



Tracing imprints of past climatic fluctuations and heterogeneous topography in *Cynips quercusfolii* (Hymenoptera: Cynipidae) in Turkey

SERDAR DİNÇ and SERAP MUTUN

Department of Biology, Faculty of Science and Art, Bolu Abant İzzet Baysal University, Bolu, 14030, Turkey;
e-mails: dinc_s@ibu.edu.tr, smutun@ibu.edu.tr

Key words. Hymenoptera, Cynipidae, *Cynips quercusfolii*, oak gall wasp, *cyt b*, ITS2, phylogeography, Pleistocene, Turkey

Abstract. *Cynips quercusfolii* is an oak gall wasp species distributed across the western Palearctic region. In this study, *C. quercusfolii* adults were used to investigate the possible effects of past climatic fluctuations and topography on the genetic structure of this species in Turkey. For this, the partial mitochondrial DNA (mtDNA) cytochrome *b* gene (*cyt b*) and nuclear DNA (nDNA) ITS2 region of 294 individuals from 38 populations covering the distribution of this species in Turkey were sequenced. The sequences generated 125 mtDNA haplotypes and 30 nDNA alleles. High variation was revealed in Turkish populations of *C. quercusfolii* ($h = 0.7328$ and $\pi = 0.0102$ for *cyt b*, and $h = 0.5865$ and $\pi = 0.0058$ for ITS2). Demographic analyses and high haplotype versus low nucleotide diversity indicated that *C. quercusfolii* populations might have undergone a series of expansions and contractions in the past. Phylogenetic analyses showed that this species is structured geographically between east and west due to a major physical barrier, the Anatolian Diagonal. Furthermore, divergence times of the lineages implied that deep splits occurred before the Pleistocene, while intermediate to shallow splits were associated with climatic changes during the Pleistocene. We conclude that ongoing geological events and associated climatic cycles were responsible for the current population genetic structure of this species in Turkey.

INTRODUCTION

Turkey is one of the most interesting places in the world due not only to the richness of species but also their high genetic diversity (Gür, 2016) and has attracted a great deal of attention because of its geographic location and complicated geological history (Okay, 2008). The Turkish landmass, composed of an Asian (Anatolia) and European part (Thrace), was formed after the collision of three main plates, and subsequent geological events resulted in the final formation of Turkey, which began in the late Tertiary (Bozkurt, 2001). Ongoing orogenesis including a series of transgressions and regressions created several mountain ranges that later acted as natural barriers to the distribution of many taxa, and thus contributed to Anatolia's richness (Okay, 2008; Şekercioğlu et al., 2011). The mountain ranges in Anatolia are the Taurus mountains located in the south, separating the Mediterranean region from the inner steppe environments, and the northern and western Anatolian mountains, separating the western Black Sea coast from central Anatolia (Demirsoy, 2002). An additional mountain range, the Anatolian Diagonal, running from northeastern Turkey to south-eastern Anatolia, is the most effective physical barrier dividing eastern Anatolia from the western part (Davis, 1971). Turkey is located at the junction of three phytogeographical regions:

the Mediterranean in the south and western part of Anatolia, the Euro-Siberian in the northern part of Turkey and the Irano-Turanian region covering eastern, south-eastern and central Anatolia (Zohary, 1973). It has been suggested that this mountain belt creates altitudinal differences on both sides that are associated with different ecological conditions, supporting distinct floristic and faunistic elements in Turkey by dividing the Irano-Turanian region into eastern and western parts (Ekim & Güner, 1986; Gür, 2016).

The active tectonic history of the Turkish landmass together with its neighbouring area was accompanied by cyclical changes in climate (Demirsoy, 2002). Recurrent events in the Late Miocene/Pliocene and the following Quaternary fluctuations greatly influenced species and the lineage distribution of many taxa in Turkey (Kaya & Çiplak, 2016). It is known that during the glacial periods species in the more northern latitudes moved to more suitable habitats in the south, and, conversely, taxa expanded their ranges towards the north when environmental conditions became more favourable (Hewitt, 2000). Likewise, species distributed in northern Europe moved to lower latitudes and survived in southern refugia (Hewitt, 2000). In addition to the Balkans, the Iberian Peninsula, Italy and Turkey acted as a refuge area in the Western Palearctic, allowing species to escape from the adverse effects of the recurrent

cycles (Hewitt, 1999; Provan & Bennet, 2008). Despite the fact that Anatolia was not covered by ice sheets, taxa distributed in Turkey also reacted to the climatic oscillations (Challis et al., 2007; Gündüz et al., 2007; Çıplak, 2008; Bellati et al., 2011). In recent years, it has been shown that at the end of the Last Glacial Maximum eastern localities, including the Iranian and Turkish populations, provided a genetic source for European populations (Dubey et al., 2007; Stone et al., 2007; Şekercioğlu et al., 2011). Consequently, high genetic diversities of Turkish populations is currently recorded for a variety of animal taxa, including oak gall wasps. Thus Anatolia is referred to as the “cradle of the oak gall wasps” (Rokas et al., 2003).

Cynips quercusfolii Linnaeus, 1758 (Hymenoptera: Cynipidae) is a heterogonous species of gall wasp with a widespread distribution across the Western Palearctic region, including Turkey (Melika, 2006). This species has a cyclically parthenogenetic lifecycle with alternation between a sexual spring generation and an asexual summer generation each year. Asexual generation galls are formed on the leaves of white oaks in the section *Quercus*. In the present study, we used sequence data of the partial mitochondrial DNA (mtDNA) cytochrome *b* gene (*cyt b*) and complete nuclear DNA (nDNA) ITS2 gene region to assess the genetic diversity of this species in Turkey. By sampling this species in a large area of Turkey and sequencing an mtDNA gene and non-coding nDNA region, we attempt to shed light on different periods of this species in the past. Within this context, we first aim to explore if there are any discontinuities in the genetic structure of this species in Turkey that can be explained by geographic barriers such as the Anatolian Diagonal as is suggested for other taxa. Secondly, we attempt to determine if past geographic and

climatic fluctuations left any imprints on the genetic structure of this species.

MATERIALS AND METHODS

Sampling, laboratory work and data preparation

Asexual generation galls that develop in early summer and mature in autumn are easily recognized and can be found throughout this species distribution. Therefore, we collected only asexual generation galls in the summers of 2010–2014 from 38 localities covering the distribution range of *C. quercusfolii* in Turkey, and a total of 294 female adults were reared from these galls. The collection sites along with their coordinates and the sample size for each population are given in Fig. 1 and Table 1, respectively. Two conspecifics, *C. divisa* Hartig (1840), and *C. longiventris* Hartig (1840) were also collected and used as outgroups in all the phylogenetic analyses.

Total genomic DNA was extracted from specimens using the DNeasy Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. CB-J-10933 (CB1) 5'-TATG-TACTACCATGAGGACAAATATC-3' and CB-N-11367 (CB2) 5'-ATTACACCTCCTAATTTATTAGGAAT-3' primers were used to amplify a 433 bp fragment of the *cyt b* gene (Simons et al., 1994). All PCR reactions were carried out following the protocols described in Stone et al. (2007). Cas5p8sFt 5'-ATGAACATCGACATTTTCGAACGCATAT-3' and Cas28sB1d 5'-TTCTTTTCTCCGCTAGTAATATGCTTAA-3' (Ji et al., 2003) primers spanning 5.8S and 28S were used for amplifying the ITS2 region using the protocols in Mutun & Atay (2015). Interestingly, the ITS2 primer pair that was successfully used in other oak gall wasp studies (Mutun & Atay, 2015; Mutun, 2016) produced background noise in the *C. quercusfolii* sequences. Therefore, we cloned PCR products (InstAclone PCR Cloning Kit, Thermo Fisher Scientific, Waltham, MA, USA, cat. K1213-14) and purified (NucleoSpin® Gel and PCR Clean-up, Macherey-Nagel, Düren, Germany, cat. 740609.50) the samples. We sequenced cloned products and designed the following internal primers: MutITSintf 5'-GTTCGTCGCGTCTCTG -3' and Mu-

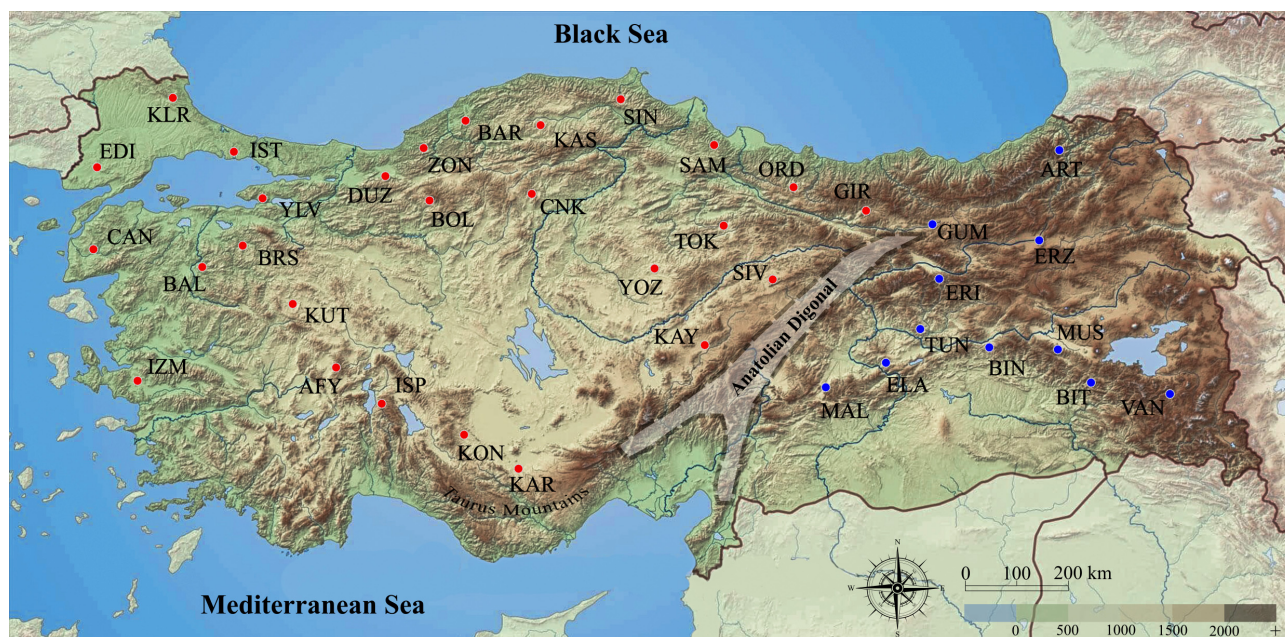


Fig. 1. Map of Turkey showing the locations of the sites where samples of *C. quercusfolii* were collected and the approximate location of the Anatolian Diagonal. Blue dots indicate populations east of the Anatolian Diagonal and red dots populations distributed west of the Diagonal. Population abbreviations are shown in Table 1.

Table 1. Names of sampling sites with abbreviations and coordinates, number of sequenced individuals and number of detected haplotypes/alleles. Haplotype (*h*) and nucleotide (π) diversity for each population are shown in the upper and the lower part of the *cyt b* and ITS2 columns.

Populations	Abb.	Coord.	Size	$N_{HAP/ALLEL}$	<i>cyt b</i> (<i>h</i> π)	ITS2 (<i>h</i> π)
AFYON	AFY	38°46.621'N 30°41.384'E	10	6/1	0.8889±0.0754 0.0229±0.0129	0.0000±0.0000 0.0000±0.0000
ARTVİN	ART	41°26.068'N 42°15.590'E	2	1/2	0.0000±0.0000 0.0000±0.0000	1.0000±0.5000 0.0132±0.0153
BALIKESİR	BAL	39°82.883'N 27°81.699'E	10	7/4	0.9333±0.0620 0.0057±0.0038	0.5333±0.1801 0.0080±0.0057
BARTIN	BAR	41°57.585'N 32°62.565'E	1	1/1	1.0000±0.0000 0.0000±0.0000	1.0000±0.0000 0.0000±0.0000
BİNGÖL	BIN	38°92.856'N 40°36.798'E	12	5/4	0.5758±0.1634 0.0088±0.0053	0.4545±0.1701 0.0028±0.0026
BİTLİS	BIT	38°43.688'N 42°55.733'E	10	7/1	0.9111±0.0773 0.0114±0.0068	0.0000±0.0000 0.0000±0.0000
BOLU	BOL	40°99.628'N 32°06.246'E	10	5/2	0.6667±0.1633 0.0031±0.0023	0.2000±0.1541 0.0017±0.0020
BURSA	BRS	40°54.139'N 29°27.262'E	9	8/3	0.9722±0.0640 0.0193±0.0112	0.4167±0.1907 0.0057±0.0044
ÇANAKKALE	CAN	39°92.080'N 27°27.564'E	10	4/5	0.7778±0.0907 0.0023±0.0019	0.8000±0.1001 0.0211±0.0127
ÇANKIRI	CNK	40°71.351'N 33°01.644'E	10	7/7	0.9111±0.0773 0.0050±0.0034	0.9111±0.0773 0.0170±0.0105
DÜZCE	DUZ	40°54.330'N 31°10.246'E	2	1/2	0.0000±0.0000 0.0000±0.0000	1.0000±0.5000 0.0398±0.0419
EDİRNE	EDI	41°27.295'N 26°78.507'E	7	1/4	0.0000±0.0000 0.0000±0.0000	0.7143±0.1809 0.0046±0.0039
ELAĞİ	ELA	38°57.990'N 38°86.239'E	5	3/2	0.7000±0.2184 0.0194±0.0126	0.4000±0.2373 0.0035±0.0035
ERZİNCAN	ERI	39°83.621'N 38°75.296'E	10	5/3	0.8000±0.1001 0.0072±0.0046	0.3778±0.1813 0.0016±0.0019
ERZURUM	ERZ	39°86.495'N 40°62.626'E	4	3/2	0.8333±0.2224 0.0234±0.0162	0.6667±0.2041 0.0029±0.0033
GİRESUN	GIR	40°26.610'N 38°94.630'E	10	9/4	0.9778±0.0540 0.0180±0.0103	0.7111±0.1175 0.0037±0.0032
GÜMÜŞHANE	GUM	40°26.209'N 39°46.407'E	11	8/3	0.8909±0.0918 0.0083±0.0051	0.5636±0.1340 0.0024±0.0024
İSPARTA	ISP	37°73.587'N 30°83.298'E	10	2/3	0.2000±0.1541 0.0023±0.0019	0.7111±0.0860 0.0031±0.0029
İSTANBUL	IST	41°19.981'N 28°94.030'E	10	9/4	0.9778±0.0540 0.0124±0.0074	0.6444±0.1518 0.0148±0.0093
İZMİR	IZM	39°25.547'N 27°06.827'E	10	3/5	0.6222±0.1383 0.0172±0.0099	0.6667±0.1633 0.0169±0.0105
KARAMAN	KAR	37°39.117'N 33°16.049'E	10	7/3	0.9111±0.0773 0.0267±0.0149	0.5111±0.1643 0.0022±0.0023
KASTAMONU	KAS	41°27.204'N 33°15.061'E	11	4/1	0.6909±0.1276 0.0047±0.0032	0.0000±0.0000 0.0000±0.0000
KAYSERİ	KAY	38°37.063'N 36°23.143'E	11	5/3	0.7091±0.1366 0.0246±0.0137	0.5818±0.1420 0.0025±0.0025
KIRKLARELİ	KLR	41°85.992'N 27°81.336'E	10	8/4	0.9333±0.0773 0.0064±0.0041	0.6444±0.1518 0.0082±0.0058
KONYA	KON	36°98.244'N 32°40.465'E	10	7/3	0.9111±0.0773 0.0275±0.0153	0.7333±0.0764 0.0032±0.0029
KÜTAHYA	KUT	39°45.540'N 30°25.914'E	10	8/3	0.9556±0.0594 0.0076±0.0048	0.6444±0.1012 0.0076±0.0054
MALATYA	MAL	38°99.107'N 37°81.619'E	3	2/1	0.6667±0.3143 0.0277±0.0216	0.0000±0.0000 0.0000±0.0000
MUŞ	MUS	38°64.304'N 41°75.369'E	11	5/5	0.6182±0.1643 0.0051±0.0034	0.7636±0.1066 0.0055±0.0042
ORDU	ORD	40°45.983'N 37°38.665'E	2	2/2	1.0000±0.5000 0.0230±0.0242	1.0000±0.5000 0.0044±0.0062
SAMSUN	SAM	41°00.933'N 35°87.766'E	11	6/3	0.8364±0.0887 0.0123±0.0072	0.6182±0.1038 0.0027±0.0026
SİNOP	SIN	41°80.846'N 34°63.325'E	10	5/3	0.8000±0.1001 0.0084±0.0052	0.3778±0.1813 0.0017±0.0020
SİVAS	SIV	39°46.451'N 38°00.067'E	6	6/3	1.0000±0.0962 0.0101±0.0067	0.6000±0.2152 0.0029±0.0029
TOKAT	TOK	40°48.323'N 36°95.483'E	1	1/1	1.0000±0.0000 0.0000±0.0000	1.0000±0.0000 0.0000±0.0000
TUNCELİ	TUN	39°59.013'N 39°86.258'E	1	1/1	1.0000±0.0000 0.0180±0.0130	1.0000±0.0000 0.0000±0.0000
VAN	VAN	38°08.991'N 43°09.806'E	10	4/2	0.6444±0.1518 0.0072±0.0046	0.2000±0.1541 0.0008±0.0013
YALOVA	YLV	40°54.139'N 29°27.262'E	10	4/6	0.5333±0.1801 0.0023±0.0019	0.8444±0.1029 0.0187±0.0115
YOZGAT	YOZ	40°55.775'N 29°28.272'E	1	1/1	1.0000±0.0000 0.0000±0.0000	1.0000±0.0000 0.0000±0.0000
ZONGUDAK	ZON	41°42.168'N 32°17.702'E	3	1/1	0.0000±0.0000 0.0000±0.0000	0.0000±0.0000 0.0000±0.0000

tITSintr 5'-CCGTCCATAATGGCCAC-3'. This primer pair was later used specifically to amplify the *C. quercusfolii* ITS2 region successfully. All PCR reactions were carried out in 35 µl volumes containing 8.75 µl of 5X buffer, 1.75 µl of MgCl₂ (25 mM), 0.7 µl of dNTPs (10 mM), 0.35 µl of template DNA, 0.35 µl of *Taq* polymerase (Vivantis, Subang Jaya, Malaysia) and 0.44 µl of each primer (20 pmol). Temperature rates, times and cycles were set as follows: 3 min at 94°C for the first denaturation, 28 cycles of 30 s at 94°C for denaturation, 30 s at 52°C for annealing, 30 s at 72°C for elongation and 5 min at 72°C for final extension. Amplified products of both mtDNA and nDNA regions were sent to a company (MACROGEN, South Korea), which supplied the sequences in the form of chromatograms.

All *cyt b* and ITS2 sequences were checked, aligned, and trimmed using the program BioEdit 7.2.5 (Hall, 1999). Similar sequences were collapsed into *cyt b* haplotypes and ITS2 alleles using GenAIEx 6.5 (Peakall & Smouse, 2006, 2012). Mitochondrial DNA sequences were analysed to ensure that the sequences were genuine and not numts (pseudogenes). Our tests showed that there were neither stop codons nor non-sense mutations after translating the *cyt b* sequences into amino acids using the program DnaSP 5.10.1 (Librado & Rozas, 2009). All haplotypes/alleles were deposited in GenBank (Accession number: MG650475-MG650599 and MG661529-MG661558 for the *cyt b* haplotypes and the ITS2 alleles, respectively).

Genetic diversity estimation and population structure

The number of polymorphic sites (*S*), nucleotide (π) and haplotype/allele (*h*) diversity (Nei, 1987), the number of substitutions and pairwise nucleotide differences (*k*) (Tajima, 1983), and population differences were calculated using the programs Arlequin 3.5.2.2 (Excoffier & Lischer, 2010) and DnaSP 5.10.1 (Librado & Rozas, 2009). For estimating the genetic diversity of *C. quercusfolii*, only those localities with sample sizes larger than two were considered for further evaluation. To test if there was a correlation between the geographic distribution of haplotypes/alleles and genetic diversity, we performed AMOVA using Arlequin 3.5.2.2 (Excoffier & Lischer, 2010) at three different hierarchical levels: among groups, among populations within groups and within each population. We grouped the populations into eastern versus western groups relative to the location of the Anatolian Diagonal as shown in Fig. 1. Furthermore, we conducted a spatial autocorrelation analysis using the program GenAIEx 6.5 (Peakall & Smouse, 2006, 2012) for both the *cyt b* gene and ITS2 region for distance classes of 100 and 150 km with 95% confidence intervals. Spatial autocorrelation analysis is a useful technique for detecting the relationship between population pairwise genetic similarity and geographic distances across the study area (Sokal et al., 1997) and has previously been applied to species of oak gall wasps, including *C. quercusfolii* (Stone & Sunnucks, 1993; Stone et al., 2002). The analysis produces a correlogram with a spatial autocorrelation coefficient (*r*), where the value of *r* is significantly greater than zero for geographically close samples and decreases with increasing geographic distance between samples isolated by distance. The distance between samples at which *r* is no longer significantly different from zero is an approximation of the distance at which the population structure can be revealed. The significance of *r* was tested by bootstrapping 95% confidence intervals (CIs) for *r* by resampling 1000 times (Peakall et al., 2003).

Inferences about the demography of the populations

Analysing population demographic structure may allow us to unravel signatures of historical events (Avice, 2000). Therefore, we applied mismatch distribution analysis using pairwise differ-

ences between haplotypes/alleles using DnaSP 5.10.1 (Librado & Rozas, 2009). In mismatch distribution analysis, expanding populations have a unimodal profile, whereas populations at demographic equilibrium have a bimodal or multimodal profile (Rogers & Harpending, 1992). We also calculated the raggedness index (*H_Rag*) (Harpending, 1994) and the sum of squared deviations (SDD) (Schneider & Excoffier, 1999) to compare the observed sample distribution with the expected probability distribution using a model of population expansion. Moreover, we used Tajima's *D* (Tajima, 1989) and Fu's *F_s* (Fu, 1997) tests implemented in Arlequin 3.5.2.2 (Excoffier & Lischer, 2010) to check for deviations from neutrality. While Tajima's *D* (Tajima, 1989) uses the frequency information for segregating sites, Fu's *F_s* uses the distribution of haplotypes/alleles (Fu, 1997). We conducted all analyses using different grouping schemes: all populations (no grouping), all populations divided into eastern versus western groups, and each population individually.

Phylogenetic and phylogeographic analyses

Maximum parsimony (MP), maximum likelihood (ML) and Bayesian analyses were used to reveal phylogenetic relationships among haplotypes/alleles using PAUP*4.0b10 (Swofford, 2002), and MrBayes 3.2.6 (Ronquist & Huelsenbeck, 2003). We performed MP analysis using the TBR branch-swapping algorithm with 1000 replicates of random addition of taxa under the heuristic search options and 1000 bootstrap replicates employed to assess support for branches (Swofford, 2002). In reconstructing the ML tree, the model best fitting our datasets was determined using JModeltest 2 (Guindon & Gascuel, 2003; Darriba et al., 2012) and the Akaike information criterion (AIC) (Akaike, 1974; Posada & Buckley, 2004). HKY + I + Γ ($-\ln = 1959.4363$) and GTR + I ($-\ln = 497.2462$) were the best fitting models for *cyt b* and the ITS2 region, respectively. BEAST version 1.5.2 was used to estimate divergence times of *cyt b* lineages and their confidence intervals using a Bayesian Markov Chain Monte Carlo (MCMC) coalescent method (Drummond et al., 2012) by applying the HKY model with I and Γ following the uncorrelated relaxed lognormal clock (Drummond & Suchard, 2010). A 1.19% sequence divergence per lineage (2.39% pairwise) per million years mutation rate was used for age calibration (Papadopoulou et al., 2010). Both the most recent common ancestors (MRCAs) and most ancient common ancestors (MACAs) calculation and operator optimizations were (Hayward & Stone, 2006) performed with BEAUTI v1.8.0. The BEAST analysis was run for 100 million generations sampling every 1000 and the convergence to stationary and the effective sample size (ESS) of the model parameters were checked using TRACER v1.6.0. The maximum clade credibility tree was built with TREEANNOTATOR v1.8.4, discarding the initial 25% of samples as burn-in. FIG-TREE version 1.3.1 (Rambaut, 2009) was used to visualize the results. An unrooted 95% probability level network analysis was constructed using the program HapStar version 0.5 (C) (Teacher & Griffiths, 2011). The pairwise distance matrix between haplotypes/alleles computed by Arlequin was used for building a minimum spanning tree. The resulting svg file was transferred to Inkscape 0.91 (www.inkscape.org) to visualize the network.

RESULTS

Genetic diversity

All sequences of the *cyt b* gene produced 125 haplotypes. The haplotypes along with their frequencies are shown in Table S1. Nucleotide frequencies were 32.93%, 43.26%, 13.01% and 10.80% for A, T, C and G, respectively. The transition/transversion rate ratios were $k1 =$

15.811 (purines) and $k2 = 27.729$ (pyrimidines), and the overall transition/transversion ratio bias was $R = 8.629$. Eighteen amino acids (12.5%) showed substitutions when all *cyt b* sequences were translated into amino acids. Most of the amino acid replacements were observed at the third codon position (37 substitutions, 39.7%), followed by the first codon position (15 substitutions, 16.1%) and then the second (3 substitutions, 3.2%). Pairwise differences between haplotypes ranged from 0.2% (1 bp) to 7.5% (30 bp). Most *C. quercusfolii* populations showed a high level of genetic diversity, where the average haplotype (*h*) and nucleotide diversity (π) for the *cyt b* gene were $h = 0.7328 \pm 0.0955$, and $\pi = 0.0102 \pm 0.0068$, respectively (Table 1). On the other hand, the ITS2 region produced 30 distinct sequences; all alleles and their frequencies are given in Table S2. Nucleotide frequencies were 28.61%, 30.75%, 18.14% and 22.49% for A, T, C and G, respectively. The transition/transversion bias was $R = 1.461$. Pairwise comparisons of alleles ranged between 0.4% (1 bp) and 4.7% (10 bp). The average allelic diversity (*h*) of *C. quercusfolii* was $h = 0.5865 \pm 0.1402$ and the nucleotide diversity was $\pi = 0.0058 \pm 0.0088$.

Geographic distribution of genetic diversity

The geographic distribution of genetic diversity across the sampling area of *C. quercusfolii* both for the *cyt b* gene and the ITS2 region was investigated using AMOVA with a statistical significance of 20,000 permutations. The Bonferroni method was used for correcting the *P* values in multiple comparisons. In the *cyt b* dataset, we detected the highest level of genetic variation at the within-population level (38.01%), followed by among populations within groups level (35.75%) (Table 2A). However, 26.24% of the variation was shared among groups in the scheme of eastern (ART, BIN, BIT, ELA, ERI, ERZ, GUM, MAL, MUS, TUN and VAN) versus western populations (all the remaining localities). The genetic differentiation between the groups was $F_{ST} = 0.61$, $P < 0.001$. On the other hand, the ITS2 dataset indicated that 71.21% of the genetic variation was within populations and 15.05% was shared between populations. Among the groups, variation was lowest (13.75%) when the same grouping scheme was used (Table 2B). The genetic differentiation between the two groups was relatively low ($F_{ST} = 0.28$, $P < 0.001$).

Table 2. AMOVA analysis of *C. quercusfolii*. A – analysis of the *cyt b* data; grouping includes Isparta and all other populations. B – analysis of the ITS2 data; each population was treated as a single group.

Source variation	d.f.	Sum of squares	Variance components	Percentage of variation
A)				
Among groups	1	254.541	1.65249	26.24
Among pops.	36	705.589	2.25169	35.75
Within pops.	256	612.778	2.39367	38.01
Total	293	1532.908	6.29784	
B)				
Among groups	1	16.910	0.12905	13.75
Among pops.	36	62.917	0.14124	15.05
Within pops.	256	171.123	0.66845	71.21
Total	293	250.949	0.93873	

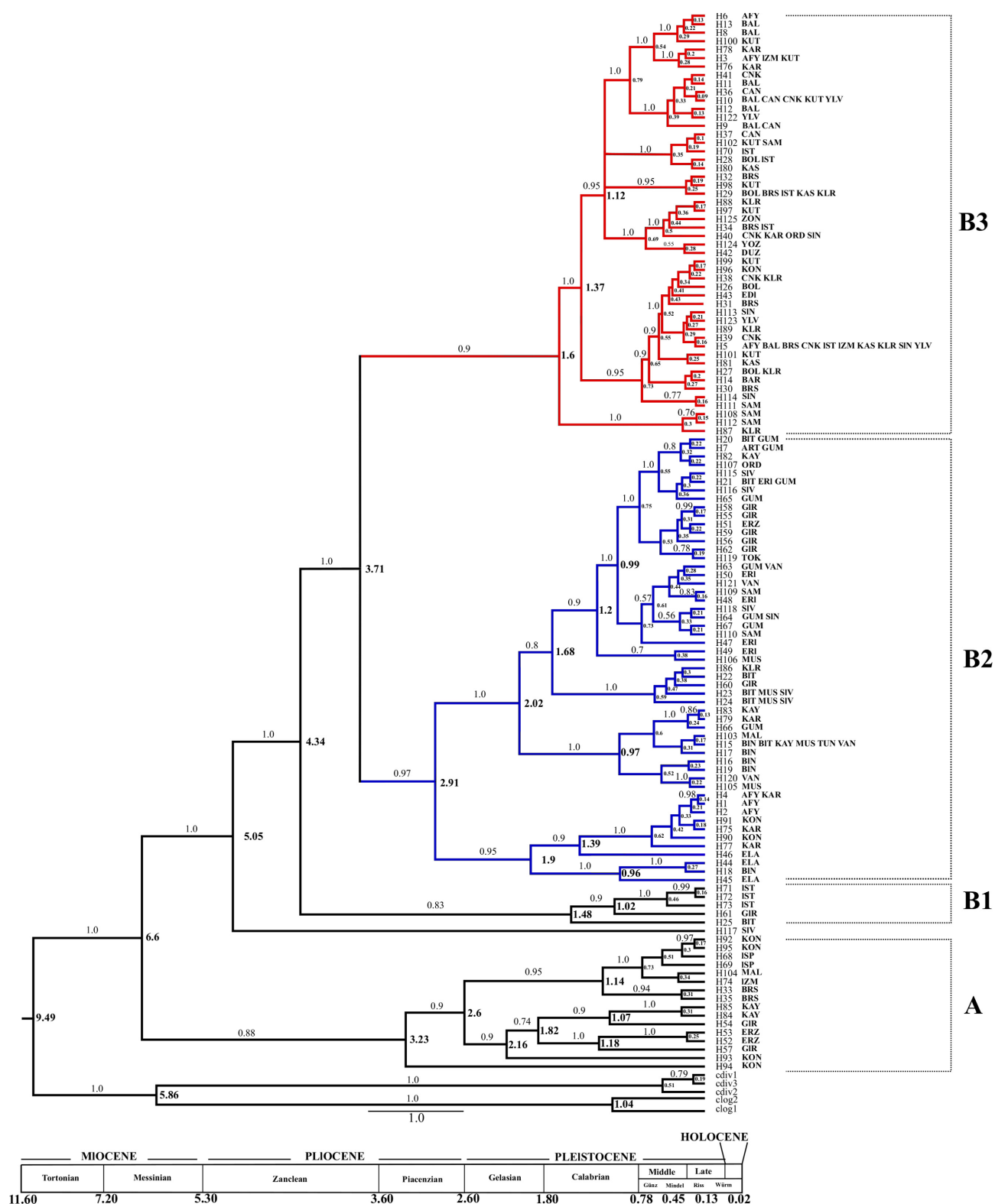


Fig. 2. The BEAST tree of *cyt b* haplotypes. Node ages (tMRCA) are shown in the interior part of each related node and posterior probability values are also on the branches.

Spatial genetic autocorrelation analyses of both *cyt b* and ITS2 datasets conducted across the study area for distance classes of 100 and 150 km with 95% confidence intervals are shown in Figs S1 and S2. Distance class correlograms for both regions showed positive and significant values ($P < 0.001$), but significantly negative values were obtained

for distances beyond the x-axis intercept 523.4 km for *cyt b* and 555.7 km for the ITS2 data. These results indicate that Turkish populations of *C. quercusfolii* geographically separated by a distance of less than 100 km are genetically similar to each other, which is higher than expected by chance, while the genetic differentiation between popula-

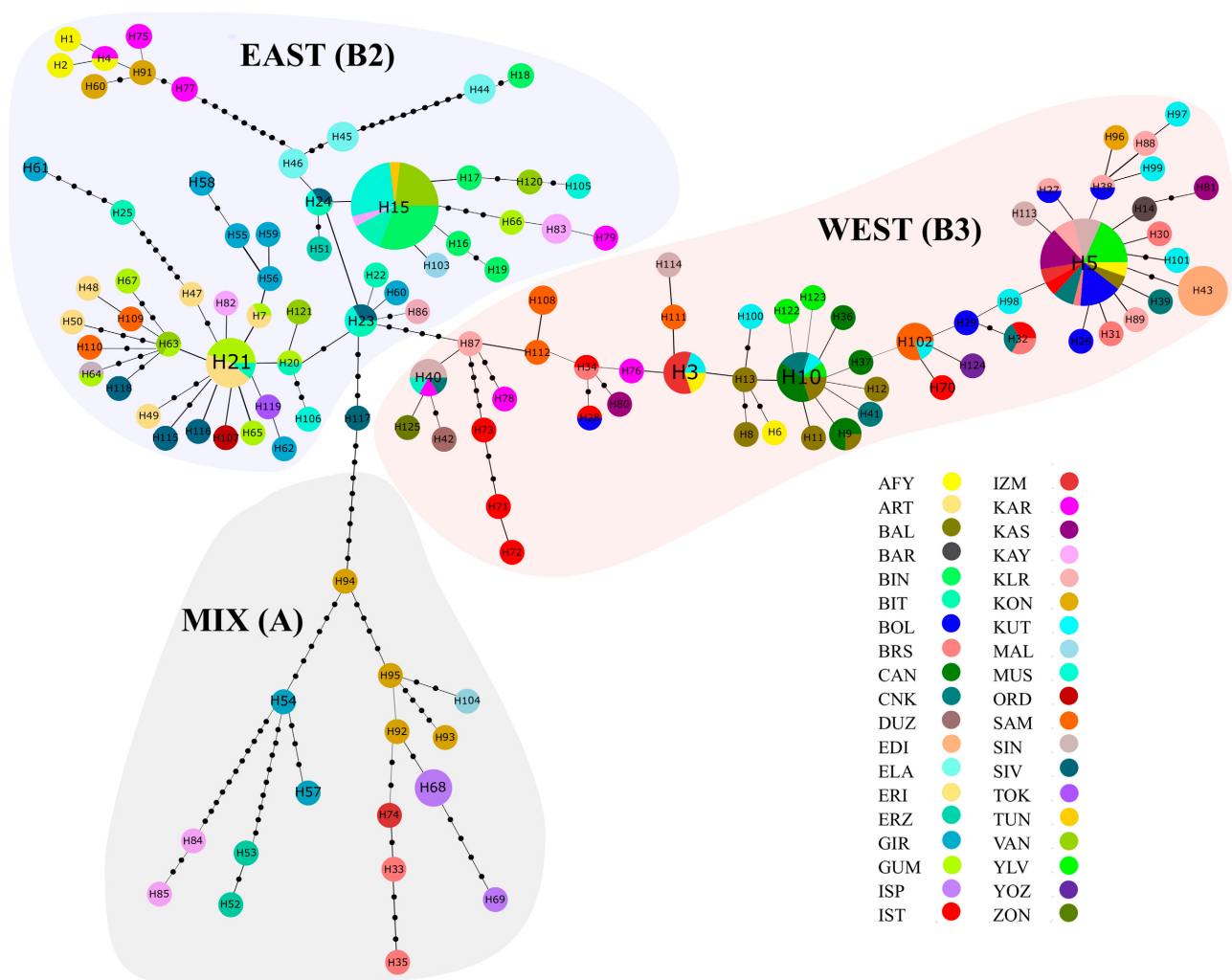


Fig. 3. Results of the network analysis of the *cyt b* haplotypes using Hapstar 0.7. Colour coded populations are shown in the circles. Size of the circles is proportional to the frequency of the haplotype.

tions separated by more than ca. 500 km is higher than expected by chance.

Phylogenetic analyses and estimation of lineage divergence

In the *cyt b* haplotypes, there were 53 parsimony informative characters, while the remaining characters were uninformative. Three different approaches were employed to analyse the evolutionary relationships among *C. quercus-folii* *cyt b* haplotypes and ITS2 alleles. Since all analyses produced mostly similar tree topologies with some differences in bootstrap/posterior probability values for the *cyt b* gene dataset, here we present only the BEAST tree in Fig. 2. Except for the presence of a small separate basal clade (the B1 subclade) composed of five haplotypes in the BI tree, all the trees produced two main clades: the first clade (Clade A), composed of a mixture of haplotypes representing eastern, central, northern and western Turkey, and the second clade (Clade B), of which H117 sampled from a central population (Sivas) was the most basal haplotype. Clade B was bifurcated into two subclades: B1 and B2+B3. Subclade B1 has an eastern basal haplotype (H25) sampled from the Bitlis population and four other

haplotypes representing northern and western localities. Subclade B2 is dominated mainly by haplotypes from eastern populations and it is more structured with deeper to intermediate-shallow splits. However, subclade B3 is dominated by haplotypes from the western and north-western localities with fewer intermediate and more shallow splits.

Application of the molecular clock to the *cyt b* dataset indicate that the ingroup haplotypes diverged from the outgroup haplotypes near 9.49 MYA (95% HPD interval 3.3886–11.9024). After diversification of *C. quercus-folii* from the other species of *Cynips*, its haplotypes were divided into two sister clades (Clade A + B) around 6.6 MYA. Subsequent to the first divergence of lineages, a series of splitting events in Clade A took place between 3.23 MYA and 0.17 MYA. On the other hand, Clade B, comprising H117 from Sivas at the basal part, diverged from the rest of the haplotypes around 5.05 MYA. The separation of the haplotypes grouped under Clade B into B1 and B2+B3 seems to have occurred around 4.34 MYA. Subclade B2, which is dominated by haplotypes from eastern populations, and B3, composed of haplotypes from western populations, diverged from each other around 3.71 MYA. Ongoing diversification events resulted in more substructure

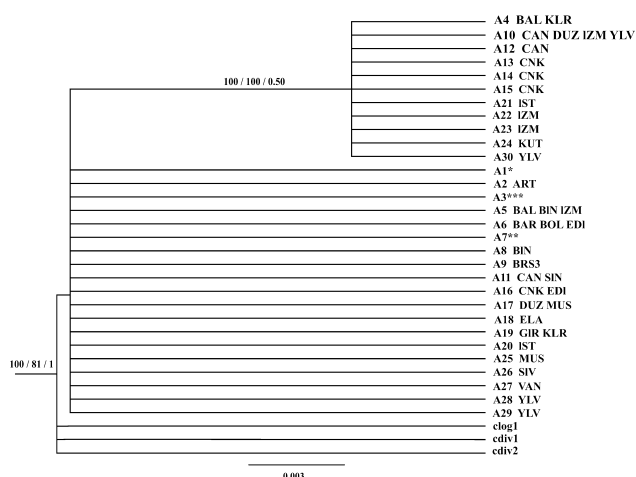


Fig. 4. Bayesian consensus tree of ITS2 alleles with branch posterior probabilities shown in the interior part of each related branch. Values on the branches indicate bootstrap values for the MP/ML and the posterior probability values, respectively.

tures that seemed to start around 2.91 MYA and continued through 0.13 MYA in subclade B2, and at 1.6 MYA to 0.09 MYA in subclade B3. Further splitting events continued with more recent diversifications between the lineages of *C. quercusfolii*.

Network analysis conducted for resolving possible reticulations and evolutionary relationships of *C. quercusfolii* haplotypes (Fig. 3) supported in most part the general pattern observed in the phylogenetic trees. The star-shaped haplotype group shown by A in the network was concordant with Clade A in the phylogenetic trees. The other haplotype group that formed B1 in the BEAST tree was found scattered between subclades B2 and B3 in this network. The rest of the haplotypes formed two haplogroups: B2

and B3. The B2 haplogroup was dominated by haplotypes from the eastern populations, while the B3 haplogroup comprised haplotypes mostly from western, central and north-western Turkey.

In the ITS2 dataset, out of 395 characters, ten parsimony informative sites and one uninformative site were detected. Separate phylogenetic analyses of the ITS2 alleles yielded mostly polytomous trees with a low resolution (Fig. 4). Nineteen alleles representing almost all sampled populations across Turkey formed a polytomy. The existence of a relatively small clade that consisted of alleles only from the western part of Turkey was also apparent in the tree. In addition, the ITS2 network analysis revealed several star-shaped allele groups indicating evolutionary connections of alleles located at the periphery to the central allele through a mutational step or a hypothetical allele (Fig. 5). The most common allele, A1, shared among 27 populations, was at the central part of the first star-shaped group and connected to two other frequent and common alleles (A3 and A7). Another small allele group with A6 (found in Bartın, Bolu, and Edirne) at its centre was connected to A23 and other alleles, all of which were parts of the small allele group in the phylogenetic tree.

Population demography

The mismatch distribution including all samples generated a nearly unimodal profile with $HRag = 0.16676$ for the *cyt b* dataset, but it was not significantly supported (Fig. 6A). Mismatch distribution analyses conducted on western versus eastern populations generated a multimodal profile for both east and west groups with a non-significant $HRag$ value for the *cyt b* dataset (Fig. 6B–C). Individually analysed populations indicated that the Balıkesir, Çankırı, Kütahya and Sivas populations showed either unimodal

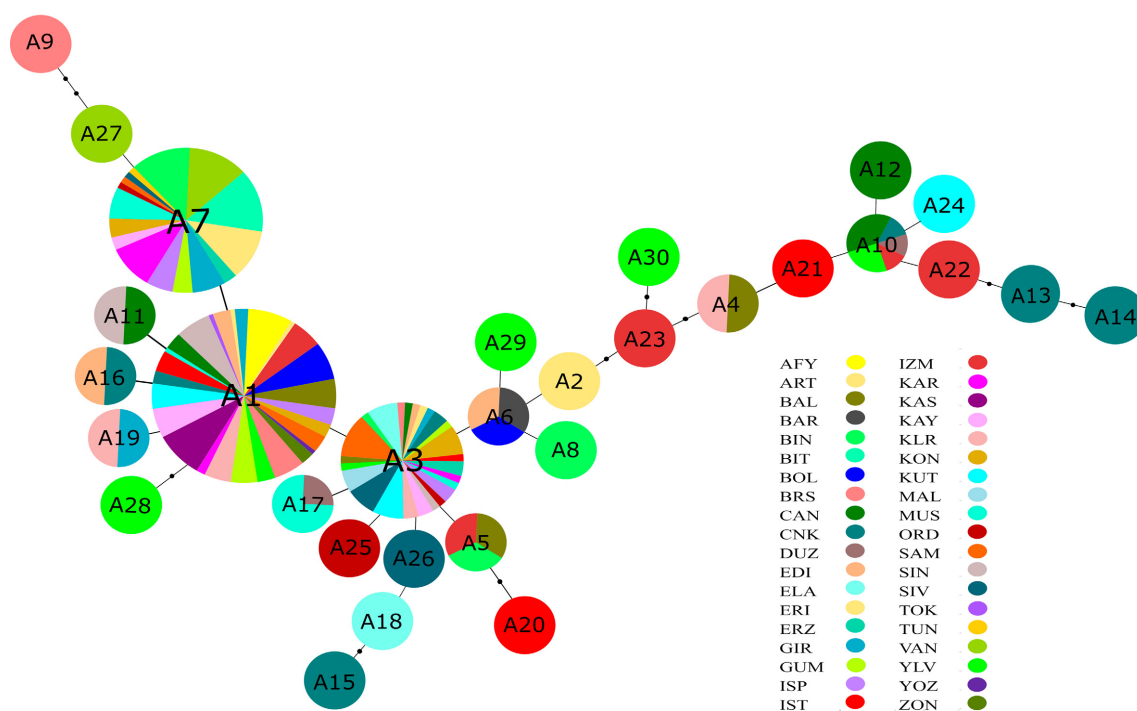


Fig. 5. Results of the network analysis of the nuclear ITS2 allele using Hapstar 0.7. Colour-coded populations are shown in the circles. Circle sizes are proportional to the frequency of the allele.

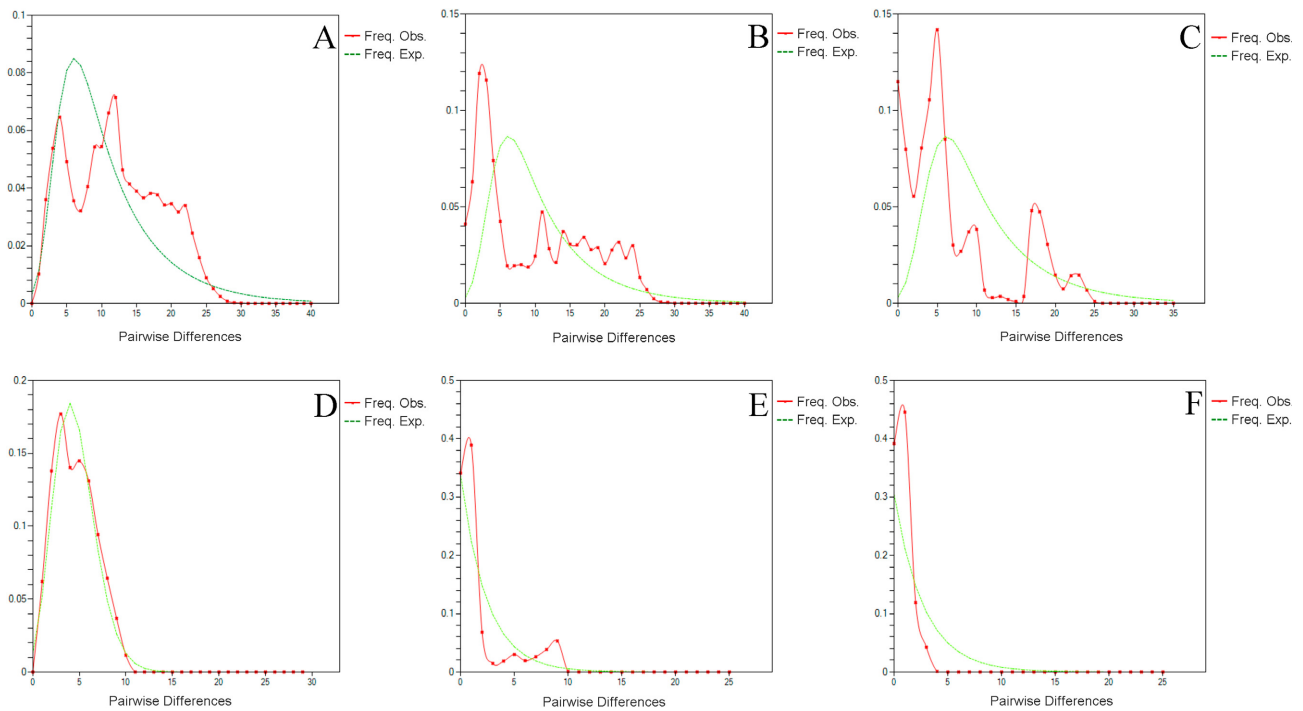


Fig. 6. Mismatch distribution of all pairwise combinations for *cyt b* haplotypes and A – all populations as a single group, B – western group, C – eastern group, and for the ITS2 alleles, D – all populations as a single group, E – western group and F – eastern group. The observed distribution is represented by a red line, and the expected frequencies by a green dotted line.

Table 3. Fu's F_s and Tajima's D , H_{rag} and SSD values calculated for each population. The values above the rows refer to the *cyt b* and the below are for the ITS2 region, respectively. (* $P \leq 0.05$, ** $P \leq 0.01$).

Pop.	Fu's F_s	Tajima's D	H_{rag}	SSD	Pop.	Fu's F_s	Tajima's D	H_{rag}	SSD
AFY	2.1165	2.2443	0.1822	0.1010	IZM	6.6069	0.0276	0.4429	0.5372
	0.0000	0.0000	0.0000	0.0000		0.8939	0.8820	0.2000	0.0821
ART	0.0000	0.0000	0.0000	0.0000	KAR	1.2189	1.0182	0.1175	0.0709
	1.0986	0.0000	0.0000	0.0000		-0.7393	1.1417	0.2617	0.0245
BAL	-2.5895	-0.5171	0.0479	0.0029	KAS	0.8098	-0.6003	0.1616	0.0622
	0.3876	-1.1163	0.1506	0.3654		0.0000	0.0000	0.0000	0.0000
BAR	0.0000	0.0000	0.0000	0.0000	KAY	4.4619	0.1864	0.0360	0.1380
	0.0000	0.0000	0.0000	0.0000		-0.4224	1.6998	0.3652	0.0466
BIN	1.4041	-1.7205*	0.1258	0.4260	KLR	-3.8244**	-1.2789	0.1081	0.0293
	-1.5898*	-1.1790	0.8830	0.0015		0.4402	-1.4664	0.1530	0.0407
BIT	-0.7583	0.7599	0.0849	0.3704	KON	1.2867	0.2552	0.0661	0.0539
	0.0000	0.0000	0.0000	0.0000		-0.1005	2.7513	0.7555	0.1387
BOL	-1.5071	-1.4928	0.1397	0.0329	KUT	-3.2237	0.1785	0.1116	0.0467
	0.5862	-1.4008	0.7200	0.0514		1.6373	-1.2685	0.2261	0.0577
BRS	-1.4237	-0.2483	0.0470	0.0448	MAL	4.4998	0.0000	1.0000	0.4645
	0.8670	-1.2298	0.2878	0.2608		0.0000	0.0000	0.0000	0.0000
CAN	-0.8305	-0.1297	0.2400	0.0347	MUS	-0.1552	-1.4911	1.0000	0.4542
	1.4503	2.1972	0.1585	0.1210		-1.4990	0.7695	0.2899	0.0554
CNK	-2.9681**	-0.9579	0.0622	0.0053	ORD	2.3025	0.0000	0.0000	0.0000
	-1.3873	0.9126	0.0483	0.0177		0.0000	0.0000	0.0000	0.0000
DUZ	0.0000	0.0000	0.0000	0.0000	SAM	0.8918	0.1940	0.1028	0.0671
	2.1972	0.0000	0.0000	0.0000		-0.3140	1.9569	0.4380	0.0627
EDI	0.0000	0.0000	0.0000	0.0000	SIN	0.7939	-1.4208	0.0893	0.0280
	-1.3898	-0.6540	0.0952	0.0067		-1.1639	-1.4008	0.1827	0.0057
ELA	3.4053	1.2185	0.4700	0.2082	SIV	-2.4199	0.0280	0.1511	0.0719
	1.0404	-0.9725	0.6800	0.3199		-0.8584	-1.1319	0.2400	0.0373
ERI	0.4232	-0.8694	0.2296	0.0753	TOK	0.0000	0.0000	0.0000	0.0000
	-1.2626**	0.1759	0.2024	0.0054		0.0000	0.0000	0.0000	0.0000
ERZ	2.6519	-0.6976	0.7500	1.0555	TUN	0.0000	0.0000	0.0000	0.0000
	0.5400	1.6329	0.5555	0.0898		0.0000	0.0000	0.0000	0.0000
GIR	-2.2481	-0.3703	0.0301	0.0286	VAN	1.6772	0.4999	0.2380	0.1161
	-1.2304	0.6273	0.2987	0.0512		-0.3393	-1.1117	0.4000	0.3310
GUM	-2.4089	-1.5198	0.1104	0.0373	YLV	-0.8305	-0.1297	0.0962	0.3802
	-0.4788	1.5712	0.3338	0.0399		0.0035	1.4631	0.0696	0.0219
ISP	2.1972	-1.7411*	0.7200	0.0577	YOZ	0.0000	0.0000	0.0000	0.0000
	-0.1561	2.5903	0.6839	0.1208		0.0000	0.0000	0.0000	0.0000
IST	-3.3361	0.4175	0.0518	0.1402	ZON	0.0000	0.0000	0.0000	0.0000
	1.8374	0.8050	0.2459	0.1134		0.0000	0.0000	0.0000	0.0000

or nearly unimodal profiles, implying expanding populations; however, other populations produced multimodal profiles (Fig. S3). On the other hand, the ITS2 dataset in all trial schemes including western versus eastern population groups produced a unimodal graph with an average of $HRag = 0.21439$ supporting overall population expansion (Fig. 6 D–F). However, the Balıkesir, Bursa, Çanakkale, Çankırı, İstanbul, İzmir, Kırklareli, Kütahya and Yalova populations generated ragged multimodal or bimodal profiles, whereas all other populations produced unimodal graphs indicating population growth (Fig. S4).

Except for Fu's F_s (-24.20524 , $P < 0.02$), the neutrality tests calculated for the *cyt b* haplotypes including all populations were not statistically significant (Tajima's $D = -0.98102$, $P > 0.05$, and $HRag = 0.0030$, $P > 0.05$). Separately analysed populations showed that the Çankırı (-2.9681 , $P < 0.02$) and Kırklareli populations (-3.8244 , $P < 0.02$) produced negatively significant Fu's F_s values, whereas other populations did not produce any significant values. Moreover, Bingöl (Tajima's $D = -1.7205$, $P < 0.05$) and Isparta (Tajima's $D = -1.7411$, $P < 0.05$) generated significant negative values. Small SSD and $HRag$ values also supported the goodness of fit between the expected and the observed distributions for all four populations. On the other hand, the ITS2 data for all populations produced a low Harpending's raggedness index ($HRag = 0.01648$, $P < 0.001$) and negative and significant Fu's F_s (-25.67093 , $P < 0.001$) suggesting population growth; however, Tajima's D (2.06816 , $P > 0.05$) was not statistically significant. Analyses conducted separately for each population showed that Fu's F_s was only negative and significant for Erzinçan (Fu's $F_s = -1.26262$, $P < 0.02$) and population growth was further suggested by the small SSD and $HRag$ values (Table 3). The neutrality tests also applied to the western versus eastern population groups indicated population growth both for the *cyt b* (Fu's $F_s = -24.1938$, $P < 0.02$ for the western and Fu's $F_s = -10.5554$, $p < 0.02$ for the eastern group) and the ITS2 dataset (Fu's $F_s = -13.8883$, $P < 0.02$ for the western group and Fu's $F_s = -5.6022$, $P < 0.02$ for the eastern group).

DISCUSSION

Genetic estimations and population demographic analyses

In the Turkish populations of *C. quercusfolii*, the high haplotype number for the *cyt b* gene is remarkable. Singleton and private haplotype (unique to a single population) numbers were also conspicuously high; there were eight private haplotypes, and, out of 125 haplotypes, 80 were singletons with only a single nucleotide substitution in pairwise comparisons. Singleton number was also high for the ITS2 alleles, where we found 18 singletons in addition to one private allele out of the 30 distinct ITS2 sequences. Furthermore, we detected a lower allele number as compared to the high haplotype number in *C. quercusfolii*, which is not unexpected since comparable results are also reported for other oak gall wasps in Turkey. For instance, there are 39 private haplotypes and 22 singletons out of 50

detected haplotypes for the *cyt b* gene and 8 singleton alleles out of 15 ITS2 alleles in *Trigonaspis synaspis* (Mutun & Atay, 2015). In another species of oak gall wasp, *Andricus curtisii*, half of the haplotypes are singletons for *cyt b* sequences (Mutun, 2016).

Excluding samples from small populations, most of the *C. quercusfolii* populations displayed high genetic variation. Specifically, some eastern populations (e.g., Erzurum, Erzincan and Gümüşhane) and some of the western localities (e.g., İstanbul, Bursa, Kütahya, Balıkesir and Kırklareli) showed high haplotype diversity. The highest nucleotide diversity was detected in two central Anatolian populations, from Konya and Karaman. The current findings are congruent with the results of other previously studied oak gall wasp taxa in Turkey (Rokas et al., 2003; Challis et al., 2007; Mutun, 2010, 2011, 2016; Dinç & Mutun, 2011; Mutun & Atay, 2015). Overall, the existence of high genetic diversity in animal and plant taxa is explained mainly by the presence of three phytogeographical regions in Turkey, the effects of the cyclical changes of glacial periods and heterogeneous topography (Bilgin, 2011); thus, our current data on *C. quercusfolii* further support this general conclusion.

Past biogeographical or climatic changes are key factors governing the contemporary population genetic structure of a species, and their effects can be traced by employing several population demographic analyses (Avise, 2000). In *C. quercusfolii*, the existence of a high number of both singletons and private haplotypes/alleles in addition to the high ratio of one or few substitutions in pairwise comparisons of sequences implies recent population growth (Crandall & Templeton, 1993). Nonetheless, lower nucleotide diversity compared to the higher haplotype diversity as recorded in *C. quercusfolii* also suggests rapid population expansion from a small population in the past (Avise, 2000; Zhang et al., 2006).

Population expansion was also indicated by the unimodal profile, non-significant $HRag$ and small SSD values, and negative statistically significant Fu's F_s for the ITS2 dataset. The unimodal or nearly unimodal profile recorded for all samples in addition to the east and west population groups strongly indicate that this species was widespread in Turkey in the ancient past and expanded its distribution. Non-significant SSD and $HRag$ values together with the unimodal profile in the mismatch distribution analyses as recorded for *C. quercusfolii* would be expected under neutrality in expanding populations (Zhang et al., 2006). However, the *cyt b* dataset produced a multimodal profile, implying that not all the samples from the populations clearly show population expansion in addition to western versus eastern groups. Despite the differences in the mismatch profiles of the ITS2 and the *cyt b* datasets, the significant and negative Fu's F_s values support expansion events in the different populations of *C. quercusfolii*. On the other hand, in our analysis the non-significant negative Tajima's D seems to contradict Fu's F_s , but the discrepancy between the two values may be due to their different statistical power, since Fu's F_s is more powerful

and sensitive to changes in population size than Tajima's D (Ramos-Onsins & Rozas, 2002; Jong et al., 2011). Furthermore, the significant and negative Fu's F_s values of both *cyt b* and ITS2 datasets for the western versus eastern populations implied historical population growth. In addition, populations individually analysed using the *cyt b* data indicated that some localities showed population growth. For instance, the Kırklareli population located in the Thrace region of Turkey yielded significantly negative Fu's F_s , and an eastern population, Bingöl, generated a significantly negative Tajima's D value. Studies on another species of oak gall wasp also indicate that the Kırklareli population is a hotspot locality and a centre of diversity for *A. caputmedusae* (Mutun, 2010). Likewise, the Bingöl population is crucial for *A. lucidus* in providing genetic diversity for nearby populations (Mutun, 2011). On the other hand, the ITS2 data produced negative and significant Fu's F_s (Fu's $F_s = -1.26262$, $P < 0.02$) only for an easterly located population, Erzincan, which was also supported by small SSD and *H*Rag values, implying that this population underwent population growth in the past. In this study *cyt b* and ITS2 data weakly supported each other; however, the former is used to explain the relatively recent past, while the latter points to the more ancient past (Zhang & Hewitt, 2003). In general, it seems that historical factors in the past have imposed their effects on *C. quercusfolii* populations in Turkey, resulting in the more ancient population expansion from eastern to western Turkey as suggested by the ITS2 data and continuing more recent historical expansion as indicated by the *cyt b* gene data.

Possible effects of topography and climate

The topographic structure within the distribution of a species is crucial for the geographic distribution of genetic diversity at the intraspecific level, and the effects of the geographic barriers that created the contemporary pattern can be assessed by employing different analyses (Avise, 2000). All phylogenetic analyses of the *cyt b* gene data support the presence of two main clades in *C. quercusfolii*. The first clade (Clade A), with haplotypes from the populations in eastern, western and central Anatolia, signifies a wide-spread and ancient distribution of this species in Turkey. A clear separation of the haplotype groups in the second clade (Clade B), corresponding to the Anatolian Diagonal, makes it apparent that the mountain line created a break between the major lineages of *C. quercusfolii*. The break between the eastern and western sides of the mountainous zone is recorded in the formation of two subclades, B2 and B3, where B2 dominated in the eastern part of the Anatolian belt and B3 in the western part. Network analysis also supported the same lineage break. Although not as robust as the results of the *cyt b* gene, the ITS2 dataset also indicate a similar but relatively less prominent separation of alleles. The difference in the power of resolution of the two markers is well known, where the *cyt b* and ITS2 datasets may shed light on different timescales due to the difference between the coalescent times as a result of their different inheritance and mutation patterns (Hoy, 1994; Zhang & Hewitt, 2003). Moreover, the difference between mtDNA

and nDNA markers may be due to the loss of phylogenetic and phylogeographic signals in ITS2 datasets since the effective population size of nDNA is four times larger than the mitochondrial DNA population size (Shaw, 2002; Marudlyn et al., 2011).

Our results indicate that the Anatolian Diagonal was effective in shaping the genetic variation in *C. quercusfolii* populations. AMOVA by clustering the populations into the eastern and the western groups revealed that a small amount of variation was partitioned between the two groups in the ITS2 dataset, where genetic variation within populations was high. On the other hand, genetic diversity among populations increased while variation within populations decreased for the *cyt b* dataset. Nonetheless, spatial autocorrelation analysis revealed a monotonic decline in the autocorrelation correlograms of both datasets, indicating that neighbouring populations are more similar, while genetic distance increases with geographic distance. As a general pattern a population can indicate either an island or a stepping-stone model. In the island model new patches are colonized by founders originating from a single source, while in the stepping-stone model migrants invade an empty and suitable patch from a geographically neighbouring population. Both models predict that genetic diversity decreases along the colonization route when range expansion occurs; however, this may not always be the case (Le Corre & Kremer, 1998). In many of the Western Palearctic species of oak gall wasps, the east, including the Iranian and Anatolian populations, is regarded as “the cradle of genetic diversity” and acting as a source for the European populations may require that the eastern part of the Turkish *C. quercusfolii* population is more diverse and an invasion route is from east to west (Rokas et al., 2003). In this study, autocorrelograms implying directional migration seem to support this pattern; however, due to the absence of samples from either Iranian populations or European localities it is not possible to draw a firm conclusion. Such a larger scale study is among our future work plans.

Past environmental changes and associated climatic fluctuations leave their imprints in the genomic components of taxa, which are reflected in their phylogenetic reconstruction, and that timing of lineage divergences can be estimated through implementing the molecular clock (Hewitt, 2001; Avise, 2004; Weir & Schluter, 2007). It is known that geological changes are often coupled with climatic oscillations (Gillispie & Roderick, 2014). Orogenesis started around the early Oligocene and continued through the Late Miocene, causing the formation of the Taurus mountains as part of the Alp-Himalayan orogenic belt, resulting in the uplift of eastern Turkey, while the central part of Anatolia was a lowland area and covered by a large central Anatolian lake system (Karadenizli, 2011). Around these times, significant climatic changes from a more wet to a more arid environment occurred (Fortelius et al., 2002), with a commonly characterized Messinian aridity crisis at the end of the Miocene (7.2–5.3 MYA) (Demirsoy, 2002). In this study, application of the molecular clock to the *cyt b* dataset indicates that the separation of the outgroups

from the ingroup occurred nearly 9.49 MYA (95% HPD bound 3.38–11.90 MYA), which was in the Tortonian age (11.6–7.2 MYA) of the Miocene epoch (23–5.3 MYA). The diversification of *C. quercusfolii* from the outgroup species may fit well with the ongoing events in the Late Miocene. A wide-scale study conducted on a variety of taxa of oak gall wasps in the Western Palearctic indicate that the genus *Cynips* split from other oak gall wasp genera around 6.81 MYA (95% HPD bound 4.5–10.09 MYA) (Stone et al., 2009). In this study, the diversification of *C. quercusfolii* from the outgroups is compatible with the upper bound of the genus split dating in the Late Miocene. The estimate of a deeper split in the Turkish ingroup/outgroup taxa recorded in the present study may be explained by the existence of more ancient lineages in Turkey. Thus, our current findings point to a common pattern in oak gall wasp phylogeny in the Western Palearctic and may imply that the more eastern lineages are more ancient than the European lineages. As a consequence, the eastern populations might have acted as source populations for the more western populations currently in Europe (Stone et al., 2012; Bunnefeld et al., 2018).

Ongoing diversification in the ingroup taxon resulted in the first bifurcation of *C. quercusfolii* lineages into two main clades, Clades A and B, around 6.6 MYA, which coincides with the Messinian age in Late Miocene. It is well documented that the changes at this transition induced speciation events in insects (Çıplak, 2004a) and other animals (Hrbek et al., 2002; Kornilios et al., 2012). Furthermore, the Pliocene oscillations around 5.05 MYA seem to structure Clade B via the split in the basally placed Sivas haplotype (H117). About 4.34 MYA, in the Pliocene epoch (early Zanclean time – split age 5.05 MYA), a major diversification event seems to have resulted in the separation of the B1 subclade. The subsequent split of the haplotypes into an eastern dominated B2 subclade and western dominated B3 subclade occurred about 3.71 MYA (Late Zanclean). The separation of subclades B2 and B3 roughly into eastern and western divisions may be explained by the effect of the formation of the Anatolian Diagonal during the Miocene/Pliocene and associated climatic changes on *C. quercusfolii*. Similar time frames causing severe splitting events during the early Pliocene due to major tectonic activities in addition to climatic fluctuations are used to explain the diversification in a group of lizards (Kapli et al., 2013). Orogenesis events associated with the wetter and hotter climate in Late Pliocene followed by the drier and colder conditions in Early Pleistocene induced fragmentation of populations (Bennett, 1990) and resulted in a widespread extinction of some taxa/populations in the western Palearctic (Hewitt, 2000). Recurrent divergence events during the pre-Pleistocene and throughout the Pleistocene appear to have resulted in relatively deep to intermediate splits creating a well-structured appearance of both the eastern (the B2 subclade) and western (the B3 subclade) lineages of *C. quercusfolii*. In the eastern subclade, splitting began prior to the Pleistocene (2.91 MYA) and continued during the interglacial stage between the Riss and the Würm (130,000

YA). On the other hand, in the western subclade splitting events seem to have started later (ca. 1.6 MYA) than in the eastern subclade and continued for 90,000 years. In fact, our findings for *C. quercusfolii* in the present study are compatible with the westward range expansion of *Andricus coriarius* around 1.6 MYA (Challis et al., 2007). The signatures found in *C. quercusfolii* also match the timing of events in *A. kollari* (Stone et al., 2007). Even though the sampling of *A. quercustozae* was limited to only seven localities, a clear separation and timing of the split between the south-western and north-eastern Turkish clades created by the Anatolian Diagonal also support our current results. Similar diversification events in other species of animals (Veith et al., 2003; Çıplak, 2004b, 2008; Gündüz et al., 2007; Bellati et al., 2011; Kaya et al., 2014; Özdemir et al., 2014) and plants (Ansell et al., 2011) in Turkey have shown that lineages diverge into more sub lineages with deep/intermediate to shallow splits. Overall these findings may provide the robust data necessary to draw the general conclusion that in the last 10 million years the dynamic geographic and climatic changes in the Turkish landmass together with its neighbouring area left strong imprints on oak gall wasps.

ACKNOWLEDGMENTS. This study was supported by the Bolu Abant İzzet Baysal University, Grant No: 2011.03.01. 380.

REFERENCES

- AKAIKE H. 1974: A new look at the statistical model identification. — *IEEE T Automat Contr.* **19**: 716–723.
- ANSSELL S.W., STENOEN H.K., GRUNDMANN M., RUSSELL S.J., KOCH M.A., SCHNEIDER H. & VOGEL J.C. 2011: The importance of Anatolian mountains as the cradle of global diversity in *Arabis alpine*, a key arctic-alpine species. — *Ann. Bot.* **108**: 241–252.
- AVISE J.C. 2000: *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA, 447 pp.
- AVISE J.C. 2004: *Molecular Markers, Natural History and Evolution*. 2nd ed. Sinauer, Sunderland, MA, 684 pp.
- BELLATI A., PELLITTERI-ROSA D., SACCHI R., NISTRI A., GALIMBERTI A., CASIRAGHI M., FASOLA M. & GALEOTTI P. 2011: Molecular survey of morphological subspecies reveals new mitochondrial lineages in *Modarcis muralis* (Squamata: Lacertidae) from the Tuscan Archipelago (Italy). — *J. Zool. Syst. Evol. Res.* **49**: 240–250.
- BENNETT M.R. 1990: The deglaciation of Glen Croulin, Knysna. — *Scott. J. Geol.* **26**: 41–46.
- BILGIN R. 2011: Back to the suture: the distribution of intraspecific genetic diversity in and around Anatolia. — *Int. J. Mol. Sci.* **12**: 4080–4103.
- BOZKURT E. 2001: Neotectonics of Turkey – a synthesis. — *Geodin. Acta* **14**: 3–30.
- BUNNEFELD L., HEARN J., STONE G.N. & LOHSE K. 2018: Whole genome data reveal the complex history of a diverse ecological community. — *Proc. Natl. Acad. Sci. U.S.A.* **115**: E6507–E6515.
- CHALLIS R.J., MUTUN S., NIEVES-ALDREY J.L., PREUSS S., ROKAS A., AEBI A., SADEGHI E., TAVAKOLI M. & STONE G.N. 2007: Longitudinal range expansion and cryptic eastern species in the western Palearctic oak gall wasp *Andricus coriarius*. — *Mol. Ecol.* **16**: 2103–2114.
- ÇIPLAK B. 2004a: Biogeography of Anatolia: the marker group Orthoptera. — *Mem. Soc. Entomol. Ital.* **82**: 357–372.

- ÇIPLAK B. 2004b: Systematics, phylogeny and biogeography of *Anterastes* (Orthoptera, Tettigonidae, Tettigoniinae): evolution within a refugium. — *Zool. Scr.* **33**: 19–44.
- ÇIPLAK B. 2008: The analogy between interglacial and global warming for the glacial relicts in a refugium: a biogeographic perspective for conservation of Anatolian Orthoptera. In Fattorini S. (ed.): *Insect Ecology and Conservation*. Research Sign Post, Trivandrum, Kerala, pp. 135–163.
- CRANDALL K.A. & TEMPLETON A.R. 1993: Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. — *Genetics* **134**: 959–969.
- DARRIBA D., TABOADA G.L., DOALLO R. & POSADA D. 2012: JModeltest 2: More models, new heuristics and parallel computing. — *Nat. Methods* **9**(8): 772, 4 pp.
- DAVIS P.H. 1971: Distribution patterns in Anatolia with particular reference to endemism. In Davis P.H., Harper P.C. & Hedge I.C. (eds): *Plant Life in South-West Asia*, Botany Society of Edinburgh, pp. 15–27.
- DEMİRİSOY A. 2002: *General Zoogeography and Zoogeography of Turkey (Geography of Animals)*. Meteksan, Ankara, 1007 pp. [in Turkish].
- DİNÇ S. & MUTUN S. 2011: PCR-RFLP variation of the oak gall wasp, *Andricus quercustozae* (Bosc, 1792) (Hymenoptera: Cynipidae) from Turkey. — *Turk. Entomol. Derg.* **35**: 47–58.
- DRUMMOND A.J. & SUCHARD A.S. 2010: Bayesian random local clocks, or one rate to rule them all. — *BMC Biol.* **8**: 114, 12 pp.
- DRUMMOND A.J., SUCHARD A.S., XIE D. & RAMBAUT A. 2012: Bayesian phylogenetics with BEAUti and the BEAST 1.7. — *Mol. Biol. Evol.* **29**: 1969–1973.
- DUBEY S., COSSON J.F., VOHRALIK V., KRYSTUFEK B., DIKER E. & VOGEL P. 2007: Molecular evidence of Pleistocene bidirectional faunal exchange between Europe and the Near East: The case of the bicoloured shrew (*Crocidura leucodon*, Soricidae). — *J. Evol. Biol.* **20**: 1799–1808.
- EKİM T. & GÜNER A. 1986: The Anatolian Diagonal: fact or fiction? — *Proc. Roy. Soc. Edinb. (B)* **89**: 69–77.
- EXCOFFIER L. & LISCHER H.E.L. 2010: Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under linux and windows. — *Mol. Ecol. Resour.* **10**: 564–567.
- FORTELIUS M., ERONEN J.T., JERNVALL J., LIU L., PUSHKINA D., RINNE J., TESAKOV A., VISLOBOKOVA I.A., ZHANG Z. & ZHOU L. 2002: Fossil mammals resolve regional patterns of Eurasian climate change during 20 million years. — *Evol. Ecol. Res.* **4**: 1005–1016.
- FU Y.X. 1997: Statistical tests of neutrality against population growth, hitchhiking and background selection. — *Genetics* **147**: 915–925.
- GILLISPIE R.G. & RODERICK G.K. 2014: Evolution: Geology and climate drive diversification. — *Nature* **509**: 297–298.
- GUINDON S. & GASCUEL O. 2003: A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. — *Syst. Biol.* **52**: 696–704.
- GÜNDÜZ İ., JAAROLA M., TEZ C., YENİYURT C., POLLY P.D. & SEARLE J.B. 2007: Multigenic and morphometric differentiation of ground squirrels (*Spermophilus scuiridae* Rodentia) in Turkey, with a description of a new species. — *Mol. Phylogenet. Evol.* **43**: 916–935.
- GÜR H. 2016: The Anatolian Diagonal revisited: testing the ecological basis of a biogeographic boundary. — *Zool. Middle East* **62**: 189–199.
- HALL T.A. 1999: BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. — *Nucl. Acid Symp. Ser.* **41**: 95–98.
- HARPENDING H.C. 1994: Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. — *Hum. Biol.* **66**: 591–600.
- HAYWARD A. & STONE G.N. 2006: Comparative phylogeography across two trophic levels: the oak gall wasp *Andricus kollari* and its chalcid parasitoid *Megastigmus stigmatizans*. — *Mol. Ecol.* **15**: 479–489.
- HEWITT G.M. 1999: Post-glacial re-colonization of European biota. — *Biol. J. Linn. Soc. Lond.* **68**: 87–112.
- HEWITT G.M. 2000: The genetic legacy of the Quaternary ice ages. — *Nature* **405**: 907–913.
- HEWITT G.M. 2001: Speciation, hybrid zones and phylogeography or seeing genes in space and time. — *Mol. Ecol.* **10**: 537–549.
- HOY M.A. 1994: *Insect Molecular Genetics: An Introduction to Principles and Applications*. Academic Press, San Diego, CA, 546 pp.
- HRBEK T., KÜÇÜK F., FRICKEY T., STÖLTING K.N., WILDEKAMP R. H. & MEYER A. 2002: Molecular phylogeny and historical biogeography of the *Apahnius* (Pisces, Cyprinodontiformes) species complex of Central Anatolia, Turkey. — *Mol. Phylogenet. Evol.* **25**: 125–137.
- Jİ Y.J., ZHANG D.X. & HE L.J. 2003: Evolutionary conservation and versatility of a new set of primers for amplifying the ribosomal internal transcribed spacer regions in insects and other invertebrates. — *Mol. Ecol. Notes* **3**: 581–585.
- JONG A.M., WAHLBERG N., EIJIK M.A., BRAKFIELD P.M. & ZWAN B.J. 2011: Mitochondrial DNA signature for range-wide populations of *Bicyclus anynana* suggest a rapid expansion from recent refugia. — *PLoS ONE* **6**(6): e21385, 5 pp.
- KAPLI P., BOTONI D., ILGAZ Ç., KUMLUTAŞ Y., AVCI A., RASTEGAR-POUYANI N., FATHINIA B., LYMBERAKIS P., AHMADZADEH F. & POULAKAKIS N. 2013: Molecular phylogeny and historical biogeography of the Anatolian lizard *Apathya* (Squamata, Lacertidae). — *Mol. Phylogenet. Evol.* **66**: 992–1001.
- KARADENİZLİ L. 2011: Oligocene to Pliocene palaeogeographic evolution of the Çankırı-Çorum Basin, Central Anatolia, Turkey. — *Sediment. Geol.* **237**: 1–29.
- KAYA S. & ÇIPLAK B. 2016: Budding speciation via peripheral isolation: The *Psorodonotus venosus* (Orthoptera, Tettigoniidae) species group example. — *Zool. Scr.* **45**: 521–537.
- KAYA S., BOZTEPE Z. & ÇIPLAK B. 2014: Phylogeography of the *Poecilimon luschni* species group (Orthoptera, Tettigoniidae): A radiation strictly correlated with climatic transitions in the Pleistocene. — *Zool. J. Linn. Soc.* **173**: 1–21.
- KORNİLİOS P., ILGAZ Ç., KUMLUTAŞ Y., LYMBERAKIS P., MORAVEC J., SINDACO R., RASTEGAR-POUYANI N., AFROOSHEH M., GİOKAS S., FRAGUEDAKIS-TSOLIS S. & CHONDROPOULOS B. 2012: Neogene climatic oscillations shape the biogeography and evolutionary history of the Eurasian blindsnake. — *Mol. Phylogenet. Evol.* **62**: 856–873.
- LE CORRE V. & KREMER A. 1998: Cumulative effects of founding events during colonization on genetic diversity and differentiation in an island and stepping-stone model. — *J. Evol. Biol.* **11**: 495–512.
- LIBRADO P. & ROZAS J. 2009: DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. — *Bioinformatics* **25**: 1451–1452.
- MARDULYN P., OTHMEZOURI N., MIKHAILOV Y.E. & PASTEELS J.M. 2011: Conflicting mitochondrial and nuclear phylogeographic signals and evolution of host-plant shifts in the Boreo-Montane leaf beetle *Chrysomela lapponica*. — *Mol. Phylogenet. Evol.* **61**: 686–696.
- MELİKA G. 2006: *Gallwasps of Ukraine: Cynipidae. Vol. 1, 2*. Schmalhausen Institute of Zoology, National Academy of Sciences of Ukraine, Kiev, 646 pp.

- MUTUN S. 2010: Intraspecific genetic variation and the phylogeography of the oak gallwasp *Andricus caputmedusae* (Hymenoptera: Cynipidae): Effects of the Anatolian Diagonal. — *Acta Zool. Acad. Sci. Hung.* **56**: 153–172.
- MUTUN S. 2011: Intraspecific genetic diversity and phylogeography of the oak gall wasp *Andricus lucidus* (Hymenoptera: Cynipidae) populations in Anatolia. — *Turk. J. Zool.* **35**: 559–570.
- MUTUN S. 2016: Molecular diversity and phylogeography of *Andricus curtisii* (Hymenoptera, Cynipidae) in Turkey. — *Biochem. Syst. Ecol.* **67**: 74–85.
- MUTUN S. & ATAY G. 2015: Phylogeography of *Trigonaspis synaspis* (Hymenoptera: Cynipidae) from Anatolia based on mitochondrial and nuclear DNA sequences. — *Eur. J. Entomol.* **112**: 259–269.
- NEI M. 1987: *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY, 512 pp.
- OKAY A.I. 2008: Geology of Turkey: A synopsis. — *Anschnitt* **21**: 19–42.
- ÖZDEMİR N., GÜL S., POYARKOV N.A., KUTRUP B., TOSUNOĞLU M. & DOĞLIO S. 2014: Molecular systematics and phylogeography of *Bufotes variabilis* (syn. *Pseudepidalea variabilis*) (Pallas, 1769) in Turkey. — *Turk. J. Zool.* **38**: 412–420.
- PAPADOPOULOU A., ANASTASIOU I. & VOGLER A.P. 2010: Revisiting the insect mitochondrial molecular clock: The mid-Aegean trench calibration. — *Mol. Biol. Evol.* **27**: 1659–1672.
- PEAKALL R. & SMOUSE P.E. 2006: GENALEX 6: Genetic analysis in excel. Population genetic software for teaching and research. — *Mol. Ecol. Notes* **6**: 288–295.
- PEAKALL R. & SMOUSE P.E. 2012: GenA1Ex 6.5: Genetic analysis in excel. Population genetic software for teaching and research – an update. — *Bioinformatics* **28**: 2537–2539.
- PEAKALL R., RUIBAL M. & LINDENMAYER D.B. 2003: Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. — *Evolution* **57**: 1182–1195.
- POSADA D. & BUCKLEY T.R. 2004: Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. — *Syst. Biol.* **53**: 793–808.
- PROVAN J. & BENNETT K.D. 2008: Phylogeographic insights into the cryptic glacial refugia. — *Trends Ecol. Evol.* **23**: 564–571.
- RAMBOUT A. 2009: *FigTree. Version 1.3.1*. URL: <http://tree.bio.ed.ac.uk/software/figtree>.
- RAMOS-ONSINS S.E. & ROZAS J. 2002: Statistical properties of new neutrality tests against population growth. — *Mol. Biol. Evol.* **19**: 2092–2100.
- ROGERS A. & HARPENDING H. 1992: Population growth makes waves in the distribution of pairwise differences. — *Mol. Biol. Evol.* **9**: 552–569.
- ROKAS A., ATKINSON R.J., WEBSTER L.M.I., CSÖKA G. & STONE G.N. 2003: Out of Anatolia: Longitudinal gradients in genetic diversity support an eastern origin for a Circum-Mediterranean oak gall wasp *Andricus quercustozae*. — *Mol. Ecol.* **12**: 2153–2174.
- RONQUIST F. & HUELSENBECK J.P. 2003: MrBayes 3: Bayesian phylogenetic inference under mixed models. — *Bioinformatics* **19**: 1572–1574.
- SCHNEIDER S. & EXCOFFIER L. 1999: Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. — *Genetics* **152**: 1079–1089.
- ŞEKERCİOĞLU Ç.H., ANDERSON S., AKÇAY E., BILGIN R., CAN Ö.E., SEMİZ G., TAŞANÖĞLU Ç., YOKES M.B., SOYUMERT A., İPEKDAL K. ET AL. 2011: Turkey's globally important biodiversity in crisis. — *Biol. Consev.* **144**: 2752–2769.
- SHAW K.L. 2002: Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. — *Proc. Natl. Acad. Sci. U.S.A.* **99**: 16122–16127.
- SIMON C., FIRATI F., BECKENBACH A., CRESPI B., LIU H. & FLOOK P. 1994: Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and compilation of conserved polymerase chain reaction primers. — *Ann. Entomol. Soc. Am.* **87**: 651–701.
- SOKAL R.R., ODEN N.L. & THOMSON B.A. 1997: A simulation study of microevolutionary inferences by spatial autocorrelation analysis. — *Biol. J. Linn. Soc.* **60**: 73–93.
- STONE G.N. & SUNNUCKS P.J. 1993: Genetic consequences of an invasion through a patchy environment – the cynipid gallwasp *Andricus quercuscalicis* (Hymenoptera: Cynipidae). — *Mol. Ecol.* **2**: 251–268.
- STONE G.N., SCHÖNRÖGGE K., ATKINSON R.J., BELLIDO D. & PUJADE-VILLAR J. 2002: The population biology of oak gall wasps (Hymenoptera: Cynipidae). — *Annu. Rev. Entomol.* **47**: 633–668.
- STONE G.N., CHALLIS R.J., ATKINSON R.J., CSÖKA G., HAYWARD A., MELIKA G., MUTUN S., PREUSS S., ROKAS A., SADEGHI E. ET AL. 2007: The phylogeographical clade trade: Tracing the impact of human-mediated dispersal on the colonization of northern Europe by the oak gall wasp *Andricus kollari*. — *Mol. Ecol.* **16**: 2768–2781.
- STONE G.N., HERNANDEZ-LOPEZ A., NICHOLLS J.A., DI PIERRO E., PUJADE-VILLAR J., MELIKA G. & COOK J.M. 2009: Extreme host plant conservatism during at least 20 million years of host plant pursuit by oak gallwasps. — *Evolution* **63**: 854–869.
- STONE G.N., LOHSE K., NICHOLLS J.A., FUENTES-UTRILLA P., SINCLAIR F., SCHÖNRÖGGE K., CSÖKA G., MELIKA G., NIEVES-ALDREY J.-L., PUJADE-VILLAR J., TAVAKOLI M., ASKEW R.R. & HICKERSON M.J. 2012: Reconstructing community assembly in time and space reveals enemy escape in a western Palearctic insect community. — *Curr. Biol.* **22**: 1–6.
- Swofford D.L. 2002: *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), Ver. 4.0 b10*. Sinauer Associates, Sunderland, MA.
- 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.
- TEACHER A.G.F. & GRIFFITHS D.J. 2011: HapStar: Automated haplotype network layout and visualization. — *Mol. Ecol. Resour.* **11**: 151–153.
- VEITH M., SCHMIDTLER J., KOSUCH J., BARAN I. & SEITZ A. 2003: Paleoclimatic changes explain Anatolian mountain frog evolution: A test for alternating vicariance and dispersal events. — *Mol. Ecol.* **12**: 185–199.
- WEIR J.T. & SCHLUTER D. 2007: The latitudinal gradient in recent speciation and extinction rates of birds and mammals. — *Science* **315**: 1574–1576.
- ZHANG D.X. & HEWITT G.M. 2003: Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. — *Mol. Ecol.* **2**: 563–584.
- ZHANG J., CAI Z. & HUANG L. 2006: Population genetic structure of crimson snapper *Lutjanus erythropterus* in east Asia, revealed by analysis of the mitochondrial control region. — *ICES J. Mar. Sci.* **63**: 693–704.
- ZOHARY M. 1973: *Geobotanical Foundations of the Middle East*. Gustav Fisher, Stuttgart, 739 pp.

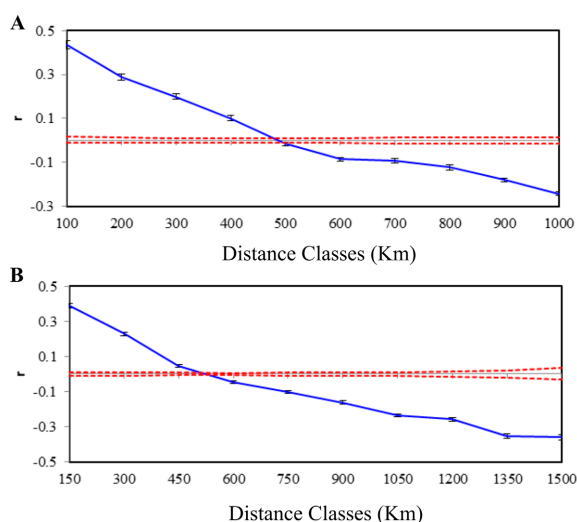
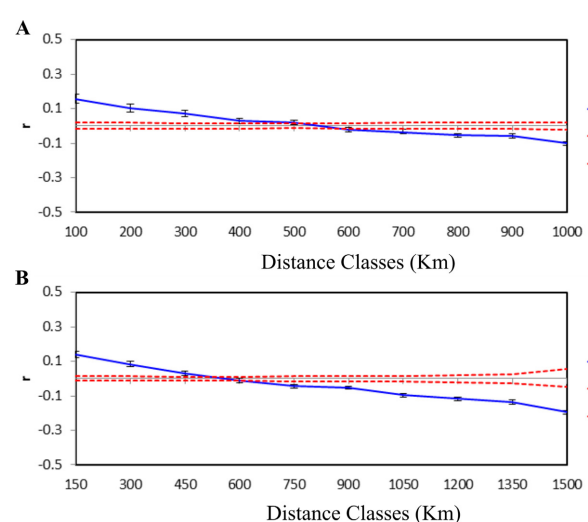
Received November 30, 2018; revised and accepted April 29, 2019
Published online May 21, 2019

Table S1. *Cyt b* haplotypes and their frequencies in *C. quercusfolii* populations.

[illegible]

Table S2. ITS2 alleles and their frequencies in *C. quercusfolii* populations.

	AFY	ART	BAL	BAR	BIN	BIT	BOL	BRS	CAN	CNK	DUZ	EDI	ELA	ERI	ERZ	GIR	GUM	ISP	IST	IZM	KAR	KAS	KAY	KLR	KON	KUT	MAL	MUS	ORD	SAM	SIN	SIV	TOK	TUN	VAN	YLV	YOZ	ZON	N _{Allot}
A1	10						9	7	4	3		4		1		3	7	4	6	6	2	11	7	6	3	5		1		4	8		1			4	1	3	128
A2		1																																					1
A3			1			1		1	1	2		1	4	1	2	1	1	2	1		1		2	2	4	4	3	1	1	6	1	4				1			49
A4			1																																				2
A5			1			1														1																			3
A6				1			1					1																											3
A7					9	10								8	2	5	3	4			7			2		3			5	1	1		1		1	9			71
A8					1																																	1	
A9								1																														1	
A10									3	1	1									1																2		8	
A11							1																									1						2	
A12									1																													1	
A13										1																												1	
A14										1																												1	
A15										1																												1	
A16											1		1																									2	
A17												1																										4	
A18													1																									1	
A19																1									1													2	
A20																																						1	
A21																																						2	
A22																																						1	
A23																																						1	
A24																																						1	
A25																																						1	
A26																																						1	
A27																																						1	
A28																																						1	
A29																																						1	
A30																																						1	
Total	10	2	10	1	12	10	10	9	10	10	2	7	5	10	4	10	11	10	10	10	10	11	11	10	10	3	11	2	11	10	6	1	1	1	10	10	1	3	294
N _{ind}	1	2	4	1	4	1	2	3	5	7	2	4	2	3	2	4	3	3	4	5	3	1	3	4	3	1	5	2	3	3	3	1	1	2	6	1	1		

**Fig. S1.** Correlograms showing spatial genetic correlation coefficient (r) as a function of geographic distances in the entire study area based on *cyt b* dataset. a – 100 km distance class, b – 150 km distance class. Dashed lines indicate upper (U) and lower (L) boundaries of 95% confidence intervals under the null hypothesis of no geographic structure is present. 95% bootstrapped confidence intervals are shown for each of the calculated r values.**Fig. S2.** Correlograms showing spatial genetic correlation coefficient (r) as a function of geographic distances in the entire study area based on ITS2 dataset. a – 100 km distance class, b – 150 km distance class. Dashed lines indicate upper (U) and lower (L) boundaries of 95% confidence intervals under the null hypothesis of no geographic structure is present. 95% bootstrapped confidence intervals are shown for each of the calculated r values.

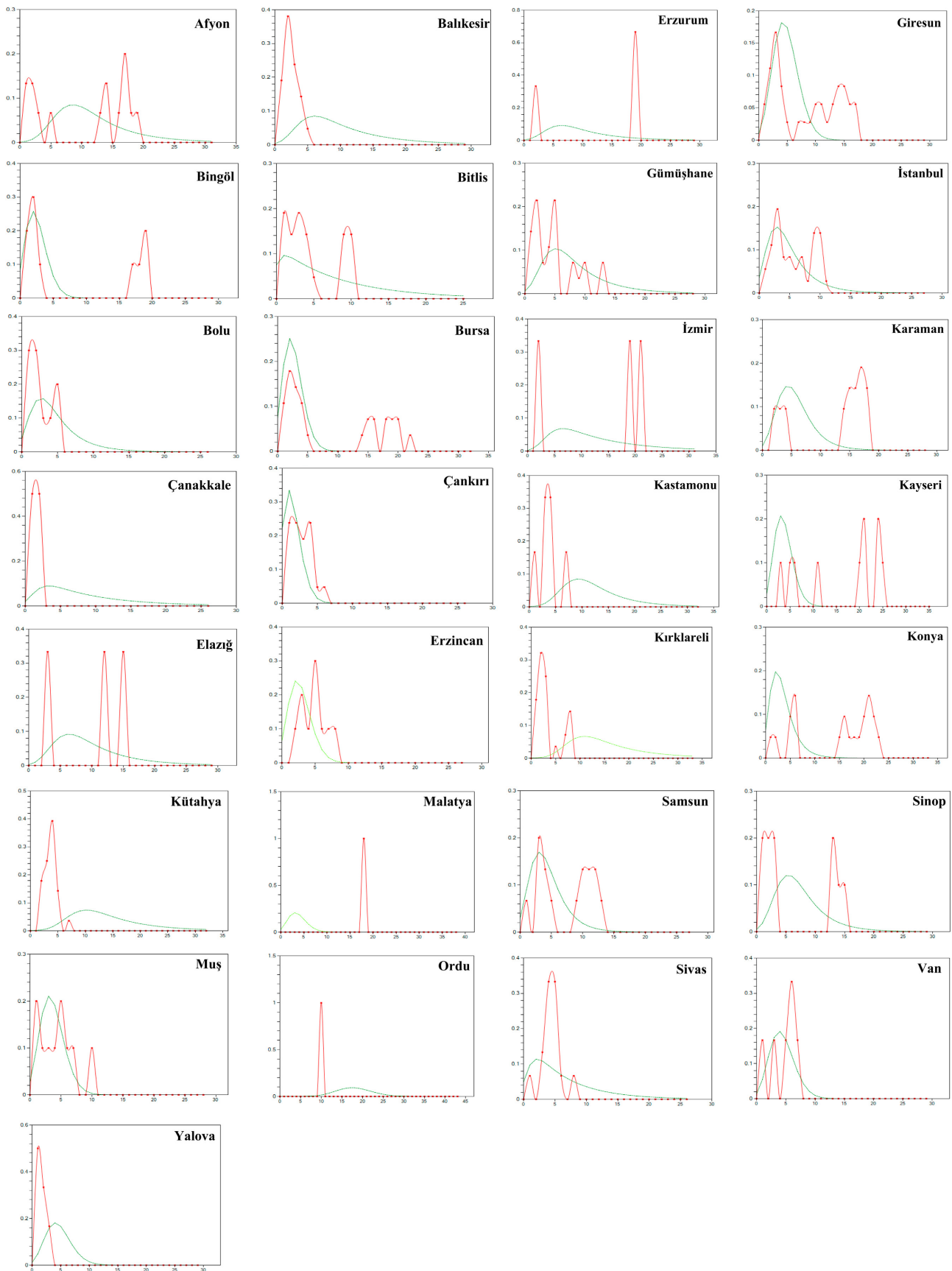


Fig. S3. Mismatch distribution of populations with their own haplotypes for the *cyt b* data set. The observed distribution is represented by a red line, and the expected frequencies by a green line.

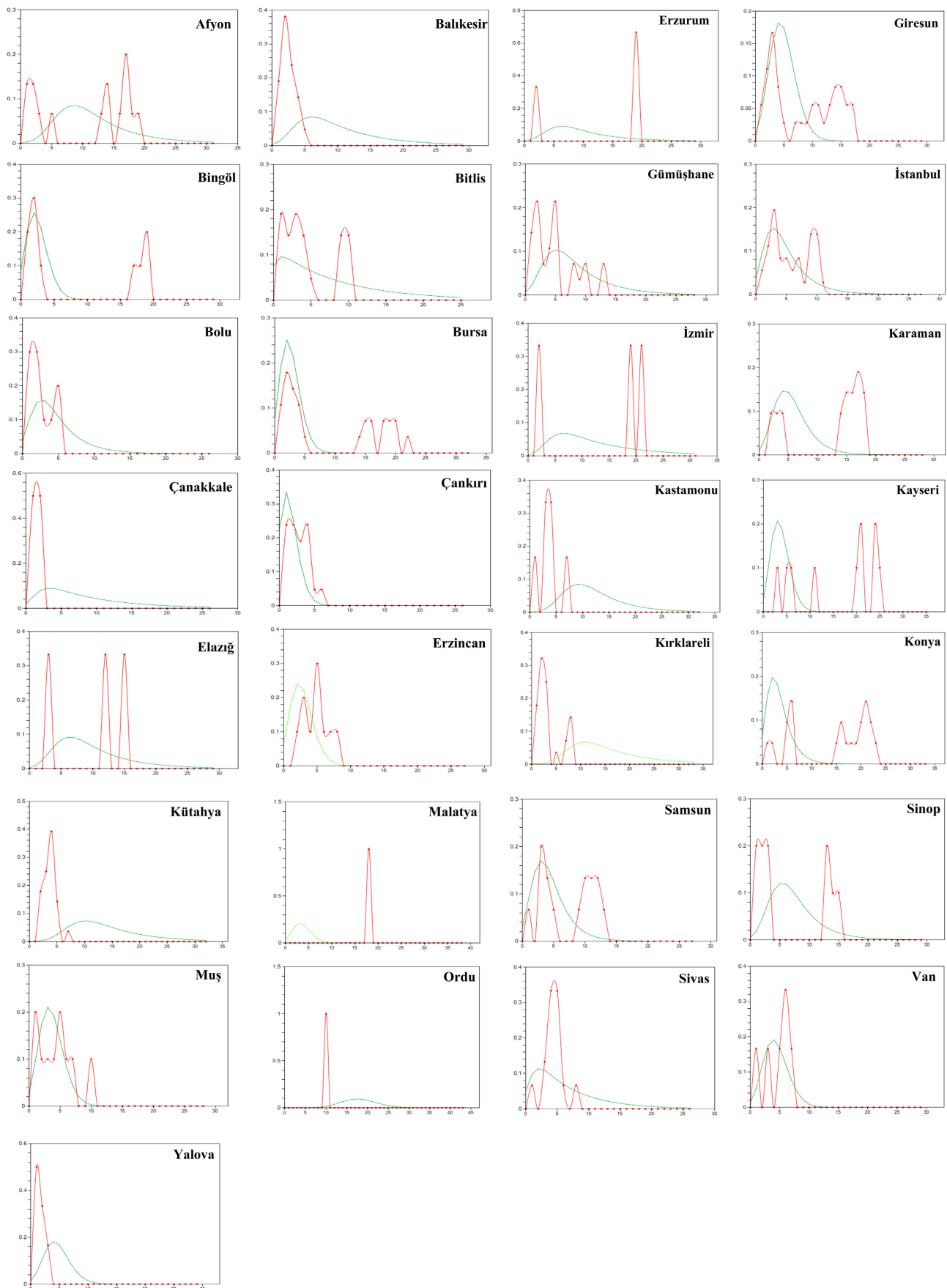


Fig. S4. Mismatch distribution of populations with their own alleles for the ITS2 data set. The observed distribution is represented by a red line, and the expected frequencies by a green line.