View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Digital.CSIC

Phylogeographic analysis of introns and mitochondrial DNA in the clam *Ruditapes decussatus* uncovers the effects of Pleistocene glaciations and endogenous barriers to gene flow

David Cordero[1], Juan B. Peña[1], Carlos Saavedra[1]

[1]Instituto de Acuicultura Torre de la Sal, Consejo Superior de Investigaciones Científicas, 12595 Ribera de Cabanes (Castellón), Spain

Correspondence:

Carlos Saavedra

Instituto de Acuicultura Torre de la Sal, 12595 Ribera de Cabanes (Castellón), Spain

E-mail: saavedra@iats.csic.es

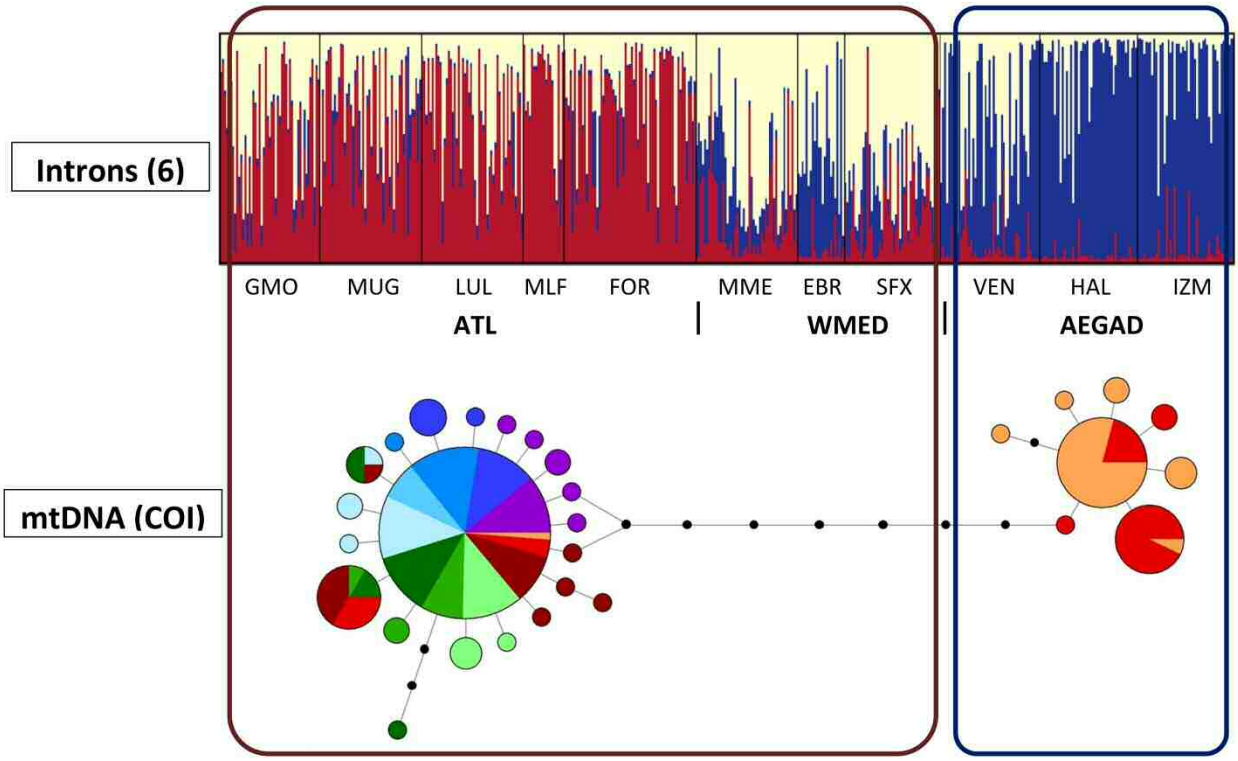
Tel: +34 964 319 500

Fax: +34 964 319 509

Highlights

- One mitochondrial gene and six introns showed genetic subdivision of *R. decussatus* populations into three regions
- COI and two introns showed haplotypes/alleles restricted to the Aegean and Adriatic Seas (AEGAD)
- COI and introns showed contrasting patterns of differentiation between Atlantic and W Mediterranean populations
- An endogenous barrier could be limiting gene flow at some nuclear loci in the Atlantic-W Mediterranean boundary

Graphical Abstract



Abstract

Studies on the phylogeography of species inhabiting the Mediterranean and the nearby coasts of the NE Atlantic Ocean (MEDAT) have found subdivision and/or phylogeographic structure in one or more of the Atlantic, western Mediterranean and eastern Mediterranean basins. This structure has been explained as the result of past population fragmentation caused by Pleistocene sea level changes and current patterns of marine circulation. However, the increasing use of nuclear markers has revealed that these two factors alone are not enough to explain the phylogeographic patterns, and an additional role has been suggested for endogenous barriers to gene flow or natural selection. In this article we examined the role of these factors in *Ruditapes decussatus*, a commercial clam species native to MEDAT. A genetic analysis of 11 populations was carried out by examining 6 introns with a PCR-RFLP technique. We found subdivision in three regions: Atlantic (ATL), western Mediterranean plus Tunisia (WMED), and Aegean and Adriatic seas (AEGAD). Two introns (*Ech* and *Tbp*) showed alleles that were restricted to AEGAD. Sequencing a subsample of individuals for these introns indicated that AEGAD-specific alleles were separate clades, thus revealing a phylogeographic brake at the WMED-AEGAD boundary. Sequencing of the mitochondrial COI locus confirmed this phylogeographic break. Dating of the AEGAD mitochondrial haplotypes and nuclear alleles with a Bayesian MCMC method revealed that they shared common ancestors in the Pleistocene. These results can be explained in the framework of Pleistocene sea level drops and patterns of gene flow in MEDAT. An additional observation was a lack of differentiation at COI between the ATL and WMED, in sharp contrast with 4 introns that showed clear genetic subdivision. Neutrality tests did not support the hypothesis of a selective sweep acting on mtDNA to explain the contrasting levels of differentiation between mitochondrial and nuclear markers across the ATL-WMED transition, and we argue that the difference between markers is

best explained by the existence of an endogenous genetic barrier, rather than by a physical barrier to larval migration alone .

Keywords: phylogeography; Mediterranean Sea; selective sweep; endogenous barriers; gene flow

1. Introduction

A majority of marine animals show life histories characterized by a long planktonic larval phase that allows for long-distance dispersal by marine currents. In these cases, a high genetic homogeneity of populations is expected. However, population genetic and phylogeographic studies have shown that genetic breaks resulting in marked genetic subdivision are common. This situation has been explained by historical population fragmentation, for instance as a result of sea level changes, or barriers resulting from present day ocean current patterns (Hellberg, 2009), and is well exemplified by species living in the Mediterranean Sea and the adjacent of NE Atlantic Ocean (MEDAT) (Fig.1). This region has a complex coastline, shows characteristic hydrographic patterns and has experienced dramatic geological and environmental changes during the last few million years, which altogether results in a particularly rich stage for the interplay of evolutionary forces. Phylogeographic and population genetic studies have provided abundant information on the initial phases of evolutionary diversification in MEDAT (reviewed in Patarnello et al., 2007). These studies have shown that phylogeographic breaks or step clines in allelic frequencies occur frequently at the boundaries between the Atlantic Ocean and the western Mediterranean Sea, between the western and eastern basins of the Mediterranean Sea, and between the eastern Mediterranean and the Black Sea (e.g. Quesada et al., 1995; Bahri-Sfar et al., 2000; Bargelloni et al., 2003; Peijnenburg et al., 2004; Luttikhuizen et al., 2008; Schunter et al., 2011). Additional genetic subdivision at sub-regional scales in the Mediterranean Sea has also been found (Pérez-Losada et al., 2007; Debes et al., 2008; Giovanotti et al., 2009; Xabier et al., 2011).

The major genetic discontinuities in MEDAT have been explained as a consequence of historical, oceanographic and climatic factors. First, a decrease of 120 m in the sea level during Pleistocene glaciations (Maggs et al., 2008; Wilson and Veraguth, 2010) would have cut the

connections between the main Mediterranean basins, and between the Mediterranean and the Atlantic, leading to population fragmentation in many marine species. Second, present-day potential barriers to migration exist at the Atlantic / Mediterranean transition, due to the Strait of Gibraltar or to the Almeria-Oran oceanographic front (AOF), and also at the transition between the western and the eastern Mediterranean, due to unidirectional water flow at the Siculo-Tunisian Strait (STS) and mesoscale features in the Ionian Sea (Millot and Taupier-Letage, 2005)(Fig. 1). These barriers could result in restrictions to gene flow among the main marine basins of the region (Bahri-Sfar et al., 2000; Schunter et al., 2011). Recent evidence also supports the view that apparently minor marine barriers within the Mediterranean could have important effects on genetic differentiation (Galarza et al., 2009; Schunter et al., 2011). Finally, a number of studies have shown that the intermittent connection of the Black Sea and the Mediterranean during the Pleistocene has had an important role in generating genetic variability in MEDAT (Magoulas et al. 1996; Ladoukakis et al., 2002; Peijnenburg et al., 2004).

Most of the current knowledge of MEDAT phylogeography come from studies of mitochondrial DNA (mtDNA) sequences. These studies have shown that phylogenetically close species, or species with similar ecological characteristics, often exhibit contrasting patterns of differentiation (Patarnello et al., 2007; García-Merchán et al., 2012). When nuclear markers were included in the same or in follow-up studies, several examples of discordance with mtDNA were found, either in the amounts of genetic differentiation observed or in the location of genetic breaks (Lemaire et al., 2005; Larmuseau et al., 2010; Durand et al., 2013). These results indicate that history and oceanography alone are not enough to account for the phylogeographic patterns observed in MEDAT. An additional role has been suggested for past population fragmentation by promoting the emergence of endogenous barriers to gene flow due to genetic incompatibilities among individuals from formerly isolated populations (Lemaire et al., 2005; Bierne et al., 2011; Sà-

Pinto et al., 2012; Gosset and Bierne, 2013), and for natural selection (Bierne et al., 2011; García-Merchán et al., 2012). Although a full understanding of the relative roles of all these factors require different approaches, including detailed field studies and laboratory crosses, inferences from the patterns of population variability across markers can help to reject some factors and to favor others. For example, demographic factors tend to show stronger effects on mtDNA since the effective population size is 2-4 times smaller for this uniparental inherited marker, so contrasting patterns showing higher differentiation at nuclear markers have to be explained in most situations by selection or endogenous barriers (Toews and Brelsford, 2012).

Bivalve mollusks, with their sessile adult life and typically long planktonic larval phases that facilitate passive dispersal by currents, are a good model to study the interactions of marine circulation with past climate change and other factors affecting phylogeography (e.g.: Hare and Avise 1998; Wares and Cunningham, 2001; Zardi et al., 2007; Tarnowska et al., 2010). In this paper we use the clam *Ruditapes decussatus* (Fig. 1) as a model to examine the relative roles of natural selection and endogenous barriers to gene flow in the Atlantic –Mediterranean transition. *R.decussatus* lives in intertidal and subtidal sandy-muddy bottoms in MEDAT, where it can be found all along the Mediterranean coasts and from the coasts of Senegal to Norway in the Atlantic (Fischer-Piette and Métivier, 1971). It is locally abundant and is frequently harvested for the shellfish market in Europe, Turkey and north African countries. It has separate sexes, reproduction takes place in spring and summer, and the larval phase lasts 10-20 days, depending on temperature (Pérez-Camacho, 1976; Borsa and Millet, 1992).

Earlier studies using allozymes showed only slight differentiation across MEDAT populations in *R. decussatus* (Borsa et al., 1994; Gharbi et al., 2011). However, in a preliminary study using three introns amplified by PCR and scored by RFLP, we found marked genetic differences in allele frequencies between two Atlantic and two western Mediterranean populations. Introns have

several advantages for phylogeographic research, one of which is that they usually contain phylogenetic information that can be studied by sequencing (Palumbi, 1994; Chenuil et al., 2010). We have now scored six introns for RFLP variability in 11 populations spanning most of the species' range, and sequenced a subsample of clams for 2 of them. In addition we have sequenced a fragment of the mitochondrial COI gene. We first show that *R. decussatus* mtDNA exhibit no population differentiation at the Atlantic – W Mediterranean transition, while most nuclear markers do, providing a possible case of natural selection acting on mtDNA or of endogenous barriers to gene flow. By testing for natural selection on mtDNA and comparing haplotype and intron allele frequency patterns and phylogenies, we have studied the relative roles of these factors in shaping the genetic differentiation in this species.

2. Materials and Methods

2.1. Sampling and DNA extraction

Clams from 11 localities from the Atlantic (5 sites), western Mediterranean (2) and eastern Mediterranean (4) were studied (Fig. 1). Samples were collected between 2004 and 2010. Tissues were preserved frozen at -80° C or in 90 % ethanol. Samples of *R. philippinarum* and *Venerupis rhomboides* from NW Spain, and *V. pullastra* and *V. aurea* from NW France, were collected for using as outgroups in phylogenetic analyses.

DNA extraction was carried out with a commercial kit (DNeasy Blood and Tissue Kit, Qiagen) or by boiling a piece of adductor muscle or mantle for 20 min in 10% Chelex 100, 200-400 mesh (Bio-Rad Laboratories).

2.2. Mitochondrial marker

A fragment of 709 bp of the COI gene was amplified with the universal primers HC02198 (5'-taaacttcagggtgaccaaaaaatca-3') and LCO1490 (5'-ggtcaacaaatcataaagatattgg-3') (Folmer et al., 1994). PCR was carried out in 20 µl reactions with 1µl template DNA, 0.2 mM of each dNTP, 0.8 µM of each primer, 1.5 mM of $MgCl_2$, and 0.5 U Taq DNA Polymerase (Invitrogen) in the buffer supplied by the manufacturer. After 4 min at 95°C, the PCR was subjected to 35 cycles of 1 min at 95°C, 30 s at 45°C, and 30 s at 72°C, followed by a final 3 min step at 72°C. PCR products were purified with the UltraClean PCR Clean-Up Kit (MO BIO) and sequenced at the University of Valencia (Spain) sequencing service. PCR products were sequenced in both directions by using the primer COI-Td-54F (5'- aaatcaacagaaaggacgtt -3') in combination with primers COI-Td-54R (5'- cttttctttctgagctggtt -3') or the original reverse primer (LCO1490).

2.3. Nuclear markers

Six nuclear markers were studied, including two that have been described elsewhere (*Srp54* and *Tbp) (*Cordero et al., 2008), and four that were newly developed for this work (*Ech, Fas, Trdmt, Ubc*) (Table S1). The new markers were obtained from a collection of ESTs (Tanguy et al., 2008) following the procedures of Cordero et al. (2008), which included PCR primer development, and cloning and sequencing a number of PCR products. PCR primers and some characteristics of the markers are given in Table S1. PCR amplification was carried out as for COI, with one cycle of 4 min at 94°C, 35 cycles of 30 s at 95°C, 45 s at 58°C, and 1 min at 72°C, and a final 3 min step at 72°C. Polymorphism at TBP was due to length variation and was detected by simply running the PCR products in a 2% agarose gel. Variation at the other markers was studied by restriction fragment

length polymorphism (RFLP) detection, using restriction enzymes *Bam*H I, *Dra* I or *Eco*R I, as described (Cordero et al., 2008). Fragments (Fig. S1) were separated by electrophoresis in 1.5% agarose gels in 1x TBE buffer, stained with SYBR SAFE (Invitrogen), and photographed under UV light. Consistency with predictions from restriction sites observed in the sequences was checked.

PCR products from selected individuals were sequenced in order to study the allelic phylogeny at loci *Ech* and *Tbp*. In the case of *Tbp*, which was a length polymorphism, bands were extracted from gels and reamplified. Then PCR products were purified as described for COI and sequenced using the same primers used for amplification. In the case of *Ech*, PCR products were sequenced using four specific sequencing primers (3F: 5'- caggagaaaatgcccagttt -3'; 3R: 5'- gcaccatttccattgctttg -3'; Seq2Fb: 5'- tcacctgacctggaatataa -3'; Seq3Rb: 5'- ggtcaacaggaaacacttta -3').

The same primers were used to amplify *Ech* and *Tbp* in outgroup species, and the identity of the PCR products was confirmed after sequencing with BLAST (Altschul et al., 1997).

2.4. Mitochondrial sequence data analysis

The sequences obtained were edited with Bioedit (Hall, 1999) and aligned with Clustal W (Thompson et al., 1994) as implemented in Bioedit. Haplotype determination, frequency computation, and gene diversity analyses were performed with the software package DnaSP v5 (Librado and Rozas, 2009). Measures of haplotype diversity (*H*d) and nucleotide diversity based on the pairwise difference between sequences (π) and on the number of segregating sites (Θ_S) (Watterson, 1975; Nei, 1987) were obtained. The average number of nucleotide differences between pairs of sequences *k* (Tajima, 1983) was also calculated.

Phylogenetic relationships among haplotypes were investigated by constructing a statistical parsimony median network (Bandelt et al., 1999) with the program Network (available at

http://www.fluxus-engineering.com). A phylogenetic tree was also obtained by means of a Bayesian methods using Mr Bayes v.3.1.2 (Huelsenbeck and Ronquist, 2001). The three codon positions were partitioned, and the substitution model was determined with jModelTest v.0.1.1 (Posada, 2008). Two independent runs, with four chains each one, were run for $1x10^6$ cycles. The phylogenetic analysis were repeated three times in order to check for the convergence of the chains.

Phylogeographic patterns were tested by measuring the degree of genealogical divergence of geographic groups of haplotypes with the genealogical sorting index (*gsi*) developed by Cummings et al. (2008), which estimates the degree of exclusive ancestry of individuals in groups on a rooted tree. This index varies between 0 (polyphyly) and 1 (monophyly). Statistical significance of *gsi* was tested by a permutation test using 10,000 permutations. All these calculations were carried out with the Genealogical Sorting Package software available from http://www.genealogicalsorting.org.

Several neutrality tests were carried out. The MK test (McDonald and Kreitman, 1991) was performed by comparing *R. decussatus* COI clades to each other, and comparing the whole *R. decussatus* data set with a set of 3 COI sequences of *V. rhomboides*, which is the closest phylogenetic relative of *R. decussatus* (Canapa et al., 2003; Kappner and Bieler ,2006). Statistical significance was tested by means of the Fisher's Exact Test (Sokal and Rohlf, 1999). Two neutrality tests based on the mutation frequency spectrum were also carried out: Tajima's *D* (1989), which tests for relative proportions of intermediate and low frequency haplotypes, and Fay and Wu's *H* (2000), which tests for departures from neutrality that result in different numbers of derived segregating sites in high frequency and intermediate frequency variants. Fu's (1997) F_S, which is based on the haplotype distribution, was also calculated. *Venerupis rhomboides* was used as the outgroup for *H* estimation. Computations were carried out with the software DnaSP v5 (Librado and

Rozas, 2009). Statistical significance was determined by examining the confidence intervals obtained from the empirical distributions calculated for each statistic in coalescent simulations conducted with DnaSP. The R_2 statistic (Ramos-Onsins and Rozas, 2002) was used to test specifically for population expansions. Confidence intervals for R_2 were calculated by parametric bootstrapping with coalescence simulations by using DnaSP. All previous methods (except MK) are affected by demographic changes in the populations. Li (2011) proposed a method based on the maximum frequency of derived mutations (MFDM) which is not affected by demography. We performed the test with the program MFDM 1.0 available at http://www.picb.ac.cn/evolgen/softwares/index.html. The program first performs the recombination test of Hudson and Kaplan (1985), and adjusts the significance level to the number of recombination events detected. Also, the possibility that deviations from neutral expectations in MFDF are due to multiple substitutions is taken into account, by indicating when the significance can be validated by at least two independent informative nucleotide sites.

Genetic differentiation and population subdivision were studied by two different approaches. Exact tests of heterogeneity of haplotype frequencies (Raymond and Rousset, 1995) were performed with the program STRUC implemented in the software GenePop '007 (Rousset, 2008). An analysis of molecular variance (AMOVA) (Excoffier et al., 1992) was also carried out. Distances among haplotypes for AMOVA were computed by means of the TrN model (Tamura and Nei, 1993). Several hierarchical levels were considered: within populations, among populations, among populations within regions, and among regions. Calculations were done with Arlequin v.3.0 (Excoffier et al., 2005).

The fit of the data to a model of isolation by distance was examined (Slatkin, 1993; Rousset, 1997). Geographic distances between pairs of populations were computed along the coastline or the shortest sea tracts. The significance of the correlation between the test statistics F_{ST} /(1-F_{ST}), or

Φ_{ST} /(1-Φ_{ST}), and geographic distances, was determined by the Mantel test (Mantel, 1967), which was performed with Arlequin.

2.5. Nuclear marker data analysis

Unbiased expected heterozygosity (*h*) was calculated for each locus and population following Nei (1978). Deviations from Hardy-Weinberg equilibrium were estimated by the F_{IS} statistic according to Weir and Cockerham (1984). Calculations were performed with software packages GenePop '007 (Rousset, 2008) and Arlequin v.3.0 (Excoffier et al., 2005). The frequency of null alleles was estimated when required under the Random Mutation Model (Kalinowski and Taper, 2006; Chapuis and Estoup, 2007) using the software INEst (Chybicki and Burczyc, 2009). Null alleles are new variants of a specific locus undetectable through PCR due to mutational changes located at the priming sites.

The geographic genetic structure was investigated by several approaches. Isolation by distance was tested as described for the mitochondrial data. A Bayesian clustering analysis was carried out using the software STRUCTURE (Pritchard et al., 2000). An admixture model with correlated allele frequencies among populations (Falush et al., 2003) was used, with 10^5 runs as burn-in period and 150,000 replicates after burn-in. Structures for *K*=1 to *K*=13 were modeled, with 10 replicates per simulation. The model with the best fit was determined by looking for the highest posterior probability of K, P(K|X), supported by the results based on the estimates of P(X|K) (Pritchard et al., 2000).

Population subdivision was further investigated by means of a hierarchical *F*-statistics analysis (Wright, 1965) performed with the AMOVA routine of ARLEQUIN v3.0 (Excoffier et al.,

1992, 2005) with the locus by locus setting. The significance of the *F*-statistics was tested using a non-parametric permutation approach described in Excoffier et al. (1992).

Recent migration rates between pairs of populations were estimated from multilocus genotypes by the Bayesian method of Wilson and Rannala (2003), which allows genotype frequencies to deviate from Hardy-Weinberg equilibrium proportions and computation of different rates in opposite directions. The program BayesAss v. 1.3 (Wilson and Rannala, 2003) was used for calculations. Details can be found in the Supporting material.

2.6. Nuclear gene sequence analysis

TBP alleles were amplified from heterozygotes and bands of interest were reamplified and sequenced with the PCR primers. These sequences were directly used for phylogenetic analysis after visual inspection of chromatograms and posterior alignment. The majority of the sequenced PCR products of *Ech* were obtained from homozygotes. Heterozygotes exhibited overlapping, long stretches of double peaks indicative of the existence of two alleles and several insertion and deletions (indels). To extract the two alleles from the chromatograms, sequences were first edited with the program Sequencher 4.10.1 (Gene Codes Corporation), and all double chromatogram peaks were checked and renamed with the IUPAC-IUB ambiguity standard code. Then the sequences with ambiguities were treated with the software OLFinder v.1.00 (Dixon, 2009), which extracts and aligns the two alleles from diploid DNA containing heterozygous indels. Reference sequences, derived from homozygotes or obtained by cloning, were used in a few ambiguous cases, following the recommendations of Dixon (2009).

Intron sequences were aligned with Clustal W (Thompson et al., 1994), as implemented in Bioedit, with different values of gap open and extension penalties, or with Dialign 2.2.1

(Subramanian et al., 2008), which compares entire segments of sequences without using gap penalty. The best-fit model of nucleotide substitution for each collection of sequences was determined under the Akaike Information Criterion with the program jModeltest v.0.1.1 (Posada, 2008). Bayesian phylogenetic analyses were conducted using Mr Bayes v.3.1.2 (Huelsenbeck and Ronquist, 2001). Each run was performed starting with a randomly chosen tree and ran for $1x10^6$ generations with sampling conducted every 100^{th} generations. Runs were repeated three times to check for convergence of the chains and reliability of the resulting trees.

2.7. Molecular dating and historical demography

Dating the main mtDNA and intron clades was carried out by determining the time of their most recent common ancestor (*tmrca*) with Bayesian MCMC methods using BEAST v 1.7 software (Drummond et al., 2012). The HKY model (Hasegawa et al., 1985) with gamma distribution was used in all cases, as determined with jModeltest v.0.1.1 (Posada, 2008). In addition for mtDNA partitions into codon positions were applied and substitution rate unlinked across partitions. Phylogenies obtained with MrBayes as described above were checked for repeatability in BEAST and imposed as priors. We examined the performance of strict clock and uncorrelated log-normal (UCLN) clock priors for rate changes, and a coalescent prior on the phylogeny. Exponential and constant models of population growth for the coalescent priors were tried. Estimates of population sizes and exponential growth rates (when applicable) were retrieved. Given the relatively low divergence among clades and the potential role of recent glaciations in the origin of the clades (see Discussion), a Pleistocene (0-2.5 Ma) uniform prior on the time of the most recent common ancestor (*Tmrca*) of all Aegean/Adriatic-specific clades was used. A neutral mutation rate prior of 1% per Ma was used for COI, based on the only fossil based substitution rate study in Bivalves (Marko, 2002). In the case

of introns a mutation rate prior of 3% was used based on intron divergence in *Drosophila* (Singh et al., 2009). Historical demography was examined by Bayesian Skyline plots (Drummond et al., 2005) with BEAST on sequences of the mitochondrial locus. The analysis was based on pooled samples from different regions. The *tmrca* of the A clade obtained in the previous analysis was used as a prior to give a realistic temporal scale, and the number of time sections studied was m=4.

In all calculations carried out with BEAST, runs of 10^7 iterations were performed with a sampling interval of 10^4. Trace diagrams were examined for convergence and a burn in of 10% was applied. Several runs (3-7) with different starting rates were carried out for each marker/ model combination. Results from different runs were checked for congruency and were combined with TRACER in the BEAST suite to obtain parameter estimates based on larger (>200) effective sample sizes (ESS).

3. Results

3.1. Mitochondrial DNA

We obtained the sequence of a 432 bp fragment of the COI gene from 327 individuals sampled from 11 localities. There were 36 polymorphic sites (Table S2), which defined a set of 29 haplotypes (GenBank accessions JX051518 - JX051549). All polymorphisms were transitions. Five were located at first codon positions, 30 at third codon positions and one at a second codon position, which resulted in a nonsynonymous change (Leu – Ser) in an individual from VEN with haplotype A21.

The haplotype frequencies are given in Table S3. The phylogenetic relationships among haplotypes are depicted in the network shown in Fig. 2. Two clades were found (A and B) which

differed by 8 fixed substitutions. Net divergence between clades was 2.3 ± 0.7 %. The haplotypes within each clade were separated by 1-5 substitutions. The two clades exhibited very different geographic distribution. Clade A was present in all the populations, while clade B was present in the Aegean Sea only (HAL and IZM samples). Clade A showed a star-like phylogeny, characterized by a most common haplotype present in all the populations (A1), and many low frequency haplotypes differing in only 1 -3 substitutions from A1 (1 substitution in the majority of the cases). Clade B had two frequent haplotypes (B1 and B2) that showed very different frequencies in each population, and a few rare haplotypes. Bayesian phylogenetic analysis confirmed the existence of clades A and B (Fig. S2). After labeling all sequences as "Aegean" or "non-Aegean" the *gsi* index was non significant for the Aegean group (*gsi* = 0, *P* = 1), but the non-Aegean group gave a clear phylogeographic signal (*gsi* = 0.667, *P* < 0.001).

Estimates of polymorphism and genetic diversity are given in Table 1. Values of π varied between 0.0001 and 0.0015, with no specific geographic pattern. The values of haplotype diversity (H_d) were highest in the two Aegean Sea populations (HAL and IZM), where the clades A and B coexist. No haplotype diversity was detected in the MLF sample, suggesting a very small population size in that locality.

Results of three neutrality tests are also shown in Table 1. Some tests could not be performed in some populations (MLF, IZM, LUL) due to lack of variability or of informative polymorphisms. An abundance of significant , negative values for H, and to a lesser extent for Fs, indicate a departure from neutrality of COI polymorphisms, which could reflect a recent change in population size or the action of natural selection. The MFDM test was performed using *V. rhomboides*, *V. pullastra* and *V. aurea* as outgroups. Two recombinant events were detected by the Hudson and Kaplan (1985) test. This observation is in line with records of recombination at mtDNA in several organisms, including bivalves such as the mussel *Mytilus* (Tsaousis et al. 2005) and a

congeneric species of *Ruditapes decussatus*, known as *R. philippinarum* (Passamonti et al. 2003). Since the test for recombination indicated a minimum number of two recombination events, the p-value necessary to reject neutrality with the MFDF test was 0.0167. The test was significant (P = 0.006), but the possibility that this was due to multiple hits can not be rejected because the results could not be validated by two informative polymorphic sites. The McDonald-Kreitman test (Table 2) indicated an excess of replacement changes at non-synonymous sites during the divergence of the COI sequence of *R. decussatus* from *V. rhomboides,* but it was non-significant (P = 0.054). The divergence of *R.decussatus* clades A and B fit the neutral expectations (P = 1).

Exact tests of population differentiation (Raymond and Rousset, 1995) showed significant differences in haplotype frequencies among populations in the total area (P < 0.001) as well as among Atlantic (P = 0.019) and Mediterranean populations (P < 0.001). Haplotype frequencies showed erratic changes among populations, but the only clear pattern was the appearance of clade B in the Aegean Sea, and no obvious changes across the Atlantic-western Mediterranean were observed (Fig. 3). The AMOVA indicated high levels of differentiation also at the nucleotide level in the total area (Φ_{ST} = 0.669, P < 0.001) as well as in the Mediterranean (Φ_{ST} = 0.663, P < 0.001) (Table 3, models 1 and 5). The Atlantic populations presented a much lower level of differentiation (Φ_{ST} = 0.028, P < 0.01). A model with subdivision in which the Aegean Sea was separated from the rest of populations (Table 3, model 7) explained almost all the variation in terms of the regional subdivision (Φ_{CT} = 0.839, P < 0.05) and showed small differentiation within regions (Φ_{SC} = 0.077, P < 0.01). Alternative subdivision models (Table 3, models 6 and 8) rendered a much higher heterogeneity within regions. The haploid analogous of the F-statistic was 0.289 (P < 0.001) for the total area, 0.027 (P < 0.05) for the Atlantic populations and 0.314 (P < 0.001) for the Mediterranean. Excluding the Aegean, the F_{ST} in the Mediterranean was 0.027 (P<0.05), and 0.031 (P < 0.05) in the total area.

The test for isolation by distance was significant when applied to both F_{ST} and Φ_{ST} (Fig. S3). But the determination coefficient R^2 was only 22% in the best case (F_{ST}), indicating that distances among sampling localities explained only a minor portion of the geographic variation in haplotype frequencies.

3.2. Nuclear markers

Two to five alleles were found at each of the six scored nuclear markers. The fragment sizes of the alleles found are given in figure S1. The restriction sites were in all cases located in the introns. The gains and losses of restriction sites were due to substitutions at a single base pair.

Twenty to 56 individuals were scored in each locality (average 43). Allelic frequencies, heterozygosity and deviations from Hardy-Weinberg equilibrium are shown in Table S4. All loci were polymorphic in all populations. Nine out of 65 tests for Hardy-Weinberg equilibrium were significant, in all cases due to a deficit of heterozygotes (positive values of F_{IS}), a typical observation in bivalves (Reece et al., 2004, and references therein). Significant heterozygote deficiencies were concentrated at locus *Ech* (6 cases). Individuals that gave no PCR product, a clear indication of the presence of null alleles, were observed at *Ech* in several populations.

Most intron alleles were present in the majority of populations, but four showed geographic distributions restricted to the Mediterranean (*Ech-C)* or to the Adriatic and Aegean samples (*Ech-D* and *Tbp-C)*. There was a single private allele (*Tbp*-D in SFX). Step clines could be seen at several loci, notably *Ech* and *Tbp*, at the transition between the Atlantic and western Mediterranean (between FOR and MME), but only shallow transitions were observed at the boundary between the western and the eastern Mediterranean (between SFX and VEN) (Fig. 3).

The differences in allelic frequencies among populations were statistically significant for all loci when tested with exact tests. *F*-statistics showed moderate differentiation in the total area (F_{ST} = 0.134, $P < 0.001$) when computed across all loci. The Mantel test for isolation by distance was significant (P< 0.004), but the determination coefficient indicated that geographic distance explained barely one third (27 %) of the genetic variation observed across the area of study (Fig. S3). When the Atlantic or the Mediterranean sets of populations were considered separately, the results were no longer significant (not shown).

The Bayesian clustering analysis assigned individual genotypes to three clusters (Fig. 4A). Three groups of populations could be distinguished, according to the most abundant cluster: Atlantic (ATL), western Mediterranean populations plus SFX (WMED), and populations from the Aegean and Adriatic seas (AEGAD) (Fig. 4B). Frequencies of clusters 1 and 3 in VEN suggest that this population has an intermediate character between WMED and AEGAD.

Results of the hierarchical analysis of population genetic differentiation with F-statistics are presented in Table 4. Several models of subdivision were tested. The best models, defined as those in which interregional differentiation (F_{CT}) was maximized and intraregional differentiation (F_{SC}) was minimized, were model 1 (ATL, WMED, AEGAD), which was the one favored by the previous Bayesian clustering analysis, and model 2 (ATL, WMED + VEN, Aegean), in which the Adriatic population was included in the WMED group. Differences between these two models were minimal, and we preferred model 1 for subsequent analyses because it reflects the results of the Bayesian structure analysis and the presence of alleles restricted to AEGAD at loci *Ech* and *Tbp*. Differences among regions accounted for 12.8 % of the total variability under this model. A model of two regions (Atlantic versus Mediterranean; model 4) and the model that best explained the distribution of variability at the mitochondrial COI (Aegean versus rest of populations; model 3) did not fully account for the genetic differentiation observed, since they gave higher F_{SC}. Comparisons

between adjacent regions (models 5-7) showed that the differentiation observed between the Atlantic and the western Mediterranean (F_{CT} = 0.128, P < 0.05) was nearly four times as high as that observed between western and eastern Mediterranean basins.

When the nuclear markers were considered individually, the 3-basin geographic structure described (model 1, Table 4) was observed in four out of the six scored loci: *Ech*, *Fas*, *Tbp* and *Ubc* (Fig. 5). Two other loci (*Srp54* and *Trdmt*) had non-significant F_{CT}. Locus-specific values of F_{ST} varied widely. Three loci (*Ech*, *TBP* and *Trdmt*) showed F_{ST} values above 0.20, which is quite high for a species characterized by a long planktonic larval phase. F_{CT} values , indicative of regional subdivision, were also contrasting across these 3 loci, with *Ech* and *Tbp* exhibiting F_{CT} above 0.2 and *Trdmt* showing no regional subdivision (F_{CT} = 0.024, P>0,05)

Estimation of migration rates between regions provided evidence for asymmetric gene flow, which was much higher towards the east (Table 5). The highest migration rate observed was from WMED to AEGAD (0.042), followed by migration rate from ATL to WMED (0.0148). Migration rates in the opposite direction were 4 and 10 times smaller, respectively. Migration rates between pairs of populations confirmed these results (Table S5).

Tests of two-locus gametic disequilibrium in the three AEGAD populations for all pairs of nuclear markers rendered only eight significant cases (P < 0.05) out of 165. Chi-square tests of association between mitochondrial clades and nuclear genotypes at individual loci in HAL and IZM were non-significant.

3.3. Phylogenetic analysis of Ech *and* Tbp *alleles*

Geographic distribution of alleles at *Ech* and *Tbp* suggested that these loci could have important phylogeographic information content. We therefore sequenced a subsample of clams (GenBank accessions JX051550- JX051604) to study the allelic phylogenies.

Thirty three alleles from 27 individuals were sequenced for locus *Ech* (Table S6). Sequence length varied between 951 bp and 957 bp. There were 64 segregating sites (61 were phylogenetically informative). Indels of 2-6 bp were abundant. Bayesian phylogenetic analysis of *Ech* alleles is shown in Fig. 6, along with biogeographic data for each allele sequenced. Allele *Ech*-D, which was found only in AEGAD samples, and *Ech* C, which was present in all the Mediterranean but was much more abundant in AEGAD (Table S4), formed separate monophyletic clades derived from one of the widespread alleles. Phylogenetic analysis also discovered cryptic variants at RFLP alleles *Ech*-A and *Ech*-B. One of these new clades (ECH-A2) was restricted to Mediterranean samples, while the other three (ECH-A1, ECH-B1 and ECH-B2) were present in samples from the whole region (Fig.6).

Twenty nine alleles of the *Tbp* locus were sequenced (Table S6). The lengths of the four allelic variants found varied between 475 and 602 base pairs. There were 20 polymorphic sites, all phylogenetically informative. Alleles were also polymorphic for indels of 104, 22, 10 and 1-3 bp. Sequences from outgroups could not be aligned properly, so we constructed an unrooted tree (Fig. 7). Alelle *Tbp*-D was almost identical to sequences of clade TBP-A1, and differed only by a 104 bp indel. As in *Ech*, the *Tbp* allele that was restricted to AEGAD (*Tbp-C*) formed a monophyletic clade, and new cryptic variants at RFLP alleles *Tbp-A* and *Tbp –B* were discovered. Clade TBP-A2 was a sister group to allele *Tbp*-C and equally seemed to be restricted to AEGAD. With the cautions that the limited number of individuals sequenced requires, the data indicate that clade TBP-A1 was abundant in ATL and WMED, but was extremely rare in AEGAD (Fig. 7). Assuming the prediction of

the coalescent model that the most variable and widespread alleles are the oldest, this would be *Tbp –B*.

3.4. Molecular dating

Dating of specific clades in the trees of the 3 sequenced markers was conducted with BEAST to infer the times of phylogeographic events in the eastern Mediterranean. Inspection on trace diagrams, likelihoods and ESS for different combinations of population demography and clock priors indicated a better performance of the *tmrca* estimation procedure under a constant population size model for all three markers, with a strict clock for *COI* and *Tbp*, and an UCLN clock for *Ech* (Table 6).

In all three markers the *tmrca* of the clades fell in the Pleistocene (0.01 – 2.5 Ma; Table 6), although confidence intervals were very wide and for the mitochondrial locus and intron *Ech* they expanded into the Pliocene. The *tmrca* estimates indicate that mitochondrial clade B and *Ech* clade C shared common ancestors with their sister clades earlier than 1 Ma. Clade D of *Ech* and clade C of *TBP* were the youngest (0.2 – 0.3 Ma). Clade A2 at *Ech* shared a common ancestor with its sister clade A1 at 0.5 Ma. These dates suggest that three different cladogenetic events could have been involved in the origin of AEGAD variants.

3.5. Historical demography

Bayesian skyline plots of *COI* clades A and B were obtained (Fig. S4). The analysis for clade B suggests a slight recent increase in population size, but confidence intervals at the onset and at the end of the curve are too wide and overlapping as to consider it significant. This plot was based in ESS values around 1000 obtained by merging 3 runs of 10 million chain length, and in all cases the

trace diagrams of parameters exhibited the typical “hairy caterpillar” shape characteristic of good quality data (Drummond et al., 2005). A reliable plot could not be obtained for Clade A, due to bad quality trace diagrams and very low ESS, even when long runs were concatenated.

4. Discussion

The results of this study using one mitochondrial and six nuclear DNA markers in 11 populations indicate that *R. decussatus* has a subdivided genetic structure, as has been observed in many other marine species living in MEDAT. However, the details are complex, and there are contrasting patterns of variation between markers.

Pooled intron RFLP data showed subdivision into three regions- the Atlantic (ATL), the western Mediterranean plus Tunisia (WMED), and the Aegean and Adriatic seas (AEGAD)- which accounted for ~13% of the genetic variability observed. However, there was a great variability across introns in the amount of regional subdivision, with some of them showing no significant subdivision (e.g.: *Trdmt*) and others showing relatively high levels (*Ech* and *TBP*). On the other hand, mitochondrial DNA showed no differentiation between ATL and WMED, and subdivision at the eastern Mediterranean was limited to the Aegean Sea. Contrasting levels of differentiation across the geographic area studied were also observed, with differentiation between ATL and WMED being four times as large as differentiation between WMED and AEGAD at introns. Moreover, mtDNA and two introns provided evidence of phylogeographic structure at the WMED-AEGAD boundary, but not at the ATL-WMED transition, where simple step clines in allelic frequencies of introns were found.

The fact that the majority of the introns showed subdivision at roughly coincident areas, and that the patterns of differentiation observed were very similar to those reported in many other

marine species of the region, point to a mechanism of genetic differentiation based on forces that act on the whole genome, such as genetic drift and gene flow restrictions. In the context of MEDAT phylogeography studies, theses forces have been typically explained as the result of population fragmentation during sea-level reductions associated with Pleistocene glaciations, and present-day marine circulation patterns, which are characterized in general by a predominant west to east marine water flow (Fig. 1) (Patarnello et al., 2007). However, the contrasting patterns across marker types and regions indicate that additional factors are acting on the molecular variability observed.

4.1. *The Atlantic-Mediterranean transition and the role of endogenous barriers*

There was a clear genetic change in clam populations across the ATL-WMED boundary, as is common in other species (Patarnello et al., 2007), but it was apparent in the nuclear markers and not at the mtDNA. Sequencing of the two most regionally differentiated nuclear markers showed that the change across the ATL-WMED transition was not due to a phylogeographic break at this point, but to simple allele frequency changes. This contrasts with the WMED-AEGAD transition, in which evidence for a phylogenetic break was found. The contrast between the two transition areas could be due to differences in the demographic characteristics of the three regions at the time of the Pleistocene subdivisions. Specifically, a large effective size of ATL and WMED populations during fragmentation at glacial times would have limited the opportunity for the establishment of new alleles in those regions. On the other hand, the amount of differentiation in allele frequencies shown by nuclear markers at this transition (F_{CT} = 0.128) was three times as large as that observed at the WMED-AEGAD transition (F_{CT} = 0.034). The lower estimates of migration rates between ATL and WMED than between WMED and AEGAD suggest that the difference could be due to a restriction to gene flow in the ATL-WMED boundary. One of the sources of gene flow restriction in

the ATL-WMED transition could be the Almeria-Oran front (AOF), although detailed sampling in its surroundings is necessary to confirm it.

Lack of genetic differentiation at mtDNA across the ATL-WMED boundary has been reported in other species (Patarnello et al., 2007, and references therein). In the few studies in which nuclear genetic markers were analysed in the same samples, the two marker types usually showed a concordant pattern (Patarnello et al., 2007). To our knowledge, our report of a marked differentiation at nuclear loci and no differentiation at mtDNA is the first one to show this striking difference at the ATL-WMED transition. It seems improbable that our results were due to a limitation of the COI gene to detect different haplotypes, since our fragment's size is similar to that used in studies of other species in which differentiation among populations was observed, levels of variability are similar to those reported in many other species, and it was effective in detecting divergent haplotypes in the Aegean Sea.

The observed discordance between introns and mtDNA could appear just by chance (Karl et al. 2012). However, since the effective size of mitochondrial genome is four times smaller than that of the nuclear genome, this seems unlikely and alternative explanations should be explored. Positive natural selection has been often invoked to explain discordant geographic patterns of variability across markers (Toews and Brelsford, 2012). A selective sweep of clade A could have eliminated other clades from the ATL and WMED regions, resulting in their genetic homogenization. However, we have found no clear evidence for a selective sweep at COI in *R. decussatus*, as the MK tests gave no significant results, and, although some neutrality tests based on intraspecific polymorphism were significant (H, MFDM), results could have been influenced by purifying selection, recent bottlenecks, or multiple hits at some nucleotide positions.

Another potential explanation for the discordance between mtDNA and introns in the ATL-WMED transitions is that the introns are affected by an endogenous barrier to gene flow. Computer

simulations showed that clines for neutral markers tend to coincide spatially when the markers are linked to incompatible loci or to genes under selection, and the points of coincidence of the clines tend to be located near gene flow barriers or environmental limits (Bierne et al., 2011). In the present case, endogenous barriers would be caused by genomic incompatibilities between Atlantic and west Mediterranean clam populations, which may have arisen during the periods of population fragmentation and isolation in the Pleistocene. The gene flow barrier/environmental limit could be the AOF or the Strait of Gibraltar (Schunter et al. 2011). The observation of significant F_{CT} at four of the six nuclear markers scored in this study suggests that incompatibilities could be abundant in the genome of *R. decussatus*.

In the context of endogenous barriers, the contrasting patterns of differentiation between mtDNA and introns across the ATL-WMED boundary could be the result of the lack of incompatibilities between the mitochondrial genes and the nuclear genetic backgrounds of ATL and WMED. This would result in no impediments for effective gene flow at the mitochondrial genome, which would have eroded the genetic differentiation caused by past fragmentation. The lack of differentiation observed at some introns (*Srp54*, *Trdmt*), would have a similar explanation. Higher intraregional F_{ST} values for mtDNA than for introns, as expected for a neutral migration-drift system, support this view.

4.2. *Genetic differentiation of the Aegean and Adriatic seas: roles of glaciations and gene flow*

The most important source of genetic subdivision in R. decussatus is the phylogeographic break detected at the limit of WMED and AEGAD. This break is revealed by the existence of specific mitochondrial and intronic clades in AEGAD which are not present in the rest of the area (mitochondrial clade B and intronic clades ECH-D, TBP-A2 and TBP-C). Clade ECH-C, which is found

all over the Mediterranean but is more frequent in the AEGAD, should be considered also in this category. The lack of significant cytonuclear and gametic disequilibria among the scored markers in that area excludes the existence of a sibling species restricted to the AEGAD.

On the other hand, the pattern of variation observed in AEGAD is similar to that found in a number of species, in which a mitochondrial clade is either specific to the Aegean, or more abundant in that area (Magoulas et al., 1996; Borsa et al., 1997; Ladoukakis et al., 2002; Nikula and Vainola, 2003; Peijnenburg et al., 2004; Papadopoulos et al., 2005; Luttikhuizen et al., 2008; Tarnowska et al., 2010). The persistence of a separate gene pool in the Black Sea during the Pleistocene has been postulated to account for such observations. *Ruditapes decussatus* is not present in the Black sea (Demir, 2003) and to our knowledge there is no evidence of its presence there in the past, so this explanation is not applicable.

The most reasonable scenario for the emergence of the genetic break is one characterized by land barriers that restricted the water flow between the Aegean and the Mediterranean seas due to the Pleistocene sea level drops, as described by Perissoratis and Conispoliatis (2003) for the Last Glacial Maximum (LGM). The clade age estimates suggest that at least two events of population isolation have taken place, one older than 1.0 Ma, responsible for the separation of mitochondrial clades A and B, and allele *Ech-C*, and a more recent one (0.2-0.5 Ma) which would have resulted in clades ECH-D and TBP-C. However, the timing of these events should be taken with caution given the potential high variability of mutation rates across introns (Singh et al., 2009;) and the possibility of time dependence of the substitution rates (Ho et al., 2007).

The presence of AEGAD-specific clades in the Adriatic Sea (VEN) cannot be explained by a Pleistocene refugium. The Adriatic was substantially reduced during the LGM, when it was no more than a wide bay at the northern coasts of the Ionian Sea (Lambeck and Purcell, 2005). Therefore, restocking of this region after the LGM would have proceeded mainly with clams from nearby

Ionian populations, which currently show a WMED genetic composition. AEGAD-specific clades could have arrived in the Adriatic through migration from the Aegean, favored by the water flow around the Peloponnese coasts towards the Adriatic (Hamad et al., 2006) (Fig. 1). The mixture of Ionian and Aegean clam stocks in the Adriatic Sea would explain the intermediate genetic composition of the VEN population reflected in the Bayesian structure analysis and in AMOVA. However, other explanations cannot be discarded without a more detailed study of clam populations in the region.

5. Conclusions

The subdivided genetic structure of *R. decussatus* is characterized by a marked differentiation between the Atlantic and the Mediterranean populations, and the existence of a phylogeographic transition between the western Mediterranean and the Adriatic and Aegean seas. Many aspects of this structure can be explained in the general framework of population subdivision produced by sea level changes during the Pleistocene glaciations and present day patterns of gene flow. However, contrasting patterns of geographic differentiation across markers, especially the absence of differentiation across the Atlantic–Western Mediterranean boundary in the mtDNA haplotype frequencies, in spite of clear allelic frequency differentiation at two intron markers, are difficult to explain as a result of past population fragmentation and present gene flow interactions alone. Neutrality tests show no clear evidence for a selective sweep at mtDNA. The existence of endogenous barriers to gene flow, due to the emergence of genetic incompatibilities at multiple loci linked to neutral genetic markers and generated during the Pleistocene phases of population isolation, provide a framework in which the genetic structure of *R. decussatus* can be more easily understood.

Acknowledgements

We are indebted to the following people that supplied clam samples: Paolo Breber (CNR), Leonor Cancela (University of Algarve), Ramón Carles (CADEMAR), Amel Chaffai (University of Sfax), Marina Delgado (CAQ-IRTA), Tolga Dincer (Ege University), Benjamín García (IMIDA), Didier Jollivet (University of Brest), Ricardo Leite (University of Algarve), Dorotea Martínez (CIMA), José Manuel Molares (CIMA), Costas and Alexandros Triantafyllidis (University of Thessaloniki). We thank Dr. Rita Castilho and two anonymous reviewers for their valuable comments during the review process. This work has been funded by grants AGL2003-04143 from MCYT (Spain) and AGL2006-08944 from MEC (Spain) to J.B. Peña, and AGL2007-60049 and AGL2010-16743 from MICINN (Spain) to C. Saavedra. Partial funding was obtained from the Marine Genomics Europe Network of Excellence (GOCE 505403).

Appendix A. Supplementary data

Supplementary data can be found in the online version of this article at http://www.elsevier.com/mpe

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nuc. Acids Res. 25, 3389-3402.

Bahri-Sfar, L., Lemaire, C., Hassine, O.K.B., Bonhomme, F., 2000. Fragmentation of sea bass populations in the western and eastern Mediterranean as revealed by microsatellite polymorphism. Proc. R. Soc. Lond. B 267, 929-935.

Bandelt, H.-J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37-48.

Bargelloni, L., Alarcon, J.A., Alvarez, M.C., Penzo, E., Magoulas, A., Reis, C., Patarnello, T., 2003. Discord in the family Sparidae (Teleostei) : divergent phylogeographical patterns across the Atlantic-Mediterranean divide. J. Evol. Biol. 16, 1149-1158.

Bierne, N., Welch, J., Loire, E., Bonhomme, F., David, P., 2011. The coupling hypothesis: why genome scans may fail to map local adaptation genes. Mol. Ecol. 20**,** 2044-2072.

Borsa, P., Blanquer, A., Berrebi, P., 1997. Genetic structure of the flounders *Platychtys flesus* and *P. stellatus* at different geographic scales. Mar. Biol. 129, 233-246.

Borsa, P., Millet, B., 1992. Recruitment of the clam *Ruditapes decussatus* in the lagoon of Thau, Mediterranean. Est., Coastal Shelf Sci. 35, 289-300.

Borsa, P., Jarne, P., Belkhir, K., Bonhomme, F., 1994. Genetic structure of the palourde *Ruditapes decussatus* L. in the Mediterranean. In: Beaumont, A.R. (Ed.), Genetics and evolution of aquatic organisms pp. 103-113, Chapman and H a ll , London.

Burton, R.S., Barreto, F.S., 2012. A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities? Mol. Ecol. 21, 4942-4957.

Canapa, A., Schiaparelli, S., Marota, I., Barucca, M., 2003. Molecular data from the 16S rRNA gene for the phylogeny of Veneridae (Mollusca: Bivalvia). Mar. Biol. 142, 1125–1130.

Chapuis, M.-P., Estoup, A., 2007. Microsatellite null alleles and estimation of population differentiation. Mol. Biol. Evol. 24, 621-629.

Chenuil, A., Hoareau, T.B., Egea, E., Penant, G., Rocher, C., Aurelle, D., Mokhtar-Jamai, K., Bishop, J.D.D., Boissin, E., Diaz, A., Krakau, M., Luttikhuizen, P.C., Patti, F.P., Blavet, N., Mousset, S., 2010. An efficient method to find potentially universal population genetic markers, applied to metazoans. BMC Evol. Biol. 10, 276.

Chybycki, J., Burczyc, J., 2009. Simultaneous Estimation of Null Alleles and Inbreeding Coefficients. J. Hered. 100, 106-113.

Cordero, D., Peña, J.B., Saavedra, C., 2008. Polymorphisms at Three Introns in the Manila Clam (*Ruditapes philippinarum*) and the Grooved Carpet-Shell Clam (*R. decussatus*). J. Shellfish Res. 27, 301-306.

Cummings, M.P., Neel, M.C., Shaw, K.L., 2008. A genealogical approach to quantifying lineage divergence. Evolution 62, 2411-2422.

Debes, P.V., Zachos, F.E., Hanel, R., 2008. Mitochondrial phylogeography of the European sprat (*Sprattus sprattus* L., Clupeidae) reveals isolated climatically vulnerable populations in the Mediterranean Sea and range expansion in the Northeast Atlantic. Mol. Ecol. 17, 3873-3888.

Demir, M., 2003. Shells of Mollusca Collected from the Seas of Turkey. Turkish J. Zool. 27, 101-140.

Dixon, C.J., 2009. OLFinder—a program which disentangles DNA sequences containing heterozygous indels. Mol. Ecol. Resources 10, 335–340.

Drummond, A.L., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol.Biol.Evol. 22, 1185-1192.

Drummond, A. J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969-1973.

Durand, J.D., Blel, H., Shen, K.N., Koutrakis, E.T., Guinand, B., 2013. Population genetic structure of *Mugil cephalus* in the Mediterranean and Black Seas: a single mitochondrial clade and many nuclear barriers. Mar. Ecol. Progr. Ser. 474, 243-261.

Excoffier, L., Smouse, P.E., Quattro, J.M. 1992. Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data . Genetics 131, 479-491.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol. Bioinformatics Online 1, 47-50.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164, 1567-1587.

Fay, J.C., Wu, C.-I., 2000. Hitchhiking under positive Darwinian selection. Genetics 155, 1405-1413.

Fischer-Piette, E., Métivier, B., 1971. Revision des Tapetinae (Molusques Bivalves). Mém. Mus. Nat. d´Histoire Naturelle, Serie A, Zoologie, 71, 1-106.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3**,** 294-297.

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

Galarza, J.A., Carreras-Carbonell, J., Macpherson, E., Pascual, M., Roques, S., Turner, G.F., Rico, C., 2009. The influence of oceanographic fronts and early-life-history traits on connectivity among littoral fish species. Proc. Nat. Acad. Sci. USA 106, 1473-1478.

García-Merchán, V.H., Robainas-Barcia, A., Abelló, P., Macpherson, E., Palero, F., García-Rodríguez, M., Gil de Sola, L., Pascual, M., 2012. Phylogeographic patterns of decapod crustaceans at the Atlantic-Mediterranean transition.Mol. Phylogenet. Evol. 62, 664-672.

Gharbi, A., Zitari-Chatti, R., van Wormhoudt, A., Dhraief, M.N., Denis, F., Said, K., Chatti, N., 2011. Allozyme variation and population genetic structure in the carpet shell clam *Ruditapes decussatus* across the Siculo-Tunisian Strait. Biochem. Genet. 49, 788-805.

Giovanotti, M., La Mesa, M., Caputo, V., 2009. Life style and genetic variation in teleosts: the case of pelagic (*Aphia minuta*) and benthic (*Gobius niger*) gobies (Perciformes: Gobiidae). Mar. Biol. 156, 239-252.

Gosset, C.C., Bierne, N., 2013. Differential introgression from a sister species explains high F_{ST} outlier loci within a mussel species. J. Evol. Biol. 26, 14-26.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Serial, **41,** 95-98.

Hamad, N., Millot, C., Taupier-letage, I., 2006. The surface circulation in the eastern basin of the Mediterranean Sea. Sci. Mar. 70, 457-503.

Hare, M.P., Avise, J.C., 1998. Population structure in the American oyster as inferred by nuclear gene genealogies. Mol. Biol. Evol. 15, 119-128.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160–174

Hellberg, M.E., 2009. Gene flow and isolation among populations of marine animals. Ann. Rev. Ecol. Evol. Syst. 40, 291-310.

Ho, S.Y.W, Shapiro, B., Phillips, M.J., Cooper, A., Drummond, A.J., 2007. Evidence for time dependency of molecular rate estimates. Syst. Biol. 56, 515-522.

Hudson, R.R., Kaplan, N.L., 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111, 147-164.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17**,** 754-755.

Kalinowski, S.T., Taper, M.L., 2006. Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. *Conserv. Genet.* 7, 991-995.

Kappner, I., Bieler, R., 2006. Phylogeny of Venus clams (Bivalvia: Venerinae) as inferred from nuclear and mitochondrial gene sequences. Mol. Phylogenet. Evol. 40, 317–331.

Karl, S.A., Toonen, R.J., Grant, W.S., Bowen, B.W., 2012. Common misconceptions in molecular ecology: echoes of the modern synthesis. Mol.Ecol. 17, 4171-4189.

Ladoukakis, E., Saavedra, C., Magoulas, A., Zouros, E, 2002. Mitochondrial DNA variation in a species with two mitochondrial genomes: the case of *Mytilus galloprovincialis* from the Atlantic, the Mediterranean and the Black Sea. Mol. Ecol. 11, 755–769.

Lambeck, K., Purcell, A., 2005. Sea-level change in the Mediterranean Sea since the LGM: modelpredictions for tectonically stable areas. Quat. Sci. Rev. 24, 1969-1988.

Larmuseau, M.H.D., Raeymaekers, J.A.M., Hellemans, B., Van Houdt, J.K.J., Volkaert, F.A.M., 2010. Mito-nuclear discordance in the degree of population differentiation in a marine goby. Heredity 105, 532-542.

Lemaire, C., Versini, J-J., Bohomme, F., 2005. Maintenance of genetic differentiation across a transition zone in the sea: discordance between nuclear and cytoplasmic markers. J. Evol. Biol. 18, 70-80.

Li, H., 2011. A new test for detecting recent positive selection that is free of confounding impacts of demography. Mol. Biol. Evol. 28, 365-375.

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

Luttikhuizen, P.C., Campo, J., van Beijswijk, J., Peijnenburg, K.T.C.A., van der Weer, H.W., 2008. Phylogeography of the common shrimp, *Crangon crangon* (L.) across its distribution range.Mol. Phylogenet. Evol. 46, 1015-1030.

Maggs, C.A., Castilho, R., Foltz, D., Henzler, C., Jolly, M.T., Kelly, J., Olsen, J., Perez, K.E., Stam, W., Väinölä, R., Viard, F., Ware, J., 2008. Evaluating signatures of glacial refugia for north Atlantic benthic marine taxa. Ecology 89, S108-S122.

Magoulas, A., Tsimenides, N., Zouros, E., 1996. Mitochondrial DNA phylogeny and the reconstruction of the population history of a species: the case of the European anchovy (*Engraulis encrasicolus*). Mol. Biol. Evol. 13, 178-190.

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27, 209-220.

Marko, P.B., 2002. Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the isthmus of Panama. Mol. Biol. Evol. 16, 2005-2021.

McDonald, J.H., Kreitman, M., 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. Nature 351, 652–654.

Millot, C., Taupier-Letage, I., 2005. Circulation in the Mediterranean Sea. Handbook Env. Che. 5, 29-66.

Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89, 583-590.

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

Nikula, R., Väinölä, R. 2003. Phylogeography of *Cerastoderma glaucum* (Bivalvia: Cardiidae) across Europe: a major break in the Eastern Mediterranean. Mar. Biol. 143, 339-350.

Palumbi, S., 1994. Nucleic acids II: the polymerase chain reaction. In: Hillis, D.M., Moritz, C., Mable, B.K. (eds), *Molecular Systematics*, 2nd ed. Sinauer, Sunderland, pp 205-247.

Papadopoulos, L.N., Peijnenburg, K.T.C.A., Luttikhuizen, P.C., 2005. Phylogeography of the calanoid copepods *Calanus helgolandicus* and *C. euxinus* suggests Pleistocene divergences between Atlantic, Mediterranean, and Black Sea populations. Mar. Biol. 147, 1353-1365.

Passamonti, M., Boore, J.L., Scali, V., 2003. Molecular evolution and recombination in gender-associated mitochondrial DNAs of the Manila clam *Tapes philippinarum.* Genetics 164**,** 603–611.

Patarnello, T., Volkaert, F.A.M.J., Castilho, R., 2007. Pillars of Hercules: is the Atlantic-Mediterranean transition a phylogeographic break?. Mol. Ecol. 16, 4426-4444.

Peijnenburg, K.T.C.A., Breeuwer, J.A.J., Pierrot-Bults, A.C., Menken, B.J., 2004. Phylogeography of the planktonic chaetognath *Sagitta setosa* reveals isolation in European seas. Evolution 58, 1472-1487.

Pérez Camacho, A., 1976. Biología de Venerupis pullastra (Montagu, 1803) y Venerupis decussata (Linne, 1767) (Mollusca, Bivalvia), con especial referencia a los factores determinantes de la producción. Bol. Inst. Esp. Oceanogr. 281, 43-76.

Pérez-Losada, M., Nolte, M.J., Crandall, K.A., Shaw, P.W., 2007. Testing hypotheses of population structuring in the Northeast Atlantic Ocean and Mediterranean Sea using the common cuttlefish *Sepia officinalis*. Mol. Ecol. 16, 2667-2679.

Perissoratis, C., Conispoliatis, N. 2003. The impacts of sea-level changes during latest Pleistocene and Holocene times on the morphology of the Ionian and Aegean seas (SE Alpine Europe). Mar. Geol. 196, 145-156.

Posada, D., 2008. jModelTest: Phylogenetic Model Averaging. *Mol. Biol. Evol.* **25,** 1253-1256.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945-959.

Quesada, H., Beynon, C.H., Skibinski, D.O.F., 1995. A mitochondrial DNA discontinuity in the mussel *Mytilus galloprovincialis* Lmk: Pleistocene vicariance biogeography and secondary intergradation. Mol. Biol. Evol. 12, 521-524.

Ramos-Onsins, S.E., Rozas, J., 2002. Statistical properties of new neutrality tests against population growth. Mol. Biol. Evol. 19, 2092-2100.

Raymond, M., Rousset, F., 1995. An exact test for population differentiation. Evolution 49, 1280-1283.

Reece, K.S., Ribeiro, V.L., Gaffney, P.M., Carnegie, R.B., Allen, S.K., 2004. Microsatellite marker development and analysis in the eastern oyster (*Crassostrea virginica*): confirmation of null alleles and non-Mendelian segregation ratios. J. Hered. 95, 346-352

Rousset, F., 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145, 1219-1228.

Rousset, F., 2008. GENEPOP' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol. Ecol. Resources 8, 103-106.

Sá-Pinto, A., Branco, M.S., Alexandrino, P. B., Fontaine, M.C., Baird, S.E., 2012. Barriers to gene flow in the marine environment: insights from two common intertidal limpet species of the Atlantic and Mediterranean. PLOS One 7, e50330. Doi:10.1371/journal.pone.0050330.

Schunter, C., Carreras-Carbonell, J., Macpherson, E., Tintoré, J., Vidal-Vijande, E., Pascual, A., Guidette, P., Pascual, M., 2011. Matching genetics with oceanography: directional gene flow in a Mediterranean fish species. Mol. Ecol. 20, 5167-5181.

Singh, N.D., Arndt, P.F., Clark, A. G. , Aquadro, C. F., 2009. Strong evidence for lineage- and sequence-specificity of substitution rates and patterns in Drosophila.Mol. Biol. Evol 26, 1591-1605

Slatkin, M., 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47, 264-279.

Sokal, R.R., Rohlf, F.J., 1999. *Biometry*. 3rd ed.,WH Freeman.

Subramanian, A.R., Kaufmann, M., Morgenstern, B., 2008. DIALIGN-TX: greedy and progressive approaches for segment-based multiple sequence alignment. Algorithms Mol. Bi. 3, 6.

Tajima, F., 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105, 437-460.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585-595.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans an chimpanzees. Mol. Biol. Evol. 10, 512–526.

Tanguy, A., Bierne, N., Saavedra, C., Piña, B., Bachère, E., Kube, M., Bazin, E., Bonhomme, F., Boudry, P., Boulo, V., Boutet, I., Cancela, L., Dossat, C., Favrel, P., Huvet, A., Jarque, S., Jollivet, D., Klages, S., Lapègue, S., Leite, R., Moal, J., Moraga, D., Reinhardt, R., Samain, J.F., Zouros, E., Canario, A., 2008. Increasing genomic information in bivalves through new EST collections in four species: development of new genetic markers for environmental studies and genome evolution. Gene 408, 27-36.

Tarnowska, K, Chenuil, A., Nikula, R., Féral, J.-P., Molowicz, M, 2010. Complex genetic population structure of the bivalve *Cerastoderma glaucum* in a highly fragmented lagoon habitat. Mar. Ecol. Progr. Ser. 406, 173-184.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nuc. Acids Res. 22, 4673-4680.

Toews, D.P.L., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol 21, 3907-3930.

Tsaousis, A.D., Martin, D. P., Ladoukakis, E. D., Posada, D., Zouros, E., 2005. Widespread recombination in published animal mtDNA sequences. Mol. Biol. Evol. 22(4), 925–933.

Wares, J.P., Cunningham, C.W., 2001. Phylogeography and historical ecology of the North Atlantic intertidal. Evolution 55, 2455-2469.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38 1358-1370.

Wilson, A.B., Veraguth, I.E., 2010. The impact of Pleistocene glaciations across the range of a widespread European coastal species. Mol. Ecol. 19, 4535-4553.

Wilson, G.A., Rannala, B., 2003. Bayesian Inference of Recent Migration Rates Using Multilocus Genotypes. Genetics 163, 1177-1191.

Wright, S., 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19**,** 395-420.

Xabier, R., Zenboudji, S., Lima, F.P., Harris, D.J., Santos, A.M., Branco, M., 2011. Phylogeography of the marine isopod *Stenosoma nadejda* (Rezig, 1989) in North African Atlantic and western Mediterranean coasts reveals complex differentiation patterns and a new species. Biol. J. Linn. Soc. 104, 419-431.

Zardi, G.I., McQuaid, C.D., Teske, P.R., Barker, N.P., 2007. Unexpected genetic structure of mussel populations in South Africa: indigenous *Perna perna* and invasive *Mytilus galloprovincialis*. Mar. Ecol. Progr. Ser. 337, 135-144.

Figure Legends

Fig. 1. Localities (dots) of *Ruditapes decussatus* (inset) sampled in this study and directions of surface water circulation in MEDAT (white arrows). GMO, Golfe du Morbihan (France); MUG, Mugardos (Spain); LUL, Lombos do Ulla (Spain); MLF, Milfontes (Portugal); FOR, Ria Formosa (Portugal); MME, Mar Menor (Spain); EBR, Ebro delta (Spain); SFX, Sfax (Tunisia); VEN, Venice (Italy); HAL, Halkidiki (Greece); IZM, Izmir (Turkey).

Fig. 2. Haplotype network of mitochondrial COI sequences of *R. decussatus*. Circle sizes represent the frequencies of each haplotype. Black dots are inferred haplotypes that were not present in the sample.

Fig. 3. Allelic and haplotype frequencies at genetic markers in *R. decussatus* populations. The geographic distance is computed as the distance from the northernmost population (GMO). The two dashed vertical lines indicate the locations of the Strait of Gibraltar and the Siculo-Tunisian Strait. Sample labels are given above the plots. (A) Allelic frequencies at 6 intronic markers . ▬□▬: Allele A. ▬●▬: Allele B. ▬ ▬◊▬ ▬: Allele C. •••• ▲ ••••: Allele D. ---○---: Null allele. (B) Haplotype frequencies at COI . — ▲ —: Haplotype A1. — • —○— • —: haplotype A2. ▬■▬: haplotype A4. ▬ ▬□▬ ▬: haplotype B1. ▬●▬: haplotype B2. •••• ∗ ••••: Private haplotypes.

Fig. 4. Bayesian cluster analysis of individual genotypes at six nuclear markers in 11 populations of *R. decussatus*. (A) Posterior probability of the data Pr(X|*K*) (circles) and corresponding values of Pr(*K*|X) (squares) for *K* = 1 to *K* = 13. (B) Estimated membership fraction for each individual and population. Proportions of memberships corresponding to each of the three clusters for each population are given above the chart. The most common cluster appears in bold. Clusters 1, 2 and 3 are shown in yellow, red and blue respectively.

Fig. 5. Locus-by-locus hierarchical *F*-statistics analysis of six nuclear markers under a model of 3-regions (model 1 in Table 6) in *R. decussatus*. *P<0.05. **P<0.01. ***P<0.001.

Fig. 6. Bayesian allelic phylogeny and regional distribution of introns *Ech* and *Tbp*. (A)Locus *Ech,* rooted tree. (B) Locus *Tbp*, unrooted tree. Numbers near nodes are posterior probability values. Allele labels include a population three-letter code and the RFLP allele name. Clade labels appear on the right side of the trees. Symbols in columns indicate the regional origin of each sequence.

Fig. 1

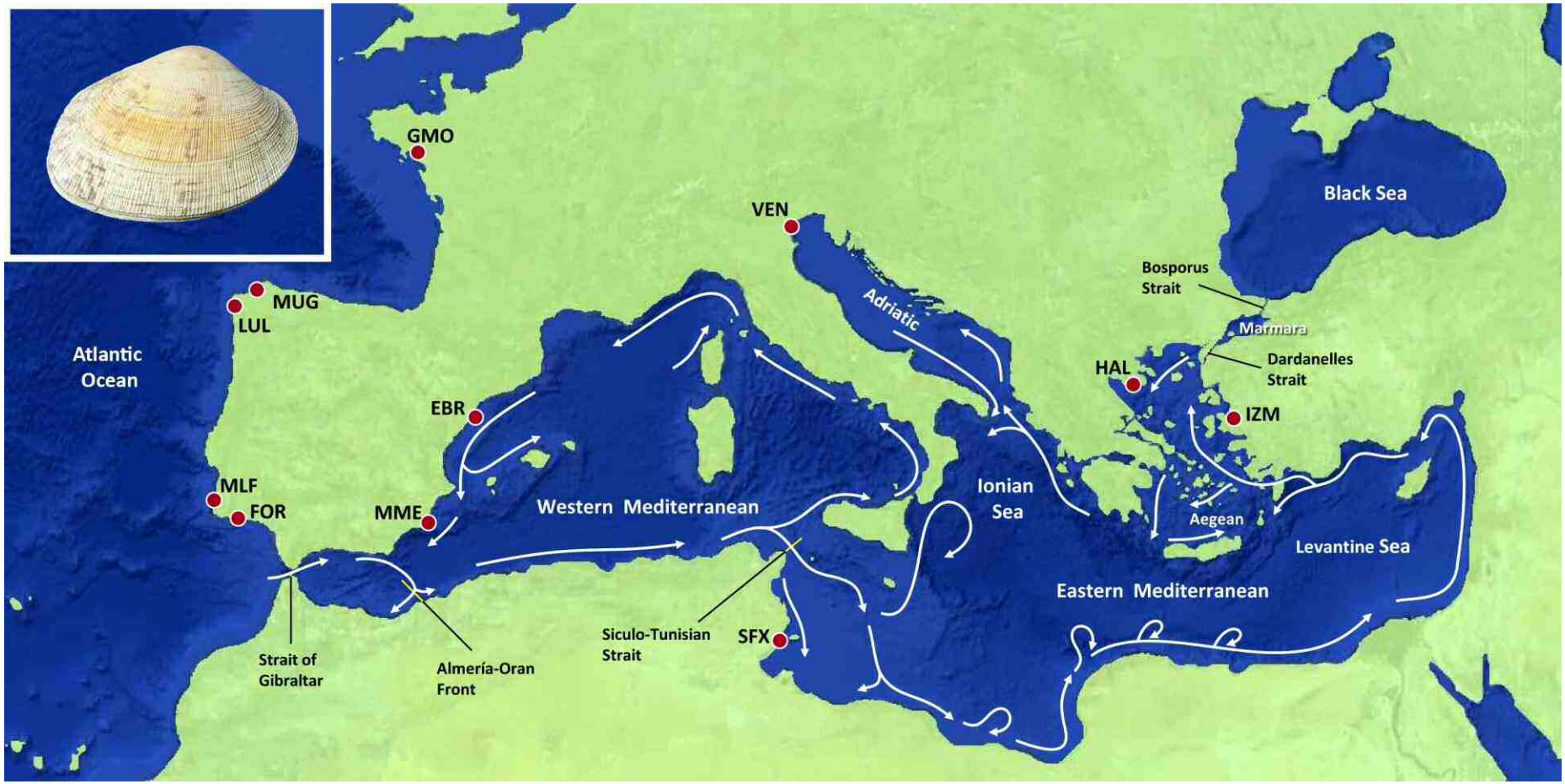

GMO
VEN
Black Sea
Bosporus
Strait
Adriatic
Marmara
Dardanelles
Strait
HAL
IZM
MUG
LUL
Atlantic
Ocean
EBR
MLF
FOR
MME
Western Mediterranean
Ionian
Sea
Aegean
Levantine Sea
Eastern Mediterranean
Siculo-Tunisian
Strait
SFX
Strait of
Gibraltar
Almería-Oran
Front

Fig. 2

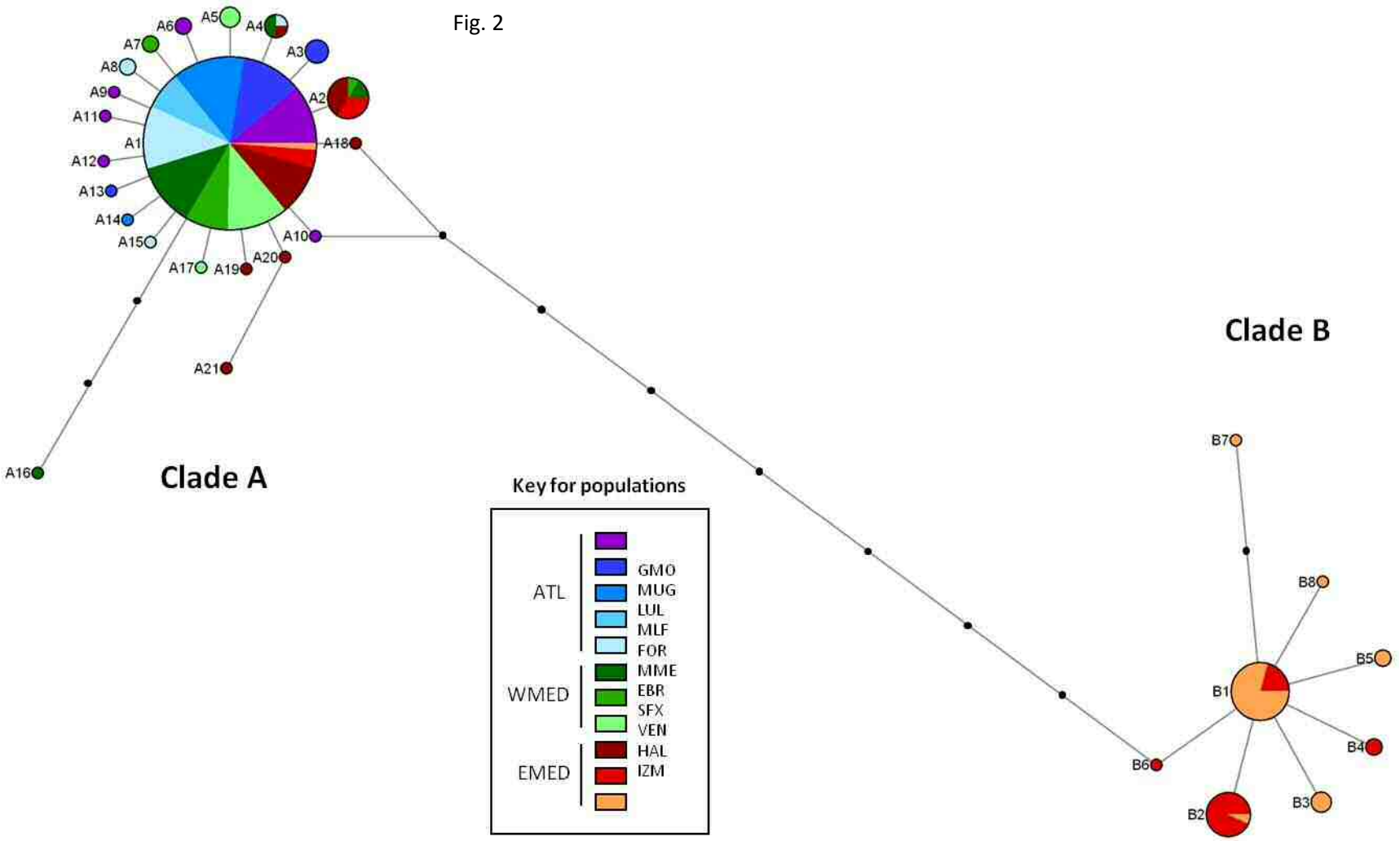

A1
A2
A3
A4
A5
A6
A7
A8
A9
A10
A11
A12
A13
A14
A15
A16
A17
A18
A19
A20
A21
Clade A
Clade B
B1
B2
B3
B4
B5
B6
B7
B8
Key for populations
ATL
GMO
MUG
LUL
MLF
FOR
WMED
MME
EBR
SFX
VEN
EMED
HAL
IZM

Fig. 3

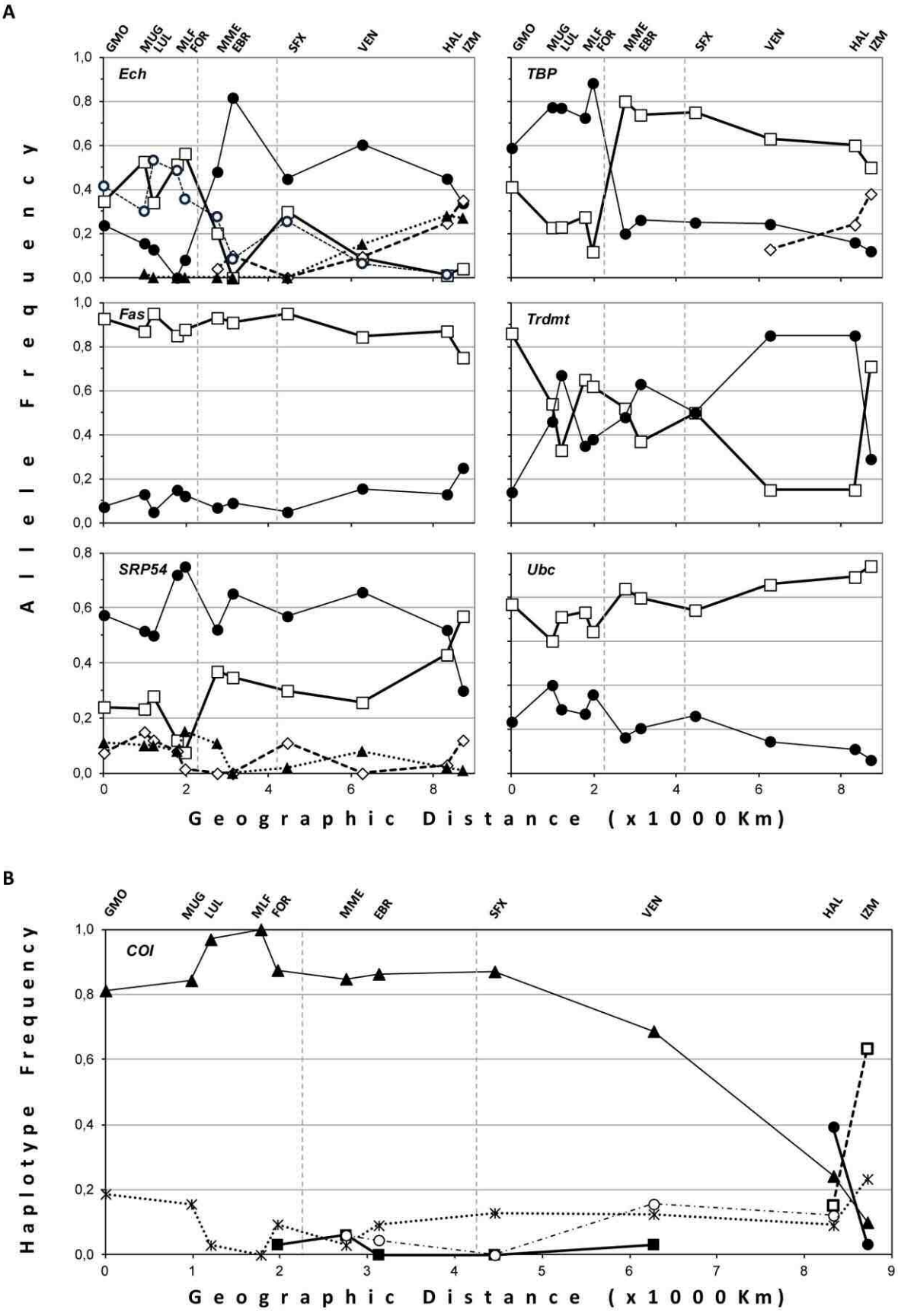

A
GMO
MUG LUL
MLF FOR
MME EBR
SFX
VEN
HAL IZM
Ech
TBP
Fas
Trdmt
SRP54
Ubc
Allele Frequency
Geographic Distance (x1000Km)
B
COI
Haplotype Frequency
Geographic Distance (x1000Km)

Fig. 4

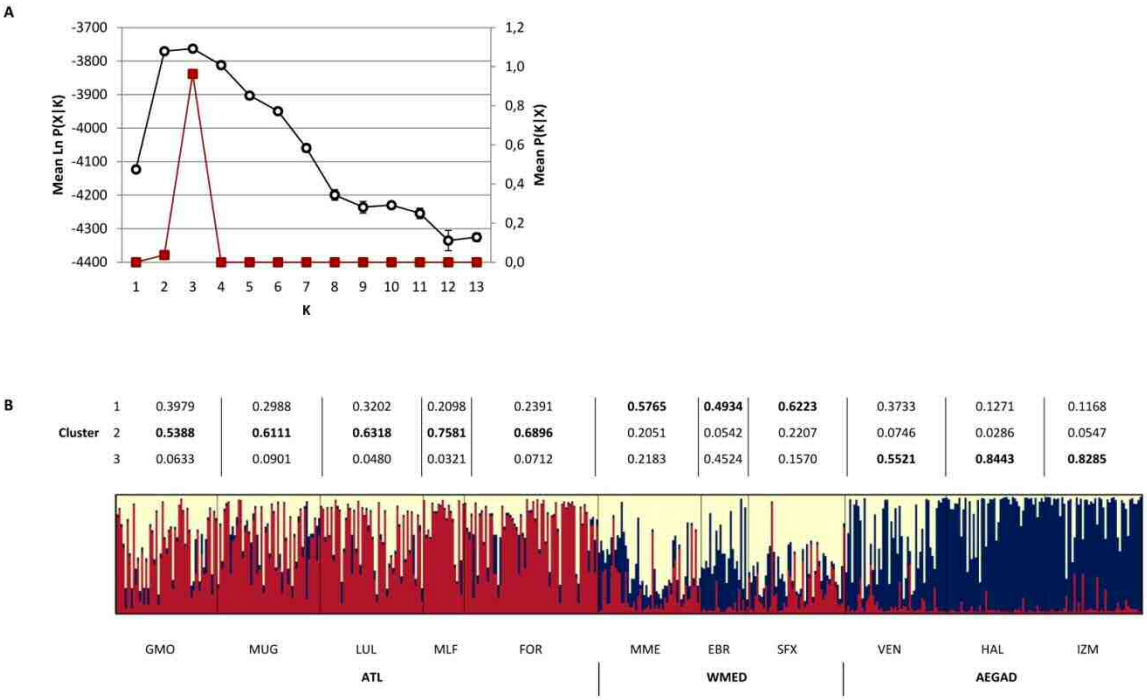

A
-3700
-3800
-3900
-4000
-4100
-4200
-4300
-4400
Mean Ln P(X|K)
1,2
1,0
0,8
0,6
0,4
0,2
0,0
Mean P(K|X)
1 2 3 4 5 6 7 8 9 10 11 12 13
K
B
Cluster
1 0.3979 0.2988 0.3202 0.2098 0.2391 0.5765 0.4934 0.6223 0.3733 0.1271 0.1168
2 0.5388 0.6111 0.6318 0.7581 0.6896 0.2051 0.0542 0.2207 0.0746 0.0286 0.0547
3 0.0633 0.0901 0.0480 0.0321 0.0712 0.2183 0.4524 0.1570 0.5521 0.8443 0.8285
GMO MUG LUL MLF FOR MME EBR SFX VEN HAL IZM
ATL
WMED
AEGAD

Fig. 5

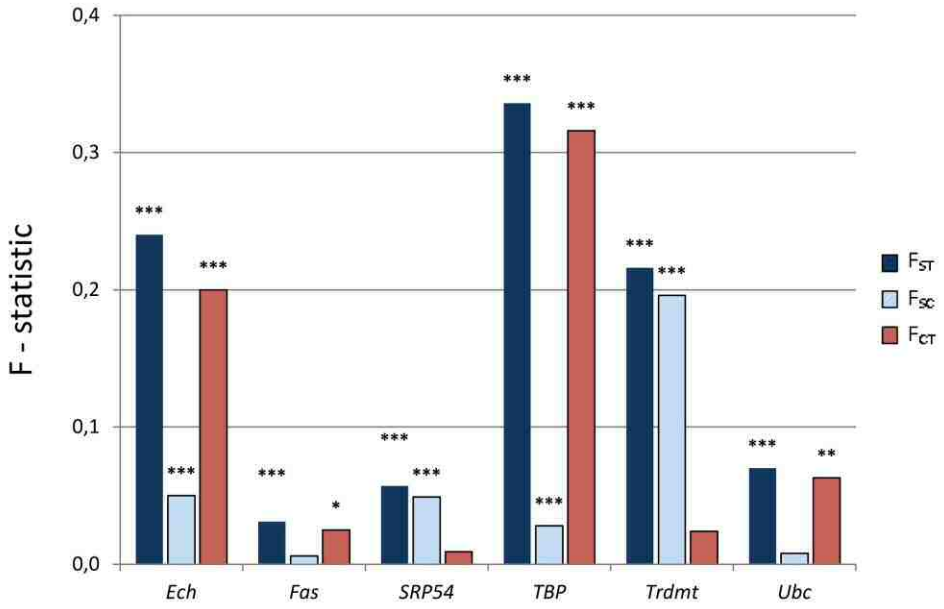

F - statistic
0,4
0,3
0,2
0,1
0,0
Ech
Fas
SRP54
TBP
Trdmt
Ubc
F_{ST}
F_{SC}
F_{CT}

Fig. 6

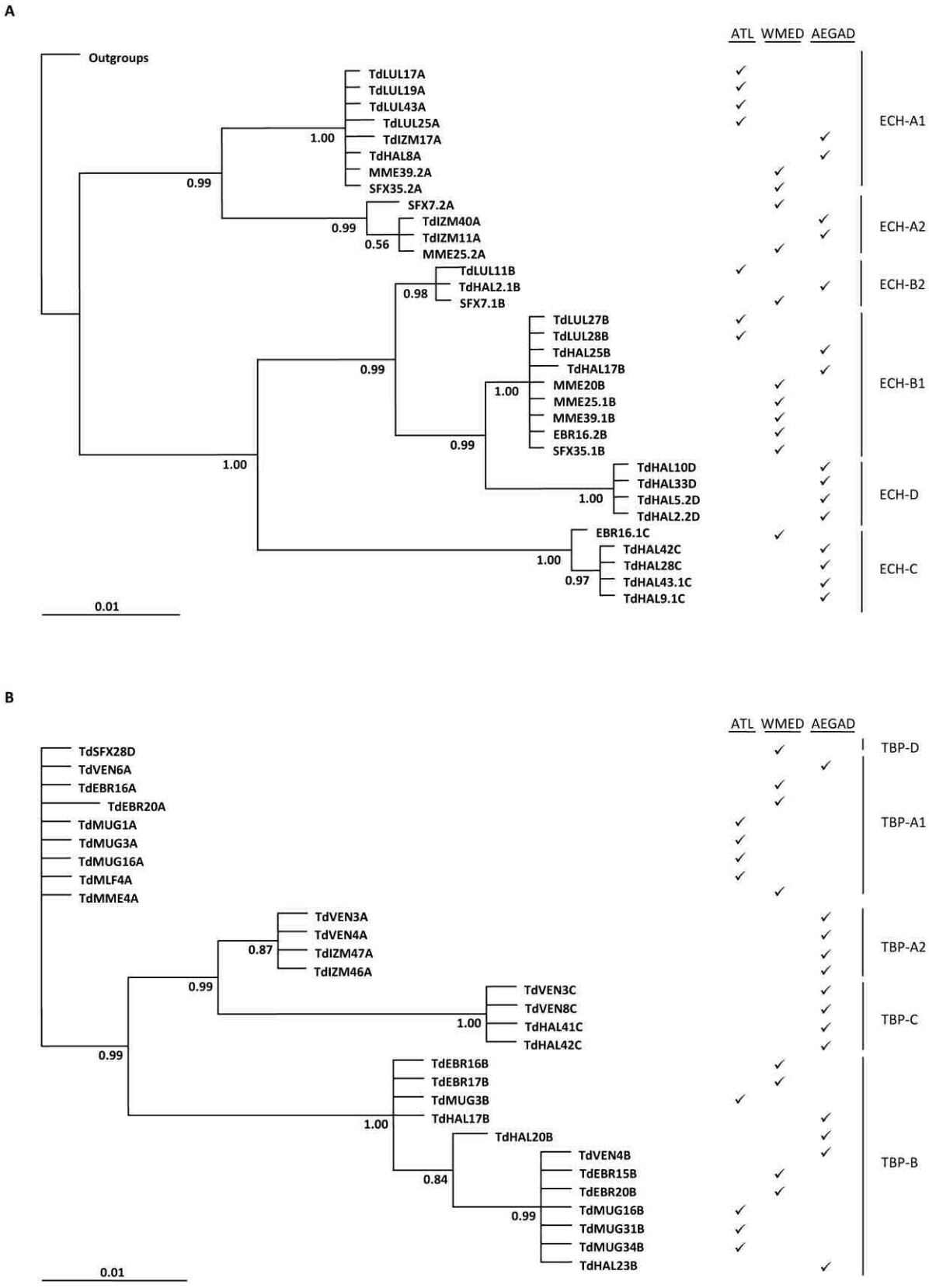

A
ATL WMED AEGAD
Outgroups
TdLUL17A
TdLUL19A
TdLUL43A
TdLUL25A
TdIZM17A
TdHAL8A
MME39.2A
SFX35.2A
SFX7.2A
TdIZM40A
TdIZM11A
MME25.2A
TdLUL11B
TdHAL2.1B
SFX7.1B
TdLUL27B
TdLUL28B
TdHAL25B
TdHAL17B
MME20B
MME25.1B
MME39.1B
EBR16.2B
SFX35.1B
TdHAL10D
TdHAL33D
TdHAL5.2D
TdHAL2.2D
EBR16.1C
TdHAL42C
TdHAL28C
TdHAL43.1C
TdHAL9.1C
1.00
0.99
0.99
0.56
0.98
0.99
1.00
0.99
1.00
1.00
0.97
0.01
ECH-A1
ECH-A2
ECH-B2
ECH-B1
ECH-D
ECH-C
B
ATL WMED AEGAD
TdSFX28D
TdVEN6A
TdEBR16A
TdEBR20A
TdMUG1A
TdMUG3A
TdMUG16A
TdMLF4A
TdMME4A
TdVEN3A
TdVEN4A
TdIZM47A
TdIZM46A
TdVEN3C
TdVEN8C
TdHAL41C
TdHAL42C
TdEBR16B
TdEBR17B
TdMUG3B
TdHAL17B
TdHAL20B
TdVEN4B
TdEBR15B
TdEBR20B
TdMUG16B
TdMUG31B
TdMUG34B
TdHAL23B
0.87
0.99
1.00
0.99
1.00
0.84
0.99
0.01
TBP-D
TBP-A1
TBP-A2
TBP-C
TBP-B

Table 1 Estimates of gene diversity and neutrality tests for the COI gene in 11 populations of *R. decussatus*. Clades A and B had been analyzed separately in the two populations where they appear toghether. *N*: number of sequences; *S*: segregating sites; *h*: number of haplotypes; *Hd*: haplotype diversity; π: nuclotide diversity estimated from pairwise differences; Θ_S: nucleotide diversity based on the number of segregating sites; *k*: average number of nucleotide differences; R_2, *D*, F_s and *H* are statistics for different neutrality tests (see Methods).

Sample	*N*	*S*	*h*	*Hd*	π (%)	Θ_S (%)	*k*	R_2	*D*	F_S	*H*
GMO	32	5	6	0.34 ± 0.11	0.09 ± 0.03	0.29 ± 0.15	0.37	0.067	-1.882*	-4.959***	-1.581*
MUG	32	2	3	0.28 ± 0.09	0.07 ± 0.02	0.12 ± 0.09	0.29	0.104	-0.843	-0.934	-1.294*
LUL	33	1	2	0.06 ± 0.06	0.01 ± 0.01	0.06 ±0.06	0.06	0.171	-1.140	-1.290	-
MLF	17	0	1	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00	-	-	-	-
FOR	32	3	4	0.24 ± 0.10	0.06 ± 0.03	0.17 ± 0.11	0.25	0.083	-1.548	-2.818**	-3.569***
MME	33	5	4	0.28 ± 0.10	0.10 ± 0.04	0.29 ± 0.15	0.42	0.106	-1.763	-1.551	-3.354**
EBR	22	2	3	0.26 ± 0.12	0.06 ± 0.03	0.13 ± 0.10	0.26	0.113	-1.175	-1.310	0.242
SFX	31	2	3	0.24 ± 0.10	0.06± 0.02	0.12 ± 0.09	0.25	0.099	-1.035	-1.244	-1.443*
VEN	32	6	7	0.51 ± 0.10	0.15 ± 0.04	0.35 ± 0.17	0.64	0.064	-1.596	-4.350***	-5.343**
HAL	33	13	6	0.77 ± 0.05	1.30 ± 0.12	0.74 ± 0.30	5.62	-	-	-	-
IZM	30	16	7	0.59 ± 0.10	0.57 ± 0.20	0.94 ± 0.36	2.44	-	-	-	-
HAL clade A	12	1	2	0.49 ± 0.11	0.11 ± 0.03	0.08 ± 0.08	0.49	0.242	1.066	1.003	0.242
HAL clade B	21	3	4	0.58 ± 0.10	0.18 ± 0.04	0.19 ± 0.12	0.77	0.133	-0.197	-0.549	-1.457
IZM clade A	3	0	1	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00	-	-	-	-
IZM clade B	27	6	6	0.50 ± 0.11	0.15 ± 0.04	0.36 ± 0.18	0.64	0.080	-1.719	-3.205**	-3.382**
Total	327	36	29	0.47 ± 0.03	0.68 ± 0.07	1.31 ± 0.34	2.92	-	-	-	-

* $P<0.05$, ** $P<0.01$, *** $P<0.001$.

Table 2

McDonald-Kreitman tests of neutrality for COI in *R. decussatus.* NI: neutrality index.

Species pair		Fixed	Polymorphic	*P* (Fisher Exact Test)	NI
R. decussatus / *V. rhomboides*	Non-synonymous	11	1	0.054	0.151
	Synonymous	58	35		
R. decussatus Clade A / R. decussatus Clade B	Non-synonymous	0	1	1.000	-
	Synonymous	6	28		

* $P < 0.05$

Table 3
Analysis of molecular variance (AMOVA) of COI in 11 populations of *R. decussatus*.

Model	Source of variation	d.f.	Sum of squares	% of Variation	Φ statistic
Models without subdivision					
1 - All populations					
	Among populations	10	323.22	66.89	Φ_{ST} = 0.669 ***
	Within populations	316	167.76	33.11	
2 - Atlantic populations					
	Among populations	4	0.78	2.80	Φ_{ST} = 0.028 **
	Within populations	141	15.08	97.20	
3 - Mediterranean populations					
	Among populations	5	262.70	66.32	Φ_{ST} = 0.663 ***
	Within populations	175	152.68	33.68	
4 - Mediterranean excluding Aegean					
	Among populations	3	1.10	2.66	Φ_{ST} = 0.027*
	Within populations	114	23.27	97.34	
5 - Aegean populations					
	Among populations	1	14.17	15.31	Φ_{ST} = 0.153***
	Within populations	61	129.41	84.69	
Models with subdivision					
6 - Atlantic vs. Mediterranean					
	Among basins	1	59.73	10.81	Φ_{ST} = 0.685 ***
	Among populations within basins	9	263.49	57.70	Φ_{SC} = 0.647 ***
	Within populations	316	167.76	31.48	Φ_{CT} = 0.108
7 - Aegean vs.rest of populations					
	Among basins	1	303.63	83.89	Φ_{ST} = 0.851 ***
	Among populations within basins	9	16.59	1.25	Φ_{SC} = 0.077**
	Within populations	316	167.76	14.87	Φ_{CT} = 0.839 *
8 - Eastern Mediterranean vs. rest of populations					
	Among basins	1	117.14	31.81	Φ_{ST} = 0.719 ***
	Among populations within basins	9	206.08	40.11	Φ_{SC} = 0.588 ***
	Within populations	316	167.76	28.08	Φ_{CT} = 0.318

* P < 0.05, ** P < 0.01, *** P < 0.001

Table 4

Hierarchical F-statistic analysis based on 6 nuclear markers in *R. decussatus*. Regional grouping of populations is as follows: ATL (GMO, MUG, LUL, MLF, FOR); AEGAD (VEN, HAL, IZM); WMED (MME, EBR, SFX) ; Aegean (HAL, IZM).

	F-statistics			Percentage of variation		
	F_{SC}	F_{ST}	F_{CT}	Among groups	Among populations (within groups)	Within populations
Whole area models with regional subdivision						
1 - ATL / WMED / AEGAD	0.044***	0.166***	0.128***	12.8	3.8	83.4
2 - ATL / WMED + VEN / Aegean	0.043***	0.168***	0.131***	13.1	3.8	83.2
3 - Aegean / rest of populations	0.109***	0.178***	0.078*	7.8	10.1	82.2
4- ATL / Mediterranean	0.057***	0.188***	0.139***	13.9	5.0	81.2
Regional models with subdivision						
5 - ATL / WMED	0.016***	0.143***	0.128*	12.8	1.4	85.7
6 - WMED / AEGAD	0.063***	0.095***	0.034	3.4	6.1	90.5
7 - WMED + VEN / Aegean	0.062***	0.098***	0.039	3.9	5.9	90.2

* P < 0.05, ** P < 0.01, *** P < 0.001

Table 5

Migration rates between regions (Means and 95% confidence intervals of posterior distribution) estimated from 6 nuclear loci with BayesAss v.1.3. Bold values in the diagonal are the proportion of individuals in each generation that are not inmigrants.

Source	ATL	WMED	AEGAD
ATL	**0.9950 (0.9803, 0.9998)**	0.0148 (8.75e-05, 0.0592)	0.0059 (0.0001, 0.0227)
WMED	0.0035 (3.43e-05, 0.0174)	**0.9807 (0.9351, 0.9994)**	0.0420 (0.0009, 0.1106)
AEGAD	0.0015 (2.35e-05, 0.0066)	0.0045 (4.73e-05, 0.0180)	**0.9521 (0.8834, 0.9969)**

Table 6

Estimates of the mean times of the most recent common ancestors (*tmrca*), with their 95% highest posterior probability (HPD) intervals, obtained with BEAST 1.7 for different clades of the three sequenced markers. A coalescent constant size prior was used in all cases. The type of molecular clock prior is indicated as strict or uncorrelated log-normal (UCLN). ESS is the effective sample size.

Marker	Clock type	Clade	*tmrca* Mean (Ka)	*tmrca* 95% HPD	ESS
COI	Strict	AB	1,575	146– 5,340	576
Ech	UCLN	A1A2*	424	35 – 1,170	1,473
		B1D*	206	19 – 573	1,634
		B1B2D	373	32 – 1,010	1,551
		B1B2CD*	1,078	133 – 2,971	1,537
TBP	Strict	A2C	273	34 - 817	2,376
		A1A2C	416	47 -1,255	2,248

Table S1
Intronic markers in *R. decussatus*. EST- Accessions refer to the Genbank accession numbers of the ESTs used to design primers.

Marker	Coded protein	EST - Accessions	Primers forward (F) and reverse (R)	PCR Product length	Nº of introns detected (length in bp)	Cloned PCR product Accessions
Ech	Enoyl Coenzyme A hydratase	AM851315, AM851542	F:5'gaaattgcgatgatgtgtga3' R:5'actgcctcatcgactacctg3'	1169	2 (558, 380)	FN428628-FN428632
Fas	Fasciclin-like protein	AM852179, AM851792	F:5'catatttgcaccatctgacg3' R:5'gtgaccaaggagaaggttcc3'	453	1(277)	FN428594-FN428600
Trdmt	tRNA aspartic acid methyltransferase	AM851822, AM852241	F:5'tctcatacagtgcggctaca3' R:5'cttgaccttctgttgccatt3'	1191	2(400, 567)	FN428612-FN428617
Ubc	Ubiquitin conjugating enzyme	AM852160, AM851337	F:5'tgtcaacaccagcaagaaga3' R:5'catcctcaaatggggtatca3'	988	2(119,717)	FN428623-FN428627

Table S2 COI nucleotide variable sites.

Haplotype	Nucleotide site																																		
	2	29	53	71	83	89	107	110	113	146	152	161	164	170	179	191	203	209	210	218	231	239	251	281	284	291	303	317	329	350	365	374	377	390	403	410
A1	A	A	C	G	G	G	A	A	T	T	A	A	T	T	C	T	T	T	T	T	T	T	T	C	T	C	T	A	T	C	C	G	T	T	T	A
A2	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A3	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.
A5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A6	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.
A8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A9	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A11	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A12	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A13	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A14	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A16	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A17	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
A20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
A21	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.
B1	.	.	T	.	A	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	C	.	.	T	C	.	.	G	.	.	T	.	.	.	.	G
B2	.	.	T	.	A	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	C	.	.	T	C	.	C	G	.	.	T	.	.	.	.	G
B3	.	.	T	.	A	.	.	.	.	.	.	.	.	.	.	C	C	.	C	.	C	.	.	T	C	.	.	G	.	.	T	.	.	.	.	G
B4	.	.	T	A	A	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	C	.	.	T	C	.	.	G	.	.	T	.	.	.	.	G
B5	.	.	T	.	A	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	C	.	C	T	C	.	.	G	.	.	T	.	.	.	.	G
B6	.	.	T	.	A	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	C	.	.	T	.	.	.	G	.	.	T	.	.	.	.	G
B7	G	.	T	.	A	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	C	.	.	T	C	.	.	G	.	.	T	.	.	C	.	G
B8	.	.	T	.	A	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	C	.	.	T	C	.	.	G	.	T	T	.	.	.	.	G

Table S3 Haplotype frecuencies at the COI locus in 11 populations of *R. decussatus*.

Haplotype	Populations											Total
	GMO	MUG	LUL	MLF	FOR	MME	EBR	SFX	VEN	HAL	IZM	
A1	26	27	32	17	28	28	19	27	22	8	3	237
A2	-	-	-	-	-	2	1	-	5	4	-	12
A3	-	4	-	-	-	-	-	-	-	-	-	4
A4	-	-	-	-	1	2	-	-	1	-	-	4
A5	-	-	-	-	-	-	-	3	-	-	-	3
A6	2	-	-	-	-	-	-	-	-	-	-	2
A7	-	-	-	-	-	-	2	-	-	-	-	2
A8	-	-	-	-	2	-	-	-	-	-	-	2
A9	1	-	-	-	-	-	-	-	-	-	-	1
A10	1	-	-	-	-	-	-	-	-	-	-	1
A11	1	-	-	-	-	-	-	-	-	-	-	1
A12	1	-	-	-	-	-	-	-	-	-	-	1
A13	-	1	-	-	-	-	-	-	-	-	-	1
A14	-	-	1	-	-	-	-	-	-	-	-	1
A15	-	-	-	-	1	-	-	-	-	-	-	1
A16	-	-	-	-	-	1	-	-	-	-	-	1
A17	-	-	-	-	-	-	-	1	-	-	-	1
A18	-	-	-	-	-	-	-	-	1	-	-	1
A19	-	-	-	-	-	-	-	-	1	-	-	1
A20	-	-	-	-	-	-	-	-	1	-	-	1
A21	-	-	-	-	-	-	-	-	1	-	-	1
B1	-	-	-	-	-	-	-	-	-	5	19	24
B2	-	-	-	-	-	-	-	-	-	13	1	14
B3	-	-	-	-	-	-	-	-	-	-	3	3
B4	-	-	-	-	-	-	-	-	-	2	-	2
B5	-	-	-	-	-	-	-	-	-	-	2	2
B6	-	-	-	-	-	-	-	-	-	1	-	1
B7	-	-	-	-	-	-	-	-	-	-	1	1
B8	-	-	-	-	-	-	-	-	-	-	1	1
Total	32	32	33	17	32	33	22	31	32	33	30	327

Table S4

Allele frequencies, sample size (N), heterozygosity (h), and deviations from Hardy Weinberg equilibrium (F_{IS}) at 6 nuclear markers in 11 populations of *R. decussatus.* F_{IS} at *Ech* were calculated before estimating the null allele frequency.

Marker	Allele		GMO	MUG	LUL	MLF	FOR	MME	EBR	SFX	VEN	HAL	IZM
Ech	A		0,347	0,526	0,341	0,514	0,563	0,202	0,000	0,299	0,085	0,011	0,041
	B		0,238	0,156	0,128	0,000	0,081	0,480	0,816	0,488	0,603	0,449	0,338
	C		0,000	0,000	0,000	0,000	0,000	0,041	0,099	0,000	0,095	0,247	0,351
	D		0,000	0,017	0,000	0,000	0,000	0,000	0,000	0,000	0,152	0,279	0,270
	Null		0,415	0,301	0,532	0,486	0,356	0,276	0,084	0,254	0,064	0,014	0,000
		N	46	40	46	20	53	38	22	39	48	44	40
		H	0,658	0,609	0,592	0,512	0,558	0,658	0,321	0,654	0,597	0,665	0,695
		F_{is}	0.689***	0.465**	0.702***	0.000	0.836***	0.428***	0.344	0.417*	0.15	0.033	-0.087
Fas	A		0.927	0.870	0.946	0.850	0.878	0.930	0.909	0.952	0.844	0.866	0.750
	B		0.073	0.130	0.054	0.150	0.122	0.070	0.091	0.048	0.156	0.134	0.250
		N	48	46	46	20	45	43	22	42	48	41	46
		h	0.137	0.229	0.104	0.262	0.217	0.131	0.169	0.092	0.266	0.235	0.379
		F_{is}	-0.068	0.053	-0.047	-0.152	-0.128	-0.063	-0.077	-0.038	-0.175	0.067	-0.091
SRP54	A		0.241	0.233	0.277	0.125	0.076	0.370	0.348	0.295	0.233	0.426	0.571
	B		0.574	0.517	0.500	0.725	0.758	0.522	0.652	0.568	0.700	0.521	0.300
	C		0.074	0.150	0.117	0.075	0.015	0.000	0.000	0.114	0.000	0.032	0.114
	D		0.111	0.100	0.106	0.075	0.152	0.109	0.000	0.023	0.067	0.021	0.014
		N	27	30	47	20	33	23	23	44	35	47	35
		h	0.606	0.657	0.655	0.459	0.403	0.592	0.464	0.583	0.502	0.552	0.578
		F_{is}	-0.04	-0.067	-0.072	-0.092	0.252**	0.047	0.064	-0.093	0.075	0.153	-0.239
Tbp	A		0.412	0.227	0.229	0.275	0.118	0.800	0.738	0.739	0.629	0.604	0.500
	B		0.588	0.773	0.771	0.725	0.882	0.200	0.262	0.239	0.243	0.156	0.115
	C		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.129	0.240	0.385
	D		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.022	0.000	0.000	0.000
		N	34	33	48	20	34	35	21	46	35	48	39
		h	0.492	0.357	0.357	0.409	0.211	0.325	0.396	0.400	0.537	0.559	0.596
		F_{is}	-0.078	-0.106	-0.051	0.147	-0.119	-0.057	0.162	-0.142	-0.119	-0.006	-0.119

Table S5
(continued)

Marker	Allele		GMO	MUG	LUL	MLF	FOR	MME	EBR	SFX	VEN	HAL	IZM
Trdmt	A		0.857	0.536	0.330	0.650	0.622	0.524	0.368	0.500	0.153	0.146	0.707
	B		0.143	0.464	0.670	0.350	0.378	0.476	0.632	0.500	0.847	0.854	0.293
		N	49	42	47	20	45	41	19	37	49	41	41
		h	0.247	0.503	0.447	0.467	0.475	0.505	0.478	0.507	0.262	0.253	0.419
		F_{is}	0.177	-0.185	-0.193	0.363	-0.123	0.568***	-0.104	-0.122	0.144	0.617***	0.304
Ubc	A		0.767	0.600	0.713	0.725	0.643	0.837	0.795	0.744	0.856	0.895	0.938
	B		0.233	0.400	0.287	0.275	0.357	0.163	0.205	0.256	0.144	0.105	0.063
		N	45	40	47	20	56	43	22	41	45	38	40
		h	0.362	0.486	0.414	0.409	0.463	0.276	0.333	0.386	0.250	0.191	0.119
		F_{is}	0.08	0.283	0.231	-0.103	-0.002	-0.012	-0.235	-0.076	-0.158	-0.105	-0.054
		Mean h	0.417	0.474	0.428	0.401	0.388	0.415	0.360	0.437	0.402	0.409	0.465
		± S.E.	0.084	0.065	0.079	0.037	0.059	0.083	0.046	0.081	0.065	0.084	0.084
		Mean F_{is}	0.126	0.074	0.095	0.033	0.119	0.152	0.025	-0.009	-0.014	0.127	-0.048
		± S.E.	0.120	0.103	0.134	0.089	0.155	0.112	0.085	0.086	0.063	0.104	0.075

*P<0.05, **P<0.01, ***P<0.001

Table S6 Pairwise F_{ST} per locus for AEGAD populations.

Population pair	*Ech*	*Fas*	*Srp54*	*Tbp*	*Trdmt*	*Ubc*
VEN-HAL	0.045	-0.009	0.056	0.007	-0.015	-0.004
VEN-IZM	0.097	-0.018	0.215	0.070	0.475	0.025
HAL-IZM	0.005	0.031	0.053	0.018	0.478	0.000

Table S7 Directional migration between populations. Means and 95% confidence intervals of posterior distribution of recent migration rates were calculated based on the 6 nuclear loci. Bold values are the proportion of individuals in each generation that are not inmigrants. Migrant sources are indicated in columm on the left side of the table.

	GMO	MUG	LUL	MLF	FOR	MME	EBR	SFX	VEN	HAL	IZM
GMO	**0.8916 (0.6746, 0.9991)**	0.0423 (0.0006, 0.1237)	0.0050 (2.73e-10, 0.0317)	0.0124 (1.93e-11, 0.0790)	0.0073 (6.74e-11, 0.0477)	0.0102 (9.39e-09, 0.0542)	0.0034 (3.59e-14, 0.0252)	0.0135 (6.59e-12, 0.0759)	0.0071 (1.12e-07, 0.0337)	0.0034 (3.92e-09, 0.0199)	0.0023 (1.92e-11, 0.0184)
MUG	0.0087 (1.54e-11, 0.0550)	**0.7151 (0.6685, 0.8555)**	0.0129 (1.88e-09, 0.0887)	0.0104 (7.03e-12, 0.0566)	0.0072 (8.16e-11, 0.0518)	0.0061 (2.63e-09, 0.0340)	0.0027 (4.02e-13, 0.0189)	0.0087 (2.08e-11, 0.0475)	0.0062 (1.61e-07, 0.029)	0.0023 (1.50e-10, 0.0137)	0.0025 (1.70e-11, 0.0193)
LUL	0.0094 (2.57e-13, 0.0527)	0.0597 (0.0019, 0.2019)	**0.9387 (0.7747, 0.9981)**	0.0151 (3.30e-12, 0.0816)	0.0228 (2.79e-12, 0.1576)	0.0067 (1.18e-09, 0.0378)	0.0028 (3.01e-14, 0.0237)	0.0164 (1.84e-10, 0.0818)	0.0075 (8.25e-07, 0.0367)	0.0026 (2.45e-09, 0.0180)	0.0019 (2.23e-11, 0.0129)
MLF	0.0255 (9.62e-11, 0.1893)	0.0405 (0.0007, 0.1270)	0.0128 (4.75e-10, 0.1293)	**0.9075 (0.6805, 0.9985)**	0.0322 (4.86e-11, 0.2190)	0.0078 (3.57e-09, 0.0427)	0.0034 (5.10e-13, 0.0242)	0.0112 (2.79e-10, 0.0607)	0.0067 (1.76e-07, 0.0332)	0.0023 (1.64e-10, 0.0142)	0.0021 (1.24e-11, 0.0164)
FOR	0.0151 (1.05e-09, 0.0913)	0.0607 (0.0010, 0.2032)	0.0062 (2.13e-10, 0.0379)	0.0183 (7.66e-11, 0.1274)	**0.9127 (0.6830, 0.9991)**	0.0051 (5.84e-09, 0.0285)	0.0026 (6.97e-13, 0.0216)	0.0096 (5.96e-10, 0.0459)	0.0065 (3.45e-07, 0.0322)	0.0024 (6.04e-10, 0.0146)	0.0018 (1.64e-12, 0.0129)
MME	0.0086 (2.45e-11, 0.0582)	0.0181 (0.0002, 0.0595)	0.0055 (5.38e-09, 0.0338)	0.0076 (1.90e-12, 0.0399)	0.0039 (4.96e-11, 0.0246)	**0.9302 (0.8209, 0.9967)**	0.0040 (1.95e-14, 0.0278)	0.0367 (1.96e-10, 0.1733)	0.0166 (4.16e-07, 0.0795)	0.0054 (5.02e-11, 0.0286)	0.0037, (1.02e-11 0.0247)
EBR	0.0056 (7.23e-11, 0.0338)	0.0107 (0.0001, 0.0395)	0.0057 (9.68e-10, 0.0361)	0.0055 (1.68e-12, 0.0326)	0.0032 (5.95e-11, 0.0206)	0.0086 (2.36e-09, 0.0521)	**0.9688 (0.8995, 0.9989)**	0.0341 (1.69e-10, 0.1735)	0.0517 (5.24e-07, 0.2132)	0.0121 (2.41e-09, 0.0619)	0.0044 (9.07e-12, 0.0267)
SFX	0.0245 (2.28e-10, 0.1682)	0.0288 (0.0004, 0.1008)	0.0062 (2.01e-10, 0.0380)	0.0081 (2.33e-11, 0.0453)	0.0032 (4.45e-11, 0.0196)	0.0117 (5.45e-09, 0.0620)	0.0036 (5.70e-15, 0.0260)	**0.8510 (0.6710, 0.9985)**	0.0128 (9.53e-07, 0.0541)	0.0036 (1.21e-09, 0.0195)	0.0040 (4.49e-12, 0.0275)
VEN	0.0041 (1.70e-10, 0.0219)	0.0094 (9.56e-05, 0.0363)	0.0024 (4.90e-10, 0.0155)	0.0058 (4.52e-12, 0.0362)	0.0033 (7.99e-13, 0.0218)	0.0058 (1.63e-09, 0.0332)	0.0030 (7.63e-13, 0.0235)	0.0090 (3.19e-11, 0.0510)	**0.8376 (0.6729, 0.9870)**	0.0234 (6.31e-09, 0.1389)	0.0038 (2.06e-11, 0.0287)
HAL	0.0037 (4.37e-11, 0.0224)	0.0065 (6.77e-05, 0.0275)	0.0027 (3.49e-10, 0.0177)	0.0046 (4.13e-11, 0.0247)	0.0024 (1.07e-12, 0.0148)	0.0041 (1.66e-09, 0.0237)	0.0029 (1.05e-13, 0.0210)	0.0057 (3.22e-11, 0.0286)	0.0415 (1.82e-06, 0.1797)	**0.9218 (0.7490, 0.9983)**	0.0079 (8.48e-11, 0.0718)
IZM	0.0030 (6.41e-11, 0.0176)	0.0079 (0.0001, 0.0329)	0.0019 (8.52e-11, 0.0123)	0.0048 (7.37e-12, 0.0327)	0.0020 (9.42e-12, 0.0122)	0.0036 (3.42e-09, 0.0230)	0.0029 (1.27e-13, 0.0225)	0.0040 (1.30e-10, 0.0210)	0.0057 (1.83e-07, 0.0295)	0.0207 (2.03e-09, 0.0748)	**0.9655 (0.8780, 0.9990)**

Supplementary Figure Legends

Fig. S1 Allelic polymorphisms in 6 nuclear markers of *R. decussatus*. For markers scored with restriction enzymes, fragment sizes are given inside boxes, and the total size is indicated at the right side of each allele. Restriction enzymes are indicated with their initial letter (*Bam*H I, *Dra* I and *Eco*R I) at the cutting point.

Fig.S2 Haplotype phylogeny of *COI*. Haplotypes as defined in Table 2. Three closely related species of the same family, *Venerupis rhomboides* (VrCo04), *Venerupis pullastra* (VpCo02) and *Venerupis aurea* (Va15) were used as outgroups.

Fig. S3 Isolation by distance in *R. decussatus*. Slatkin's linearized genetic distances are plotted against geographic distances for all population pairs and coastal marine distances. (a) Φ_{ST} from mitochondrial COI. (b) F_{ST} from mitochondrial COI. (c) F_{ST} estimated from 6 nuclear loci. Statistic significance of Mantel tests were computed over 1000 permutations.

Fig. S4 Bayesian skyline plots for mitochondrial clades A (a) and B (b) obtained with BEAST.

Fig S1

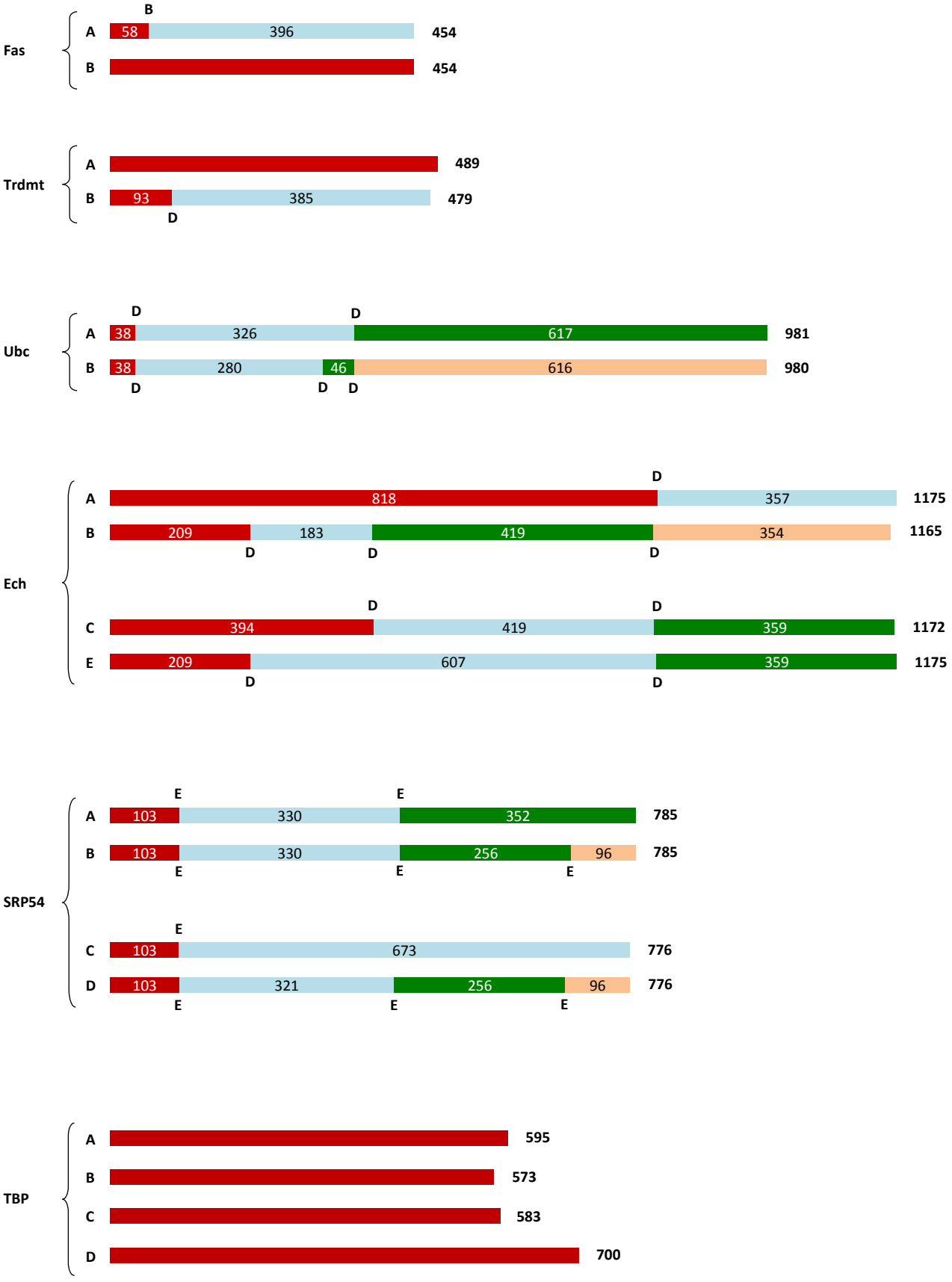

Fas
A
B
58
396
454
454
Trdmt
A
B
93
385
489
479
Ubc
A
B
38
326
617
981
38
280
46
616
980
Ech
A
B
C
E
818
357
1175
209
183
419
354
1165
394
419
359
1172
209
607
359
1175
SRP54
A
B
C
D
103
330
352
785
103
330
256
96
785
103
673
776
103
321
256
96
776
TBP
A
B
C
D
595
573
583
700

Fig S2

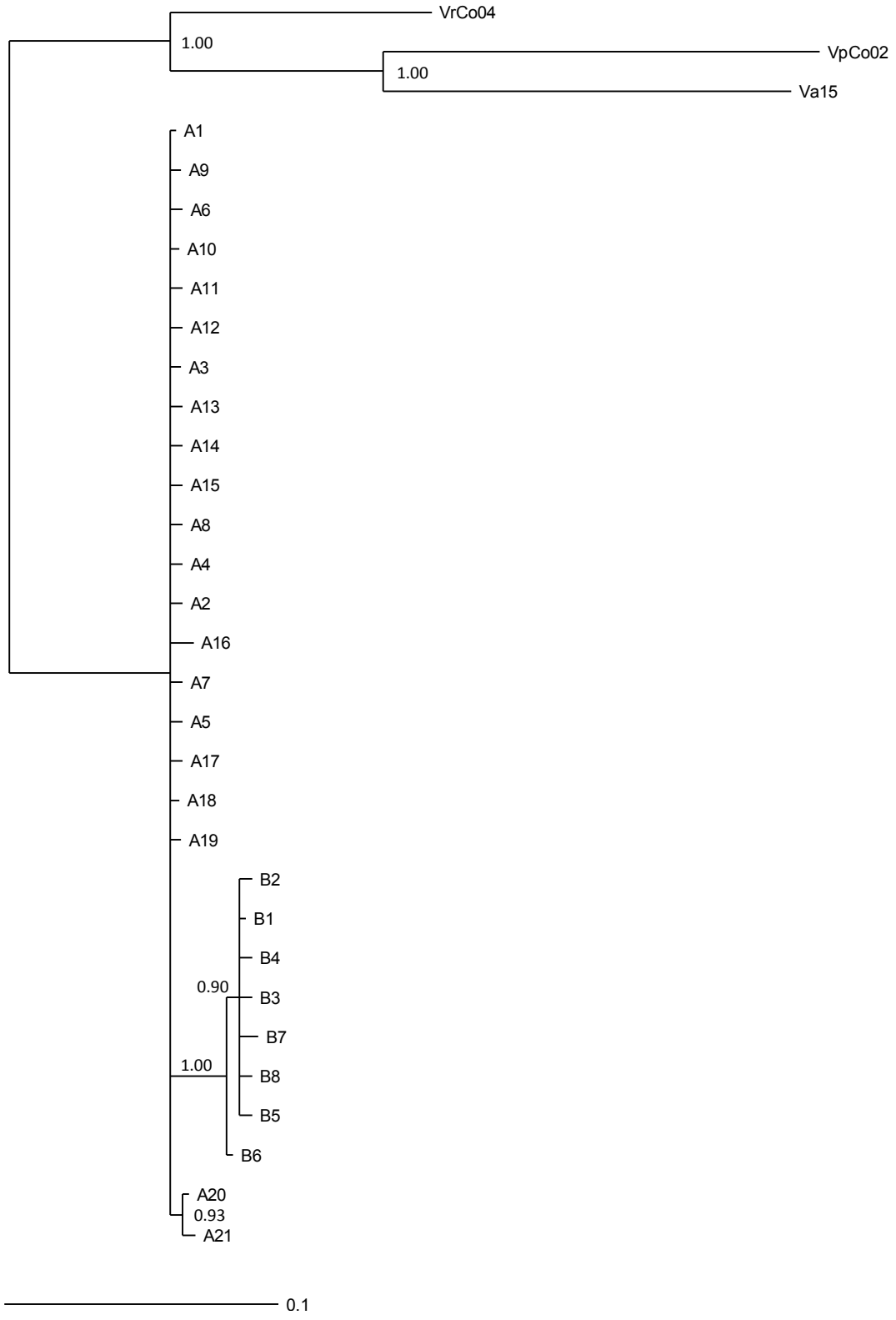

VrCo04
1.00
VpCo02
1.00
Va15
A1
A9
A6
A10
A11
A12
A3
A13
A14
A15
A8
A4
A2
A16
A7
A5
A17
A18
A19
B2
B1
B4
0.90
B3
B7
1.00
B8
B5
B6
A20
0.93
A21
0.1

Fig S3

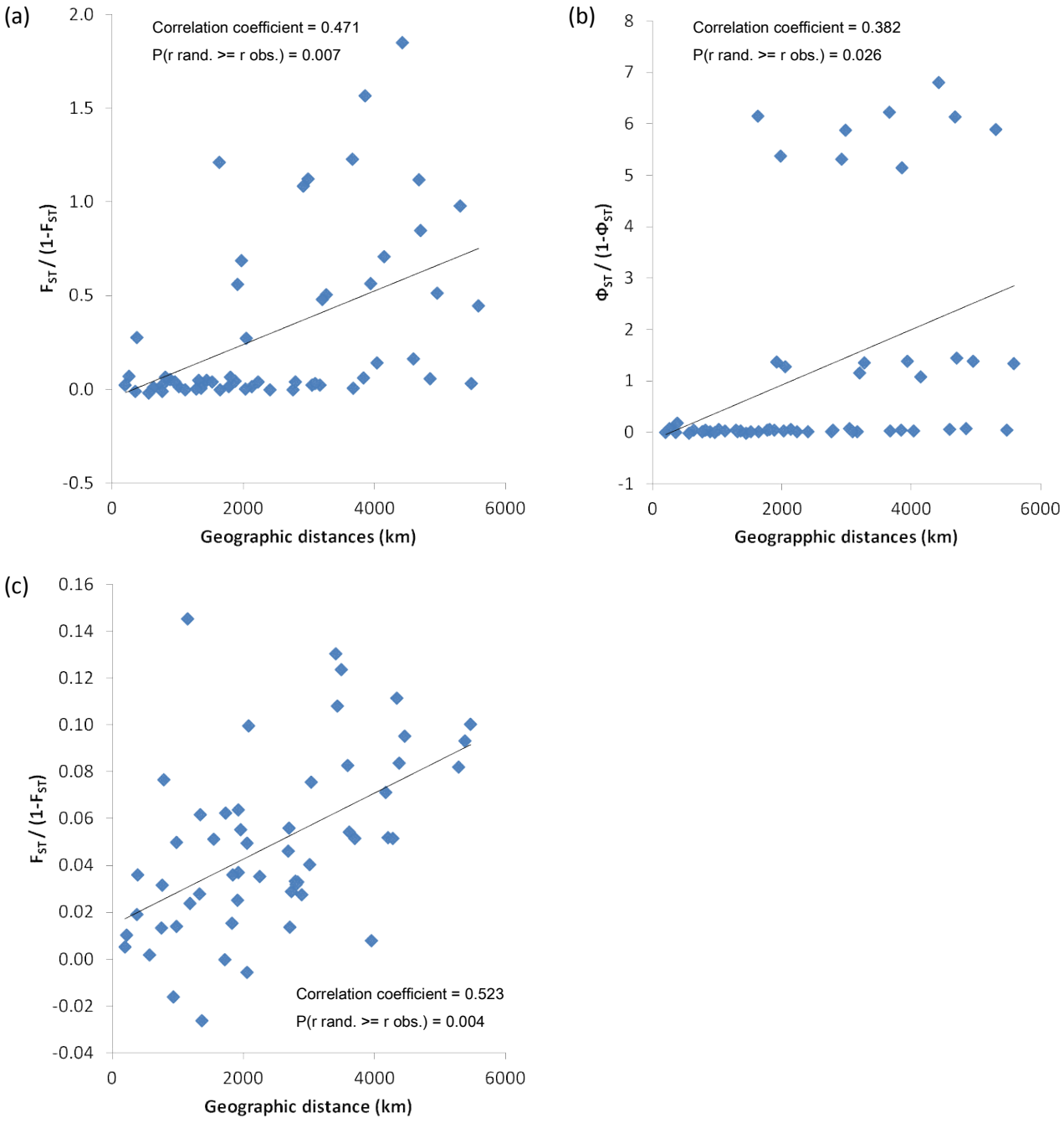
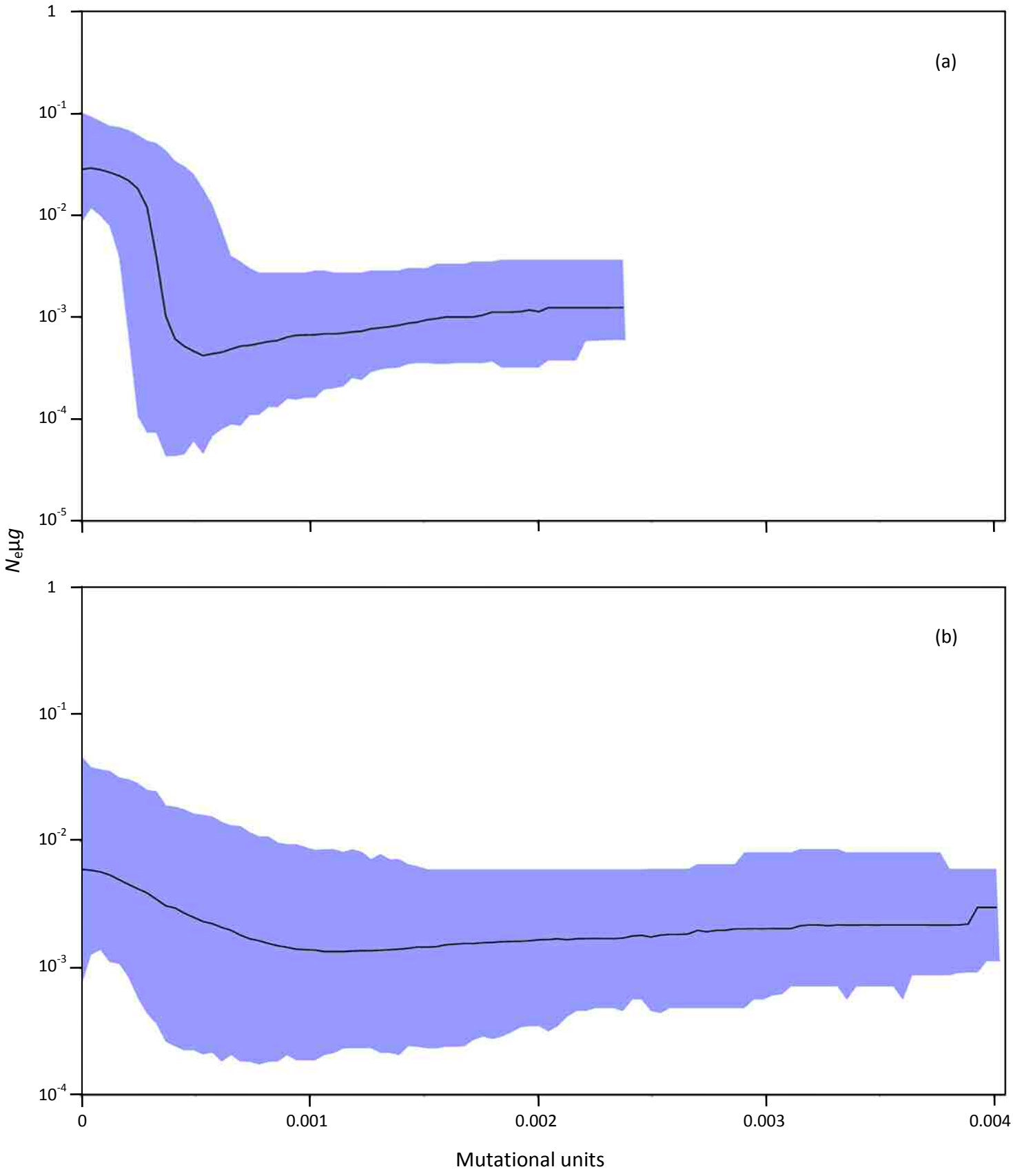

(a)
Correlation coefficient = 0.471
P(r rand. >= r obs.) = 0.007
$F_{ST} / (1-F_{ST})$
Geographic distances (km)
(b)
Correlation coefficient = 0.382
P(r rand. >= r obs.) = 0.026
$\Phi_{ST} / (1-\Phi_{ST})$
Geograpphic distances (km)
(c)
$F_{ST} / (1-F_{ST})$
Correlation coefficient = 0.523
P(r rand. >= r obs.) = 0.004
Geographic distance (km)

Fig S4


(a)
(b)
1
10-1
10-2
10-3
10-4
10-5
Neμg
0
0.001
0.002
0.003
0.004
Mutational units