Evolutionary history of the brown trout, *Salmo trutta* L., in France

Jillian Oliver

First-year MEME research project
Université Montpellier II
April-June 2014

Supervisors: Dr. Christelle Tougard, Dr. Patrick Berrebi

Institut des Sciences de l'Evolution-Montpellier

Abstract

Salmo trutta is a species of Palearctic fish with a wide, ubiquitous distribution across Europe. *S. trutta* is divided into five widely accepted lineages: Atlantic, Danubian, Mediterranean, Adriatic, and Marbled. These five lineages are posited to have originated and diverged during the glaciation cycles of the Pleistocene, and have thus been heavily affected by these events. This study uses the mitochondrial cytochrome *b* and control region, a novel and a well-established phylogeographic marker for this species, respectively, to unravel the evolutionary history of *S. trutta* in and around France. We found that the use of both markers together is the most powerful way to tease apart the phylogeny and history of this species. The results demonstrate that the Atlantic lineage's current incongruous distribution can be traced back to the glaciations, which have affected it since its origin as the oldest lineage. The Mediterranean and Adriatic lineages, both younger, have evolved in more stable conditions, and display strong geographic structuring. We are still lacking a phylogenetic map of high resolution for *S. trutta*, but our results indicate that both cytochrome *b* and the control region should be used in future studies of this species' evolutionary history.

Introduction

The brown trout, *Salmo trutta* L., is a freshwater species of the Salmonidae family native to the Palearctic (Bernatchez, 2001). Its range extends from northern Scandinavia south towards the Atlas Mountains of northern Africa, and from Iceland east towards Afghanistan (MacCrimmon and Marshall, 1968). As is the case with many Eurasian freshwater fishes, the distribution and population structure of the brown trout has been shaped extensively by glaciation cycles throughout the Pleistocene. Glacial periods saw brown trout restricted to refugia in southern European peninsulas (Hewitt, 2000), with migration upon glacial retreat (Cortey *et al.*, 2004). Divergence of the major lineages is thought to have taken place within the last 0.5 to 2 million years (Bernatchez, 2001), thus overlapping with several severe glaciation events of the Pleistocene.

There are currently five widely accepted lineages of brown trout identified with mitochondrial DNA studies (Bernatchez *et al.*, 1992; Bernatchez, 2001). Each lineage is associated with a specific basin: Atlantic (AT), Danubian (DA), Mediterranean (ME), Adriatic (AD), and Marmoratus (MA, marbled trout, north Adriatic rivers). In addition, a lineage endemic to the Duero river basin in Spain, termed Duero (DU), has been proposed based on a unique set of haplotypes uncovered in this basin (Suárez *et al*., 2001). A Tigris River lineage (TI) found in Turkey has also been proposed as distinct (Bardakci *et al.,* 2006). Bernatchez (2001) suggested an initial allopatric divergence between Atlantic, Danubian, and Mediterranean lineages, with a subsequent divergence within the Mediterranean lineage into ME, AD, and MA lineages. Today, the AT lineage is primarily distributed across the Atlantic basin, DA in the Ponto-Caspian basin, ME in the western Mediterranean basin, AD in eastern Mediterranean tributaries with populations in northern Spain and Corsica, and MA is restricted primarily to rivers in Italy and the western Balkans (Cortey *et al*., 2004; Pustovrh *et al*., 2011).

The phylogenetic structure of brown trout in its native range has been studied extensively, and broad-scale patterns of distribution and divergence have been established. Most phylogeographic studies of brown trout have used the mitochondrial control region (CR). There have been some Eurasia-wide studies using the CR (for example, Bernatchez, 2001), but most focus on a specific area in Europe. The Iberian Peninsula has received a lot of attention due to its high level of diversity as a former

glacial refugium, and its geographic location on the edge of the brown trout's range (Suárez *et al.,* 2001; Sanz *et al.,* 2006). Other studies have been conducted on brown trout populations in the Balkans (for example, Marić *et al.*, 2006), and in France (for example, Aurelle and Berrebi, 2001). Phylogenies constructed with CR sequence data are often polytomic. Therefore, we are still lacking a phylogeographic map of high resolution.

In order to address this issue, we used a novel marker for brown trout, the mitochondrial cytochrome *b* gene (Cytb). Cytb has been used in very few instances with *S. trutta* (eg. Patarnello *et al*., 1994; Turan *et al*., 2009), but has proven to be well suited to uncovering high-resolution phylogenies in mammal populations (for example, Jaarola and Searle, 2002; Tougard *et al.,* 2013). Cytb's usefulness in elucidating phylogeography at haplotype-level resolution may be attributed to its slower rate of evolution, which is more suited to these data than is the CR (Tang *et al.,* 2006).

In the present study, the phylogeographic, demographic, and diversity patterns of brown trout across France and elsewhere in Europe were elucidated with a dataset containing the complete Cytb and CR sequences. This data was used as well to compare the usefulness of Cytb and CR in uncovering the phylogeographic patterns of *S. trutta*.

Materials and Methods

Sampling

Three hundred fifty-one tissue samples were taken from ISEM's collection of trout fin samples, and selected so as to cover as wide a distribution as possible in France. Samples from the United Kingdom [4], Ireland [3], Germany [5], Croatia [6], Russia [2], Romania [4], and Turkey [6] were also included as peripheral references.

DNA processing and sequence alignment

Total DNA was extracted from tissue samples according to the protocol of the GenElute Mammalian Genomic DNA Kit (Sigma). Cytb was PCR-amplified in two overlapping fragments, and CR was PCR-amplified to supplement existing sequence data (**Table 1**).

	Primer	Sequence	Reference
Cytb (1)	SalmoCBF	5'-CATAATTCCTGCCCGGACTCTAACC-3'	Crête-Lafrenière *et al*., 2012
	SalmoCB1R	5'-GGTTATTAGAGCCGGTTTCATG-3'	Not published
Cytb (2)	SalmoCB1F	5'-CGACAACGCCACCCTAACACGA-3'	Not published
	SalmoCBR	5'-TTTAACCTCCGATCTCCGGATTACA-3'	Crête-Lafrenière *et al*., 2012
CR	StrDLF	5'-GCACCGACTACACTATCAATT-3'	Not published
	StrDLR	5'-TTTATATGTTTGATTGAGA-3'	Not published

Table 1. Sequences of primers used to amplify CytB and CR, and references.

PCR products were either sent to Macrogen Inc. (Amsterdam, the Netherlands) to be sequenced, or locally processed (Plateforme Génomique Environmentale du Labex CeMEB at UM2). In the latter case, PCR products were purified and products were then sequenced by an ABI Prism (Applied Biosystems) sequencer in forward and reverse directions. Collected sequences were aligned with sequences from both GenBank and from a previous study (Reynaud, 2011) using MEGA6.06 (Tamura *et al.,* 2013).

Phylogenetic analyses

Phylogenetic trees were constructed from six datasets: (1) Cytb [537 sequences] (2) CR [628 sequences], (3) concatenated Cytb and CR [428 sequences] (**Appendix II, Table E1, E2**), and corresponding haplotypes. *Salmo salar* was used as outgroup. Phylogenetic trees were constructed using both Maximum Likelihood (ML) with PhyML3.0 (Guindon *et al*., 2010), and Bayesian inference (BI) with MrBayes3.1.2 (Ronquist and Huelsenbeck, 2003). In both cases, the best-fitting evolutionary models were defined *a priori* by the minimum value of the Akaike Information Criterion, and were obtained with MrModeltest2.3 (Nylander, 2008). In both ML and BI, all datasets except for CR sequences (2) and the CR portion of the concatenated sequences (3) under BI were run under the GTR model (Yang, 1994), with the proportion of invariable sites (I) and the gamma distribution (G) taken into account. The two excluded datasets were run under the HKY + I + G model (Hasegawa *et al*., 1985). Node robustness in ML was estimated by bootstrap percentages (BP) from 100 pseudo-replicates. In BI, three Markov Chain

Monte Carlo samplings were run for three million generations. Node robustness was estimated by posterior probabilities (PP), which were obtained by sampling trees every 100^{th} generation, and by discarding the first 25,000 trees as a burn-in stage. Alternative hypotheses were examined with Paup4.0b10 by the Shimodaira and Hasegawa test, which uses ML to explore all the possible tree hypotheses of consensus sequences from each lineage (Shimodaira and Hasegawa, 1999; Swofford, 2002). A topology with $p>0.05$ indicates that it is not worse than the best ML topology.

Genetic diversity indices

Genetic distances were estimated within and between lineages using MEGA. DnaSP5.10 (Librado and Rozas, 2009) was used to calculate for each lineage the number of haplotypes (*nh*), nucleotide diversity (π), haplotype diversity (*h*), and mean number of pairwise differences (*k*). Lineages of high diversity will have a high *nh* compared to number of total sequences, high π (close to 1), high h (close to 1), and high k.

Demographic analyses

Demographic history was inferred from a mismatch distribution (Rogers and Harpending, 1992), and three neutrality tests using DnaSP: (1) Fu's Fs (Fu, 1997), (2) Tajima's D (Tajima, 1989), and (3) R_2 (Ramos-Onsins and Rozas, 2002). Significantly negative Tajima's D and Fs, and a significantly positive R_2 indicate that the population is in the process of expanding. Mismatch analysis produces a curve displaying the distribution of pairwise differences in a population or lineage, which is compared to an expected curve under a model of population growth-decline. A curve with a single peak at a low number of pairwise differences indicates expansion, while a curve with two or multiple peaks indicates stability. Deviations of observed from expected distributions were compared using the (4) sum of squared deviations (SSD) and the (5) raggedness index (r) of Harpending (1994) with 1000 permutations as implemented by Arlequin3.5.1.2 (Excoffier and Lischer, 2010). For both indices, rejection of the null hypothesis of expansion ($p<0.05$) indicates stability.

Population structure

Median-joining haplotype networks were constructed from the concatenated dataset with Network4.2.1.6 (http://www.fluxus-engineering.com/sharenet.htm; Bandelt *et al*., 1999) for each lineage for a spatial reconstruction of how the haplotypes are related to each

other and have diverged.

Divergence time estimates

Time to the most recent common ancestor (TMRCA) was estimated for selected clades from the concatenated dataset with a Bayesian coalescent analysis using BEAST1.8.0 (Drummond *et al.*, 2012) under a GTR + I + G model. Three molecular clocks (strict, uncorrelated lognormal, uncorrelated exponential) and three coalescent models (constant size, exponential growth, expansion growth) were compared using Bayes Factor to determine the best fitting model to the data (Brandley *et al.,* 2005). The models were run for fifty million generations with a 10% burn-in stage. A calibration point of 11.6 ± 2.3 Mya for the origin of the *Salmo* genus was used (Crête-Lafrenière *et al.,* 2012).

Results

Phylogenetic analysis

The aligned Cytb contains 1140 sites with 107 parsimony-informative sites (PIS), CR contains 976 sites (including indels) and 95 PIS, and the concatenated dataset contains 2116 sites (including indels) with 171 PIS.

Six sequence trees and six haplotype trees were produced [Cytb: 96 haplotypes (**Figures S1, S3**), CR: 205 (**Figure S2, S4**), concatenated: 153 (**Figures 1, S5**)]. Haplotype trees were used to readily interpret the large amount of sequence data.

In all haplotype trees constructed with both ML and BI, the monophyly of *S. trutta* is significantly supported (100% BP; 0.98-1.00 PP) (**Figure 1**; **Figures S1-5**). The five established lineages (AT, ME, AD, MA, DA) are monophyletic in all BI trees, but only in the concatenated ML tree (**Table 2**). A monophyletic TI lineage is supported in all trees except for the Cytb BI tree (**Table 2**). In the Cytb BI tree, one haplotype each from ME and AD are not grouped together with the other haplotypes (Cytb_Hap_35 and Cytb_Hap_91, respectively) (**Figure S1**). These haplotypes were assigned based on position within the concatenated BI tree. Likewise in the CR BI and ML trees, one AT haplotype (CR_Hap_205), assigned based on the concatenated BI tree, is excluded from its lineage's clade, but AT remains monophyletic in both cases (**Figures S2, S4**). DU haplotypes included in the CR dataset are strongly supported as monophyletic (PP=0.99, BP=33%), but within AT in both BI and ML trees (**Figures S2, S4**).

	BI			ML		
Lineage/Tree	**Cytb (PP)**	**CR (PP)**	**Concat. (PP)**	**Cytb (BP)**	**CR (BP)**	**Concat. (BP)**
AT	0.55	0.95*	0.84	N/A	N/S	NS
ME	0.51*	0.91	0.92	N/A	N/S	55
AD	0.94*	0.55	0.91	N/A	N/A	N/S
DA	1	0.96	1	87	57	92
MA	0.87	0.55	0.97	74	56	91
TI	N/A	1	0.51	N/S	100	78

Table 2. Posterior probabilities (PP, >0.50) for BI trees and bootstrap percentages (BP, >50%) for ML trees with all datasets. N/A indicates lineage non-monophyly, while N/S indicates unsupported node.
*the value given is the robustness value for the clade except for one excluded individual in each case.

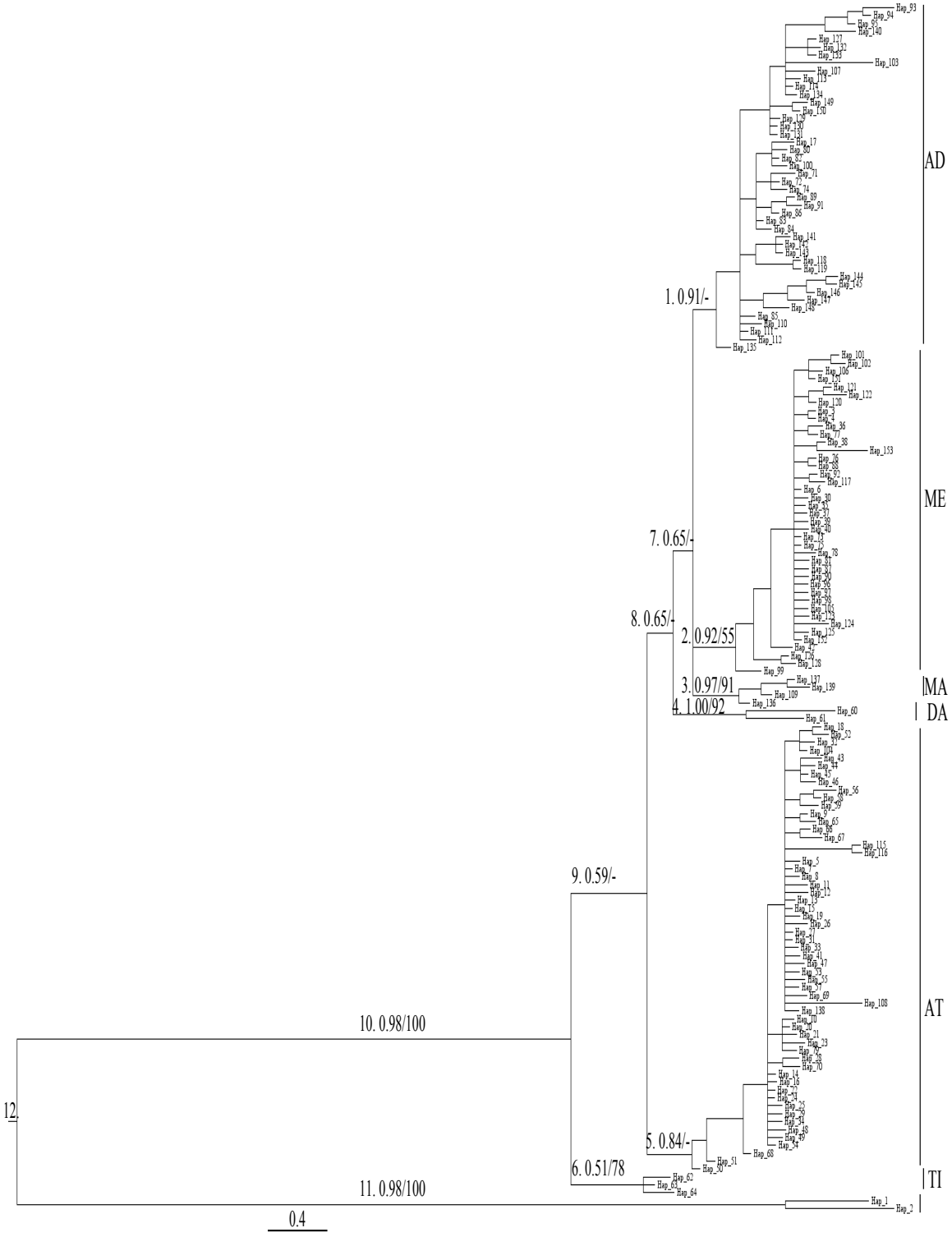

Figure 1. Bayesian inference tree reconstructed from concatenated Cytb and CR haplotypes. Haplotype labels are detailed in **Table E1**. The labels on the right refer to the lineages. Numbers at the nodes refer to posterior probabilities and bootstrap percentages (PP>5.0/BP>50%). (-) indicates BP<50%. Node numbers are provided before PP on nodes used for dating.

The concatenated ML and BI trees show similarities in that in both trees, AD, ME, and MA are clustered and polytomic (PP=0.65; BP<50%) (**Figure 1**; **Figure S5**). This cluster has a common ancestor with DA (PP=0.65) in the BI tree (**Figure 1**), but with AT (BP=31%) in the ML tree (**Figure S5**). In both concatenated trees, TI has a basal position in *S. trutta*. In contrast, in the CR BI tree, TI is nested within the tree and shares a common ancestor with DA (PP=0.55).

All of the 10,395 possible alternative hypotheses uncovered with the Shimodaira and Hasegawa test were not significantly different from the highest likelihood tree ($p \geq 0.05$). Of these trees, 944 placed *S. salar* as outgroup and within these hypotheses, 105 rooted *S. trutta* on AT, and 105 rooted *S. trutta* on TI.

Genetic diversity

Intra- and inter-lineage genetic distances were calculated with each dataset (**Table 3; Table S1 a-b**). Levels of intra-lineage distance are comparable across datasets, but levels of inter-lineage distance are consistently higher in the Cytb dataset, and lowest in the CR dataset.

	d (SE)	**ME**	**MA**	**AT**	**AD**	**TI**	**DA**	***S. salar***
ME	0.001 (0.000)	-	0.002*	0.002*	0.002*	0.002*	0.002*	0.007*
MA	0.001 (0.001)	0.008	-	0.002*	0.002*	0.002*	0.002*	0.007*
AT	0.002 (0.000)	0.010	0.010	-	0.002*	0.002*	0.002*	0.007*
AD	0.002 (0.001)	0.008	0.007	0.009	-	0.001*	0.002*	0.007*
TI	0.002 (0.001)	0.011	0.009	0.010	0.008	-	0.002*	0.007*
DA	0.004 (0.001)	0.011	0.010	0.012	0.010	0.012	-	0.008*
S. salar	0.004 (0.001)	0.060	0.060	0.058	0.059	0.058	0.0620	-

Table 3. Intra- and inter-lineage genetic distances with the concatenated dataset. *SE=standard error.

All datasets contain comparable levels of genetic diversity for ME, AT, AD, TI, and DA (**Table 4; Table S2 a-b**). MA displays a much lower level of diversity in all indices only with the CR dataset (**Table S2b**).

	N	*nh*	π (SD)	*h* (SD)	*k* (%)
ME	103	36	0.00146 (0.00014)	0.933 (0.016)	0.14
MA	13	4	0.00117 (0.00018)	0.731 (0.079)	0.12
AT	206	55	0.00177 (0.00009)	0.929 (0.009)	0.18
AD	92	29	0.00223 (0.00012)	0.947 (0.011)	0.21
TI	6	3	0.00175 (0.00073)	0.600 (0.215)	0.17
DA	6	2	0.00406 (0.00131)	0.533 (0.02963)	0.40

Table 4. Indices of genetic diversity with the concatenated dataset. N= number of sequences, *nh*= number of haplotypes, π=nucleotide diversity, *h*=haplotype diversity, *k*= average percent of pairwise differences, SD=standard deviation.

Demography

A population can be considered to be in expansion if three of the five measured indices (D, Fs, R_2, SSD, and r) indicate such. Data from all three datasets suggest that ME, AD, and AT are in expansion (**Table 5; Table S3 a-b**). The three lineages' mismatch curves do indeed match the expected curves under expansion (**Figure 2a; Figures S6-8**). The Cytb and concatenated datasets also suggest TI expansion. TI has only one CR haplotype, so these indices could not be calculated from this dataset. MA and DA are stable according to the Cytb and concatenated datasets, but in expansion in the CR dataset. DA's CR curve is inconclusive, but MA's CR mismatch curve follows the curve expected under expansion (**Figure 2b**).

	Neutrality Tests			**Mismatch Analysis**		**Model**
	Fs	**D**	**R_2**	**SSD**	**r**	
ME	**-28.267*****	**-1.94373***	**0.0364***	**0.00589**	**0.03750**	**Expansion**
MA	1.605	0.34603	0.1717	**0.09590**	**0.27268**	(Stable)
AT	**-46.182*****	**-1.66763***	**0.0369***	**0.00680**	**0.03004**	**Expansion**
AD	**-11.509****	**-1.43356***	0.0544	**0.00356**	**0.00855**	**Expansion**
TI	2.242	**-1.44477***	0.2372	**0.18715**	**0.31111**	**(Expansion)**
DA	7.400	1.34683	0.2667	0.56889***	**0.78667**	(Stable)

Table 5. Demographic indices with the concatenated dataset. Fs=Fu's F statistic, D=Tajima's D statistic, R_2= Ramos-Onsins and Rozas statistic, SSD=sum of standard deviations of mismatch distribution, r=raggedness index of mismatch distribution. Models in brackets indicate low sample size. ***p<0.001, **p<0.01, *p<0.05.

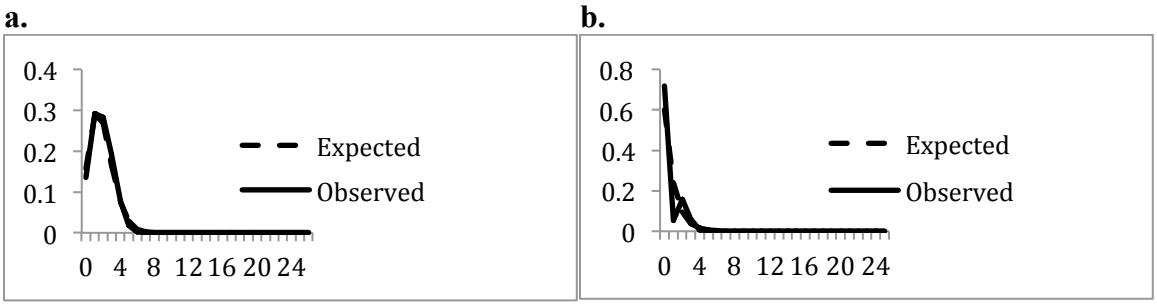

Figure 2. Mismatch distributions of (**a**) the AD lineage, and (**b**) the MA lineage from the CR dataset.

Population structure

Three haplotype networks were constructed from the concatenated dataset: (1) ME, (2) AD + MA + DA + TI, (3) AT. Networks are presented in this way to facilitate visualization. The ME network shows very strong spatial segregation (**Figure S9**). The most prevalent haplotype (CytB+CR_Hap_6) consists of sequences from individuals found in the southeast of France. From this haplotype, additional southeast French, Spanish, Corsican, Italian, and Croatian haplotypes have diverged.

The AD lineage also shows strong spatial differentiation, but without a predominant haplotype (**Figure S10**).

The AT network does not show geographic correlation (**Figure S11**). The most present haplotype (CytB+CR_Hap_7) is found in rivers all across France, Corsica, the UK, and Germany. The second and third most present haplotypes (CytB+CR_Hap_15 and CytB+CR_Hap_24, respectively) are more localized to mainland France, but still do not display any geographic correlation.

Divergence Time Estimates

A strict clock under a constant size coalescent model fit the data best. The origin of *Salmo* is placed at 10.47 Mya, with *S. trutta* diverging 1.16 Mya (**Table S4**). AT is the oldest lineage, with its origin dated to 0.61 Mya, and MA is the youngest lineage, with its origin dated to 0.19 Mya. **Figure S12** shows the degree of support of each lineage's TMRCA estimate. Curves with one peak represent confident estimates.

Discussion

Utility of Cytb and CR in resolving S. trutta *phylogeny*

The mitochondrial control region has been used extensively in phylogenetic studies of *S. trutta*, but has not been able to resolve inter-lineage relationships. This study introduced the use of the mitochondrial cytochrome *b* gene to wide-scale phylogeographic mapping of *S. trutta*. Cytb has a slower rate of evolution than CR (Tang *et al.,* 2006), so it was predicted that Cytb would provide information lost in C R to phylogenetic analysis, and thus be able to better resolve lineages of *S. trutta*. Our phylogenies constructed from datasets containing only Cytb sequences suggest that it is not better at resolving lineages than CR. Both the BI and the ML haplotype trees reconstructed from the Cytb dataset contain poorly supported nodes, if any, and do not resolve relationships between lineages. The trees constructed from the CR provide more resolution than do the Cytb trees, but the trees constructed from the concatenated dataset provide the clearest inter-lineage resolution. Thus, our results suggest that while Cytb has not proven to be a useful phylogenetic marker on its own with *S. trutta*, it produces much more robust trees than either gene alone when concatenated with CR.

Inter-lineage relationships

Bernatchez's (2001) results from a cross-Europe study of *S. trutta* phylogeography using the CR suggest that the AT lineage was the first to diverge. Our concatenated BI tree, which is the best tree in terms of inter-lineage resolution and nodal support, suggests that the TI lineage (not considered by Bernatchez, 2001) is the first to diverge, followed by AT. Indeed, a search of all possible lineage consensus ML hypotheses revealed that hypotheses in which TI diverges first are not worse than the best possible hypothesis, or than hypotheses in which AT diverges first. Therefore, the hypothesis that TI is the first lineage to diverge within *S. trutta* is plausible. Aside from this discrepancy, the concatenated BI tree agrees with Bernatchez's (2001) topology. It must be noted, however, that the nodes resolving inter-lineage relationships in the concatenated phylogeny are not very strongly supported, so lineage divergenc is eessentially polytomic.

Dating S. trutta

In our analysis of divergence times, we used a coalescent model to infer the TMRCA of a series of selected nodes, calibrated by an estimation of *Salmo*'s origin from Crête-Lafrenière *et al*. (2012). This calibration point was chosen due to the very limited *Salmo* fossil record. Crête-Lafrenière *et al*. (2012) inferred an origin of 11.6 ± 2.3 Mya for *Salmo* by using a relaxed molecular clock and five fossil calibration points in a phylogenetic study of Salmonidae. In our analysis, this was the only calibration point in a lineage that spans upwards of ten million years. The youngest estimate in our analysis belongs to the MA lineage, whose origin is dated to 0.19 Mya. Estimates at such a different scale than the calibration point are informative, but should be interpreted with caution. Our use of coalescent model selection, however, rather than a pre-defined divergence rate for the entire tree, as in Bernatchez (2001), provides a more realistic estimate of lineage ages (Drummond *et al.*, 2006). While Bernatchez (2001) did not perform a dating analysis, he used a rate of 1-2% sequence divergence per million years in other analyses with CR. He based this estimate on findings of the rate of evolution of mitochondrial DNA in salmonids (Smith, 1992). Such an estimate is not specific enough to *Salmo*, or to CR, which evolves at a higher rate than the rest of the mitochondrial genome (Tang *et al.*, 2006), to provide plausible estimates.

Evolutionary history of S. trutta

(1) The Atlantic lineage

According to dating estimates, AT is the oldest lineage, originating 0.61 Mya, right before the Günz glaciation (0.6-0.5 Mya; Randi, 2007). Along with ME, AD, and TI lineages, AT is in the process of recovery from a population collapse. The AT lineage contains, for the most part, populations situated at more northern latitudes than do the other lineages (**Table E1**). Thus, AT populations were more heavily affected by glaciation cycles in the Pleistocene, and have existed in stable environmental conditions for a shorter period of time (Bernatchez and Wilson, 1998). This instability is reflected in the AT network, which shows little geographic structuring. Prominent haplotypes are found all over and outside France, and closely related haplotypes do not cluster according to geography. This pattern suggests a process of population mixing throughout the Pleistocene. Rivers and lakes in the Atlantic basin underwent more drastic changes

throughout the glaciations than did rivers and lakes in other basins, and went through cycles of population collapse, expansion, and recolonization (Bernatchez and Wilson, 1998; Ball-Llosera *et al*., 2002). Additionally, AT fish took refuge in the southern peninsulas during glaciations, leading to their current distribution outside of the AT basin. The geographically disorganized AT haplotype network is therefore a strong reflection of the lineage's eventful history.

(2) The Mediterranean and Adriatic lineages

Unlike AT, the ME and AD lineages are present mostly in areas that acted both as hotspots for endemism, and as refugia during glaciation events (Bilton *et al*., 1998; Hewitt, 1996). They were thus exposed to a different set of pressures during glaciation. When global temperatures increased during interglacial periods, many populations living in glacial refugia decreased in size or went extinct (Bennet, 1990; Taberlet *et al.,* 1998). ME and AD lineages show very strong geographic correlation: closely related haplotypes are clustered according to spatial location. The ME network also unravels a pattern of diversification within the lineage. The most common haplotype in the lineage occurs in the Rhône and Roya rivers in southeast France, and all other haplotypes stem from this one according to location. This suggests that ME originated in southern France, agreeing with Persat and Berrebi's (1990) suggestion of Roussillon as a centre of dispersion.

(3) A comparison across AT, ME, and AD lineages

Based on a pattern of post-glaciation northward dispersal, one would expect the AT lineage to be characterized a lower genetic diversity than the other lineages (Randi, 2007), but our results suggest that it is not lower than in the other lineages. The comparable levels of diversity suggest that all three lineages have undergone population collapses at similar times, leading to similar patterns of demography and diversity despite being exposed to opposing forces during glaciation cycles. An alternative explanation for the comparable levels of genetic diversity between AT, and ME and AD is that the latter two lineages are younger than AT, and are still in the process of expanding and differentiating from initial colonization.

(3) The TI lineage

It is difficult to make any conclusions about the TI lineage's history from our dataset. While it follows a similar pattern of genetic diversity and demography as AT, ME, and

AD, the results must be interpreted with caution. We only used six samples, and the lineage's monophyly is not very robust (0.51 PP). Thus, any indications of the lineage's distinctness and evolutionary history must be further investigated with a concatenated dataset of many more individuals.

(4) The Danubian lineage

Out of all the lineages, DA has the highest levels of diversity, in accordance with Bernatchez (2001). Our low sample size for this lineage does not, however, allow for any confident conclusions to be made about this lineage.

(5) The Marbled lineage

MA was found to be considerably younger than the other lineages, dated to 0.19 Mya, during an interglacial period (Randi, 2007). This result is unexpected according to Bernatchez's (2001) hypothesis of a nearly simultaneous split between AD, ME, and MA. If his hypothesis is correct, then our result may reflect a limitation in the sample set. The only MA samples included in this study are from Slovenia and Corsica, which are not representative of total MA diversity (Pustovrh *et al.,* 2011). Our results suggest that these MA populations are much younger than the hypothesized origin of the lineage, which would be placed along with ME and AD between 0.3-0.5 Mya, during the Mindel glaciation (0.45-0.35 Mya; Randi, 2007). Accordingly, genetic diversity and demographic indices, and the mismatch curve produced from the CR dataset all suggest that MA is in expansion (Patarnello *et al*., 2007). These data indicate a founder's effect resulting from the recent establishment of the populations.

These results bring up the question of the different mitochondrial genes' efficacy in detecting recent sudden expansion events. Due to the recent age of the Pleistocene glaciation events, many markers evolve at too slow a rate to detect demographic events during the time span. The CR is unique in its very high rate of evolution compared to both mitochondrial and nuclear genes, and thus might be uniquely suited to detecting such events that would be missed by Cytb and other markers (Randi, 2007).

References

Aurelle D, Berrebi P. 2001. Genetic structure of brown trout (*Salmo trutta,* L.) populations from south-western France: data from mitochondrial control region variability. *Molecular Ecology* **10:** 1551-1561.

Ball-Llosera NS, García-Marín JL, Pla C. 2002. Managing fish populations under mosaic relationships. The case of brown trout (*Salmo trutta*) in peripheral Mediterranean populations. *Conservation Genetics* **3:** 385-400.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining network for inferring intraspecific phylogenies. *Molecular Biology and Evolution.* **16:** 37-48.

Bardakci F, Degerli N, Ozdemir O, Basibuyuk HH. 2006. Phylogeography of the Turkish brown trout *Salmo trutta* L.: mitochondrial DNA PCR-RFLP variation. *Journal of Fish Biology* **68**: 36-55.

Bennet KD. 1990. Milankovitch cycles and their effects on species in ecological and evolutionary time. *Paleobiology*. **16:** 11-21.

Bernatchez L, Guyomard R, Bonhomme F. 1992. DNA sequence variation of the mitochondrial control region among geographically and morphologically remote European brown trout *Salmo trutta* populations. *Molecular Ecology* **1:** 161-173.

Bernatchez L, Wilson CC. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology* **7:** 431-452.

Bernatchez L. 2001. The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution* **55:** 351-379.

Bilton DT, Mirol PM, Mascheretti S, Fredga K, Zima J, Searle JB. 1998. Mediterranean Europe as an area of endemism for small mammals rather than a source of northwards postglacial colonization. *Proceedings of the Royal Society B* **265:** 1219-1226.

Brandley MC, Schmitz A, Reeder TW. 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. *Systematic Biology* **54(3):** 373-390.

Cortey M, Pla C, García-Marín JL. 2004. Historical biogeography of Mediterranean trout. *Molecular Phylogenetics and Evolution* **33:** 831-844.

Cortey M, Vera M, Pla C, García-Marín JL. 2009. Northern and southern expansions

of Atlantic brown trout (*Salmo trutta*) populations during the Pleistocene. *Biological Journal of the Linnean Society* **97:** 904-917.

Crête-Lafrenière A, Weir LK, Bernatchez L. 2012. Framing the Salmonidae family phylogenetic portrait: A more complete picture from increased taxon sampling. *PLoS One* **7(10):** e46662.

Drummond AJ, Ho SYW, Philips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS One* **4(5).** e88.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969-1973.

Duftner N, Weiss S, Medgyesy N, Sturmbauer C. 2003. Enhanced phylogeographic information about Austrian brown trout populations derived from complete mitochondrial control region sequences. *Journal of Fish Biology* **62(2):** 427-435.

Excoffier L, Lischer HEL. 2010. Arlequin suite version 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10:** 564-567.

Fu Y-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147:** 915-925.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithm and methods to estimate maximum likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59:** 307–321.

Harpending HC. 1994. Signature of ancient population growth in a low resolution mitochondrial DNA mismatch distribution. *Human Biology* **66:** 591-600.

Hasegawa M, Kishino H, Yano T. 1985. Dating the human-ape split by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22:** 160-174.

Hewitt GM. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58:** 247-276.

Hewitt GM. 2000. The genetic legacy of Quaternary ice ages. *Nature* **405:** 907-913.

Jaarola M, Searle JB. 2002. Phylogeography of field voles (*Microtus agrestis*) in Eurasia inferred from mitochondrial DNA sequences. *Molecular Ecology* **11:** 2613-2621.

Librado P, Rozas J. 2009. DNAsp v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:** 1451-1452.

MacCrimmon HR, Marshall TL. 1968. World distribution of brown trout, *Salmo trutta*. *Journal Fisheries Research Board of Canada* **25(12):** 2527-2548.

Marić S, Sušnik S, Simonović P, Snoj A. 2006. Phylogeographic study of brown trout from Serbia, based on mitochondrial DNA control region analysis. *Genetics Selection Evoluion.* **38:** 411-430.

Meraner A, Baric S, Pelster B, Via JD. 2007. Trout (*Salmo trutta*) mitochondrial DNA polymorphism in the centre of the marble trout distribution area. *Hydrobiologia*. **579:** 337-349.

Nylander JAA. 2008. MrModeltest 2.3. Program distributed by the author. Evolutionary Biology Centre. Uppsala University.

Patarnello T, Bargelloni L, Caldara F, and Colombo L. 1994. Cytochrome *b* and 16S rRNA sequence variation in the *Salmo trutta* (Salmonidae, Teleostei) species complex. *Molecular Phylogenetics and Evolution*. **3(1):** 69-74.

Patarnello T, Volckaert FAMJ, Castilho R. 2007. Pillars of Hercules: is the Atlantic-Mediterranean transition a phylogeographical break? *Molecular Ecology* **16:** 4426-4444.

Persat H, Berrebi P. 1990. Relative ages of present populations of *Barbus barbus* and *Barbus meridionalis* (Cyprinidae) in southern France: preliminary considerations. *Aquatic Living Resources* **3:** 253-263.

Pustovrh S, Bajec SS, Snoj A. 2011. Evolutionary relationship between marble trout of the northern and the southern Adriatic basin. *Molecular Phylogenetics and Evolution* **59:** 761-766.

Ramos-Onsins SE, Rozas J. 2002. Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19:** 2092-2100.

Randi E. 2007. 'Chapter 3 Phylogeography of south European mammals', in S. Weiss, N. Ferrand (eds.), *Phylogeography of Southern European Refugia*. Springer. Dordrecht.

Reynaud N. 2011. Structuration géographique de la truite commune (*Salmo trutta*) en France basée sur le séquençage de la région de contrôle mitochondriale. *Rapport OSU OREME*: 46 p.

Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9:** 552-589.

Ronquist F, Huelsenbeck JP. 2003. MrBayes3: Bayesian phylogenetic inference under

mixed models. *Bioinformatics* **19:** 1572–1574.

Sanz N, Cortey M, Pla C, García-Marín JL. 2006. Hatchery introgression blurs ancient hybridization between brown trout (*Salmo trutta*) lineages as indicated by complementary allozymes and mtDNA markers. *Biological Conservation* **130:** 278-289.

Shimodaira H, Hasegawa M. 1999. Mutiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* **16:** 1114-1116.

Smith GR. 1992. Introgression in fishes: significance for paleontology, cladistics, and evolutionary rates. *Systematic Biology* **41:** 41-57.

Suárez J, Bautista JM, Almodóvar A, Machordom A. 2001. Evolution of the mitochondrial control region in Palaearctic brown trout (*Salmo trutta*) populations: the biogeographical role of the Iberian Peninsula. *Heredity* **87:** 198-206.

Swofford DL. 2002. PAUP* version 4.0. *Phylogenetic analysis using parsimony (and other methods)*, 4.0 edn. Sinauer, Sunderland, MA.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7:** 453-464.

Tajima F. 1989. Statistical model for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123:** 585-595.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30:** 2725-2729.

Tang Q, Liu H, Mayden R, Xiong B. 2006. Comparison of evolutionary rates in the mitochondrial DNA cytochrome *b* gene and control region and their implications for phylogeny of the Cobitoidea (Teleostei: Cypriniformes). *Molecular Phylogenetics and Evolution* **39:** 347-357.

Tougard C, Montuire S, Volobouev V, Markova E, Contet J, Aniskin V, Quere J-P. 2013. Exploring phylogeography and species limits in the Altai vole (Rodentia: Cricetidae). *Biological Journal of the Linnean Society* **108:** 434-452.

Turan D, Kottelat M, Engin S. 2009. Two new species of trouts, resident and migratory, sympatric in streams of northern Anatolia (Salmoniformes: Salmonidae). *Ichthyological Exploration of Freshwaters* **20:** 333-364.

Yang Z. 1994. Estimating the pattern of nucleotide substitution. *Journal of Molecular Evolution.* **39:** 105-111.

Appendix I

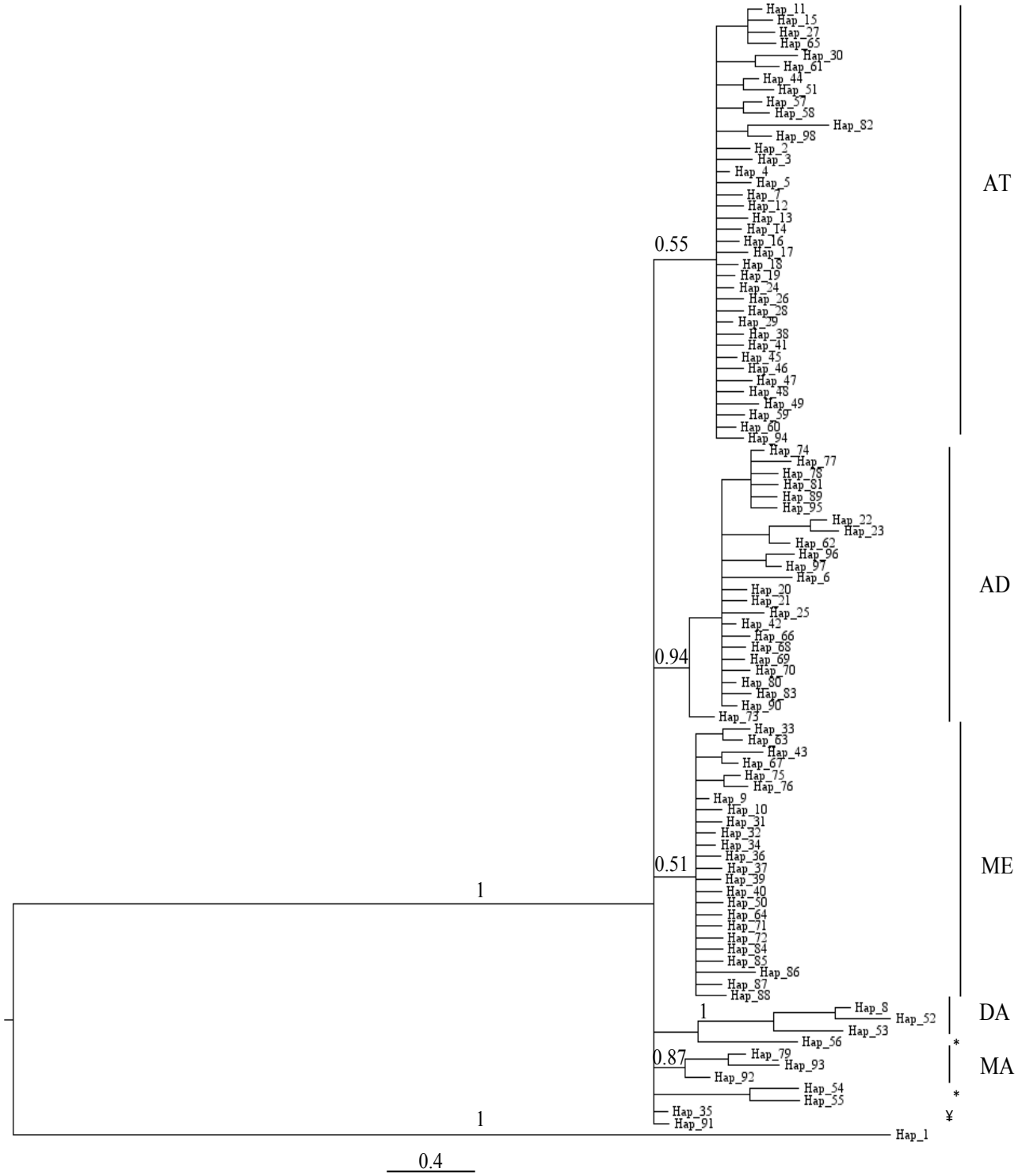

Figure S1. Bayesian inference tree reconstructed from Cytb haplotypes. Haplotype labels are detailed in Appendix II. The labels on the right refer to the lineages, and the numbers at the nodes refer to posterior probabilities (>0.50).
*Hap_56, Hap_54, and Hap_55 are TI
¥Hap_35 is ME; Hap_91 is AD

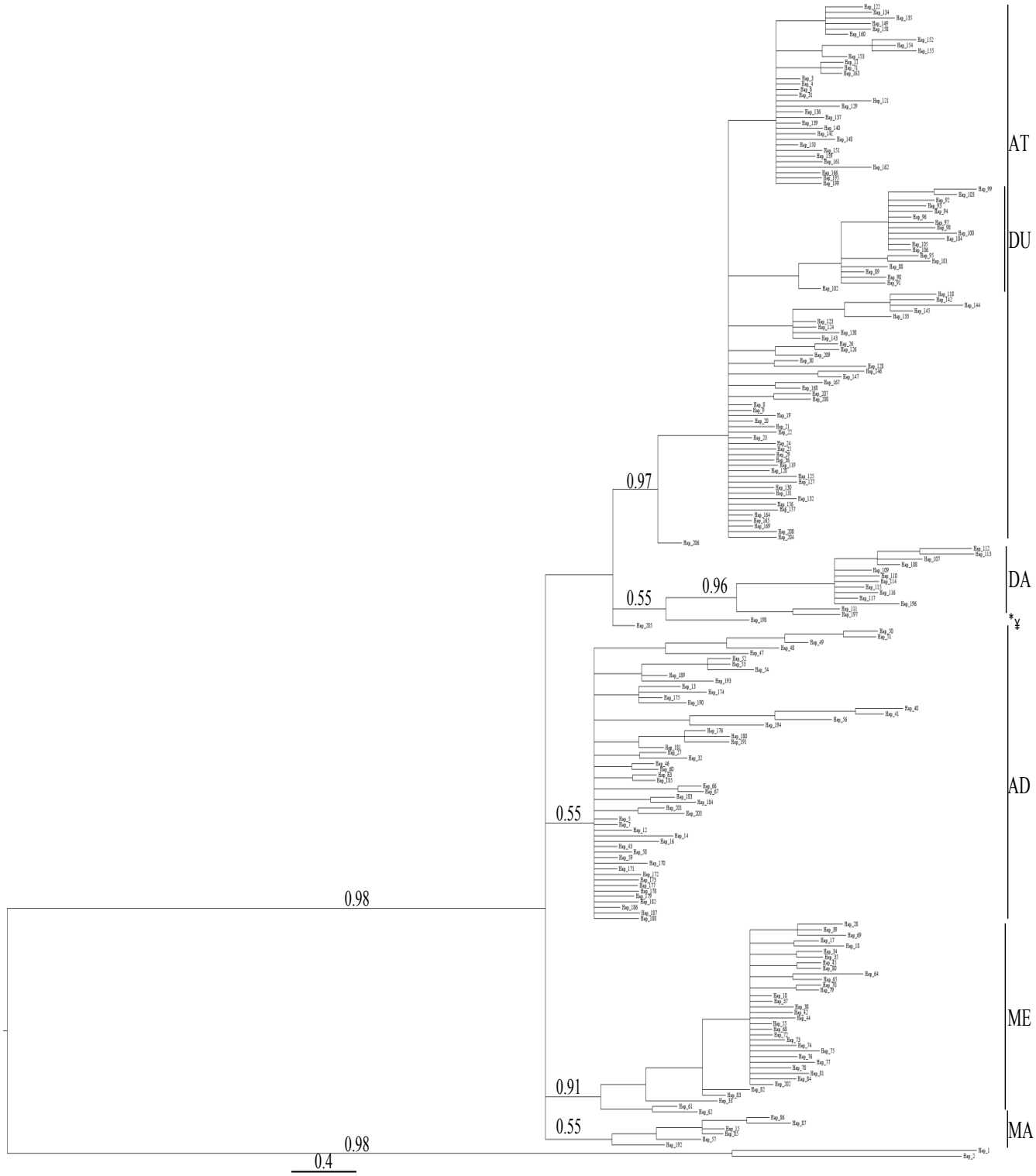

Figure S2. Bayesian inference tree reconstructed from CR haplotypes. Haplotype labels are detailed in Appendix II. The labels on the right refer to the lineages, and the numbers at the nodes refer to posterior probabilities (>0.50).
*Hap_198 is TI
¥Hap_205 is AT

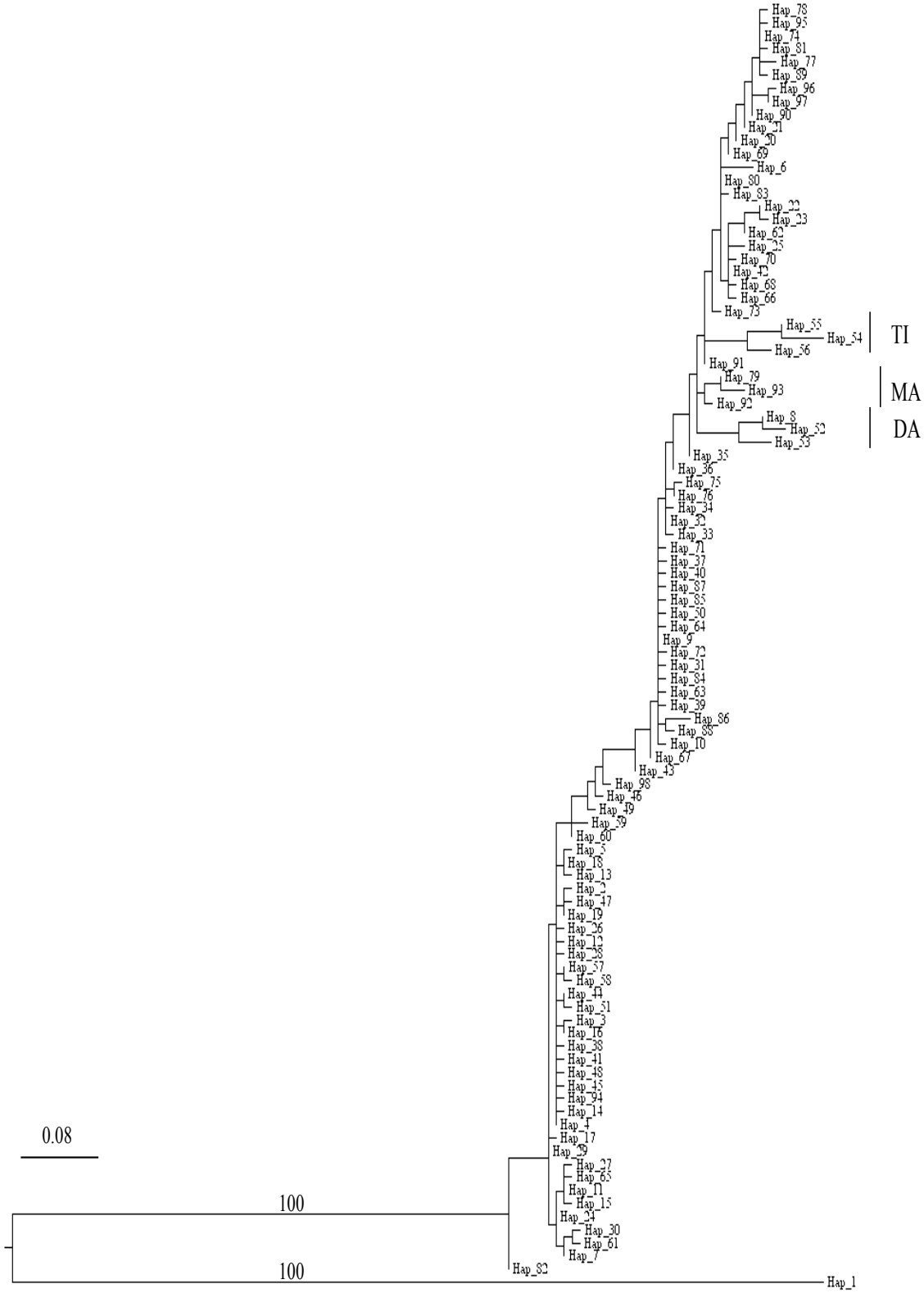

Figure S3. Maximum likelihood tree reconstructed from Cytb haplotypes. Haplotype labels are detailed in Appendix II. The labels on the right refer to the lineages, and the numbers at the nodes refer to bootstrap percentages (>50%).

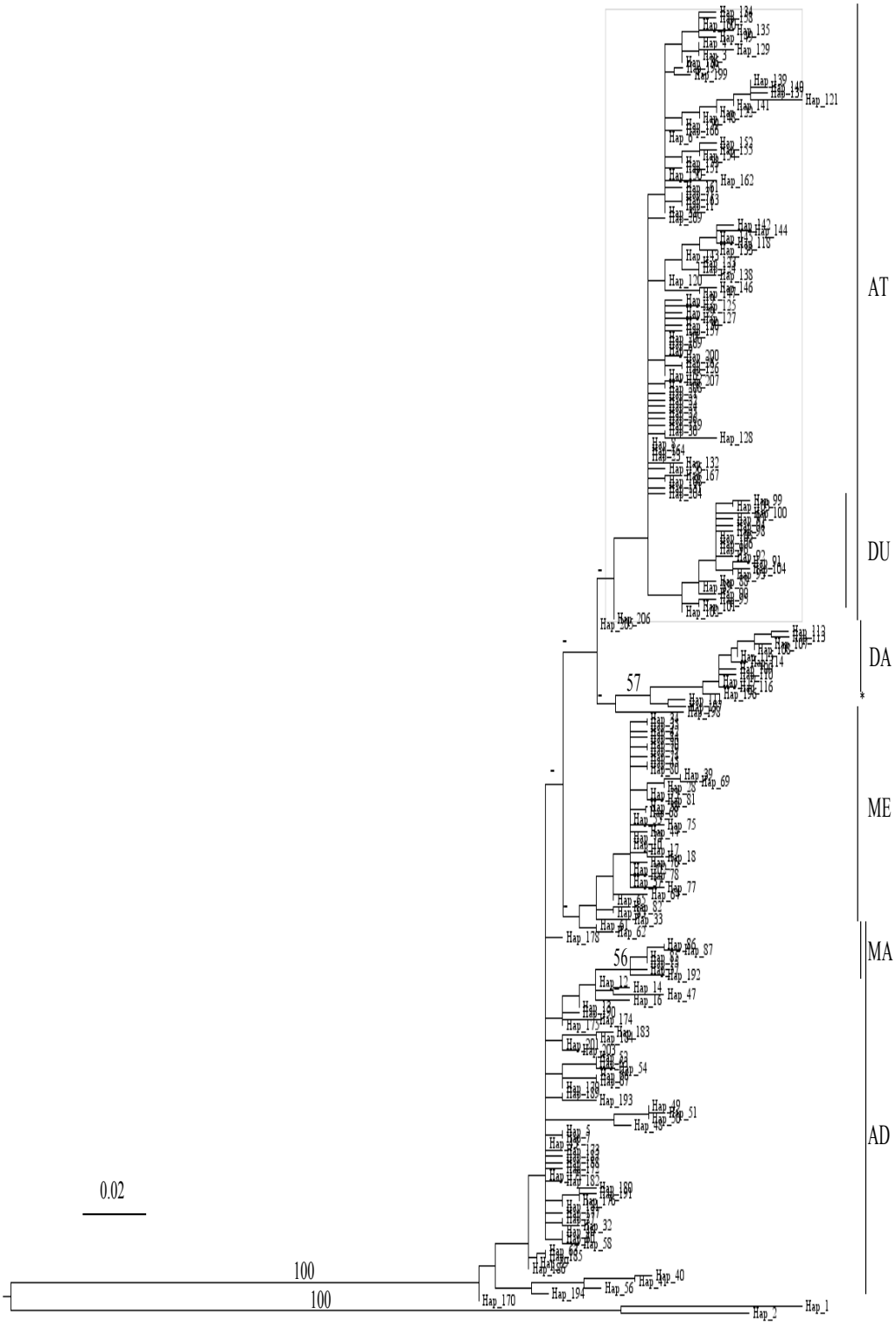

Figure S4. Maximum likelihood tree reconstructed from CR haplotypes. Haplotype labels are detailed in Appendix II. The labels on the right refer to the lineages, and the numbers at the nodes refer to bootstrap percentages (>50%).
*Hap_198 is TI.

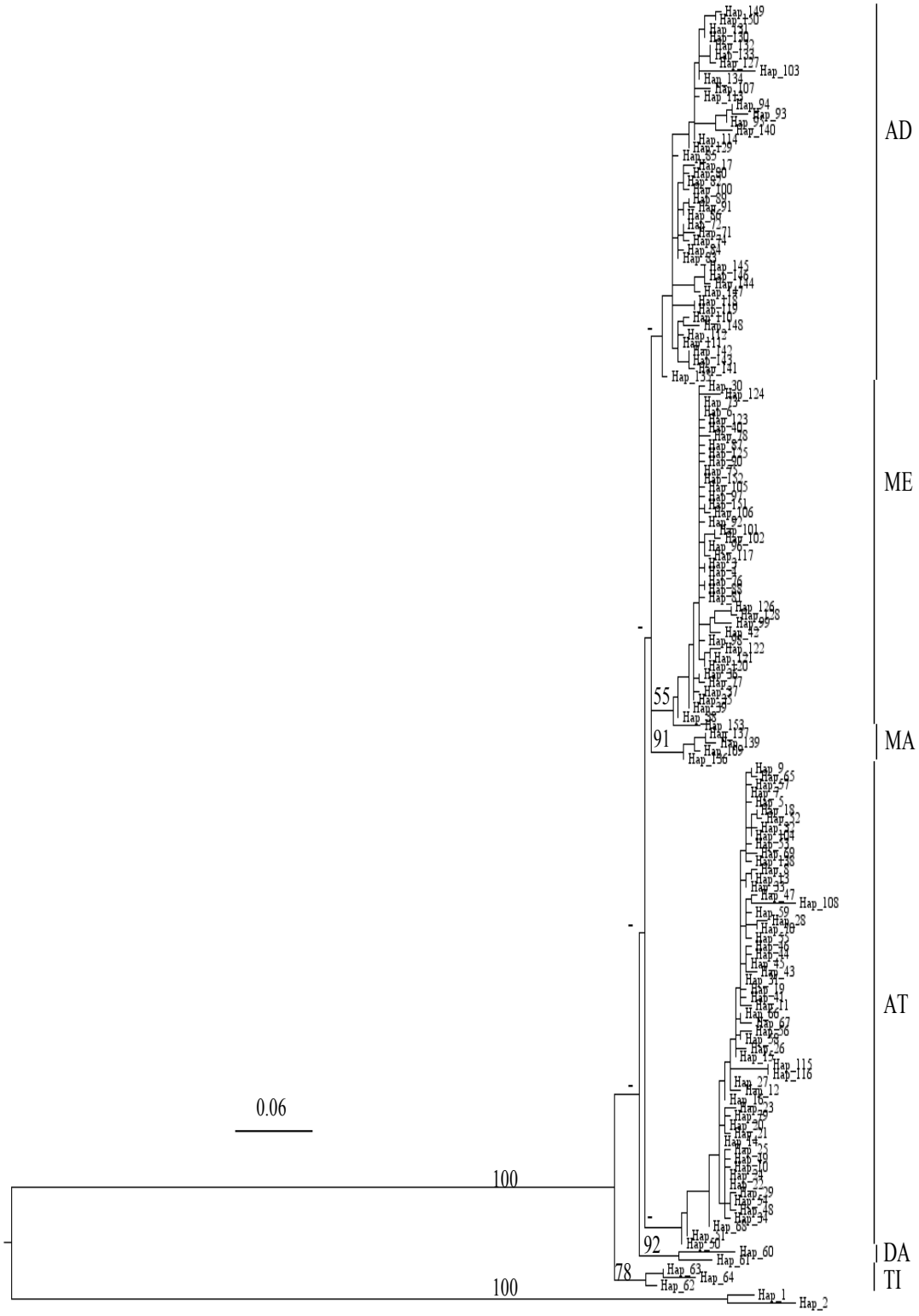

Figure S5. Maximum likelihood tree reconstructed from concatenated Cytb and CR haplotypes. Haplotype labels are detailed in Appendix II. The labels on the right refer to the lineages, and the numbers at the nodes refer to bootstrap percentages (>50%).

a.

	d (SE)	ME	MA	AT	AD	TI	DA	*S. salar*
ME	0.001 (0.000)	-	0.002*	0.003*	0.002*	0.003*	0.003*	0.007*
MA	0.002 (0.001)	0.009	-	0.003*	0.002*	0.003*	0.003*	0.008*
AT	0.002 (0.001)	0.012	0.012	-	0.002*	0.003*	0.003*	0.007*
AD	0.003 (0.001)	0.009	0.009	0.011	-	0.002*	0.003*	0.007*
TI	0.003 (0.001)	0.012	0.012	0.013	0.010	-	0.003*	0.008*
DA	0.005 (0.002)	0.013	0.012	0.015	0.012	0.015	-	0.007*
S. salar	0.000 (0.000)	0.070	0.075	0.069	0.071	0.069	0.074	-

b.

	d (SE)	ME	MA	AT	AD	TI	DA	*S. salar*
ME	0.001 (0.000)	-	0.003*	0.002*	0.002*	0.003*	0.003*	0.007*
MA	0.000 (0.000)	0.007	-	0.003*	0.002*	0.002*	0.003*	0.006*
AT	0.000 (0.001)	0.008	0.008	-	0.002*	0.003*	0.003*	0.006*
AD	0.002 (0.001)	0.006	0.004	0.007	-	0.002*	0.003*	0.007*
TI	0.000 (0.000)	0.008	0.006	0.006	0.007	-	0.002*	0.007*
DA	0.005 (0.001)	0.011	0.010	0.009	0.009	0.006	-	0.007*
S. salar	0.010 (0.003)	0.046	0.040	0.046	0.043	0.044	0.047	-

Table S1. Intra- and inter-lineage genetic distances with (**a**) the Cytb dataset, and (**b**) the CR dataset. *SE=standard error.

a.

	N	*nh*	π (SD)	*h* (SD)	*k* (%)
ME	147	26	0.00106 (0.00012)	0.701 (0.042)	0.11
MA	15	3	0.00155 (0.00029)	0.561 (0.095)	0.16
AT	251	37	0.00206 (0.00012)	0.820 (0.021)	0.21
AD	113	24	0.00264 (0.00014)	0.903 (0.014)	0.26
TI	6	3	0.00322 (0.00134)	0.600 (0.215)	0.32
DA	7	3	0.00476 (0.00104)	0.667 (0.160)	0.48

b.

	N	*nh*	π (SD)	*h* (SD)	*k* (%)
ME	134	35	0.00187 (0.00019)	0.771 (0.037)	0.18
MA	33	6	0.00044 (0.0020)	0.236 (0.097)	0.04
AT	328	99	0.00270 (0.00015)	0.876 (0.011)	0.26
AD	113	53	0.00226 (0.00018)	0.867 (0.00069)	0.19
TI	6	1	N/A	N/A	N/A
DA	17	13	0.00499 (0.00051)	0.949 (0.044)	0.49

Table S2. Indices of genetic diversity with (**a**) Cytb dataset, and (**b**) CR dataset. N= number of sequences, *nh*= number of haplotypes, π=nucleotide diversity, *h*=haplotype diversity, *k*= average number of pairwise differences, SD=standard deviation.

a.

	Neutrality Tests			Mismatch Analysis		Model
	Fs	**D**	**R_2**	**SSD**	**r**	
ME	**-25.972*****	**-2.12617*****	**0.0241***	**0.00184**	**0.06775**	**Expansion**
MA	2.346	0.51268	0.1806	0.21878*	0.57850***	(Stable)
AT	**-26.465*****	**-1.62898***	**0.0339***	**0.00357**	**0.01890**	**Expansion**
AD	**-9.478***	-1.06047	0.0591	**0.00886**	**0.02412**	**Expansion**
TI	2.242	**-1.44477***	0.2372	**0.18715**	**0.31111**	**(Expansion)**
DA	3.760	1.77071	0.2714	**0.20926**	0.49206*	(Stable)

b.

	Neutrality Tests			Mismatch Analysis		Model
	Fs	**D**	**R_2**	**SSD**	**r**	
ME	**-17.753*****	**-1.70871***	0.0361	0.11341***	**0.02505**	**Expansion**
MA	-1.950	**-1.68103***	**0.0716***	**0.02459**	**0.28486**	**Expansion**
AT	**-121.646*****	**-2.28336*****	**0.0178***	0.17443***	**0.01730**	**Expansion**
AD	**-27.397*****	**-1.88928***	**0.0338***	**0.00471**	**0.01922**	**Expansion**
TI	N/A	N/A	N/A	N/A	N/A	N/A
DA	**-5.388***	-0.19890	0.1272	**0.01352**	**0.02363**	**(Expansion)**

Table S3. Demographic indices with (**a**) Cytb dataset, and (**b**) CR dataset. Fs=Fu's F statistic, D=Tajima's D statistic, R_2= Ramos-Onsins and Rozas statistic, SSD=sum of standard deviations of mismatch distribution, r=raggedness index of mismatch distribution. Models in brackets indicate low sample size. ***p<0.001, **p<0.01, *p<0.05.

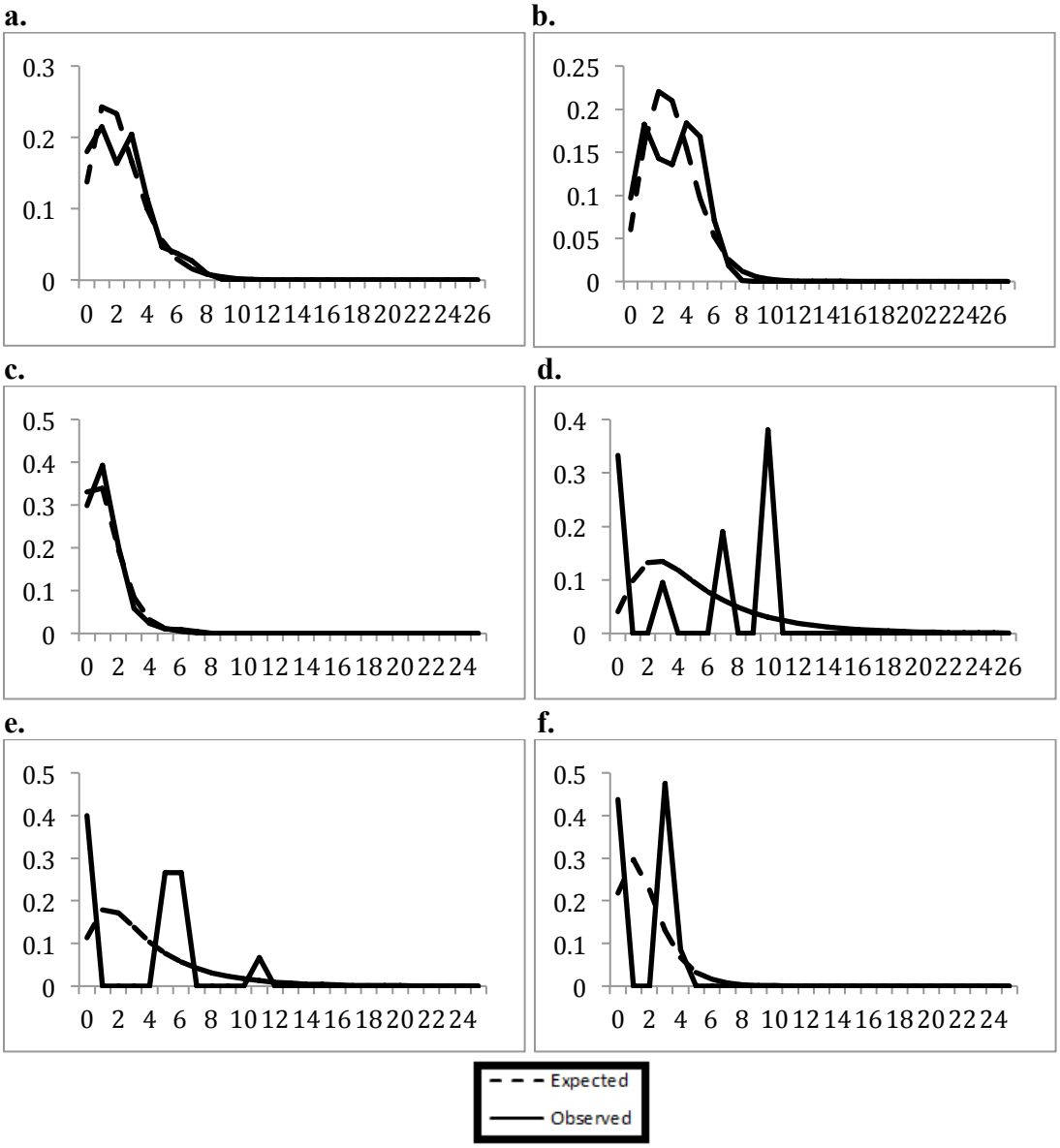

Figure S6. Mismatch distributions of (**a**) AT, (**b**) AD, (**c**) ME, (**d**) DA, (**e**) TI, and (**f**) MA with the Cytb dataset. The curves plot the frequencies of numbers of pairwise differences in the sequence data.

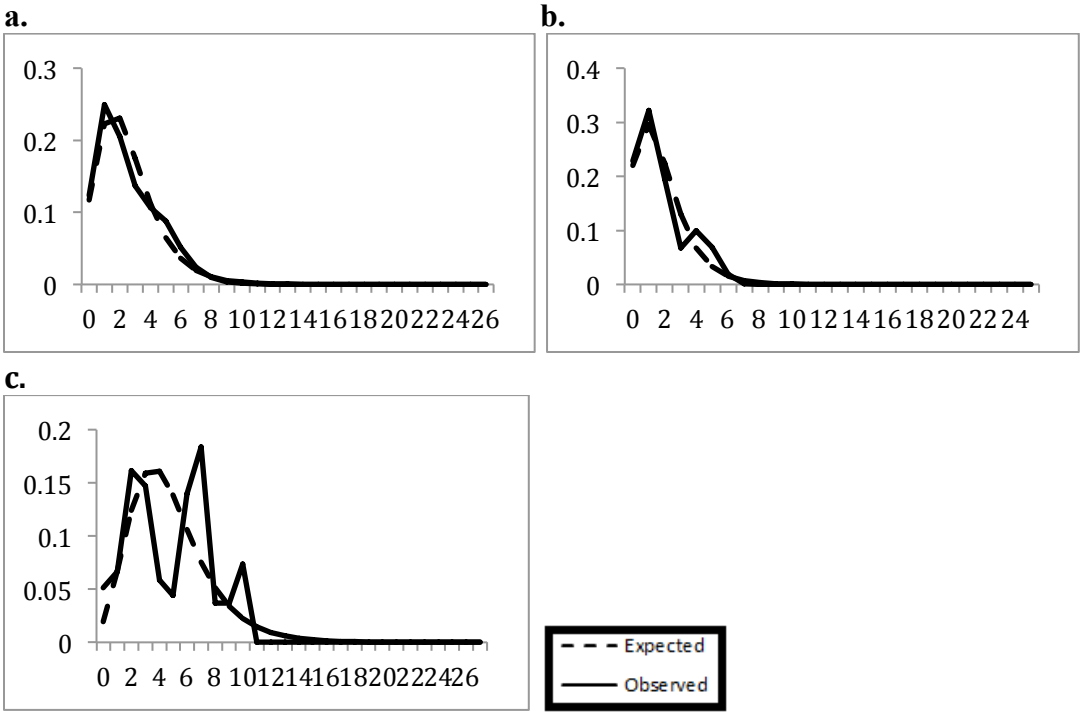

Figure S7. Mismatch distributions of (**a**) AT, (**b**) ME, and (**c**) DA with the CR dataset. The curves plot the frequencies of numbers of pairwise differences in the sequence data.

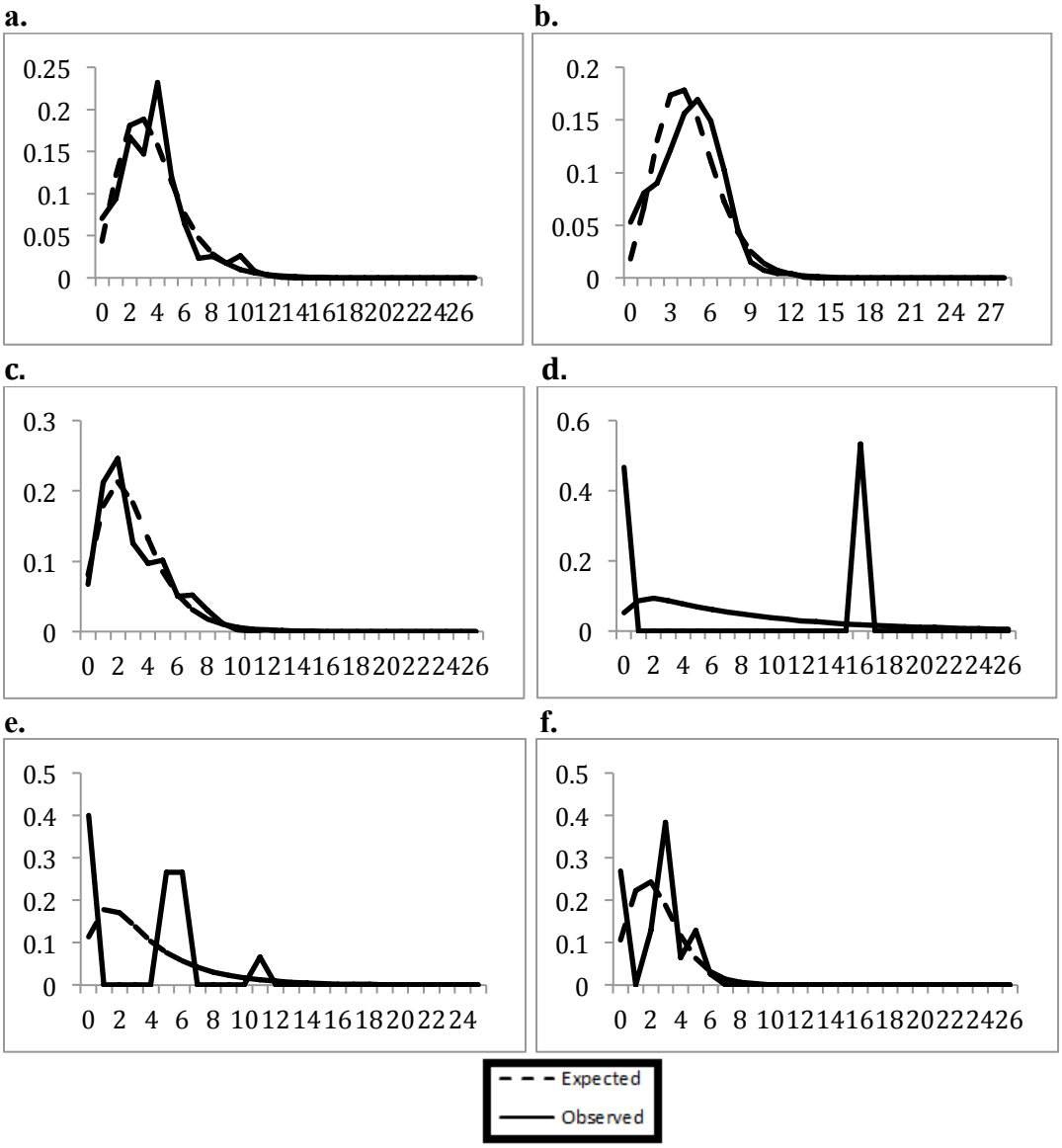

Figure S8. Mismatch distributions of (**a**) AT, (**b**) AD, (**c**) ME, (**d**) DA, (**e**) TI, and (**f**) MA with the concatenated Cytb and CR dataset. The curves plot the frequencies of numbers of pairwise differences in the sequence data.

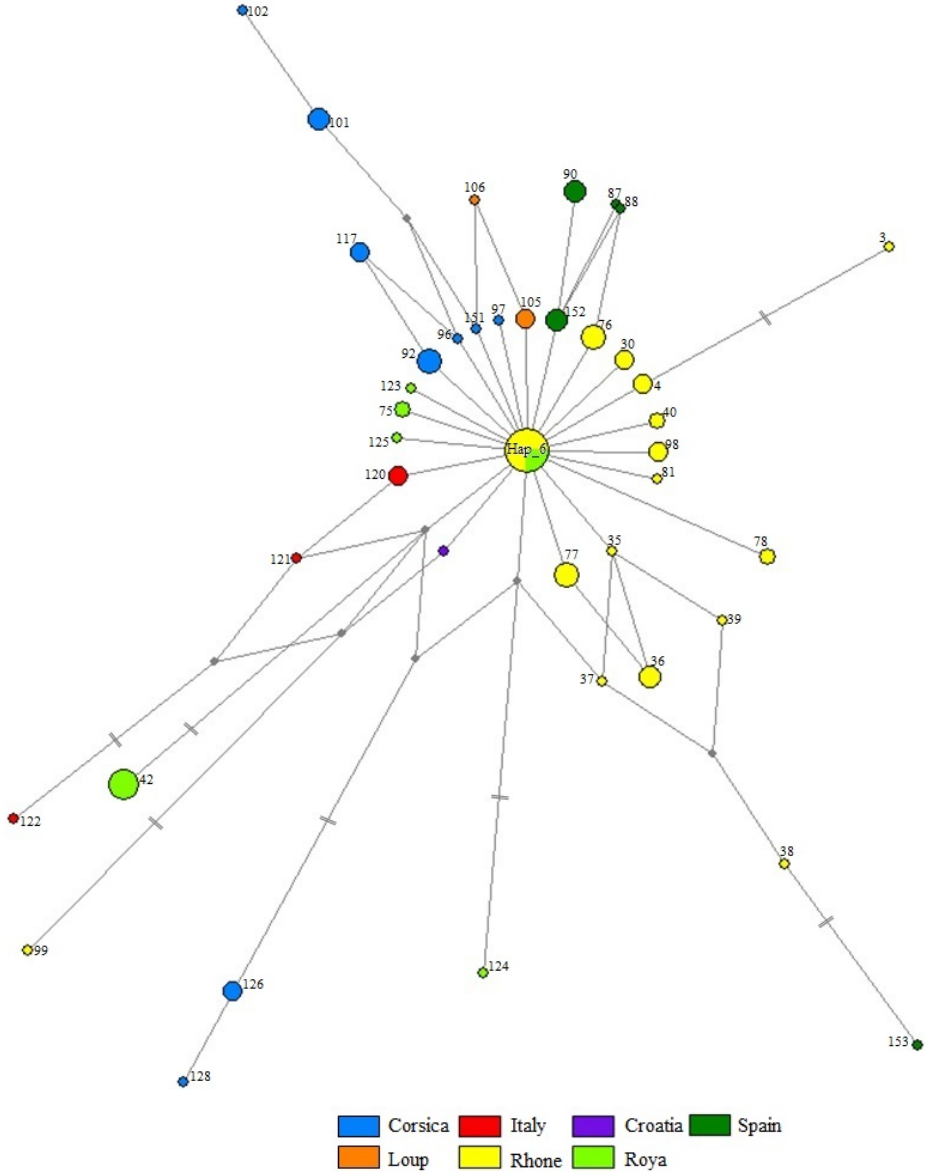

Figure S9. Median-joining network of the concatenated Cytb and CR haplotypes of the Mediterranean (ME) lineage. Numbers indicate haplotypes (See **Appendix II**).

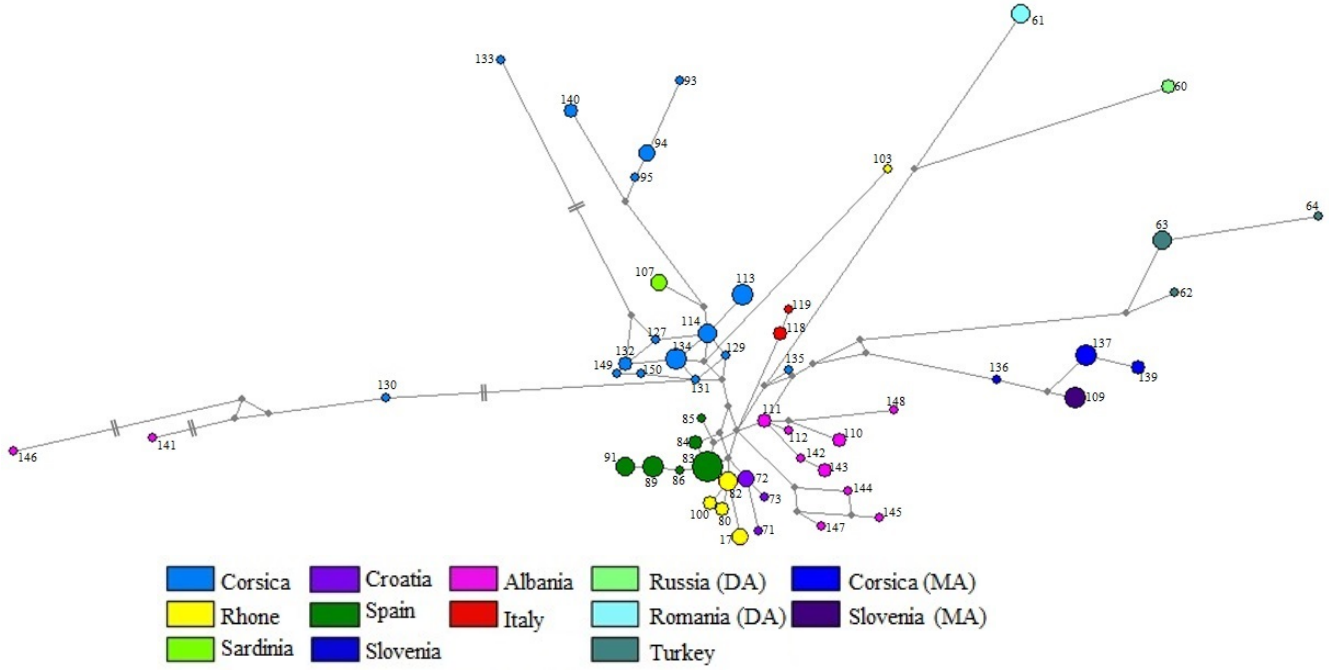

Figure S10. Median-joining network of the concatenated Cytb and CR haplotypes of the Adriatic (AD), Marmoratus (MA), Danubian (DA), and Turkish (TI) lineages. Numbers indicate haplotypes (See **Appendix II**).

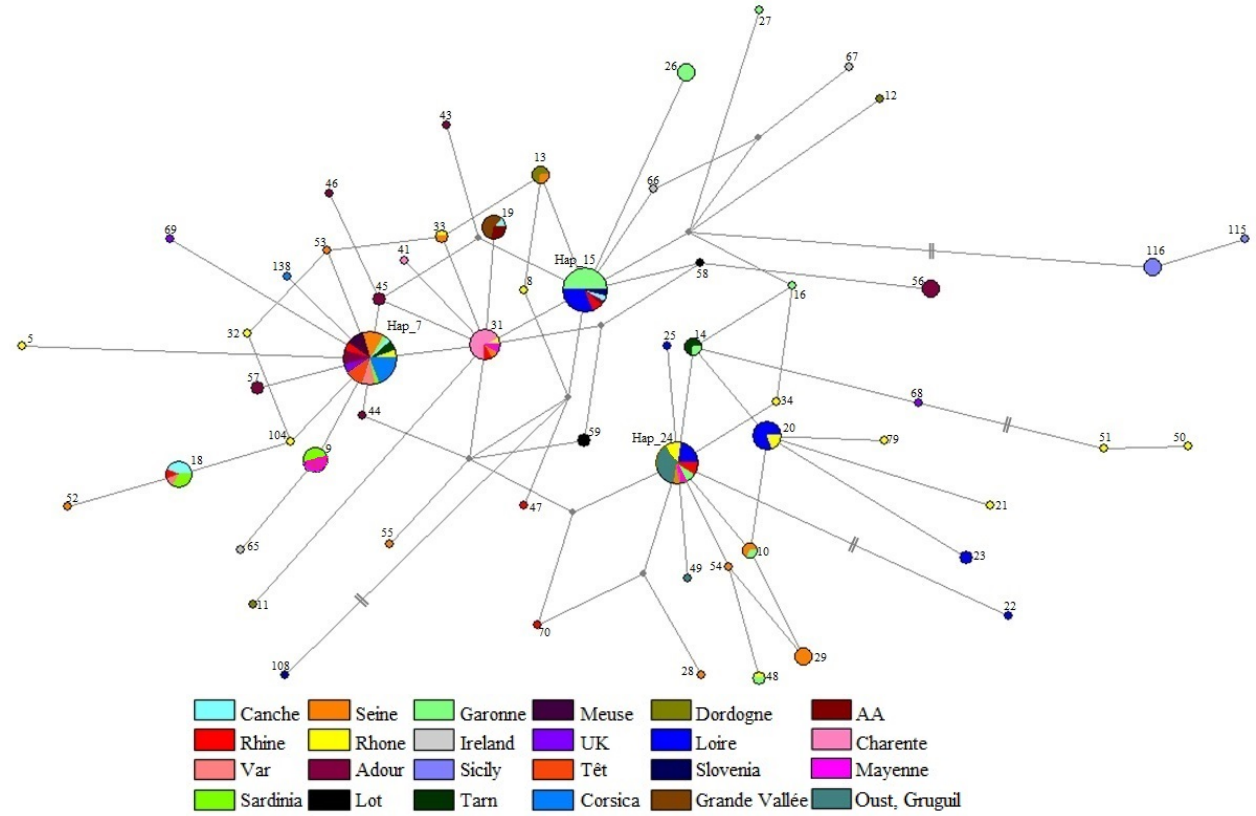

Figure S11. Median-joining network of the concatenated Cytb and CR haplotypes of the Atlantic (AT) lineage. Numbers indicate haplotypes (See **Appendix II**).

Node	TMRCA (my)	Lower 95% CI (my)	Upper 95% CI (my)
1. AD	**0.482**	0.2012	0.7953
2. ME	**0.3694**	0.1461	0.6542
3. MA	**0.1873**	0.0470	0.3649
4. DA	**0.4974**	0.185	0.8481
5. AT	**0.6105**	0.255	0.9996
6. TI	**0.3895**	0.1016	0.7312
7. AD+ME+MA	**0.7192**	0.3176	1.1608
8. AD+ME+MA+DA	**0.9138**	0.41	1.4559
9. AD+ME+MA+DA+AT	**0.9996**	0.4334	1.5726
10. *S. trutta*	**1.1624**	0.5156	1.8729
11. *S. salar*	**0.6337**	0.2181	1.1221
12. *Salmo*	**10.47**	5.834	15.2911

Table S4. Time to most recent common ancestor (TMRCA) estimates for several nodes on the concatenated BI phylogenetic tree with upper and lower 95% confidence intervals, given in millions of years. See **Figure 1** for locations of nodes.

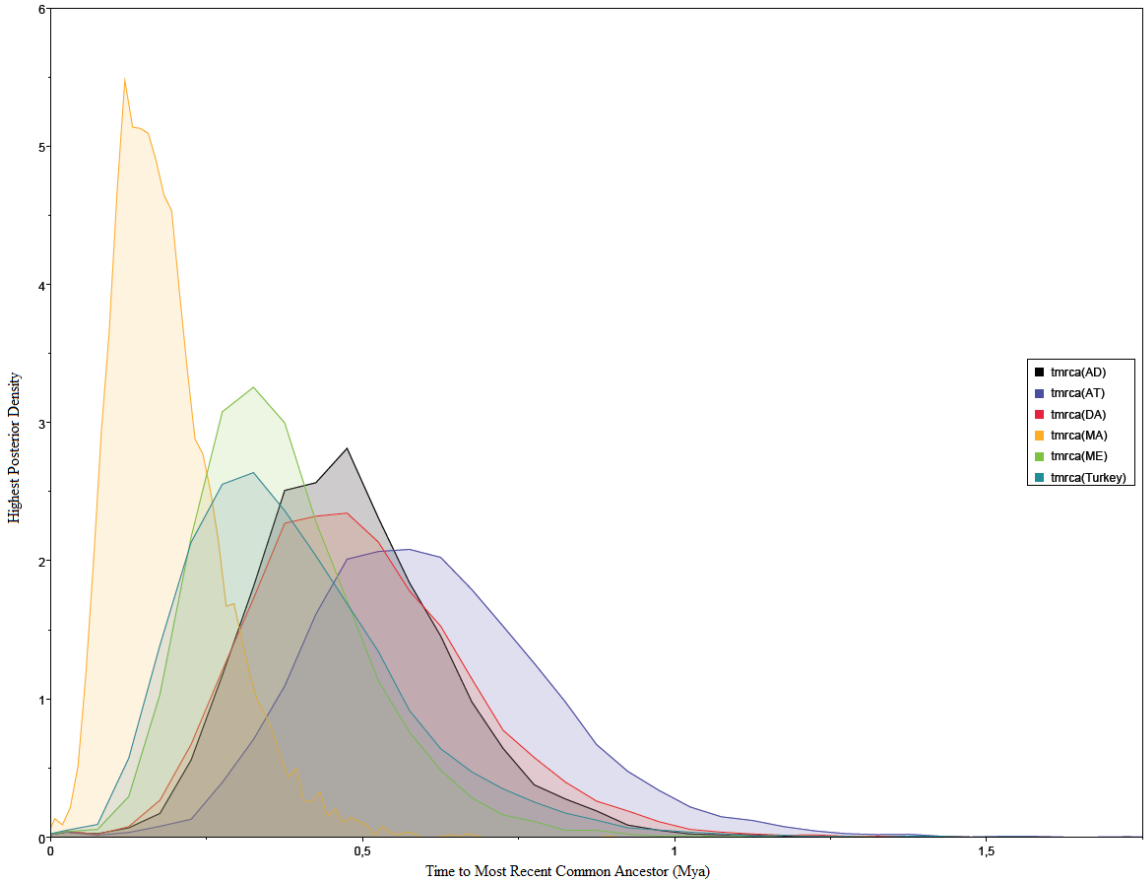

Figure S12. Time to most recent common ancestor of each lineage. Highest posterior density is a measure of estimate confidence.

Appendix II

Lineage	Haplotype (Cytb)	**Haplotype (concatenated)**	Haplotype (CR)	Site	Basin/ Country	Samples (ISEM)	Accession Number (Reference)
AT	Hap_2; AT_12	**N/A**	N/A	N/A	N/A	N/A	EU492282 (Noren *et al*., not published)
		Hap_47	Hap_9	Petit Fecht à Stosswihr	Rhine	T20479	
	Hap_3; AT_11	**N/A**	N/A	N/A	N/A	N/A	FJ435623 (Espineira *et al*., not published)
				N/A	N/A	N/A	D58400 (Matsuda *et al*., not published)
	Hap_4; AT_2	**N/A**	N/A	N/A	N/A	N/A	EU492348 (Noren *et al*., not published); FJ435622 (Espineira *et al*., not published); HQ167696 (Keskin, not published)
		N/A	N/A	N/A	N/A	N/A	
		N/A	N/A	N/A	N/A	N/A	
		Hap_7	Hap_6	Lavezon	Rhône	T12910, T12919	
		N/I	N/I	Llipoudère aval (amont cady)	La Têt	T12919, T13416, T13418, T13419, T13420	
		N/I	N/I	Lentilla amont à valmanya (amont los maos)	La Têt	T13437	
		N/I	N/I	Coumélade aval (amont chapelle st guillem)	Tech	T13498	
		Hap_7	Hap_6	Agout St. Pierre	Tarn	T13556, T13559	
		Hap_15	Hap_9	Garbet Station 7	Garonne	T15278, T15281, T15282	
		Hap_7	Hap_6	Créquoise	Canche	T15681	
		Hap_7, Hap_15	Hap_6, Hap_9	Fouzette	Garonne	T15797, T15810, T15811, T15815	
		Hap_15	Hap_9	Desges	Loire	T15847, T15853, T15857	
		Hap_15	Hap_9	Béthuzon	Garonne	T15894, T15895, T15896, T15897, T15898	

		Hap_15	Hap_9	Chantelouve	Loire	T15922, T15928, T15929, T15933, T15934	
		Hap_7	Hap_6	Aubette	Seine	T16653, T16659	
		Hap_31	Hap_8	Bessay	Rhône	T16725	
		Hap_15	Hap_9	Oulas	Garonne	T16845, T16850, T16858, T16863	
		Hap_31	Hap_8	Touvre	Charente	T16898, T16899, T16910, T16920	
		N/I	N/I	Mouline	La Têt	T17269	
		N/I	N/I	Arn	Garonne	T18303	
		N/I	N/I	Coumelade pont Banat	Tech	T20008	
		Hap_45	Hap_208	Gave de pau à gavarnie	Adour	T20184, T20186	
		Hap_7, Hap_15	Hap_6, Hap_9	Dorlon à Charency-Vézin	Meuse	T20337, T20338, T20339, T20340, T20341	
		Hap_7, Hap_15, Hap_31	Hap_6, Hap_9, Hap_8	Petit Fecht à Stosswihr	Rhine	T20480, T20481, T20482, T20483	
		Hap_31	Hap_8	Argence	Charente	T20502, T20503, T20504, T20505	
		Hap_7, Hap_31	Hap_6, Hap_8	Groëme à Terrefondrée	Seine	T21470, T21471, T21472	
		Hap_31	Hap_8	Egrenne à Beauchêne	Mayenne	T21632	
		Hap_7	Hap_6	Cléry	Seine	T24102	
		N/I	N/I	Boulzane à Caudiès	Agly	T24169	
		Hap_7, Hap_57	Hap_6	Hestapeko erreka	Adour	T24771, T24772, T24773, T24774, T24775	
		Hap_58	Hap_195	Vert	Lot	T25782	
		Hap_7	Hap_6	Camel R	UK	T26073, T26076	
		Hap_7	Hap_6	Sieg R	Rhine	T26080	

		Hap_7	Hap_6	Nohède amont à Pla del Gorg	La Têt	T13457, T13458, T13459, T13461, T13465	
		Hap_104	Hap_11	Vidourle	Rhône	T16109	
		Hap_7	Hap_6	Cians	Var	T16152, T16158, T16161	
		Hap_7	Hap_10	Sadall-Esterzili	Sardinia	T16279	
		Hap_15	Hap_9	Volaja	Slovenia	T16493	
		Hap_7	Hap_6	U Furcone	Corsica	T19959	
		Hap_15	Hap_9	Rocce	Corsica	T08318	
		Hap_7	Hap_6	Bocca Bianca	Corsica	T08335, T08336, T08338, T08343, T08348, T08353	
	Hap_5; AT_13	**N/A**	N/A	N/A	N/A	N/A	EU492108 (Pelt Heerschap *et al*., not published
				N/A	N/A	N/A	EU492109 (Pelt Heerschap *et al*., not published
				N/A	N/A	N/A	FJ435621 (Espineira *et al*., not published
	Hap_7; AT_6	**N/A**	N/A	N/A	N/A	N/A	JX960836 (Crête-Lafrenière *et al*., 2012)
		N/I	N/I	Lentilla amont à valmanya	La Têt	T13440	
		Hap_22	Hap_23	Andrable	Loire	T15953	
		Hap_24, Hap_25	Hap_8, Hap_22	Lignon	Loire	T16016, T16021	
		Hap_24	Hap_8	Vienne	Loire	T16050, T16052, T16056, T16062	
		Hap_24, Hap_34	Hap_8, Hap_9	Bessay	Rhône	T16729, T16731	
		N/I	N/I	Mouline	La Têt	T17268	
		N/I	N/I	Tet à Llagonne (amont) Pla dels Aveillans	La Têt	T18300	

		Hap_24, Hap_34	Hap_8, Hap_9	Gruguil	Gruguil	T20717, T20718, T20719, T20720, T20721	
		Hap_24, Hap_34	Hap_8, Hap_9	Sedon	Oust	T20746, T20747, T20748, T20749, T20750	
		Hap_24	Hap_205	Doubs à L'Abergement SM	Rhône	T21438, T21439	
		Hap_24	Hap_8	Yonne à Arleuf	Seine	T21505	
		Hap_24, Hap_34	Hap_8, Hap_9	Egrenne à Beauchene	Garonne	T21631, T21633	
		Hap_24	Hap_8	Mayenne à Lalacelle	Garonne	T21659, T21660	
		Hap_34	Hap_9	Cléry	Seine	T24101	
		Hap_24	Hap_8	Sieg R	Rhine	T26077, T26078	
	Hap_11	**N/I**	N/I	Slovenia	Slovenia	T11071	-
		Hap_20	Hap_8	Desges	Loire	T15848, T15951, T15956, T15963	
		Hap_20	Hap_8	Lignon	Loire	T16010	
		Hap_20	Hap_8	Loire	Loire	T16373, T16379, T16383, T16391, T16396	
		Hap_20, Hap_21	Hap_8	Gluyère	Rhône	T09713, T09720	
	Hap_12	**Hap_5**	Hap_31	Lavezon	Rhône	T12901	-
	Hap_13	**Hap_8**	Hap_9	Lavezon	Rhône	T12920	-
		N/I	N/I	Mouline	La Têt	T17271	
	Hap_14	**Hap_9**	Hap_6	Dronne	Garonne	T12921, T12923, T12924, T12925	-
				Mayenne à Lalacelle	Garonne	T21658, T21661, T21662	
				Boyne R	Ireland	T26070	
	Hap_15	**Hap_10**	Hap_8	Dronne	Garonne	T12922	-
		Hap_10, Hap_29	Hap_8, Hap_6	Aubette	Seine	T16656, T16657	
		Hap_10, Hap_29	Hap_8, Hap_6	Yonne à arleuf	Seine	T21501, T21502	
		N/I	N/I	Scie à ND du Parc	Scie	T21782	

		Hap_29	Hap_6	Cléry	Seine	T24103, T24104	
	Hap_16	**Hap_11**	Hap_30	amont TCC	Dordogne	T13129	-
	Hap_17	**Hap_12**	Hap_29	amont TCC	Dordogne	T13130	-
	Hap_18	**Hap_13**	Hap_9	amont TCC	Dordogne	T13131, T13132, T13133	-
		N/I	N/I	Lentilla amont à valmanya (amont los maos)	La Têt	T13446	
		N/I	N/I	Coumélade aval (amont chapelle St Guillem)	Tech	T13496	
		Hap_32, Hap_33	Hap_11, Hap_8	Bessay	Rhône	T16726, T16728	
		Hap_13	Hap_9	Groëme à Terrefondrée	Seine	T21468	
		Hap_33, Hap_53	Hap_8, Hap_6	Yonne à Arleuf	Seine	T21503, T21504	
	Hap_19	**N/I**	N/I	Lilipoudre aval (amont cady)	La Têt	T13417	-
		N/I	N/I	Coumélade aval (amont chapelle St Guillem)	Tech	T13499, T13500	
		Hap_18	Hap_11	Créquoise	Canche	T15674	
		N/I	N/I	Hem	AA	T16681, T16688, T16690	
		N/I	N/I	Las Illas amont Maureillas	Tech	T20022, T20024, T20025	
		Hap_44		Gave de pau a gavarnie	Adour	T20183	
		Hap_19		Vert	Lot	T25783, T25784, T25785	
		Hap_18	Hap_11	Sieg R	Rhine, Germany	T26079	
		Hap_18	Hap_11	Cians	Var	T16146	
		Hap_18	Hap_11	Sadall-Esterzili	Sardinia	T16276, T16278, T16280	

		Hap_19, Hap_108	Hap_8, Hap_15	Volaja	Slovenia	T16490	
		N/I	N/I	Gravezon à Joncel	Tarn	T19020, T19022	
	Hap_24	**Hap_14**	Hap_8	Agout st pierre	Tarn	T13557, T13558, T13560	-
		Hap_14, Hap_16	Hap_8, Hap_9	Garbet Station 7	Garonne	T15279, T15280	
	Hap_26	**Hap_19**	Hap_8	Créquoise	Canche	T15696	-
				Grande Vallée	Grande Vallée	T15816, T15821, T15824, T15831	
				Hem	AA	T16686, T16689	
	Hap_27	**Hap_23**	Hap_21	Lignon	Loire	T16014, T16036	-
	Hap_28	**Hap_26**	Hap_19	Lingas	Garonne	T16070, T16071, T16075, T16082	-
	Hap_29	**Hap_27**	Hap_20	Lingas	Garonne	T16073	-
	Hap_30	**Hap_28**	Hap_8	Aubette	Seine	T16654	-
	Hap_38	**Hap_41**	Hap_8	Touvre	Charente	T16909	-
	Hap_41	**N/I**	N/I	Tet à Llagonne (aval) Pla de Barrès	La Têt	T18273	-
	Hap_44	**Hap_43**	Hap_207	Gave de pau a gavarnie	Adour	T20182	-
	Hap_45	**Hap_46**	Hap_208	Gave de pau a gavarnie	Adour	T20185	-
	Hap_46	**Hap_50, Hap_51**	Hap_205, Hap_206	Doubs à L'Abergement SM	Rhône	T21440, T21441, T21442	-
	Hap_47	**Hap_52**	Hap_11	Groëme à Terrefondrée	Seine	T21469	-
		Hap_55	Hap_8	Cléry	Seine	T24105	
	Hap_48	**N/I**	N/I	Scie a ND du Parc	Scie	T21783	-
	Hap_49	**N/I**	N/I	Groëme à Terrefondrée	Seine	T24170	-
	Hap_51	**Hap_56**	Hap_26	Souye	Adour	T24751, T24752, T24753, T24754	-
	Hap_57	**Hap_66**	Hap_9	Boyne R	Ireland	T26071	-
	Hap_58	**Hap_67**	Hap_200	Boyne R	Ireland	T26072	-

	Hap_59	**Hap_68**	Hap_8	Camel R	UK	T26074	-
	Hap_60	**Hap_60**	Hap_196	Camel R	UK	T26075	-
	Hap_61	**Hap_70**	Hap_8	Sieg R	Rhine, Germany	T26081	-
	Hap_65	**Hap_79**	Hap_8	Gluyere	Rhône	T09700	-
	Hap_82	**Hap_115, Hap_116**	Hap_4, Hap_3	Sicily	Sicily	T25009, T25010, T25011, T25012, T25013, T25016	-
	Hap_94	**Hap_138**	Hap_6	Bocca Bianca	Corsica	T08334	-
AD	**Hap_98**	**Hap_153**	**Hap_72**	**Vallauca**	Spain	5	MEcs1
	Hap_6; AD_22	**N/A**	N/A	N/A	N/A	N/A	JX960835 (Crête-Lafrenière *et al*., 2012)
	Hap_20	**N/I**	N/I	Lentilla aval à Valmanya (passerelle Los Maos)	La Têt	T13448	-
	Hap_21	**N/I**	N/I	Lentilla aval à Valmanya (passerelle Los Maos)	La Têt	T13449	-
	Hap_22	**N/I**	N/I	Lentilla aval à Valmanya (passerelle Los Maos)	La Têt	T13450	-
				Lladure à Formiguèr es (aval Estany de Mig)	La Têt	T13489, T13490, T13491	
	Hap_23	**N/I**	N/I	Lladure à Formiguèr es (aval Estany de Mig)	La Têt	T13486	-
	Hap_25	**Hap_17**	Hap_27	Veyer	Rhône	T15554, T15578	-
				Ubaye à Gleizolles	Durance	T04432	-
	Hap_42	**N/I**	N/I	Tet à Llagone (aval) Pla de Barrès	La Têt	T18274	-
		N/I	N/I	Tet à Llagone (amont)	La Têt	T18301	

				Pla dels Aveillans			
		Hap_72	Hap_201	Butiznica R	Croatia	T26084	
		Hap_72, Hap_74	Hap_201, Hap_203	Kosovcica R	Croatia	T26086, T26087, T26088	
		Hap_82	Hap_27	Plampinet	Rhône	T9892, T9893	
		Hap_83	Hap_171; ADcs1	Castril	Spain	2, 5	AY836330 (Cortey *et al.*, 2004)
		Hap_83	Hap_171; ADcs1	Conangles	Spain	7, 8	AY836330 (Cortey *et al.*, 2004)
		Hap_83	Hap_171; ADcs1	Cardos	Spain	3, 5	AY836330 (Cortey *et al.*, 2004)
		Hap_83	Hap_171; ADcs1	Madera	Spain	1, 2, 3, 4, 5	AY836330 (Cortey *et al.*, 2004)
		Hap_100	Hap_32	St.11 pont Batie	Rhône	T12537, T12556	- - -
		Hap_82	Hap_27	Ubaye à Gleizolles	Durance	T04416, T04424	
		N/I	N/I	Gravezon à Joncel	Tarn	T19018	
	Hap_62	**Hap_71**	Hap_201	Butiznica R	Croatia	T26083	-
	Hap_66	**Hap_80**	Hap_27	Plampinet	Rhône	T9889, T9890	-
	Hap_68	**Hap_84**	Hap_171; ADcs1	Castril	Spain	3, 4	AY836330 (Cortey *et al.*, 2004)
	Hap_69	**Hap_85**	Hap_171; ADcs1	Conangles	Spain	10	AY836330 (Cortey *et al.*, 2004)
	Hap_70	**Hap_86**	Hap_171; ADcs1	Congales	Spain	5	AY836330 (Cortey *et al.*, 2004)
		Hap_89	Hap_181; ADcs19	Dilar	Spain	1, 2, 3, 4, 5	AY836348 (Cortey *et al.*, 2004)
		Hap_90	ADcs5	Trévelez	Spain	1, 2, 3, 4	AY836334 (Cortey *et al.*, 2004)
	Hap_73	**Hap_93**	Hap_40	Lette	Corsica	T10588	-
	Hap_74	**Hap_94, Hap_95**	Hap_40, Hap_41	Lette	Corsica	T10592, T10597, T10602, T10606	-
		Hap_114	Hap_5	U Furcone	Corsica	T19958, T19960	
		Hap_114, Hap_134	Hap_5, Hap_46	Manica	Corsica	T08134, T08138,	

						T08143, T08153	
		Hap_114, Hap_134	Hap_5, Hap_46	Rocce	Corsica	T08328, T08333	
		Hap_134	Hap_46	Carnevale	Corsica	T08355	
	Hap_77	**Hap_103**	Hap_6	Vidourle	Rhône	T16100	-
	Hap_78	**Hap_107**	Hap_16	Is Abius (Camboni)	Sardinia	T16246, T16247, T16248	-
	Hap_80	**Hap_110, Hap_111, Hap_112**	Hap_14, Hap_13, Hap_12	Lake Ohrid	Albania	T16585, T16586, T16587, T16588, T16589	-
		Hap_141, Hap_142, Hap_143	Hap_54, Hap_53, Hap_52	Shkumbini	Albania	T09059, T09060, T09061, T09065, T09067, T09069, T09072, T09074	
		Hap_144, Hap_145	Hap_51, Hap_50	Valbona-Dragobi	Albania	T09161, T09163	
		Hap_146	Hap_49	Valbona	Albania	T09165	
		Hap_147	Hap_48	Valbona-Gashi	Albania	T09167	
		Hap_148	Hap_47	Valbona-Drini	Albania	T09169	
	Hap_81	**Hap_113**	Hap_7	Lataga	Corsica	T18605, T18607, T18609, T18611, T18613	-
	Hap_83	**Hap_118, Hap_119**	Hap_66, Hap_67	Fibreno	Italy	T05727, T05729, T05733, T05741, T05751, T05761	-
	Hap_89	**Hap_127, Hap_132, Hap_133**	Hap_63, Hap_58, Hap_59	Chjuvone	Corsica	T07712	-
		N/I	N/I	Veraculungu (Cuscione)	Corsica	T08117, T08118, T08119	
	Hap_90	**Hap_129, Hap_130, Hap_131**	Hap_5, Hap_60, Hap_46	Uccialinu	Corsica	T08015, T08016, T08017, T08018, T08019	-
	Hap_91	**Hap_135**	Hap_46	Manica	Corsica	T08148	-
	Hap_95	**Hap_140**	Hap_56	Haut Borato	Corsica	T08403, T08408	-

				N2000			
	Hap_96	**Hap_149**	Hap_46	Marmanu	Corsica	T09212	-
	Hap_97	**Hap_150**	Hap_46	Marmanu	Corsica	T09214, T09215, T09216	-
ME	Hap_9; ME_2	**N/A**	N/A	N/A	N/A	N/A	JX960839 (Crête-Lafrenière *et al.*, 2012); AY836350 (Cortey *et al.*, 2004); AY836364 (Cortey *et al.*, 2004);
		Hap_152	Hap_72; MEcs1	Vallauca	Spain	1, 2, 3, 4	
		Hap_88	Hap_80; MEcs15	Cardos	Spain	4	
		N/I	N/I	Drome	Rhône	T16806, T16821	
		N/I	N/I	Loup	Loup	T16143	
		N/I	N/I	Zoico	Corsica	T09352, T09357, T09361	
		Hap_6	Hap_10	Lavezon	Rhône	T12905	
		N/I	N/I	Lentilla amont a valmanya (amont los maos)	La Têt	T13436, T13438, T13439	
		N/I	N/I	Lentilla aval à Valmanya (passerelle Los Maos)	La Têt	T13447	
		N/I	N/I	Lladure à Formiguèr es (amont refuge Lladura)	Aude	T13476, T13477, T13478, T13479, T13480	
		N/I	N/I	Coumélade aval (amont chapelle St Guillem)	Tech	T13497	
		Hap_6	Hap_10	Veyer	Rhône	T15549, T15551, T15581	
		Hap_6	Hap_10	Riotet	Rhône	T15984, T15991, T15994, T15998, T16007	
		Hap_6	Hap_10	Drome	Rhône	T16797, T16803, T16805	
		N/I	N/I	Llech à Estoher	La Têt	T17252, T17255	
		N/I	N/I	Mouline	La Têt	T17270	
		Hap_42	Hap_33	Breil pont Arbouset	Roya	T17284, T17285,	

						T17286, T17287, T17288	
		N/I	N/I	Tet à Llagone (aval) Pla de Barrès	La Têt	T18275	
		N/I	N/I	Tet à Llagone (amont) Pla dels Aveillans	La Têt	T18299	
		N/I	N/I	Bruyante aux anciennes forges	Aude	T18823, T18827	
		N/I	N/I	Coumélade pont Banat	Tech	T20007, T20009, T20010, T20011	
		N/I	N/I	Boulzane a Caudies	Agly	T24167, T24168	
		Hap_73	Hap_202	Butiznica R	Croatia	T26085	
		Hap_6, Hap_75	Hap_10, Hap_55	Sallevieille	Roya	T08844, T08850, T08856, T08861, T08869	
		Hap_76	Hap_45	Sumène - Pt St Prix	Rhône	T09393, T09398, T09403, T09409, T09418	
		Hap_92	Hap_42	Radule	Corsica	T10159, T10160, T10161, T10162, T10163	
		Hap_96, Hap_97	Hap_39, Hap_38	Corbica	Corsica	T10788, T10792	
		Hap_98, Hap_99	Hap_34, Hap_27	Biaysse amont	Rhône	T11559, T11566, T11571, T11585	
		Hap_42	Hap_33	Fontan (aval)	Roya	T11639, T11647	
		Hap_105	Hap_17	Loup	Loup	T16126, T16131, T16140	
		Hap_117	Hap_69	Haut Golu (Valdoniellu)	Corsica	T03088, T03089, T03091	
		Hap_6	Hap_10	Fontaine de	Roya	T07426	

				Vaucluse			
		N/I	N/I	Gravezon à Joncel	Tarn	T19024, T19026	
	Hap_10	**Hap_3, Hap_4**	Hap_37, Hap_10	Thine	Rhône	T11044, T11050, T11055, T11061	
	Hap_31	**Hap_30**	Hap_10	Mouge	Rhône	T16699, T16704, T16718	
	Hap_32	**Hap_35**	Hap_10	Albarine	Rhône	T16759	-
	Hap_33	**Hap_36**	Hap_10	Albarine	Rhône	T16763, T16764	-
	Hap_34	**Hap_37**	Hap_10	Ouvèze	Rhône	T16825	-
	Hap_35	**Hap_38**	Hap_10	Ouvèze	Rhône	T16827	-
	Hap_36	**Hap_39**	Hap_10	Ouvèze	Rhône	T16828	-
	Hap_37	**Hap_40**	Hap_10	Ouvèze	Rhône	T16829, T16834	-
	Hap_39	**N/I**	N/I	Carança (réserve amont)	La Têt	T17222, T17223, T17224	-
		N/I	N/I	Llech à Estoher	La Têt	T17253, T17254, T17256	
		N/I	N/I	Bruyante aux anciennes forges	Aude	T18822	
		N/I	N/I	Les Illas amont Maureillas	Tech	T20023, T20026	
	Hap_40	**N/I**	N/I	Eyne aval Orri de Baix	Sègre	T17815, T17816, T17817, T17818, T17819	-
	Hap_43	**N/I**	N/I	Tet a Llagonne (amont) Pla dels Aveillans	La Têt	T18302	-
	Hap_50	**N/I**	N/I	Boulzane a Caudies	Agly	T24171	-
	Hap_63	**Hap_77**	Hap_10	Ardeche Amt - Pt Mercier	Rhône	T09512, T09519, T09526, T09533, T09540	-
	Hap_64	**Hap_78**	Hap_44	Gluyère	Rhône	T09694, T09707	-
	Hap_67	**Hap_81**	Hap_10	Névache	Rhône	T09891	-
	Hap_71	**Hap_87**	Hap_72; MEcs1	Cardos	Spain	2	AY836350 (Cortey *et al.*, 2004)

	Hap_72	**Hap_90**	Hap_72; MEcs1	Endrinales	Spain	1, 2, 4, 6	AY836350 (Cortey *et al.*, 2004)
	Hap_75	**Hap_101**	Hap_28	Paratella	Corsica	T15508, T15509, T15510, T15512	-
	Hap_76	**Hap_102**	Hap_28	Paratella	Corsica	T15511	-
	Hap_84	**Hap_120, Hap_121, Hap_122**	Hap_10, Hap_65, Hap_64	Volturno	Italy	T05787, T05788, T05789, T05790, T05791	-
	Hap_85	**Hap_123**	Hap_10	Fontaine de Vaucluse	Roya	T07424	-
	Hap_86	**Hap_124**	Hap_10	Fontaine de Vaucluse	Roya	T07425	-
	Hap_87	**Hap_125**	Hap_10	Fontaine de Vaucluse	Roya	T07427	-
	Hap_88	**Hap_126, Hap_128**	Hap_61, Hap_62	Chjuvone	Corsica	T07703, T07707, T07717, T07721	-
DA	Hap_8; MA_4	N/A	N/A	N/A	N/A	N/A	JX960837 (Crête-Lafrenière *et al.*, 2012)
	Hap_52	Hap_60	Hap_196	Chirak R	Russia	T26053, T26054	-
	Hap_53	Hap_61	Hap_197	Agres R	Romania	T26055, T26056, T26057, T26058	-
MA	Hap_79	Hap_109	Hap_15	Volaja	Slovenia	T16492, T16494, T16495, T16496	-
	Hap_53	Hap_109 Hap_137	Hap_15 Hap_15	Trebuscica -1	Slovenia	T06836	-
	Hap_93	Hap_109 Hap_137 Hap_137, Hap_139	Hap_15 Hap_15 Hap_15, Hap_57	Paratella	Corsica	T08198, T08199, T08200, T08201	-
				Carnevale	Corsica	T08356, T08357, T08361, T08362, T08363	-
	Hap_92	Hap_136	Hap_15	Paratella	Corsica	T08197	-
TI	Hap_54	Hap_62	Hap_198	Gatak R	Turkey	T26061	-

	Hap_55	Hap_63	Hap_198	Gatak R	Turkey	T26062, T26064, T26065, T26066	-
	Hap_56	Hap_64	Hap_198	Gatak R	Turkey	T26063	-

Table E1. Samples used in the Cytb and concatenated datasets with accession numbers and references for GenBank sequences. Concatenated haplotypes are highlighted, with corresponding Cytb and CR haplotypes when available. N/A indicates missing information for GenBank haplotypes; N/I indicates that the samples are not included in the concatenated and CR datasets; - indicates that the haplotype is not on GenBank.

Lineage	Haplotype	Site	Basin/Country	Samples (ISEM)	Accession number (reference)
AT	Hap_3	Anapo à Palazzolo Acreide	Sicily	T25010-T25013	-
	Hap_4	Anapo à Palazzolo Acreide	Sicily	T25009	-
	Hap_6; ER127	N/A	N/A	N/A	AF253541 (Suarez *et al*., 2001)
		U Furcone	Corsica	T19959	
		Aubette	Seine	T16653, T16657, T16659	
		Sadali-Esterzili	Corsica	T16279	
		Cians	Var	T16152, T16158, T16161, T16164	
		Vidourle	Rhone	T16100	
		Fouzette	Garonne	T15810	
		Créquoise	Canche	T15681	
		Agout St Pierre	Tarn	T13556, T13559	
		Nohède amont à Pla del Gorg	La Têt	T13457, T13459, T13461, T13465	
		Lavezon	Rhone	T12910, T12919	
		Bocca Bianca	Corsica	T08334-T08336, T08338, T08343, T08348, T08353	
		Hestapeko erreka	Adour	T24773-T24775	
		Camel R	UK	T26075,	

				T26076	
		Sieg R	Rhine, Germany	T26080	
		Dronne	Dordogne	T12921, T12923-T12925	
		Petit Fecht à Strosswihr	Rhine	T20481	
		Yonne à Arleuf	Seine	T21501, T21504	
		Egrenne à Beauchêne	Mayenne	T21629	
		Mayenne à Lalacelle	Mayenne	T21658, T21661, T21662	
		Cléry	Seine	T24101-T24104	
		Groëme à Terrefondrée	Seine	T21471, T21472	
		Dorlon à Charency-Vézin	Meuse	T20337, T20338, T20340, T20341	
	Hap_8; U54	N/A	N/A	N/A	AF253543 (Suarez *et al*., 2001)
		Touvre	Charente	T16898, T16899, T16906, T16910, T16920	
		Hem	AA	16686, 16689	
		Aubette	Seine	T16654, T16656	
		Loire (Cros Romeau)	Loire	T16373, T16379, T16383, T16391, T16396	
		Sadali-Esterzili	Corsica	T16281, T16288	
		Vidourle	Rhone	T16101, T16114	
		Vienne	Loire	T16039, T16050, T16052, T16056, T16062	
		Lignon	Loire	T16060, T16016	
		Andrable	Loire	T15956, T15960, T15963	
		Desges	Loire	T15848	
		Grande Vallée	Grande Vallée	T15816, T15821, T15824,	

				T15831, T15835, T15848	
		Créquoise	Canche	T15696	
		Garbet St 7	Garonne	T15280	
		Agout St Pierre	Tarn	T13557, T13558, T13560	
		Nohède amont à Pla del Gorg	La Têt	T13456	
		Fontan (aval)	Roya	T11632	
		Gluyère	Rhone	T09700, T09713, T09720	
		Camel R	UK	T26074	
		Sieg R	Rhine, Germany	T26077, T26078, T26081	
		Dronne	Dordogne	T12922	
		Bessay	Rhone	T16725, T16728, T16729	
		Petit Fecht à Strosswihr	Rhine	T20480	
		Goyen	Goyen	T20620-T20622, T20624	
		Gruguil	Gruguil	T20717, T20718, T20720, T20721	
		Sedon	Oust	T20746, T20747, T20749, T20750	
		Doubs à l'Abergement SM	Rhone	T21438, T21439	
		Yonne à Arleuf	Seine	T21502, T21503, T21505	
		Egrenne à Beauchêne	Mayenne	T21630-T21632	
		Mayenne à Lalacelle	Mayenne	T21659, T21660	
		Cléry	Seine	T24105	
		Argence		T20501-T20505	
		Groëme à Terrefondrée	Seine	T21470	

	Hap_9	Oulas		T16841, T16845, T16850, T16858, T16863	-
		Hem	AA	T16677	-
		Volaja	Slovenia	T16491, T16493	-
		Chantelouve		T15922, T15928, T15929, T15933, T15934	-
		Béthuzon		T15893, T15894, T15896, T15898, T15899	-
		Desges	Loire	T15844, T15847, T15853, T15857	-
		Fouzette	Garonne	T15797, T15802, T15811, T15815	-
		Garbet St 7	Garonne	T15278, T15279, T15281, T15282	-
		amont TCC	Dordogne	T13131-T13133	-
		Lavezon	Rhone	T12920	-
		Rocce	Corsica	T08318	-
		Boyne R	Ireland	T26071	-
		Bessay	Rhone	T16731	-
		Petit Fecht à Strosswihr	Rhine	T20479, T20482, T20483	-
		Groëme à Terrefondrée	Seine	T21468	-
		Dorlon à Charency-Vézin	Meuse	T20339	-
	Hap_11; T7	N/A	N/A	N/A	AF253542 (Suarez *et al*., 2001)
		Hem	AA	T16681, T16688, T16690	
		Sadali-Esterzili	Corsica	T16276-T16278, T16280, T16282, T16289	

		Cians	Var	T16146	
		Vidourle		T16109	
		Créquoise	Canche	T15674	
		Sieg R	Rhine, Germany	T26079	
		Bessay	Rhone	T16726	
		Gruguil	Gruguil	T20719	
		Egrenne à Beauchêne	T21633	-	
		Groëme à Terrefondrée	Groëme à Terrefondrée	T21469	
	Hap_19	Linguat	Garonne	T16070, T16071, T16075, T16082	-
	Hap_20	Linguat	Garonne	T16073	-
	Hap_21	Lignon	Loire	T16014, T16036	-
	Hap_22	Lignon	Loire	T16021	-
	Hap_23	Andrable	Loire	T15953	-
	Hap_24	Andrable	Loire	T15951	-
	Hap_25	Créquoise	Canche	T15680	-
	Hap_26	Créquoise	Canche	T15668	-
		Souye	Adour	T24751-T24754	
		Vert	Lot	T25782	
	Hap_29	amont TCC	Dordogne	T13130	-
	Hap_30	amont TCC	Dordogne	T13129	-
	Hap_31	Lavezon	Rhone	T12901	-
	Hap_36	Thine	Rhone	T11067	-
	Hap_118; T2	N/A	N/A	N/A	AF253554 (Suarez *et al*., 2001)
	Hap_119; PIG5	N/A	N/A	N/A	AF253559 (Suarez *et al*., 2001)
	Hap_120; ATcs25	N/A	N/A	N/A	EF530487 (Cortey *et al*., 2009)
	Hap_121; STMAR2	N/A	N/A	N/A	AF253556 (Suarez *et al*., 2001)
	Hap_122; LM20	N/A	N/A	N/A	AF253555 (Suarez *et al*., 2001)
	Hap-122; ATcs23	N/A	N/A	N/A	EF530485 (Cortey *et al*., 2009)
	Hap_123; JE1	N/A	N/A	N/A	AF253557 (Suarez *et al*., 2001)
	Hap_124; JA387	N/A	N/A	N/A	AF253553 (Suarez *et al*., 2001)
	Hap-124; ATcs32	N/A	N/A	N/A	EF530494 (Cortey *et al*., 2009)
	Hap_125; ATcs52	N/A	N/A	N/A	EF530512 (Cortey *et al*., 2009)

	Hap_126; ATcs51	N/A	N/A	N/A	EF530511 (Cortey *et al*., 2009)
	Hap_127; ATcs50	N/A	N/A	N/A	EF530510 (Cortey *et al*., 2009)
	Hap_128; ATcs49	N/A	N/A	N/A	EF530509 (Cortey *et al*., 2009)
	Hap_129; ATcs48	N/A	N/A	N/A	EF530508 (Cortey *et al*., 2009)
	Hap_130; ATcs47	N/A	N/A	N/A	EF530507 (Cortey *et al*,. 2009)
	Hap_131; ATcs46	N/A	N/A	N/A	EF530506 (Cortey *et al*., 2009)
	Hap_132; ATcs45	N/A	N/A	N/A	EF530505 (Cortey *et al*., 2009)
	Hap_133; ATcs43	N/A	N/A	N/A	EF530504 (Cortey *et al*., 2009)
	Hap_134; ATcs42	N/A	N/A	N/A	EF530503 (Cortey *et al*., 2009)
	Hap_135; ATcs41	N/A	N/A	N/A	EF530502 (Cortey *et al*., (2009)
	Hap_136; ATcs39	N/A	N/A	N/A	EF530501 (Cortey *et al*., 2009)
	Hap_137; ATcs38	N/A	N/A	N/A	EF530500 (Cortey *et al*., 2009)
	Hap_138; ATcs37	N/A	N/A	N/A	EF530499 (Cortey *et al*., 2009)
	Hap_139; ATcs33	N/A	N/A	N/A	EF530495 (Cortey *et al*., 2009)
	Hap_139; ATcs36	N/A	N/A	N/A	EF530498 (Cortey *et al*., 2009)
	Hap_140; ATcs35	N/A	N/A	N/A	EF530497 (Cortey *et al*., 2009)
	Hap_141; ATcs34	N/A	N/A	N/A	EF530496 (Cortey *et al*., 2009)
	Hap_142; ATcs31	N/A	N/A	N/A	EF530493 (Cortey *et al*., 2009)
	Hap_143; ATcs30	N/A	N/A	N/A	EF530492 (Cortey *et al*., 2009)
	Hap_144; ATcs29	N/A	N/A	N/A	EF530491 (Cortey *et al*., 2009)
	Hap_145; ATcs28	N/A	N/A	N/A	EF530490 (Cortey *et al*., 2009)
	Hap_146; ATcs27	N/A	N/A	N/A	EF530489 (Cortey *et al*., 2009)
	Hap_147; ATcs26	N/A	N/A	N/A	EF530488 (Cortey *et al*., 2009)
	Hap_148; ATcs24	N/A	N/A	N/A	EF530486 (Cortey *et al*., 2009)
	Hap_149; ATcs22	N/A	N/A	N/A	EF530484 (Cortey *et al*., 2009)
	Hap_150; AT2	N/A	N/A	N/A	AF273087 (Cortey & Garcia-Marin, not published)
	Hap-150; ATcs21	N/A	N/A	N/A	EF530483 (Cortey *et al*., 2009)

	Hap_151; ATcs20	N/A	N/A	N/A	EF530482 (Cortey *et al*., 2009)
	Hap_152; ATcs19	N/A	N/A	N/A	EF530481 (Cortey *et al*., 2009)
	Hap_153; ATcs18	N/A	N/A	N/A	EF530480 (Cortey *et al*., 2009)
	Hap_154; ATcs17	N/A	N/A	N/A	EF530479 (Cortey *et al*., 2009)
	Hap_155; ATcs16	N/A	N/A	N/A	EF530478 (Cortey *et al*., 2009)
	Hap_156; ATcs15	N/A	N/A	N/A	EF530477 (Cortey *et al*., 2009)
	Hap_157; ATcs14	N/A	N/A	N/A	EF530476 (Cortey *et al*., 2009)
	Hap_158; ATcs13	N/A	N/A	N/A	AY836329 (Cortey *et al.*2004)
	Hap_159; ATcs12	N/A	N/A	N/A	AY836328 (Cortey *et al.*2004)
	Hap-159; AT10	N/A	N/A	N/A	AY185577 (Duftner *et al*., 2003)
	Hap_160; ATcs11	N/A	N/A	N/A	AY836327 (Cortey *et al.,* 2004)
	Hap_161; AT6	N/A	N/A	N/A	AF274577 (Cortey & Garcia-Marin, not published)
	Hap_162; AT5	N/A	N/A	N/A	AF274576 (Cortey & Garcia-Marin, not published)
	Hap_163; AT4	N/A	N/A	N/A	AF274575 (Cortey & Garcia-Marin, not published)
	Hap_164; AT3	N/A	N/A	N/A	AF274574 (Cortey & Garcia-Marin, not published)
	Hap_165; AT1f	N/A	N/A	N/A	DQ841193 (Meraner *et al*., 2007)
	Hap_166; AT1e	N/A	N/A	N/A	-
	Hap_167; AT11b	N/A	N/A	N/A	-
	Hap_168; AT11a	N/A	N/A	N/A	-
	Hap_169; AT1, ATDU246	N/A	N/A	N/A	-
	Hap_195	Hestapeko erreka	Adour	T24771, T24772	-
	Hap_199	Boyne R	Ireland	T26070	-
	Hap_200	Boyne R	Ireland	T26072	-
	Hap_204	Sedon	Oust	T20748	-
	Hap_205	Doubs à l'Abergement SM	Rhone	T21440	-
	Hap_206	Doubs à	Rhone	T21441	-

		l'Abergement SM			
	Hap_207	Gave de Pau à Gavarnie	Adour	T20182	-
	Hap_208	Gave de Pau à Gavarnie	Adour	T20183-T20186	-
	Hap_209	Vert	Lot	T25784, T25785	-
DU	Hap_88; TI2	N/A	N/A	N/A	AF253545 (Suarez *et al*., 2001)
	Hap_88; DUcs6	N/A	N/A	N/A	EF530518 (Cortey *et al*., 2009)
	Hap_88; DUcs5	N/A	N/A	N/A	EF530517 (Cortey *et al*., 2009)
	Hap_89; DUcs9	N/A	N/A	N/A	EF530521 (Cortey *et al*., 2009)
	Hap-89; DUcs4	N/A	N/A	N/A	EF530516 (Cortey *et al*., 2009)
	Hap_89; DUcs2	N/A	N/A	N/A	EF530514 (Cortey *et al*., 2009)
	Hap_89; AT8	N/A	N/A	N/A	AF273088 (Cortey & Garcia-Marin, not published)
	Hap_89; AT10_AF	N/A	N/A	N/A	AF274580 (Cortey & Garcia-Marin, not published)
	Hap_90; DUcs8	N/A	N/A	N/A	EF530520 (Cortey *et al*., 2009)
	Hap_91; DUcs7	N/A	N/A	N/A	EF530519 (Cortey *et al*., 2009)
	Hap_91; DUcs3	N/A	N/A	N/A	EF530515 (Cortey *et al*., 2009)
	Hap_91; AT9	N/A	N/A	N/A	AF274579 (Cortey & Garcia-Marin, not published)
	Hap_92; DUcs23	N/A	N/A	N/A	EF530535 (Cortey *et al*., 2009)
	Hap_93; DUcs22	N/A	N/A	N/A	EF530534 (Cortey *et al*., 2009)
	Hap_94; DUcs21	N/A	N/A	N/A	EF530533 (Cortey *et al*., 2009)
	Hap_95; DUcs20	N/A	N/A	N/A	EF530532 (Cortey *et al*., 2009)
	Hap_95; DUcs13	N/A	N/A	N/A	EF530525 (Cortey *et al*., 2009)
	Hap_96; DUcs19	N/A	N/A	N/A	EF530531 (Cortey *et al*., 2009)
	Hap_96; AT7	N/A	N/A	N/A	AF274578 (Cortey & Garcia-Marin, not published)
	Hap_97; DUcs18	N/A	N/A	N/A	EF530530 (Cortey *et al*., 2009)

	Hap_98; DUcs17	N/A	N/A	N/A	EF530529 (Cortey *et al*., 2009)
	Hap_99; DUcs16	N/A	N/A	N/A	EF530528 (Cortey *et al*., 2009)
	Hap_100; DUcs15	N/A	N/A	N/A	EF530527 (Cortey *et al*., 2009)
	Hap_101; DUcs14	N/A	N/A	N/A	EF530526 (Cortey *et al*., 2009)
	Hap_102; DUcs12	N/A	N/A	N/A	EF530524 (Cortey *et al*., 2009)
	Hap_103; DUcs11	N/A	N/A	N/A	EF530523 (Cortey *et al*., 2009)
	Hap_104; DUcs10	N/A	N/A	N/A	EF530522 (Cortey *et al*., 2009)
	Hap_105; DUcs1	N/A	N/A	N/A	EF530513 (Cortey *et al*., 2009)
DA	Hap_106; CE365	N/A	N/A	N/A	AF253544 (Suarez *et al*., 2001)
	Hap_107; DA9a	N/A	N/A	N/A	GQ222380 (Jadan *et al*., not published)
	Hap_108; DA9	N/A	N/A	N/A	AY185572 (Duftner *et al*., 2003)
	Hap_109; DA3	N/A	N/A	N/A	AY185571 (Duftner *et al*., 2003)
	Hap_110; DA26	N/A	N/A	N/A	DQ841194 (Meraner *et al*., 2007)
	Hap_111; DA24	N/A	N/A	N/A	AY185576 (Duftner *et al*., 2003)
	Hap_112; DA23b	N/A	N/A	N/A	AY185575 (Duftner *et al*. (2003)
	Hap_113; DA23a	N/A	N/A	N/A	AY185574 (Duftner *et al*., 2003)
	Hap_114; DA22	N/A	N/A	N/A	AY185573 (Duftner *et al*. (2003)
	Hap_115; DA2	N/A	N/A	N/A	AY185570 (Duftner *et al*., 2003)
	Hap_116; DA1b	N/A	N/A	N/A	AY185569 (Duftner *et al*. (2003)
	Hap_117; DA1a	N/A	N/A	N/A	AY185568 (Duftner *et al*., 2003)
	Hap_196	Chirak R	Russia	T26053, T26034	-
	Hap_197	Agres R	Romania	T26055, T26056, T26057, T26058	-
TU	Hap_198	Gatak R	Turkey	T26061, T26062, T26063, T26064, T26065, T26066	-
AD	Hap_5	U Furcone	Corsica	T19957, T19958,	-

				T19960, T19961	
		Sadali-Esterzili	Corsica	T16290	
		Rocce	Corsica	T08328	
		Manica	Corsica	T08143	
		Uccialinu	Corsica	T08017	
	Hap_7	Lataga	Corsica	T18605, T18607, T18609, T18611, T18613	-
	Hap_12	Lac Ohrid	Albania	T16584, T16587, T16589	-
		piscicultur e Lac Ohrid	Albania	T08924, T08925, T08926, T08928	
	Hap_13	Lac Ohrid	Albania	T16588	-
	Hap_14	Lac Ohrid	Albania	T16585, T16586	-
	Hap_16	Is Abius	Sardinia	T16246-T16250	-
	Hap_27	Veyer	Rhone	T15554, T15578	-
		St. 11 pont Bâtie	Buëch	T12549	
		Biaysse amont	Rhone	T11571	
		Plampinet	Rhone	T09889-T09893	
		Ubaye à Gleizolles	Rhone	T04416, T04424, T04432, T04437	
	Hap_32	St. 11 pont Bâtie	Buëch	T12537, T12543, T12556, T12563	-
	Hap_40	Lette	Corsica	T10588, T10592, T10602, T10606	-
	Hap_41	Lette	Corsica	T10597	-
	Hap_43	Kranska-3	FYROM	T10037, T10040-T10043	-
	Hap_46	Marmanu	Corsica	T09212-T09214	-
		Carnevale	Corsica	T08355	
		Rocce	Corsica	T08333	
		Manica	Corsica	T08134, T08138, T08148, T08153	

		Veraculungu	Corsica	T08019	-
	Hap_47	Valbona-Drini	Albania	T09169	-
	Hap_48	Valbona-Gashi	Albania	T09167	-
	Hap_49	Valbona	Albania	T09165	-
	Hap_50	Valbona-Dragobi	Albania	T09163	-
	Hap_51	Valbona-Dragobi	Albania	T09161	-
	Hap_52	Shkumbini	Albania	T09065, T09067	-
	Hap_53	Shkumbini	Albania	T09060	-
	Hap_54	Shkumbini	Albania	T09059	-
	Hap_56	haut Borato	Corsica	T08403, T08408	-
	Hap_58	Veraculungu	Corsica	T08117, T08119	-
	Hap_59	Veraculungu	Corsica	T08118	-
	Hap_60	Uccialinu	Corsica	T08018	-
	Hap_63	Chjuvone	Corsica	T07712	-
	Hap_66	Fibreno	Italy	T05727, T05733	-
	Hap_67	Fibreno	Italy	T05729	-
	Hap_170; AdN	N/A	N/A	N/A	DQ297172 (Melkic *et al.*, not published)
	Hap_171; GA22	N/A	N/A	N/A	AF253552 (Suarez *et al.*, 2001)
	Hap_171; ADcs1	N/A	N/A	N/A	AY836330 (Cortey *et al.*, 2004)
	Hap_172; ADcs9	N/A	N/A	N/A	AY836338 (Cortey *et al.*, 2004)
	Hap_173; ADcs8	N/A	N/A	N/A	AY836337 (Cortey *et al.*, 2004)
	Hap_174; ADcs7	N/A	N/A	N/A	AY836336 (Cortey *et al.*, 2004)
	Hap_175; ADcs6	N/A	N/A	N/A	AY836335 (Cortey *et al.*, 2004)
	Hap_176; ADcs5	N/A	N/A	N/A	AY836334 (Cortey *et al.*, 2004)
	Hap_177; ADcs4	N/A	N/A	N/A	AY836333 (Cortey *et al.*, 2004)
	Hap_178; ADcs3	N/A	N/A	N/A	AY836332 (Cortey *et al.*, 2004)
	Hap_179; ADcs20	N/A	N/A	N/A	-
	Hap_180; ADcs2	N/A	N/A	N/A	-
	Hap_181; ADcs19	N/A	N/A	N/A	-
	Hap_182;	N/A	N/A	N/A	-

	ADcs17				
	Hap_183; ADcs17	N/A	N/A	N/A	-
	Hap_184; ADcs16	N/A	N/A	N/A	-
	Hap_185; ADcs15	N/A	N/A	N/A	-
	Hap_186; ADcs14	N/A	N/A	N/A	-
	Hap_187; ADcs13	N/A	N/A	N/A	-
	Hap_188; ADcs12	N/A	N/A	N/A	-
	Hap_189; ADcs11	N/A	N/A	N/A	-
	Hap_190; ADcs10	N/A	N/A	N/A	-
	Hap_191; ADAUA5	N/A	N/A	N/A	-
	Hap_193; AD_M1	N/A	N/A	N/A	-
	Hap_194; AD_C1	N/A	N/A	N/A	-
	Hap_201	Butiznica R	Croatia	T26083, T26084	-
	Hap_201	Kosovcica R	Croatia	T26087, T26088	-
	Hap_203	Kosovcica R	Croatia	T26086	-
ME	Hap_10	Ouvèze	Rhone	T16825, T16827, T16828, T16829, T16834	-
		Drôme	Rhone	T16797, T16803, T16806, T16821	
		Albarine	Rhone	T16754, T16759, T16762-T16764	
		Mouge	Rhone	T16699, T16701, T16704, T16713, T16718	
		Riotet	Rhone	T15984, T15991, T15994, T15998, T16007	
		Veyer	Rhone	T15549, T15551, T15581	

		Lavezon	Rhone	T12905	
		Thine	Rhone	T11050, T11055, T11061	
		Corbica	Corsica	T10789	
		Plampinet	Rhone	T09891	
		Ardèche Amt - Pt Mercier		T09512, T09519, T09526, T09533, T09540	
		Zoïco	Corsica	T09352	
		Sallevieille	Roya	T08850, T08861, T08869	
		Fontaine de Vaucluse		T07424-T07428	
		Volturno	Italy	T05787, T05789, T05790	
	Hap_17	Loup	Loup	T16126, T16131, T16140, T16143	-
	Hap_18	Loup	Loup	T16136	-
	Hap_28	Paratella	Corsica	T15508-T15512	-
	Hap_33	Fontan (aval)	Roya	T11618, T11625, T11639, T11647	-
		Breil pont Arbouset	T17284-T17288	-	
	Hap_34	Biaysse amont	Rhone	T11559, T11566, T11585	-
	Hap_35	Biaysse amont	Rhone	T11577	-
	Hap_37	Thine	Rhone	T11044	-
	Hap_38	Corbica	Corsica	T10792	-
	Hap_39	Corbica	Corsica	T10788, T10790, T10791	-
	Hap_42	Haut Golu	Corsica	T10183, T10188	-
		Radule	Corsica	T10159-T10163	-
	Hap_44	Gluyère	Rhone	T09694, T09707	-
	Hap_45	Sumène - Pt St Prix		T09393, T09398, T09403, T09409, T09418	-

	Hap_55	Sallevieille	Roya	T08844, T08856	-
	Hap_61	Chjuvone	Corsica	T07703, T07707, T07721	-
	Hap_62	Chjuvone	Corsica	T07717	-
	Hap_64	Volturno	Italy	T05791	-
	Hap_65	Volturno	Italy	T05788	-
	Hap_68	Ubaye à Gleizolles	Rhone	T04444	-
	Hap_69	Haut Golu	Corsica	T03088-T03091	-
	Hap_70	Bistrica Danube	Slovenia	T02456-T02459, T02465-T02468	-
	Hap_72; T5	N/A	N/A	N/A	AF253549 (Suarez *et al*., 2001)
	Hap_72; S61	N/A	N/A	N/A	AF253548 (Suarez *et al*., 2001)
	Hap-72; J53	N/A	N/A	N/A	AF253547 (Suarez *et al*., 2001)
	Hap_72; MEcs2	N/A	N/A	N/A	AY836351 (Cortey *et al.*, 2004)
	Hap_72; MEcs1	N/A	N/A	N/A	AY836350 (Cortey *et al.*, 2004)
	Hap_73; MEDU240	N/A	N/A	N/A	AF253550 (Suarez *et al*., 2001)
	Hap_73; MEcs14	N/A	N/A	N/A	AY836363 (Cortey *et al.*, 2004)
	Hap_74; MEcs9	N/A	N/A	N/A	AY836358 (Cortey *et al.*, 2004)
	Hap_75; MEcs8	N/A	N/A	N/A	AY836357 (Cortey *et al.*, 2004)
	Hap_76; MEcs7	N/A	N/A	N/A	AY836356 (Cortey *et al.*, 2004)
	Hap_77; MEcs6	N/A	N/A	N/A	AY836355 (Cortey *et al.*, 2004)
	Hap_78; MEcs5	N/A	N/A	N/A	AY836354 (Cortey *et al.*, 2004)
	Hap_78; MEcs4	N/A	N/A	N/A	AY836353 (Cortey *et al.*, 2004)
	Hap_79; MEcs3	N/A	N/A	N/A	AY836352 (Cortey *et al.*, 2004)
	Hap_80; MEcs15	N/A	N/A	N/A	AY836364 (Cortey *et al.*, 2004)
	Hap_81; MEcs13	N/A	N/A	N/A	AY836362 (Cortey *et al.*, 2004)
	Hap_82; MEcs12	N/A	N/A	N/A	AY836361 (Cortey *et al.*, 2004)
	Hap_83; MEcs11	N/A	N/A	N/A	AY836360 (Cortey *et al.*, 2004)
	Hap_84; MEcs10	N/A	N/A	N/A	AY836359 (Cortey *et al.*, 2004)
	Hap_202	Butiznica R	Croatia	T26085	-

MA	Hap_15	Volaja	Slovenia	T16489, T16490, T16492, T16494, T16496-16500	-
		Carnevale	Corsica	T08361	
		Paratella	Corsica	T08197-T08201	
		Trebuscica -1	Slovenia	T06832-T06836	
		Svenica-1	Slovenia	T06752, T06770, T06772, T06774, T06782	
		Svenica-2	Slovenia	T06782	-
	Hap_61	Chjuvone	Corsica	T07703, T07707, T07721	-
	Hap_62	Chjuvone	Corsica	T07717	-
	Hap_85; MAcs1	N/A	N/A	N/A	AY836365 (Cortey *et al.*, 2004)
	Hap_85; MA1a	N/A	N/A	N/A	DQ841191 (Meraner *et al.*, 2007)
	Hap_86; MA2b	N/A	N/A	N/A	DQ841190 (Meraner *et al.*, 2007)
	Hap_87; MA2a	N/A	N/A	N/A	DQ841189 (Meraner *et al.*, 2007)
	Hap_192; AD_Z1	N/A	N/A	N/A	DQ381565 (Susnik *et al.*, 2007)

Table E2. Samples used in the CR dataset with haplotypes, and accession numbers and references for GenBank sequences. N/A indicates missing information for GenBank haplotypes; - indicates that the haplotype is not on GenBank.

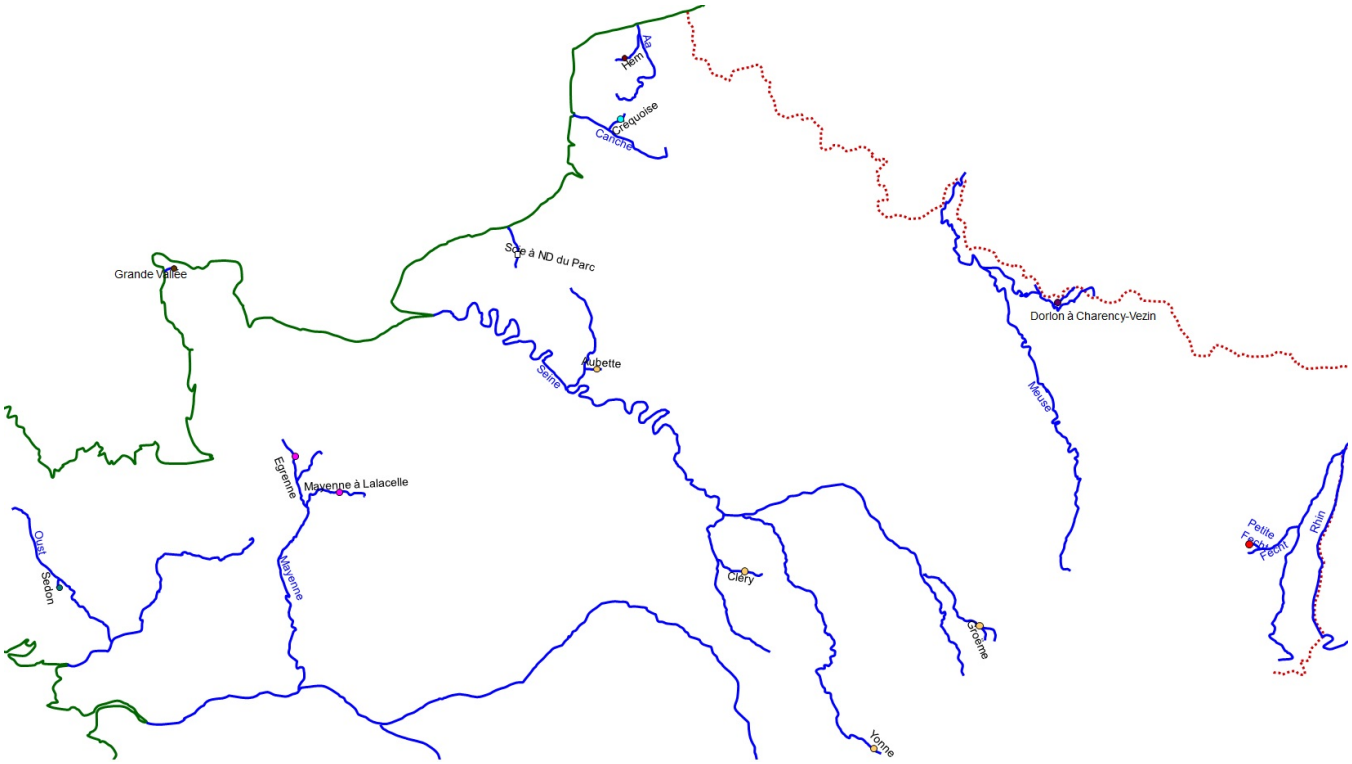

Figure E1. Map of the north of France showing sample sites. Coloured circlescorrespond with colours in haplotype networks (**Figures S9-11**). White squares indicate that the sample is included in the Cytb dataset, but not the concatenated dataset.

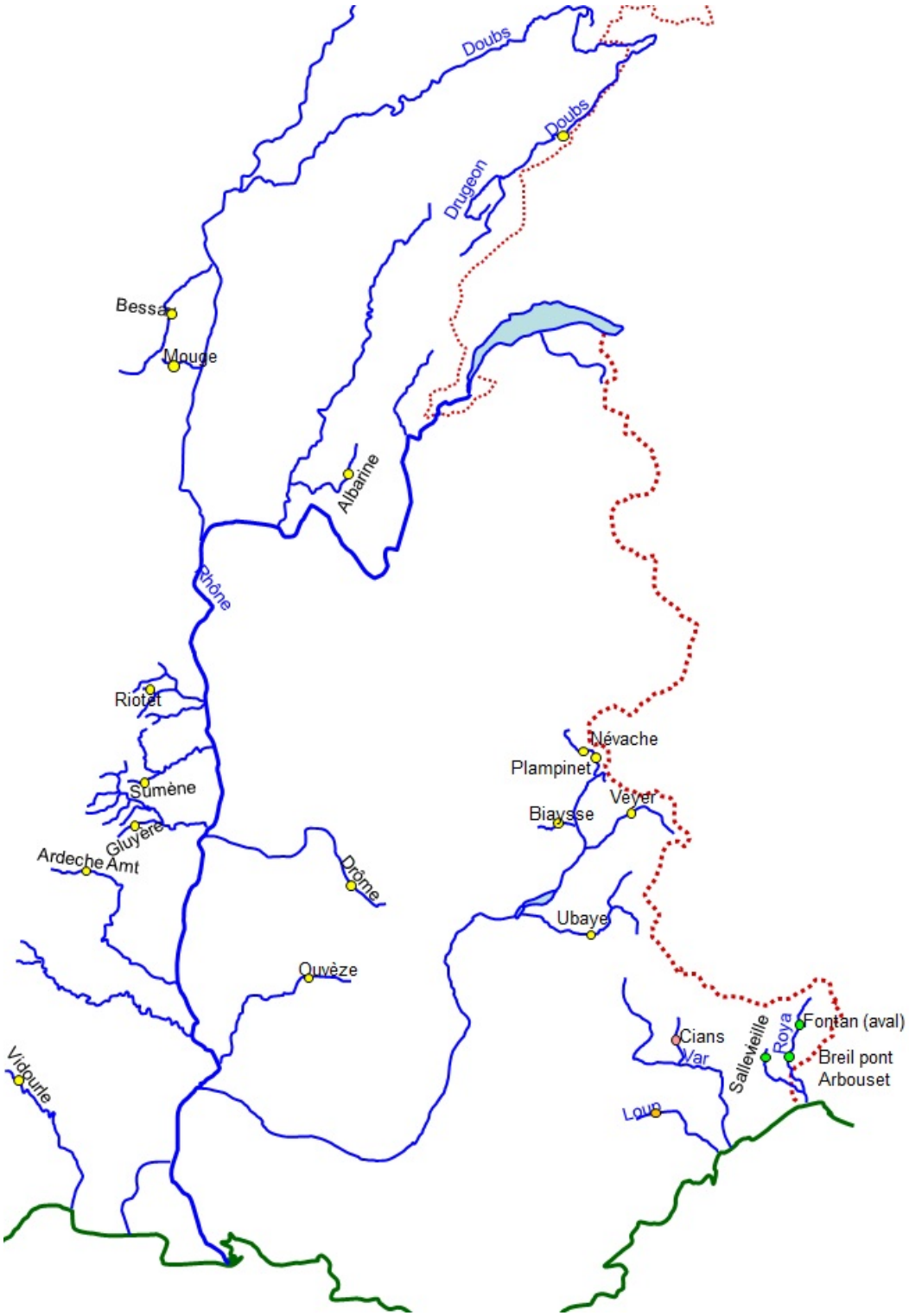

Figure E2. Map of the east of France showing sample sites. Coloured circles correspond with colours in haplotype networks (**Figures S9-11**). White squares indicate that the sample is included in the Cytb dataset, but not the concatenated dataset.

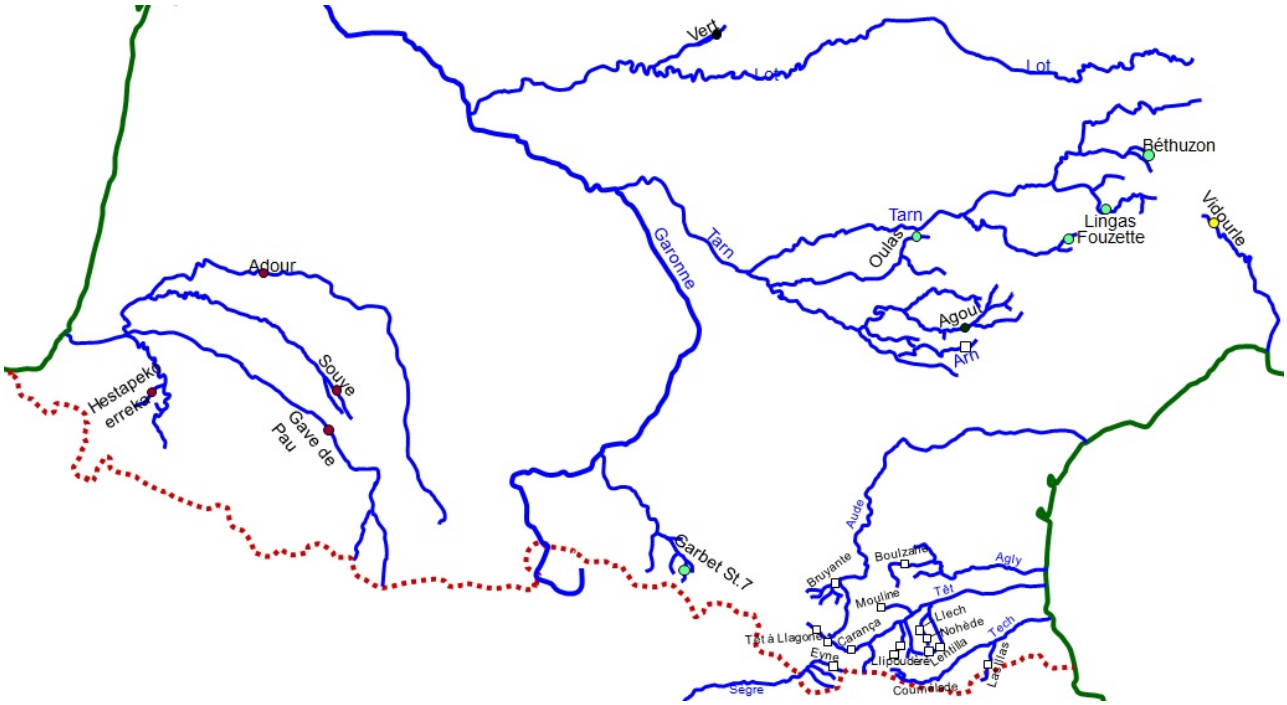

Figure E3. Map of the southwest of France showing sample sites. Coloured circles correspond with colours in haplotype networks (**Figures S9-11**). White squares indicate that the sample is included in the Cytb dataset, but not the concatenated dataset.

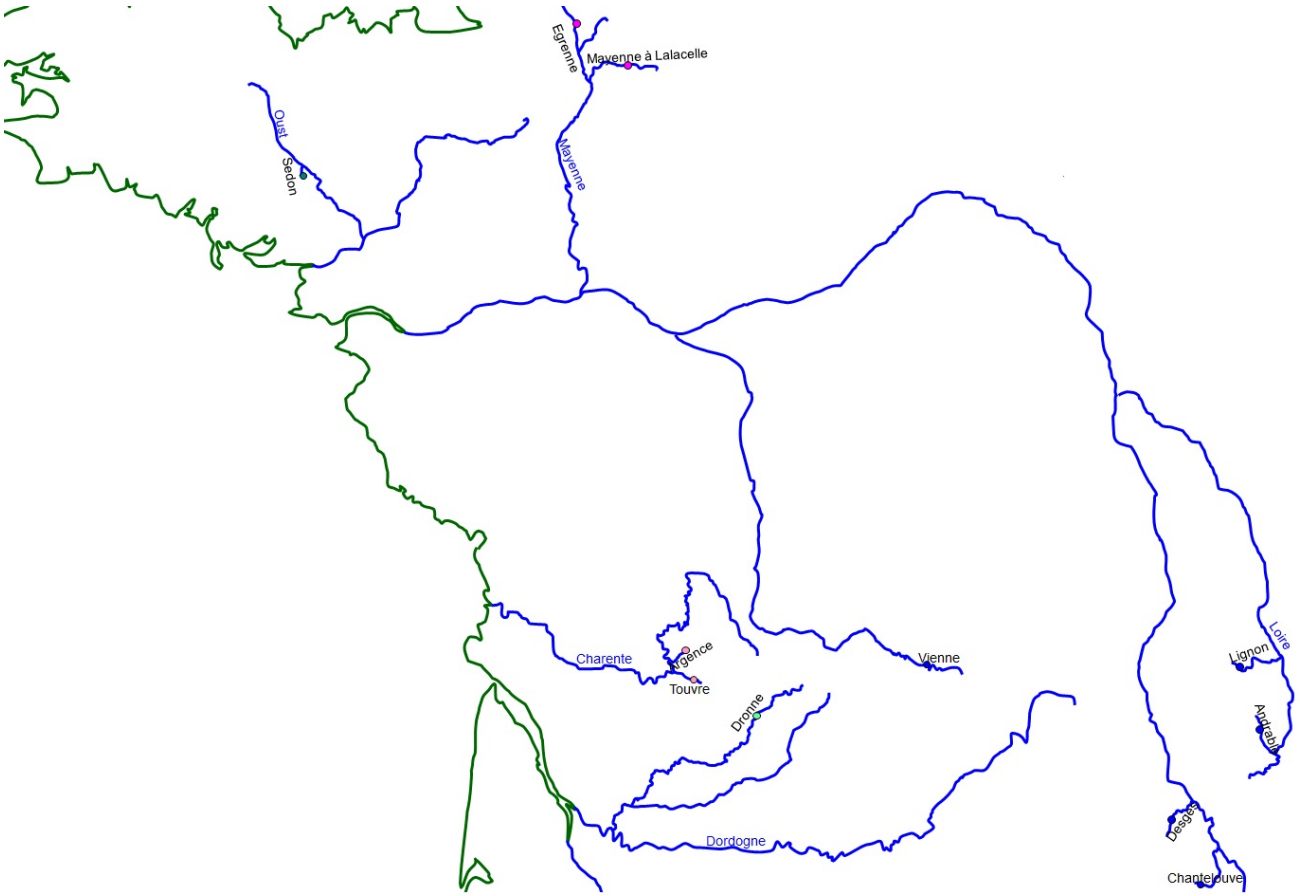

Figure E4. Map of the west of France showing sample sites. Coloured circles correspond with colours in haplotype networks (**Figures S9-11**). White squares indicate that the sample is included in the Cytb dataset, but not the concatenated dataset.