CYTONUCLEAR DISCORDANCE IN THE MARBLED CRAB, *PACHYGRAPSUS MARMORATUS* (FABRICIUS, 1787) ALONG THE MEDITERRANEAN COASTS OF TURKEY

by

Cansu Çetin

BS. in Molecular Biology and Genetics, Boğaziçi University, 2010

Submitted to the Institute of Environmental Sciences in partial fulfillment of the requirements for the degree of

Master of Science

in

Environmental Sciences

Boğaziçi University

2015

CYTONUCLEAR DISCORDANCE IN THE MARBLED CRAB, *PACHYGRAPSUS MARMORATUS* (FABRICIUS, 1787) ALONG THE MEDITERRANEAN COASTS OF TURKEY

APPROVED BY:

Assoc. Dr. İ. Raşit Bilgin (Thesis Supervisor)	……..
Prof. Dr. Andrzej Furman (Thesis Co-supervisor)	……..
Assist. Prof. Dr. Berat Z. Haznedaroğlu	……..
Assist. Prof. Dr. M. Baki Yokeş	……
Assist. Prof. Dr. Ulaş Tezel	..……..

DATE OF APPROVAL: 11/August/2015

To the memory of my grandfather,
Mehmet Ali Anbaroğlu

ACKNOWLEDGEMENTS

This thesis would not have been possible without the help of numerous people that either mediated, arranged, or collected crab samples. First of all, I owe many thanks to Evrim Kalkan who shared with me the joy of marine life, who taught me how to collect crabs, and who helped me whenever I needed from the beginning. Several crab samples were collected and sent to me from Muğla by Mert Elverici for which I am very grateful. I am especially thankful to my mother Ayşe Canan Çetin, my grandmother Zeliha Anbaroğlu and my father Gülşen Çetin who helped me a lot and joined me in some of the trips. Cevat Çetin, Feriha Çetin, Nazan Yürür, Mehmet Yürür, Özer Çekmer and Beyhan Çekmer also believed in me and provided logistic assistance in collecting crabs. I also want to thank several of my friends who were especially helpful in the field: Sena Özbek, Burcu Demirağ, Şebnem Özaltın, and Ceren Çekmer. I am especially grateful to my sister Aysu Çetin, who made me stronger by her presence and always kept my spirit and also crabs' spirit high by occasionally singing to them.

I would also like to thank Kanat Gürün, Öncü Maracı, Kübra Karaman and Ayşe Mergenci who helped me technically and emotionally in the laboratory. I would also like to thank Emek Çelik who responded to all of my questions. Emrah Çoraman motivated me to finish my thesis and gave me some important insights. I would also like to thank my thesis supervisor Raşit Bilgin who was always available and kind to me. Without his support and trust in me from the beginning, I would not have enjoyed field trips and laboratory this much. Lastly, I am very grateful to Andrzej Furman who was really patient with me, showed me how to deal with my data, was a great teacher, and most importantly, who encouraged me a lot.

This study was supported by a grant from the Boğaziçi University Research Fund (No: 11Y00P2) to RB.

CYTONUCLEAR DISCORDANCE IN THE MARBLED CRAB, *PACHYGRAPSUS MARMORATUS* (FABRICIUS, 1787) ALONG THE MEDITERRANEAN COASTS OF TURKEY

The marbled crab, *Pachygrapsus marmoratus* is a common species in the Mediterranean and along the coasts of Turkey. Its global distribution extends to the north-eastern part of the Atlantic Ocean. In this study, a total of 384 specimens from 32 sites were collected from around coasts of Turkey, and genetically examined in mitochondrial CO1 gene and five nuclear microsatellites for a better understanding of its evolutionary history. High levels of genetic differentiation was found between populations of *P. marmoratus* distributed all around the Turkish coasts (Population C) and those constrained to the Mediterranean coasts (Population M). Both populations, however, shared the same or have very similar mtDNA CO1 haplotypes. Different scenarios are discussed referring to the biogeographical history of the Mediterranean for explaining this pattern. Two alternative explanations for the *cytonuclear discordance* as retention of ancestral polymorphisms or gene flow after secondary contact could not be distinguished. The use of complementary mitochondrial markers with faster mutation rates than CO1 subunit in addition to microsatellites, and sampling individuals from the rest of the Mediterranean and from the Atlantic Ocean can be useful in resolving the observed pattern.

MERMER YENGECİ *PACHYGRAPSUS MARMORATUS* (FABRICIUS, 1787) TÜRÜNÜN TÜRKİYE'NİN AKDENİZ KIYILARINDAKİ SİTONÜKLEER UYUMSUZLUĞU

Mermer yengeci olarak da bilinen *Pachygrapsus marmoratus* Akdeniz'de ve Türkiye kıyılarında yaygın görülen bir yengeç türüdür. Akdeniz dışında küresel dağılımı Atlantik Okyanus'unun kuzeydoğusuna kadar uzanır. Toplam 32 noktadan 384 örnek toplanmış ve bu örnekler mitokondrial DNA'nın CO1 bölgesi ve beş nükleer mikrosatelit lokus kullanılarak incelenmiştir. Tüm kıyılara yayılmış *P. marmoratus* popülasyonu (Popülasyon C) ve Akdeniz'de sınırlı olan popülasyon (Popülasyon M) arasında büyük ölçüde genetik farklılık bulunmuştur. Buna ragmen, iki popülasyon mtDNA CO1 haplotipleri açısından belirgin farklılık göstermez. İki *P.marmoratus* popülasyonun ne zaman oluştuğu veya oluşma süreci belirsiz kalmıştır. Akdeniz'in biyocoğrafi tarihine değinilerek farklı senaryolar tartışılmıştır. Eski polimorfizmlerin iki popülasyonda farklılaşamaması veya ikincil karşılaşma sonrası gen akışı sitonükleer uyuşmazlığı açıklayabilir. CO1'den daha hızlı evrimleşen başka mitokondriyal bölgelerin kullanılması, Akdeniz'in geri kalanından ve Atlantik Okyanusu'ndan örneklerde kapsamlı bir çalışma yapılması türün Türkiye kıyılarındaki paternini anlamaya yardımcı olabilir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

Symbol	Explanation	Units used
AMOVA	Analysis of molecular variance	
bp	Base pair	
CO1	Cytochrome c oxidase subunit 1	
dNTP	Deoxyribonucleotide triphosphates	
D	Tajima's D	
DNA	Deoxyribonucleic acid	
Fs	Fu's Fs	
h	Haplotype diversity	
He	Expected heterozygosity	
Ho	Observed heterozygosity	
k	Mean number of pairwise differences	
µl	Microliter	
µM	Micromolar	10^{-3}mol/m^3
mtDNA	Mitochondrial DNA	
$MgCl_2$	Magnesium Chloride	
MSC	Messian Salinity Crisis	
Mya	Million years ago	
n	Number of samples	
Na	Number of alleles	
Nh	Number of haplotypes	
Np	Number of polymorphic sites	
SoM	Sea of Marmara	
TSS	Turkish Straits System	
ya	Years ago	

yr	Year
π	Nucleotide diversity
PCR	Polymerase chain reaction

1. INTRODUCTION

1.1. Marine Population Connectivity

A population can be considered the fundamental unit of evolutionary change and population genetics deals with the change in the genetic composition of populations over time. The main evolutionary factors acting on populations are mutation, migration (gene flow), selection, and genetic drift (Hartl & Clark, 1997). Migration, movement of gametes, extinction and colonization of entire populations, or the movement of non-nuclear DNA segments like plasmids, mitochondria and viruses result in gene flow. It can, on the one hand, act as a homogenizing force between populations and thus serve to keep the biological species integrated as a unit. In addition, gene flow can be an adaptive force by allowing populations to respond to environmental fluctuations (Slatkin, 1985). On the other hand, populations accumulate genetic differences more independently as the gene flow between them decreases.

Dispersal is permenant movement away from the place of origin. Marine species with planktonic larval stages generally have good dispersal capabilities. However, this does not always translate into extensive gene flow. In addition, the exchange of genetic material between populations does not guarentee that migrant individuals or gametes will survive to reproduce in the new population (Hedgecock, 1986). In marine invertebrates with planktonic larvae, gene flow can be affected by migration capability of the species (such as the duration of larval stage), physical factors (ocean currents, temperature, salinity, etc.), the type of coastal distribution of species, ecological barriers not correlating with any obvious physical barriers, capacity of pelagic larvae to affect its own dispersal (larval behaviour and swimming capability), and temporal variation in settlement patterns and genetic make-up of larval recruits (Hedgecock, 1986). In addition to these factors, historical events such as glaciations, previous bottlenecks, sea level fluctuations, and fragmentation can leave lasting signatures on the current population genetic structure (Patarnello et al., 2007). For a better understanding of intraspecific genetic structure, the information provided by genetic markers can be valuable.

1.2. Genetic Markers

1.2.1. Microsatellite DNA

Microsatellites are DNA regions which are segments of short repeats of mostly one to six nucleotides more frequently found in non-coding or intergenic regions (Metzgar et al., 2000). They are codominant markers so that heterozygotes can be distinguished from homozygotes, and are widely used in population genetic studies owing to their high polymorphism, their assumed neutrality and high mutation rates (Schlötterer, 2000). Their polymorphism provides better resolution in terms of revealing population structure. However, due to their high mutation rate, their use in related taxa may be limited because primer binding sites may not be conserved (null alleles) or loci having same lengths may not correspond to same evolutionary histories (homoplasy) (Selkoe & Toonen, 2006). Microsatellites mutate differently from other nuclear regions, therefore the main mechanism for their mutation is replication slippage in contrast to point mutations where a nucleotide is substituted by another (Ellegren, 2004). During DNA replication or repair, DNA polymerase slips so that one strand of DNA strand dissociates from the other and instantenously binds to another position (Figure 1.1).

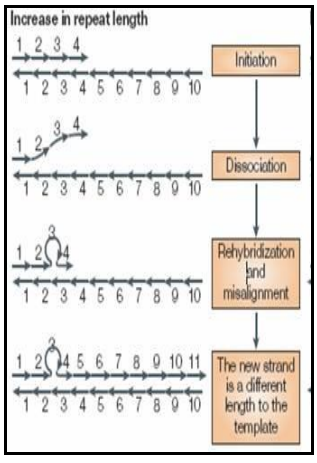

Figure 1.1. Main mutation mechanism of the microsatellites (Ellegren, 2004).

1.2.2. Mitochondrial DNA – CO1 Subunit

Due to its uniparental inheritance and its general lack of recombination, mitochondrial DNA (mtDNA) has been widely used in phylogenetic analyses (Avise, 2000). Since mtDNA has reduced effective population size, its haplotypes are more likely to get extinct due to genetic drift and this may lead to infer oversimplified population histories. In addition, since it is inherited only by females in most animals and plants, it gives information only about the dispersal of one sex. Because it has higher mutation rate than most nuclear genes (around 2 % per million years) and mostly conserved nucleotide order, it can be used to infer phylogenetic relationships (Avise, 2000). Still, mtDNA can have insufficient variation for understanding recent phylogeographic history of a species (Funk & Omland, 2003). The mitochondrial marker used in this study, CO1, has relatively slow evolution rate (1.66 % per million years) in crustaceans (Schubart et al., 1998).

1.2.3. Different Signals From mtDNA And Microsatellites

When populations first become reproductively isolated, they may share copies of the same ancestral alleles for a period of time. After a while, in neutral loci, as a result of genetic drift and the accumulation of mutations, alleles may become population specific leading to reciprocal monophyly so that all of the alleles in one population are more similar to each other than alleles in another population (Funk & Omland, 2003). However, the effect of genetic drift and the rate of mutations are not the same across all loci. For example, genetic drift affects nuclear loci to a lessser extent than mtDNA, due to its biparental inheritance and codominance. In contrast, mtDNA is haploid and uniparentally inherited, so that reproductively isolated populations are expected to reach monophly in mtDNA four times faster than in microsatellites. On the other hand, since microsatellites accumulate mutations faster than mtDNA in general, the information they provide for can help to infer more recent processes.

Furthermore, historical processes such as periods of geographic isolation followed by secondary contact can lead to cytonuclear discordance due to divergent patterns of gene flow

in mtDNA and nuclear markers (Toews & Brelsford, 2012). In this thesis, the term *cytonuclear discordance* is used to refer different mutation rates of mitochondrial DNA CO1 subunit and microsatellites. When populations have been recently diverged and found in sympatry, it may be especially hard to differentiate the roles of retention of ancestral polymorphisms or gene flow. An important clue is that gene flow after secondary contact will result in a geographic pattern in distribution of mtDNA clades among nuclear groups whereas retention of ancestral polymorphisms will not produce such a discernable biogeographical pattern (Funk & Omland, 2003). When different species are concerned, the introduction of alleles from one species to the gene pool of another is known as introgression. Currat *et al.* (2008) and Excoffier *et al.* (2009) showed that the direction of gene flow is generally from the local taxon to the incoming one during mtDNA introgression after secondary contact.

1.3. Biogeography and Paleogeography of the Study Area

The Mediterranean Sea is a relatively isolated sea with a connection to the Atlantic Ocean via the Straits of Gibraltar. Sicily Strait separates Eastern Mediterranean from the Western Mediterranean. The Red Sea has also been connected to the Mediterranean since its opening by humans in 1869 which brought up saline water and some species from the area (Patarnello et al., 2007). Surface ocean currents in the Mediterranean coasts of Turkey are shown in Figure 1.4. The Black Sea is a large, mostly isolated and deep body of inland water linked to the Mediterranean Sea through the Bosporus, the Sea of Marmara (SoM), and the Dardanelles (collectively known as 'The Turkish Straits System' (TSS)) and to the Azov Sea through the Kerch Strait (See Figure 1.2) (Zaitsev & Mamaev, 1997). Around 87 % of its waters are anoxic and it contains a hydrogen sulfide layer below 100 to 200 meters of the surface. Thus, deep pelagic and benthic organisms are very rare. Due to river runoff and precipitation, the water level in the Black Sea is higher than the SoM. The average salinity of its surface waters is around 17 and 18 psu (Zaitsev & Mamaev, 1997). The surface water circulation in the Black Sea is dominated by three anti-clockwise gyres corresponding to western, central and the eastern Black Sea (Figure 1.3).

Figure 1.2. Coastal waters of Turkey.

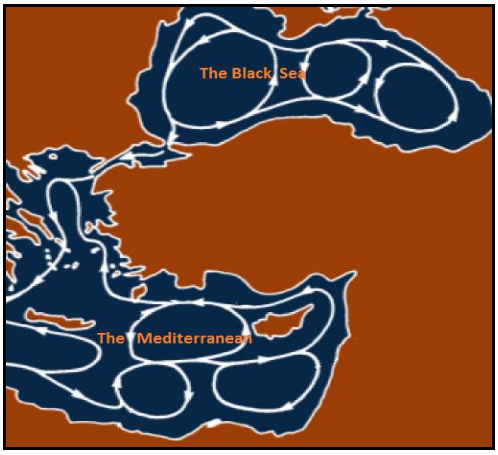

Figure 1.3. Surface ocean currents in the coasts of Turkey (modified from Tomczak & Godfrey (2001).

The SoM is an inland marine basin connected to the Mediterranean by the Dardanelles and to the Black Sea by the Bosphorus. The Dardanelles strait is more shallow, longer, and wider than that of the Bosphorus (Poulos et al., 1997). The SoM has been affected to a great extent from Pleistocene glacials and interglacials such that it has been occupied by waters of both the Mediterranean and the Black Sea in different periods (Görür et al., 1997). Due to the density differences between the less saline Black Sea water (18 psu) and highly saline Mediterranean water (38-39 psu), there is a stratification both in the Black Sea basin and in the TSS. While the cooler (5-15°C) and brackish (17-20 psu) Black Sea waters flow down to the Aegean on the upper layer, the warmer (15-20°C) and more saline (38-39 psu) Mediterranean waters flow in the opposite direction on the lower layer (Zaitsev & Mamaev, 1997)

The Messian Salinity Crisis (MSC) in the late Miocene (around 5.5 mya) has severed the connection of the Mediterranean Sea to the Atlantic Ocean via the Strait of Gibraltar, and the Mediterranean was turned into a series of saline lakes during that period. The Atlantic waters have invaded the Mediterranean at the end of the MSC. Since then, throughout the Quaternary, their connection has been interrupted for short periods and re-established corresponding to glacials and interglacials (Patarnello et al., 2007). These cycles have caused fluctuations in sea

levels in the Mediterranean and affected the distribution and abundance of marine populations (Magoulas et al., 2006).

Glaciatiation patterns have also been affecting the Black Sea geology over the past three million years. Its composition alternated between the isolated fresh water lake and brackish-marine environment due to several connections and splits with the Mediterranean. In the Pliocene, approximately 1,5 - 3 million years ago, the Black Sea was part of the Pontian Sea which was a fresh water lake spanning the present day Black Sea and Caspian Sea. At that time, the Black Sea was connected to SoM, but separated from the Mediterranean. During Riss-Würm Interglacial (100,000-150,000 BP), the Dardanelles opened for the first time and Karangat basin was formed, the salinity of which was higher than that of the current Black Sea. The Black Sea was mainly composed of marine fauna and flora during that time. In the end of Last Würm Glaciation, around 18,000 – 20,000 years ago, the connection between the Black Sea and the Mediterranean was lost again. During that time, both the Black Sea and the SoM were isolated from each other. By 15,000 years ago, Eurasian and Alpine glaciers retreated and resulted in the loss of meltwater resources to the Black Sea. Sea levels in the Mediterranean, on the other hand, has risen during that time and slowly invaded the Black Sea beginning from 9,400 BC (or 12-14 kya) (Zaitsev & Mamaev, 1997) (Figure 1.4).

In population genetics studies entailing the study area, a recurrent pattern is the differentiation of the Black Sea populations from those in the Mediterranean and the Atlantic, in different magnitudes (Patarnello et al., 2007). For example, in two studies of dolphins, both morphology, microsatellites and mtDNA pattern suggested that Black Sea population was differentiated from the western Atlantic populations (Fontaine et al., 2007; Viaud-Martínez et al., 2007) suggesting a long history of differentiation. Also, in a study on anchovy (Magoulas et al., 2006), lower genetic variability was observed in the Black Sea and this was attributed to the fact that Black Sea was an almost isolated sea and the surface current direction in the TSS is from the Black Sea towards the Mediterranean. In addition, in a study of plankton using both mtDNA and microsatellite markers, populations in the Black Sea were found to be differentiated from the Mediterranean and Atlantic populations (Peijnenburg et al., 2006).

Kalkan *et al.* (2011) has found no differentiation between the Black Sea and the SoM mussel (*Mytilus galloprovancialis*, Lamarck, 1819) populations by using microsatelites and mtDNA COIII subunit. On the other hand, by inspecting three marine invertebrate populations (*Mytilus galloprovancialis*, *Palaemon elegans* (Rathke 1837), and *Pachygrapsus marmoratus* (Fabricius, 1787)) around Turkish coasts, Kalkan (2013) found different haplotype frequencies between the Black Sea and Mediterranean populations in mtDNA corresponding to the Dardanelles strait in each species. In addition, in *Mytilus galloprovancialis* and *Palaemon elegans*, using mtDNA COIII subunit, different clades, one predominantly corresponding to the TSS and the Black Sea and another corresponding to the Mediterranean Sea were detected. Microsatellites did not detect any population structure in the mussel species. In terms of *Pachygrapsus marmoratus* populations, although haplotype frequencies were different in the Mediterranean and in North of the Dardanelles, the difference between main haplogroups were only one or two few base pairs, in addition the sample size in the Black Sea was low, decreasing the resolution of the inferences.

1.4. The Marbled Rock Crab, *Pachygrapsus Marmoratus*

1.4.1. Description and Life History

The focus species for this thesis, the marbled rock crab, *Pachygrapsus marmoratus* or "Marmara pavuryası" as its local name in Turkish (Dumitrache, 1999) has a squarish carapace up to 4.5 cm in width with three pointed teeth on each side (Figure 1.4). Its color ranges from a carapace of deep purple on a whitish background to blackish background, with a marbled yellowish-brown pattern with lighter coloured limbs (Sweet, 2011). It is difficult to catch due to its quick movements and flattened slippery carapace helping it to slip into very narrow crevices. It lives on the rocks of the middle to upper shores (Edwards, 2005).

Figure 1.4. An individual of *Pachygrapsus marmoratus* (Photo credit Ünsal Karhan)

1.4.2. Distribution and Habitat

P. marmoratus is one of the most abundant intertidal crab species throughout the Mediterranean basin (Flores & Paula, 2002), the fauna and flora of which are known to be mainly of Atlantic origin (Patarnello et al., 2007). Other than the Mediterranean, the distribution of *P. marmoratus* also includes the Black Sea, and the north-eastern part of the Atlantic Ocean from Morocco to the English Channel. Furthermore, it is present in all of the Turkish Seas (the Sea of Marmara, the Black Sea, the Aegean, and the Mediterranean) (Kocatas et al., 2004). Also, probably due to climate change, its distribution has been expanding northward (Dauvin, 2012). Figure 1.5 shows the approximate distribution of *P. marmoratus* (The Global Biodiversity Information Facility: GBIF Backbone Taxonomy), however it does not include all of the Turkish Black Sea coasts although the species is present throughout the region and also along the Turkish Mediterranean coasts (Kocatas et al., 2004).

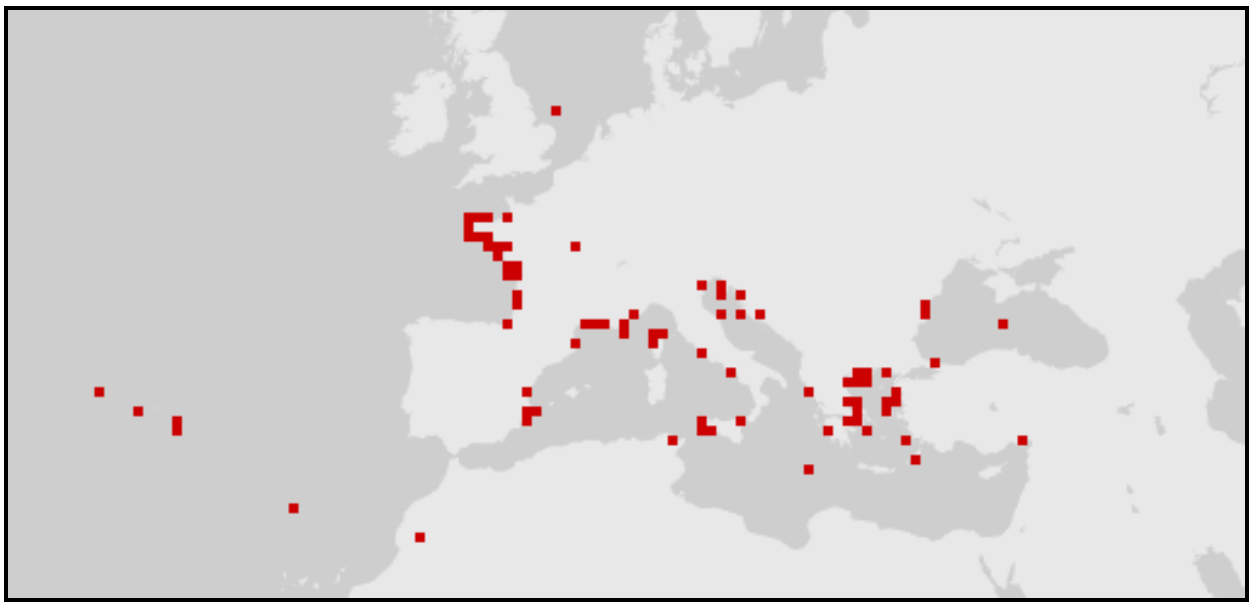

Figure 1.5. Distribution of *P. marmoratus*

1.4.3. Intertidal Life Style And Feeding Habits

The intertidal species are active during aerial phases, while they shelter themselves in crevices during high tide or avoid submersion by going away from the water's edge. Studies suggest that *P. marmoratus* is also a very adaptable rocky shore crab, capable of modifying its use of time and space according to the different characteristics of the shores it inhabits. The plastic behavior of the species seems to be mainly shaped by inter-related factors such as temperature, refuge availability, competition and escape from predators (Cannicci et al., 1999). *P. marmoratus* also has the capability to rely on many trophic sources, and therefore it can consume large amounts of algae based on its needs, and can also feed on invertebrates (Cannicci et al., 2007). As invertebrates, they mostly prey upon mussels and limpets, followed by conspecific crabs (Cannicci et al., 1999). Consequently, given that this species is abundant along the Turkish coasts, it is likely that it has an important role in the ecology of the local coastal habitats especially in terms of sessile populations.

1.4.4. Life History Traits and Their Role in Dispersal

P. marmoratus has a planktonic larval stage lasting about a month, and they can potentially disperse during that stage. In contrast, adults are mostly sedentary, inhabiting specific regions, which they protect from con-specific invasions (Cannicci et al., 1999). Breeding season of *P. marmoratus* extends from late spring to late summer (Fratini et al., 2011). Wide distribution range (Figure 3.1), long planktonic larval stage (lasting about one month), high abundance in rocky shores, capacity to consume various trophic sources (see section 1.2.3) and resilience to anthropogenic impact imply high dispersal ability for the species (Fratini et al., 2013). Although it has planktonic dispersal, populations of this species are predicted to have a smaller effective population size than the actual population size, as few individuals are inferred to be successful in reproducing at each reproduction cycle (Fratini et al., 2013; Hedgecock et al., 2007).

A previous study using mitochondrial cytochrome oxidase 1 (mtDNA CO1) subunit showed that geographically distant Mediterranean and Atlantic populations of *P. marmoratus* were not differentiated from each other. In terms of smaller geographical scales of settlement, hydrological features in coastal areas were inferred to be important whereas in terms of larger scales, oceanographic processes were likely to be involved in the dispersal and settlement of its larvae (Fratini et al., 2011). Studies which employed microsatellites (Fratini et al., 2013, 2011, 2008; Silva et al., 2009) found weak genetic differentiation between populations, sometimes partly corresponding to different properties of the seas they inhabit. Fratini et al. (2008) showed that its populations exhibited weak genetic structure, with the existence of three genetically distinct groups of populations corresponding to three parts of the Tuscany coastline; northern, central, and southern. Geographic barriers present along the Tuscany coast such as the Corsican Channel separating the two distinct Mediterranean basins were consistent with the clusters corresponding to Northern and Southern Tuscany. They also found many common genotypes among all the populations. The long larval dispersal phase of this species also supported their results, as a shared pool may contribute to each new generation of populations in the area. Another study (Fratini et al., 2013) also performed in the same area (in

the islands of Tuscan Archipelago) found slightly stronger differentiation among populations, and indicated a split into two groups in the Tuscan coast populations, which was weakly attributed to different biogeographic properties of the two seas of Ligurian and Tyrrhenian. Silva et al. (2009) found evidence for lack of panmixia throughout Portuguese coastline, with statistically significant genetic heterogeneity with no geographical clines. The results of this study suggested that the genetic patchiness observed by temporal variation in genetic composition of recruits coupled with currents, post-settlement selection or low effective population size. In addition, local retention of larvae was found to be not an important factor in explaining its genetic structure (Fratini et al., 2013).

2. OBJECTIVES

A principal objective of this study is to try to get a better understanding of the genetic structure in the intertidal crab species *Pachygrapsus marmoratus* around Turkish coasts using nuclear microsatellites and increasing the sample size and the sampling range in Kalkan (2013). Accordingly, five microsatellite loci and the mtDNA CO1 subunit, with to different evolutionary properties were employed with the main objectives of:

1. Determining contemporary population structure of *P. marmoratus* around Turkish coasts.
2. Combining information from mtDNA CO1 subunit and microsatellites to get a more comprehensive understanding of the role of Pleistocene glaciations in the Mediterranean Sea on the current population structure of the species.

3. MATERIALS AND METHODS

3.1. Sampling process

A total of 384 specimens of *P. marmoratus* were collected from 32 sites including nine sites from the Black Sea (94 specimens), seven from the TSS (95 specimens), 12 from the Aegean Sea (162 specimens) and four from the Levantine Sea (33 specimens) between the years 2010 – 2014. The geographical locations of crabs are shown in Figure 3.1, and their corresponding coordinates are provided in Appendix A, Table A1. The samples were collected manually at a depth range of 0-0.5 m, from the upper-littoral zone. The sampling sites consisted of pier pilings, breakwaters made of rock, and similar habitats. Nocturnal sampling was preferred because crabs were more active at night, making the collection of individuals easier. A LED torch with 150 lumen power was used during the sampling process to prevent crabs from escaping; when exposed to light, most crabs maintained their positions for a while. Subsequently, they either continued their activity or began to escape. The sampled individuals included those outside crevices or refuges.

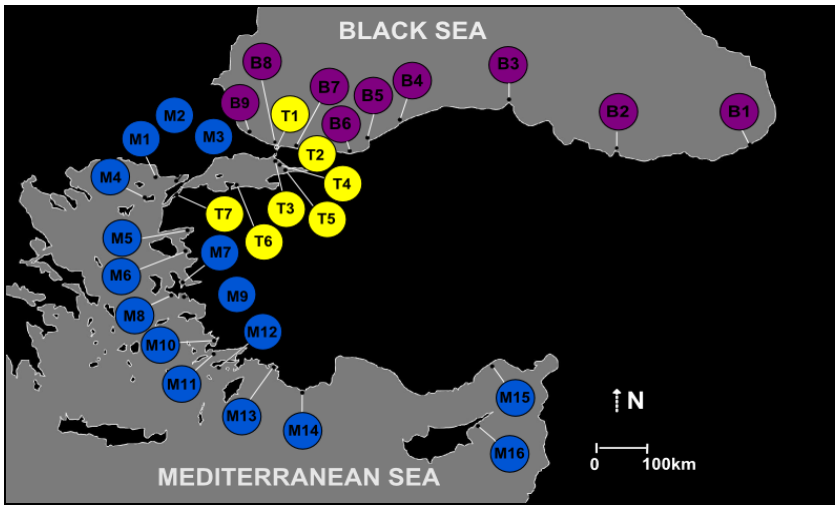

Figure 3.1.Sampling locations of *P. marmoratus* around Turkish coastal waters.

3.2. Isolation of DNA and PCR Amplification

After collection, crab samples were immediately put into 80-99% ethanol. In the laboratory, they were preserved in 95-99% ethanol. For DNA extraction mostly gill tissues and for a few individuals, pereipod muscle tissues were used. Removed tissues were washed with distilled water before beginning the extraction to remove residual ethanol. Genomic DNA was extracted using either Roche High Pure PCR Template Preparation Kit (Indianapolis, USA) or Invitrogen Kit (Carlsbad, CA) following the manufacturer's protocol.

3.2.1. Microsatellite Loci Amplification

A total of 384 individuals were amplified at five microsatellite loci (Table 3.1), designed by Fratini et al. (2006). Forward primer for each locus was 5'-labelled with one of the four fluorophores (VIC, HEX, FAM, NED).

Table 3.1. Nuclear microsatellite markers used in the study.

Locus (EMBL accession #)	Forward primer	Reverse primer	Repeat motif	Size range (bp)
Pm79 (DQ155410)	HEX-CATGCCCAAAGAGAAACACA	TGCACGCCCGTTATCACTA	$(TG)_nAG(TG)_n$	212-216
Pm99 (DQ155418)	FAM-CGGGAATTGTGTTTGAGTT	AGGCACGTGGACAGAGAGA	$(TC)_n$	289-331
Pm101 (DQ155416)	FAM-GGCTGAGGAGAATGAGTGA	GGCAGCTATAGCGTGTATGC	$(TC)_n$~$(TC)_n$	206-238
Pm108 (DQ155414)	VIC-CTATCCCTACTGCTGCTGCT	CCACCCTCGGCTGTTTATTA	$(TGA)_n$	201-210
Pm183 (DQ155423)	FAM-ATCACCTTGCCTCCCTCTTT	GGAGCATCGATTGGGTCAC	$(CA)_nCC(CA)_n$	206-264
Pm187 (DQ155419)	NED-AGCCTGAGGAGAGCCAACAC	CTATCGGCCGCATCTTTAT	$(CA)_n$	180-190

Four of the six loci (Pm101, Pm108, Pm187, and Pm99) were amplified in a multiplex reaction undertaken using QIAGEN Multiplex PCR Kit (Hilden, Germany) in a volume of 12.5 µl, containing 1 µl of DNA, 6.25 µl multiplex mix, 0.25 µl of each primer (10 µM), and 4.75 µl distilled water following the manufacturer's protocols. PCR conditions for these four loci are shown in Table 3.2.

Table 3.2. PCR conditions of four primers (Pm101, Pm108, Pm187, and Pm99).

PCR steps	Duration	Temperature	
Inital Denaturation	15 minutes	95°C	
Denaturation	30 seconds	94°C	35 cycles
Annealing	90 seconds	60°C	
Extension	60 seconds	72°C	
Final Extension	30 minutes	60°C	

The other two loci, Pm79 and Pm183, were amplified according to PCR conditions used by Fratini et al. (2006), with some modifications. The final PCR reactions were undertaken so that each of 20 µl of reaction volume contained 1.2 µl DNA (for Pm79) or 1.6 µl DNA (for Pm183), 0.5 µl of each primer (10 µM), 2 µl (2.5mM) of $MgCl_2$, 0.4 µl of dNTPs (10 mM), and 0.1 µl of Taq DNA Polymerase (5 U/µl). Their PCR thermal profile for each locus is shown in Table 3.3.

Table 3.3. PCR conditions of primers Pm79 and Pm183.

PCR steps		Duration		Temperature		
Pm79	**Pm183**	**Pm79**	**Pm183**	**Pm79**	**Pm183**	
Inital Denaturation		5 minutes		94°C		
Denaturation		30 seconds		94°C		
Annealing		60 seconds		53°C	58°C	35 cycles
Extension		180 seconds		72°C		
Final Extension		60 minutes		72°C		

3.2.2. mtDNA CO1 Amplification

New samples added to Evrim Kalkan's previous collection (Kalkan, 2013) were also amplified for the CO1 region. A total of 136 PCR reactions were performed, each in a 25 µl reaction volume containing 2.5 µl of 10X high fidelity buffer, 2.5 µl of $MgCl_2$ (25 mM), 0.5 µl of dNTPs (10 mM), 0.3 µl of each primer (20 µM) (Table 3.4), 0.25 units of Taq DNA polymerase (5U/µl), 2 µl of DNA (approximately 50 ng/µl), and 18,45 µl of H_2O. The PCR amplifications were performed with the following conditions: 10 min at 94°C for initial denaturation followed by 35 cycles of 45s at 94°C for denaturation, 1 min at 50°C for annealing, and 1 min at 72°C for extension, followed by a final extension of 10 min at 72°C.

Table 3.4**.** CO1 primers used for PCR amplification and sequencing.

Name	Sequences from 5' to 3' end	Reference
LCO1490	GGTCAACAAATCATAAAGATATTGG	Folmer et al. 1994
HC02198	TAAACTTCAGGGTGACCAAAAAATCA	Folmer et al. 1994

3.2.3. Visualization and Analysis of PCR Products

Approximately half of the PCR products were visualized by agarose gel electrophoresis before the fragment analysis or sequencing to confirm that PCR reactions worked. PCR products were run on a 1 % agarose gel for mtDNA COI fragments, and on a 2 % gel for microsatellite fragments. For microsatellites, the solution was prepared by mixing 1 grams of agarose in 100 ml of 1X TAE buffer (0.04 M Tris – Acetate 0.001 M EDTA), and melting it in a microwave until agarose granules disappeared. For mtDNA, 0.5 grams of agarose was used instead. After the gel cooled down a little, 2.5 ul of Ethidium Bromide was added, and the solution was casted in a tray with combs. Upon solidification of the gel, PCR products were loaded into the wells by mixing with equal volumes of loading buffer. The gels were run for 15 minutes at 120V and visualized under UV light. Pm101, Pm108, Pm187, and Pm99 were sent to Cornell University, and Pm79 and Pm183 were sent to Macrogen Holland for the sizing of the fragments commercially. Fragments were scored manually at first using Peak

Scanner v.1 (Applied Biosystems, USA), and then also by GeneMarker v.2.6.3 (SoftGenetics, USA).

3.3. Statistical Methods

3.3.1. Microsatellite DNA

3.3.1.1. Null Alleles and Hardy-Weinberg Equilibrium. The allele sizes deduced from the visual analysis of the peaks with the softwares GeneMarker and Peak Scanner were binned automatically using the excel macro FlexiBin (Amos et al., 2007). The genotyping errors due to null alleles, large allele drop-out or the scoring of stutter peaks that can potentially lead to deviations from Hardy–Weinberg proportions were analysed using MICRO-CHECKER (Van Oosterhout et al., 2004). In order to compensate for the low sample sizes when present, samples from geographically proximate locations were pooled. In addition, STRUCTURE v.2.3.4 (Pritchard et al., 2000) analysis of nuclear microsatellite data was used to further divide populations according to whether they belonged to one of the two structure clusters (see section 4.1.2 below). In this way, a total of 12 subpopulations were defined for comparisons (Appendix A, Table A1). To estimate null allele frequencies, the Equation Brookfield 1 (Brookfield, 1996) was used according to the flow chart in the Help section of MICRO-CHECKER based on the properties of the data.

Out of 384, a total of 351 individuals were used in the subsequent analyses. The individuals having missing data in at least three loci out of six were removed from the data set. Also, the locus pm183 was removed from the analyses, due to its non-uniform allele distribution with a lack of alleles of intermediate sizes, and its compound character (Estoup et al., 2002). Hardy-Weinberg Equilibrium (HWE) was tested in Genepop version 4.2 (Raymond & Rousset, 1995; Rousset, 2008) for each locus and each population using an exact test with a dememorisation number of 10000, and 10000 Markov chain iterations for 100 batches in 12 pooled populations. Linkage Disequilibrium (LD) was also tested in Genepop version 4.2

(Raymond & Rousset, 1995; Rousset, 2008) for all pairs of loci for each population using 10000 dememorisation steps and 10000 Markov chain iterations per 1000 batches.

3.3.1.2. Clustering Method. STRUCTURE (Pritchard et al., 2000) was used to determine the number of differentiated clusters in the data set. This program estimates the likely number of clusters (K) using multiloci data via minimizing within group Hardy-Weinberg and linkage disequilibrium, and it assigns the probability of each individual to predefined clusters. Several runs were performed with different burn-in lengths and MCMC repeats (50000, 10000, and 20000 for each). These runs were performed with different K values ranging from 1 to 12. To infer the number of clusters (K) which fits data the best, the program STRUCTURE HARVESTER (Earl & vonHoldt, 2012) was used. This program uses the 'Evanno method' (Evanno et al., 2005) to detect the number of groups (K). Once the best K was inferred with STRUCTURE HARVESTER, 20 iterations were performed using this value of K. Then, mean Q values for these iterations were calculated with CLUMPAK (Cluster Markov Packager Across K, (Kopelman et al., 2015) by using a Markov clustering algorithm that relies on a similarity matrix between replicate runs, as computed by the software CLUMPP (Cluster Matching and Permutation Program) (Jakobsson & Rosenberg, 2007). Afterwards, with CLUMPAK, a DISTRUCT (Rosenberg, 2004) graph was produced for the graphical display of the results. In all of the analyses, uncorrelated allele frequencies were assumed allowing for admixture.

3.3.1.3. Ordination Method. Principal Coordinate Analysis (PCoA) was performed by GenAlEx v.6.5 (Peakall & Smouse, 2006, 2012) based on the distance method of Peakall & Smouse (1999). In this method, eigenvalues and eigenvectors (major axes of variation) are calculated from similarity and dissimilarity matrices. Then, this distance matrix is used to create a PCoA graph by choosing the covariance-standardized method, as implemented in GenAlEx (Peakall & Smouse, 2006, 2012). The analysis was made based on both the two main Structure clusters, and using the 12 pooled populations.

3.3.1.4. Genetic Variation and Population Genetic Differentiation. The number of alleles (Na), the effective number of alleles (Ne), the expected heterozygosity (the proportion of heterozygosity expected under random mating) (He), unbiased expected heterozygosity (uHe), observed heterozygosity (the proportion of heterozygous samples to sample size at a given locus) (Ho), number of private alleles (Pa), and Shannon's information index (I) were calculated using GenAlEx. In calculating allelic richness, allelic richness (RS) and private allele richness independent from sample size (PR) were computed by the rarefaction method (Hurlbert, 1971) as implemented in HPRARE software (Kalinowski, 2005), as sample sizes of populations were different from each other. These calculations were performed for each of 12 pooled populations and also separately for the two structure clusters.

To perform a multilocus analysis of molecular variance (AMOVA, Excoffier et al., 1992), populations were first grouped according to their Structure clusters as population M (Mediterranean only) and population C (Common, all around Turkey's coasts). AMOVA analysis using ARLEQUIN v.3.5 (Excoffier & Lischer, 2010) was performed both for the two and 12 population data sets. The aim was to partition the relative amount of genetic differentiation within populations, among populations within structure clusters, and between structure clusters. These 12 pooled populations were also used to compute pairwise F_{ST} values and their corresponding significance values. F_{ST} value showing variance among subpopulations relative to the total population can be between zero-no differentiation and one-complete differentiation (subpopulations fixed for different alleles) (Wright, 1954). Significance values were tested with 10000 permutations. Bonferroni correction (Rice, 1989) was used to adjust the significance of P values, such that since there were 12 populations and five loci, P value of 0.05 was divided by 60 (12x5) and the resulting P value of 0.00083 was used to call significance of pairwise comparisons.

3.3.2. Mitochondrial DNA – CO1 Subunit

Sequence data from 187 individuals were combined with 129 individuals sampled and analysed previously by Kalkan (2013), and 38 CO1 fragment sequences from GenBank with accession numbers JF930650.1-JF930682.1, JQ305921.1, JQ306087.1-JQ306090.1 were also included. Sequences were edited manually with the software Sequencer 5.2.4 (2013 Gene Codes Corporation, Ann Arbor, USA) and aligned with the software ClustalX (Thompson et al., 1997). The length of sequences used in the final analyses was 490 bp.

For each of the 12 populations pooled based on the microsatellite data, haplotype diversity (h) and nucleotide diversity (π) were calculated with the software DnaSP v5.0 (Librado & Rozas, 2009). For the pairwise Φ_{st} analysis and AMOVA, software ARLEQUIN v.3.5 was used. A total of 15 individuals having Structure assignment probabilities (Q values) lower than 70% were omitted from pairwise Φ_{st} mtDNA analysis. Also for 33 individuals out of 335, PCRs did not work. As a result, 302 individuals were compared in terms of pairwise Φst values. Differently from pairwise F_{ST} calculations for the nuclear data where heterozygosities were used to calculate the F_{ST} values, nucleotide diversity (π) was used for Φ_{st} calculations made by using haploid mitochondrial DNA. The distances between haplotypes were calculated using Tamura and Nei distances (Tamura & Nei, 1993), as it is the most inclusive substitution model, which accounts for different transition and transversion rates, and also different transition rates between purines and pyrimidines (Laurent Excoffier & Lischer, 2010). A haplotype network was constructed with the software Network v.4.6.1.2 (Bandelt et al., 1999) that uses the median-joining algorithm with sequences from Genbank and those generated in this study (Figure 4.6).

Deviations from neutral molecular evolution were tested with ARLEQUIN, by calculating Tajima's D (Tajima, 1989) and Fu's F_S (Fu & Li, 1993; Fu, 1997) . Significant negative values of these tests may suggest population expansion. Also, R_2 test (Ramos-Onsins & Rozas, 2002) was applied using DnaSP v.5.0 (Librado & Rozas, 2009) to detect population expansion, if any. Statistical significance was tested with coalescent simulations based on 5000 simulated

resampling replicates. DnaSP was also used to calculate mismatch distribution plots of populations based on the number of differences between pairs of DNA sequences.

4. RESULTS

4.1. Microsatellites

4.1.1. Descriptive Statistics

As explained in the section 3.3.1.1, a total of 12 subpopulations were defined to compare based on proximity in location and STRUCTURE clusters (see section 4.1.2, below, for details of STRUCTURE results). In terms of linkage disequilibrium, for each pair of loci across all subpopulations, Pm99 and Pm101 were linked (*P<0.1*) according to Fisher's exact test without Bonferroni adjustment. However, after Bonferroni correction, no linkage disequilibrium was observed either across all populations or in separate subpopulations in any pairs of loci. According to HWE test performed on 12 pooled subpopulations (Table 4.1), after Bonferroni correction, the locus Pm99 was significantly out of HWE in all of the subpopulations inhabiting the Mediterranean (C7, C8, C9, M7, M8, and M9). Pm101 was out of equilibrium in C5, C9, and M8. Lastly, Pm187 was out of equilibrium in the subpopulation C4. The loci deviating from HWE in the whole data set indicated the existence of null alleles. Although there were null alleles in C6 and C1 at loci Pm187, Pm99, and Pm79, they did not result in any Hardy-Weinberg disequilibrium. According to the Brookfield equation 1 (Van Oosterhout et al., 2004), the frequency of null alleles was higher than 20% for subpopulations C7, C8, C9, and M9 at locus Pm99. The two populations defined by STRUCTURE (Table 4.1) deviated from HWE, again probably due to the presence of null alleles. The frequency of null alleles was higher than 10 % at locus Pm99 in the M cluster (19 %). In the C cluster, the frequency of null alleles was higher than 10 % at loci Pm99 (13 %), Pm187 (13 %), and Pm101 (10 %). Stuttering was suggested as a possible cause at Pm79 and Pm187 in population C, and at Pm99 in population M. Genotypic errors due to large-allele drop-out were not observed in either population.

Table 4.1. The loci with significant null alleles in two clusters (C and M) and in 12 subpopulations together with significant loci in terms of HWE test. Percantages of loci with more than 20 % null alleles are in bold.

Pop	Null allele (Frequency according to Brookfield Eq. 1)				HWE exact test			
C1		Pm79 (10%)	Pm187(15%)					
C2								
C3								
C4	Pm99 (11%)		Pm187 (15%)				Pm187	
C5	Pm99 (12%)	Pm101 (13%)	Pm187 (15%)			Pm101		
C6		Pm101 (9%)	Pm187 (14%)					
C7	Pm99 **(25%)**	Pm101 (13%)	Pm187 (15%)		Pm99			
C8	Pm99 **(31%)**	Pm101 (17%)			Pm99			
C9	Pm99 **(24%)**	Pm101 (15%)	Pm187 (20%)		Pm99	Pm101		
C	Pm99 (13%)	Pm101 (10%)	Pm187 (13%)	Pm79 (5%)	Pm99	Pm101	Pm187	Pm79
M7	Pm99 (18%)	Pm101 (6%)	Pm187 (10%)		Pm99			
M8	Pm99 (15%)	Pm101 (10%)			Pm99	Pm101		
M9	Pm99 **(29%)**				Pm99			
M	Pm99 (19%)	Pm101 (8%)	Pm187 (5%)		Pm99	Pm101	Pm187	

Pm 108, with a simple trinucleotide repeat, differed between the two populations in allele frequency distributions (Figure 4.1). Two other loci, Pm99 and Pm187, have simple, dinucleotide repeats. Between two populations, allele frequency distribution of Pm99 differed whereas that of Pm187 did not. Pm101 and Pm79 have compound diallelic repeats. The locus Pm 79 had different allele frequency distributions in the two populations: population C had one predominant allele size (214 bp); on the other hand the population M had alleles of 212 bp and 214 bp in similar frequencies. The other compound microsatellite locus, Pm101, showed similar allelic frequency distributions in the two populations. Pm99 and Pm101 were the most polymorphic loci, with 21 and 17 alleles, respectively. Next, pm187 had six alleles, followed by Pm108 with four alleles. Pm79 was almost diallelic (allele sizes of 212 and 214) with only one individual with the allele size of 216 bp.

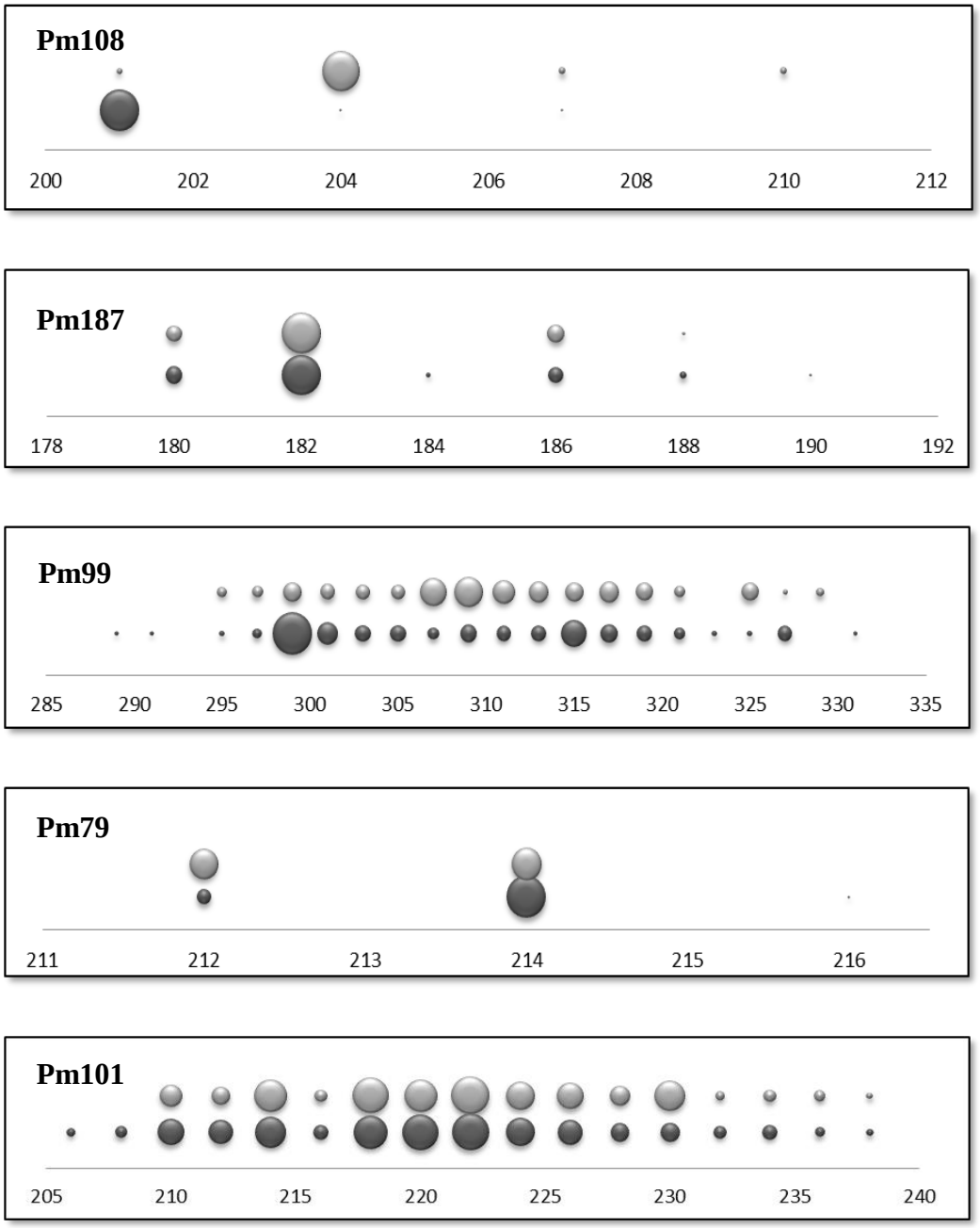

Figure 4.1. Allele frequency distributions of five microsatellite loci. In each figure, the upper row belongs to the population M and the lower row belongs to the population C. Numbers in the x-axis correspond to base pairs. The diameter of each circle is proportional to the number of individuals with the corresponding allele.

For both nuclear Common (C) and Mediterranean (M) populations, the loci Pm99 and Pm101 were highly polymorphic, having between 15 – 20 alleles per population. The other three loci were also polymorphic in both populations, but had a smaller number of alleles, ranging between two to six. Average sample sizes of population C and M were 206.4 and 113, respectively. For all loci, population C had more alleles than population M (9.8 vs. 8.4, respectively). In both populations, M and C, the effective number of alleles (Ne) was maller than actual number of alleles. This implies that allele frequencies were not evely distributed, as can also be seen in the allelic frequency distributions plot above (Figure 4.1). However, average Ne of population M was slightly higher than that of population C (5.13 vs. 4.26 respectively), suggesting that the former had more equal allele frequencies when compared to latter. In terms of allelic richness (RS) and the number of private alleles (PR), population C had slightly higher values than the population M (Table 4.2).

Considering the 12 subpopulations pooled separately, it was observed that subpopulations in the north of the Dardanelles (C1-C6), together with North Aegean subpopulation C7 and Levantine subpopulation C9 were fixed for the locus Pm108. Subpopulations M8 and C1 had the highest Ne (5.13) and the lowest Ne (3.01), respectively (Table A2). In terms of allelic richness, C2 had the highest RS (5.54) and PR (0.56), whereas C9 had the lowest RS (4.83) and PR (0) values (Table A2). Expected heterozygosity (He) values of the subpopulations of population C in the north of the Dardanelles (C1-C6) ranged between 0.42-0.45, and those of subpopulations in the Mediterranean coasts (C7-C9) ranged between 0.49-0.54. On the other hand, population M had higher He values than population C, ranging between 0.56-0.58 (Table A2).

Table 4.2. Genetic variability of the five microsatellite loci for *P. marmoratus* in two populations. N, sample size; Na, number of alleles; RS, allelic richness; Ne, Number of effective alleles; I, Shannon's diversity index; Ho, observed heterozygosity; PR, number of private alleles; He, expected heterozygosity; uHe, unbiased expected heterozgosity; F, fixation index.

Population		**Pm99**	**Pm101**	**Pm108**	**Pm187**	**Pm79**	**Av.**
Common	**N**	189	203	216	211	213	206.4
	Na	20	17	3	6	3	9.8
	RS	17.85	16.49	1.91	5.38	2.46	8.82
	Ne	7.01	10.23	1.01	1.77	1.27	4.26
	I	2.38	2.49	0.03	0.87	0.38	1.23
	Ho	0.61	0.71	0.01	0.25	0.15	0.34
	PR	2.47	2.01	0	1.52	0.46	1.29
	He	0.86	0.9	0.01	0.44	0.21	0.48
	uHe	0.86	0.9	0.01	0.44	0.21	0.48
	F	0.29	0.21	0	0.42	0.31	0.25
Mediterranean	**N**	115	119	98	114	119	113
	Na	17	15	4	4	2	8.4
	RS	16.85	14.79	4	3.86	2	8.3
	Ne	11.23	9.53	1.16	1.72	1.99	5.13
	I	2.58	2.39	0.34	0.77	0.69	1.35
	Ho	0.56	0.74	0.12	0.34	0.48	0.45
	PR	1.47	0.3	2.09	0	0	0.77
	He	0.91	0.9	0.14	0.42	0.5	0.57
	uHe	0.91	0.9	0.14	0.42	0.5	0.57
	F	0.39	0.17	0.1	0.18	0.04	0.18

4.1.2. Cluster Analysis and Geographical Distribution

STRUCTURE analysis of nuclear microsatellite data without *apriori* classification of populations predominantly (in 2/20 runs) resulted in discrimination of two populations. Q values specifying the assignment percentages of individuals to each cluster were higher than 90 percent in most of the individuals, and there were only few indivudals which showed mixed ancestry. One of the inferred populations, which was present all around the Turkish coasts, was the only one in the Black Sea and the Turkish Straits System (TSS) except in the Dardanelles. This cluster will be referred to as the 'Common' population or population C from

this point forward (Figure 4.2). Population C appeared in low frequencies throughout the Mediterranean Sea, except the Eastern Mediterranean including Kaş, Mersin, and Kıbrıs (n=1) where it became the dominant population again. The geographical range of the second population was limited to the Mediterranean coasts of Turkey, and is referred as the 'Mediterranean' population or population M (Figure 4.2). This cluster was not observed in the Sea of Marmara (SoM), the Bosphorus and the Black Sea, thus, it did not extend into the north of the Dardanelles. Throughout the Mediterranean coasts, this population was never the only population present in sampling points except few sites with low sample sizes (Fethiye, Seferihisar, Çeşme and Ayvalık where sample sizes are one or two individuals). Between the Dardanelles and Fethiye, population M was observed more frequently than population C. As mentioned above, from Kaş and to the east, the pattern was reversed and the frequency of population C became higher than that of population M.

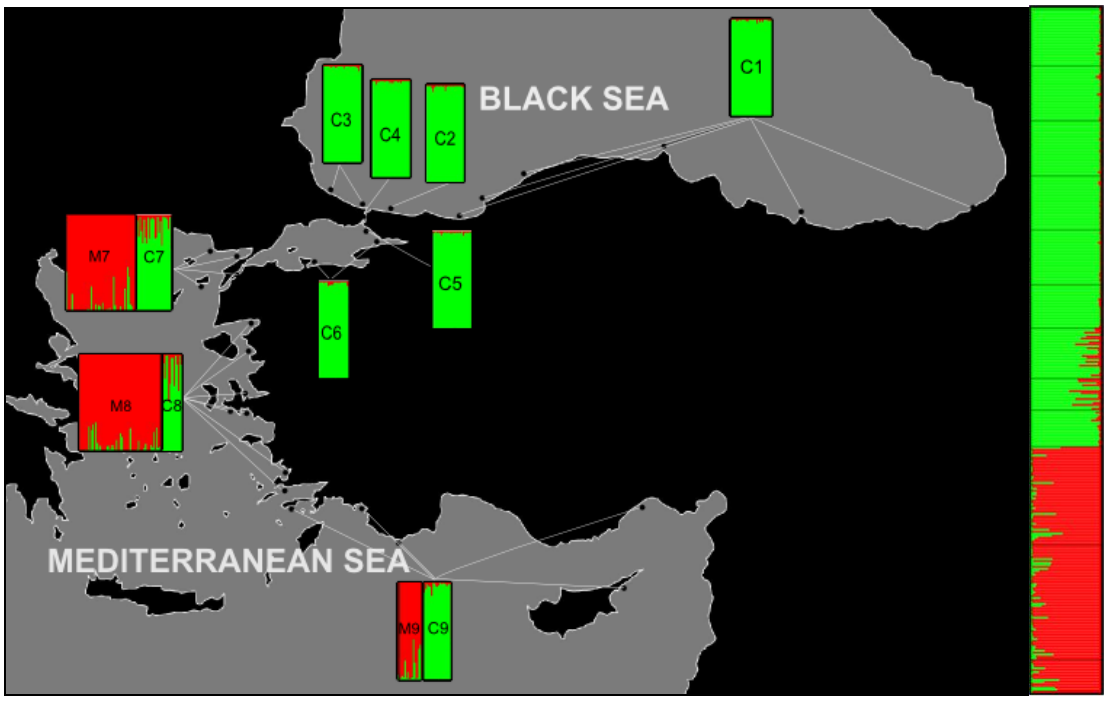

Figure 4.2. STRUCTURE results at K=2 for *P. marmoratus* on the right. Colors indicate percantage contribution of individuals to assigned clusters (x axis), individuals represented by each line (y axis), black lines separate 12 subpopulations. The populations M and C are shown in red and green, respectively. The geographical distribution of subpopulations is on the map.

4.1.3. Ordination Method

In a parallel manner to STRUCTURE, the results of the Principal Coordinates Analysis (PCoA) showed that two populations were separated from each other (Figure 4.3). The individuals of population C were closer to each other (Figure 4.3), suggesting that its individuals were more similar to each other in terms of their genetic distances. On the other hand, Coordinate 2 separated each population into three 'clusters' but they didn't correspond to any immediate geographical information.

In the PCoA results for the 12 pooled subpopulations, three clusters were observed (Figure 4.4). First, population M was separated from population C. In addition, population C was divided into two groups, one corresponding to the Black Sea and the TSS (C1-C6), and the other to the Mediterranean (C7, C8, and C9). Also, the clustering of C1-C6 was more compact than the other two clusters (C7-C9 and M7-M9) (Figure 4.4).

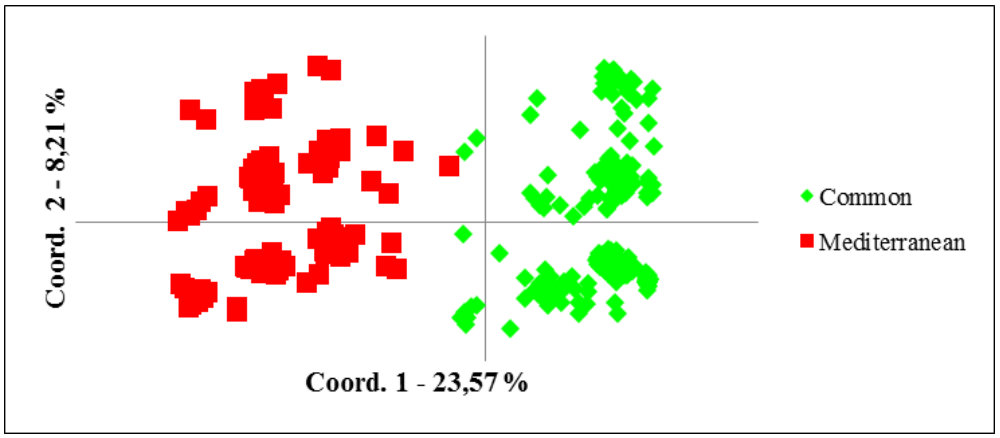

Figure 4.3. PCoA of *P. marmoratus* individuals coloured based on two STRUCTURE clusters. Each square indicates an individual.

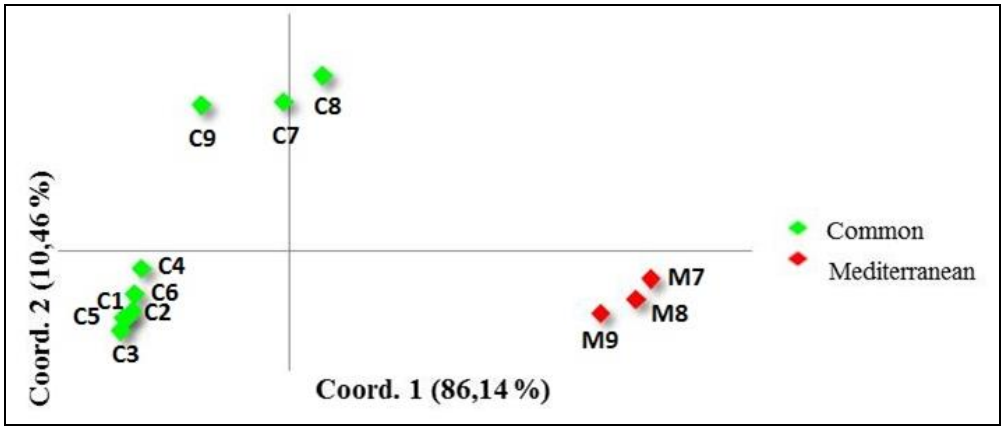

Figure 4.4. PCoA of 12 subpopulations of *P. marmoratus.*

4.1.4. Genetic Differentiation

In terms of the pairwise F_{ST} values, the highest differentiation was between population M and population C. The subpopulations M7 (North Aegean) and M9 (Levantine) were most differentiated from population C, with average F_{ST} values of 0.304 and 0.302, respectively (Table 4.3). On the other hand, subpopulations M7, M8, and M9 were not differentiated from each other. This pattern is also concordant with the results of STRUCTURE and PCoA that grouped these three subpopulations together as population M. However, population C was not genetically homogeneous, in that its subpopulations inhabiting the Mediterranean Sea (C7-C9) were slightly differentiated from those inhabiting the Black Sea, the Bosphorus and the SoM (C1-C6). C8 and C7 were differentiated from the subpopulations C1-C6 with average F_{ST} values of 0.15 and 0.08, respectively. C9 was differentiated from C3, C2, and C5 only, in the order of decreasing F_{ST} values with an average of 0.08. In addition, subpopulations of C to the north of the Dardanelles (C1-C6) were not differentiated from each other. Subpopulations of population C in the Mediterranean (C7-C9) were also not significantly differentiated from each other.

Table 4.3. Pairwise comparisons of nuclear genetic differentiation estimated from the pairwise F_{ST} values (below diagonal) and of mitochondrial genetic differentiation (Φ_{ST}) estimated from sequence divergence data (above diagonal, see section 4.2 for more detailed resuts of mtDNA analysis). Significant values are in bold (*P* threshold <0.00083 for F_{ST} and <0.05 for Φ_{ST}).

	C1	C2	C3	C4	C5	C6	C7	C8	C9	M7	M8	M9
C1	0	-0.01	-0.01	-0.01	-0.01	-0.03	**0.54**	**0.34**	**0.61**	**0.59**	**0.54**	**0.69**
C2	0.01	0	-0.03	0.00	0.00	-0.02	**0.44**	**0.21**	**0.50**	**0.52**	**0.47**	**0.59**
C3	-0.01	-0.01	0	0.00	0.00	-0.01	**0.44**	**0.21**	**0.50**	**0.52**	**0.47**	**0.59**
C4	0.00	0.01	-0.02	0	-0.02	-0.02	**0.52**	**0.32**	**0.59**	**0.59**	**0.53**	**0.67**
C5	0.01	0.01	0.00	0.01	0	-0.02	**0.54**	**0.34**	**0.61**	**0.60**	**0.54**	**0.69**
C6	0.00	-0.01	0.00	0.01	-0.01	0	**0.53**	**0.35**	**0.61**	**0.59**	**0.53**	**0.70**
C7	**0.06**	**0.09**	**0.09**	**0.06**	**0.09**	**0.07**	0	0.04	-0.05	-0.02	-0.03	0.18
C8	**0.14**	**0.14**	**0.16**	**0.13**	**0.17**	0.11	0.03	0	0.09	0.15	0.10	-0.03
C9	0.07	**0.08**	**0.09**	0.06	**0.08**	0.07	0.02	0.04	0	-0.04	-0.04	-0.04
M7	**0.32**	**0.34**	**0.34**	**0.32**	**0.34**	**0.32**	**0.23**	**0.25**	**0.28**	0	-0.01	-0.04
M8	**0.29**	**0.30**	**0.31**	**0.29**	**0.30**	**0.29**	**0.21**	**0.23**	**0.25**	0.00	0	-0.03
M9	**0.32**	**0.33**	**0.34**	**0.32**	**0.35**	**0.32**	**0.23**	**0.23**	**0.28**	0.00	-0.01	0

An analysis of molecular variance (AMOVA) was performed using ARLEQUIN v.3.5 (Excoffier & Lischer, 2010) to see the contribution of hierarchical levels of grouping to the variance in the data. As a result, F statistics were significant at all levels and the largest percentage of variance was explained by variation within 12 pooled populations (68 %) (Table 4.4). The differentiation among 12 subpopulations was 0.32 (F_{ST}). In addition, F_{CT} value shows differentiation among groups (population C and population M) relative to the total population. The F_{CT} value of 0.30 corresponded to 30.2 % of the variance explained due to the difference of two groups. F_{SC} value was 0.03 within subpopulations in each group corresponding to 1.84 % of the variance explained.

Table 4.4. Analysis of molecular variance for distribution of genetic variation using five microsatellite loci. Sum of squares (SS), variance, percentage of total variation, and F statistics are shown. Significant *P* values are in bold.

Source of variation	SS	Variance components	Variance (%)	F statistics	P value
Among groups	150.87	0.56	30.21	$F_{CT} = 0.30$	**< 0.001**
Among populations/groups	28.69	0.03	1.84	$F_{SC} = 0.03$	**< 0.001**
Within populations	743.82	1.25	67.95	$F_{ST} = 0.32$	**< 0.001**
Total	923.38	1.84			

4.2. Mitochondrial DNA – CO1 Subunit

4.2.1. Descriptive Statistics

With the addition of haplotypes from Genbank, the software DnaSP indicated a total of 50 haplotypes, 31 of which were present in the sampled area of this study. Among 490 nucleotides, 44 polymorphic sites with a total of 46 mutations were observed. Among these 27 were singleton variable sites and 17 parsimony informative. The three most common haplotypes detected in the sampled individuals around Turkish coasts were H9 (51 %), H1 (31 %), and H2 (0.07 %). The haplotype network showed a star-like where most of the haplotypes were derived from the most common haplotypes with distances corresponding to one or two bp. There were only two bp difference between the two most common haplotypes (H1 and H9), corresponding to a 0.4% divergence (2/490 bp). GC content of the data was 38%.

The individuals with the H9 haplotype were present throughout the Turkish coasts. They were in the Black Sea, the Bosphorus, the SoM with high frequencies and in the Mediterranean with lower frequencies (Figure 4.5 and Figure 4.7). H9 and its closest halotypes (haplogroup H9) were not present in the Atlantic Ocean. The range of the second most common haplotype (H1) encompassed again almost all of the Turkish coasts. However it was mainly present in the Aegean Sea followed by the Mediterranean, and it had low frequencies in the Black Sea and in the SoM. Haplogroup H1 (Figure 4.5) also included individuals from the Atlantic Ocean coasts of Portugal and Fuerteventura and the Mediterranean coasts of Italy,

Spain, Corsica, and Morocco (Fratini et al., 2011). Although it was the most frequent haplotype in the Eastern Mediterranean populations, its frequency was not as high as that of H9 in the Black Sea and in the SoM (Figure 4.7).

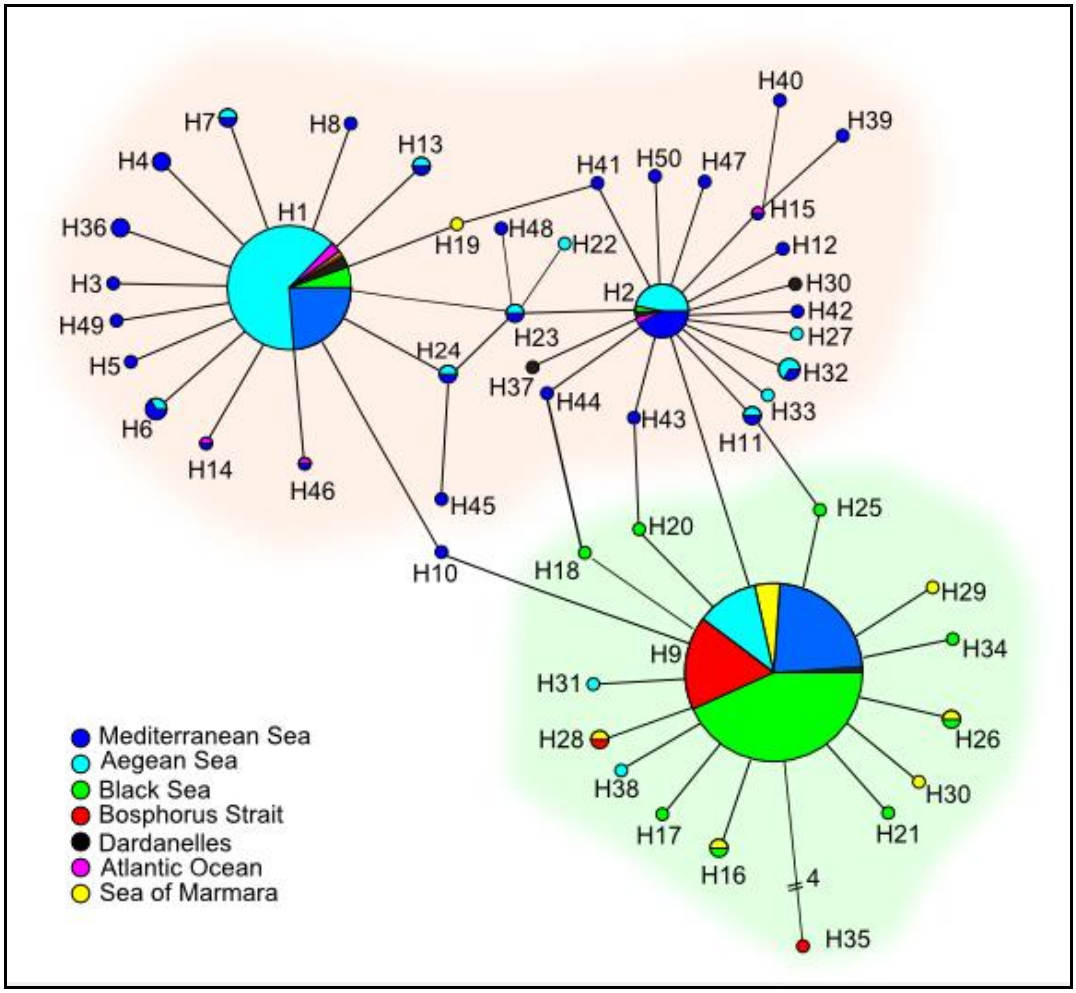

Figure 4.5. The haplotype network of mtDNA CO1 sequences of *P. marmoratus* used in this study and in Kalkan (2013) together with Genbank sequences. Each line corresponds to one base-pair difference and the diameter of each circle is proportional to the frequency of that haplotype. The background shading around several haplotypes define haplogroups, with the pink and green shading corresponding to haplogroups of H1 and H9, respectively.

4.2.2. mtDNA Pattern in Microsatellite Clusters

Considering the two main populations based on the results of STRUCTURE (M and C), population M had 110 individuals with mtDNA CO1 subunit sequence data. 14 sites among 490 were variable. A total of 16 haplotypes were present with haplotype diversity and nucleotide diversity values of 0.607 and 0.002, respectively. Population C had 206 individuals with CO1 data, which showed a total of 21 haplotypes with 23 polymorphic sites. M exhibited a higher haplotype diversity than C (0.607 vs. 0.488, respectively) despite its smaller sample size. However, the nucleotide diversities of the two populations were comparable (Table 4.5). Population C was mainly composed of haplogroup of H9 whereas population M was mainly composed of haplogroup of H1 (Figure 4.8). Number of haplotypes in C was a slightly higher than that of M (21 vs. 16), as might be expected due to different sample sizes (206 and 110, respectively). According to AMOVA results, the differentiation among the two populations was significant with an F_{CT} value of 0.39, explaining 38.79% of total variation in the data. The variation within 12 subpopulations and among 12 subpopulations within groups (two populations) explained 48.34% and 12% of the variation, respectively (Table 4.6).

Considering individual populations, Central Aegean subpopulation C8 showed the highest haplotype diversity (0.818), followed by North Aegean subpopulations C7 and M7, respectively. In terms of nucleotide diversity, C7 and C8 again were the highest with values of 0.0027 and 0.0026, respectively (Table 4.5). For population C, all of population expansion tests (Tajima's D, Fu's F_S and R_2) suggested population expansion whereas for M, only Fu's F_S and Tajima's D hinted for it (Table 4.5). Mismatch distribution graphs (Figure 4.6) of both populations also resembles a pattern for a very recent bottleneck or an expansion, as they showed a unimodal distribution of pairwise differences. When population C was divided further into two groups as suggested in microsatellite analyses (PCoA and pairwise F_{ST}), its members to the North of the Dardanelles (C1-C6) showed signs of population expansion but its members in the Mediterranean (C7-C9) did not (Table 4.5).

In comparing sympatric subpopulations in the Eastern Mediterrnean Sea, population C in the Eastern Mediterranean (C7-C9) was not differentiated from population M in the same region (M7-M9) in terms of mtDNA CO1 subunit (Table 4.3). Members of population C (C7C9) had higher haplotype and nucleotide diversities than members of M in the Mediterranean, with the exception of C9 (which had similar values of h and π with members of population M, Table 4.5). Haplotype frequencies of sympatric populations (Figure 4.7) were also similar (C8 and M8 seemed to be different but sample sizes were not comparable, 12 vs. 51, respectively).

Table 4.5. Number of individuals (N), number of haplotypes (Nh), haplotype (h) and nucleotide (π) diversities per site, Tajima's D parameter, Fu's F_S, and R_2 parameter for each subpopulation (C1-C9, M7-M9) and for two populations (C and M).

Pop	**N**	**Nh**	**Np**	**h**	**π**	**D**	**F_S**	**R_2**
C1	29	5	5	0.261 (0.11)	0.0008 (0.0004)			
C2	27	5	5	0.393 (0.11)	0.0013 (0.0004)			
C3	27	5	5	0.393 (0.11)	0.0013 (0.0004)	**-2.28**	**-18.8**	0.02
C4	27	4	7	0.214 (0.10)	0.0011 (0.0006)			
C5	27	6	6	0.342 (0.12)	0.0009 (0.0004)			
C6	22	3	4	0.177 (0.11)	0.0007 (0.0005)			
C7	20	6	6	0.705 (0.09)	0.0027 (0.0005)			
C8	12	4	4	0.818 (0.07)	0.0026 (0.0004)	-0.8	-2.7	0.08
C9	15	4	3	0.543 (0.13)	0.0019 (0.0005)			
C	206	21	23	0.488 (0.04)	0.0019 (0.04)	**-2.08**	**-19.15**	**0.02**
M7	43	9	7	0.630 (0.07)	0.0022 (0.0003)			
M8	51	9	8	0.602 (0.07)	0.0022 (0.0003)			
M9	16	5	4	0.600 (0.13)	0.0020 (0.0004)			
M	110	16	14	0.607 (0.05)	0.0020 (0.0002)	**-1.69**	**-12.03**	0.04

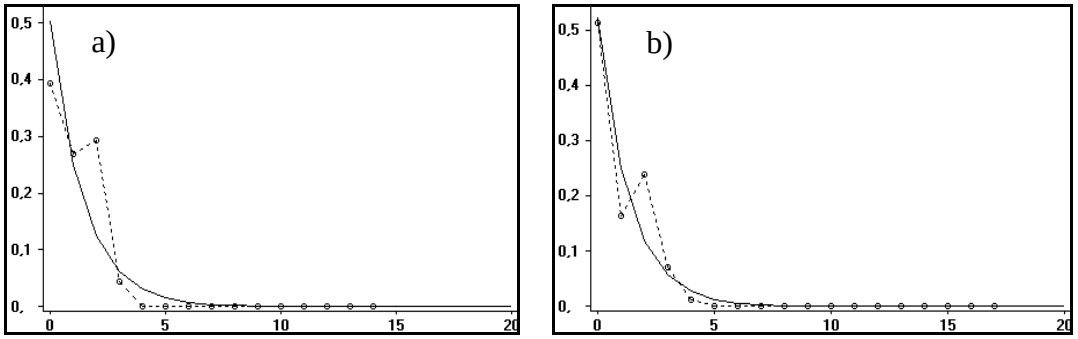

Figure 4.6. Mismatch distribution plot of a) population M b) population C. The lines with empty circles represent the observed distribution and continuous lines represent the expected distribution under a sudden expansion model.

Table 4.6. Analysis of molecular variance for distribution of genetic variation using mtDNA CO1 subunit. Sum of squares (SS), variance, percentage of total variation, F statistics and *P* values are shown. Significant *P* values are in bold.

Source of variation	SS	Variance components	Variation (%)	F statistics	P value
Among groups	44.51	0.30	38.79	$F_{CT} = 0.39$	< **0.005**
Among populations/groups	27.34	0.10	12.87	$F_{SC} = 0.21$	< **0.001**
Within populations	108.62	0.37	48.34	$F_{ST} = 0.52$	< **0.001**
Total	180.47	0.77			

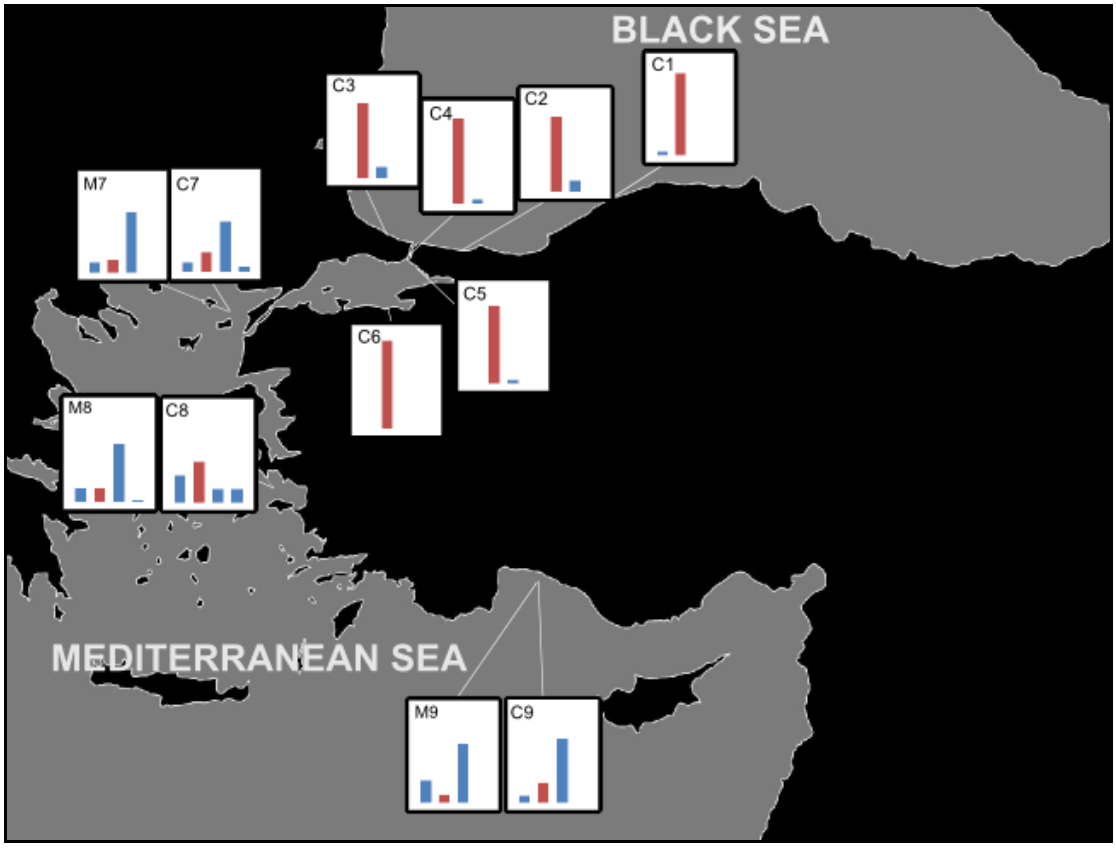

Figure 4.7. Geographical distribution of the four most common haplotypes (H2, H9, H1, and H6, respectively) in 12 subpopulations. Sampling locations of pooled populations are approximated as most populations consist of several sampling sites.

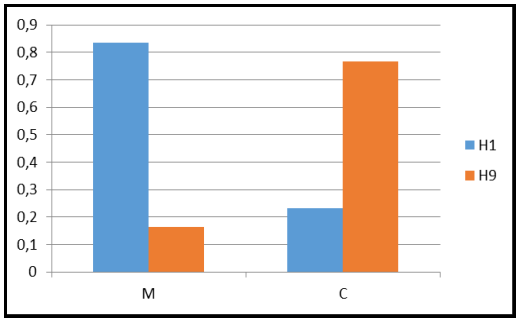

Figure 4.8. Frequencies of haplogroups of H1 and H9 in two populations defined by STRUCTURE (M and C).

4.2.3. Sympatric Groups in the Mediterranean

In the Mediterranean coasts of Turkey, there were two sympatric groups according to STRUCTURE clusters based on microsatellite data. One of them was composed of subpopulations C7, C8, and C9, which belonged to population C. The other population (M) was composed of M7, M8, and M9 subpopulations. The sample size of population M was approximately twice that of population C in this basin (126 and 61, respectively). To note, although microsatellites revealed two distinct sympatric groups in that region, mtDNA COI subunits were similar as shown in the haplotype frequency distribution (Figure 4.7), and in Table 4.3 with very low/minus insignificant pairwise Φst values between sympatric subpopulations C7-C9 and M7-M9.

To illustrate the distribution of individuals that exhibit this cytonuclear mismatch, a STRUCTURE run bar graph was produced at K=2 by ordering individuals according to their assignment probabilities (Q values) for each of the 12 subpopulations. Then, in another structure run, individuals were labeled according to their mtDNA haplogroups (H1 and H9, see Figure 4.9) in the same order with previous STRUCTURE microsatellite bar graph. The resulting bar chart was matched with that of microsatellite results and subsequently it was possible to visualize the individuals with cytonuclear mismatch in each of the 12 subpopulations.

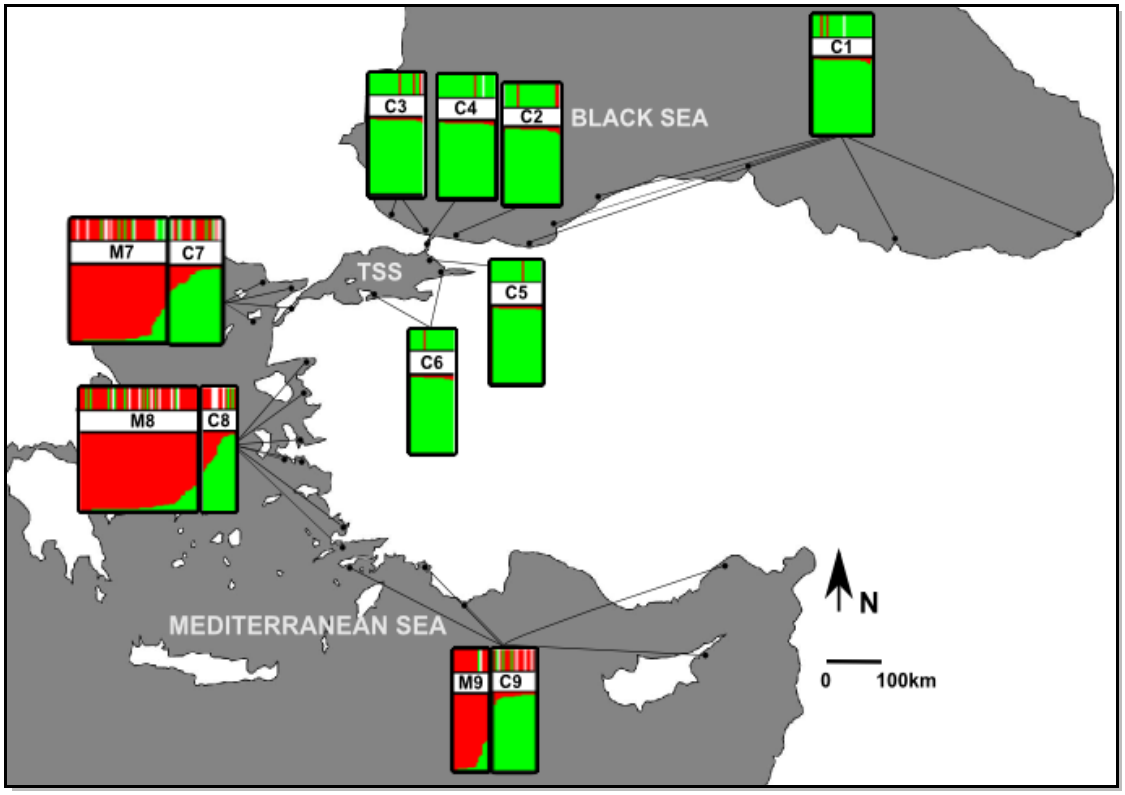

Figure 4.9. Visual representation of microsatellite data combined with mtDNA data of 12 subpopulations on the map. STRUCTURE Q values were ordered in each subpopulation (green-population C, red-population M) and corresponding mtDNA haplogroups (H1-red or H9-green, no data-white) were indicated on the top row. Each bar represents a single individual.

5. DISCUSSION

The study of phylogeography aims at inferring historical and contemporary forces which have produced the current genetic architecture of populations or closely related species. To this aim, various complementary genetic markers can be used. However, due to distinct evolutionary processes affecting them, they can produce discordant differentiation patterns (see section 1.4 above). Here, two *P. marmoratus* populations revealed by nuclear microsatellites share the same or have very similar mtDNA CO1 haplotypes. *Cytonuclear discordance* in sympatric populations is discussed referring to historical biogeography of the TSS focusing on three hypothetical scenarios. Before focusing on those scenarios, I will summarize patterns revealed by microsatellite and mtDNA CO1 results.

5.1. Population Structure Using Nuclear Microsatellites

Because microsatellites have high mutation rates and high levels of polymorphism, they are expected to provide a better understanding of larval dispersal and connectivity of populations (see section 3.3.1). Using STRUCTURE and five microsatellite loci, two clusters (referred to as population C and M) were detected, having 225 and 126 individuals, respectively. Among the five loci used, three (pm108, pm79, and pm99) had visibly different allele frequencies in two populations. Population C was present all around the Turkish coasts (C1-C9), and population M was constrained to Turkish Mediterranean coasts (M7-M9). Significant genetic differentiation between these two main clusters was concordantly identified by the results of pairwise F_{ST}, AMOVA, and PCoA. Additionally, the results hinted for another, albeit less extensive, differentiation of population C into two groups, C1-C6 and C7-C9. Whereas subpopulations C1-C6 are present in north of the Dardanelles, C7-C9 are in Turkish Mediterranean coasts.

5.2. mtDNA CO1 subunit

Because of the low mutation rate, CO1 gene (Schubart et al., 1998) might not be informative in intraspecific inferences. Differences between main haplotypes (H1 and H9) are not high in terms of numbers of mutations and both of them are observed around most of the sampled basins. However, certain patterns are observed. First of all, there are frequency differences of the CO1 gene revealed by high pairwise Φst values between subpopulations north of the Dardanelles (C1-C6) and those in the Mediterranean coasts of Turkey (C7-C9, M7-M9). As well, subpopulations C1-C6 have lower haplotype and nucleotide diversities than those in the Mediterranean. Secondly, haplogroup of H1 has a distribution including all of the Turkish coasts, western Mediterranean and also the Atlantic Ocean but it has very low frequency in the Black Sea and in the TSS. Haplogroup of H9 has the highest frequency in the Black Sea and is not present in samples from the Atlantic Ocean. In addition, although H9 is present in the Mediterranean basin, the distribution of its associated haplotypes (differentiated from H9 by one bp) are absent in the Mediterranean.

5.3. Sympatric Populations along the Mediterranean Coasts of Turkey

As explained in section 4.2.3, along the Mediterranean coasts of Turkey, there are two sympatric populations (C7-C9 vs. M7-M9) based on nuclear microsatellites, with population M having more individuals than that of population C in the region (126 vs. 61). The microsatellite data strongly differentiate sympatric populations, so the present day gene flow between them should be limited. In terms of mitochondrial data, however, no differentiation is observed between Mediterranean subpopulations C7-C9 and M7-M9. Similarity of mitochondrial haplotypes in these sympatric populations which are distinct in nuclear microsatellites suggests *cytonuclear discordance*. This pattern can be due to either the gene flow between these two subpopulations or alternatively these haplotypes may represent ancestral polymorphisms that did not yet get sorted (Funk & Omland, 2003). In each of three scenarios below, the putative role of these alternative processes is discussed referring to the biogeographical history of the Mediterranean.

5.3.1. Population C: A Local Population along the Turkish Coasts

One of the scenarios that might have resulted in the observed *cytonuclear discordance* could be explained by a uniform population that existed in the Eastern Mediterranean (population C) before the last opening of the Dardanelles, about 9,000 years BP (Zaitsev & Mamaev, 1997). Under this scenario, the Black Sea population of the species would have gone extinct due to low temperatures and brackish conditions during the Last Würm Glaciation (24,000-10,000 BP). Subsequently, when the Mediterranean Sea levels exceeded that of the Black Sea around 9,200-7,500 years BP, Mediterranean waters began to flow towards the Black Sea. In about 1,000-1,500 years, its salinity became sufficient to support various Mediterranean species (Zaitsev & Mamaev, 1997) (see section 1.4). *P. marmoratus*, with its pelagic larvae could have populated the Black Sea during that period, through a founder event. During this range expansion, the expanding population which is supposedly ancestral to population C is likely to have undergone a population bottleneck, leading to the loss of genetic variation in the Black Sea. The average He values of the subpopulations to the north of the Dardanelles (C1-C6) being slightly lower (0.44±0.01) than those of the Mediterranean subpopulations of C (0.51±0.03) support this scenario, as low diversity values are indicative of strong bottlenecks due to founder events (Roman, 2006). The lower haplotype diversity values (0.297) of subpopulations C1-C6 in the CO1 gene, when compared to subpopulations C7-C9 (0.689) in the Mediterranean also suggest low diversity in populations to the north of the Dardanelles. Low diversity in Black Sea subpopulations with respect to those in the Mediterranean has also been shown in many marine species, including anchovy (Magoulas et al., 2006) and plankton (Peijnenburg et al., 2006) in terms of mtDNA (Patarnello et al., 2007).

Previously, population M might already have been present in the western Mediterranean or northeastern Atlantic and have not yet reached eastern Mediterranean. In contrast, the population C has been present along Turkish Mediterranean coasts including the Black Sea as suggested in the above paragraph. Subsequently, a more recent range expansion might have resulted in population M inhabiting the Mediterranean coasts of Turkey. Around 7,000 years BP, the current pattern in the TSS changed, and the Black Sea waters began to flow towards

the Mediterranean. Thus, if population M reached eastern Mediterranean after 7,000 years BP, the current pattern may have impeded the incoming population M from inhabiting the north of the Dardanelles. Selection tests both showing signs of population expansion for M in the Eastern Mediterranean (M7-M9), but not for the C subpopulations in the same region (C7-C9) also supports this scenario.

Alternatively, ecological differences between different water bodies may have played a role in keeping population M from entering sea basins to the north of the Dardanelles. Same factors (current patterns and/or ecological differences) may have also prevented the population C in the Mediterranean from reestablishing on northern parts of the Dardanelles. In addition, higher STRUCTURE assignment values of C9 may imply that population M incoming from the western Mediterranean was able to populate the Southern Aegean (C8) and Northern Aegean (C7) regions more easily when compared to the Levantine basin (C9). The fact that the number of individuals that belonged to population M was higher than that of C in the Mediterranean suggests that, either population M was able to successfully or repeatedly dominate local population C in the area or that M was the local population in the area and C later inhabited there, albeit unsuccessfully. However, the population C is present all along the Turkish coasts in contrast to population M which is present only in the Mediterranean, so the local population around Turkish coasts is more probably the population C. In addition, if M was the population dispersing into the Turkish Mediterranean coasts, unidirectional gene flow of local population C's mtDNA into M could be an explanation to *cytonuclear discordance*, creating similarity in terms of previously differentiated mtDNA haplotypes (Currat et al., 2008; Excoffier et al., 2009). For example, in a study of two introduced green crab *Carcinus maenas* (Linnaeus, 1758) populations using the mtDNA CO1 gene and nine microsatellite loci, (Darling et al. (2014) found unidirectional mitochondrial introgression, from the resident population to the recent dispersers into the area. The study also showed a geographical pattern as evidenced by an admixture zone.

In contrast, no such pattern can be distinguished in the *P. marmoratus* subpopulations of the eastern Mediterranean. When there is not any geographical pattern detected, the retention

of polymorphisms is more likely as an explanation for the discordance because in that case polymorphisms in the common ancestor would arbitrarily be distributed among populations regardless of the geography (Funk & Omland, 2003). Thus, CO1 of two sympatric populations could be similar due to incomplete lineage sorting of ancestral polymorphisms (Funk & Omland, 2003). In other words, given the uncertainty in the duration of allopatry before their encounter in the eastern Mediterranean (previous than 7,000 BP), it is possible that two populations did not have stayed isolated long enough to diverge in CO1. It should be noted that, among the three scenarios mentioned, the longest isolated time period is assumed in this scenario. Thus, if incomplete lineage sorting is a valid explanation for the discordance in this scenario, it can be even more likely in the two scenarios below.

5.3.2. A Long History of Sympatry in the Eastern Mediterranean

Considering another potential scenario, the two sympatric populations in the Mediterranean may have actually co-inhabited Turkish coasts long before the LGM. Population C and M might have originated in allopatry during glaciations in the Pleistocene in the Atlantic or in the western Mediterranean and might have expanded their range towards the eastern Mediterranean either consecutively or simultaneously. Since then, two populations might have been present in sympatry throughout the Mediterranean. In that case, occasional gene flow between populations and slow evolution rate of CO1 might not have allowed for differentiation in sympatric populations. In contrast, nuclear microsatellites would reflect the differentiation as they evolve much faster, providing evidence for the more recent genetic differentiation. If nuclear differentiation in two populations is the result of long term reproductive isolation between them, two populations might as well have diverged to different species. However, two populations share most of their CO1 haplotypes. Thus, a much more possible hypothesis is that two populations were never completely mixed during glacial fluctuations in the Pleistocene. For example, population C might have reached the eastern Mediterranean through the northern African coasts but larvae of population M might have inhabited the Aegean coasts by following a different current pattern. Previous studies about the species could not find different microsatellite clusters in the western Mediterranean

(Fratini et al., 2013) and in the Portuguese coasts (Silva et al., 2009) without being able to detect sympatric groups. Thus, these sympatric populations in the Mediterranean coasts of Turkey might have reached there via different independent routes. Populations of *P. marmoratus* in the western Mediterranean and northeastern Atlantic (Fratini et al., 2011) also showed signs of population expansion with a unimodal mismatch distribution using the same marker, mtDNA CO1. Thus, two populations could have reached eastern Mediterranean from the western Mediterranean and northeastern Atlantic. However, as stated earlier, the inferences about CO1 should be taken with caution due to its insufficient variation.

According to this scenario, population C might have expanded into the Black Sea between 9,200-7,500 years ago during the Mediterranean inflow into the Black Sea because of some adaptational advantage whereas population M could not. However, presumed adaptational advantage of population C which would foster its dispersal into the Black Sea could not be tested although significant Tajima's D results in subpopulations of C in north of the Dardanelles could also be due to selection other than bottleneck or population expansion. To test this adaptational advantage scenario, RAD (Restriction site associated DNA) sequencing could be performed.

5.3.3. Population C: Surviving the Last Glacial Maximum in the Black Sea

According to another scenario, the population M could have been the local population in the Mediterranean and population C, differentiating from population M in the Black Sea during the LGM could have populated the region more recently, after the establishment of the current system in the TSS around 7,000 years ago. The similarity in CO1 data of sympatric populations could be explained by secondary gene flow of mtDNA of local population M into that of C, although the alternative hypothesis of incomplete lineage sorting in mtDNA is still relevant as two populations would be isolated approximately for only about 15,000 years.

After the opening of the Dardanelles first time since the formation of the Tethys Sea in the Riss-Wurm Interglacial (100,000-150,000 ya), ancestral *P. marmoratus* populations might

have dispersed from the Mediterranean into the Black Sea, which had higher salinity than present (Zaitsev & Mamaev, 1997). During the peak of Last Würm Glaciation (24,000-20,000 ya), the Black Sea population of the species could have survived in isolation, and differentiated from the Mediterranean populations. Under this isolation, population C and population M differentiated in the Black Sea and in the Mediterranean, respectively. Kalkan (2013) also proposed a similar scenario for the differentiation of *Mytilus galloprovincialis* mitochondrial clades in the Black Sea and in the Aegean Sea during the LGM. Afterwards, population C could expand its range into the Mediterranean when the Dardanelles opened and the current system was established. During its expansion, its mtDNA would have been replaced by that of the local population M, resulting in the lack of differentiation in mtDNA of sympatric populations (Excoffier et al., 2009).

An objection to this scenario is that selection tests do not give significant results to support population expansion of C in the Mediterranenan. In addition, subpopulations in the Levantine basin (C9 and M9) show higher assignment values to population C than other Mediterranean subpopulations although C9 is the farthest population to the Black Sea. Thus, this resemblance of C9 to subpopulations in the Black Sea and in the TSS may suggest connectivity between them which cannot readily be explained by the current patterns and geography in the region. Another disadvantage of this scenario is the assumption that populations in the Black Sea have survived the Last Würm Glaciation during which the salinity of the Black Sea was around 1 psu in compared to its current salinity, 18 psu (Zaitsev & Mamaev, 1997). The species might not survive in habitats with such low salinities since its distribution does not involve lower salinities than the Black Sea (Kocatas et al., 2004). However, a refugial population might have survived in the SoM which may have had relatively higher salinity during the LGM as also mentioned regarding a planktonic species around the Black Sea (Peijnenburg et al., 2004) but was not given much credit as an explanation.

6. CONCLUSIONS

Here, using putatively neutral and fast evolving microsatellite markers, we found high levels of genetic differentiation between populations of *P. marmoratus* distributed all around the Turkish coasts and those constrained to the Mediterranean coasts. Both populations, however, share the same or have very similar mtDNA CO1 haplotypes. It appears that mitochondrial marker used in this study was slow for the relevant time scale. The exact timing or the process of the disjunction of *P. marmoratus* populations remains unclear. However, the connection of the Black Sea with the Mediterranean which has been interrupted several times might have influenced the pattern detected in this study. Among the three scenarios discussed, the more plausible one seems to be the first one in which population C is an ancestral local population around the Turkish coasts and population M later colonized Mediterranean coasts of Turkey. This scenario is supported by expansion tests for both populations in the Mediterranean and also lower genetic diversity values of subpopulations in the Black Sea than those in the Mediterranean. Two alternative explanations for the *cytonuclear discordance* as retention of ancestral polymorphisms or gene flow after secondary contact could not be distinguished. Most probably, both processes have contributed the pattern during the process although the former is more likely due to slow mutation rate of CO1. The use of complementary mitochondrial markers with faster mutation rates than CO1 subunit in addition to microsatellites and sampling individuals from the rest of the Mediterranean and from the Atlantic Ocean can be useful in resolving the pattern. In addition, RAD sequencing might be effective in resolving the phylogeography of the species. Lastly, the subpopulations to the north of the Dardanelles deserves further attention as all of the private alleles belonged to them although they had low genetic diversity. Thus, future studies should also focus on the subpopulations in the Black Sea and in the Sea of Marmara to understand their role in genetic diversity of the species in the Mediterranean Sea.

REFERENCES

Amos, W., Hoffman, J. I., Frodsham, A., Zhang, L., Best, S., & Hill, A. V. S. (2007). Automated binning of microsatellite alleles: Problems and solutions. *Molecular Ecology Notes*, *7*, 10–14. http://doi.org/10.1111/j.1471-8286.2006.01560.x

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

Bandelt, H.-J., Forster, P., & Rohl, A. (1999). Median-Joining Networks for Inferring Intraspecific Phylogenies. *Molecular Biology and Evolution*, *16*(1), 37–48.

Brookfield, J. F. Y. (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, *5*(3), 453–455. http://doi.org/10.1111/j.1365-294X.1996.tb00336.x

Cannicci, S., Gomei, M., Dahdouh-Guebas, F., Rorandelli, R., & Terlizzi, A. (2007). Influence of seasonal food abundance and quality on the feeding habits of an opportunistic feeder, the intertidal crab Pachygrapsus marmoratus. *Marine Biology*, *151*, 1331–1342. http://doi.org/10.1007/s00227-006-0570-3

Cannicci, S., Paula, J., & Vannini, M. (1999). Activity pattern and spatial strategy in Pachygrapsus marmoratus (Decapoda: Grapsidae) from Mediterranean and Atlantic shores. *Marine Biology*, *133*, 429–435. http://doi.org/10.1007/s002270050481

Currat, M., Ruedi, M., Petit, R. J., & Excoffier, L. (2008). The hidden side of invasions: Massive introgression by local genes. *Evolution*, *62*(8), 1908–1920. http://doi.org/10.1111/j.1558-5646.2008.00413.x

Darling, J. A., Tsai, Y. E., Blakeslee, A., & Roman, J. (2014). Are genes faster than crabs ? Mitochondrial introgression exceeds larval dispersal during population expansion of the invasive crab Carcinus maenas. *Royal Society Open Science*, *1:140202*. http://doi.org/http://dx.doi.org/10.1098/rsos.140202

Dauvin, J.-C. (2012). Establishment of a population of marbled crab Pachygrapsus marmoratus (Crustacea: Brachyura: Grapsoidea) on the coast of northern Cotentin, Normandy, on the western English Channel. *Marine Biodiversity Records*, *5*(2002), 1–4. http://doi.org/10.1017/S1755267212000334

Dumitrache, C. (1999). Pachygrapsus marmoratus Fabriciu, 1787. Retrieved January 3, 2015, from http://www.grid.unep.ch/bsein/redbook/txt/pachygr.htm? CRUSTACEA

Earl, D. A., & VonHoldt, D. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, (4), 359–361.

Edwards, R. (2005). Pachygrapsus marmoratus. A marbled rock crab. Retrieved August 5, 2014, from http://www.marlin.ac.uk/speciesinformation.php?speciesID=4005

Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nature Reviews. Genetics*, *5*(June), 435–445. http://doi.org/10.1038/nrg1348

Estoup, A., Jarne, P., & Cornuet, J. M. (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, *11*, 1591–1604. http://doi.org/10.1046/j.1365-294X.2002.01576.x

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*, 2611–2620. http://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic Consequences of Range Expansions. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 481–501. http://doi.org/10.1146/annurev.ecolsys.39.110707.173414

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. *Molecular Ecology Resources*, *10*, 564–567. http://doi.org/10.1111/j.1755-0998.2010.02847.x

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.

Flores, A. A. V., & Paula, J. (2002). Population dynamics of the shore crab Pachygrapsus marmoratus (Brachyura: Grapsidae) in the central Portuguese coast. *Journal of the Marine Biological Association of the UK*, *82*, 229–241. http://doi.org/10.1017/S0025315402005404

Fontaine, M. C., Baird, S. J. E., Piry, S. et al. (2007). Rise of oceanographic barriers in continuous populations of a cetacean: the genetic structure of harbour porpoises in Old World waters. *BMC Biology*, *5*, 30. http://doi.org/10.1186/1741-7007-5-30

Fratini, S., Ragionieri, L., Cutuli, G., Vannini, M., & Cannicci, S. (2013). Pattern of genetic isolation in the crab pachygrapsus marmoratus within the tuscan archipelago (Mediterranean Sea). *Marine Ecology Progress Series*, *478* (Levin 2006), 173–183. http://doi.org/10.3354/meps10247

Fratini, S., Ragionieri, L., Papetti, C., Pitruzzella, G., Rorandelli, R., Barbaresi, S., & Zane, L. (2006). Isolation and characterization of microsatellites in Pachygrapsus marmoratus (Grapsidae; Decapoda; Brachyura). *Molecular Ecology Notes*, *6*, 179–181. http://doi.org/10.1111/j.1471-8286.2005.01184.x

Fratini, S., Schubart, C. D., & Ragionieri, L. (2011). Population genetics in the rocky shore crab Pachygrapsus marmoratus from the western Mediterranean and eastern Atlantic : complementary results from mtDNA and microsatellites at different geographic scales. In: Held, C., Koenemann, S., Schubart, C.D., (Eds.), 191−213, Crustacean issues 19: Phylogeography and Population Genetics in Crustacea, CRC Press, Boca Raton.

Fratini, S., Zane, L., Ragionieri, L., Vannini, M., & Cannicci, S. (2008). Relationship between heavy metal accumulation and genetic variability decrease in the intertidal crab Pachygrapsus marmoratus (Decapoda; Grapsidae). *Estuarine, Coastal and Shelf Science*, *79*, 679–686. http://doi.org/10.1016/j.ecss.2008.06.009

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

Fu, Y. X., & Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics*, *133*, 693–709. http://doi.org/evolution

Funk, D. J., & Omland, K. E. (2003). Annual Review of Ecology, Evolution, and Systematics,. *Annual Review of Ecology and Systematics*, *34*, 397–423.

Görür, N., Cagatay, M. N., Sakinc, M., Sümengen, M., Sentürk, K., Yaltirak, C., & Tchapalyga, A. (1997). Origin of the Sea of Marmara as deduced from Neogene to Quaternary paleogeographic evolution of its frame. *International Geology Review*, *39*(March 2014), 342–352. http://doi.org/10.1080/00206819709465276

Hartl, L. D. & Clark, G. A. (1997). *Principles of Population Genetics*. *Sinauer Associates Inc.* (3rd Edition). Sunderland, Massachusetts. http://doi.org/10.1016/0378-1119(90)90076-4

Hedgecock, D. (1986). Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates? *Bulletin of Marine Science*, *39*(2), 550–564. Retrieved from http://www.ingentaconnect.com/content/umrsmas/bullmar/1986/00000039/00000002/art00033

Hedgecock, D., Launey, S., & Pudovkin, A. (2007). Small effective number of parents (Nb) inferred for a naturally spawned cohort of juvenile European flat oysters Ostrea edulis. *Marine Biology*, *150*, 1173−1182.

Hurlbert, S. H. (1971). The nonconcept of species diversity: a critic and alternative parameters. *Ecology*, (52), 577–586.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801–1806. http://doi.org/10.1093/bioinformatics/btm233

Kalinowski, S. T. (2005). HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, (5), 187–189.

Kalkan, E. (2013). *Investigating the role of the Turkish Straits System as a phylogeographical break in the Mediterranean-Black Sea transition*, Ph.D. Thesis, Bogazici University.

Kalkan, E., Kurtuluş, A., Maraci, Ö., & Bilgin, R. (2011). Is the Bosphorus Strait a barrier to gene flow for the Mediterranean mussel, Mytilus galloprovincialis (Lamarck, 1819)? *Marine Biology Research*, *7*(July 2014), 690–700. http://doi.org/10.1080/17451000.2011.558098

Kocatas, A., Katagan, T., & Ates, A. S. (2004). Atlanto-Mediterranean Originated Decapod Crustaceans in the Turkish Seas. *Pakistan Journal of Biological Sciences*, *7*(10), 1827–1830. http://doi.org/10.3923/pjbs.2004.1827.1830

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A, & Mayrose, I. (2015). Clumpak : a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, n/a–n/a. http://doi.org/10.1111/1755-0998.12387

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11), 1451–1452. http://doi.org/10.1093/bioinformatics/btp187

Magoulas, A., Castilho, R., Caetano, S., Marcato, S., & Patarnello, T. (2006). Mitochondrial DNA reveals a mosaic pattern of phylogeographical structure in Atlantic and Mediterranean populations of anchovy (Engraulis encrasicolus). *Molecular Phylogenetics and Evolution*, *39*, 734–746. http://doi.org/10.1016/j.ympev.2006.01.016

Metzgar, D., Bytof, J., & Wills, C. (2000). Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Research*, *10*(1), 72–80. http://doi.org/10.1101/gr.10.1.72

Patarnello, T., Volckaert, F. A. M. J., & Castilho, R. (2007). Pillars of Hercules: Is the Atlantic-Mediterranean transition a phylogeographical break? *Molecular Ecology*, *16*, 4426–4444. http://doi.org/10.1111/j.1365-294X.2007.03477.x

Peakall, R., & Smouse, P. E. (1999). Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity*, (82), 561–573.

Peakall, R., & Smouse, P. E. (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*, 288–295. http://doi.org/10.1111/j.1471-8286.2005.01155.x

Peakall, R., & Smouse, P. E. (2012). GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, *28*(19), 2537–2539. http://doi.org/10.1093/bioinformatics/bts460

Peijnenburg, K. T. C. A, Breeuwer, J. A. J., Pierrot-Bults, A. C., & Menken, S. B. J. (2004). Phylogeography of the planktonic chaetognath Sagitta setosa reveals isolation in European seas. *Evolution*, *58*(7), 1472–1487. http://doi.org/10.1554/03-638

Peijnenburg, K. T. C. A, Fauvelot, C., Breeuwer, J. A. J., & Menken, S. B. J. (2006). Spatial and temporal genetic structure of the planktonic Sagitta setosa (Chaetognatha) in European seas as revealed by mitochondrial and nuclear DNA markers. *Molecular Ecology*, *15*, 3319–3338. http://doi.org/10.1111/j.1365-294X.2006.03002.x

Poulos, S. E., Drakopoulos, P. G., & Collins, M. B. (1997). Seasonal variability in sea surface oceanographic conditions in the Aegean Sea (Eastern Mediterranean): an overview. *Journal of Marine Systems*, *13*, 225–244. http://doi.org/10.1016/S0924-7963(96)00113-3

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959. http://doi.org/10.1111/j.1471-8286.2007.01758.x

Ramos-Onsins, S. E., & Rozas, J. (2002). Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, *19*(12), 2092–2100. http://doi.org/10.1093/molbev/msl052

Raymond, M., & Rousset, F. (1995). Genepop (Version-1.2) - Population-Genetics Software for Exact Tests and Ecumenicism. *J. Hered.*, *86*, 248–249. Retrieved from <Go to ISI>://A1995RB30200017

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution*. http://doi.org/10.2307/2409177

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, *4*, 137–138. http://doi.org/10.1046/j.1471-8286.2003.00566.x

Rousset, F. (2008). GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, *8*, 103–106. http://doi.org/10.1111/j.1471-8286.2007.01931.x

Schlötterer, C. (2000). Evolutionary dynamics of microsatellite DNA. *Chromosoma*, *109*, 365–371. http://doi.org/10.1007/s004120000089

Schubart, C. D., Diesel, R., & Hedges, B. (1998). Rapid evolution to terrestrial life in Jamaican crabs. *Nature*, *393*, 363–365.

Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters*, *9*, 615–629. http://doi.org/10.1111/j.1461-0248.2006.00889.x

Silva, I. C., Mesquita, N., Schubart, C. D., Alves, M. J., & Paula, J. (2009). Genetic patchiness of the shore crab P*achygrapsus marmoratus* along the Portuguese coast. *Journal of Experimental Marine Biology and Ecology*, *378*(1-2), 50–57. http://doi.org/10.1016/j.jembe.2009.07.032

Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics*, *16*(May), 393–430. Retrieved from http://www.jstor.org/stable/2097054

Sweet, N. (2011). Marbled crab, *Pachygrapsus marmoratus*. Retrieved January 2, 2015, from http://www.nonnativespecies.org/factsheet/factsheet.cfm?speciesId=2491

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 and chimpanzees. *Molecular Biology and Evolution*, *10*(3), 512–526. http://doi.org/10.1093/molbev/msl149

The Global Biodiversity Information Facility: GBIF Backbone Taxonomy. (n.d.). Retrieved March 15, 2015, from http://www.gbif.org/species/2225815

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., & Higgins, D. G. (1997). The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, *25*(24), 4876–4882. http://doi.org/10.1093/nar/25.24.4876

Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, *21*, 3907–3930. http://doi.org/10.1111/j.1365-294X.2012.05664.x

Tomczak, M., & Godfrey, J. S. (2001). Adjacent Seas of the Atlantic Ocean. In *Regional Oceanography: An Introduction* (Vol. 0, pp. 271–298).

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*, 535–538. http://doi.org/10.1111/j.1471-8286.2004.00684.x

Viaud-Martínez, K.A., Vergara, M.M., Goldin, P.E., Ridoux, V., Öztürk, A.A., Öztürk, B., Rosel, P.E., Frantzis, A., Komnenou, A., Bohonak, A.J. (2007). Morphological and genetic differentiation of the Black Sea harbour porpoise *Phocoena phocoena*. *Marine Ecology Progress Series, 338*, 281–294. http://doi.org/10.3354/meps338281

Wright, S. (1954). The Genetical Structure of Populations. *The Annals of Human Genetics*.

Zaitsev, Y., & Mamaev, V. (1997). *Biological Diversity in the Black Sea: A Study of Change and Decline*.

Appendix A: LOCATIONS AND NUCLEAR GENETIC DIVERSITY OF 12 SUBPOPULATIONS

Table A1. Sampling sites with their codes used in figure 3.1 and their corresponding coordinates with relevant number of individuals used in mtDNA or microsatellite analyses.

					number of individuals	
Pooled Populations	Codes	Sampling Sites	Basin	Coordinates	microsatellites	mtDNA
C1	B1	Hopa	Black Sea	41°28'55.27"N 41°31'12.09"E	2	2
	B2	Persembe	Black Sea	41° 8'0.17"N 37°40'53.87"E	1	1
	B3	Sinop	Black Sea	42° 0'57.92"N 35°10'56.89"E	1	1
	B4	Gideros	Black Sea	41°51'35.62"N 32°51'25.13"E	1	1
	B5	Karadeniz Ereglisi	Black Sea	41°16'53.87"N 31°24'57.06"E	24	23
	B6	Akcakoca	Black Sea	41° 5'23.65"N 31° 7'20.22"E	1	1
C2	B7	Sile	Black Sea	41°10'54.42"N 29°36'22.97"E	28	27
C3	B8	Kilyos	Black Sea	41°14'53.18"N 29° 2'2.07"E	25	25
	B9	Karaburun	Black Sea	41°20'36.16"N 28°41'12.00"E	3	2
C4	T1	Hamsi Limani	Turkish Straits System	41°12'18.82"N 29° 6'12.38"E	19	19
	T2	Buyuk Liman	Turkish Straits System	41°10'51.80"N 29°6'22.21"E	9	8
C5	T3	Caddebostan	Turkish Straits System	40°57'46.28"N 29° 3'30.94"E	24	24
	T4	Kinaliada	Turkish Straits System	40°54'44.88"N 29° 3'14.98"E	4	3
C6	T5	Yalova	Turkish Straits System	40°40'18.25"N 29°19'46.57"E	12	12
	T6	Bandirma	Turkish Straits System	40°21'29.14"N 27°57'33.29"E	10	10
C7/M7	T7	Canakkale Bogazi	Turkish Straits System	40° 9'8.30"N 26°24'18.14"E	4/6	2/5
	M1	Enez	Mediterranean/Northern Aegean	40°41'45.48"N 26° 3'16.60"E	4/10	3/7
	M2	Saros-İbrice	Mediterranean/Northern Aegean	40°36'7.51"N 26°32'30.69"E	0/1	0/1
	M3	Saros-Guneyli	Mediterranean/Northern Aegean	40°30'36.99"N 26°42'12.76"E	11/22	11/22
	M4	Gokceada	Mediterranean/Northern Aegean	40°13'54.53"N 25°53'39.18"E	7/11	4/8
C8/M8	M5	Ayvalik	Mediterranean/Southern Aegean	39°18'47.85"N 26°41'16.49"E	0/1	0/0
	M6	Dikili	Mediterranean/Southern Aegean	39° 3'51.14"N 26°52'51.62"E	4/11	3/10
	M7	Urla	Mediterranean/Southern Aegean	38°21'51.04"N 26°46'16.66"E	4/16	3/14
	M8	Cesme	Mediterranean/Southern Aegean	38°20'12.64"N 26°23'19.13"E	0/2	0/2
	M9	Seferihisar	Mediterranean/Southern Aegean	38°11'36.27"N 26°46'24.38"E	0/1	0/1
	M10	Milas District	Mediterranean/Southern Aegean	37°16'36.92"N 27°33'23.34"E	4/10	3/10
	M11	Bodrum	Mediterranean/Southern Aegean	37° 0'20.28"N 27°15'23.41"E	4/18	3/14
C9/M9	M12	Palamutbuku	Mediterranean/Levantine	36°40'9.97"N 27°30'9.62"E	2/6	2/6
	M13	Fethiye	Mediterranean/Levantine	36°38'25.32"N 29° 6'3.12"E	0/2	0/2
	M14	Kas	Mediterranean/Levantine	36°11'48.18"N 29°38'39.41"E	14/7	11/6
	M15	Tasucu	Mediterranean/Levantine	36°16'25.78"N 33°48'55.10"E	2/2	1/2
	M16	Kibris	Mediterranean/Levantine	35°22'34.97"N 34° 5'33.45"E	1/0	1/0

Table A2. Genetic variability of the five microsatellite loci in subpopulations for *P. marmoratus.* N, sample size; Na, number of alleles; RS, allelic richness; Ne, Number of effective alleles; I, Shannon's diversity index; Ho, observed heterozygosity; PR, number of private alleles; He, expected heterozygosity; uHe, unbiased expected heterozgosity.

		Pm99	Pm101	Pm108	Pm187	Pm79	Avg.
C1	N	25	30	30	28	30	28.6
	Na	13	11	1	3	2	6
	RS	8.93	8.84	1	2.98	1.93	4.74
	Ne	4.19	6.82	1	1.88	1.18	3.01
	I	1.92	2.12	0	0.82	0.29	1.03
	He	0.76	0.85	0	0.47	0.15	0.45
	uHe	0.78	0.87	0	0.48	0.16	0.46
	Ho	0.68	0.77	0	0.25	0.03	0.35
	PR	0.48	0.14	0	0	0	0.12
C2	N	21	25	28	26	28	25.6
	Na	12	15	1	5	3	7.2
	RS	8.62	11.74	1	4.08	2.25	5.54
	Ne	3.68	10.78	1	1.71	1.16	3.66
	I	1.79	2.51	0	0.86	0.3	1.09
	He	0.73	0.91	0	0.42	0.13	0.44
	uHe	0.75	0.93	0	0.42	0.14	0.45
	Ho	0.71	0.8	0	0.35	0.14	0.4
	PR	0.58	0.86	0	0.92	0.43	0.56
C3	N	26	22	28	26	28	26
	Na	13	12	1	4	2	6.4
	RS	9.07	9.55	1	3.41	1.43	4.89
	Ne	5.1	7.81	1	1.63	1.04	3.31
	I	2.02	2.21	0	0.75	0.09	1.01
	He	0.8	0.87	0	0.39	0.04	0.42
	uHe	0.82	0.89	0	0.39	0.04	0.43
	Ho	0.85	0.86	0	0.27	0.04	0.4
	PR	0.24	0	0	0.46	0	0.14
C4	N	22	27	28	27	25	25.8
	Na	11	12	1	3	2	5.8
	RS	8.67	10.28	1	2.91	1.73	4.92
	Ne	5.83	9	1	1.64	1.08	3.71
	I	2.01	2.33	0	0.7	0.17	1.04
	He	0.83	0.89	0	0.39	0.08	0.44
	uHe	0.85	0.91	0	0.4	0.08	0.45
	Ho	0.64	0.81	0	0.19	0.08	0.34
	PR	0	0.42	0	0	0	0.08

Table A2. Genetic variability of the five microsatellite loci in subpopulations for *P. marmoratus.* N, sample size; Na, number of alleles; RS, allelic richness; Ne, Number of effective alleles; I, Shannon's diversity index; Ho, observed heterozygosity; PR, number of private alleles; He, expected heterozygosity; uHe, unbiased expected heterozgosity. (Cont.)

		Pm99	Pm101	Pm108	Pm187	Pm79	Avg.
C5	N	21	24	28	26	28	25.4
	Na	13	11	1	4	1	6
	RS	10.34	9.37	1	3.92	1	5.13
	Ne	5.84	7.02	1	2.22	1	3.42
	I	2.16	2.15	0	1.04	0	1.07
	He	0.83	0.86	0	0.55	0	0.45
	uHe	0.85	0.88	0	0.56	0	0.46
	Ho	0.62	0.63	0	0.35	0	0.32
	PR	0.42	0	0	0.01	0	0.09
C6	N	21	20	22	20	21	20.8
	Na	12	12	1	4	2	6.2
	RS	9.26	9.84	1	3.82	1.93	5.17
	Ne	3.87	7.62	1	1.81	1.15	3.09
	I	1.88	2.21	0	0.87	0.26	1.04
	He	0.74	0.87	0	0.45	0.13	0.44
	uHe	0.76	0.89	0	0.46	0.14	0.45
	Ho	0.62	0.7	0	0.25	0.14	0.34
	PR	0.57	0	0	0	0	0.12
C7	N	23	24	21	25	23	23.2
	Na	13	12	1	3	2	6.2
	RS	11.43	9.37	1	2.97	2	5.35
	Ne	10.91	7.25	1	1.85	1.91	4.58
	I	2.47	2.17	0	0.8	0.67	1.22
	He	0.91	0.86	0	0.46	0.48	0.54
	uHe	0.93	0.88	0	0.47	0.49	0.55
	Ho	0.43	0.63	0	0.24	0.35	0.33
	PR	0.08	0.02	0	0	0	0.02
C8	N	14	15	12	14	13	13.6
	Na	10	10	3	3	2	5.6
	RS	9.52	9.32	3	2.71	2	5.31
	Ne	6.88	6.92	1.19	1.16	1.9	3.61
	I	2.09	2.09	0.34	0.31	0.67	1.1
	He	0.85	0.86	0.16	0.14	0.47	0.49
	uHe	0.89	0.89	0.16	0.14	0.49	0.51
	Ho	0.29	0.53	0.17	0.14	0.46	0.32
	PR	0	0.08	0	0	0	0.02

Table A2. Genetic variability of the five microsatellite loci in subpopulations for *P. marmoratus.* N, sample size; Na, number of alleles; RS, allelic richness; Ne, Number of effective alleles; I, Shannon's diversity index; Ho, observed heterozygosity; PR, number of private alleles; He, expected heterozygosity; uHe, unbiased expected heterozgosity.(Cont.)

		Pm99	Pm101	Pm108	Pm187	Pm79	Avg.
C9	N	16	16	19	19	17	17.4
	Na	12	8	1	3	2	5.2
	RS	10.85	7.33	1	2.99	2	4.83
	Ne	8.83	4.13	1	1.83	1.56	3.47
	I	2.31	1.69	0	0.8	0.55	1.07
	He	0.89	0.76	0	0.45	0.36	0.49
	uHe	0.92	0.78	0	0.47	0.37	0.51
	Ho	0.44	0.5	0	0.16	0.35	0.29
	PR	0	0	0	0	0	0
M7	N	44	47	40	45	47	44.6
	Na	15	14	4	4	2	7.8
	RS	10.46	10.24	2.48	3.19	2	5.67
	Ne	8.74	9.48	1.17	1.74	1.99	4.62
	I	2.39	2.39	0.34	0.79	0.69	1.32
	He	0.89	0.89	0.14	0.43	0.5	0.57
	uHe	0.9	0.9	0.14	0.43	0.5	0.58
	Ho	0.55	0.79	0.15	0.29	0.6	0.47
	PR	0.17	0	0.34	0	0	0.1
M8	N	56	56	43	52	56	52.6
	Na	17	13	3	3	2	7.6
	RS	11.98	9.52	1.76	2.96	2	5.64
	Ne	12.08	8.77	1.07	1.75	1.97	5.13
	I	2.63	2.3	0.17	0.77	0.69	1.31
	He	0.92	0.89	0.07	0.43	0.49	0.56
	uHe	0.93	0.89	0.07	0.43	0.5	0.56
	Ho	0.63	0.7	0.07	0.37	0.41	0.43
	PR	0.41	0.06	0.16	0	0	0.13
M9	N	15	16	15	17	16	15.8
	Na	11	13	3	3	2	6.4
	RS	10.61	11.56	2.96	2.98	2	6.02
	Ne	9	9.48	1.41	1.53	1.88	4.66
	I	2.3	2.38	0.56	0.65	0.66	1.31
	He	0.89	0.89	0.29	0.35	0.47	0.58
	uHe	0.92	0.92	0.3	0.36	0.48	0.6
	Ho	0.33	0.75	0.2	0.41	0.38	0.41
	PR	0.01	0.01	0	0	0	0