| Node | Region | Species pair | Div. | Contact | Transition | Reference |
|------|--------|--------------|------|---------|------------|-----------|
| A    | W-France | *H. meridionalis*/arborea-*molleri* | 20   | Antiquity | sympatry without gene flow<sup>a</sup> | 22        |
| B    | S-Anatolia | *H. orientalis*/savignyi* | 8.8  | ≥LGM    | steep, hybridization/admixture not detected<sup>b</sup> | 38,39    |
| C    | NE-Italy/Slovenia | *H. arborea*/perrini* | 6.6  | ≥LGM    | steep, faint traces of admixture<sup>b</sup> | 40        |
| C    | W-Switzerland | *H. arborea*/perrini* | 6.6  | 1950s   | hybrid swarm<sup>c</sup> | 50        |
| D    | Balkans | *H. arborea*/orientalis* | 5.1  | ≥LGM    | steep (w = 30–32 km)<sup>b</sup> | 51        |
| D    | Poland | *H. arborea*/orientalis* | 5.1  | post-glacial | wide, admixture across a large area (>200 km)<sup>d</sup> | 51 |
| D    | W-France | *H. arborea*/molleri* | 5.1  | post-glacial | wide (w = 98 km)<sup>d</sup> | this study |
| E    | Central Italy | *H. perrini*/intermedia* | 3.5  | post-glacial | wide (w = 96 km)<sup>d</sup> | 50 |
| F    | Levant | *H. savignyi*/felixarabica* | 6.4  | ≥LGM    | steep, geographically restricted admixture<sup>b</sup> | 38,48    |
| G    | E-Algeria | *H. meridionalis*/carthaginienst* | 4.1  | ≥LGM    | steep, admixture detected at a single locality<sup>b</sup> | 37 |

---

**Figure 3.** Distribution of W-Palearctic tree frogs based on genetically barcoded populations, combining 25 studies. At transitions, populations were considered as a mix (represented in half circles) if they featured both mtDNA haplotypes and/or nuclear genotypes consistent with admixture (STRUCTURE’s Q < 0.9 in multilocus analyses; co-segregation of species-diagnostic alleles). The map was created in QGIS 3.4 (https://qgis.org). Photo: *Hyla meridionalis* from southeastern France, by CD.
H. perrini, but not for *H. intermedia*, which appeared scattered throughout the Peninsula (Fig. 4). In N-Africa, the Maghreb coast offered favorable conditions for both *H. meridionalis* and *H. carthaginiensis* (Fig. 5). Finally, the present and LGM predictions were globally similar for the insular *H. sarda*, and the Middle-Eastern *H. savignyi* and *H. felixarabica* (Fig. 5).

Integrating the projected species distributions with the well-documented phylogeography of W-Palearctic *Hyla* (File S2), the wide transitions/hybrid swarms were all presumably formed after the LGM (Holocene or Anthropocene contact): *H. arborea/mollerii, H. arborea/orientalis* in Poland, *H. intermedia/perrini, H. meridionalis/*
perrini/arborea in W-Switzerland (Fig. 6). In contrast, the steep transitions all encompass putative glacial refugia (Pleistocene contact): \textit{H. savignyi/orientalis}; \textit{H. arborea/orientalis} in the Balkans; \textit{H. arborea/perrini}; \textit{H. savignyi/felixarabica}; \textit{H. meridionalis/carthaginensis} (Fig. 6). Finally, the sympatric (but not admixing), early-diverged pair \textit{H. meridionalis/arborea–molleri} result from a human introduction (File S2). Accordingly, the relative age of contact (Holocene/Anthropocene vs Pleistocene) was a significant predictor of these three categories of transition (wide, steep, or sympatry without gene flow), based on an ordinal logistic regression ($\chi^2 = 3.85, n = 10, df =$...
1, \( P = 0.049 \), and as illustrated in Fig. 7. The effect was also significant when comparing only wide vs steep hybrid zones with a \( \chi^2 \) test of independence (\( \chi^2 = 5.41, n = 9, \text{df} = 1, P = 0.020 \)) – two categories where species pairs do not significantly differ in divergence times (Mann-Whitney-Wilcoxon test, \( n = 9, W = 13.5, P = 0.45 \)).

**Discussion**

**A wide tree frog transition across W-France.** The widespread introgression between *H. arborea* and *H. molleri* documented across W-France confirms previous studies that reported hybrid specimens in several scattered localities21–23. Their genomes are thus mostly compatible, so selection against hybrids is probably weak (see below). The phylogeographic history of these two species suggests a somewhat classic scenario of post-glacial contact. According to our predicted distributions (Fig. 4, see also46), the Iberian *H. molleri* could have expanded from some Atlantic refugia into western France (where private mtDNA haplotypes are found21,46), while *H. arborea* arrived from the more distant Balkan refugium 26. Subsequently, these two species introgressed massively along the Atlantic coast, but the dry Garonne valley may nowadays act as an effective bioclimatic barrier preventing northwards dispersal of *H. molleri* (as in other biota62). This hypothesis is supported by the asymmetric transition (Fig. 2): *H. arborea* allele frequencies decrease continuously with latitude, up to a sharp drop between the northern (loc. 4–20) and southern sides of the valley (loc. 21–29). Hence, the hybrid zone is probably not stable, and *H. arborea* might eventually fully backcross the hybrid populations over time. This particular scheme of hybrid zone movement is somewhat reminiscent of the *Bufotes viridis*/*balearicus* green toad transition in Italy, where hybrids became isolated north of the Po River, and cannot escape inflocks of *B. viridis* expanding southward63. Here, the *H. arborea*/*molleri* species pair adds a valuable point to our comparative framework of admixture vs history of secondary contacts.

**Tree frogs admix freely in young hybrid zones, but not in old ones.** The dense phylogeographic framework, including ten contacts from the fully-resolved radiation of W-Palearctic tree frogs, supports that the width of species transitions drastically differs depending on when these were first established (Figs. 6 and 7, Table 1). Lower-Miocene lineages do not introgress (node A), as expected given the deep genetic (and phenotypic) divergence. However, Upper Miocene (nodes B–D and F) and Pliocene (nodes E–G) taxa admix significantly more across post-glacial and introduced ranges than in refugial areas (Figs. 6 and 7). Note that the divergence times reported here (from20, see Methods) should be taken with caution, since the genus *Hyla* might be younger than previously assumed64.

Because they feature hybrid zones involving the same pair, or between species that coalesced at the same time, two nodes are of particular interest for direct comparisons. First, we highlighted striking differences among the *H. arborea*/orientalis contacts (node D). In their Balkan refugium, mtDNA and nuclear transitions do not exceed 30 km33, while they extend over hundreds of kilometers in the Polish post-glacial ranges51. The latter mirrors the wide post-glacial hybrid zone documented herein for *H. arborea/molleri*, which share the same amount of divergence (node D). Second, the slightly older *H. arborea* and *H. perrini* (node C) do not admix along their natural margins in NE-Italy/Slovenia69, where subrefugia of both species are putatively located30,26 (see File S2). Nevertheless, this pair widely hybridized in W-Switzerland50 following introductions of *H. perrini* in the 1950s65. All other tree frogs meeting in ecologically stable regions featured steep mitochondrial and nuclear transitions (File S2), i.e. *H. savignyi/orientalis* in southern Anatolia (note B38,39), *H. savignyi/feilxarabica* in the Levant (node

![Figure 6](https://qgis.org). Phylogeny and distribution of Western-Palearctic tree frogs, illustrating post-glacial expansions and the extent of introgression in secondary contact zones in respect to their amount of divergence (capital letters: phylogenetic nodes) and age since contact (subscripts: “young” as post-LGM (y) or “old” as pre-LGM (o)). Species distributions are adapted from77. See Table 1 for details on the patterns of admixture and relevant references. The map was created in QGIS 3.4 (https://qgis.org). Photo: *H. arborea* from southeastern France, by CD.
Figure 7. Species transitions in respect to divergence time in W-Palaearctic tree frogs. Color codes refer to Figs. 3 and 6, and replicate contacts between the same species are specifically labelled. Letters indicate the phylogenetic nodes shown in Fig. 6. Symbols differentiate the contacts according to when they were putatively formed.

Figs. 3 and 6, and replicate contacts between the same species are specifically labelled. Letters indicate the phylogenetic nodes shown in Fig. 6. Symbols differentiate the contacts according to when they were putatively formed.

while we assume that local geographic barriers played a minor effect on the patterns of introgression (see above), this remains to be thoroughly addressed by habitat analyses, notably in the context of landscape changes during the Anthropocene. In parallel, the relative contribution of pre- vs post-zygotic isolation in limiting admixture across the narrow hybrid zones is unclear; characterizing pre-mating barriers (breeding calls, pheromones) between supposedly cryptic incipient species remains a major gap in the amphibian literature. Accounting for all the factors potentially affecting the age and admixture patterns at secondary contact zones is already colossal for a single transition, and it becomes logistically impossible in a comparative framework involving multiple contacts.

While acknowledging these limitations, the qualitative data in hand thus tend to support our working hypothesis that the contrasted hybrid zones in *Hyla* partly result from the phylogeographic dynamics that led to their establishment. Compared to refuge areas, colonizers at the front of post-glacial expansions should hybridize more with related species, due to limited mate choice, multiple contacts across unsettled parapatric ranges, and few opportunities to evolve reinforcement given the much shorter time since first contact. While the prevalence of reinforcement in the wild is debated, this mechanism is known to contribute to speciation in hyliid frogs.

For the same species pair, different genomic backgrounds (i.e. intraspecific lineages) may also affect patterns of introgression across the narrow hybrid zones is unclear; characterizing pre-mating barriers (breeding calls, pheromones) between supposedly cryptic incipient species remains a major gap in the amphibian literature. Accounting for all the factors potentially affecting the age and admixture patterns at secondary contact zones is already colossal for a single transition, and it becomes logistically impossible in a comparative framework involving multiple contacts.

While acknowledging these limitations, the qualitative data in hand thus tend to support our working hypothesis that the contrasted hybrid zones in *Hyla* partly result from the phylogeographic dynamics that led to their establishment. Compared to refuge areas, colonizers at the front of post-glacial expansions should hybridize more with related species, due to limited mate choice, multiple contacts across unsettled parapatric ranges, and few opportunities to evolve reinforcement given the much shorter time since first contact. While the prevalence of reinforcement in the wild is debated, this mechanism is known to contribute to speciation in hyliid frogs.
incompatibilities and admixture; here this could apply to the two *H. arborea*/*orientalis* contacts, which involve independent refugial lineages. Therefore, replicate hybrid zone analyses appear necessary to fully appreciate where two taxa lay on the speciation continuum.

In this context, our findings have profound consequences for species delimitation. Phylogeographic lineages are increasingly considered in taxonomy, and when possible, their specific or subspecific rank may be decided based on whether they are (partially) reproductively isolated, following the universal biological species concept. Accordingly, wide transitions are interpreted as a lack of reproductive barriers, as opposed to narrow transitions. An ad hoc way to quantify reproductive isolation from hybrid zones is to calculate selection against hybrids from cline width and dispersal, with the equations of. Without extrinsic selection (as expected for ecologically-similar cryptic species), the coefficient corresponds to the fitness difference between populations from the center and the edge of the hybrid zone (under heterozygote disadvantage), given 

\[ w \approx 2s/\sigma \]

In *Hyla*, taking a dispersal rate \( \sigma \approx 0.5 \text{ km/yr} \) (adapted from, who found average movements of 1.5 km over a three-year capture-mark-recapture survey), and average ages at maturity and longevity of 2 and 6 years, respectively (reviewed by), the refugial contact between *H. arborea* and *H. orientalis* (\( w = 30 \text{ km} \)), reflect ten-fold higher estimates (\( s^* = 0.05 \)) than the post-glacial contact between *H. arborea* and *H. molleri* (\( s^* = 0.005 \), from \( w = 98 \text{ km} \)), despite similar genetic divergence. The extent of admixture between incipient species might thus misguide taxonomic conclusions if the history of contacts is not considered. When possible, the grey zone of speciation would be best inferred from refugial areas, where selection had time to operate and species transitions putatively reached an equilibrium. Reciprocally, narrow post-glacial hybrid zones would strongly indicate incipient speciation. To conclude, we recommend that taxonomists interpret introgression patterns at range margins in the light of two key timings of speciation, i.e. the time of divergence and the time since first contact. Finally, the present results call for caution when assessing the hybridization potential of invasive species. Genetic pollution is a major issue of bioinvasions, and the risk should not be underestimated even if the local and invasive species rarely hybridize in natural ranges. This is well-illustrated here by the introduction of *H. perrini* north of the Alps, which extensively admixed with local *H. arborea*, despite nearly absent introgression at their natural phylogeographic transition. In the syntopic fishes *Alosa pseudoharengus* and *A. aestivalis*, human-driven habitat fragmentation disrupted their well-established reproductive barriers, leading to whole-hybrid populations. If the diverged genomes of incipient species are partly incompatible, such “forced” hybridization can severely affect the sustainability of local taxa, and thus precipitate their extinctions.

**Conclusions**

Combining new data with fifteen years of research on the W-Palaearctic radiation of tree frogs (*Hyla*), we emphasized how admixture patterns at species transitions can drastically vary depending on their biogeographic history. For similar evolutionary divergence, species pairs form wider hybrid zone in post-glacial and introduced ranges, compared to regions spanning their putative glacial refugia. We thus conclude that both the divergence time and the time since contact significantly affect the extent of admixture between incipient species. Future studies should focus on identifying the proximate mechanisms, and especially test whether reinforcement evolved in long-term contacts. Informed taxonomic decisions are thus best taken from replicate hybrid zone analyses that account for the late-Quaternary dynamics of phylogeographic transitions.

**Data availability**

The data of this study is available as supplementary information.

Received: 2 October 2019; Accepted: 11 March 2020;
Published online: 26 March 2020

**References**

1. Avise, J. *Phylogeography: The history and formation of species*. Cambridge, (Harvard University Press) (2000).
2. Coyne, J. A. & Orr, H. A. *Speciation*. (Sinauer Associates) (2004).
3. Mayr, E. *Systematics and the origin of species, from the viewpoint of a zoologist*. (Harvard University Press), (1942).
4. Roux, C. *et al.* Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biol.* **14**, e20000234 (2016).
5. Singhal, S. & Moritz, C. Reproductive isolation between phylogeographic lineages scales with divergence. *Proc. R. Soc. Lond.* **280**, 20132246 (2013).
6. Artsen, J. W., Wielstra, B. & Wallis, G. P. The modality of nine *Triturus* newt hybrid zones assessed with nuclear, mitochondrial and morphological data. *Biol. J. Linn. Soc.* **113**, 604–622 (2014).
7. Dufresnes, C. *et al.* Phylogeography of a cryptic speciation continuum in Eurasian spadefoot toads (*Pelobates*). *Mol. Ecol.* **28**, 3257–3270 (2019).
8. Padial, J., Miralles, A., De la Riva, I. & Vences, M. The integrative future of taxonomy. *Front. Zool.* **7**, 16 (2010).
9. Barton, N. H. & Hewitt, G. M. Analysis of hybrid zones. *Annu. Rev. Ecol. Evol. Syst.* **16**, 113–148 (1985).
10. Dufkova, P., Macholan, M. & Pialek, J. Inference of selection and stochastic effects in the house mouse hybrid zone. *Evolution* **65**, 993–1010 (2011).
11. Beyersd. N. & Heckel, G. Structure and dynamics of hybrid zones at different stages of speciation in the common voles (*Microtus arvalis*). *Mol. Ecol.* **23**, 673–687 (2014).
12. Wielstra, R. *et al.* A genomic footprint of hybrid zone movement in crested newts. *Evol. Lett.* **1**, 93–101 (2017).
13. Hewitt, G. Quaternary phylogeography: the roots of hybrid zones. *Genetica* **139**, 617–638 (2011).
14. Butlin, R. Speciation by reinforcement. *Trends Ecol. Evol.* **2**, 8–13 (1987).
15. Servedio, M. R. & Noor, M. A. F. The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Evol. Syst.* **34**, 339–364 (2003).
16. Saestre, G.-P., Kral, K., Bures, S. & Ims, R.-A. Dynamics of a clinal hybrid zone and a comparison with island hybrid zones of flycatchers (*Ficedula hypoleuca* and *F. albicollis*). *J. Zool.* **247**, 53–64 (1999).
17. Saestre, G.-P. *et al.* A sexually selected character displacement in flycatchers reinforces premating isolation. *Nature* **387**, 589–592 (1997).
Acknowledgements

We thank J Jacquiéry and O Drillon for providing samples, as well as M Stöck and G Dufresnes for their help on the field. V Gvoždík and two anonymous reviewers provided useful comments on this manuscript. This study was funded by a grant from the Swiss National Science Foundation (n° 31003A_166323) to NP and a PhD fellowship from the Faculty of Biology and Medicine of the University of Lausanne to CD; SLN was supported by a grant from the Russian Foundation for Basic Research (n° RFBR.20-04-00918).

Author contributions

C.D. and N.P. designed the study. C.D. and M.B. conducted fieldwork. C.D. conducted labwork and analyzed the data with S.D. and S.N.L. C.D. wrote the manuscript, improved by all co-authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-62382-4.

Correspondence and requests for materials should be addressed to C.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020