Strong genetic structure among populations of the tick Ixodes ricinus across its range

Pedro Poli, Jonathan Roger Michel Henri Lenoir, Olivier Plantard, Steffen Ehrmann, Knut Røed, Hans Petter Leinaas, Marcus Panning, Annie Guiller

Strong genetic structure among populations of the tick *Ixodes ricinus* across its range

Pedro Poli[1], Jonathan Lenoir[1], Olivier Plantard[2], Steffen Ehrmann[3], Knut H. Røed[4], Hans Petter Leinaas[5], Marcus Panning[6], Annie Guiller[1]

[1]Université de Picardie Jules Verne, UMR « Ecologie et Dynamique des Systèmes Anthropisés » (EDYSAN, UMR 7058 CNRS), 33 Rue Saint Leu, 80000 Amiens CEDEX 1, France

[2]BIOEPAR, INRAE, Oniris, 44307, Nantes, France

[3]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany

[4]Department of Basic Sciences and Aquatic Medicine, Norwegian University of Life Sciences, N-0033, Oslo, Norway (knut.roed@nmbu.no)

[5]Department of Biosciences, University of Oslo, Box 1066 Blindern, N-0316 Oslo, Norway (h.p.leinaas@ibv.uio.no)

[6]Institute of Virology, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Hermann-Herder-Str, 11 79104, Freiburg, Germany

Corresponding author: Pedro Poli, UMR « Ecologie et Dynamique des Systèmes Anthropisés » (EDYSAN, UMR 7058 CNRS-UPJV), Université de Picardie Jules Verne, 33 Rue Saint Leu, 80000 Amiens CEDEX 1, France, pvpoli@gmail.com

Abstract: *Ixodes ricinus* is the most common and widely distributed tick species in Europe, responsible for several zoonotic diseases, including Lyme borreliosis. Population genetics of disease vectors is a useful tool for understanding the spread of pathogens and infection risks. Despite the threat to the public health due to the climate-driven distribution changes of *I. ricinus*, the genetic structure of tick populations, though essential for understanding epidemiology, remains unclear. Previous studies have demonstrated weak to no apparent spatial pattern of genetic differentiation between European populations. Here, we analysed the population genetic structure of 497 individuals from 28 tick populations sampled from 20 countries across Europe, the Middle-East, and northern Africa. We analysed 125 SNPs loci after quality control. We ran Bayesian and multivariate hierarchical clustering analyses to identify and describe clusters of genetically related individuals. Both clustering methods support the identification of three spatially-structured clusters. Individuals from the south and north-western parts of Eurasia form a separated cluster from northern European populations, while central European populations are a mix between the two groups. Our findings have important implications for understanding the dispersal processes that shape the spread of zoonotic diseases under anthropogenic global changes.

Keywords: gene flow; infection risks; range shift

Introduction

Ixodes ricinus (Acari, Ixodidae) is the most widespread tick species occurring across Europe and an important vector of multiple tick-borne diseases, both to humans and livestock. Commonly reported pathogens transmitted by *I. ricinus* include: the bacteria *Borrelia burgdorferi* sensu lato, responsible for the Lyme borreliosis, which is the most prevalent tick-borne disease in temperate Europe (ECDC, 2015); arboviruses (genus *Flavivirus*) causing tick-borne encephalitis (TBE) and louping-ill (LI); the protozoan *Babesia microti*, responsible for the babesiosis; and the bacterium *Candidatus* Neoehrlichia mikurensis, responsible for neoehrlichiosis, an emerging tick-borne pathogen (Portillo et al., 2018; Welinder-Olsson et al., 2010). The current climate-driven redistribution of hematophagous arthropods such as ticks and mosquitoes may lead to severe challenges to public health and husbandry, by carrying a wide range of vector-borne diseases to new areas (Dantas-Torres, 2015; Pecl et al., 2017). For instance, many studies have demonstrated that the range of *I. ricinus* is already shifting northward and to higher elevations (e.g. Hvidsten et al., 2020; Jore et al., 2011; Lindgren and Gustafson, 2001) and those shifts are expected to continue in the future (Alkishe et al., 2017; Medlock et al., 2013).

Despite the threats of emerging infectious diseases following the redistribution of *I. ricinus*, little is known about the genetic structure of tick populations across the entire species range. Population genetic differentiation and spatial structuring can, however, impact the vector fitness and distribution, and therefore disease transmission (Blanchong et al., 2016; Wonham et al., 2006).

Population genetics approaches such as individual genetic clustering and assignment methods enable inference on migrants (exchange of genes between populations) and the risk of pathogen spread between populations (Kozakiewicz et al., 2018). For example, Lang and Blanchong (2012) applied clustering and distance-based methods to assess gene flow and disease spread risk between populations of white-tailed deer in the USA. Similarly, Van Zee et al. (2015) identified different genetic clusters between the southern and northern range of the tick *Ixodes scapularis* while the prevalence of borreliosis is known to be lower in the southern range. The authors suggest that this pattern of spatial genetic structure might be linked to differences in questing behaviour as ticks from the northern range would be more likely to bite humans. Differences in several life history traits of *I. ricinus* – such as the temperature at which nymphs begin to quest – have been reported along a latitudinal gradient (Gilbert et al., 2014), suggesting a spatially explicit phenotypic plasticity or adaptation. Yet, such basic knowledge about the distribution of genetic variation in *I. ricinus* and the migration processes involved in disease transmission remain largely unknown, albeit being essential to design better vector control strategies (Araya-Anchetta et al., 2015; Gooding, 1996; Tabachnick and Black, 1995).

The genetic structure of parasites' populations is known to be influenced by the distribution of the hosts (Kempf et al., 2009; Wessels et al., 2019). In general, it is assumed that generalist parasites relying on a wide range of hosts tend to show weak or no genetic structure, as shown in many studies on various parasite species (e.g. Archie and Ezenwa, 2011; Wessels et al., 2019). The tick species *I. ricinus* is a

85 generalist ectoparasite infesting a wide range of hosts, such as reptiles, mammals,
86 and birds (Casati et al., 2008; Norte et al., 2012). It has been proposed that tick
87 abundance and population genetic structure are dependent on the species' biology
88 (such as reproduction strategies and life cycle), but also on the host distribution
89 and behaviour (Kempf et al., 2011; McCoy et al., 2001; Rizzoli et al., 2009; Norte et
90 al., 2012). Large ungulates, such as deer, bovidae, and wild boar may be highly
91 efficient carriers of ticks for long distances, as long as there are no severe barriers
92 to their migration (Handeland et al., 2013; Kriz et al., 2014). By contrast,
93 transportation of ticks by migrating birds seems to be less efficient across
94 contiguous landmasses (Hasle et al., 2009; Røed et al., 2016). Based on these
95 findings, it is expected that *I. ricinus* populations should show a weak spatial genetic
96 structure.

97 Regarding previous works on population structure and dispersal of *I.*
98 *ricinus*, Noureddine et al. (2011) found a clear differentiation between European
99 and African populations using sequences from three nuclear and three
100 mitochondrial markers. Regarding the results from that study, it was later
101 suggested by Estrada-Peña et al. (2014) that those northern African samples could
102 correspond to *Ixodes inopinatus*, a sibling species of the *I. ricinus* complex within
103 the *Ixodes* subgenus. Considering only European populations, some studies showed
104 weak to no differentiation, but an extensive genetic diversity was observed within
105 each local population (Casati et al., 2008; Noureddine et al., 2011; Porreta et al.,
106 2013; Carpi, 2016). Other investigations analysing the frequency of mitochondrial
107 haplotypes showed a marked phylogeographical structure in northern Europe,

notably when considering populations from the north of the UK (Scotland) and Scandinavia (Al Khafaji et al., 2019; Dinnis et al., 2014; Røed et al., 2016). Although none of the mitochondrial haplotypes was exclusive to any of those populations, their frequencies varied significantly between populations from different regions. Interestingly, the British clade identified by Røed et al. (2016) coincides with the occurrence of a particular subtype of the louping-ill virus, which is closely related to other Irish and Spanish subtypes. Other studies focusing on the genetic structure of *I. ricinus* populations were based on microsatellite loci (Kempf et al., 2009; Kempf et al., 2011). Microsatellite variations have led to the identification of significant levels of genetic structure at different spatial scales, deviation from panmixia in *I. ricinus* populations likely due to assortative mating and patterns of host use (see Araya-Anchetta et al., 2015 for a review). However, those studies have also assessed patterns of genetic variation from localised samples that cover only a subset of the species range and thus likely do not capture the entire species genetic structure at the continental level.

Here, we aim to elucidate the population genetic structure of the tick *I. ricinus* based on single nucleotide polymorphisms (SNPs). To the best of our knowledge, no other study on the population genetic structure of *I. ricinus* throughout the Eurasian continent was based on the variation detected by this type of marker. Although generally having a weaker mutation rate than microsatellites, SNPs offers the possibility of building a larger range of markers and have been suggested to be more reliable markers for population genetic studies (Helyar et al., 2011; Smouse, 2010). Our main objective is to describe the genetic

structure of *I. ricinus* populations to infer the geographical and environmental factors shaping this structure. Particularly we hypothesized that (i) *I. ricinus* from the western parts of Europe might have genetic similarities to the Great Britain lineage (Røed et al., 2016) while (ii) there should be a pronounced genetic differentiation between ticks south and north of the extensive mountain areas covering central Europe (i.e., the Eastern Alps, the Western Alps, the Carpathian Mountains, and the Balkan Mountains).

Materials and Methods

Sampling

A total of 28 tick populations from 20 countries were sampled covering most of the species' range, including populations close to the northern (Norwegian, Sweden, Ireland, and England) and southern (Iran, Spain, and northern Africa) range limit of *I. ricinus* (Figure 1). Samples were collected by flagging inside or near forest fragments from the ground vegetation and were preserved in alcohol. A significant subset of the sampled populations we used, covering 8 regions across Europe (southern and northern France; Belgium; western and eastern German; southern and central Sweden; and northern Estonia), originated from a single project (smallFOREST, BiodivERsA 2010-2011 Joint call: https://www.biodiversa.org/491/download) and was sampled by the same person during the same year 2013 (See Ehrmann et al., 2018 for details). The remaining samples were collected for different projects (for details on those

projects see Røed et al., 2016 for the Norwegian samples and Noureddine et al., 2011 for the remaining samples). The coordinates of the sampled populations are provided in Table S1 (see Supporting Information). Aside from smallFOREST samples, sampling dates varied among the sampled populations (Table S1).

Ticks sampled for those projects were identified at the laboratory using standard morphological keys provided in Babos (1964), Hillyard (1996), or Perez-Eid (2007). As most samples we used were identified before the description of *I. inopinatus* (Estrada-Peña et al., 2014) and considering that it was impossible to re-evaluate the identification of samples based on morphological features, we conducted an *a posteriori* evaluation of the potential presence of *I. inopinatus* among our samples. To fulfil this aim, northern African *I. inopinatus* samples analysed by Noureddine et al. (2011) were included in the present study.

DNA extraction and SNP genotyping

Since ticks and DNA samples analysed in this study had different origins and therefore different storage methods, three different methods were used to ensure DNA extraction. Ticks were either: (i) frozen and crushed with a pestle in individual tubes before extracting DNA using DNeasyTM Tissue Kit (Qiagen); (ii) disrupted using a Tissue Lyser (Qiagen) before DNA extraction using the Wizard Genomics DNA Purification Kit (Promega, USA); or (iii) crushed with Lysing matrix H (MP Biomedicals, Santa Ana, USA) before extracting DNA with MagNA Pure LC Total Nucleic Acid Isolation Kit (Roche, Basel, Switzerland).

We genotyped 192 SNPs as described by Quillery et al. (2014). The list of SNPs, variant basis, and primers are presented in Table S2. All samples were amplified by whole genome amplification (WGA) before genotyping. The PEP-PCR WGA kit (LGC-Biosearch Technologies) was used for whole genome amplification of each sample. The WGA protocol associated with KASP genotyping has already been tested by Quillery et al. (2014) and showed a reduced number of "no-call" data (missing values) during genotyping. The WGA and genotyping steps were subcontracted by the GENTYANE platform (INRA, Clermont-Ferrand, France: http://gentyane.clermont.inra.fr/). The GENTYANE platform is an INRAE (French National Research Institute for Agriculture, Food and Environment) research facility located in Clermont-Ferrand (France) which offers sequencing and genotyping services. Genotyping was conducted in a Biomark HD System (Fluidigm) and KASPar assays. The KASPar method is a KBiosciences competitive allele-specific PCR amplification. A PCR mix containing two allele-specific forward primers and one common reverse primer was carried out. Each forward primer had a 5' tail sequence homologous to universal secondary oligos labelled with a fluorophore (FAM or HEX). If a particular locus is homozygous, only one fluorescent signal is generated. Bi-allelic loci generate both fluorescent signals.

Quality control

Data was filtered after genotyping and before statistical analysis. First, all invariant SNPs were removed. After this first filtering step, all individuals with more than 20% of non-amplified sites (missing data) were removed. Finally, all remaining SNPs with more than 20% missing data were also removed. The remaining dataset consisted

of both individuals and SNPs with less than 20% missing data. After quality control steps, 125 SNP loci and 497 individuals were kept for further analyses.

Cluster analysis and genetic structure

Two complementary clustering methods were used to access the genetic structure of *I. ricinus* populations. First, we investigated the genetic clustering by performing a discriminant analysis of principal components (DAPC, Jombart et al., 2010) with the package 'adegenet' (Jombart, 2008) in R (R Core Team, 2019). The optimal *k* number of clusters was identified by the *k*-means algorithm using the *find.cluster()* function based on BIC values. A maximum of 28 clusters was allowed, i.e. the total number of sampled populations. Next we performed a Bayesian analysis in STRUCTURE (Pritchard et al., 2000) with the parameter *K*, i.e. the optimal number of clusters, varying from 1 to 10, according to the results from the DAPC. We used a non-admixture model with the sampling locations as prior. Twenty repetitions of 80,000 MCMC iterations with a burning length of 20,000 iterations were run for each value of *K*. The results were analysed with Structure Harvester (Earl and vonHoldt, 2012). The best *K* value for the optimal number of clusters was identified by comparing the estimates of log probabilities of the data (i.e. ln[Pr($X|K$)]) for each *K* value as well as Evanno's delta *K* method (Evanno et al., 2005). Pritchard et al. (2007) suggested aiming for the smallest value of *K* that captures most of the genetic structure in the data. Assigning probabilities for individuals and populations across repetitions were then averaged in CLUMPP (Jakobsson and Rosenberg, 2007). We applied a hierarchical clustering analysis (e.g. Vähä et al., 2007) in each identified cluster to detect more refined patterns of

221 genetic structure. Hierarchical analysis in STRUCTURE was realised with ten
222 repetitions and the same other parameters as the first round of analysis. We
223 realised a similar analysis for each cluster identified by DAPC.

224 To test our data for isolation by distance (IBD), pairwise F_{ST} values were
225 estimated with the package 'hierfstat' (Goudet and Jombart, 2018) in R (R Core
226 Team, 2019) as Weir and Cockerham unbiased parameter θ (Weir and Cockerham,
227 1984). The IBD pattern was first tested across all pairs of Eurasian samples and
228 second only between pairs of samples collected during the same year to avoid
229 potential biases due to temporal variability in dispersal and genetic structure. Those
230 corresponded to samples from southern and northern France, Belgium, western
231 and eastern German, northern Estonia, southern and central Sweden, a total of 8
232 samples (28 pairs). Since the 25 Eurasian samples are distributed across a large
233 continental extent, pairwise geographical distances were calculated with the
234 'geosphere' package (Hijmans, 2017) in R (R Core Team, 2019) to account for the
235 curvature of the Earth. The strength of the IDB was evaluated as the relationship
236 between $\theta/(1-\theta)$ and the natural logarithm of the geographic distance as
237 described by Rousset (1997). In a two dimensions population, the slope parameter
238 b of the linear regression $\theta/(1-\theta) = a + bD_{Geo}$ is inversely proportional to the
239 average neighbourhood size $Nb = 1/b$, and $b = 1/(4D_e\pi\sigma^2)$, where D_e is the sub-
240 population density and σ^2 is the averaged square axial distances between adults
241 and their parents and σ is half the average adult-parent distance (Séré et al., 2017).
242 In this case, a proxy of dispersal can be calculated as $\delta \approx 2 * \sqrt{(4\pi Deb)}$
243 (Manangwa, 2018). The population density was calculated as $D_e = N_e/S\pi$, where

244 *S* is the smallest distance between sites considered and included in the IBD analysis.
245 We used NeEstimator version 2.1 to calculate effective population sizes (*Ne*) by
246 applying two different methods, one based on linkage disequilibrium and another
247 based on molecular co-ancestry (Do et al, 2014). We calculated the mean of *Ne*
248 estimated with these two methods after the exclusion of 'infinity' results. The
249 obtained mean value was weighted by the number of times one of the two methods
250 generated a non-infinity value. The significance of the IBD pattern was assessed by
251 Mantel tests as implemented in the 'vegan' package (Oksanen et al., 2019) in R (R
252 Core Team, 2019).
253 *Genetic diversity*
254 For each locus, we estimated the observed heterozygosity (Ho), the gene diversity
255 (Hs), and Wright's fixation indices F_{IS}, F_{ST}, and F_{IT}. Wright's statistics measure
256 inbreeding in three levels of population structure: F_{IS} is the inbreeding coefficient
257 of individuals relative to subpopulations; F_{ST} is the inbreeding coefficient of
258 subpopulations relative to populations; and F_{IT} is a measure of the inbreeding of
259 individuals relative to populations. All metrics were calculated with the package
260 'hierfstat' (Goudet and Jombart, 2018) in R (R Core Team, 2019). A Monte-Carlo
261 permutation test (999 replicates) was conducted to test for the significance of the
262 differences of mean gene diversity and F_{IS} values over loci between pairs of genetic
263 clusters identified. For each replicate, individuals were randomly assigned to one
264 genetic cluster and the simulated statistics were calculated. We ran
265 the *randtest()* function from the 'ade4' package (Dray and Dufour, 2007) to access
266 the significance of the observed differences.

To investigate null alleles and possible Wahlund effect on genotype frequencies, we followed the procedure proposed by De Meeûs (2018). According to that study, the presence of null alleles could be identified by a suit of comparisons of F_{IS}, F_{ST}, and the number of missing data. In case of null alleles, we would observe: (i) a high positive correlation between F_{IS} and F_{ST}; (ii) high variation of both F_{IS} and F_{ST} across loci; (iii) F_{IS} standard errors (StrdErrFIS) much bigger than F_{ST} standard errors (StrdErrFst); and (iv) F_{IS} values mainly explained by the presence of missing data. For the Wahlund effect, the correlation between F_{IS} and F_{ST} should approximate zero, a small variation of F_{ST} and a moderate variation of F_{IS} should be observed across loci, F_{IS} standard errors (StrdErrFIS) should be higher than F_{ST} standard errors (StrdErrFst) and no or rare missing data should be obtained. To test those relations, values of F_{IS}, F_{ST}, StrdErrFst, and StrdErrFIS were calculated in the FSTAT software version 2.9.4 (Goudet, 2003), the latter values calculated by Jackknife. The Spearman's rank correlation test was applied to test for correlations. Finally, De Meeûs (2018) suggested a linear regression between F_{IS} and missing data to quantify, using the R^2 value, the contribution of missing data in F_{IS} values. Because the Wahlund effect can produce between-locus dependencies, we also tested linkage disequilibrium for each pair of loci by using G-based tests implemented in FSTAT 2.9.4. Since *p*-values from each test are not independent, we applied the procedure described by Benjamini and Yekutieli (2001) to calculate the false discovery rate (FDR) and correct *p*-values.

Results

Clustering analysis, genetic differentiation and isolation by distance

The DAPC analysis identified two possibilities for the number of clusters, one suggesting three different genetic clusters and the other suggesting four genetic clusters (the BIC difference is 0.842 between $K = 3$ and $K = 4$, Figure S1). Choosing $K = 4$ clusters created two overlapping groups, while $K = 3$ grouped individuals into 3 well-separated clusters (Figure 2). Hence, we decided to set the number of clusters to $K = 3$ with the DAPC approach. Bayesian analysis performed with STRUCTURE also identified a $K = 3$ differentiated genetic clusters (Figure 2b and Figure S2) whose compositions are very similar to the three clusters retained with the DAPC approach. In both analyses, northern African (yellow colour in Figures 2 and 3) and Eurasian populations (the other colours) were highly differentiated. Two main groups were identified within Eurasia, one corresponding mainly to northern and continental middle European populations (grey colour in Figures 2 and 3), the other corresponding mainly to southern and western populations in Eurasia (blue colour in Figures 2 and 3). The DAPC approach separated northern African populations from Eurasian ones along the first axis, while Eurasian clusters were mostly separated along the second axis (Figure 2a). Regarding clustering analyses with STRUCTURE, individual probabilities of different K values ranging from 2 to 10, excepted for $K = 3$ which is already depicted in Figure 2b, are presented in the Supporting Information (see Figure S3).

Finer genetic structure was identified from our hierarchical analyses (Figures S4 and S5). These analyses, either carried out with DAPC (Figure S4) or the STRUCTURE approach (Figure S5), were able to isolate Iran and/or Turkey from the other sampled sites within the southern Eurasian cluster. Atlantic sites (Spain, southern and western France, Ireland, and England) were further isolated from the remaining sites in this group (Italy, Romania, Hungry, and Slovakia). The northern European sites showed a more admixture structure, and separation in further clusters varied between the DAPC and STRUCTURE approaches (see the 'Hierarchical analysis' section in the Supplementary Information for more details).

A pattern of isolation by distance (IBD) was observed across all sampled populations (Mantel r = 0.726, $p < 0.001$). Restricting the IBD analysis to the set of sites sampled during the same year, we found an even stronger pattern of IDB (Mantel r = 0.870, $p < 0.0001$, Figure 4). In the latter case, the coefficient estimate of the slope parameter (b) in the regression was $b = 0.01$ with a 95% confidence interval (CI) ranging from 0.007 to 0.013. Neighbourhood size (Nb) reached $Nb = 99$ individuals, on average (95% CI = [71-140]), and immigration rate (N_em) was estimated to reach $N_em = 16$ (95% CI = [11-22]) individuals per generation and subpopulation.

We found a mean effective population size of 62 individuals. The closest sampled sites were North France and Belgium, separated 119 km from one another. We found surface and population densities to reach, on average, $S^2 = 11.3$ km^2 and

331 *De* = 5.4 individuals/m^2, respectively. We found the dispersal rate to reach, on
332 average, $\delta \approx 76$ km/generation (95% CI = [65-90]).

333 *Genetic diversity*

334 The observed heterozygosity (*Ho*), gene diversity (*Hs*), and F_{IS} were highly variable
335 across loci (Table S3). The observed F_{ST} values were, however, more constant than
336 F_{IS} ones. For most loci, gene diversity was higher than the observed heterozygosity.
337 Consequently, the overall gene diversity across all loci was significantly higher than
338 the observed heterozygosity (Wilcox Signed-Rank Test, V = 6959, $p < 0.0001$). The
339 mean gene diversity per sampled population was still higher than the observed
340 heterozygosity (Wilcox Signed-Rank Test, V = 406, $p < 0.0001$) and mean F_{IS} was
341 always positive. Mean values of observed heterozygosity, gene diversity, and F_{IS} for
342 each population are shown in Figure S6 (Supporting Information). The highest mean
343 gene diversity and F_{IS} values over loci were identified in the southern Eurasian
344 cluster (Hs = 0.355, F_{IS} = 0.275), followed by the northern European cluster (Hs =
345 0.340, F_{IS} = 0.2708) and the cluster from northern Africa (Hs = 0.171, F_{IS} = 0.191)
346 (Figure 5). The Monte-Carlo test showed a significant difference in gene diversity
347 values for all pairs of clusters (p = 0.001 for all three comparisons), but none for F_{IS}
348 values (p = 0.199 and 0.239 when comparing northern Africa to the northern
349 European cluster and northern Africa to the southern Eurasian cluster, respectively;
350 while p = 0.644 when comparing the southern Eurasian cluster to the northern
351 European cluster). Populations from northern Africa showed a high deficit in
352 heterozygosity, of which 71 out of 125 loci with *Hs* values of zero.

After *p*-value correction (Benjamini and Yekutieli, 2001), no pair of locus showed significance values of linkage disequilibrium. No correlation was found between F_{IS} and F_{ST} (ρ = -0.0206, p = 0.8198) and missing data were positively correlated to F_{IS} values (ρ = 0.5804, $p < 0.001$). The linear regression of F_{IS} against the number of missing data estimated an adjusted R^2 of 0.19, suggesting that around one-fifth of F_{IS} variance is explained by the number of missing data. Finally, StrdErrFIS was around 4 times bigger than StrdErrFst (0.033 and 0.008, respectively).

Discussion

We investigated the genetic structure of populations from the tick *I. ricinus* in much of its range, *i.e.* in Eurasia and in northern Africa. In addition to a strong and expected divergence between northern African and Eurasian populations, the two Eurasian genetic clusters described here showed clear spatial patterns. The isolation by distance patterns we found, either throughout the entire dataset or restricted to samples from the same period, suggest an association between the genetic structure of *I. ricinus* populations and the geographical location of these populations. Hierarchical analyses confirmed the genetic affinity between western European populations, from the UK and Ireland in the north to Spain in the south, supporting our first hypothesis regarding genetic similarities in western continental Europe and the British Isles. Also consistent with our second hypothesis stating a genetic signature of central European mountains, we found a clear differentiation between populations from southern Eurasia and populations from northern

375 Europe. Indication of migration of individuals between the two clusters is suggested
376 by the different degrees of affinity from central Europe with one cluster or another
377 (e.g. in Romania, Hungary, Slovakia, and Moldova).

378 *Ixodes ricinus* and *I. inopinatus* have recently been suggested to be
379 sympatric both in northern Africa (Younsi et al., 2020) and in Europe (Estrada-Peña
380 et al., 2014; Chitimia-Dobler et al., 2018). Our results are clear concerning the
381 genetic identity of northern African samples. According to the results from both the
382 DAPC and STRUCTURE analysis, there is no possibility of any individuals from those
383 populations to belong to any other genetic clusters. Also, no individual from Eurasia
384 had any probability of identity with the northern African cluster. Converging results
385 of both analyses indicate with a great deal of certitude that: (i) all samples from
386 northern Africa belong to the same species and have the same ancestry; (ii) no
387 sample in Eurasia share ancestry with northern African ones. Northern African
388 samples were also a particular case as more than half loci were monomorphic
389 across all three populations, which was not found in Eurasian populations. Again, it
390 is important to note that individuals from the three northern African populations
391 analysed here were identified before the description of *I. inopinatus* (Estrada-Peña
392 et al., 2014). If *I. inopinatus* was present in the Eurasian samples, we would expect
393 at least small probabilities of identity of Eurasian samples with the northern African
394 cluster, which was not the case. The clear-cut genetic differentiation we obtained
395 between Eurasian and northern African populations strongly suggests that all the
396 individuals from the three northern African populations analysed here correspond

to *I. inopinatus.* Those results also illustrate the potential of using some of the SNPs analysed here to differentiate at a molecular level the two *Ixodes* species.

Two previous studies covering a large spatial extent of *I. ricinus'* range (Noureddine et al., 2011; Porreta et al., 2013) did not find such a clear geographical structure between Eurasian populations. Several reasons may explain this difference. First, a somewhat reduced number of individuals per population (sometimes a single individual per population in Nourredine et al., 2011) may explain a lack of spatially structured signal in former studies. Second, those former studies were based on mitochondrial and nuclear sequences. This said, a marked genetic differentiation into two distinctive clades has already been reported (Dinnis et al., 2014; Røed et al., 2016), suggesting a split in *I. ricinus* populations between northern continental Europe and Great Britain. Our results confirm and extend this pattern to most of the Eurasian range of the species by suggesting that Scandinavian populations are genetically closer to the populations from the north-eastern continental parts of Europe. Although there is a certain degree of gene flow between the two clusters, the north vs. south-eastern exchange may be hampered by mountain areas in central Europe. This reinforces the argument that large animals efficiently maintain high gene flow between tick populations across contiguous and permeable landscapes, while intense transportation by birds, during spring and autumn migration across sea or mountains (Hasle et al., 2009; Røed et al., 2016), may not be as sufficient to break down boundaries between established genetic entities.

419 Surprisingly, we found a close genetic affinity between all Atlantic samples
420 (i.e. Ireland, England, western and southern France, and Spain) and the
421 geographically separated populations from Turkey and Iran. This genetic affinity
422 among distant populations in Eurasia was supported by the two different clustering
423 methods we used (DAPC and STRUCTURE). Besides these results, the refined
424 hierarchical analyses isolated Iran and Turkey in their particular clusters in the first
425 (DAPC) and second (STRUCTURE) round of hierarchical clustering analyses. This
426 suggests that an east-west transport of ticks across southern Eurasia must be
427 sufficient to maintain a genetically identifiable cluster across this extensive area.
428 Interestingly, louping-Ill like viruses are also known from Greece and Turkey (Gao
429 et al., 1993; Marin et al., 1995), which might further support our findings and a link
430 between tick lineages and *Flavivirus*, although the causation is not known.
431 Since migratory birds carry *I. ricinus* across long distances, different
432 migratory routes could also contribute to the north-south genetic differentiation
433 we observed (Hasle et al., 2009; Røed et al., 2016). However, birds mainly carry
434 larvae and nymphs. Since surviving rates between development states are low, the
435 overall reproductive success of tick transported by birds is likely smaller than that
436 of adult ticks carried by large mammals. This may explain the maintenance of
437 genetic differentiation e.g. between the UK and Norway despite massive transport
438 of ticks' larvae in both directions (Røed et al., 2016).
439 Regarding the population structure observed within samples, the deviation
440 from Hardy-Weinberg equilibrium we found is in agreements with previous studies
441 on population genetics of *I. ricinus* based on SNPs (Quillery et al., 2014) and

442 microsatellites (Kempf et al., 2009; Kempf et al., 2011; Røed et al., 2006), as well as
443 other tick species (Dharmarajan et al., 2011). Possible causes of the observed
444 deviation from the Hardy-Weinberg equilibrium are assortative mating (or
445 assortative pairing), Wahlund effect, or errors in the genotyping. A tendency of
446 mating between phenotypically or genetically similar individuals may effectively
447 increase the inbreeding and thus heterozygote deficiency within populations (Jiang
448 et al., 2013). Kempf et al. (2009) suggested that assortative mating might occur in
449 *I. ricinus*, mostly via host selection (Kempf et al., 2011). Inbreeding in ticks could be
450 a result of host infestation by related individuals, which leads to high breeding
451 success of sibling groups (Araya-Ancheta et al., 2015). The highly aggregated egg
452 masses in *I. ricinus* (1000 to 3000 eggs) and the limited active dispersal of larvae
453 and nymphs may lead to a high likelihood of mating between related individuals
454 and thus inbreeding. Finally, the parasite-host relationship specificities could also
455 play an important role in establishing or maintaining population structure in *I.*
456 *ricinus*. If different host populations are present locally and exhibit behaviours
457 favouring mating within (and not between) each host population, this may induce
458 a Wahlund effect and explains the heterozygote deficiency observed. The existence
459 of such a host population behaviour has been characterized in *I. uriae*, a tick
460 associated with sea birds (Mc Coy et al., 2001) but also suggested in *I. ricinus* (Kempf
461 et al., 2009, 2011). Even though we did not conceive this study to test for such a
462 hypothesis, our results support at least partially non-random mating in *I. ricinus*
463 populations and the consequent Wahlund effect. Dharmarajan et al. (2011) facing
464 a similar result for the American species *I. texanus* showed that subdivided

breeding groups and high variance in individual reproductive success can correctly explain Hardy-Weinberg equilibrium deviation.

It is widely acknowledged that more or less isolated populations could develop particular adaptations in response to environmental differences between habitats. Nonetheless, very few studies to date have clearly observed phenotypic variations among *I. ricinus* populations from different geographical origins. In Estrada-Peña et al. (1996 and 1998), differences in cuticular hydrocarbon composition among European populations of *I. ricinus* were observed according to the geographical origin of those populations. Interestingly, the multivariate phenotypic analysis presented in those studies showed a somewhat similar pattern to our hierarchical genetic clustering analysis, notably concerning what the authors call 'peripheral populations'. Aside from chemical differentiation, behavioural differences between ticks' populations have also been documented, such as mismatches in questing peaks (Schulz et al., 2014) and questing responses to temperature (Gilbert et al., 2014; Tomkins et al., 2014). In controlled conditions, Gilbert et al. (2014) and Tomkins et al. (2014) showed that *I. ricinus* nymphs from cooler climates begin questing at lower temperatures than nymphs from warmer climates. They also start questing sooner when the temperature was kept constant. In any case, local adaptations could impact the spatial redistribution of the species range in response to changes in abiotic conditions. In a global changing context, such consequences could be explored by environmental niche modelling to identify areas of potential future expansion. It remains to be investigated if the different

clusters we identified here could pose different threats for human health and the potential risk of tick-borne disease transmission to humans.

Our findings on isolation by distance suggest small population densities and large dispersal distances among the sampled populations. The large dispersal distance is not a surprising result since ticks can parasitize highly mobile species. In a changing climate context, this result indicates that ticks could easily colonize new suitable habitats outside the current limits of the species geographical range in a few generations.

Despite being a generalist ectoparasite, our results highlight geographically distinct and genetically structured populations in *I. ricinus*. More research on host preference and dispersal capacity is needed to better understand those patterns. The differentiation of Eurasian populations into two geographically distinct clusters (northern Europe vs. southern Eurasia) could have important implications for the redistribution of *I. ricinus* in response to anthropogenic climate change. Ticks from a given genetic cluster could be more or less prone to increase in abundance in some regions. Combining tick and pathogen population genetics with knowledge on host distribution could help in the early detection of the spread of tick-borne diseases and thus improve the responsiveness of public authorities to limit major public health concerns.

Acknowledgements

We thank all the invaluable contribution of Dr. Brigitte Degeilh (Laboratory of Parasitology and Mycology, Rennes University Hospital, France), Dr. Francesco

Nazzi (Università Degli Studi di Udine), Dr. Mohammad Abdigoudarzi (Razi Institute, Iran), Dr Ali Bouattour (Institut Pasteur de Tunis, Tunisia), Dr Loubna Dib (Institut Vétérinaire, Centre Universitaire El Tarf, Algeria), M'hammed Sarih (Institut Pasteur du Maroc, Morocco), Dr. Irina Golovljova (National Institute for Health Development, Estonia), Nicole Voss (University of Erlangen-Nürnberg, Germany), Dr. Zati Vatansever (Kafkas University, Turkey), Dr. Eoin Healy (University College, Ireland), Dr. Davide Sassera (University of Pavia, Italy), Dr. Albert Agoulon (Oniris, France), Dr. Ionut Pavel (Regional Institute of Oncology, Romania), Dr. György Csóka (Hungary Forest Research Institut, Hungary), Dr. Elena Kocianová (Slovak Academy of Sciences, Slovakia), Dr. Alexandru Movila (Academy of Sciences of Moldova, Moldavia) for sharing some of the samples used in this study. We thank Dr. Michael Scherer-Lorenzen (Freiburg University, German) for the general smallFOREST study design, from which many samples were made available. The smallFOREST project was funded by the ERA-Net BiodivERsA, via the national funders ANR (France), FORMAS (Sweden), ETAG (Estonia), DFG (Germany), BELSPO (Belgium) and the European Union through the European Regional Development Fund (the EcolChange Centre of Excellence). We thank the Centre de Ressources Régionales en Biologie Moléculaire (CRRBM) from the Université de Picardie Jules Verne. This work was also supported by the Région Hauts-de-France, the Centre National de la Recherche Scientifique (CNRS, INEE) and the European Regional Development Fund.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Al-Khafaji, A.M., Clegg, S.R., Pinder, A.C., Luu, L., Hansford, K.M., Seelig, F., Dinnis, R.E., Margos, G., Medlock, J.M., Feil, E.J., Darby, A.C., McGarry, J.W., Gilbert, L., Plantard, O., Sassera, D., Makepeace, B.L., 2019. Multi-locus sequence typing of *Ixodes ricinus* and its symbiont *Candidatus* Midichloria mitochondrii across Europe reveals evidence of local co-cladogenesis in Scotland. Ticks Tick-Borne Dis. 10, 52–62. https://doi.org/10.1016/j.ttbdis.2018.08.016

Alkishe, A.A., Peterson, A.T., Samy, A.M., 2017. Climate change influences on the potential geographic distribution of the disease vector tick *Ixodes ricinus*. PLOS ONE 12, e0189092. https://doi.org/10.1371/journal.pone.0189092

Araya-Anchetta, A., Busch, J.D., Scoles, G.A., Wagner, D.M., 2015. Thirty years of tick population genetics: A comprehensive review. Infect. Genet. Evol. 29, 164–179. https://doi.org/10.1016/j.meegid.2014.11.008

Archie, E.A., Ezenwa, V.O., 2011. Population genetic structure and history of a generalist parasite infecting multiple sympatric host species. Int. J. Parasitol. 41, 89–98. https://doi.org/10.1016/j.ijpara.2010.07.014

Babos S. 1964. Die Zeckenfauna Mitteleuropas. Budapest: Akadémiai Kiadó.

Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat. 29(4), 1165-1188. www.jstor.org/stable/2674075

Blanchong, J.A., Robinson, S.J., Samuel, M.D., Foster, J.T., 2016. Application of genetics and genomics to wildlife epidemiology: Genetics and Wildlife Epidemiology. J. Wildl. Manag. 80, 593–608. https://doi.org/10.1002/jwmg.1064

Carpi, G., Kitchen, A., Kim, H.L., Ratan, A., Drautz-Moses, D.I., McGraw, J.J., Kazimirova, M., Rizzoli, A., Schuster, S.C., 2016. Mitogenomes reveal diversity of the European Lyme borreliosis vector *Ixodes ricinus* in Italy. Mol. Phylogenet. Evol. 101, 194–202. https://doi.org/10.1016/j.ympev.2016.05.009

Casati, S., Bernasconi, M.V., Gern, L., Piffaretti, J.-C., 2008. Assessment of intraspecific mtDNA variability of European *Ixodes ricinus* sensu stricto (Acari: Ixodidae). Infect. Genet. Evol. 8, 152–158. https://doi.org/10.1016/j.meegid.2007.11.007

Chitimia-Dobler, L., Rieß, R., Kahl, O., Wölfel, S., Dobler, G., Nava, S., Estrada-Peña, A., 2018. *Ixodes inopinatus* − Occurring also outside the Mediterranean region. Ticks Tick-Borne Dis. 9, 196–200. https://doi.org/10.1016/j.ttbdis.2017.09.004

Dantas-Torres, F., 2015. Climate change, biodiversity, ticks and tick-borne diseases: The butterfly effect. Int. J. Parasitol. Parasites Wildl. 4, 452–461. https://doi.org/10.1016/j.ijppaw.2015.07.001

De Meeûs, T., 2018. Revisiting FIS, FST, Wahlund Effects, and Null Alleles. J. Hered. 109, 446–456. https://doi.org/10.1093/jhered/esx106

Dharmarajan, G., Beasley, J.C., Rhodes, O.E., 2011. Heterozygote deficiencies in parasite populations: an evaluation of interrelated hypotheses in the raccoon tick, *Ixodes texanus*. Heredity 106, 253–260. https://doi.org/10.1038/hdy.2010.84

Dinnis, R.E., Seelig, F., Bormane, A., Donaghy, M., Vollmer, S.A., Feil, E.J., Kurtenbach, K., Margos, G., 2014. Multilocus sequence typing using mitochondrial genes (mtMLST) reveals geographic population structure of *Ixodes ricinus* ticks. Ticks Tick-Borne Dis. 5, 152–160. https://doi.org/10.1016/j.ttbdis.2013.10.001

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., Ovenden, J.R., 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. Mol. Ecol. Resour. 14, 209–214. https://doi.org/10.1111/1755-0998.12157

Dray, S., Dufour, A.-B., 2007. The ade4 Package: Implementing the Duality Diagram for Ecologists. J. Stat. Softw. 22. https://doi.org/10.18637/jss.v022.i04

Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4, 359–361. https://doi.org/10.1007/s12686-011-9548-7

Ehrmann, S., Liira, J., Gärtner, S., Hansen, K., Brunet, J., Cousins, S.A.O., Deconchat, M., Decocq, G., De Frenne, P., De Smedt, P., Diekmann, M., Gallet-Moron, E., Kolb, A., Lenoir, J., Lindgren, J., Naaf, T., Paal, T., Valdés, A., Verheyen, K., Wulf, M., Scherer-Lorenzen, M., 2017. Environmental drivers of *Ixodes ricinus* abundance in forest fragments of rural European landscapes. BMC Ecol. 17. https://doi.org/10.1186/s12898-017-0141-0

Estrada-Peña, A., Gray, J.S. & Kahl, O. 1996. Variability in cuticular hydrocarbons and phenotypic discrimination of *Ixodes ricinus* populations (Acarina: Ixodidae) from Europe. *Exp Appl Acarol* 20**,** 457–466. https://doi.org/10.1007/BF00053309

Estrada-Peña, A., Daniel, M., Frandsen, F., Gern, L., Gettinby, G., Gray, J.S., Jaenson, T.G.T., Jongejan, F., Kahl, O., Korenberg, E., Mehl, R., Nuttall, P.A., 1998. *Ixodes ricinus* Strains in Europe. Zentralblatt für Bakteriologie 287, 185–189. https://doi.org/10.1016/S0934-8840(98)80119-9

Estrada-Peña, A., Nava, S., Petney, T., 2014. Description of all the stages of *Ixodes inopinatus* n. sp. (Acari: Ixodidae). Ticks Tick-Borne Dis. 5, 734–743. https://doi.org/10.1016/j.ttbdis.2014.05.003

European Centre for Disease Prevention and Control and European Food Safety Authority. Tick maps 2019. https://ecdc.europa.eu/en/disease-vectors/surveillance-and-disease-data/tick-maps (accessed 12 January 2020).

613 Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of
614 individuals using the software structure: a simulation study. Mol. Ecol. 14,
615 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x
616 Gao, G.F., Hussain, M.H., Reid, H.W., Gould, E.A., 1993. Classification of a new
617 member of the TBE *Flavivirus* subgroup by its immunological, pathogenetic
618 and molecular characteristics: identification of subgroup-specific
619 pentapeptides. Virus Res. 30, 129–144. https://doi.org/10.1016/0168-
620 1702(93)90002-5
621 Gilbert, L., Aungier, J., Tomkins, J.L., 2014. Climate of origin affects tick (*Ixodes*
622 *ricinus*) host-seeking behavior in response to temperature: implications for
623 resilience to climate change? Ecol. Evol. 4, 1186–1198.
624 https://doi.org/10.1002/ece3.1014
625 Gooding, R.H., 1996. Genetic Variation in Arthropod Vectors of Disease-Causing
626 Organisms: Obstacles and Opportunities. Clin. Microbiol. Rrev. 9, 301-320.
627 Goudet, J. (2003). FSTAT (version 2.9.4), a program to estimate and test population
628 genetics parameters. Available from:
629 https://www2.unil.ch/popgen/softwares/fstat.htm
630 Goudet, J. & Jombart, T. (2018). hierfstat: Estimation and Tests of Hierarchical F-
631 Statistics. http://www.r-project.org, http://github.com/jgx65/hierfstat
632 Handeland, K., Qviller, L., Vikøren, T., Viljugrein, H., Lillehaug, A., Davidson, R.K.,
633 2013. *Ixodes ricinus* infestation in free-ranging cervids in Norway—A study
634 based upon ear examinations of hunted animals. Vet. Parasitol. 195, 142–149.
635 https://doi.org/10.1016/j.vetpar.2013.02.012
636 Hasle, G., Bjune, G., Edvardsen, E., Jakobsen, C., Linnehol, B., Røer, J.E., Mehl, R.,
637 Røed, K.H., Pedersen, J., Leinaas, H.P., 2009. Transport of ticks by migratory
638 passerine birds to Norway. J. Parasitol. 95, 1342–1351.
639 https://doi.org/10.1645/GE-2146.1
640 Helyar, S.J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M.I., Ogden, R., Limborg,
641 M.T., Cariani, A., Maes, G.E., Diopere, E., Carvalho, G.R., Nielsen, E.E., 2011.
642 Application of SNPs for population genetics of nonmodel organisms: new
643 opportunities and challenges. Mol. Ecol. Resour. 11, 123–136.
644 https://doi.org/10.1111/j.1755-0998.2010.02943.x
645 Hijmans, R. J. (2017). geosphere: Spherical Trigonometry. R package version 1.5-7.
646 https://CRAN.R-project.org/package=geosphere
647 Hillyard PD. 1996. Ticks of North-West Europe. In: Barnes RSK, Crothers JH, editors.
648 Synopses of the British Fauna (New Series).
649 Hvidsten, D., Frafjord, K., Gray, J.S., Henningsson, A.J., Jenkins, A., Kristiansen, B.E.,
650 Lager, M., Rognerud, B., Slåtsve, A.M., Stordal, F., Stuen, S., Wilhelmsson, P.,
651 2020. The distribution limit of the common tick, *Ixodes ricinus*, and some
652 associated pathogens in north-western Europe. Ticks Tick-Borne Dis. 11,
653 101388. https://doi.org/10.1016/j.ttbdis.2020.101388
654 Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and
655 permutation program for dealing with label switching and multimodality in

analysis of population structure. Bioinformatics 23, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jiang, Y., Bolnick, D.I., Kirkpatrick, M., 2013. Assortative Mating in Animals. Am. Nat. 181, E125–E138. https://doi.org/10.1086/670160

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11, 94. https://doi.org/10.1186/1471-2156-11-94

Jore, S., Viljugrein, H., Hofshagen, M., Brun-Hansen, H., Kristoffersen, A.B., Nygård, K., Brun, E., Ottesen, P., Sævik, B.K., Ytrehus, B., 2011. Multi-source analysis reveals latitudinal and altitudinal shifts in range of *Ixodes ricinus* at its northern distribution limit. Parasit. Vectors 4. https://doi.org/10.1186/1756-3305-4-84

Kempf, F., de Meeûs, T., Arnathau, C., Degeilh, B., McCoy, K.D., 2009. Assortative Pairing in *Ixodes ricinus* (Acari: Ixodidae), the European Vector of Lyme Borreliosis. J. Med. Entomol. 46, 471–474. https://doi.org/10.1603/033.046.0309

Kempf, F., De Meeûs, T., Vaumourin, E., Noel, V., Taragel'ová, V., Plantard, O., Heylen, D.J.A., Eraud, C., Chevillon, C., McCoy, K.D., 2011. Host races in *Ixodes ricinus*, the European vector of Lyme borreliosis. Infect. Genet. Evol. 11, 2043–2048. https://doi.org/10.1016/j.meegid.2011.09.016

Kriz, B., Daniel, M., Benes, C., Maly, M., 2014. The Role of Game (Wild Boar and Roe Deer) in the Spread of Tick-Borne Encephalitis in the Czech Republic. Vector-Borne Zoonotic Dis. 14, 801–807. https://doi.org/10.1089/vbz.2013.1569

Kozakiewicz, C.P., Burridge, C.P., Funk, W.C., VandeWoude, S., Craft, M.E., Crooks, K.R., Ernest, H.B., Fountain-Jones, N.M., Carver, S., 2018. Pathogens in space: Advancing understanding of pathogen dynamics and disease ecology through landscape genetics. Evol. Appl. 11, 1763–1778. https://doi.org/10.1111/eva.12678

Lang, K.R., Blanchong, J.A., 2012. Population genetic structure of white-tailed deer: Understanding risk of chronic wasting disease spread. J. Wildl. Manag. 76, 832–840. https://doi.org/10.1002/jwmg.292

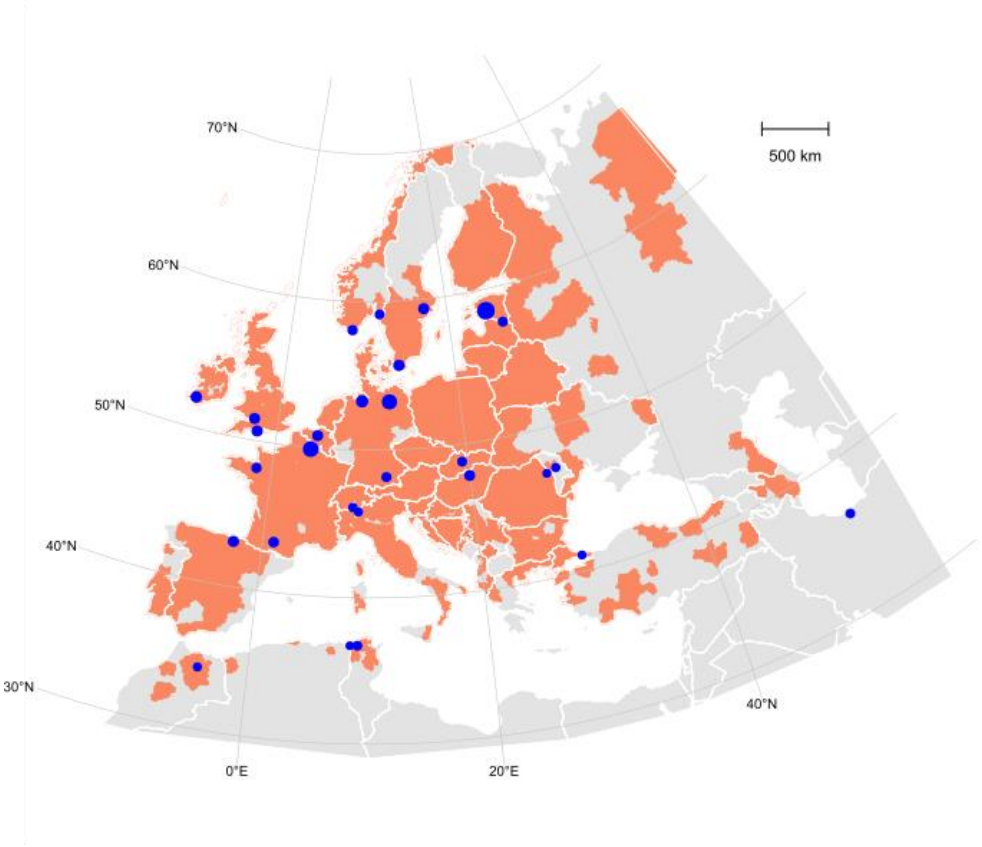
Lindgren, E., Gustafson, R., 2001. Tick-borne encephalitis in Sweden and climate change. The Lancet 358, 16–18. https://doi.org/10.1016/S0140-6736(00)05250-8

Manangwa, O., De Meeûs, T., Grébaut, P., Ségard, A., Byamungu, M., Ravel, S., 2019. Detecting Wahlund effects together with amplification problems: Cryptic species, null alleles and short allele dominance in *Glossina pallidipes* populations from Tanzania. Mol. Ecol. Resour. 19, 757–772. https://doi.org/10.1111/1755-0998.12989

698 Marin, M.S., McKenzie, J., Gao, G.F., Reid, H.W., Antoniadis, A., Gould, E.A., 1995.
699 The virus causing encephalomyelitis in sheep in Spain: a new member of the
700 tick-borne encephalitis group. Res. Vet. Sci. 58, 11–13.
701 https://doi.org/10.1016/0034-5288(95)90081-0
702 Mccoy, K.D., Boulinier, T., Tirard, C., Michalakis, Y., 2001. Host specificity of a
703 generalist parasite: genetic evidence of sympatric host races in the seabird tick
704 *Ixodes uriae*. J. Evol. Biol. 14, 395-405. https://doi.org/10.1046/j.1420-
705 9101.2001.00290.x
706 Medlock, J.M., Hansford, K.M., Bormane, A., Derdakova, M., Estrada-Peña, A.,
707 George, J.-C., Golovljova, I., Jaenson, T.G., Jensen, J.-K., Jensen, P.M., 2013.
708 Driving forces for changes in geographical distribution of *Ixodes ricinus* ticks in
709 Europe. Parasit. Vectors 6, 1. https://doi.org/10.1186/1756-3305-6-1
710 Norte, A.C., de Carvalho, I.L., Ramos, J.A., Gonçalves, M., Gern, L., Núncio, M.S.,
711 2012. Diversity and seasonal patterns of ticks parasitizing wild birds in western
712 Portugal. Exp. Appl. Acarol. 58, 327–339. https://doi.org/10.1007/s10493-012-
713 9583-4
714 Noureddine, R., Chauvin, A., Plantard, O., 2011. Lack of genetic structure among
715 Eurasian populations of the tick *Ixodes ricinus* contrasts with marked
716 divergence from north-African populations. Int. J. Parasitol. 41, 183–192.
717 https://doi.org/10.1016/j.ijpara.2010.08.010
718 Oksanen, F., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D.,
719 Minchin, F., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, H.,Szoecs, E.,
720 Wagner, H., 2019. vegan: Community Ecology Package. R package version 2.5-
721 5. https://CRAN.R-project.org/package=vegan
722 Pecl, G.T., Araújo, M.B., Bell, J.D., Blanchard, J., Bonebrake, T.C., Chen, I.-C., Clark,
723 T.D., Colwell, R.K., Danielsen, F., Evengård, B., Falconi, L., Ferrier, S., Frusher,
724 S., Garcia, R.A., Griffis, R.B., Hobday, A.J., Janion-Scheepers, C., Jarzyna, M.A.,
725 Jennings, S., Lenoir, J., Linnetved, H.I., Martin, V.Y., McCormack, P.C.,
726 McDonald, J., Mitchell, N.J., Mustonen, T., Pandolfi, J.M., Pettorelli, N.,
727 Popova, E., Robinson, S.A., Scheffers, B.R., Shaw, J.D., Sorte, C.J.B., Strugnell,
728 J.M., Sunday, J.M., Tuanmu, M.-N., Vergés, A., Villanueva, C., Wernberg, T.,
729 Wapstra, E., Williams, S.E., 2017. Biodiversity redistribution under climate
730 change: Impacts on ecosystems and human well-being. Science 355, eaai9214.
731 https://doi.org/10.1126/science.aai9214
732 Pérez-Eid, C., 2007. Les tiques - Identification, biologie, importance médicale et
733 vétérinaire. Lavoisier, Provigny.
734 Portillo, A., Santibáñez, P., Palomar, A.M., Santibáñez, S., Oteo, J.A., 2018.
735 '*Candidatus* Neoehrlichia mikurensis' in Europe. New Microbes New Infect. 22,
736 30–36. https://doi.org/10.1016/j.nmni.2017.12.011
737 Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure
738 using multilocus genotype data. Genetics. 155, 945-959.
739 https://www.genetics.org/content/155/2/945

740 Pritchard, J.K., Wen., X., Falush, D. 2007. Documentation for structure software:
741 version 2.2.
742 https://web.stanford.edu/group/pritchardlab/software/structure22/readme.
743 pdf (accessed 20 November 2019)
744 Quillery, E., Quenez, O., Peterlongo, P., Plantard, O., 2014. Development of
745 genomic resources for the tick *Ixodes ricinus*: isolation and characterization of
746 single nucleotide polymorphisms. Mol. Ecol. Resour. 14, 393–400.
747 https://doi.org/10.1111/1755-0998.12179
748 R Core Team (2019). R: A language and environment for statistical computing. R
749 Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-
750 project.org/
751 Rizzoli, A., Hauffe, H.C., Tagliapietra, V., Neteler, M., Rosà, R., 2009. Forest Structure
752 and Roe Deer Abundance Predict Tick-Borne Encephalitis Risk in Italy. PLoS
753 ONE 4, e4336. https://doi.org/10.1371/journal.pone.0004336
754 Røed, K.H., Hasle, G., Midthjell, V., Skretting, G., Leinaas, H.P., 2006. Identification
755 and characterization of 17 microsatellite primers for the tick, *Ixodes ricinus*,
756 using enriched genomic libraries: PRIMER NOTE. Mol. Ecol. Notes 6, 1165–
757 1167. https://doi.org/10.1111/j.1471-8286.2006.01475.x
758 Røed, K.H., Kvie, K.S., Hasle, G., Gilbert, L., Leinaas, H.P., 2016. Phylogenetic
759 Lineages and Postglacial Dispersal Dynamics Characterize the Genetic
760 Structure of the Tick, *Ixodes ricinus*, in Northwest Europe. PLOS ONE 11,
761 e0167450. https://doi.org/10.1371/journal.pone.0167450
762 Rousset, F., 1997. Genetic Differentiation and Estimation of Gene Flow from
763 FStatistics Under Isolation by Distance. Genetics 145, 1219-1228.
764 https://www.genetics.org/content/145/4/1219
765 Séré, M., Thévenon, S., Belem, A.M.G., De Meeûs, T., 2017. Comparison of different
766 genetic distances to test isolation by distance between populations. Heredity
767 119, 55–63. https://doi.org/10.1038/hdy.2017.26
768 Schulz, M., Mahling, M., Pfister, K., 2014. Abundance and seasonal activity of
769 questing *Ixodes ricinus* ticks in their natural habitats in southern Germany in
770 2011. J. Vector Ecol. 39, 56–65. https://doi.org/10.1111/j.1948-
771 7134.2014.12070.x
772 Smouse, P.E., 2010. How many SNPs are enough? Mol. Ecol. 19, 1265–1266.
773 https://doi.org/10.1111/j.1365-294X.2010.04555.x
774 Tabachnick, W.J., Black, W.C., 1995. Making a case for molecular population genetic
775 studies of arthropod vectors. Parasitol. Today 11, 27–30.
776 https://doi.org/10.1016/0169-4758(95)80105-7
777 Tomkins, J.L., Aungier, J., Hazel, W., Gilbert, L., 2014. Towards an Evolutionary
778 Understanding of Questing Behaviour in the Tick *Ixodes ricinus*. PLoS ONE 9,
779 e110028. https://doi.org/10.1371/journal.pone.0110028

780 Vähä, J.-P., Erkinaro, J., Niemelä, E., Primmer, C.R., 2007. Life-history and habitat
781 features influence the within-river genetic structure of Atlantic salmon. Mol.
782 Ecol. 16, 2638–2654. https://doi.org/10.1111/j.1365-294X.2007.03329.x
783 Van Zee, J., Piesman, J.F., Hojgaard, A., Black IV, W.C., 2015. Nuclear Markers Reveal
784 Predominantly North to South Gene Flow in *Ixodes scapularis*, the Tick Vector
785 of the Lyme Disease Spirochete. PLOS ONE 10, e0139630.
786 https://doi.org/10.1371/journal.pone.0139630
787 Weir, B.S., Cockerham, C.C., 1984. Estimating F-Statistics for the Analysis of
788 Population Structure. Evolution 38, 1358-1370. https://doi:10.1111/j.1558-
789 5646.1984.tb05657.x
790 Younsi, H., Fares, W., Cherni, S., Dachraoui, K., Barhoumi, W., Najjar, C., Zhioua, E.,
791 2020. *Ixodes inopinatus* and *Ixodes ricinus* (Acari: Ixodidae) Are Sympatric Ticks
792 in North Africa. J. Med. Entomol. 57, 952–956.
793 https://doi.org/10.1093/jme/tjz216
794 Welinder-Olsson, C., Kjellin, E., Vaht, K., Jacobsson, S., Wenneras, C., 2010. First
795 Case of Human "*Candidatus* Neoehrlichia mikurensis" Infection in a Febrile
796 Patient with Chronic Lymphocytic Leukemia. J. Clin. Microbiol. 48, 1956–1959.
797 https://doi.org/10.1128/JCM.02423-09
798 Wessels, C., Matthee, S., Espinaze, M.P.A., Matthee, C.A., 2019. Comparative
799 mtDNA phylogeographic patterns reveal marked differences in population
800 genetic structure between generalist and specialist ectoparasites of the
801 African penguin (*Spheniscus demersus*). Parasitol. Res. 118, 667–672.
802 https://doi.org/10.1007/s00436-018-6150-x
803 Wonham, M.J., Lewis, M.A., Renclawowicz, J., van den Driessche, P., 2006.
804 Transmission assumptions generate conflicting predictions in host-vector
805 disease models: a case study in West Nile virus. Ecol. Lett. 9, 706–725.
806 https://doi.org/10.1111/j.1461-0248.2006.00912.x



807

808 **Figure 1**. Distribution of the sampled populations of *Ixodes ricinus* across its
809 putative range. The range of *I. ricinus* is displayed in dark orange on the map and
810 was adapted from the European Centre for Disease Prevention and Control – ECDC
811 (January 2019). The size of each blue dot on the map is proportional to the sample
812 size of each sampled population.

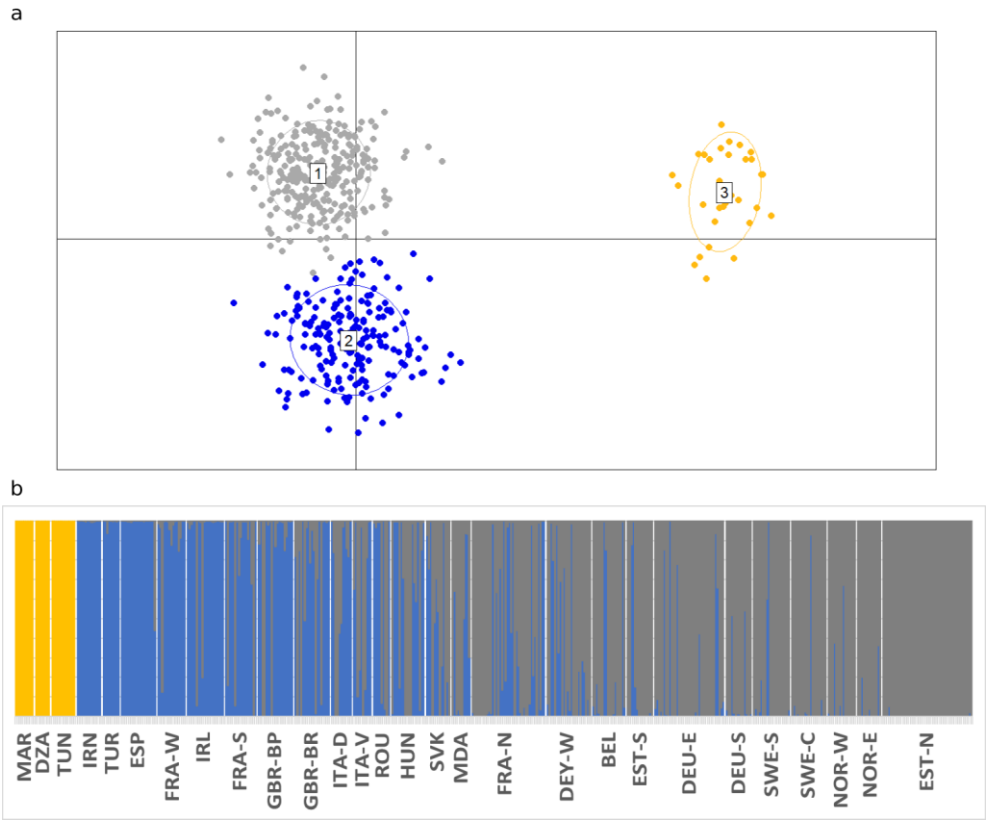
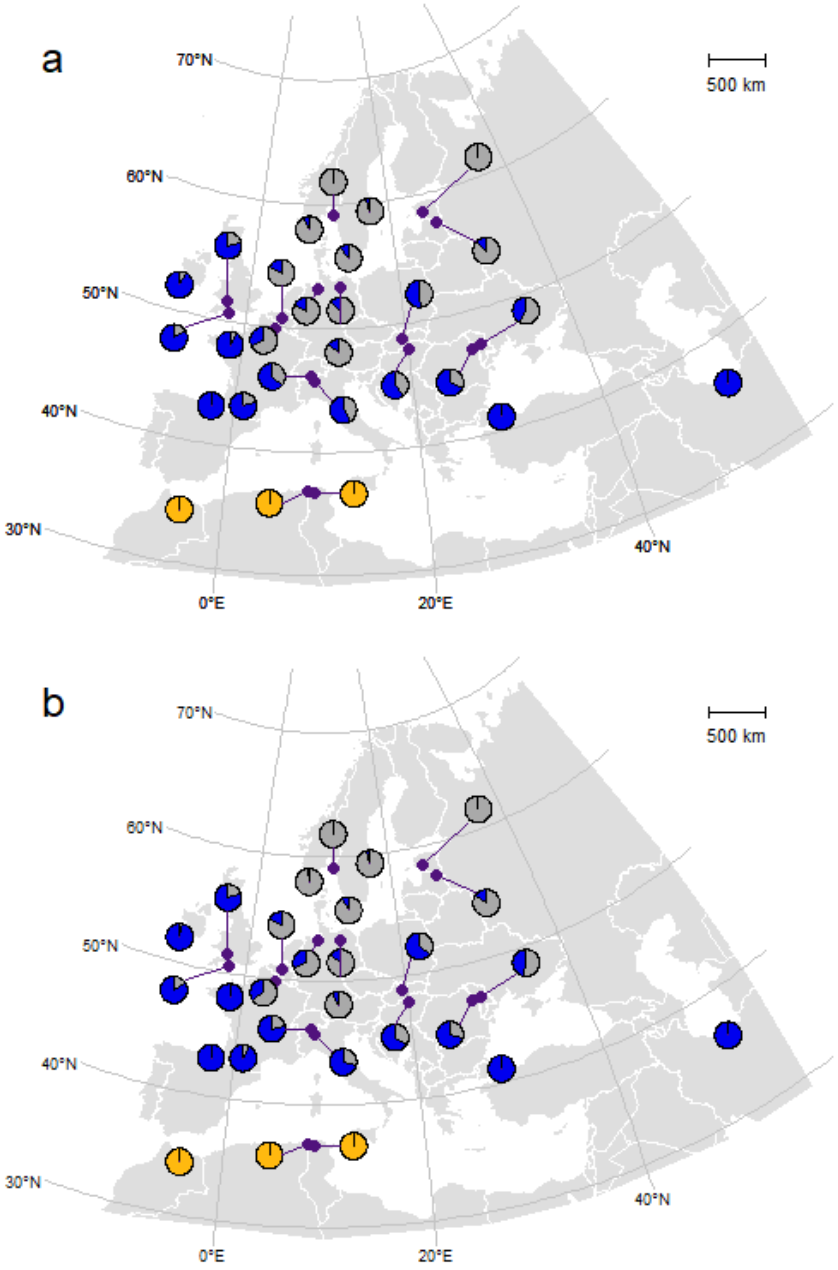


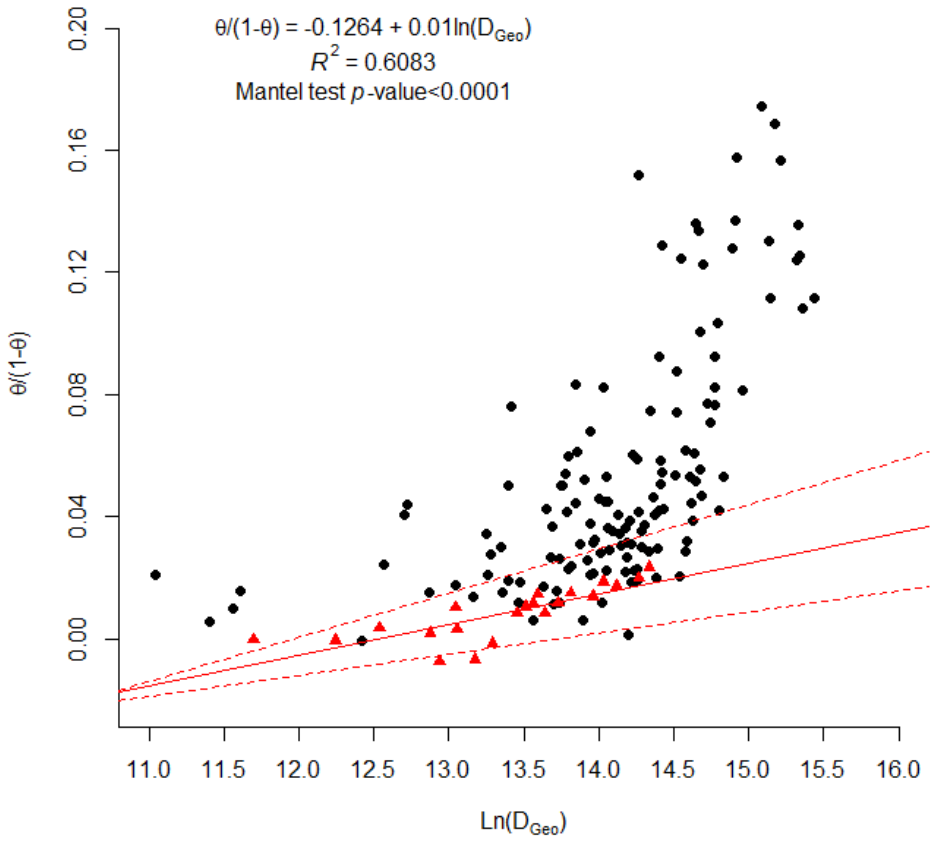
Figure 2. Cluster assignment analysis results based on either the DAPC scatter plot of individual memberships for *K* = 3 (a) or the STRUCTURE individual membership probabilities for *K* = 3 as described by Evanno et al. (2005) (b).The sampled populations are coded as follows: MAR: Morocco; DZA: Algeria; TUN: Tunisia; ESP: Spain; IRN: Iran; TUR: Turkey; FRA-W: West France; IRL: Ireland; FRA-S: South France; GBR-BP: England Blue Pool; GBR-BR: England Bristol; ITA-D: Italy Domodossola; ITA-V: Italy Varese; ROU: Romania; HUN: Hungary; SVK: Slovakia; MDA: Moldavia; FRA-N: North France; DEU-W: West Germany; BEL: Belgium; EST-S: South Estonia: DEU-E: East Germany; DEU-S: South Germany; SWE-S: South Sweden; SWE-C: Central Sweden; NOR-So: Norway Søgne; NOR-Gr: Norway

824 Grønnsundfjellet; EST-N: North Estonia. Coordinates of sampled populations are
825 presented in Table S1.



826

827 **Figure 3.** Distribution of the relative importance of each cluster on each sampled
828 population (see Figure 2 for the groups which colors are matching). Results are
829 provided for both the DAPC (a) and the STRUCTURE (b) analysis.

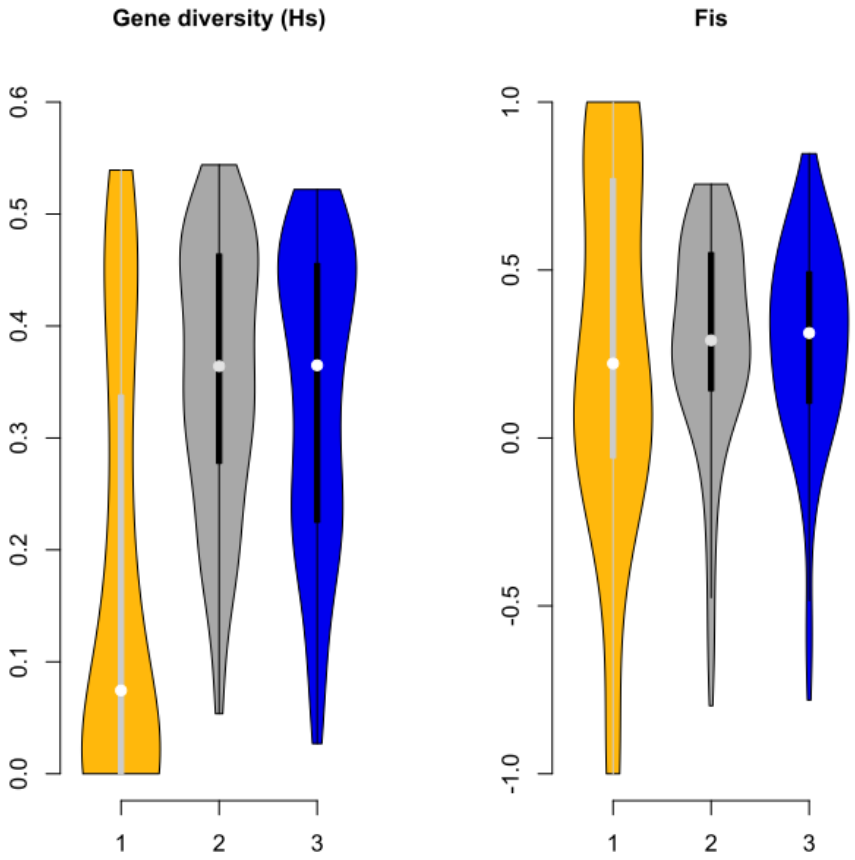


830

831 **Figure 4**. Isolation by distance between all Eurasian samples. Red triangles
832 represent the pair of samples from the same year: South and North France,
833 Belgium, West and East German, North Estonia, South and Central Sweden. The
834 regression line (plain line), 95% confidence interval (CI) calculated by bootstrap
835 (dashed lines), Mantel test significance and regression equation corresponds only

836 to red triangles pairs of samples are also shown. Black points correspond to all other

837 pairs of samples not used for further IBD analysis.



838

839 **Figure 5.** Values of gene diversity (a) and F_{IS} (b) for each of the three genetic clusters

840 identified by DAPC. Yellow: northern Africa cluster; Blue: southern Eurasia; Grey:

841 northern Europe. Permutation test (Monte-Carlo test, 999 replicates) between all

842 pairs of clusters was significant for gene diversity ($p = 0.001$) but no significance was

843 identified for F_{IS}. Eurasian clusters show a more pronounced heterozygote excess

844 than the northern African one. A variation of F_{IS} values across loci was observed in

the three clusters, even though this variation was much larger in the northern African cluster.

Supporting Information for:

Strong genetic structure among populations of the tick Ixodes ricinus across its range: insights from population genetics

Pedro Poli, Jonathan Lenoir, Olivier Dr. Plantard, Steffen Ehrmann, Knut H. Røed, Hans Petter Leinaas, Marcus Panning, Annie Guiller

Table of Contents:

Materials and Methods

Table S1. Sample coordinates. The Reference column indicates from which source samples were made available. PC = Personal collection.

Sample locality	Code	Longitude	Latitude	Number of samples	Sample Date	Reference
Morocco	MAR	4221933.21	1519759.51	10	Before 2010	Dr. Plantard, PC
Algeria	DZA	4165854.78	1520079.18	8	Before 2010	Dr. Plantard, PC
Tunisia	TUN	4287083.09	1370080.62	13	Before 2010	Dr. Plantard, PC
Spain	ESP	3292343.37	2302053.84	19	Before 2010	Dr. Plantard, PC
Iran	IRA	7920535.19	2511813.36	13	Before 2010	Dr. Plantard, PC
Turkey Istambul	TUR	5907775.11	2200447.26	9	Before 2010	Dr. Plantard, PC
North France	FRA-N	3872010.67	2994279.45	40	2013	Erhmann et al., 2018
West France	FRA-W	3465235.38	2853298.78	15	2016	Dr. Degeilh, PC
South France	FRA-S	3593881.21	2296634.56	17	2013	Erhmann et al., 2018
Ireland	IRL	3013710.61	3385835.15	20	Before 2010	Dr. Plantard, PC
England Blue Pool	GBR-BP	3470079.25	3130233.33	19	Before 2010	Dr. Plantard, PC
England Bristol	GBR-BR	3450947.31	3224484.53	19	Before 2010	Dr. Plantard, PC
Italy Domodossola	ITA-D	4188665.99	2556599.15	11	Before 2010	Dr. Plantard, PC
Italy Varese	ITA-V	4229419.76	2523525.45	10	Before 2010	Dr. Plantard, PC
Romania	ROU	5643875.12	2813096.13	9	Before 2010	Dr. Plantard, PC
Hungary	HUN	5064737.95	2796444.23	18	Before 2010	Dr. Plantard, PC
Slovakia	SVK	5008087.64	2900574.08	13	Before 2010	Dr. Plantard, PC
Moldavia	MDA	5711169.6	2856440.17	10	Before 2010	Dr. Plantard, PC

Sample locality	Code	Longitude	Latitude	Number of samples	Sample Date	Reference
West Germany	DEU-W	4257417.83	3352915.67	24	2013	Erhmann et al., 2018
East Germany	DEU-E	4462732.5	3348531.08	38	2013	Erhmann et al., 2018
South Gemany	DEU-S	4440340.3	2784710.43	14	2013	Dr. Plantard, PC
Belgium	BEL	3924610.12	3095109.35	18	2013	Erhmann et al., 2018
North Estonia	EST-N	5186688.83	4032319.73	49	2013	Erhmann et al., 2018
South Estonia	EST-S	5313297.86	3950296.69	14	Before 2010	Dr. Plantard, PC
South Swqeen	SWE-S	4533959.53	3622513.31	20	2013	Erhmann et al., 2018
Central Sweden	SWE-C	4720133.45	4047795.89	19	2013	Erhmann et al., 2018
Norway West	NOR-W	4186225.49	3886420.36	15	2006	Dr. Leinaas, PC
Norway East	NOR-E	4389275.28	4003811.98	13	2006	Dr. Leinaas, PC

Table S2. List of SNPs, variant basis and primers used in the study (from Quillery et al., 2014)

locus	variant	Primer1-specific to allele1	Primer2-specific to allele2	Primer3-common to both alleles
1133	T/C	GCTTGGCCACTTCCACTGCTTT	GCTTGGCCACTTCCACTGCTTC	ACAACAGAGAAGGCAGCCCACA
3705	A/C	AGCATGGCGCACTGTGAAAGCTC	AGCATGGCGCACTGTGAAAGCTA	TCCTAGTCGGCTGGCTGGAG
6283	T/C	AATGAGGCGTCAGTGACAGCATAAC	AATGAGGCGTCAGTGACAGCATAAT	CGTGACGTCAAGGCAGAATGCTAT
6363	A/G	TCGTCCTCCGTCACGTAGCCG	TCGTCCTCCGTCACGTAGCCA	CCATTGAACCCTGGTGGGTCATCA
10041	A/G	GTTGTTCCCTTGGCAGACG	GTTGTTCCCTTGGCAGACA	AACATACCCGAGACTGTCAAC
19998	A/G	CAGAAGTGGAGATTGTTGCGTGTG	CAGAAGTGGAGATTGTTGCGTGTA	TACATACATTGAGCATCGACCAA
21130	C/T	GCTGCTGCAACCGGTTTATCTTC	GCTGCTGCAACCGGTTTATCTTT	AGGCACGTAGATCACGAGAATTATT TC
30736	C/G	GCTAGGTGACGAGGACTGGACG	GCTAGGTGACGAGGACTGGACC	GTTGTTCCACCTTTCGCAGGAGAT
31200	A/G	CGTTCAGGTTGACCGAGAAGTAA	GTTCAGGTTGACCGAGAAGTAG	GCCTCTCGTTACTGTCGTATC
32114	C/T	GACTAATCACCAGGAAATCCATTCTG C	GACTAATCACCAGGAAATCCATTCTG T	GGCTATACTCGGACGTATGTTGA
32551	T/C	TTCGGTGGCAACAGCTCGTCCATC	TTCGGTGGCAACAGCTCGTCCATT	CCAGCCTCATAGCCGAGCACCA
34502	G/A	CGGATTCGAACCAGTTATCAATGGG	CGGATTCGAACCAGTTATCAATGGA	GCCTCTCTAGAAAACAGTTGCTCTC
42351	A/G	CTTGTAGGAATGGAGGTCATCTTCG	CTTGTAGGAATGGAGGTCATCTTCA	CTTCTGTGTCGCAGGTGGCATCAT
57206	C/G	GCACTATGAGCCATCGAAGCCAAG	GCACTATGAGCCATCGAAGCCAAC	ACGTGACAACACTTACACGGCATTT C
60684	C/T	TGCACATAGTCGCGCAATACGTTC	TGCACATAGTCGCGCAATACGTTT	CGAGCCGTTGCAACCGATCCG
61606	G/A	ACATAGGACATCTCAAGGTCATTCG	ACATAGGACATCTCAAGGTCATTCA	GAAGAAACCGAGGATGAGTGTCATG
66390	C/T	GCCGAACAGCCGTGCAACCC	GCCGAACAGCCGTGCAACCT	TCGCTGCTGTATACCCATTG
68328	G/C	CAGGCAGTTTGCGGTTCACAG	CAGGCAGTTTGCGGTTCACAC	TAGAGGTTTCCCAAGTATTTATCGT A
68391	A/G	CAGCGTCAAGTTGTGGTGTT	CAGCGTCAAGTTGTGGTGTC	GCATCGCGTGACATTAGTTACA
72226	G/A	GAGGTTCCTGACATGCAGGAAACG	GAGGTTCCTGACATGCAGGAAACA	GCTCTGCAGATGCAAGTTCCAA
77668	G/A	GGAACGTCGTGACAGCCGTAG	GGAACGTCGTGACAGCCGTAA	GGATGGCTTCGAGTTGGACTACTA
78934	G/C	AAAGAAGCGTTTCCCGGTCG	AAAGAAGCGTTTCCCGGTCC	TCTGGCAAAGCAAGCACTCACC
81501	T/C	GTCCTTTCGAAGGTGTATGCATTC	GTCCTTTCGAAGGTGTATGCATTT	ACGATGCTAGTTTGTCAAATAGTG

locus	variant	Primer1-specific to allele1	Primer2-specific to allele2	Primer3-common to both alleles
81758	G/A	ACAAATCTGAAGCAGGCGCGAAAG	ACAAATCTGAAGCAGGCGCGAAAA	AGGACGTCGCCGAGTCGTAGAT
87199	T/C	GCTGGATTGCGTCGTCGCCT	GCTGGATTGCGTCGTCGCCC	CGGCTCTGGCCAGGACCTGATG
93695	G/A	GTCCTAGCCGCTGTCCCGTG	GTCCTAGCCGCTGTCCCGTA	CTGGGACAAACTCTTTCTCGAAGTG
96296	G/A	GCATAAGCAAACTTCAAAGCTTCCACG	GCATAAGCAAACTTCAAAGCTTCCACA	ACGAGGCGGCTCTCATGTACCA
105385	T/G	CCGCGAGCATTTTTGCCACATG	CCGCGAGCATTTTTGCCACATT	TTGACGTCACGACCTATTTGACGAA
113142	A/T	GAGCTCATAGTCCTGAAGACCACA	GAGCTCATAGTCCTGAAGACCACT	TTACGTTGGTCACTATGGGAACGCT
114791	G/C	CGCTGCTAGCAGACGGGAGG	CGCTGCTAGCAGACGGGAGC	GAGAGCGTACACGATTTGCCACGA
116335	A/C	GTGCGTCGAATGTCCAGGTTTATCC	GTGCGTCGAATGTCCAGGTTTATCA	CAAGTTGCGCAAGAGGTGGCAAA
125671	C/T	GTCTGCTTCTGCTATGCTCTGTTTC	GTCTGCTTCTGCTATGCTCTGTTTT	AGCGTCTGCTGCGGAACATCGTA
129322	T/A	CAAGGCAGCGCAGTTCTGACACT	CAAGGCAGCGCAGTTCTGACACA	ATCTGCGTAGCATAAGCCGTGCC
133049	G/A	ACGGGTCGTACAGCGACAAGAG	ACGGGTCGTACAGCGACAAGAA	CGAACATTACAAACGCCGCAAGAGG
137096	T/G	GTGAATGGCAATGCCAGAGTGTAT	GTGAATGGCAATGCCAGAGTGTAG	CTCGGTATTCTGCGGAGCACAA
143089	G/A	GGCACAGGATTTGCTGGTTATAGAGG	GGCACAGGATTTGCTGGTTATAGAGA	GGTGCTATGTGTACCTCACGCC
144259	C/T	GTTGAGTGTCGTGTCCTTCGCC	GTTGAGTGTCGTGTCCTTCGCT	AACAGCTCCTCGTAGACTGCGTAC
145634	C/T	CGGACGCGTGGACGTGACTC	CGGACGCGTGGACGTGACTT	TGGTGACCGTGTGTTGCGCAG
150669	T/C	TGTGCACAAGATGATTCCATAATT	TGTGCACAAGATGATTCCATAATC	GTCATCGGTGATTGTGTCAGTTTAT
155043	G/A	GAATGTGATCGTGGGAGAAGATATAGG	GAATGTGATCGTGGGAGAAGATATAGA	GCTGTGGAAGCTAAGTGCTCGTTG
159151	C/G	AGACAACGTACGCGCGATTTCAC	AGACAACGTACGCGCGATTTCAG	TGCTAACTGCCAGCGCGTGG
166766	A/G	ATCGACCGGCTGGCTGGCTA	ATCGACCGGCTGGCTGGCTG	GCCTGTTCTTCTGTAAGTCGCTCTA
167418	T/A	TGTCCGATACCTGCCTCCAATTTGTT	TGTCCGATACCTGCCTCCAATTTGTA	TTACCTCCACCGGGTGTCCCAT
175115	T/C	ATGGCAGTGTCAAGAAGGCCAAGT	ATGGCAGTGTCAAGAAGGCCAAGC	CAATGGCAGTGTCAAGGTCGATCTC
176991	C/A	AGAAGCTAGACGCAGAGTTAGGGC	AGAAGCTAGACGCAGAGTTAGGGA	AGGAAGAGTCCAATGTGTGCGCAA
180239	G/T	GTCCTGTGCTGTTGCCGCCG	GTCCTGTGCTGTTGCCGCCT	TGTTCCTGGACGCAAGTCACG
189207	T/A	TGGGCGTTGCAGTAATGCAACAGTT	TGGGCGTTGCAGTAATGCAACAGTA	TCTAAGGCTCCTGGTGTAAGCACACG
197784	C/T	GTTCATTAGAAGCTGTCAGTTGACTC	GTTCATTAGAAGCTGTCAGTTGACTT	CAGTGGCGTAACACGAGAAACTAG
198227	C/T	GACAACATCCAGGGCGAGTTCTAC	GACAACATCCAGGGCGAGTTCTAT	TTGCTATAACCAGTCTTCGACGC

locus	variant	Primer1-specific to allele1	Primer2-specific to allele2	Primer3-common to both alleles
205578	G/A	GATGTAGCCCCAGATATACTCAAAG	GATGTAGCCCCAGATATACTCAAAA	ACAGGTACTAAACCAATTTCGGC
207995	A/C	CGAGGTAAGATTGCCACTTATCTTTCC	CGAGGTAAGATTGCCACTTATCTTTCA	ACCACCTGCCAGTGTTCGACGAT
208593	C/G	GGTCTGGTGCCTGGAAAGTGC	GGTCTGGTGCCTGGAAAGTGG	GGACGCAGTAAACAGAGCAGTCATA
209761	C/T	ACATCATAAGTCACGTGGCCTGAC	ACATCATAAGTCACGTGGCCTGAT	ACGCCGTGACGTCTCCTGAT
210654	T/C	GTGATTCTGCTGGTGATCTTTGTGATC	GTGATTCTGCTGGTGATCTTTGTGATT	AGCACGCCCAACAAGATCAACGG
212829	C/G	GGCATCTGAACGACATCGTCCACC	GGCATCTGAACGACATCGTCCACG	CGTGTGTCAGGAATGAGAGATAATC
214684	T/C	GTAACGCCGTCACACGGTAAGAC	GTAACGCCGTCACACGGTAAGAT	CTGTCTGATCCAGGCTTTACGCAA
221603	T/C	AGTCGATCATACCTTACTGCTGTGT	AGTCGATCATACCTTACTGCTGTGC	TTCGCGAGTCCGAGTTGCACAGA
224277	C/A	ACAGCTAGGAGCAAAGTCCAGTTCCC	ACAGCTAGGAGCAAAGTCCAGTTCCA	CTATTCCCCTTTCGATCGAACATCGG
225377	G/A	TAAAGAGTCGCCTTGGGGAATCTGG	TAAAGAGTCGCCTTGGGGAATCTGA	CACGGACAACAACATTGAACGAG
230247	T/G	GTTTCCAGCTCGCGGTCGATT	GTTTCCAGCTCGCGGTCGATG	GACTGCGTAGAGTGCGCTTTTCAA
233961	A/C	GTCATGCATTTGACAAACTTTGTTA	GTCATGCATTTGACAAACTTTGTTC	GACACTACTAGGGCCTCAATCAA
234508	C/T	TGCTGTCGCTACGCTCGACC	TGCTGTCGCTACGCTCGACT	GAGAGCAGCTCCTGGGAGTCCTTG
236290	T/C	GATGCAATATGTTTACTGGATTCGC	GATGCAATATGTTTACTGGATTCGT	TAGAAATCGGGGCCCCAACGG
243436	T/C	CTTGTGCCTGGCGTCATCTGT	CTTGTGCCTGGCGTCATCTGC	AGGCCCGTGCTCGCTCG
251320	T/A	AGGATCACGTTATACGAAGGCAAGT	AGGATCACGTTATACGAAGGCAAGA	CAAGGATGACAGCACCGGTACGA
255757	T/G	TTCATCGGCGTATCCTTTGAGCGAT	TTCATCGGCGTATCCTTTGAGCGAG	ATGATGGCGACGTAGAGGTAGTTCA
259770	C/G	ACCCTTTTTGAAAGATGAACGTTGTC	ACCCTTTTTGAAAGATGAACGTTGTG	CGTTGCTCAAAGTCAAATGCCAGTG
281206	T/G	GACACTACTAGGGCCTCAATCAAGCAT	GACACTACTAGGGCCTCAATCAAGCAG	CAGTCATGCATTTGACAAACTTTG
283680	T/A	GGCGAAACCTTTGAAGCGTTCTTCAT	GGCGAAACCTTTGAAGCGTTCTTCAA	GACAGCGTGATGACTGTTCTTGTG
287805	T/G	CTGCCGCCTGTAATTCCCGACT	CTGCCGCCTGTAATTCCCGACG	TAGGTTCACGACACGAGGTTGATTC
292025	C/T	AACGCCGTGAAAGCCGCGAAC	AACGCCGTGAAAGCCGCGAAT	GCACACCGTACATCACCGAAGCC
296275	C/A	CTGCGTAGAGTGCGCTTTTCAAGGTC	CTGCGTAGAGTGCGCTTTTCAAGGTA	TCGTTTGGTTTCCAGCTCGCGGT
298125	A/G	TTTGTTCAGTTGTCAGAGGTGGCAGTA	TTTGTTCAGTTGTCAGAGGTGGCAGTG	CCTTGTGGCATGCTCCAGTGATTC

locus	variant	Primer1-specific to allele1	Primer2-specific to allele2	Primer3-common to both alleles
299627	C/T	GGTATCCGCTCGCTCGATATGTATATC	GGTATCCGCTCGCTCGATATGTATATT	CGTGTGCAGCTATCCAAAGACTCG
300752	T/G	AGATGCTGAACTGTCAGATGACGAAT	AGATGCTGAACTGTCAGATGACGAAG	ACCACTGTAGTTGTGTCTCGCTCTG
303781	C/G	CTCCAATTAGCTTCAAATGAATGTTC	CTCCAATTAGCTTCAAATGAATGTTG	CTTGGTTAGTTTCTGCTGGCGTTTTC
305888	C/A	GTTTCCTCCACGCAGAGCGAAAGA	GTTTCCTCCACGCAGAGCGAAAGC	CATGCGCTTCGCACTGTCG
307361	T/C	GCGGTATTTTCGGTCAGGC	GCGGTATTTTCGGTCAGGT	GACAAATGTTCGTCGTTCTCAACAG
313057	A/C	AATAGCGGCCAGCAGTTCCTCATA	AATAGCGGCCAGCAGTTCCTCATC	CGAATCCGATAGTGCCGTGAGAGA
320000	A/C	CAAATTTCGTGTTCGTCCATGGCGTGA	CAAATTTCGTGTTCGTCCATGGCGTGC	CGTGACTTGACGTGACGTGCCA
329834	T/G	TAGAAAGCCGGCCCGGATCTT	TAGAAAGCCGGCCCGGATCTG	CTTTCCCAGTTCAAGCACTCTTTTAG
333882	T/C	GCTCCTCCATGTCTTGTCGTCGTTTCT	GCTCCTCCATGTCTTGTCGTCGTTTCC	CACGGTGGCAGCGGGAA
336267	G/T	GCGTTGTCTGTACATCCGCCAT	GCGTTGTCTGTACATCCGCCAG	GAGCGCAGCGGATACTCTGTTCA
339272	A/G	CCGCACCGGCTTTTACGACA	CCGCACCGGCTTTTACGACG	TCTCGTCGCTGGAGGCGTCAT
340581	C/T	CTGAACCCAACGTTGGCTGAACT	CTGAACCCAACGTTGGCTGAACC	ACTGAGTGGTTCTAGTAACGATGGCT
356074	G/A	AAGTATGGGGGAACCCGTGTGA	AAGTATGGGGGAACCCGTGTGG	TAGGAGTTGGAACACTGCGACG
356395	G/A	CATTTGCGATAGGTCGATCACGATATG	CATTTGCGATAGGTCGATCACGATATA	CCGACTTCCGACGCATGTAAAATG
371093	A/G	AGCGATGGCGTCTACCAGCGGA	AGCGATGGCGTCTACCAGCGGG	TTCTGGACTAGCAGCGAGCGAC
374382	T/C	CATGCTTTGTCAACTTTCGAGAT	CATGCTTTGTCAACTTTCGAGAC	TTATGCTGTCAGCTGAGTCCCG
376474	T/C	AGGTGGCCACTCTGACATGGATC	AGGTGGCCACTCTGACATGGATT	TGTAGAGTGTAGATGCCAGCTTCCTC
380487	C/T	CAGCCGTTCGACGGGATC	CAGCCGTTCGACGGGATT	TCGCTCGTGTCCCTCGTGT
393248	T/G	CTGCATGTCTTGGCGTCTGATGTCTTCT	CTGCATGTCTTGGCGTCTGATGTCTTCG	GGTTCACTGGCCAAACGCTCCTCTAC
399212	A/G	GTTCAATGGGGCTTCTGCTATCA	GTTCAATGGGGCTTCTGCTATCG	GCGTGAATTCAACGTTCGCTAAG
411541	G/A	AGTCGTTGTGGGCGCGCATGGG	AGTCGTTGTGGGCGCGCATGGA	GTCAGGCTGTTCGGCTTGACGTATG

locus	variant	Primer1-specific to allele1	Primer2-specific to allele2	Primer3-common to both alleles
419658	T/G	TGTCCTCGTACGTGCTCGTTGTGACT	TGTCCTCGTACGTGCTCGTTGTGACG	AGCAGATGGCCTGGTAGCGGTCC
428503	G/A	CATGCAGGATACCGTGTGAGTTCAG	CATGCAGGATACCGTGTGAGTTCAA	GATGCTGTGCGCGTTGGACTG
438644	A/G	GCACTGCAAACACCTCTGCTCAAGTATG	GCACTGCAAACACCTCTGCTCAAGTATA	CTATGAATGCTCTTGCTAGCAGGCTTTA
441042	A/G	GAATTCCAAACGCGGTTCATAAACCACG	GAATTCCAAACGCGGTTCATAAACCACA	TCGAAGATAGTGTGCTCAATGGCGGTTA
446758	T/A	TTGTTGCGAACATAGAGTACAGAGGAGCA	TTGTTGCGAACATAGAGTACAGAGGAGCT	GCTACAACGTGGGAATTGCCGAGGA
450975	T/G	TGCGGTTACGCAGTCGAAGCTATT	TGCGGTTACGCAGTCGAAGCTATG	ATGGGCACTCAAGGTGCGCACG
465604	A/T	CCTAAACGTCTCGGCGCTAATA	CCTAAACGTCTCGGCGCTAATT	AACTAAGACCACATTCCCGACATTG
465892	G/A	CCCACTGACGAGCGTGCTGAAGA	CCCACTGACGAGCGTGCTGAAGG	CATGCTCTTTCCTGTTGTCCGGTTCA
468480	A/G	CATAACGCTGAATTATCTTCGCCGACTA	CATAACGCTGAATTATCTTCGCCGACTG	GTAAGGGGCCCACAAGCCTGG
480915	A/G	CTAATTCTCGTTCTACTGCCGCATG	CTAATTCTCGTTCTACTGCCGCATA	GGACACATCTCAGAACCAGATTG
487540	T/C	CACGGGAACGACGGGCACT	CACGGGAACGACGGGCACC	GGCACGTGAAGCTCCGAGATTTCAT
493429	A/G	TAGTGGGTTCGCTGAAGAACTACAAGAA	TAGTGGGTTCGCTGAAGAACTACAAGAG	CGCGCAGCTTTCTGAAGTAGTTGT
552113	T/A	TCATAGTTGGTTCACAGGCGACCT	TCATAGTTGGTTCACAGGCGACCA	GTTCTGGACTAAGTATGATTCGCTCCA
558063	A/G	CAGCTCCTGGGAGTCCTTGAGA	CAGCTCCTGGGAGTCCTTGAGG	AGTGGCTGCTGTCGCTACGCT
561492	T/C	ATCTTGCGACTGCTCGAT	ATCTTGCGACTGCTCGAC	TTCTCGCCCAGGAATGCCAT
580716	T/C	TCGGCGTTCAGCAGGCTTGAC	TCGGCGTTCAGCAGGCTTGAT	GCACCAGACCGCCGGCGA
583125	T/G	TGTTCTGAGGAAATGAGATGACTGTT	TGTTCTGAGGAAATGAGATGACTGTG	CAACACACGTCAACAGCAACAT
585284	T/A	GCTTCAGTTATCAGCTGTAAACCTA	GCTTCAGTTATCAGCTGTAAACCTT	TTCGGTAATGCGTGTATTACTCA
585318	A/G	GTACATCACCGAAGCCGAACAG	GTACATCACCGAAGCCGAACAA	TTAGCCGCAACGCCGTGAAA
589219	C/T	ATGCCGCACGTGCTTGAGGTC	ATGCCGCACGTGCTTGAGGTT	CAAGAAACGGCAACAGCGGACAATGAAC
627150	A/C	CAATACAGCGGTATTTGCACTA	CAATACAGCGGTATTTGCACTC	CAATGGAGCAGACGCATCT
751708	G/A	TTGAAGCACAGCTCTTAGAGAAGG	TTGAAGCACAGCTCTTAGAGAAGA	GACTCCGTCAGCTGGTTTATG

locus	variant	Primer1-specific to allele1	Primer2-specific to allele2	Primer3-common to both alleles
754496	C/G	GCCTCGGCGTCGGAACTCG	GCCTCGGCGTCGGAACTCC	TGGCTGAAACCAGGGACCTCAA
761047	G/A	CAACATGGACGTTTTCAAGATTGCCA	CAACATGGACGTTTTCAAGATTGCCG	GAGCCTCGCTCAGCACGGAA
763022	T/C	CACAAAGGGCACGATTTCCTCT	CACAAAGGGCACGATTTCCTCC	AGATGAGTCTGCCATCGTGTCT
764527	T/A	GGGCGTTGCAGTAATGCAACAGTA	GGGCGTTGCAGTAATGCAACAGTT	AAGGCTCCTGGTGTAAGCACACG
767569	A/G	AAACACACCTTGAACTCAGCCTCA	AAACACACCTTGAACTCAGCCTCG	GGACGACAGCTATCAACATTAGCC
768618	C/G	GAACAATTCAAAACCATGATTGAAAC AC	GAACAATTCAAAACCATGATTGAAAC AG	TACACTCCCAAGTGAGTTGATGC
771828	T/C	GATCCAAAGTGATCATGCCGATAGT	GATCCAAAGTGATCATGCCGATAGC	ATATCACAGTATCACGTCACGG
775381	A/G	TGTGCAGCTATCCAAAGACTCGG	TGTGCAGCTATCCAAAGACTCGA	ATGGTATCCGCTCGCTCGATATGT
777961	C/G	CTCAGCACAAGTGAATGTCAAG	CTCAGCACAAGTGAATGTCAAC	GGGCATTTGTAAGCATCTTATCGC
781023	G/T	GGCTCTATGTAGAACCAAAGATAAGT GAG	GGCTCTATGTAGAACCAAAGATAAGT GAT	ATTCTGCGGCTTCAACGAATCA
783090	G/A	ACCCGTACAGCAAACCACTACG	ACCCGTACAGCAAACCACTACA	CGACTGATTTCTCGCAACCCA
792422	T/C	TGCCACGGTAGTTTTGCTTAGT	TGCCACGGTAGTTTTGCTTAGC	ATGTTCCACGAGGCCCGTTG
43247	C/T	AGTAGACTTAAAGGCCACGCTCGAC	AGTAGACTTAAAGGCCACGCTCGAT	CCTTATATTCTCTGTCAGCGTAAG
84140	T/C	CAATCGAAATCGTGACCAATGGGATT C	CAATCGAAATCGTGACCAATGGGATT T	ACCAAGTGCCGCGCAAAGCAT
117944	C/T	CGAATTCGAAGGCGGAGATCCTC	CGAATTCGAAGGCGGAGATCCTT	CGGCTTGGCGAAGCGACG
316915	T/G	CGCTTCGCCGAGCACTCG	CGCTTCGCCGAGCACTCT	ACCGGTTGTGCTACGCGTAGGT
197588	T/G	CAAGCGCATCCCCATTCTGATCTT	CAAGCGCATCCCCATTCTGATCTG	CTTAGAAAGGCAAGACCTCCTTCA
2932	C/T	CTCCTACGAGGGGTGCCTGT	CTCCTACGAGGGGTGCCTGC	TTGTGACGTTCCTCGTGCTCCCT
112567	C/T	GCTCATGCGCATTGGAAGC	GCTCATGCGCATTGGAAGT	TTGCACGTACTACGTGCCTCTG
207179	T/C	CGCACGGAGATGGCATTCCTC	CGCACGGAGATGGCATTCCTT	ACACGATCTTCGGCGAGAACGTCA
165428	G/C	GTCCGCCACGTCGGTTCCAGAG	GTCCGCCACGTCGGTTCCAGAC	AAGCGGGGCTCTGCTTCCGCCT
109194	C/T	AGGCCCACAACTCCACTCTTC	AGGCCCACAACTCCACTCTTT	TACGGTAGCTATGTAACAGACACTA
139650	C/T	TACGACGGCACCGAGATC	TACGACGGCACCGAGATT	ATCTCCGGCGAGGCGTACA
56083	T/C	CCAGGCGCTCCTCCTCGGTC	CCAGGCGCTCCTCCTCGGTT	CGCCGGAGTTGGCCCAGGA
143860	A/G	ACAGGTACACGAACGATCGCAGAA	ACAGGTACACGAACGATCGCAGAG	TGCGTTCGTGCTTGTGTCATGT
152555	A/G	GCTCCAGGACAACCGTTTACCTCA	GCTCCAGGACAACCGTTTACCTCG	ATGGAAACATCGCTACACATGG

locus	variant	Primer1-specific to allele1	Primer2-specific to allele2	Primer3-common to both alleles
51899	A/G	GAGGTGTACGAGTGTCACTCGAAG	GAGGTGTACGAGTGTCACTCGAAA	GTATCTAGGAGGCTCGGGCGAAA
225801	C/T	GACTTCTGACATTTGATAGAATGCTC	GACTTCTGACATTTGATAGAATGCTT	TGCGGGTCAGCCATCTTACAAGTA
190468	G/A	TGAACGAAGCTGAGAGGCGCTATGA	TGAACGAAGCTGAGAGGCGCTATGG	TACGCCCAGACACTCTTGTTCAGT
31277	C/G	ATCATAGACCAACTCGCCTGCATC	ATCATAGACCAACTCGCCTGCATG	GATTCTGGAAGACAGCTTTTTCGC
455987	G/C	AATGTACGCGACGTACGCACAAG	AATGTACGCGACGTACGCACAAC	GGATTTCCGAGAGAAGCCATTTTCAG
27147	T/G	CGCAATTGTGACACCACTAG	GCGCAATTGTGACACCACTAT	CGGCTTTTGATACTCCCATCA
751588	G/T	CCGCATTTCTTCACTGCTGTTTGAAAG	CCGCATTTCTTCACTGCTGTTTGAAAT	TCGCAAATCCTGGCGCGGTAA
313642	T/A	GTGCAGTTGGCAATGGAGGTGA	GTGCAGTTGGCAATGGAGGTGT	CCGGACAACTGAAGGTGGTGC
182969	G/A	AAGACGCACTTGCCCTGGAAACATG	AAGACGCACTTGCCCTGGAAACATA	GGTCTGAGTCTTGGTTGTGTCGCAT
186625	A/G	GAGGAGCTGCGATGCAGAAGTGGTA	GAGGAGCTGCGATGCAGAAGTGGTG	ATGCTGATGACGCAACGCTGACTTC
191703	A/G	CCGCCGTCTTTGCAGCCTCA	CCGCCGTCTTTGCAGCCTCG	GGGGCCCCGATTTCTAGAAC
438440	A/G	GTTGAGCGCATGCGCAGGGAA	GTTGAGCGCATGCGCAGGGAG	ACTCCCTGACGTAGCCTTCGTAGGA
82163	T/C	TAAGGCTTCCAGGTGACTTC	CTAAGGCTTCCAGGTGACTTT	GGTGTGTTGCTTCTATATTG
788521	C/T	ACCCGAACTTTGCAGGCCAT	ACCCGAACTTTGCAGGCCAC	AATGAACGACCGAGCGAATCCAGA
233756	C/G	TCTACAAACCAGGCGGTTGTAAGC	TCTACAAACCAGGCGGTTGTAAGG	TCTGTTTGGGACTCCTTCCACCG
201653	G/A	GCAGTCATCAAACGTGATTTCGTCCG	GCAGTCATCAAACGTGATTTCGTCCA	AAATTGGAGAGATCACTTGACCCGC
259800	C/G	CGTGTGCCTCGCTGGCATC	CGTGTGCCTCGCTGGCATG	GCGCATTCCAGAGGCTTCC
370147	C/T	GACACCCTAGCAAAGCAAAGCGTTCTC	GACACCCTAGCAAAGCAAAGCGTTCTT	TTTCGTTCACGGCTCCCGCAA
153000	G/A	CCTACCTGCTTCCAACATTCTTTAGG	CCTACCTGCTTCCAACATTCTTTAGA	TGCACATTAGGTCAGAGATGCGGA
500950	A/T	CCACAACTCATCGCACCGAAGACT	CCACAACTCATCGCACCGAAGACA	AGACGATTATTCGGCTGTGACACATT
170547	C/G	GGTGAATACGCGTCGCGTGAGTC	GGTGAATACGCGTCGCGTGAGTG	GTGACCTTTGGTAGGACGGCAGC
466967	C/G	GAATATTTATGATGTGACCACGGCAAAC	GAATATTTATGATGTGACCACGGCAAAG	AACGCCCTGCCGCATAGTCC
246408	T/C	GGAAACAGTTATAACTATCTAGAACT	GGAAACAGTTATAACTATCTAGAACC	CACACCGAGAAATCAGACGTACC
5630	G/A	CAGCAAGCAGAGAACGTCGTCGATG	CAGCAAGCAGAGAACGTCGTCGATA	TTCAGGGTGAGACCGTCGGC
561563	A/G	TGAAGGATCTCGTACACAATACACAG	TGAAGGATCTCGTACACAATACACAA	CGAGTACTTCACGACCACGCA

locus	variant	Primer1-specific to allele1	Primer2-specific to allele2	Primer3-common to both alleles
338495	T/C	GGTTCTCGAAGCCGCGTTTC	GGTTCTCGAAGCCGCGTTTT	TCTGCAGCTGCTGTAGAGTCCTG
166887	A/G	TCGGCCGCCAGCAGCGTCA	CGGCCGCCAGCAGCGTCG	CCCGTCGGGAGCAATGCAG
766292	T/C	TGCCGAAGCTGGGTTTCGT	TGCCGAAGCTGGGTTTCGC	CTGGGCTGCTCCGAGGACTA
176206	G/T	ACTGCGATTGAAGTGCGTCCCG	ACTGCGATTGAAGTGCGTCCCT	ATCCTCTTGAAATTTGCTGCGGGTG
245496	T/C	TTCCAGCGTGCACCGTACC	TTCCAGCGTGCACCGTACT	GAAAATGCAATTTTTGTGAGCCT
199727	T/A	GGCTTCTTGTCTCGTTATTATCGT	GGCTTCTTGTCTCGTTATTATCGA	CAGTGCCACTTTATGTGAGTTG
524153	G/A	CTCTATCAAACGATGTGCTACTGTGA	CTCTATCAAACGATGTGCTACTGTGG	ACTAATTCATTGTAACCCATTTCACGAT
54140	A/G	GGTAGACACAATCTGCTCATAATGG	GGTAGACACAATCTGCTCATAATGA	ATGACTGTTACAATCTTTTGAATGC
18708	A/G	CTCCGCGTGTATGCGAGTGAA	CTCCGCGTGTATGCGAGTGAG	GGCGCGTATCATCCCAGAGC
546612	G/C	TTTCCCGCGCAGGCCGCTAG	TTTCCCGCGCAGGCCGCTAC	TCAAGGCCAACGGCGCGCA
523859	C/A	CTGGACCTGTGCTACCGTGAGTCC	CTGGACCTGTGCTACCGTGAGTCA	GCTCAGGATGTCGTACGCGCGG
160279	A/T	ATCAGCAGCGCACACGCTCA	ATCAGCAGCGCACACGCTCT	CGTCGACGGGCGATCGTGA
624322	A/G	TATCAGCTAAAGCCTCCTTCTCAGTCA	TATCAGCTAAAGCCTCCTTCTCAGTCG	GAACTGAAGCACCAGCGCCT
410904	C/G	GTCAGAGTAAGGATCTGCTAGATACCG	GTCAGAGTAAGGATCTGCTAGATACCC	TAAGAAGGTTGGCCCGAATTTGTGAA
71660	C/A	GAAATTAGAATGGTACCTGGATTACC	GAAATTAGAATGGTACCTGGATTACA	CCTTTGGGGTGCGCTTATGTAAT
87165	G/A	GAATCCACGTGTCAGAGCCCTGG	GAATCCACGTGTCAGAGCCCTGA	GGTTGTATTTACAACTGACTCCTCGG
61479	A/G	GGCTAATCCTGCTTCTTGGCCTT	GGCTAATCCTGCTTCTTGGCCTC	CGATCCTGAAATCGAGCAAAGCC
571455	T/A	GTTCTGCCAGCAATTCTATCACT	GTTCTGCCAGCAATTCTATCACA	GGATGGATGCAAAGTGATATTTTAG
270863	C/T	GCAATTATAGGATCTCCGTAAACTCT	GCAATTATAGGATCTCCGTAAACTCC	CCTTTTTACGGACACTCACTTTCCTG
185472	C/G	ATTCGCCAGACCACTTGGATTCTC	ATTCGCCAGACCACTTGGATTCTG	CGTTTTCAATGAGTCTTGATTCTCG
200386	T/C	GATGGAATTAGGTACGGTCATTTCAT	GATGGAATTAGGTACGGTCATTTCAC	GTTCAGCGCATACTATGACTGACAA
40367	A/G	CACATGTGGCAAGCATTCAA	CACATGTGGCAAGCATTCAG	GCAGCAACGTTTGCTTCAGA
494898	T/C	AGCGTTGCACGCCATACATTCTCT	AGCGTTGCACGCCATACATTCTCC	TCCACAGGGTCACGTGACGCA
14134	C/T	CATACATTCCCTGAATACCTAGAGC	CATACATTCCCTGAATACCTAGAGT	ATTAGCCAAGCGCCCCG
361495	T/G	ATAACACAGGCAGACATTGGAGGCAG	ATAACACAGGCAGACATTGGAGGCAT	GCTCACATGCATTGAAACTGATGTC

Results

Table S3. Basic statistics per locus.

Locus	Observed heterozygosity	Gene diversity	Fst	Fis
1133	0.1101	0.2324	-0.0040	0.5262
31200	0.4319	0.467	0.0629	0.075
66390	0.1269	0.291	0.0539	0.5639
129322	0.0700	0.0873	0.0061	0.1974
159151	0.2649	0.3446	0.1504	0.2312
198227	0.5741	0.4987	-0.0093	-0.1512
221603	0.2760	0.3557	0.0932	0.2243
251320	0.1778	0.3271	0.1478	0.4564
298125	0.5543	0.4636	0.0127	-0.1956
329834	0.0918	0.2123	-0.0187	0.5675
374382	0.176	0.4705	0.0317	0.626
3705	0.1782	0.4146	0.0673	0.5704
32114	0.0759	0.1787	-0.0022	0.5753
68328	0.2477	0.3736	0.0396	0.3369
93695	0.2975	0.4519	0.0335	0.3418
133049	0.1991	0.3770	0.1375	0.4718
255757	0.2072	0.2468	0.2394	0.1604
299627	0.1109	0.2400	-0.0107	0.5381
376474	0.2141	0.3006	0.0491	0.288
6283	0.3140	0.4164	0.0634	0.2459
32551	0.262	0.343	0.0092	0.2361
96296	0.2678	0.3734	0.2144	0.2829
137096	0.5322	0.4652	0.0382	-0.1438
207995	0.8246	0.45	0.0726	-0.8323
225377	0.2191	0.3855	0.097	0.4316
259770	0.1676	0.2408	0.0235	0.3041
300752	0.4814	0.3724	0.1811	-0.2928
336267	0.2115	0.2254	0.0116	0.0618
380487	0.3168	0.4498	0.0807	0.2957
6363	0.3144	0.4011	0.1326	0.2162
34502	0.1219	0.3181	0.0031	0.6166
105385	0.3326	0.3783	0.0686	0.1208
143089	0.7837	0.4849	0.0132	-0.6163
208593	0.1799	0.2333	0.0843	0.2288
230247	0.2052	0.3581	0.223	0.427
281206	0.2277	0.3757	0.2359	0.394
303781	0.1121	0.2912	-0.027	0.6152

Locus	Observed heterozygosity	Gene diversity	Fst	Fis
339272	0.1745	0.1785	0.0237	0.0222
393248	0.0722	0.1882	-0.0019	0.6163
176991	0.1966	0.3406	0.029	0.4228
144259	0.1649	0.2687	0.0127	0.3863
113142	0.2391	0.3714	0.2421	0.3563
77668	0.4137	0.4424	0.0191	0.0648
42351	0.128	0.1783	0.439	0.2821
10041	0.1733	0.4534	0.0484	0.6178
399212	0.0839	0.3201	0.1127	0.7377
340581	0.0682	0.0911	0.0379	0.2513
305888	0.1105	0.1722	0.0151	0.3583
283680	0.6714	0.4485	0.0759	-0.497
233961	0.2212	0.3625	0.2481	0.3898
209761	0.0828	0.3481	0.0662	0.7621
180239	0.094	0.151	0.0109	0.3773
145634	0.3787	0.4585	0.018	0.1741
114791	0.2429	0.3708	0.0262	0.3449
57206	0.1859	0.3233	0.0253	0.425
19998	0.0914	0.1196	0.0199	0.2359
411541	0.227	0.3092	0.0348	0.2658
356074	0.2633	0.3343	0.0754	0.2123
307361	0.0887	0.1801	0.0333	0.5074
287805	0.0435	0.0744	0.0317	0.4154
234508	0.2516	0.3103	0.059	0.1891
210654	0.228	0.2674	0.4378	0.1475
189207	0.1252	0.3037	0.0042	0.5879
150669	0.2338	0.4764	0.0025	0.5092
116335	0.1823	0.3998	0.0415	0.544
81501	0.2806	0.4505	0.101	0.377
60684	0.2221	0.442	0.0547	0.4974
21130	0.1723	0.412	0.0976	0.5818
356395	0.3341	0.4604	0.0791	0.2745
313057	0.1085	0.388	0.0497	0.7203
292025	0.1001	0.1208	-0.0134	0.1714
236290	0.0798	0.1775	-0.0126	0.5505
212829	0.2648	0.4684	0.0637	0.4347
197784	0.3144	0.2259	0.1932	-0.392
155043	0.1033	0.1762	0.5208	0.4136
125671	0.3398	0.4652	0.0539	0.2695

Locus	Observed heterozygosity	Gene diversity	Fst	Fis
81758	0.1436	0.2886	0.1153	0.5023
61606	0.2479	0.3716	0.2166	0.3331
428503	0.3744	0.395	0.0206	0.0521
320000	0.2643	0.2579	0.1614	-0.0248
296275	0.1384	0.3744	0.2015	0.6303
243436	0.1853	0.2832	0.3311	0.3455
214684	0.2222	0.4963	-0.0018	0.5522
438644	0.1799	0.4129	0.149	0.5644
487540	0.0981	0.1404	0.288	0.3014
767569	0.1635	0.2148	0.0246	0.2389
165428	0.211	0.2398	0.0199	0.1199
191703	0.1966	0.3729	0.2489	0.4729
153000	0.0919	0.4793	0.0403	0.8082
166887	0.1921	0.3481	0.0653	0.4483
441042	0.3243	0.4491	0.0996	0.2777
84140	0.1825	0.3554	0.0587	0.4865
438440	0.0944	0.1089	0.1972	0.1325
766292	0.738	0.4745	0.0398	-0.5553
523859	0.3518	0.4397	0.1064	0.1999
270863	0.0571	0.23	0.0146	0.7516
446758	0.0327	0.1022	0.0336	0.6802
552113	0.144	0.219	0.0062	0.3423
627150	0.1918	0.4255	0.074	0.5491
117944	0.8044	0.4649	0.0642	-0.7304
139650	0.2503	0.2708	0.1388	0.0757
176206	0.3211	0.5036	-0.0175	0.3624
185472	0.1899	0.3965	0.0348	0.5211
450975	0.2154	0.3343	0.1091	0.3557
558063	0.2358	0.2929	0.0731	0.195
751708	0.0754	0.1006	-0.026	0.2506
775381	0.0911	0.2703	-0.0258	0.6629
27147	0.1022	0.3726	0.0636	0.7257
200386	0.1562	0.4361	0.1004	0.6418
777961	0.0977	0.3904	0.1195	0.7499
754496	0.3408	0.4748	0.051	0.2822
561492	0.0813	0.3268	0.0409	0.7512
465604	0.0536	0.2293	0.0085	0.7664
410904	0.1761	0.2763	-0.0034	0.3628
199727	0.1067	0.4443	-0.0179	0.7599

Locus	Observed heterozygosity	Gene diversity	Fst	Fis
751588	0.1149	0.3961	0.0719	0.7099
152555	0.1713	0.4355	0.058	0.6067
2932	0.0525	0.0604	0.0063	0.1315
781023	0.1682	0.4324	0.0954	0.6109
761047	0.1318	0.173	-0.0271	0.2378
580716	0.316	0.3822	0.0479	0.1731
465892	0.113	0.1621	-0.0162	0.3033
5630	0.7425	0.4783	0.0441	-0.5525
313642	0.1283	0.1442	0.0071	0.1102
783090	0.2976	0.3471	0.0268	0.1426
763022	0.1247	0.302	0.0606	0.5872
583125	0.534	0.4398	0.097	-0.2141
468480	0.3318	0.4527	0.0932	0.267
14134	0.1556	0.345	0.0144	0.549
259800	0.125	0.1671	0.0074	0.2517
182969	0.2349	0.2092	0.005	-0.1228
225801	0.1255	0.254	0.3865	0.5058
792422	0.402	0.4864	-0.0019	0.1736
764527	0.1107	0.2841	0.0287	0.6103
585284	0.1384	0.2335	-0.0063	0.4072
480915	0.0849	0.2483	0.032	0.6581
338495	0.2448	0.3455	0.0157	0.2915
186625	0.5284	0.4536	0.0852	-0.1649

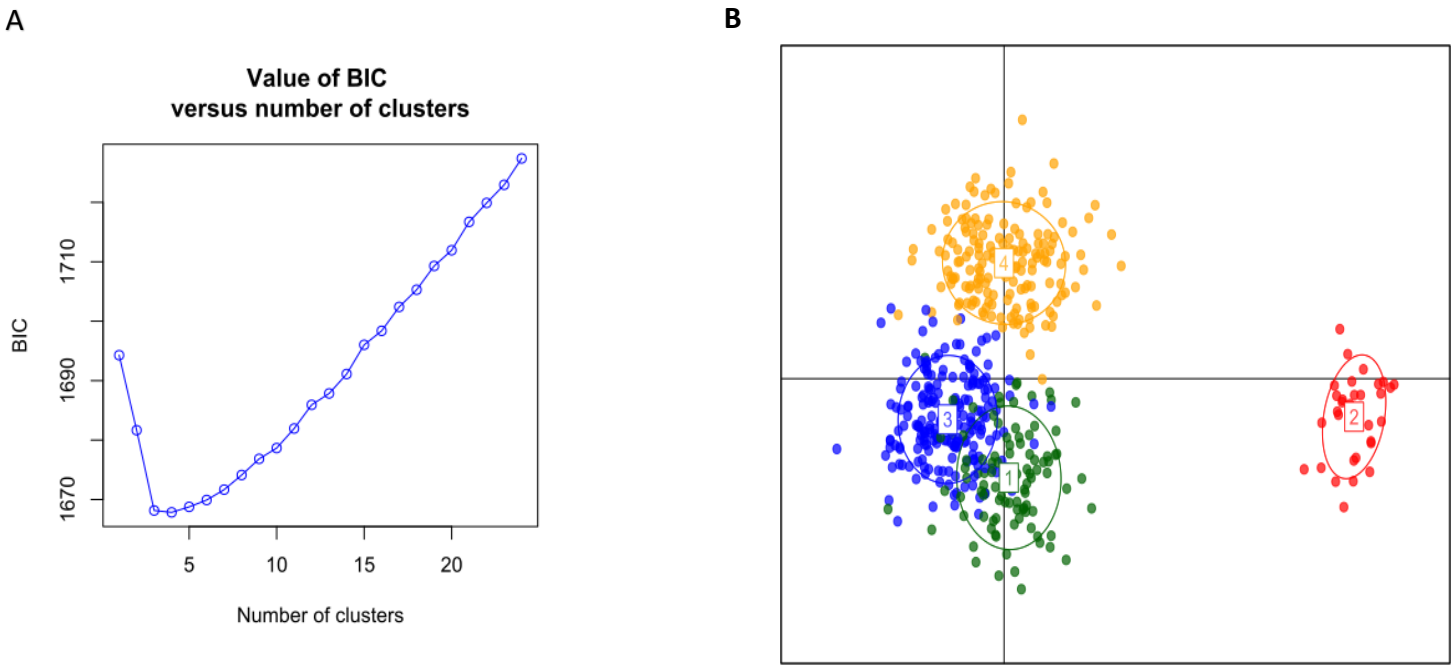


Figure S1. Discriminant analysis of principal component (DAPC) of *Ixodes ricinus* based on 497 individuals using 125 SNPs. A. BIC values as a function of the number of clusters *k*. The difference in BIC values between *k* = 3 and *k* = 4 is 0.842. B. Scatterplot of individuals on the two principal components of DAPC. The graph represents the individuals as dots and the groups as inertia ellipses. Two of the clusters overlap, while when *k* = 3 we identify 3 well separated groups (figure 3). Red : North African cluster; yellow : only individuals from southern Eurasian cluster; green : only individuals from the Northern European cluster; blue: admixture cluster with mainly individuals from the northern European cluster in fugure 3.

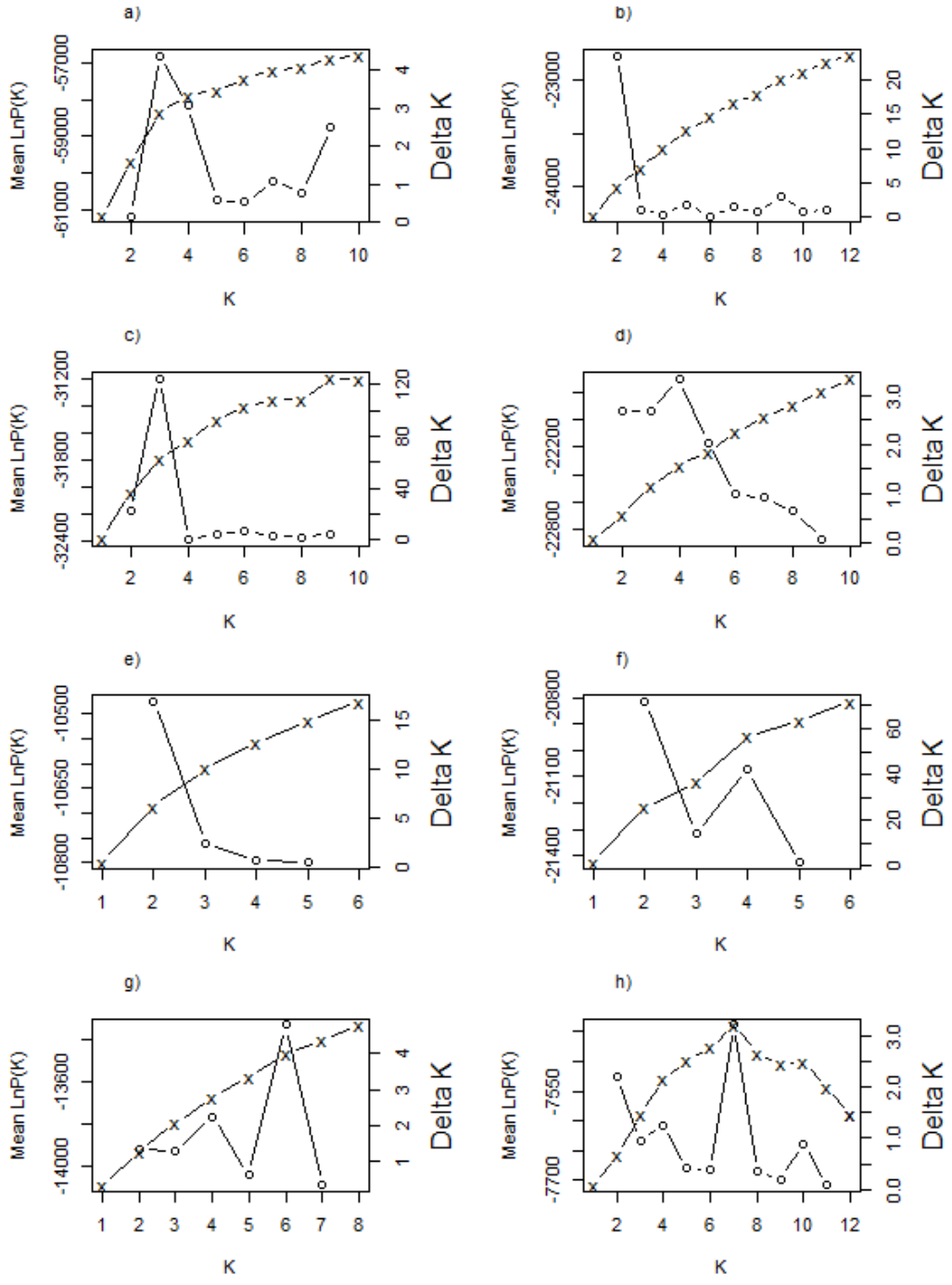


Figure S2. Probabilities ln P(*X*|*K*) for each level of hierarchical analysis. First round of analysis: a); Second round: b) southern Eurasian cluster and c) northern European clusters; Third round: d) Southern European cluster without Iran, e) Central Sweden, Norwegian West and East and North Estonia, f) Moldavia, North France, West German, Belgium, South Estonia, East German, South German and South Sweden; Forth round: g) Atlantic samples (Spain, South and West France, Ireland and England, h) South-west samples (Italy, Romania, Slovakia and Hungary), i) and i): fourth round of analysis. Details of each level of Hierarchical analysis are present in the corresponding session.

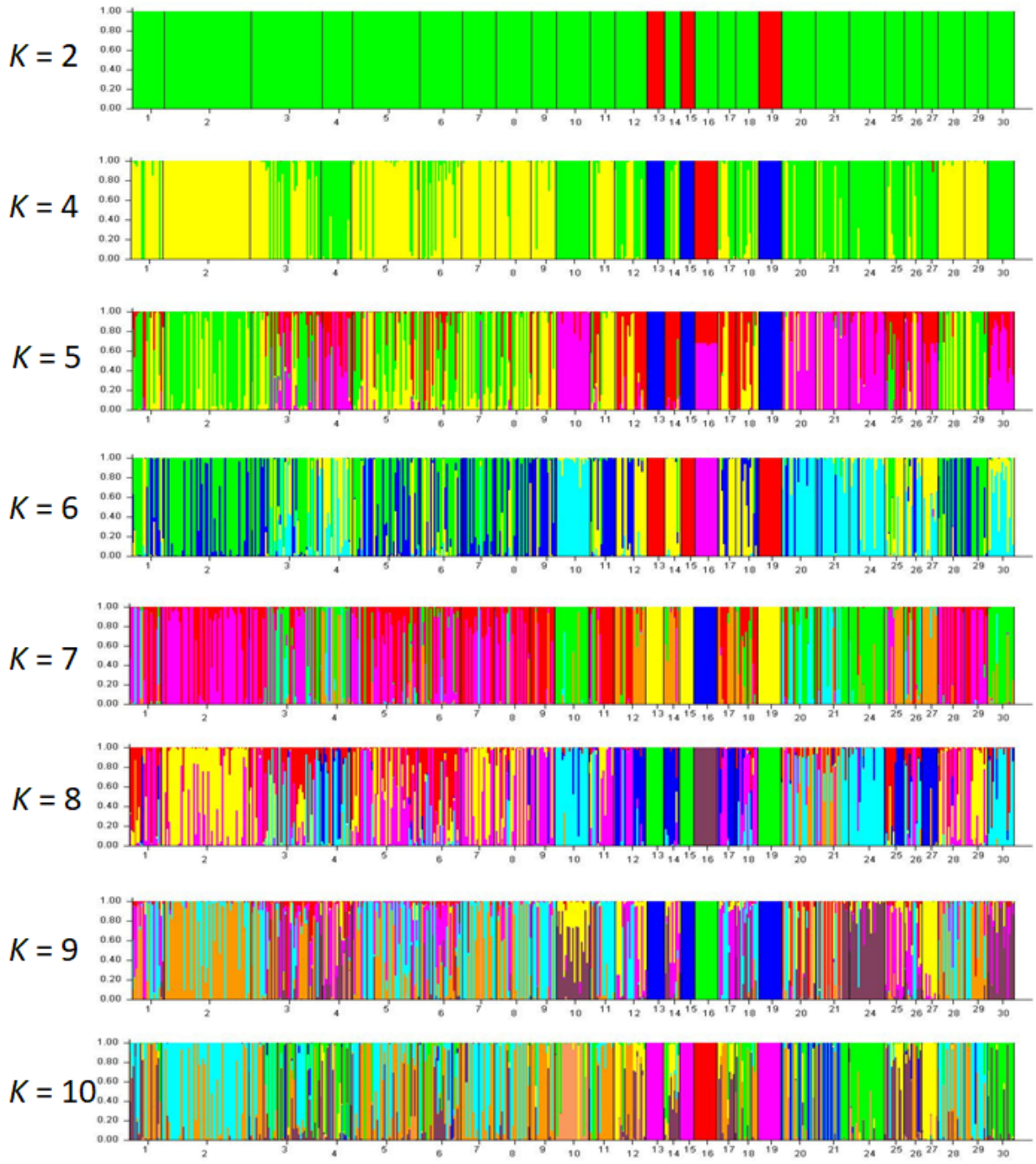


Figure S3. STRUCTURE Individual probabilities for each value of *K* from 2 to 10.

Hierarchical analysis

Finer genetic structure was identified from hierarchical analysis (Figure S6 and S7 for STRUCTURE and DAPC analysis, respectively). The southern Eurasian cluster was further separated into two differentiated clusters, irrespective of the approach used (STRUCTURE or DAPC). The STRUCTURE approach separated Iran from the remaining samples, while the DAPC approach assigned most individuals from both Iran and Turkey samples to the same cluster (violet). The northern European cluster was further separated into two to three clusters depending on the methods, DAPC and STRUCTURE, respectively. Clusters identified by the DAPC approach were distributed almost equally among the different sampled locations. Of the three clusters identified by STRUCTURE, the orange and green ones showed a clear affinity to certain sample locations, while the grey cluster was represented in all sampled locations. No further structure was identified for the African cluster in both methods.

The DAPC's third round of analysis was unable to identify further genetic structure in the northern European cluster. It did however identify two groups inside the southern Eurasian cluster (without Turkey and Iran as a result of previous analyse). It appears that individuals from Spain, Western France and Ireland were mainly assigned to one (light blue) cluster. No other cluster was identified by the DAPC approach regarding refined hierarchical analysis. The STRUCTURE's third round of analysis was able to identify a K = 4 in the southern European cluster. Individuals from Turkey were assigned to an exclusive cluster (grey). Individuals from south-western Europe and from Italy were mainly assigned to one cluster (orange), while those from Spain, West France and Ireland were grouped in a different cluster (blue). The fourth cluster (green) was distributed across all sampling locations with few individuals (11 out of 179) exhibiting more than 50% of assigning probability. In the northern European cluster, for this third

round of hierarchical analysis, individuals were regrouped according to population probabilities *of the two almost exclusive clusters from last step, green and orange ones. From this third round until the last one, Evanno's method (*Evanno et al. 2005) always identified two clusters, but the analyses of ln[Pr(X|K)] was not clear in identifying those clusters (Figure SX). Also, individual probabilities of inside those K = 2 clusters show very mixed populations. The results for those subsequent rounds with a K = 2 are presented in Supplementary Information (figure SXX). We did a fourth and last round of hierarchical analysis for the two main southern Eurasian clusters identified in the previous round: (i) one cluster composed of Spain, West and South France, Ireland and England samples and (ii) the other cluster composed of Italy, Romania, Hungary and Slovakia. For the first one, Evanno's method identified K = 6, but the analysis of ln[Pr(X|K)] does not indicate any structure. For the later, both methods clearly identified a K = 7 structuring. In both cases, clusters are mainly distributed in all sample sites and very rarely a single individual had ~100% probability of being assigned to a particular cluster. The exceptions were individuals from West France and Ireland for which probability values to be assigned to the same cluster reached one.

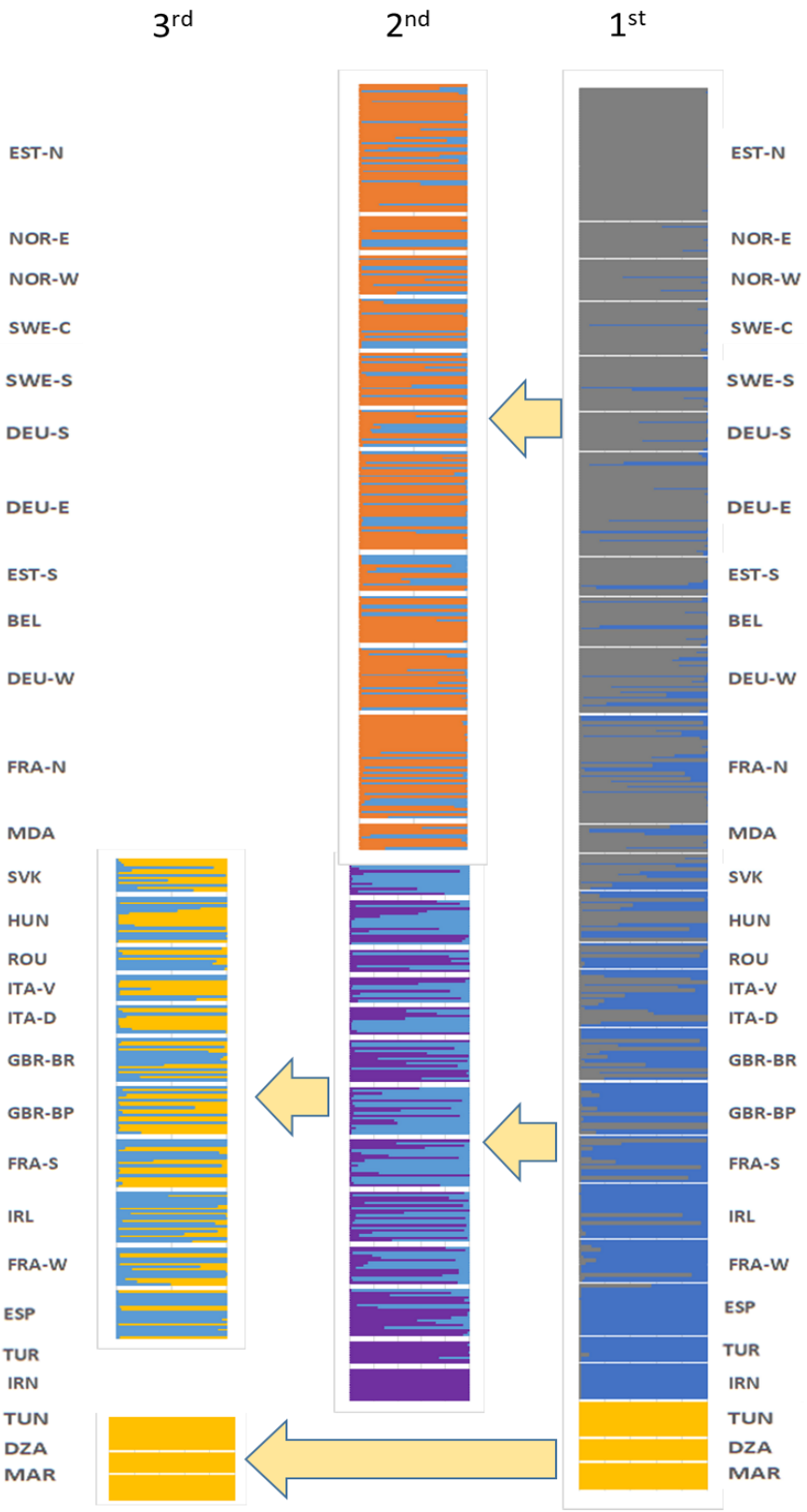


Figure S4. DAPC Hierarchical analysis. Each column corresponds to one level of analysis.

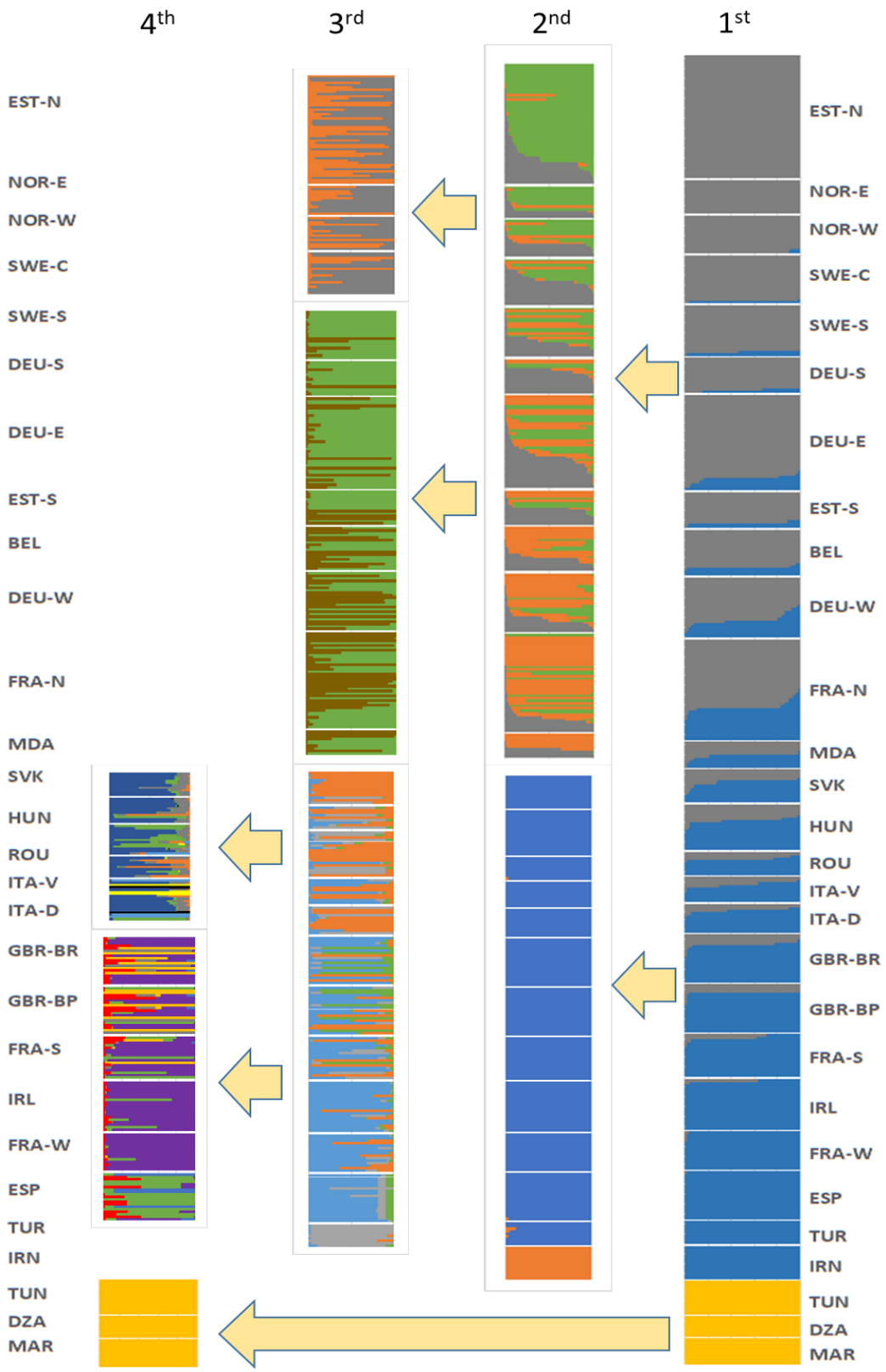


Figure S5. STRUCTURE Hierarchical analysis. Each column corresponds to one level of analysis.

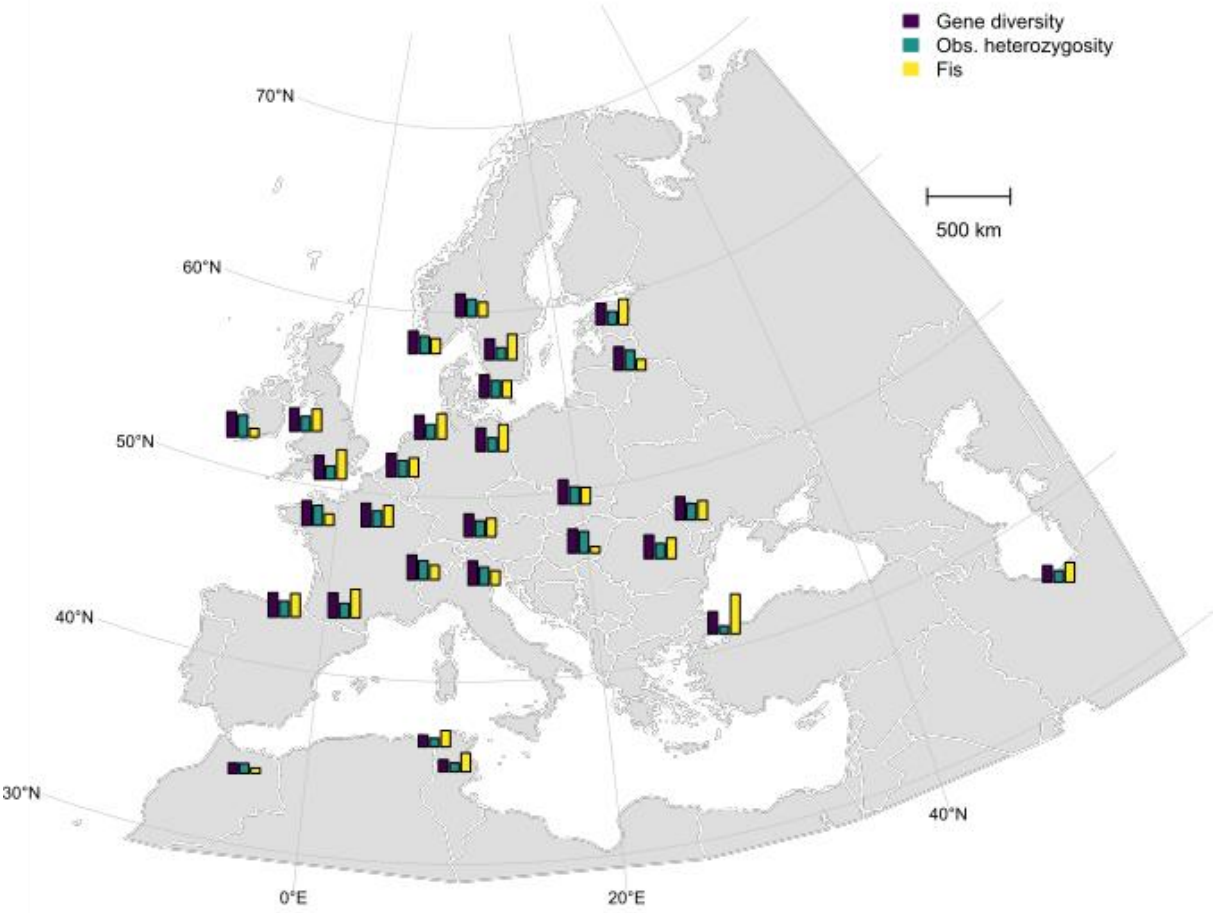


Figure S6. Mean gene diversity, observed heterozygosity, and *F*is per population. Mean population gene diversity was always greater than the observed heterozygosity and *F*is was always positive.