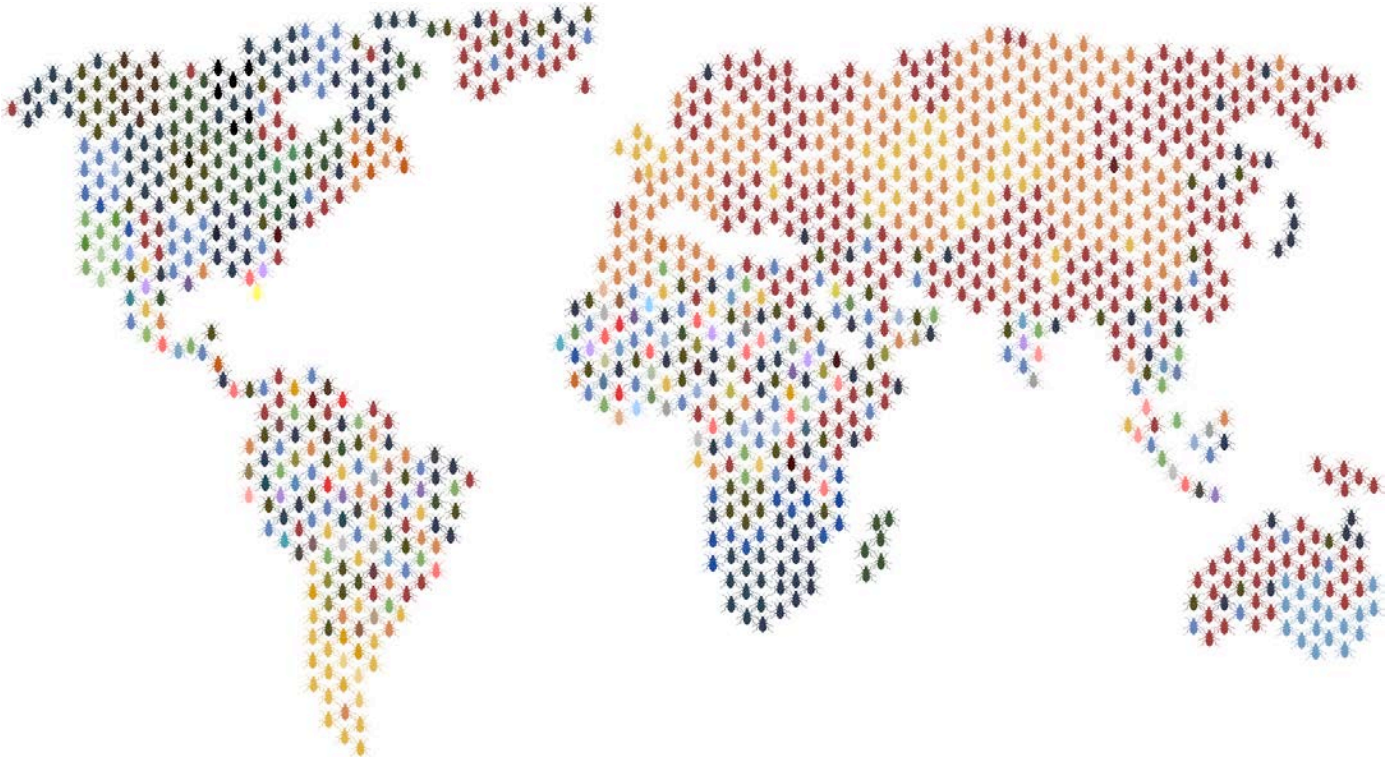
UNIVERSITAT DE BARCELONA

Departament de Biologia Evolutiva,
Ecologia i Ciències Ambientals

Species persist while haplotypes perish

Global patterns of genetic diversity and structure across freshwater insect species

Anna Sofia Salinas Ivanenko



Director/tutor: Dr. Cesc Múrria

Màster en Ecologia, Gestió i Restauració del Medi Natural
September, 2018

UNIVERSITAT DE BARCELONA

Departament de Biologia Evolutiva,
Ecologia i Ciències Ambientals

Species persist while haplotypes perish

Global patterns of genetic diversity and structure across freshwater insect species

Anna Sofia Salinas Ivanenko

Student:
Anna Sofia Salinas Ivanenko

Director/tutor:
Dr. Cesc Múrria

Màster en Ecologia, Gestió i Restauració del Medi Natural
September, 2018

Abstract

Intraspecific evolutionary histories are driven by a number of intrinsic (e.g. life-history traits) and extrinsic (e.g. biogeography) factors which result in vast variability of patterns. Usually, comparative phylogeographic studies use few regionally co-occurring species to uncover patterns for inferring biogeographic history, which may lead to a misinterpretation of generalizable processes driving these patterns. Here, we implement a macrophylogeographic approach assessing multiple freshwater insect species belonging to Ephemeroptera and Coleoptera from Panama and across a latitudinal gradient in Europe. Our aim was to unravel the processes shaping intraspecific genetic diversity and structure across these species in order to understand the role of long-term habitat stability. We found genetic diversity and structure were positively correlated, with Panamanian species showing the highest genetic diversity and structure due to long-term and also contemporary environmental stability. In Europe, the latitudinal derived patterns expected from climate-driven demographic fluctuations during the Pleistocene were rejected, suggesting a weak influence of historical events but a strong contribution of contemporary processes. Lineage-specific dispersal ability for European species was the main driver of intraspecific genetic structure, with mayflies showing more structured populations than co-distributed beetles. Nonetheless, phylogenetically close-related species within *B. rhodani* cryptic species complex revealed a wide range of haplotype network structures despite their limited, conserved dispersal ability. This idiosyncrasy in Europe was attributed to local stochasticity associated to high habitat instability. Moreover, the similar pre-pleistocene clade age of species from both continents decoupled from the radically different population genetic patterns found revealed European haplotypes locally perishing under contemporary habitat instability yet species persist at larger spatial scale. The macrophylogeographic approach outlined here will be increasingly useful for determining the existence, or lack thereof, of generalizable intraspecific genetic patterns across species which should uncover the extent at which macroecological processes influence patterns at the population level.

Resum

Les històries evolutives a nivell d'espècie són influenciades per factors intrínsecs (e.g. trets d'història de vida) i extrínsecs (e.g. biogeografia) que donen lloc a patrons variables. La filogeografia comparativa utilitza poques espècies co-distribuïdes regionalment per tal de detectar patrons i inferir la història biogeogràfica, el que pot provocar una mala interpretació dels processos que condueixen a aquests patrons. Aquí implementem un enfocament macrofilogeogràfic examinant diverses espècies d'insectes aquàtics que pertanyen als ordres Ephemeroptera i Coleoptera de Panamà i al llarg d'un gradient latitudinal europeu. El nostre objectiu és investigar els processos que han portat a la diversitat i estructura genètica intraespecífica, per tal de comprendre el paper de l'estabilitat de l'hàbitat a llarg termini. Vam trobar que la diversitat i estructura genètica es correlacionen positivament, les espècies panamenyes mostren més diversitat i estructura genètica a causa de l'estabilitat ambiental tan històrica com contemporània. A Europa, es van rebutjar els patrons latitudinals previstos per les fluctuacions climàtiques del Pleistocè, fet que suggereix una influència feble dels esdeveniments històrics, però una forta contribució dels processos contemporanis. La capacitat de dispersió dels diferents llinatges europeus va ser el principal promotor de l'estructura genètica intraspecífica, amb els efemeròpters presentant poblacions més estructurades que els escarabats. No obstant, el complex d'espècies críptiques *B. rhodani* va manifestar una àmplia gamma d'estructures haplotípiques malgrat la seva limitada capacitat de dispersió , que atribuïm a la estocasticitat associada a la inestabilitat local de l'hàbitat. L'edat pre-pleistocènica similar de les especies d'ambdós continents, desvinculada dels patrons genètics radicalment diferents de les poblacions, va revelar que els haplotips europeus poden desaparèixer localment per la inestabilitat de l'hàbitat, però les especies persisteixen en escales espacials més grans. L'enfocament macrofilogeogràfic que es descriu aquí serà cada cop més útil per determinar la influencia dels processos macroecològics a nivell poblacional.

Acknowledgments

First and foremost, I would like to thank my tutor, Cesc Múrria, for the unconditional support, continuous encouragement and invaluable contribution to my intellectual and academic development. This investigation has taken me through a journey of great discoveries and tough challenges and your positive outlook and active problem-solving has made this an amazing experience. To the rest of the Macro&Evo team, Aida, Aina, Gabone, Simmona and Dani, I thank you for your inputs and kindness, you have been very helpful. I also thank the FEHM research team for creating a great working environment.

Thank you to Alfried Vogler for facilitating resources and collaborating in the creation of the huge DNA dataset that made this study possible, and also for the stimulating questions on evolutionary biology. I would also like to thank Carmen Zamora and Núria Bonada, the directors of the Richabun project, as well as all collaborators that participated in the collection and identification of the samples in Europe. In the same way, I would like to thank Matt Whiles and Amanda Rugenski for the Panama dataset.

Thank you to Miguel Arnedo for your help with the phylogenies and your valuable insights. Also, thank you to the members of our journal club for the thought-provoking discussions.

Thank you, Pol, for putting through countless conversations about genetic diversity and genetic structure and for keeping me motivated. Your support has meant everything to me.

To my family, thank you for your patience and care throughout this process. Mami you are always able to keep me calm and Pali and Gigi you can make a stressful situation humorous, thank you for that.

I am dedicating this investigation and all the hard work put into it to my Dad.

Table of Contents:

1. Introduction

1.1. Scales in ecology

Ecological patterns are scale dependent. When researching distinct hierarchical levels, from genotypes to species and communities, or different spatial and temporal ranges, each scale may provide unique ecological information (Vellend, 2016). Just as when one admires a Jackson Pollock painting, glancing from a far and visualizing the work of art as a whole versus stepping closer and exploring in detail the dynamic between strokes of paint, each perspective adds information to the viewer and contributes to the visual experience. Specifically, in evolutionary biology, the different organizational levels and scales have been historically treated as separate and independent disciplines, with microevolution (within species, population level) seen only in the fields of population ecology and genetics, while macroevolution (among species, community level) was exclusive to the fields of community ecology and phylogenetics (Avise et al., 1987; Vellend, 2003, 2010; Vellend & Geber, 2005). Yet, the acknowledgment of an undoubtable link between both scales was simply put by Avise et al. (1987) who explained how "organism have parents, who in turn had parents, and so on back through time" (pp .490), then corroborated with coalescence theory (Hickerson et al., 2010). The merger of within and among species levels across different spatial scales (Antonovics, 2003) would result in the contemporary field of phylogeography (Hickerson et al., 2010) which concerns itself with determining population genetic patterns conceptualized geographically (i.e., phylogeographic patterns) and examining the ecological and evolutionary processes that drive them (Avise, 2000; Hickerson et al., 2010).

Patterns often seen in one scale do not extrapolate to the next (Múrria et al., 2012; Vamosi et al., 2009; Vellend et al., 2014). Yet, the main four processes that drive evolution (drift, mutation/ speciation, selection and gene flow, Vellend, 2010) can be seen acting sequentially, from the intraspecific to the interspecific scale, and are in turn the processes that set the patterns (Marske et al., 2013). For this reason, to advance our understanding of evolution, phylogeographical analysis integrating intraspecific and interspecific perspectives is essential (Baselga et al., 2015; Craft et al., 2010; Dexter et al., 2012). Unfortunately, given the financial and technological limitations for massive sequencing, most comparative phylogeographical studies have been conducted with one or few species (Avise, 2009). This entails an absence of information and an inability to reliably extrapolate the historical and

contemporary macroecological processes that drive evolution at the whole-community level (Baselga et al., 2013; Marske et al., 2013; Múrria et al., 2015, 2017).

1.2. Macrophylogeography

Here, we aim to expand comparative phylogeography (i.e., test if species co-distributed in space and time share a common history, Gutiérrez-García & Vázquez-Domínguez, 2011; Page & Hughes, 2014) by using a multi-species approach among phylogenetically close related species and between lineages, across large spatial and temporal scales (macroecological perspective), an approach we call macrophylogeography. This should allow an appreciation of both the detail of ever stroke of paint as well as the art work as a whole. To determine macrophylogeographic patterns, population genetic tools such as population structure and genetic diversity estimates are fundamental. Genetic diversity measures the heritable variation found within a particular biotic entity (individual, population, species) (Lowe, 2004), while population structure measures the spatial distribution of the magnitude of genetic diversity among populations (Allendorf et al., 2013; Bohonak, 1999; Lowe, 2004). Given their usefulness to address a range of ecological and evolutionary questions (Knowles, 2009), relationships between genetic diversity or population structure and life-history traits (i.e. relationship between dispersal ability and population structure (Hughes, 2007; Miller et al., 2002; Papadopoulou et al., 2009; Zeller et al., 2006; Zickovich & Bohonak, 2007;) or abiotic variables (i.e. relationship between environmental instability and genetic diversity (Bertheau et al., 2013; Qu et al. 2014; Vangestel et al., 2012) have commonly been examined. Yet a relationship between genetic diversity and population structure has, to our knowledge, never been tested.

Moreover, given the shared factors that may simultaneously influence both of these genetic variability estimates (e.g., dispersal abilities, habitat instability), genetic diversity and population structure are not independent from one another thus their relationship may be useful to disentangle macroecological processes shaping current phylogeographic patterns. To uncover these patterns, many phylogeographers have historically relied on mitochondrial DNA (mtDNA) as a molecular marker for tracking demographic histories (Avise, 1989; Zink & Barrowclough, 2008). Due to its lack of recombination, uniparental mode of inheritance, and rapidly evolving nature, mtDNA has been extremely useful at constructing cleaner genealogies which reveal clearer population histories. (Avise, 2009; Lowe, 2004). mtDNA has also generally been considered a nearly neutral marker (Baselga et al., 2015; Nabholz

et al., 2008;), which means the expected rate of nucleotide substitution is equal to the mutation rate and reveals no fitness related effect (Hedrick, 2011), simplifying the analysis of phylogeographic patterns. Discounting the shuffling effect of syngamy and meiosis and the directional effect of selection, mtDNA variation is fully dependent on mutations alone (Avise, 2009). Since mtDNA mutation rate is assumed to be constant and lineage-specific, this facilitates the comparison of genetic diversity and structure among species (Brower, 1994). Although mtDNA's effectiveness as a historical demographic tracker has been put into question (Bazin et al., 2006), studies on mammals (Nabholz et al., 2008) and avian lineages (Zink & Barrowclough, 2008) have countered these claims, validating its use as a reliable indicator of phylogeographic patterns (Avise, 2009; Zink and Barrowclough 2008).

1.3. Testing stability

Assuming neutrality with respect to fitness (Kimura, 1983), variation in population structure is primarily affected by barriers to gene flow (allelic exchange between population), whether through physical or ecological factors and/or functional traits (Lowe, 2004) (Fig. 1). In parallel, genetic diversity within a population is influenced by demographic fluctuations in effective population size (Ne) (Ellegren & Galtier, 2016), mutation rate (Hedrick, 2011) and time to coalescence (Avise, 2004; Fischer, 1960) as well as population structure also through barriers to gene flow (Zickovich & Bohonak, 2007) (Fig. 1). All factors contribute to the increase or reduction of genetic diversity and structure in a similar way, with the exception of genetic drift which leads to an opposite effect (Fig. 1). When effective population size is at optimum, and therefore the effect of drift is minimized, intraspecific genetic structure and diversity should be positively correlated. This scenario would reflect long-term habitat stability achievable through a lack of perturbations causing potential demographic fluctuations (Baselga, 2008; Carnaval et al., 2009; Hawkins et al., 2006; Ribera et al., 2003). Yet, in unstable scenarios with low effective population size possibly caused by population bottlenecks or founder effects associated to habitat instability, stronger genetic drift may lead to a loss of genetic diversity in a population (Lacy, 1987) in addition to an increase in structure between populations (Hughes et al., 2009; Young et al., 1996).

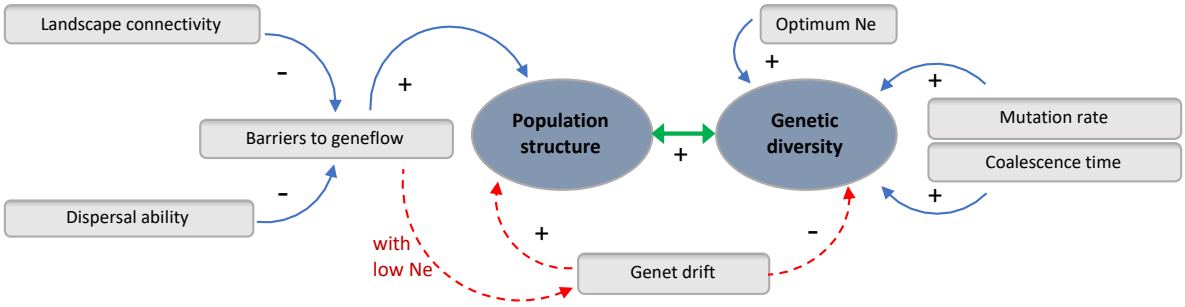


Figure 1 | Simplification scheme outlining the main factors influencing genetic diversity and population structure (in reference to neutral loci). Blue arrows reveal correlations assuming a stable scenario while red dashed arrows reveal instability. Gene flow, effective population size, mutation rate, time to coalescence and genetic drift are the main influencing factors. The direction of correlation is indicated by positive and negative symbols. Modified from Ellegren and Galtier, 2016.

In order to test the effect of long-term habitat stability and intrinsic functional traits as macroecological processes on the relationship between intraspecific genetic diversity and structure, aquatic macroinvertebrate within a variety of lineages and presumably different states of historic stability were analyzed. At the intra- and interspecific levels, stability may be reflected across a latitudinal genetic and biodiversity gradient (Martin & McKay, 2004) which may be partly due to northern impoverishment through Pleistocene glaciations (Young et al.,1996). Species ranges in northern latitudes experienced an unstable history of extensive extinction and recolonization events while southernmost latitudes in Europe reveal complex refugia (Hewitt, 2004; Hortal et al., 2011). At the tropics, Pleistocene glaciations should favor altitudinal rather than large distance latitudinal movements suggesting more long-term stability than in European regions (Múrria & Hughes, 2008; Rull, 2005, 2006). For this reason, genetic variability across a latitudinal gradient from northern Morocco to northern Europe (Múrria et al., 2017) and longitudinally across highland streams in Panama (Múrria et al., 2015) were analyzed.

In general, phylogeographic patterns in aquatic macroinvertebrates may be affected by lineage-specific life-history traits as well as biogeographical history and contemporary ecological conditions (Carnaval et al., 2014; Hughes et al., 2009; Pauls et al., 2014). The relative influence of these factors would hence determine a generalizable pattern of population genetic diversity and structure across lineages. Therefore, analysis of genetic variability of multiple-species across a gradient of long-term habitat stability provides an opportunity to examine the macroecological signature which may have been imprinted by

past glaciations (Lehrian et al., 2009) or more recent processes (such as stream intermittency, Múrria et al., 2010) versus patterns produced by species-specific life-history traits, such as, dispersal ability (Alp et al., 2012) or landscape connectivity (Múrria et al., 2013). To test these macroecological expectations, species within several lineages with distinct dispersal abilities (Sarremejane et al., 2017) and territorial ranges across Europe (Múrria et al., 2017) were compared to counterpart lineages from Neotropical highland streams in Panama (Múrria et al., 2015) (Fig. 2).

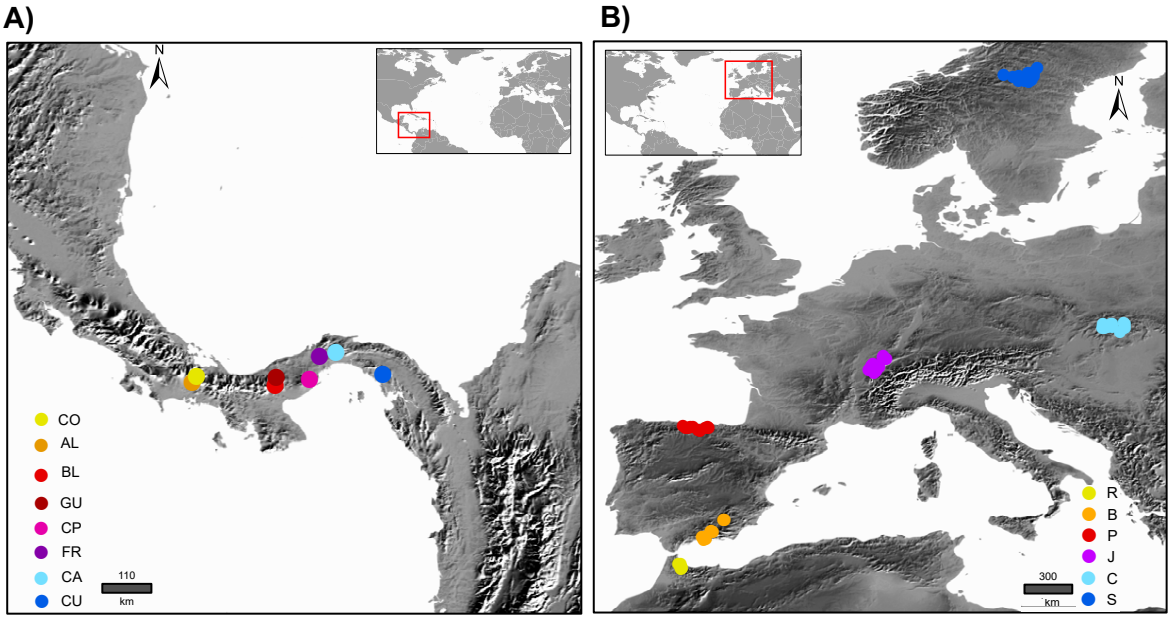


Figure 2 | **(A)** Distribution of study sites across a latitudinal gradient from northern Morocco to northern Europe (mountain regions: Riff (R), Betic (B), Picos (P), Jura (J), Carpathians (C) and Jämtland (J) and **(B)** a longitudinal gradient in Panama (regions: Chorro (CO), Aleman (AL), Blanco (BL), Guabal (GU), Capira (CP), Frijolitos (FR), Cerro Azul (CA), Chuchantí (CU).

From a classic view point, Hewitt's (1996) assertion of a strong influence from Pleistocene glaciations on current species assemblages in Europe has been relatively unchallenged and well accepted for trees and vertebrates (Svenning et al., 2011; Svenning & Skov, 2007). If this extrapolates onto freshwater macroinvertebrates at the intraspecific scale and biogeographical history is the main driver of genetic patterns (proven not to be the case at the community level, Múrria et al., 2017), we hypothesize that (1) species in the north would show less intraspecific genetic diversity and structure than in the south given expected bottlenecks and extinctions as well as few founder events, which would be evident with latitudinal clusters (Fig. 3 A). Alternatively, if life-history traits are the main drivers of genetic patterns, (2) genetic structure would be similar among phylogenetically close related

species, given their shared dispersal ability, and independent to genetic diversity which would correspond to the species evolutionary age (Fig. 3 B). Yet, genetic patterns may also be determined by a combination of both factors (Fig. 3 C). In this case, (3) species with low dispersal ability would be structured independent of the latitude they inhabit or their genetic diversity, while good dispersive species would show more structure and diversity in southern stable environments and be less structured and diversified in northern latitudes. European species from southern latitudes driven by expected long-term stability (hypothesis 1 and 3) would also reveal a positive correlation between genetic diversity and structure given the presence of stable refugia, yet this correlation is expected to be lower and weaker than in Neotropical areas of Panama (Fig.3).

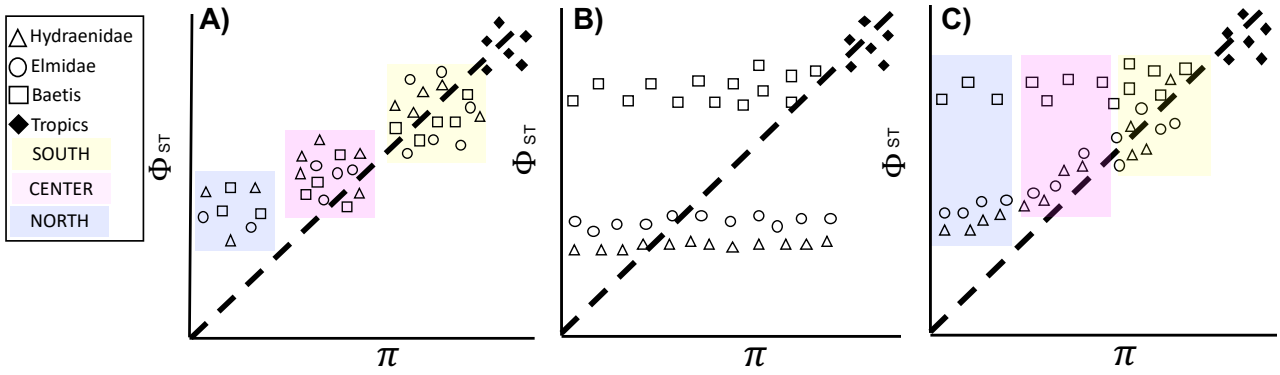


Figure 3 | Schematic representation of the three hypotheses used to investigate the role of biogeographic history in the context of long-term stability and dispersal ability on intraspecific genetic patterns. The dotted diagonal line represents the expected positive correlation between genetic diversity and structure in a scenario of long-term stability. Genetic patterns influenced primarily by biogeographical history **(A)**, dispersal ability **(B)**, or equally influenced by both **(C)** are hypothesized with changes in estimates of Φ_{ST} and π. European lineages and species from the tropics are represented with distinct symbols and presence in different regions of Europe is represented with different colors.

Through this investigation we aim to (1) determine whether genetic diversity and structure are correlated within and across lineages considering dispersal abilities and biogeographic history associated to long-term habitat stability. We will also, (2) uncover the factors that may lead to deviations from this correlation (higher genetic diversity relative to structure and vice versa). Finally, we will examine the micro and macroevolutionary patterns for each species in order to (3) assess the consistency of intraspecific genetic patterns, visualized with haplotype networks, within each lineage, and (4) identify interspecific divergence time, with time calibrated trees, to assess its relationship with genetic diversity.

2. Material and Methods

2.1. Study sites and subjects

Genetic data for European macroinvertebrates was obtained from Múrria et al. (2017), which sampled 62 streams (150 meters reaches per stream) in the fall and spring of 2008 across six mountain regions within three climatic zones. Sampled sites exemplifying each climatic zone where as follows: Mediterranean climate in Rif (R), Morocco and Betic (B), southern Spain; temperate climate in Picos de Europa (P), northern Spain, Jura (J), France and Carpathians (C), Slovakia; and subarctic climate in Jämtland (S), Sweden (Fig. 2). All sites were spatially independent (no connection through flow) pristine headwater streams (first/second order) characterized by permanent flow regimes and calcareous soils. Intraregional sampling patterns were very similar among regions with comparable spacing distances between sites (mean of 53.27-71.85 km) and number of sites within region (10-11), with the exception of Rif which, due to limited calcareous zones, had less sites (9) and shorter intersite distances (mean of 15.65 km) (Múrria et al., 2017). This same factor affected interregional sampling distance and resulted in a sampling gap between Carpathians and Jämtland. Yet, despite these limitations, the distribution of sampled regions and climatic zones were effective for establishing a latitudinal gradient and are thus pertinent for this study.

Macroinvertebrate genetic data from Panama in Múrria et al. (2015), was used as an example of species from the tropics and a control for the stability hypothesis. Like in Europe, 150-meter reaches were sampled from February to March of 2011 in 8 streams along a longitudinal gradient starting from west-most site at Chorro (CO) to Aleman (AL), Blanco (BL), Guabal (GU), Capira (CP), Frijolitos (FR), Cerro Azul (CA) and finishing in east-most Chucantí (CU) (Fig. 2). All streams were sampled at similar altitudes (700-1200m), with the exception of lowland Frijolitos, and were spatially independent, located within five rainforests separated by mountains, with only Aleman and Chorro connected by water. Distances between most sites ranged from 50 to 180 km and conditions within each site were fairly consistent with similar habitats, substrate, stream width, temperature, and discharge.

Genetic data of European species of the mayfly (Ephemeroptera, Insecta) genus *Baetis* (Baetidae) and the aquatic beetle families (Coleoptera, Insecta) Elmidae (seven genera) and Hydraenidae (three genera) were compared to Panamanian species of the aquatic beetle

genus *Anchytarsus* (Ptilodactylidae) and the mayfly families Leptophlebiidae (two genera), Baetidae (two genera) and Leptohyphidae (one genus).

For Europe, species were first sorted out by number of individuals sampled and their site of provenance. Only species fulfilling the criteria of at least 5 individuals sampled per site and found in at least 3 sites (5x3) were used for genetic variability analyses. This was implemented in order to obtain valid estimates given genetic diversity requires a minimum number of individuals while population structure requires a comparison between more than one population. For Panama, estimates of genetic diversity and structure were also calculated with similar criteria, but it was reduced to 4 individuals per site in at least 2 sites given the high diversity and structure of tropical species.

2.2. Genetic variability estimates

Cytochrome *c* oxidase subunit I (*cox1*) segments of mtDNA (fragments that differ from one another through mutations accumulated from their last shared female ancestor) are commonly used to obtain estimates of genetic diversity and structure (Avise, 2009). Moreover, genetic variability of *cox1* is frequently used in DNA-taxonomy to delimit species as sequence variation clusters haplotypes into molecular entities that roughly approximate morphological Linnaean species and discriminates between morphologically similar species (Hebert et al., 2003). 658-bp of the *cox1* gene were amplified for all lineages in Panama and for *Baetis* in Europe, whereas fragments of 822-bp were obtained for the two beetle families in Europe (Múrria et al., 2015, 2017). Fragments were collapsed into unique haplotypes and used for molecular species delimitation with the generalized mixed Yule-coalescent (GMYC) model applied separately to each lineage (Pons et al., 2006). The GMYC model refines the boundary between intra and interspecific genetic variation and delineates populations from species, hence avoiding possibly inflated estimates of genetic variability caused by cryptic species (Fujita et al., 2012; Fujisawa & Barraclough, 2013). Therefore, the species unit will hereafter refer to molecular GMYC entities/species.

Genetic diversity was calculated as the average number of nucleotide differences per site between two sequences, symbolized by π (Nei,1987) with software DNASP v6.11.01 (Rozas et la., 2017). Population genetic structure was calculated as the ratio of genetic diversity between and within populations, also referred to as the fixation index F_{ST} or Φ_{ST}, the latter

more applicable to mtDNA as it incorporates a measure of distance between haplotypes (Excoffier et al., 1992). As the value of the fixation index increases, the genetic variation attributed to haplotype differences between populations is larger than the genetic variation found within populations (Holsinger & Weir, 2009). This reveals a higher demographic structure and lower gene flow between populations (Lowe, 2004). Here, genetic structure was measured as pairwise Φ_{ST} using an analysis of molecular variance (AMOVA) with ARLEQUIN v3.1 (Excoffier et al., 2005). For species found in several European regions, two values of genetic diversity and structure were calculated, one by including all haplotypes derived from that species independent of region and another by partitioning haplotypes into regions yielding multiple values of π in order to analyze interregional differences. Assuming the 5x3 criteria, when species were found in 3 sites per region and in multiple regions an Φ_{ST} value for the total sites as well as Φ_{ST} values for sites within each region were calculated, given how the criteria was met both regionally and in a multiregional context. However, when a species was found in many regions but did not meet the criteria regionally (5x3 per region) a total Φ_{ST} value was calculated but Φ_{ST} regional values were excluded.

The potential correlation between π and Φ_{ST} was analyzed for Europe and Panama using a linear regression model in R package (R Core Team, 2016)

2.3. Micro and macroevolutionary analysis

To identify and visually represent the microevolutionary (intraspecific) genetic patterns of haplotype diversity and geographic structure, haplotype networks were estimated for species in Europe fulfilling the 5x3 criteria with 95% statistical parsimony (Posada & Crandall, 2001) using TCS v.1.21 (Clement et al., 2000). To solve network ambiguities (multiple connections between haplotypes), we used the geographical criteria (Crandall & Templeton, 1993) which asserts that haplotypes within the same population are more likely to be connected. Haplotype network configurations were then subjectively compared among phylogenetically close and distant related species in order to determine the consistency, or lack thereof, of genetic patterns and phylogenetic signals within sister species.

To uncover macroevolutionary patterns, phylogenetic trees of European and Panamanian species complexes (identified in Múrria et al., 2015, 2017) were estimated using a Maximum Likelihood method (RAXML- CPC2 v.8.2.10; Stamatakis, 2014) and a Bayesian method

(MRBAYES v.3.2.6; Ronquist & Huelsenbeck, 2003) through CIPRES SCIENCE GETAWAY v.3.3 (a national science foundation XSEDE resource (Miller et al., 2010). To run these phylogenetic analyses the codon positions of the sequences for each lineage were identified with MESQUITE v3.40 (Maddison & Maddison, 2018) and the evolutionary models of each codon position (necessary for Bayesian analysis) were determined with a Bayesian information criterion (BIC) in PARTITION FINDER v2.1.1 using all haplotypes sequences within each lineage (Lanfear et al., 2017). Strong node probabilities and congruency between Maximum Likelihood and Bayesian trees were verified in order to proceed to time calibration.

Since time to coalescence may influence the accumulation of mutations (Fischer, 1960), genetic diversity should presumably be correlated to divergence time between species. For this reason, all phylogenies were time calibrated in order to evaluate this relationship with a Bayesian MCMC analysis using BEAST v.1.10.0 (Suchard et al., 2018). Due to the lack of dated fossils or biogeographical events that could be used for insect calibration, the molecular dating analysis was carried out with mutation rate priors testing four approaches to identify the priors which yielded better results. First, in BEAUTI v.1.10.0 (Suchard et al., 2018) a high mutation rate of 0.0177/ site/ million years (Papadopoulou et al., 2010) and then a low mutation rate of 0.115/site/million years (Brower, 1994) were fixed under a speciation yule tree model with a lognormal relaxed clock. Then, using the same molecular clock and tree prior settings, the fixed mutation rate was changed to a normal distribution (mean 0.0146 ± 0.002 to incorporate the high and low mutation rate within the range of values analyzed). Finally, the last approach consisted in a uniform distribution with an initial value of 0.0146 (as the mean of 0.0115 and 0.0177) and an upper and lower value of 0.018 and 0.0112, respectively. All models were run for 20 million generations, sampling every 1000 generations (time period in which a tree is saved).

The optimum evolutionary models obtained were applied to each codon position by unlinking substitution models in BEAUTI v.1.10.0, however given the complexity of the evolutionary models and the limited number of species used, models needed to be simplified (excluding gamma or invariant properties at times) in order to converge and to reveal effective sample size (ESS) values higher than 200. Additionally, to give further robustness and improve the analysis (increase values of ESS), additional species from Genbank (https://www.ncbi.nlm.nih.gov/genbank/) were added to some species complexes that included few species. Convergence and ESS values for each analysis were then evaluated

with TRACER v.1.6 with a 10% burn-in (Suchard et al., 2018) to identify which approach produced the best results. The consensus tree from the chosen approach was compiled with TREEANNOTATOR v.1.10.0 (Suchard et al. 2018) and visualized in FIGTREE v.1.4.2 (Rambaut, 2014).

3. Results

3.1. Correlation between genetic diversity and structure

A positive correlation between π and Φ_{ST} both with regional (Fig. 4 A) and total (Fig. 4 B) genetic variability estimates was found. Total estimates showed a normal distribution of residuals (Shapiro Wilks test p value > 0.05), however, given how residual normality for regional values was only confirmed when log transforming π, an exponential model was deemed a better fit for the regression analysis (Total values: t_{31}= 4.817; p value < 0.05; R^2=0.41 and regional values t_{36}= 5.773 p value < 0.05; R^2=0.46 (Appendix S2). In general, Panama species exhibited higher genetic diversity and structure than those in Europe, with the exception of M16, M17 and B6 in the *Hydrosmilidon* and *Baetodes* genera (Ephemeroptera) and *Anchytarsus* genus, respectively (Appendix S3). For genetic variability estimates in a regional context (Fig. 4 A), the vast majority of European species showed little variation and a narrow range of genetic diversity π across species, with most estimates lying below 0.01. Several northern or central species showed higher genetic diversity than species from the south, therefore π appeared to have no apparent association to latitude. On the other hand, regional Φ_{ST} estimates indicated high interspecific variation in structure, yet once again did not correspond to a latitudinally derived pattern. Φ_S variation seems to be related to lineage-specific patterns given the generally higher estimates found in *Baetis* species compared to Hydraenidae and Elmidae species. Nevertheless, several exceptions (E14, B9, Appendix S3) imply a limited influence of lineage-specific life-history traits.

Total genetic diversity π and structure Φ_{ST} estimates, which included all populations of a species distributed in several regions (Fig. 4 B), were on average higher than regional π and Φ_{ST}. This trend was vastly influenced by Elmidae species distributed in several regions which showed the highest Φ_{ST}, both in a regional and total context (>0.6). This indicates limited gene flow between regions in spite of Elmidae's strong dispersal ability. The high variability of the genetic structure Φ_{ST} estimates combined with relatively low genetic diversity π in

most European species suggests that genetic diversity is being affected by different factors which are likely independent across species and should cause deviations from the regression line.

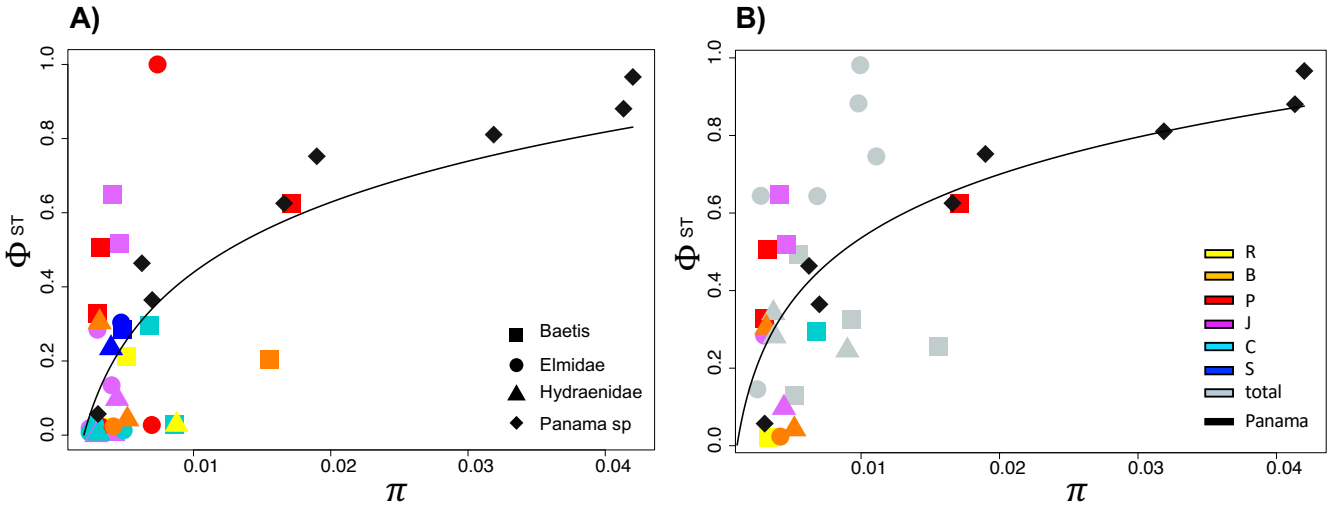


Figure 4 | Correlation between π and Φ_{ST} of species from Europe and Panama. Europeans species of *Baetis,* Hydraenidae and Elmidae and all Panamanian lineages are represented with different symbols. Colors indicate regions from northern Morocco to northern Europe where species were present. Black indicates species from Panama irrespective of region of presence. **A** indicates regional estimate of Φ_{ST} and π while **B** indicates total estimates Φ_{ST} and π. When species are found in more than one region this is represented with grey.

3.2. Intraspecific genetic patterns

A total of 29 haplotype networks were generated that followed three generalized patterns. Common *star-like* networks recognized by a dominant ancestral haplotype connected to many low frequency haplotypes with little genetic differentiation from the ancestral sequence were identified (Fig. 5 A). This pattern is commonly attributed to population expansion, and it was found in species from northern, central and southern localities (Appendix S4). Species in all lineages also revealed haplotypic geographic restrictions evident in networks with *parochial structure* (Fig. 5 B). This pattern is suggested to be driven by low dispersal ability and identified through geographic clusters with haplotypes within regions showing closer genetic links than haplotypes between regions. Most species exemplifying parochial networks were Elmidae species (Appendix S4), which also showed high Φ_{ST} estimates. The third haplotype networks pattern is defined as having a *diffuse structure*, which is exemplified by a large number of haplotypes occurring in a wide range of localities (Fig. 5 C). This pattern

was also identified across lineages and latitudes with a surprising high number of examples for bad dispersers such as *Baetis* species.

There were no apparent haplotype network structures specific for each latitude or lineage. Species from either north, central or south of Europe revealed networks characterized by few or vast haplotypes as well as regionally conserved or inter-regionally linked patterns. Moreover, networks patterns were variable within each lineage which indicated species-specific evolutionary histories decoupled from expected patterns based only on dispersal ability (e.g. mayflies showing diffuse structures (B8, B9) and beetles showing parochial structures (E14, H3, Fig.5 A and C).

3.3. Relationship between coalescence time and genetic diversity

For Bayesian time calibration, substitution models for each codon positions (1,2,3) were simplified by replacing model F81 to GTR and removing invariants. The final used models were as follows: For Europe: *Baetis:* 1: SYM; 2: HKY and 3: GTR; and Hydraendae/ Elmidae: 1,2 and 3: GTR; and for Panama: *Anchytarsus* / Farrodes/ Haplohyphes: 1: SYM; 2 and 3: GTR; and Baetodes/ Moribaetis: 1 and 2: GTR; 3: HKY. All phylogenetic trees revealed strong posterior support at both basal and terminal nodes and HDP confidence intervals were wider in basal nodes (Fig. 6; Appendix S5). There was no clear difference in divergence times between Panamanian and European lineages or between beetles and mayflies. Within both regions most species showed divergence times prior to the beginning of the Pleistocene glaciations (~ 2.5 My) and to times of mayor climatic oscillations (~700 ky). No species revealed a divergence time posterior to the last glacial maxima (~20 ky) and to subsequent northern expansion (12 ky).

The genus *Baetis* was divided into three species complexes due to the high number of molecular entities delimited within the Linnean morphological species *Baetis rhodani* (10 GMYCs) and *Baetis alpinus* (6 GMYCs) (Appendix S5). Despite each complex encompassing only one Linnean species, phylogenies revealed very old divergence times, with the oldest split between the molecular species B9 and B10 in *B. rhodani* complex occurring around 8.78 Mya and *B. alpinus* molecular species B21 and B25 diverging around 5.26 Mya (Appendix S5). Identical morphology coupled with old divergence times revealed potential cryptic species for *B. rhodani* and *B. alpinus*. *Baetis alpinus* complex included high-

elevation specialist species that were geographically restricted to one region, whereas the more generalist *B. rhodani* contained species limited to one region as well as species inhabiting multiple regions, e.g., B8 was found in 5 out of the 6 regions. For Elmidae, similar patterns of old divergence between molecular entities delimited from unique Linnean species revealed potential cryptic species (E12-E14, E18-E19 and E6-E7, Fig 6 A). The majority of Elmidae species were collected in Picos de Europa, however many of these species were distributed across multiple regions (10 out of 21), revealing a wider distributional range than species from any other lineage and therefore indicating high dispersal ability. All Elmidae species present at northern latitudes were also found in central and southern latitudes (Fig. 6 A). Similarly, a large distributional range across several regions was also found for Hydraenidae, yet most species were exclusive to one region (16 out of 24) and the wide distributed species had a narrower distribution than Elmidae species. In contrast to Elmidae, Hydraenidae species inhabiting mainly southern latitudes (Rif, Betic and Picos) were phylogenetically clustered and separated from northern latitudes species (Jura, Carpathians and Sweden) (Appendix S5).

In general, species in Panama were mostly exclusive to one site (Appendix S5) and only one *Anchytarsus* species (B6) was found in 6 out of 8 sampled sites. Hence, species distribution in Panama was much more structured than those distributed across Europe (Fig. 6 A-B).

When comparing divergence time and genetic diversity π there was no clear relationship. Most European species showed very low genetic diversity (Fig. 4) and high variability in interspecific divergence time (Fig 6 A; Appendix S5). For example, *B. rhodani* B8 split much more recently than *B. rhodani* B2 (divergence time of 0.84 and 4.3 million years, respectively) nonetheless their genetic diversity was very similar (π =0.015 and 0.017, respectively). In fact, B8 together with its sister species B7, are the youngest species in this species complex (divergence times= 0.84) (Appendix S5). However, B8 has a much higher genetic diversity (one order of magnitude higher) than B7 or any of the other species within the complex (Appendix S1). Similar patterns were found across all studied lineages.

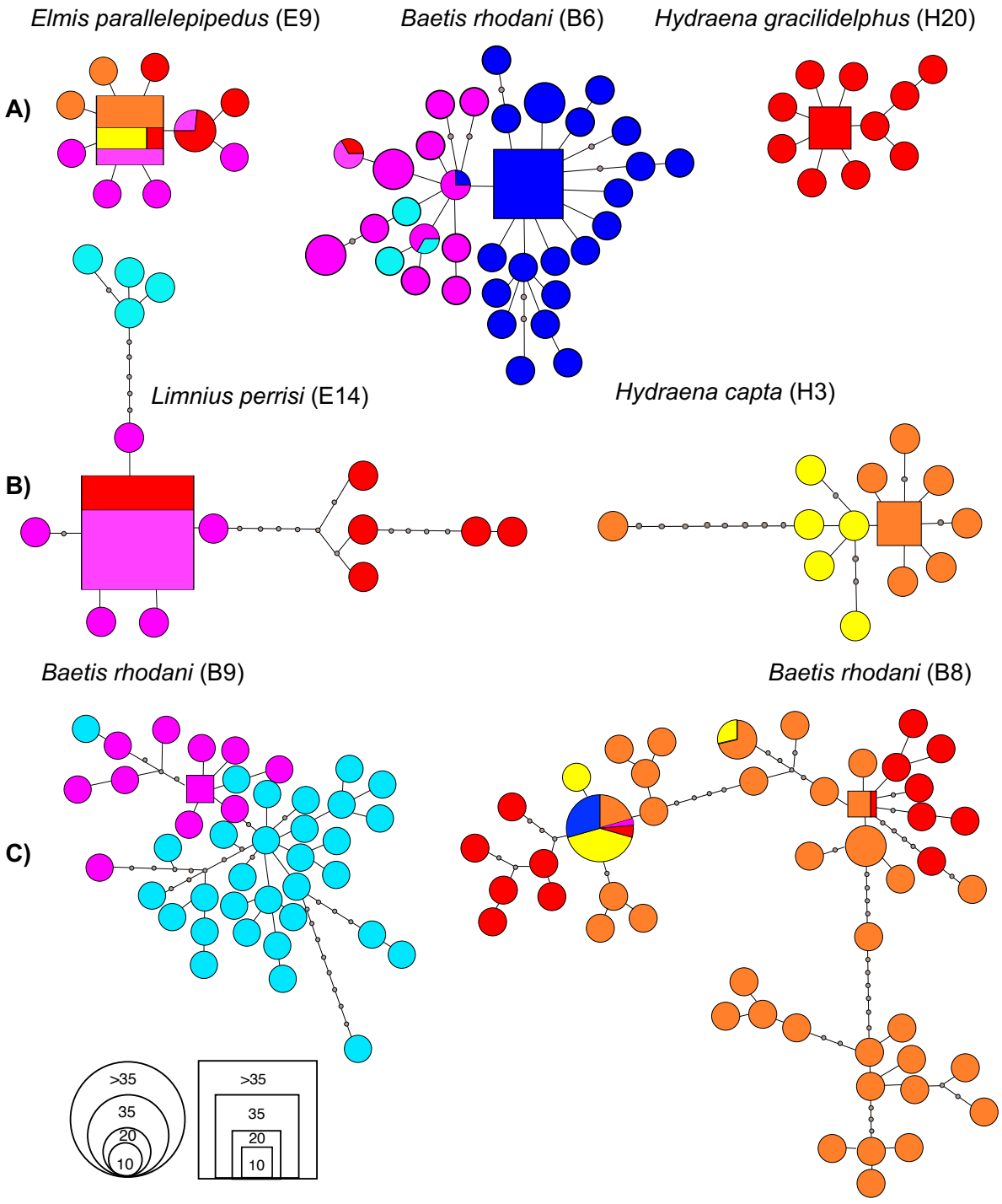


Figure 5 | Examples of European *Baetis*, Elmidae and Hydraenidae haplotype networks with morphological and molecular species labels. **(A)** Networks E9, B6 and H20 exemplify *star-like* patterns. **(B)** Networks E14 and H3 exemplify a geographically restricted *parochial* pattern while **(C)** networks B9 and B8 exemplify a geographically *diffuse* pattern. Colors corresponds to regions of presence (see Fig. 2). Squares represent haplotypes most likely to be ancestral (highest outgroup probability correlated to haplotype age) as determined by TCS. Size of circles and squares are proportional to haplotype frequency. Small grey dots represent unsampled or missing intermediate mutations.

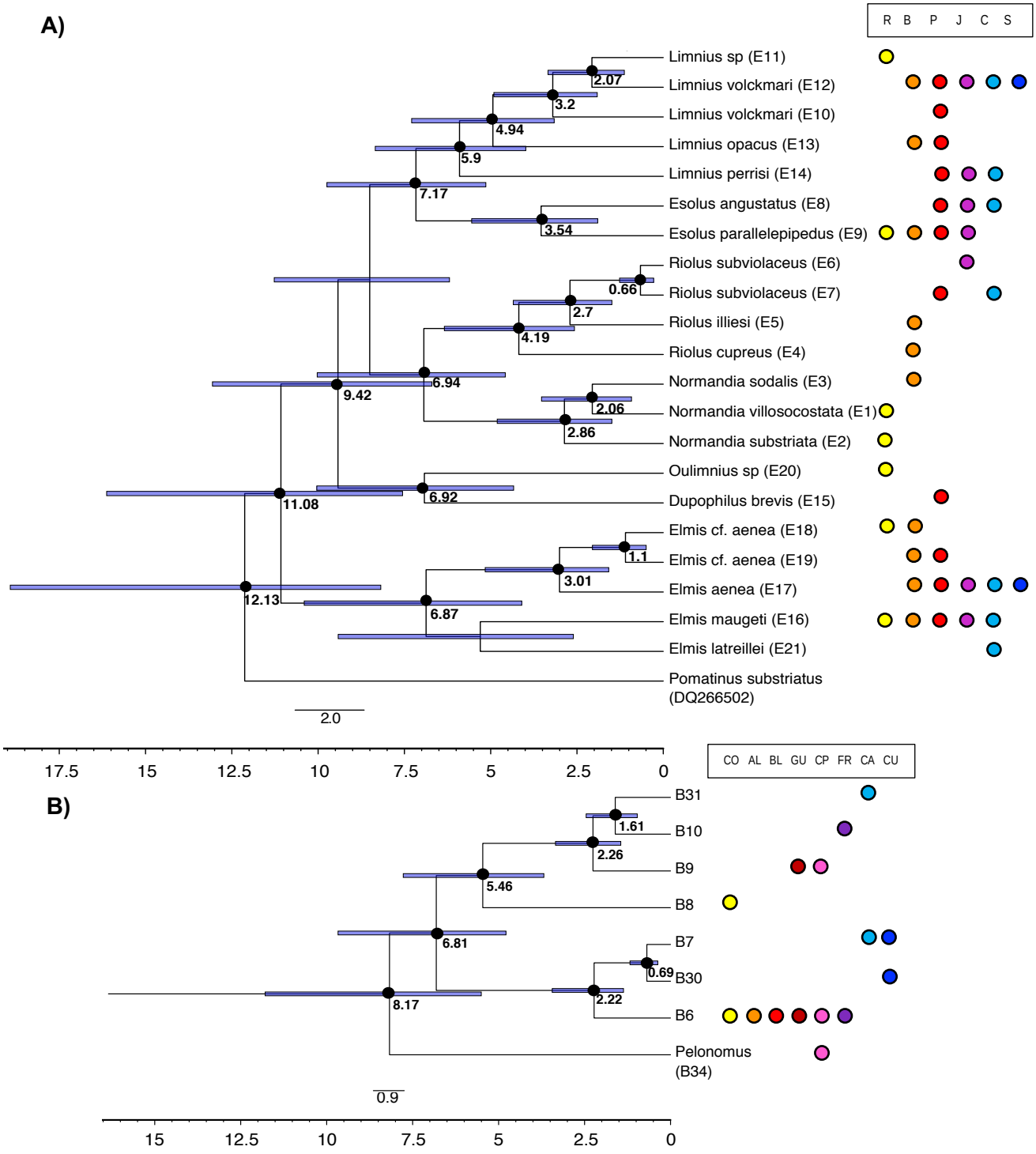


Figure 6 | *Cox1* time calibrated phylogenetic trees of **(A)** the beetle family Elmidae in Europe and **(B)** genus *Anchytarsus* in Panama. In Europe and in Panama beetles were taxonomically identified to the species and the genus level respectively. Codes correspond to the molecular species labeled in Múrria et al., 2015, 2017. Colored circles signify region of presence. Scales and numbers left of nodes indicate divergence time. Bars on nodes indicate the 95% highest posterior density (HPD) confidence intervals of the divergence time while black dots reveal posterior probabilities >0.95.

4. Discussion

Our macrophylogeographic approach uncovered global generalizable patterns across aquatic freshwater lineages as well as species-specific histories within each lineage. The discovery of a positive relationship between intraspecific genetic diversity and structure provides insight to the implications of long-term habitat stability at the population level. However, for reasons henceforth discussed, all our hypotheses were rejected suggesting genetic variability is driven by more factors other than biogeographic history or dispersal ability.

4.1. Genetic diversity within and across continental regions.

Genetic diversity patterns revealed a strong Quaternary signal across continents but not within Europe. Most Panamanian species were more genetically diverse than species from Europe, reaching π estimates up to one order of magnitude higher. At large scale, the low values of genetic diversity π (<0.01) in Europe, together with the high diversity found in Panama, produced a positive exponential correlation between genetic diversity and structure. The wide variance of genetic diversity estimates between species from Panama and Europe may be due to an array of factors. Theoretically, genetic diversity increases by stochastic mutations caused by DNA replication errors and by mutagenic DNA damage (Ellegren & Galtier, 2016). Consequently, the final amount of accumulated mutations in a population depends on the nucleotide substitution rate, the time span from the speciation event, and the effective population size (Fischer, 1960). Several studies suggested variation in mutation rate within invertebrate taxa, even among phylogenetic related lineages (Hebert et al., 2002; Thomas et al., 2006), thus standardized mutation rate for all arthropods was rejected (Brower, 1994). Although variable mutation rates could therefore explain differences in genetic diversity π, the higher variance in π found across regions than within lineages and the similar pre-pleistocene clade age found for European and Panamanian species discards varying mutation rate or divergence time as explanations for the inter-continental differences. Hence, only effective population size related to habitat stability seems to explain the variance of nucleotide diversity π across continents. In Panama, neotropical sites have been less affected by Pleistocene glaciations (Hewitt, 2000, Rull, 2005) and by contemporary environmental fluctuations (Martin & McKay, 2004) than in Europe. Thus, the Neotropics should conserve more unaltered population sizes and therefore yield higher values of genetic diversity π. This finding supports a parallel effect of

long-term environmental stability on species (at community level) (Pianka, 1966) and genetic (at population level) diversities in tropical regions and suggests a correlation between both levels of diversity (Múrria et al., 2015).

Across Europe, contrary to Panama, biogeographic history in terms of latitudinal differences in habitat stability associated to the cyclic Pleistocene glaciations was not seen to be the major influencing factor on current population genetic patterns. The absence of latitudinal clusters for all lineages related to estimates of either intraspecific genetic diversity or structure invalidates both biogeographical hypothesis 1 and 3 (Fig. 3 A & C). Therefore, our findings revealed a lesser effect of Pleistocene climatic shifts than predicted (Hewitt, 2000). Alternatively, the low genetic diversity among European species irrespective of latitude may be due to a lack of neutrality within mtDNA loci, as previously suggested (Bazin et al., 2006; Galtier et al., 2009). Under a neutral model (unaffected by selection) the genetic diversity of a population will be proportional to the total number of individuals within that population (Roberts, 2015), hence genetic diversity of neutral loci has been assumed to reflect demographic effects (Nabholz et al., 2008). However, while variability in population size can reach many orders of magnitude, genetic diversity has a much narrower distribution and shows no clear correlation with population size (Lewontin, 1974). Decoupled genetic diversity and population size, a discrepancy called “Lewontin’s paradox”, has been identified within many different phyla and attributed to an indirect influence of natural selection on these presumably “neutral” loci (Bazin et al., 2006; Ellegren & Galtier, 2016; Galtier et al., 2009; Roberts, 2015).

Since mtDNA is involved in essential metabolic functions, it has been considered that mitochondrial genes are less likely to be influenced by adaptive processes. Bazin et al. (2006) challenged this with a meta-analysis using around 3000 species, finding presence of adaptive selection in invertebrate mtDNA loci. Briefly, selection on functional loci may influence nearby neutral genes and sequences that are linked together through genetic hitch-hiking (Ellegren & Galtier, 2016). Given mtDNA non-recombining nature, the fixation of a beneficial mutation would lead to the loss of pre-existing polymorphisms in the entire mitochondrial genome, a process known as a selective sweep (Bazin et al., 2006; Ellegren & Galtier, 2016). As a result, mitochondrial genetic diversity would not exclusively reflect a population’s demographic history (e.g., bottleneck or founder event) but rather the time since the last selective sweep, given how each sweep pushes a diversity reset button (Bazin et al., 2006). Recurrent selective sweeps, known as genetic draft, would result in frequent

drops of the population's genetic diversity (Ellegren & Galtier, 2016), regardless of its effective population size or environmental stability. Genetic draft could potentially explain the similar genetic diversity estimates across European populations, since it seems effective in large and stable populations given their high amount of potential adaptive mutations, and thus compensates the lack of genetic drift (Bazin et al., 2006). Therefore, the genetic draft-drift trade-off acting on larger and smaller populations respectively could lead to similar low estimates of genetic diversity as those found for both southern European long-term stable populations and northern populations expected to be characterized by population bottlenecks and/or founder events (Nabholz et al., 2008). However, genetic draft-drift trade-off does not explain the high genetic diversity found in stable population in Panama compared to Europe, and consequently, the suggested non-neutrality of mtDNA is rejected as an explanatory factor for the global genetic diversity patterns found.

Overall, if neither biogeographic history, neutrality of mtDNA loci nor the genetic draft-drift trade-off explained the global pattern found, the influence of contemporary environmental conditions may alternatively be considered, as previously suggested at the community level by Múrria et al. (2017). *Star-like* haplotype network patterns commonly exemplify northern recolonization events through founder effects and also population expansions after a bottleneck event (Allcock & Strugnell, 2012; Avise, 2000). However, several *star-like* patterns were found in species located in southern latitudes, which are predicted to be long-term stable habitats, and therefore indicates rapid population expansion likely recovering from contemporary disturbances that caused population bottlenecks (Fig.5 A). Towards the north, some species showed *diffuse* patterns, which are usually recorded for long-term stable populations with high genetic diversity (Abellán et al., 2007; Allcock & Strugnell, 2012) (Fig.5 C). Therefore, contemporary environmental fluctuations associated to climate, local habitat heterogeneity and variability, seasonal changes in streamflow such as droughts or floods, may be overshadowing the patterns expected by the effects of the biogeographical history (Múrria et al., 2017), leading to low genetic diversity irrespective of latitude. More importantly, the generalizable low genetic diversity across freshwater European species reveals how under contemporary instability species persist at a regional scale although haplotypes perish locally.

4.2. The effect of dispersal ability on genetic structure

The mismatch found between European macroinvertebrate genetic structure and expected Pleistocene derived patterns elucidates the importance of lineage-specific dispersal ability. Nonetheless, hypothesis 2 (Fig. 3. B) is rejected given how, as suggested earlier, haplotype diversity and species age are decoupled. Comparative phylogeographers have primarily focused on testing concordance across co-distributed taxa in order to understand the role of extrinsic (i.e. geographic barriers, biogeographic history) processes on intraspecific patterns. However, discordance found across co-distributed taxa, while usually regarded as irrelevant or uninteresting, may suggest a strong contribution from intrinsic (i.e. species-specific life-history traits) and/or stochastic processes and provide insights on species dynamics (Gutiérrez-García & Vázquez-Domínguez, 2011; Papadopoulou & Knowles, 2016). The macrophylogeography approach used here revealed several incongruences within community-wide genetic patterns, which can be attributed to species-specific genetic trajectories and therefore each species' unique response to contemporary environmental conditions (Lorenzen et al., 2011), however general patterns are also apparent. In fact, the poorly dispersive mayflies (Sarremejane et al., 2017) did generally show higher genetic structure estimates than co-distributed beetles primarily due to their short adult life span. However, haplotype networks across mayfly species revealed a high variety of evolutionary histories. For instance, distinct potentially cryptic species within *B. rhodani* exemplified *star-like* (B4, B6) *parochial* (B2, B5, B7) and *diffuse* (B8, B9) haplotype network patterns (Fig.5 A-C; Appendix S4). In central and northern latitudes, the *diffuse* structure in B9 and the high haplotypic diversity in B6 may be attributed to abiotic/biotic factors favoring the pioneer post-glacial colonizer species (Hewitt, 2000; Taberlet et al., 1998). However, the complex relationships and high number of haplotypes in B8, in contrast to all other southern *B. rhodani,* indicates the presence of local stochasticity influencing evolutionary history in addition to weak dispersal ability.

Similarly, the widespread haplotype networks of aquatic beetles were consistent with the lineage's strong dispersal ability (Sarremejane et al., 2017), yet the haplotype networks showed *parochial* and *star-like* structures and therefore revealed different evolutionary histories likely due to distinct rates of post-glacial colonization (Svenning & Skov, 2007). Most Elmidae species that were found in different regions were highly structured and showed parochial haplotype networks caused by genetic differentiation across regions

(Allcock & Strugnell, 2012; Pálsson et al., 2014). Several factors may explain this pattern such as a slow colonization rate given how various areas of the entire species range appear to be colonized at different times. (Hewitt, 2000). Nonetheless within Elmidae, similarly to most Hydreanidae, species evidenced rapid recolonizations with *star-like* networks patterns with branches corresponding to different regions derived from one southern ancestral haplotype that colonized central and northern latitudes (e.g., E17, E12) (Avise, 2000). Despite common life-history traits, the varying intraspecific genetic diversity and structure found across aquatic beetles suggested different species-specific evolutionary histories likely driven by untested factors. Again, species persist after disturbances but haplotypes perish at each local disturbance, such as droughts in Mediterranean regions (Múrria et al., 2010).

5. Conclusions

While genetic structure varies according to dispersal ability as well as stochasticity of local disturbance, the misalignment between high genetic structure and low genetic diversity for European freshwater insects implies contemporary environmental instability in southern, central and northern latitudes. Contemporary demographic fluctuations are reducing genetic diversity expected from stable refugia, as seen in Neotropical highland streams, and leading to the loss of haplotypes in old pre-Pleistocene species. Conversely, the high genetic structure in Panama species is paralleled with high genetic diversity indicating consistently stable habitats. Beyond clear deterministic patterns, species with similarly old divergence age, biogeographic origins and history, morphology or dispersal ability revealed a wide variety of evolutionary histories. Our macrophylogeographic approach allows for a better appreciation of this inter-species variability at the population level and therefore comes much closer to reality (Page & Hughes, 2014). And reality is much less deterministic than comparative phylogeographic studies concluded, as seen in this study with the high variability in genetic structure across *B. rhodani* cryptic species complex as a product of “the baroque of nature” (Margalef, 1980). “Representation of the world, like the world itself, is the work of men; they describe it from their own point of view, which they confuse with the absolute truth” Simone de Beauvoir. To avoid falling into a deterministic pitfall, one must embrace the idiosyncrasy of each species, not as unwanted noise or deviations from a deterministic pattern but as strokes of paint portraying a fuller picture.

To improve the quality and resolution of this picture we propose incorporating next generation sequencing (NGS) using a restriction-site associated DNA (RADSeq) strategy as a cost-effective mechanism to increase the number of molecular markers analyzed in future macrophylogeographic investigations (Davey & Blaxter, 2010). We also advise the use of unlinked nuclear neutral markers as well as mtDNA for comparative purposes.

6. References

Abellán, P., Gómez-Zurita, J., Millán, A., Sánchez-Fernández, D., Velasco J., & Ribera, I. (2007). Conservation Genetics in Hypersaline Inland Waters: Mitochondrial Diversity and Phylogeography of an Endangered Iberian Beetle (Coleoptera: Hydraenidae). *Conservation Genetics* 8(1): 79–88.

Allcock, A.L., & Strugnell, J.M. (2012). Southern Ocean Diversity: New Paradigms from Molecular Ecology. *Trends in Ecology & Evolution* 27(9): 520–528.

Allendorf, F.W., Luikart, G., & Aitken, S.N. (2013). *Conservation and the genetics of populations*. 2nd edn. John Wiley & Sons, Hoboken, NJ

Alp, M., Keller, I., Westram, A. M., & Robinson, C. T. (2012). How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. *Freshwater Biology* 57(5): 969-981.

Antonovics, J. (2003). Toward Community Genomics?. *Ecology* 84(3): 598–601.

Avise, J.C. (1989). Gene Trees and Organismal Histories: A Phylogenetic Approach to Population Biology. *Evolution* 43(6): 1192–1208.

Avise, J.C. (2000). *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA.

Avise, J.C. (2004). *Molecular markers, natural history, and evolution*. 2nd edn. Sinauer Associates, Inc., Sunderland, MA.

Avise, J.C. (2009). Phylogeography: Retrospect and Prospect.*Journal of Biogeography* 36(1): 3–15.

Avise, J.C., Arnold, J., Martin Ball, R., Bermingham, E., Lamb, T., Neigel, J.E., … Saunders, N.C. (1987). Intraspecific Phylogeography: The Mitochondrial DNA Bridge between Population Genetics and Systematics. *Annual Review of Ecology and Systematics*18(1): 489–522.

Baselga, A. (2008). Determinants of Species Richness, Endemism and Turnover in European Longhorn Beetles. *Ecography* 31(2): 263-271.

Baselga, A., Gómez-Rodríguez, C., & Vogler, A.P. (2015). Multi-Hierarchical Macroecology at Species and Genetic Levels to Discern Neutral and Non-Neutral Processes: Multi-Hierarchical Macroecology. *Global Ecology and Biogeography*. 24(8): 873–882.

Baselga, A., Fujisawa, T., Crampton-Platt, A., Bergsten, J., Foster, P.G., Monaghan, M.T., & Vogler, A.P. (2013). Whole-Community DNA Barcoding Reveals a Spatio-Temporal Continuum of Biodiversity at Species and Genetic Levels. *Nature Communications* 4:1892.

Bazin, E., Glémin, S., Galtier, N. (2006). Population Size Does Not Influence Mitochondrial Genetic Diversity in Animals. *Science* 312(5773): 570–572.

Bertheau, C., Schuler, H., Arthofer, W., Avtzis, D.N., Mayer, F., Krumböck, S., Moodley, Y., & Stauffer, C. (2013). Divergent Evolutionary Histories of Two Sympatric Spruce Bark Beetle Species. *Molecular Ecology* 22(12): 3318–32.

Bohonak, A. J. (1999). Dispersal, Gene Flow, and Population Structure. *The Quarterly Review of Biology* 74(1): 21–45.

Brower, A. V. Z. (1994). Rapid Morphological Radiation and Convergence among Races of the Butterfly Heliconius Erato Inferred from Patterns of Mitochondrial DNA Evolution. *Proceedings of the National Academy of Sciences* 91(14): 6491–6495.

Carnaval, A. C., Hickerson, M. J., Haddad, C. F. B., Rodrigues, M. T., & Moritz, C. (2009). Stability Predicts Genetic Diversity in the Brazilian Atlantic Forest Hotspot. *Science* 323 (5915): 785–789.

Carnaval, A. C., Waltari, E., Rodrigues, M.T., Rosauer, D., VanDerWal, J., Damasceno, R., … Moritz C. (2014). Prediction of phylogeographic endemism in an environmentally complex biome. *Proceeding of the Royal Society B* 281(1792):20141461

Clement, M., Posada, D., & Crandall, K.A. (2000). TCS: A Computer Program to Estimate Gene Genealogies. *Molecular Ecology* 9(10):1657–1659.

Craft, K. J., Pauls, S. U., Darrow, K., Miller, S. E., Hebert, P. D. N., Helgen, L. E., … Weiblen, G. D. (2010). Population Genetics of Ecological Communities with DNA Barcodes: An Example from New Guinea Lepidoptera. *Proceedings of the National Academy of Sciences* 107(11): 5041–5046.

Crandall, K.A., & Templeton, A.R (1993). Empirical Tests of Some Predictions from Coalescent Theory with Applications to Intraspecific Phylogeny Reconstruction. *Genetics* 134(3): 959–969.

Davey, J.W., Blaxter, M.L. (2010). RADSeq: next-generation population genetic. *Briefings in functional genomics* 9(5):416-423.

Despres, L., Loriot, S., & Gaudeul, M. (2002). Geographic Pattern of Genetic Variation in the European Globeflower Trollius Europaeus L. (Ranunculaceae) Inferred from Amplified Fragment Length Polymorphism Markers. *Molecular Ecology* 11(11): 2337–2347.

Dexter, K. G., Terborgh, J. W., & Cunningham, C. W. (2012). Historical Effects on Beta Diversity and Community Assembly in Amazonian Trees. *Proceedings of the National Academy of Sciences* 109(20): 7787–7792.

Ellegren, H., & Galtier, N. (2016). Determinants of Genetic Diversity. *Nature Reviews Genetics* 17(7): 422–433.

Excoffier, L., Smouse, P.E., & Quattro, J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131(2): 479-491.

Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin (Version 3.0): An Integrated Software Package for Population Genetics Data Analysis. *Evolutionary Bioinformatics* 1: 47-50

Fischer, A.G. (1960). Latitudinal Variations in Organic Diversity. *Evolution* 14(1): 64–81.

Fujisawa, T., & Barraclough, T.G. (2013). Delimiting Species Using Single-Locus Data and the Generalized Mixed Yule Coalescent Approach: A Revised Method and Evaluation on Simulated Data Sets. *Systematic Biology* 62(5): 707–724.

Fujita, M.K., Leaché, A.D., Burbrink, F.T., McGuire, J.A., & Moritz, C. (2012). Coalescent-Based Species Delimitation in an Integrative Taxonomy. *Trends in Ecology & Evolution* 27(9): 480–488.

Galtier, N., Nabholz, B., GléMin, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a Marker of Molecular Diversity: A Reappraisal. *Molecular Ecology* 18(22): 4541–4550.

Gutiérrez-García, T.A., & Vázquez-Domínguez, E. (2011). Comparative Phylogeography: Designing Studies While Surviving the Process. *BioScience* 61(11): 857–868.

Hawkins, B. A., Diniz-Filho, J.A.F., Jaramillo, C.A., & Soeller, S.A. (2006). Post-Eocene Climate Change, Niche Conservatism, and the Latitudinal Diversity Gradient of New World Birds. *Journal of Biogeography* 33(5): 770–780.

Hebert, P.D.N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B* 270(1512): 313–321.

Hebert, P.D.N., Remigio, E.A., Colbourne, J.K., Taylor, D.J., & Wilson, C.C. (2002). Accelerated Molecular Evolution in Halophilic Crustaceans. *Evolution* 56(5): 909–26.

Hedrick, P.W. (2011). *Genetics of populations*. 4th edn. Jones and Bartlett Publishers, Sudbury, MA.

Hewitt, G.M. (1996). Some Genetic Consequences of Ice Ages, and their role in Divergence and Speciation. *Biological Journal of the Linnean Society* 58(3): 247-276

Hewitt, G.M. (2000). The Genetic Legacy of the Quaternary Ice Ages. *Nature* 405(6789): 907–13.

Hewitt, G.M. (2004). Genetic Consequences of Climatic Oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B* 359(1442):183–95.

Hickerson, M.J., Carstens, B.C., Cavender-Bares, J., Crandall, K.A., Graham, C.H., Johnson, J.B., Rissler, L., Victoriano, P.F., & Yoder, A.D. (2010). Phylogeography's Past, Present, and Future: 10 Years after Avise, 2000. *Molecular Phylogenetics and Evolution* 54(1): 291–301.

Holsinger, K.E., & Weir, B.S. (2009). Genetics in Geographically Structured Populations: Defining, Estimating and Interpreting FST. *Nature Reviews Genetics* 10(9): 639–50.

Hortal, J., Diniz-Filho, J.A.F., Bini, L.M., Rodríguez, M.A., Baselga, A., Nogués-Bravo, D., Rangel, T.F., Hawkins, B.A., & Lobo, J.M. (2011). Ice Age Climate, Evolutionary Constraints and Diversity Patterns of European Dung Beetles: Ice Age Determines European Scarab Diversity. *Ecology Letters* 14(8): 741–748.

Hughes, J.M. (2007). Constraints on Recovery: Using Molecular Methods to Study Connectivity of Aquatic Biota in Rivers and Streams. *Freshwater Biology* 52(4): 616–631.

Hughes, J.M., Schmidt, D.J., & Finn, D.S. (2009). Genes in Streams: Using DNA to Understand the Movement of Freshwater Fauna and Their Riverine Habitat. *BioScience* 59 (7): 573–583.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press, New York, US.

Knowles, L. L. (2009). Statistical Phylogeography. *Annual Review of Ecology, Evolution, and Systematics* 40 (1): 593–612.

Lacy, R. (1987). Loss of Genetic Diversity from Managed Populations: Interacting Effects of Drift, Mutation, Immigration, Selection and Population Subdivision. *Conservation Biology* 1 (2):143-158

Lanfear, R., Frandsen, P.B., Wright, A.M., 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(3): 772–773.

Lehrian, S., Pauls, S. U., & Haase, P. (2009). Contrasting patterns of population structure in the montane caddisflies Hydropsyche tenuis and Drusus discolor in the Central European highlands. *Freshwater Biology* 54(2): 283-295.

Lewontin, R.C. (1974). *The Genetic Basis of Evolutionary Change*. Columbia University Press.

Lorenzen, E. D., Nogués-Bravo, D., Orlando, L., Weinstock, J., Binladen, J., Marske, K. A., ... Ho, S. Y. (2011). Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* 479(7373): 359-364.

Lowe, A., Harris, S., & Ashton, P. (2004). *Ecological Genetics: Design, Analysis, and Application*. Blackwell Publishing, Oxford, UK.

Maddison, W. P., & Maddison, D.R. (2018). *Mesquite: a modular system for evolutionary analysis.* Version 3.40. Available at: http://www.mesquiteproject.org

Margalef, R. (1980). *La biosfera: entre la termodinámica y el juego*. Omega, Barcelona, España

Marske, K.A., Rahbek, C., & Nogués-Bravo, D. (2013). Phylogeography: Spanning the Ecology-Evolution Continuum. *Ecography* 36(11):1169–1181.

Martin, P.R., & McKay, J.K. (2004). Latitudinal variation in genetic divergence of populations and the potential for future speciation. *Evolution* 58(5):938-945

Miller, M.A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees. *Proceeding of the Gateway Computing Environments Workshop (GCE)*. New Orleans, LA, 1-8

Miller, M.P., Blinn, D.W., & Keim, P. (2002). Correlations between Observed Dispersal Capabilities and Patterns of Genetic Differentiation in Populations of Four Aquatic Insect Species from the Arizona White Mountains, U.S.A. *Freshwater Biology* 47(9): 1660–1673.

Múrria, C., & Hughes, J. M. (2008). Cyclic Habitat Displacements during Pleistocene Glaciations Have Induced Independent Evolution of Tasimia Palpata Populations (Trichoptera: Tasimiidae) in Isolated Subtropical Rain Forest Patches. *Journal of Biogeography* 35(9):1727–1737.

Múrria, C., Bonada, N., Ribera, C & Prat, N. (2010). Homage to the Virgin of Ecology or why an aquatic insect unadapted to desiccation may maintain populations in very small temporary Mediterranean streams?. *Hydrobiologia* 653(1): 179-190.

Múrria, C, Rugenski, A.T., Whiles, M.R. & Vogler, A.P. (2015). Long-Term Isolation and Endemicity of Neotropical Aquatic Insects Limit the Community Responses to Recent Amphibian Decline. *Diversity and Distributions* 21(8): 938–949.

Múrria, C., Bonada, N., Arnedo, M.A., Prat, N., & Vogler, A.P. (2013). Higher β-and γ-diversity at species and genetic levels in headwaters than in mid-order streams in H ydropsyche (T richoptera). *Freshwater Biology* 58(11): 2226-2236.

Múrria, C., Bonada, N., Arnedo, M.A., Zamora-Muñoz, C., Prat, N., & Vogler, A.P. (2012). Phylogenetic and Ecological Structure of Mediterranean Caddisfly Communities at Various Spatio-Temporal Scales: Evolution of Aquatic Insect Assemblages. *Journal of Biogeography* 39(9):1621–1632.

Múrria, C., Bonada, N., Vellend, M., Zamora-Muñoz, C., Alba-Tercedor, J., Elisa Sainz-Cantero, C., … Vogler, A.P. (2017). Local Environment Rather than Past Climate Determines Community Composition of Mountain Stream Macroinvertebrates across Europe. *Molecular Ecology* 26(21): 6085–6099.

Nabholz, B., Mauffrey, J.F., Bazin, E., Galtier, N., & Glémin, S. (2008). Determination of Mitochondrial Genetic Diversity in Mammals. *Genetics* 178(1): 351–361.

Nei, M. (1987). *Molecular evolutionary genetics*. Columbia University Press, New York, US.

Page, T.J., & Hughes, J.M. (2014). Contrasting Insights Provided by Single and Multispecies Data in a Regional Comparative Phylogeographic Study: Australian Freshwater Phylogeography. *Biological Journal of the Linnean Society* 111(3): 554–69.

Pálsson, S., Magnúsdóttir, H., Reynisdóttir, S., Jónsson, Z.O., & Örnólfsdóttir, E.B (2014). Divergence and Molecular Variation in Common Whelk Buccinum Undatum (Gastropoda: Buccinidae) in Iceland: A Trans-Atlantic Comparison. *Biological Journal of the Linnean Society* 111(1): 145–159.

Papadopoulou, A., & Knowles, L.L. (2016). Toward a Paradigm Shift in Comparative Phylogeography Driven by Trait-Based Hypotheses. *Proceedings of the National Academy of Sciences* 113(29): 8018–8024.

Papadopoulou, A., Anastasiou, I. & Vogler, A.P. (2010) Revisiting the insect mitochondrial molecular clock: The mid-Aegean trench calibration. *Molecular Biology and Evolution*, 27(7):1659-1672

Papadopoulou, A., Anastasiou, I., Keskin, B., & Vogler, A.P. (2009). Comparative Phylogeography of Tenebrionid Beetles in the Aegean Archipelago: The Effect of Dispersal Ability and Habitat Preference. *Molecular Ecology* 18(11): 2503–2517.

Pauls, S. U., Alp, M., Bálint, M., Bernabò, P., Čiampor Jr, F., Čiamporová-Zaťovičová, Z., ... Paz-Vinas, I. (2014). Integrating molecular tools into freshwater ecology: developments and opportunities. *Freshwater Biology* 59(8):1559-1576.

Pianka, E.R. (1966). Latitudinal Gradients in Species Diversity: A Review of Concepts. *The American Naturalist* 100(910): 33–46.

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., ... Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* 55(4): 595-609

Posada, D., & Crandall, K.A. (2001). Intraspecific Gene Genealogies: Trees Grafting into Networks. *Trends in Ecology & Evolution* 16(1): 37–45.

Qu, Y., Ericson, P.G.P., Quan, Q., Song, G., Zhang, R., Gao, B., & Lei, F. (2014). Long-Term Isolation and Stability Explain High Genetic Diversity in the Eastern Himalaya. *Molecular Ecology* 23(3): 705–720.

R Core Team (2016). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. Vienna, Austria. Available at: https://www.R-project.org/.

Rambaut, A. (2014). FIGTREE Version 1.4.2. Available at: http:// tree.bio.ed.ac.uk/software/figtree/

Ribera, I., Foster, G. N., & Vogler, A. P. (2003). Does habitat use explain large scale species richness patterns of aquatic beetles in Europe? *Ecography* 26(2): 145–152.

Roberts, R.G. (2015). Lewontin's Paradox Resolved? In Larger Populations, Stronger Selection Erases More Diversity. *PLOS Biology* 13(4): e1002113.

Ronquist, F., & Huelsenbeck, J.P. (2003). Summary: MrBayes 3 Performs Bayesian Phylogenetic. *Bioinformatics* 19(12): 1572–1574.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sánchez-Gracia, A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Datasets. *Molecular Biology and Evolution* 34(12): 3299-3302

Rull, V. (2005). Biotic Diversification in the Guayana Highlands: A Proposal: Guayana Highlands Diversification. *Journal of Biogeography* 32(6): 921–927.

Rull, V. (2006). Quaternary Speciation in the Neotropics". Molecular Ecology 15(13):4257–4259.

Sarremejane, R., Mykrä, H., Bonada, N., Aroviita, J., & Muotka, T. (2017). Habitat Connectivity and Dispersal Ability Drive the Assembly Mechanisms of Macroinvertebrate Communities in River Networks. *Freshwater Biology* 62(6): 1073–1082.

Stamatakis, A. (2014). RAxML Version 8: A Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* 30(9): 1312–1313.

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.L., & Rambaut, A. (2018). Bayesian Phylogenetic and Phylodynamic Data Integration Using BEAST 1.10. *Virus Evolution* 4(1): vey016

Svenning, J.C., & Skov, F. (2007). Could the Tree Diversity Pattern in Europe Be Generated by Postglacial Dispersal Limitation?. *Ecology Letters* 10 (6): 453–460.

Svenning, J.C., Fløjgaard, C., & Baselga, A. (2011). Climate, History and Neutrality as Drivers of Mammal Beta Diversity in Europe: Insights from Multiscale Deconstruction: Deconstructing Mammal Beta Diversity. *Journal of Animal Ecology* 80(2): 393–402.

Taberlet, P., Fumagalli, L, Wust-Saucy, A.G., & Cosson, J.F. (1998). Comparative Phylogeography and Postglacial Colonization Routes in Europe. *Molecular Ecology* 7(4): 453–464.

Thomas, J. A., Welch, J. J., Woolfit, M. & Bromham, L. (2006). There Is No Universal Molecular Clock for Invertebrates, but Rate Variation Does Not Scale with Body Size. *Proceedings of the National Academy of Sciences* 103(19): 7366–7371.

Vamosi, S. M., Heard, S. B., Vamosi, J. C., & Webb, C. O. (2009). Emerging Patterns in the Comparative Analysis of Phylogenetic Community Structure. *Molecular Ecology* 18(4): 572–592.

Vangestel, C., Mergeay, J., Dawson, D.A., Callens, T., Vandomme, V., & Lens, L. (2012). Genetic Diversity and Population Structure in Contemporary House Sparrow Populations along an Urbanization Gradient. *Heredity* 109(3): 163–172.

Vellend, M. (2003). Island Biogeography of Genes and Species. *The American Naturalist* 162(3): 358–365.

Vellend, M. (2010). Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology* 85(2): 183–206.

Vellend, M. (2016*). The Theory of Ecological Communities*. Princeton University Press, Princeton, US.

Vellend, M., & Geber, M.A. (2005). Connections between Species Diversity and Genetic Diversity: Species Diversity and Genetic Diversity. *Ecology Letters* 8(7): 767–781.

Vellend, M., Lajoie, G., Bourret, A., Múrria, C., Kembel, S.W., & Garant, D. (2014). Drawing Ecological Inferences from Coincident Patterns of Population- and Community-Level Biodiversity. *Molecular Ecology* 23(12): 2890–2901.

Young, A., Boyle, T., & Brown, T. (1996). The Population Genetic Consequence of Habitat Fragmentation for Plants. *Trends in Ecology & Evolution* 11(10):413-418

Zeller, M., Reusch, T.B.H., & Lampert, W. (2006). A Comparative Population Genetic Study on Calanoid Freshwater Copepods: Investigation of Isolation-by-Distance in Two Eudiaptomus Species with a Different Potential for Dispersal. *Limnology and Oceanography* 51(1): 117–124.

Zickovich, J.M., & Bohonak, A.J. (2007). Dispersal Ability and Genetic Structure in Aquatic Invertebrates: A Comparative Study in Southern California Streams and Reservoirs. *Freshwater Biology* 52(10): 1982–1996.

Zink, R.M., &. Barrowclough, G.F. (2008). Mitochondrial DNA under Siege in Avian Phylogeography. *Molecular Ecology* 17(9): 2107–2121.

7. Supplementary Material

Appendix S1: Genetic diversity π and structure Φ_{ST} of freshwater macroinvertebrates fulfilling the 5x3 (Europe) and 4x2 (Panama) criteria. Linnean and molecular (GMYC) species labels are included. For species found in various European regions, total and regional genetic variability estimates were calculated and specified in GMYC labels. Location categories correspond to regions of presence (see Fig. 2). Divergence time was obtained from nodes corresponding to the closest common ancestor and represented only when posterior probabilities >0.95(dashed lines represent <0.95).

Europe

Order	Family/Genera	Linnean label	GMYC label	Φ_{ST}	π	Location	Divergence (Mya)
Ephemeroptera	*Baetis*	*B.rhodani*	B02	0.62	1.71E-02	P	4.31
		B.rhodani	B04	0.51	3.25E-03	P	4.31
		B.rhodani	B05	0.02	3.34E-03	R	6.33
		B.rhodani	B06_total	0.49	5.46E-03	C; J; S	2.88
		B.rhodani	B06_J	0.01	4.28E-03	J	
		B.rhodani	B06_S	0.28	4.83E-03	S	
		B.rhodani	B07_total	0.13	5.18E-03	B; R	0.84
		B.rhodani	B07_R	0.21	5.14E-03	R	
		B.rhodani	B08_total	0.26	1.56E-02	B; P; R; S	0.84
		B.rhodani	B08_B	0.20	1.55E-02	B	
		B.rhodani	B09_total	0.33	9.34E-03	C; J	8.78
		B.rhodani	B09_C	0.03	8.64E-03	C	
		B.alpinus	B20	0.65	4.13E-03	J	--
		B.alpinus	B21	0.33	3.04E-03	P	2.3
		B.alpinus	B22	0.30	6.81E-03	C	1.14
		B.alpinus	B24	0.52	4.61E-03	J	0.71
Coleoptera	Elmidae	*R.subviolaceus*	E06	0.28	3.00E-03	J	0.66
		R.subviolaceus	E07_total	0.64	6.84E-03	C; P	0.66
		R.subviolaceus	E07_C	0.01	4.91E-03	C	
		E.angustatus	E08_total	0.88	9.81E-03	C; P; J	3.54
		E.angustatus	E08_P	0.02	3.26E-03	P	
		E.parallelepipedus	E09_total	0.15	2.51E-03	B; J	3.54
		L.volckmari	E12_total	0.64	2.76E-03	B; J	2.07
		L.perrisi	E14_total	0.98	9.94E-03	C; P; J	5.9
		L.perrisi	E14_P	1.00	7.38E-03	P	
		L.perrisi	E14_J	0.02	2.43E-03	J	
		L.perrisi	E14_C	0.01	2.43E-03	C	
		E.aenea	E17_total	0.75	1.11E-02	S; P; J; C	3.01
		E.aenea	E17_S	0.30	4.73E-03	S	
		E.aenea	E17_P	0.03	6.96E-03	P	
		E.aenea	E17_J	0.13	4.03E-03	J	

Order	Family/Genera	Linnean label	GMYC label	Φ_{ST}	π	Location	Divergence time
		E.aenea	E17_C	0.03	2.83E-03	C	
		E.cf.aenea	E19	0.02	4.17E-03	B	1.1
	Hydraenidae	*H.capta*	H03	0.04	5.18E-03	B	1.91
		H.stussineri/rigua/rufipennis	H05_total	0.25	9.01E-03	B; R	2.19
		H.stussineri/rigua/rufipennis	H05_R	0.03	8.77E-03	R	
		H.truncata	H10	0.34	3.67E-03	C; J; P	2.57
		H.lapidicola	H13	0.10	4.43E-03	J	2.21
		H.exasperata	H18	0.30	3.16E-03	B	
		H.gracilis	H19_total	0.28	3.76E-03	S; J; C	0.42
		H.gracilis	H19_S	0.23	3.99E-03	S	
		H.gracilis	H19_J	0.00	2.92E-03	J	
		H.gracilis	H19_C	0.00	3.09E-03	C	
			Panama				
Ephemeroptera	*Leptohypes*	*Haplohyphes*	M34; M35	0.75	4.60E-02	CA; CP	--
	Leptophlebiidae	*Farrodes*	M12; M13; M79; M80	0.81	3.19E-02	CP; GU; AL	4.11
		Farrodes	M11	0.63	1.66E-02	CA; CU	4.11
		Hydrosmilidon	M16	0.46	6.24E-03	FR; CA	--
	Baetidae	*Moribaetis*	M24; M25; M26	0.97	4.20E-02	GU; AL; CO	2.65
		Baetodes	M18; M19	0.88	4.13E-02	GU; FR; CA	--
		Baetodes	M17	0.06	3.04E-03	AL; CO	7.81
Coleoptera	Ptilodactylidae	*Anchytarsus*	B6	0.36	6.99E-03	CP; CO; GU	2.22

Appendix S2: Linear regression of π and Φ_{ST} using European and Panamanian species. Total (**A**) and regional (**B**) estimates were analyzed independently.

A)

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.62657	0.24802	6.558	2.53e-07 ***
log(π)	0.23691	0.04918	4.817	3.63e-05 ***

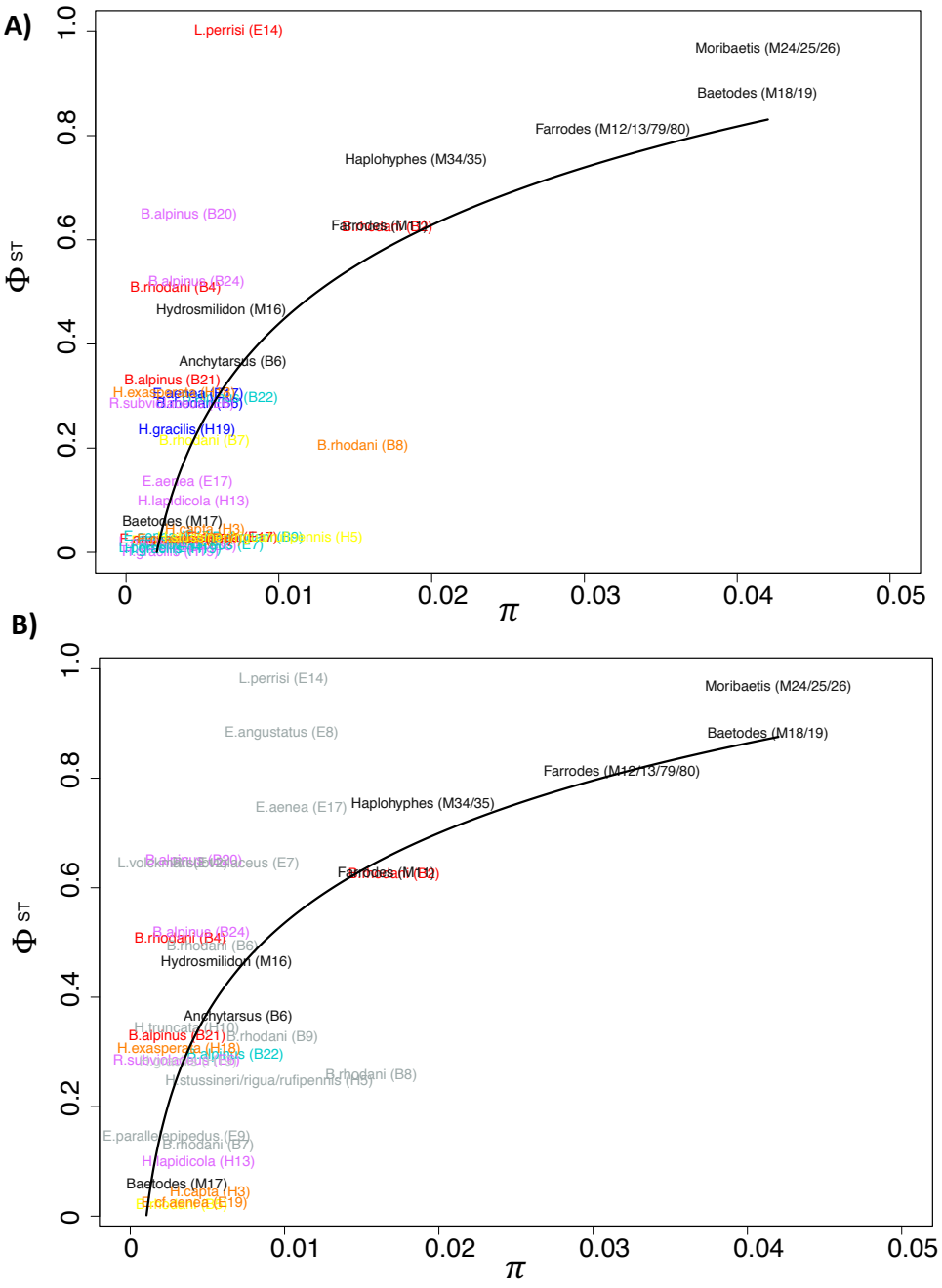
Residual standard error: 0.2211 on 31 degrees of freedom
Multiple R-squared: 0.4281, Adjusted R-squared: 0.4096
F-statistic: 23.2 on 1 and 31 DF, p-value: 3.63e-05

B)

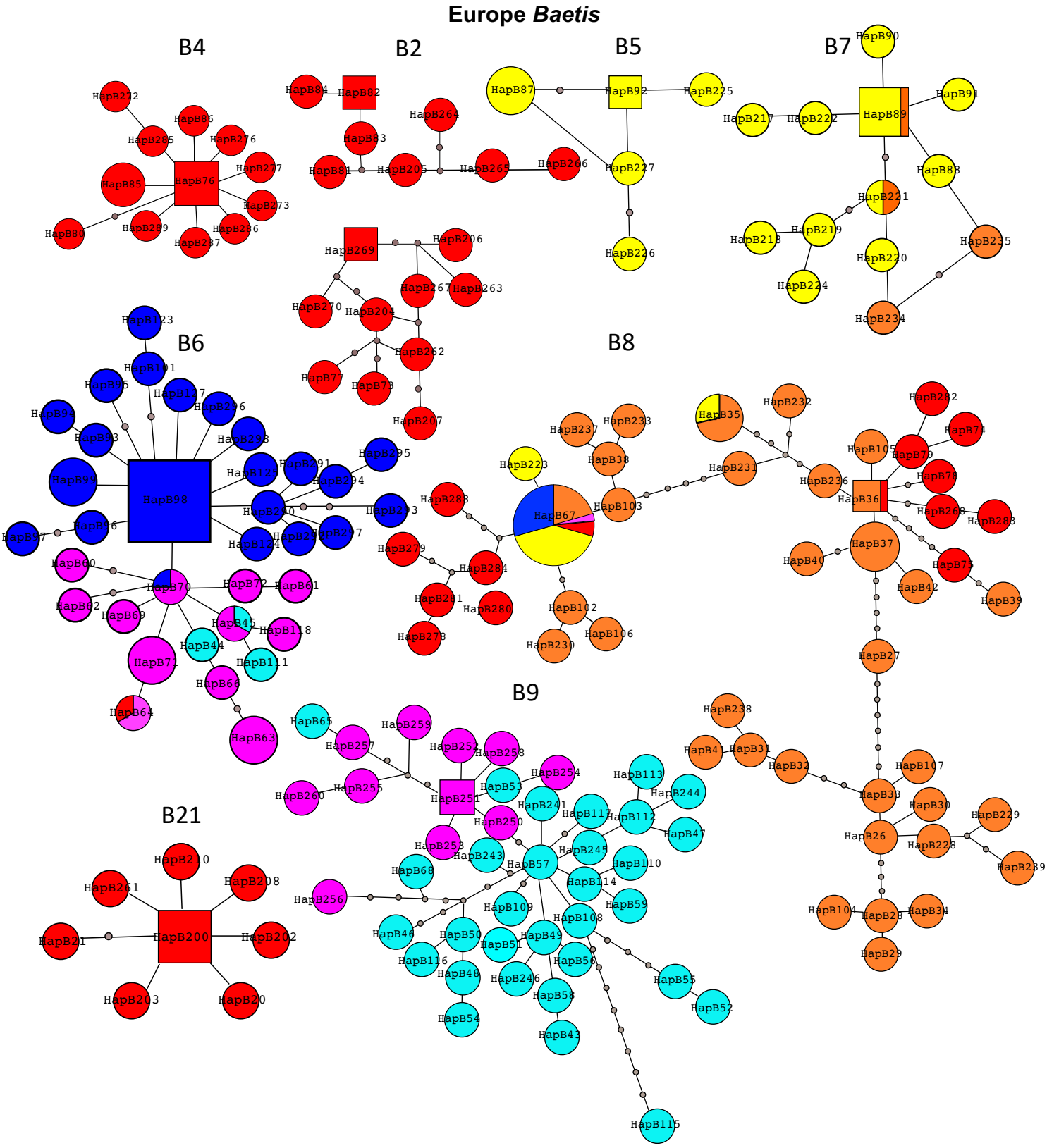
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.69691	0.24567	6.907	4.35e-08 ***
log(π)	0.27318	0.04732	5.773	1.40e-06 ***

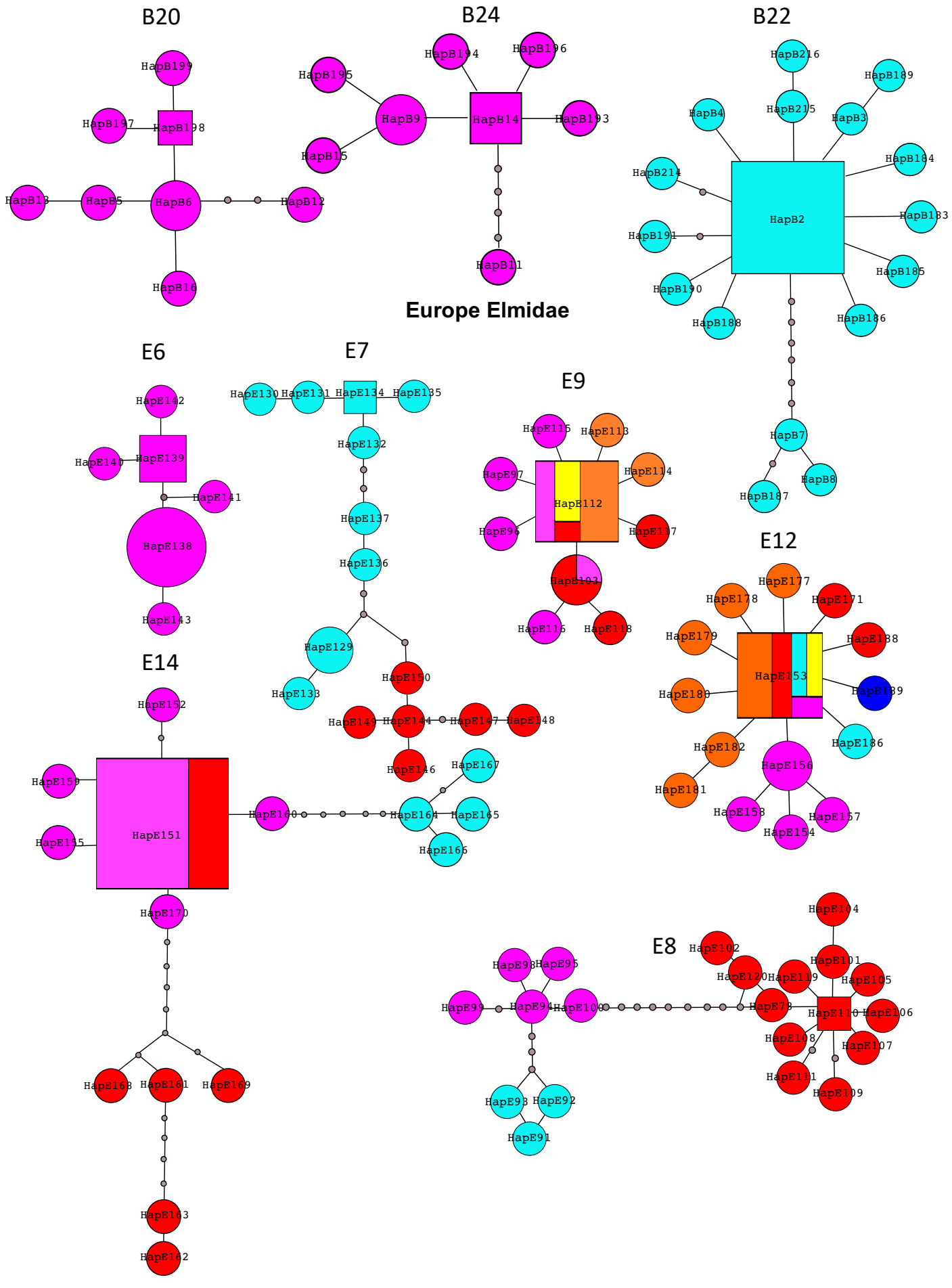
Residual standard error: 0.2228 on 36 degrees of freedom
Multiple R-squared: 0.4807, Adjusted R-squared: 0.4663
F-statistic: 33.32 on 1 and 36 DF, p-value: 1.399e-06

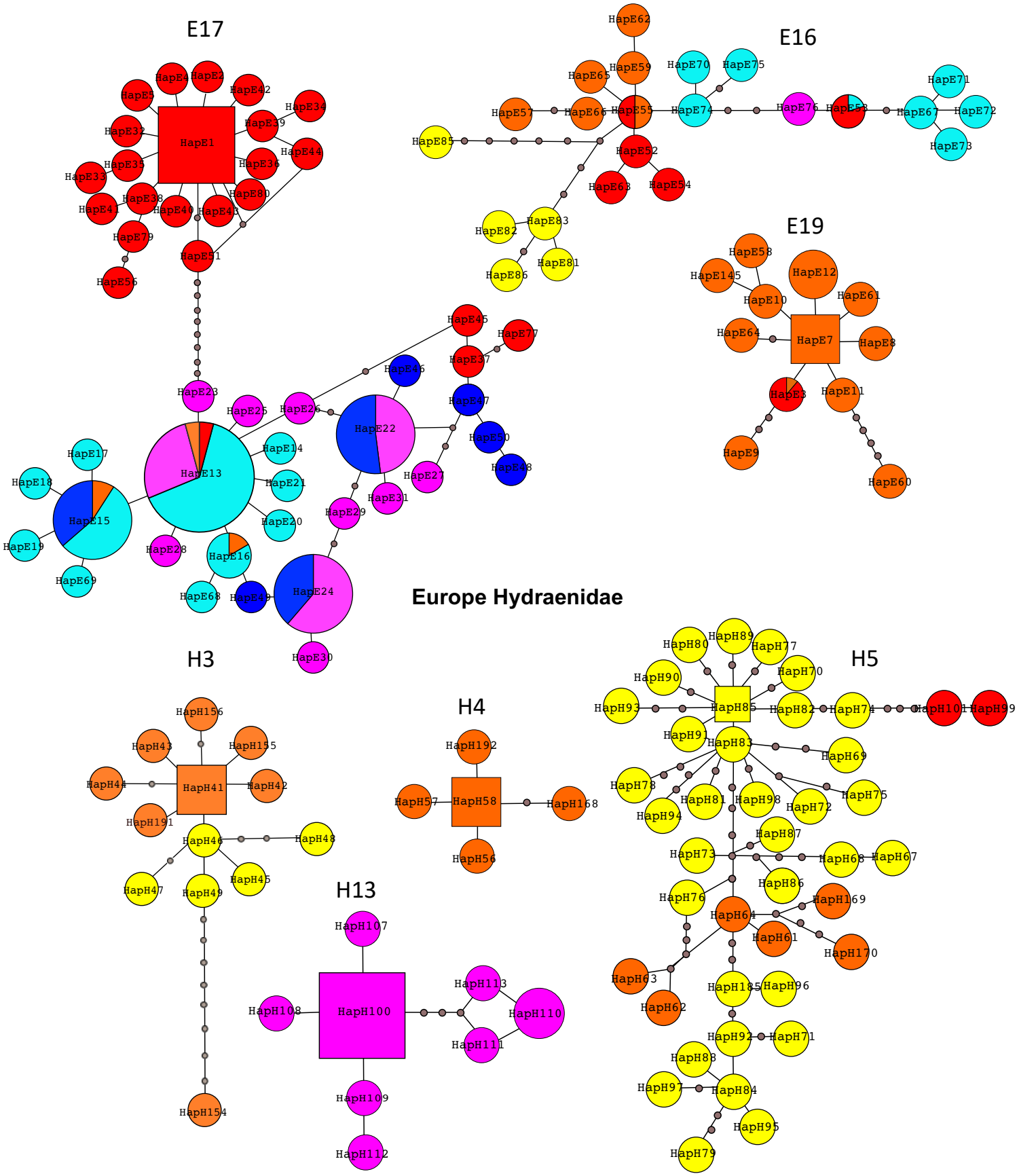
Appendix S3. Correlations between π and Φ_{ST} of species from Europe and Panama. Colors indicate regions from northern Morocco to northern Europe where species were present (see Fig. 2). Black indicates species from Panama irrespective of region of presence. **A** indicates regional estimate of π and Φ_{ST} while **B** indicates total estimates Φ_{ST} and π. When species are found in more than one region this is represented with grey.

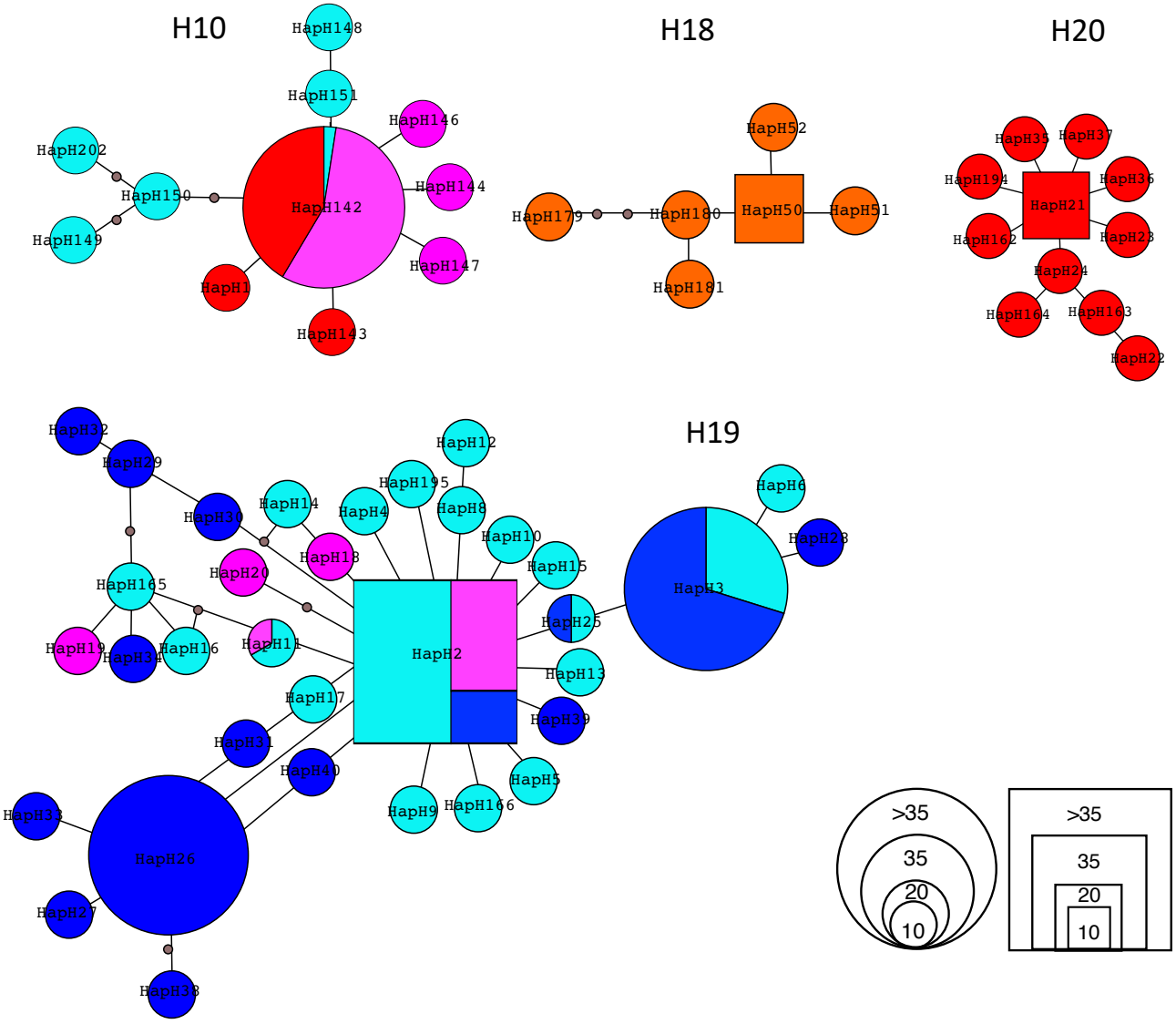


Appendix S4: European *Baetis*, Elmidae and Hydraenidae haplotype networks with molecular species labels. Colors corresponds to regions of presence (see Fig. 2). Squares represent haplotypes most likely to be ancestral (highest outgroup probability correlated to haplotype age) as determined by TCS. Size of circles and squares are proportional to haplotype frequency. Small grey dots represent unsampled or missing intermediate mutations.

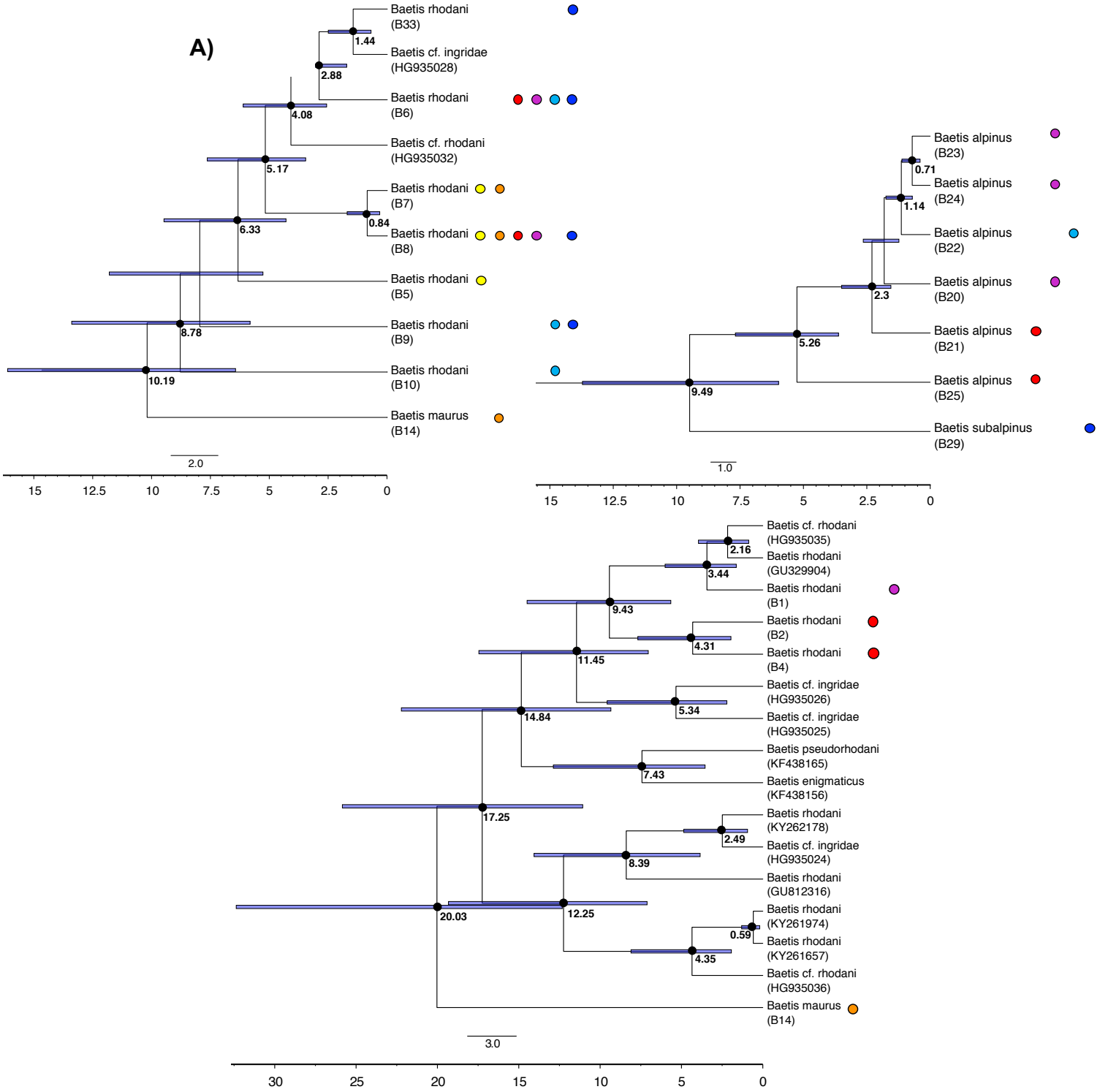



B20
HapB199
HapB197
HapB198
HapB13
HapB5
HapB6
HapB12
HapB16
B24
HapB194
HapB196
HapB195
HapB9
HapB14
HapB193
HapB15
HapB11
B22
HapB216
HapB189
HapB4
HapB215
HapB3
HapB184
HapB214
HapB2
HapB183
HapB191
HapB185
HapB190
HapB188
HapB186
HapB7
HapB8
HapB187
Europe Elmidae
E6
HapE142
HapE140
HapE139
HapE141
HapE138
HapE143
E7
HapE130
HapE131
HapE134
HapE135
HapE132
HapE137
HapE136
HapE129
HapE150
HapE133
HapE149
HapE144
HapE147
HapE148
HapE146
HapE167
HapE164
HapE165
HapE166
E9
HapE115
HapE113
HapE97
HapE114
HapE112
HapE96
HapE117
HapE103
HapE116
HapE118
E12
HapE177
HapE178
HapE171
HapE179
HapE188
HapE153
HapE189
HapE180
HapE182
HapE156
HapE186
HapE181
HapE158
HapE157
HapE154
E14
HapE152
HapE159
HapE151
HapE155
HapE160
HapE170
HapE168
HapE161
HapE169
HapE163
HapE162
E8
HapE104
HapE102
HapE101
HapE120
HapE119
HapE105
HapE98
HapE95
HapE99
HapE94
HapE100
HapE78
HapE110
HapE106
HapE108
HapE107
HapE111
HapE109
HapE93
HapE92
HapE91

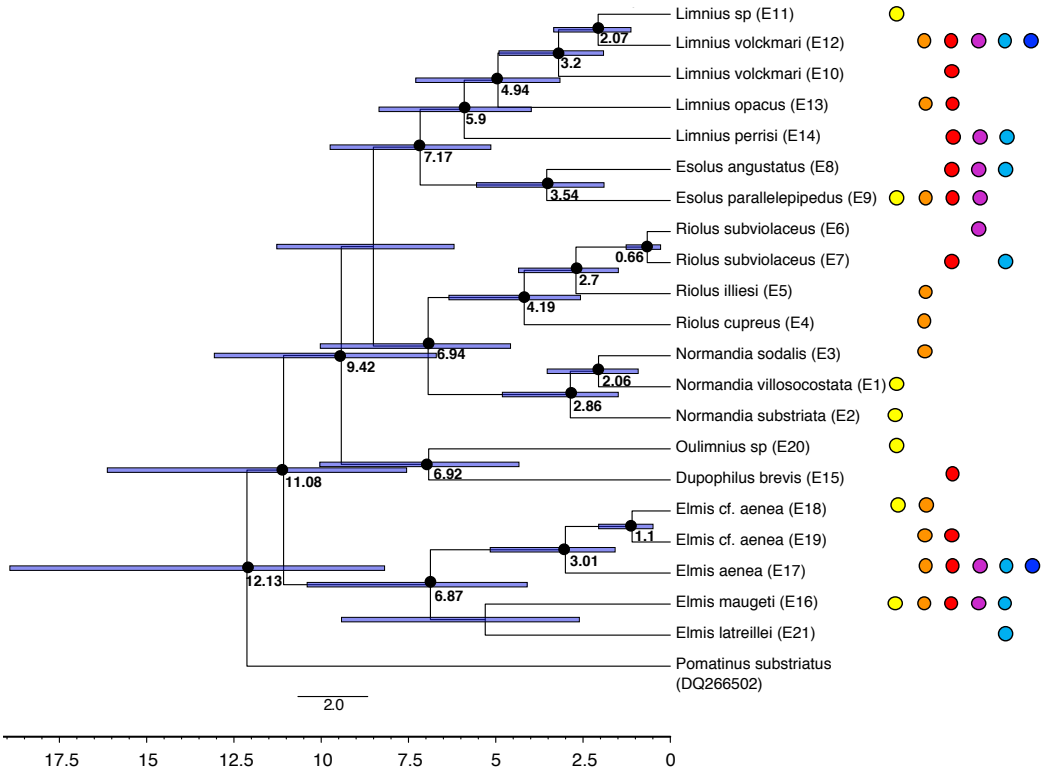

E17
HapE1
HapE2
HapE4
HapE5
HapE42
HapE34
HapE39
HapE32
HapE44
HapE35
HapE36
HapE33
HapE80
HapE38
HapE41
HapE40
HapE43
HapE79
HapE91
HapE56
HapE23
E16
HapE62
HapE59
HapE65
HapE70
HapE75
HapE71
HapE57
HapE66
HapE55
HapE74
HapE76
HapE53
HapE67
HapE72
HapE85
HapE52
HapE73
HapE63
HapE54
HapE83
HapE82
HapE81
HapE86
E19
HapE58
HapE12
HapE145
HapE10
HapE61
HapE64
HapE7
HapE8
HapE3
HapE11
HapE9
HapE60
HapE45
HapE77
HapE37
HapE46
HapE47
HapE25
HapE26
HapE22
HapE50
HapE14
HapE48
HapE27
HapE17
HapE13
HapE21
HapE31
HapE18
HapE29
HapE15
HapE20
HapE19
HapE28
HapE16
HapE69
HapE68
HapE49
HapE24
HapE30
Europe Hydraenidae
H3
HapH156
HapH43
HapH155
HapH44
HapH41
HapH42
HapH191
HapH46
HapH48
HapH47
HapH49
HapH45
HapH154
H4
HapH192
HapH57
HapH58
HapH168
HapH56
H5
HapH80
HapH89
HapH77
HapH90
HapH70
HapH93
HapH85
HapH82
HapH74
HapH101
HapH99
HapH91
HapH83
HapH69
HapH78
HapH81
HapH98
HapH75
HapH94
HapH72
HapH87
HapH73
HapH68
HapH67
HapH86
HapH76
HapH169
HapH64
HapH61
HapH170
HapH63
HapH185
HapH96
HapH62
HapH92
HapH71
HapH88
HapH97
HapH84
HapH95
HapH79
H13
HapH107
HapH113
HapH108
HapH100
HapH110
HapH111
HapH109
HapH112


H10
HapH148
HapH151
HapH146
HapH202
HapH144
HapH150
HapH142
HapH149
HapH147
HapH1
HapH143
H18
HapH52
HapH179
HapH180
HapH50
HapH51
HapH181
H20
HapH35
HapH37
HapH194
HapH36
HapH21
HapH162
HapH23
HapH24
HapH164
HapH163
HapH22
H19
HapH32
HapH29
HapH12
HapH195
HapH30
HapH14
HapH4
HapH8
HapH10
HapH18
HapH15
HapH20
HapH165
HapH25
HapH19
HapH34
HapH16
HapH11
HapH2
HapH13
HapH17
HapH39
HapH31
HapH40
HapH5
HapH9
HapH166
HapH6
HapH28
HapH3
HapH33
HapH26
HapH27
HapH38
>35
35
20
10
>35
35
20
10

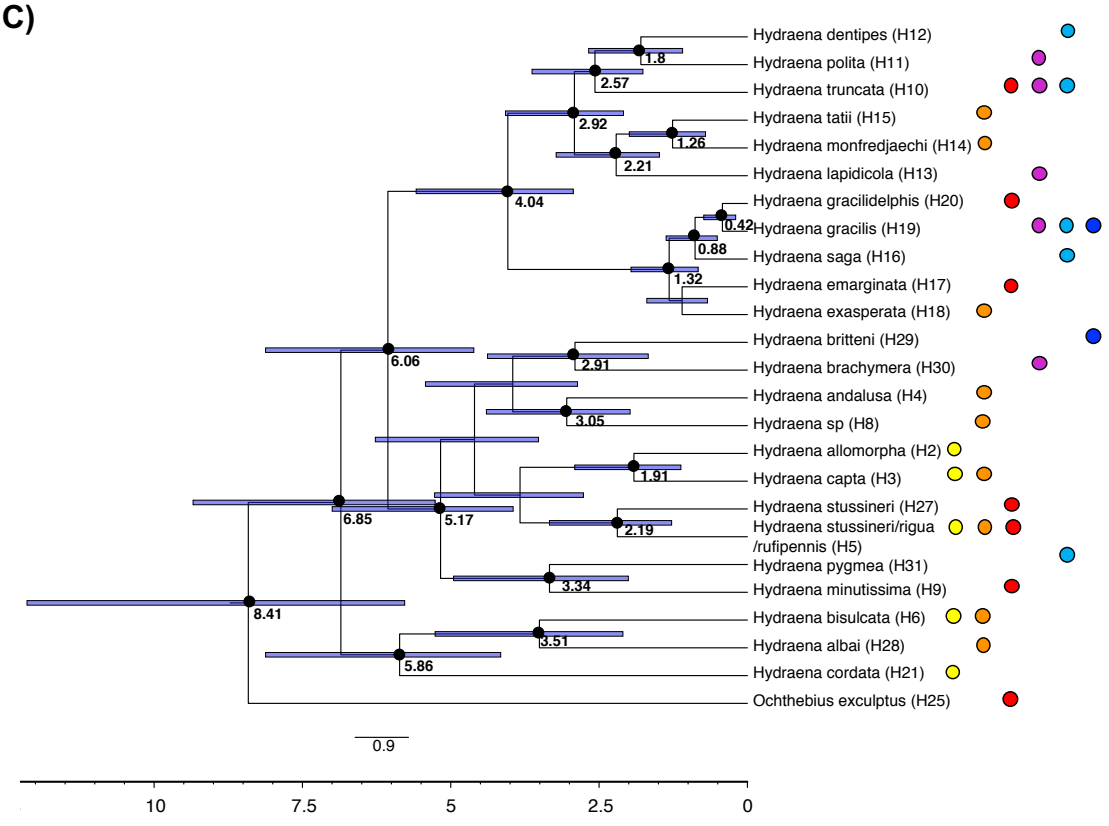
Appendix S5: *Cox1* time calibrated phylogenetic trees of the European *Baetis (A),* Elmidae (B) and Hydraenidae (C) and Panamanian Haplohyphes (D), Farrodes (E), Moribaetis (F), Baetodes (G) and *Anchytarsus* (H). Codes correspond to the molecular species labeled in Múrria et al., 2015, 2017. Colored circles signify region of presence (see Fig.2). Scales and numbers left of nodes indicate divergence time. Bars on nodes indicate the 95% highest posterior density (HPD) confidence intervals of the divergence time while black dots reveal posterior probabilities >0.95.

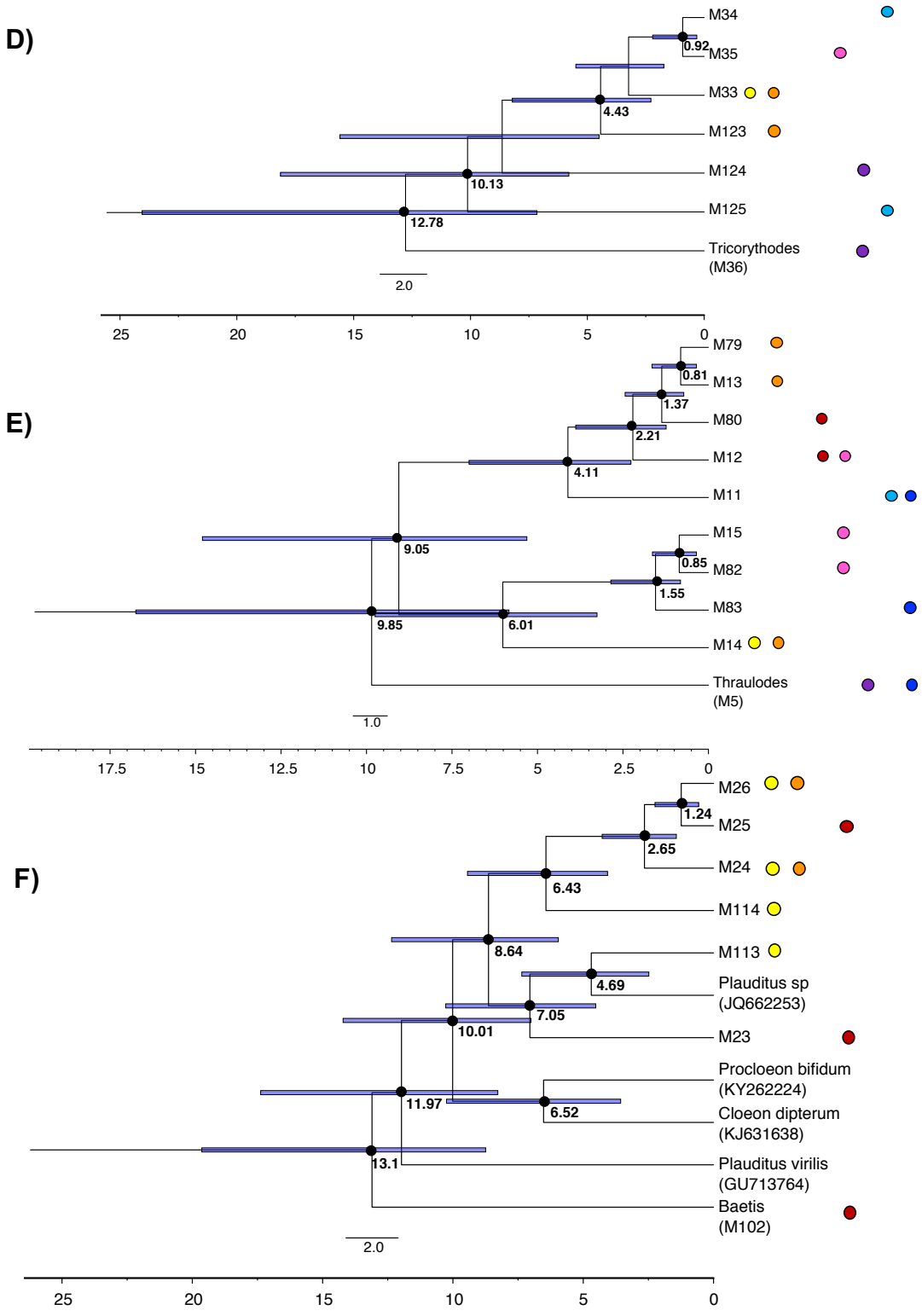


B)

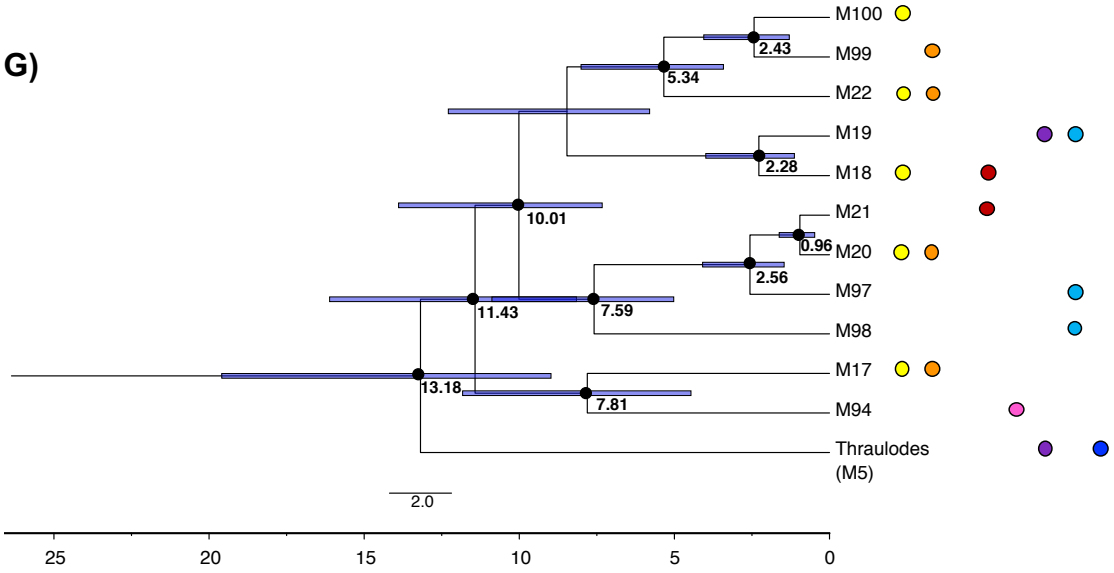

Limnius sp (E11)
Limnius volckmari (E12)
Limnius volckmari (E10)
Limnius opacus (E13)
Limnius perrisi (E14)
Esolus angustatus (E8)
Esolus parallelepipedus (E9)
Riolus subviolaceus (E6)
Riolus subviolaceus (E7)
Riolus illiesi (E5)
Riolus cupreus (E4)
Normandia sodalis (E3)
Normandia villosocostata (E1)
Normandia substriata (E2)
Oulimnius sp (E20)
Dupophilus brevis (E15)
Elmis cf. aenea (E18)
Elmis cf. aenea (E19)
Elmis aenea (E17)
Elmis maugeti (E16)
Elmis latreillei (E21)
Pomatinus substriatus (DQ266502)
2.07
3.2
4.94
5.9
7.17
3.54
0.66
2.7
4.19
6.94
9.42
2.06
2.86
6.92
11.08
1.1
3.01
12.13
6.87
2.0
17.5
15
12.5
10
7.5
5
2.5
0

C)

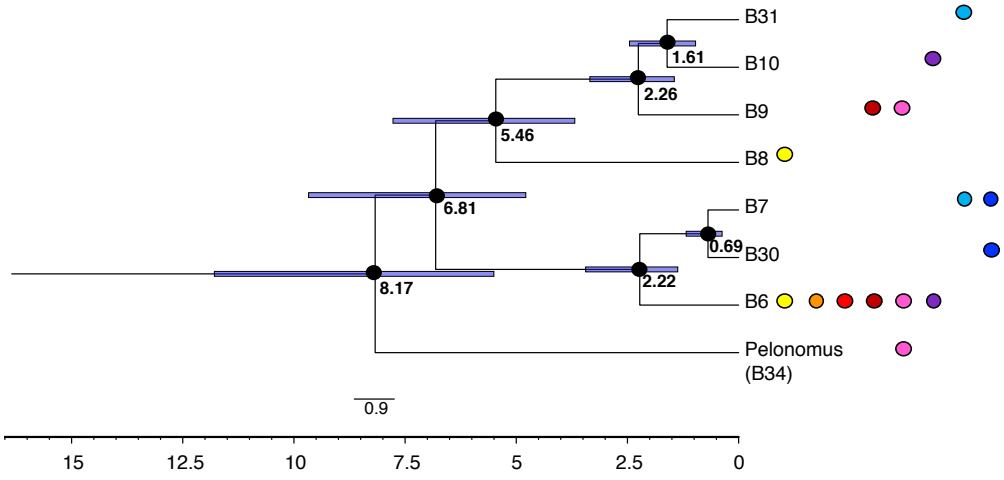

Hydraena dentipes (H12)
Hydraena polita (H11)
Hydraena truncata (H10)
Hydraena tatii (H15)
Hydraena monfredjaechi (H14)
Hydraena lapidicola (H13)
Hydraena gracilidelphis (H20)
Hydraena gracilis (H19)
Hydraena saga (H16)
Hydraena emarginata (H17)
Hydraena exasperata (H18)
Hydraena britteni (H29)
Hydraena brachymera (H30)
Hydraena andalusa (H4)
Hydraena sp (H8)
Hydraena allomorpha (H2)
Hydraena capta (H3)
Hydraena stussineri (H27)
Hydraena stussineri/rigua /rufipennis (H5)
Hydraena pygmea (H31)
Hydraena minutissima (H9)
Hydraena bisulcata (H6)
Hydraena albai (H28)
Hydraena cordata (H21)
Ochthebius exculptus (H25)
1.8
2.57
2.92
1.26
2.21
4.04
0.42
0.88
1.32
6.06
2.91
3.05
1.91
6.85
5.17
2.19
3.34
8.41
3.51
5.86
0.9
10
7.5
5
2.5
0


D)
M34
0.92
M35
M33
4.43
M123
M124
10.13
M125
12.78
Tricorythodes
(M36)
2.0
25
20
15
10
5
0
E)
M79
0.81
M13
1.37
M80
2.21
M12
4.11
M11
M15
9.05
0.85
M82
1.55
M83
9.85
6.01
M14
Thraulodes
(M5)
1.0
17.5
15
12.5
10
7.5
5
2.5
0
F)
M26
1.24
M25
2.65
M24
6.43
M114
8.64
M113
4.69
Plauditus sp
(JQ662253)
7.05
10.01
M23
Procloeon bifidum
(KY262224)
11.97
6.52
Cloeon dipterum
(KJ631638)
13.1
Plauditus virilis
(GU713764)
Baetis
(M102)
2.0
25
20
15
10
5
0

G)


M100
M99
M22
M19
M18
M21
M20
M97
M98
M17
M94
Thraulodes
(M5)
2.43
5.34
2.28
10.01
0.96
2.56
11.43
7.59
13.18
7.81
2.0
25
20
15
10
5
0

H)


B31
B10
B9
B8
B7
B30
B6
Pelonomus
(B34)
1.61
2.26
5.46
6.81
0.69
8.17
2.22
0.9
15
12.5
10
7.5
5
2.5
0