Faculty of Science and Technology

Department of Life and Environmental Sciences

2018/2019

POPULATION GENETICS AND CONSERVATION OF THE ENDEMIC *MUS CYPRIACUS*

Francesca Riccioli

Thesis submitted in fulfilment of the requirements for the degree “Master of Research”, awarded by Bournemouth University

October 2019

Abstract

Endemic species have a higher risk of extinction due to habitat destruction, introduction of invasive species, pollution, or overexploitation. *Mus cypriacus* was first described in 2006 and is one of the two endemic rodents from Cyprus. It diverged from *Mus macedonicus* 0.53 million years ago, probably during the Mindel glaciation. Nowadays, *M. cypriacus* is mostly found in areas with vast cultivation at moderate altitudes (300-900 metres). Although, it could share habitat with *Mus musculus domesticus*, it is almost absent from urban areas or in areas with massive anthropogenic pressure. Even though *M. cypriacus* has been described to be of least concerned in the IUCN red list, there is lack of information on its ecology and demography, as well as a poor understanding of its genetic population structure.

Using the mitochondrial D-loop, single nucleotide polymorphisms and microsatellite data, I investigated the genetic diversity of *M. cypriacus*, the genetic structure of different *M. cypriacus* populations and tested for possible hybridisation between *M. m. domesticus* and *M. cypriacus.*

As expected, *M. cypriacus* was found to be closely related to *M. macedonicus using mitochondrial DNA*. No phylogeographic pattern was found for *M. cypriacus* on Cyprus with all the markers tested (mitochondrial D-loop, microsatellites and SNPs). The level of genetic diversity of *M. cypriacus* was comparable to the one found in *M. m. domesticus* (e.g. average number of alleles per loci 2.8 for *M. m. domesticus* and 2.5 for *M. cypriacus, based on the* SNPs). No genetic signature of hybridisation between *M. m. domesticus* and *M. cypriacus* was detected.

Overall, the data suggested that *M. cypriacus* is comprised of a single stable panmictic population. However, due to the small sample size, more research is needed to confirm these results. Furthermore, only little is known on the population size, population trends and the distribution of this species. Future work needs to estimate population sizes, provide a detailed species distribution map and be complemented with mark-release-recapture work to better understand the dispersal of the species.

Contents

List of Figures

List of Figures of Appendices - Supplementary material

List of Tables

List of Tables of Appendices - Supplementary material

Acknowledgements

First and foremost, I would like to express my thanks to my supervisors Emilie Hardouin, Demetra Andreou and Miguel Baltazar Soares for their advice, supports, and guidance throughout both the duration of this project and my degree.

Secondly, I am also grateful to Sven Künzel, for the RADseq library preparation; George P. Mitsainas, Eleftherios Hadjisterkotis, and Andreas Vakis for their help in the field, and Oxala Garcia Rodriguez for the msat and the mtDNA preparation. Without their help, this project would not have been possible.

A special thank goes to my friends and mostly to my family, who has helped and supported me in a number of different ways.

1. Introduction

1.1 Threats to endemic species

Human activity is considered one of the primary causes of environmental change, altering various ecological systems at a global level (Brooks *et al.*, 2006; Pysek *et al.*, 2010; Strassburg *et al.*, 2012). It has been predicted that up to 50% of species will be extinct in the next 50 years (Koh *et al.*, 2004; Thomas *et al*., 2004) due to human mediated climate change and destruction which has also been termed the “Anthropocene” (Crutzen, 2002; Zalasiewicz *et al*. 2011). The destruction of natural habitat has led to population declines and species extinctions (Burlakova *et al.,* 2011; Ceballos *et al*., 2015). For example, in the islands of Oceania, 1800 species of birds are predicted to go extinct in ~2000 years due to human colonization (Steadman, 2006). Such destruction is not only limited to terrestrial systems, for example, marine ecosystems, such as estuaries, coral reefs, and coastal and oceanic fish communities are rapidly losing populations and species (Worm *et al.,* 2006).

In 2004, the IUCN (International Union for Conservation of Nature) Red List reported that 7266 animal species are threatened with extinction (Baillie *et al.,* 2004). The various reasons for the decline and in particular the multiple threats to biodiversity including habitat destruction, invasive species, pollution, population and overexploitation (Baillie *et al.,* 2004; Frankham *et al.,* 2010; Primack, 2014). Worldwide, many species are threatened by one, several or all these factors, with endemic species being particularly vulnerable to these threats. According to Purvis *et al.* (2000) and Wilson *et al.* (2006), the percentage of extinction in a geographical place depends not on its total number of species but mostly on the presence of endemic species. According to different studies, endemic species have a higher risk of extinction (Vié *et al.,* 2008). For example, in the island of Madagascar, 100% of the lemurs, 99% of the frogs and 92% of plant species are endemic species (Vences *et al.*, 2009). However, nearly 80% of the land has been threatened by human activity, putting almost half of the species at risk of extinction (Myers *et al.,* 2000). A further threat to endemic species is invasive species. Alteration of environments and dispersal of invasive species by human activity have influenced the geographical and taxonomic trend of biological invasions at a global level (Gillespie *et al.*, 2007; Wilson *et al.*, 2009).

Cosmopolitan species have wide ranging distributions covering almost all continents (Fenchel and Finlay, 2004), whereas endemic species are distributed into a specific and particular area (Pimm *et al.*, 2014). Depending on their range of distribution, endemic species are named differently: continental endemic, regional endemic, national endemic, provincial endemic or local endemic (Primack, 2006). In remote islands, the majority of species are local endemic (Amori *et a*l., 2008; IŞIK, 2011; Irl *et al.,* 2017). The Hawaiian Islands make the perfect case study as more than 90% of plants and land birds are unique to the archipelago (Chen and He, 2009). However, endemic species can also be found in continental areas (Stebbins, 1942). Such is the case of the flowering plants fynbos in South Africa, where almost 70% are endemic (Hall, and Veldhuis, 1985; Rebelo, and Siegfried, 1990).

Endemic species are often vulnerable to extinction because they have a narrow geographic range (Bennett *et al.*, 2007; IŞIK, 2011) and fewer and smaller populations (Diethart *et al.,* 2004; IŞIK, 2011). To protect biodiversity, it is essential to identify those species most

vulnerable to extinction (Malcolm *et al.,* 2006). In order to facilitate this, the IUCN has established a number of conservation categories which can be used to determine the conservation status and need of a species. For example, species categorised as critically endangered (CR), endangered (EN), and vulnerable (VU) are considered to be threatened with extinction.

1.2 Hybridisation as a threat to endemic species

Hybridisation can occur between species that are not fully reproductively isolated and can in some cases lead to the formation of hybrid zones (Bouchemousse *et al.*, 2016). Hybridisation is determined as the interbreeding of individuals from two distinct species, which have one or even more distinct heritable traits (Harrison, & Larson, 2014; Taylor *et al.,* 2015). Hybridisation affects biodiversity in various ways, including the introgression of genetic variation within a new species, to the origin of novel hybrid species (Brennan *et al.,* 2014). Hybridisation has been reported in mostly all taxa (Pastorini *et al.,* 2009), and plays a fundamental role in the process of speciation (Dierking *et al.,* 2014). It decelerates or reverses differentiation; expedites speciation thanks to introgression or in the case of plants, leads to near-immediate speciation through allopolyploidisation (Abbott *et al.,* 2013).

Hybridisation can also lead to extinction, and it has been proved that the introduction of invasive species can cause hybridisation with natives leading to the decline or extinction of parental species (Todesco *et al.,* 2016). The literature has both examples of hybrid vigour (where hybrids are fitter compared to their parents (Gröning and Hochkirch, 2008) and outbreeding depression and consequent loss of locally adapted genotypes to the invasive species (Allendorf *et al.*, 2001; Perry *et al.* 2002; Muhlfeld *et al*. 2009). Endemic species can often have incomplete reproductive barrier to closely related species due to them evolving in the absence of such close relatives. As a result, the introduction of closely related species can lead to hybridisation with endemic species. Where the hybrids are fitter, these can lead to the decline and even the extinction of pure endemic species. Thus, hybridisation can be a serious threat to endemic species.

1.3 Cyprus and its endemic fauna and flora

Cyprus is the third largest island of oceanic origin located in the Eastern part of the Mediterranean basin. It emerged from the sea around twenty million years ago (Hadjisterkotis, 2001, Kryštufek and Vohralík, 2001). It was connected with mainland only during the Messinian salinity crisis, and therefore, it has been isolated for more than 5.3 million years (Kryštufek and Vohralík, 2001).

During late Pleistocene, Cyprus was home for pygmy hippopotamus (*Phanourios minutus*) and pygmy elephant (*Elephas cypriotes*); however, the human population caused of the final extinctions of the two endemic species (Hadjisterkotis et al., 2000; Marra, 2005). In the Neolithic and Chalcolithic ages, Cyprus started to host new mammals: the Mesopotamian deer (*Dama dama mesopotanica*), the European deer (*Dama dama*), the common deer *(Cervus elephus*), the ferret (*Mustela nivalis*), the wildcat (*Felis Silvestris*), the fox (*Vulpes vulpes*), the mufflon (*Ovis gmelini ophion*), the hare (*Lepus europaeus*) and various mice

species. However, many of these species did not survive to present times (Hadjisterkotis, 2001; Gippoliti and Amori, 2004).

Even if it has been considered a biodiversity hot-spot area, little was known regarding the mammalian fauna (Kryštufek and Vohralík, 2001). The first book regarding the mammals of Cyprus was written by Unger and Kotschy (1865), reporting only 8 species; however, some species were confused with close relatives (e.g. *Erinaceus europaeus* and *Hemiechinus auritus*; *Lepus timidus* and *Lepus europaeus*). In 1879, *Pipistrellus kuhlii* and *Rattus rattus* were added to the list of mammals of Cyprus by Günther. In 1903, the native spiny mouse *Acomys nesiotes* was added by Bate D.

Almost 40% of the land is covered by forests, along with garigue and maquis vegetations. Cyprus hosts numerous endemic species, due to its long-time of isolation (Hadjikyriakou and Hadjisterkotis, 2002). One hundred twenty-eight species of plants are known to be endemic to Cyprus (Tsintedes and Kourtellarides 1998). More than 350 species of birds can be found on Cyprus, most of them are migratory, and about ten species are endemic (Whaley and Dawes, 2003). The Cyprus whip snake (*Dolichophis cypriensis*) and the troodos lizard (*Phoenicolacerta troodica*) are the only endemic reptiles. Thirty mammal species (Kryštufek and Vohralík, 2001), 25 amphibian and reptile species, 11 lizard species (4 of which are endemic) (Baier *et al.,* 2009) 2 turtle species (McGowan et al., 2001), 250 fish species and about 6000 insects (Violaris *et al.,* 2009), have been counted in Cyprus. Over the years, more species have been added to its list of species, including numerous species of bats, shrew, the Cypriot mouse*,* the house mouse, owl pellets, deer (Kryštufek and Vohralík, 2001).

Cyprus has a Mediterranean climate, with hot and dry summers, and rainy and mild winter. However, due to climate change there has been an increase in temperature and a decrease of rainfall resulting in the desertification of the environment (Tsangari *et al.*, 2016). There is also a shortage of water, and, inside the forest, spring water increased its evaporation (Hadjinicolaou *et al.,* 2011). All those issues are leading to a lack of food, which consequently forces the animals to leave their natural habitat inside the forest, and go in areas where there are domestic sheep and goats (which can be a source of transmittable diseases) (Hadjisterkotis, 2001). Further threats to biodiversity in Cyprus are due to high development efforts in hotel building, luxury apartments, villas and golf club, abolishing or altering natural environment mostly near the sea (Hadjimitsis, 2010; Zachariadis, 2012; Welz, 2015). Other issues for the island are the extensive fires, which destroy broad areas of the forest and agricultural land (Ciesla *et al.,* 2004). Lastly, the introduction of alien species of plants and animals is also taking place in Cyprus, putting at risk the ecosystems and the extinction of native species. For example, in 1990, five wild boars (*Sus scrofa*) were introduced in Cyprus from Greece for game farming (Hadjikyriakou and Hadjisterkotis, 2002). However, in 1995 they were released into the wild. In 1995, 60-90 individuals were estimated to be present in the forest, and nowadays the number is not known; according to the Red List of Threatened Species, wild boar threatened 19 taxa.

1.4 Description of *Mus cypriacus*

The Cypriot mouse, *Mus cypriacus,* is endemic to Cyprus. It diverged from *Mus macedonicus* 0.53 million years ago, probably during the Mindel glaciation (Cucchi *et al.*

2006, Macholán et al 2007). Before being recognized as *M. cypriacus*, it was reported as *Mus spicilegus "South"* (Orsini *et al.,* 1983), *M. spretoides* (Bonhomme et al., 1984; Auffray *et al.,* 1990), *M. abbotti* (Cheylan, 1991), and *M. macedonicus* (Harrison & Bates, 1991; Kryštufek & Vohralík, 2001; Musser & Carleton, 2005). Only in 2004, the Cypriot mouse was recognized as an independent phylogenetic lineage by Bonhomme *et al*. *Mus cypriacus* was only described in 2006 by Cucchi *et al.,* based on a comparative analysis with other mouse species from Europe, using both D-Loop mitochondrial sequences, cranial and dental morphometry.

Mus spp. fossil, dated back to the Pleistocene, were found along with fossils of large endemic mammals in Cape Pyla (Reese, 1999; Kryštufek and Vohralík, 2001); *M. cypriacus*, was the only small mammal endemic of Cyprus (Kryštufek and Vohralík 2009; Cucchi *et al.,* 2012), with the exception for dwarfs hippos (*Phanourios minutus*) and elephants (*Elephas cypriotes*) as well as by a genet (*Genetta* cf*. plesictoides*). Nowadays, three of the small mammals' species on Cyprus are endemic, *M. cypriacus, A. nesiotes* and *Crocidura cypria* (Kryštufek *et al.,* 2009).

M. cypriacus is morphologically similar to *M. macedonicus*; however, *Mus cypriacus* is, on average, larger and has relatively longer tail, but ranges overlap. It has been noticed that the tail of the Cypriot mouse is mostly longer than head and body (Macholán *et al.,* 2007). The upper part of the body is frequently brown, the belly is greyish-cream or greyish-buff; the feet are white and ears are brown; regarding the tail, it varied for each individual, grey brown, greyish or buff-grey above, pale grey to greyish- white below (Figure 1; Kryštufek *et al.,* 2009).

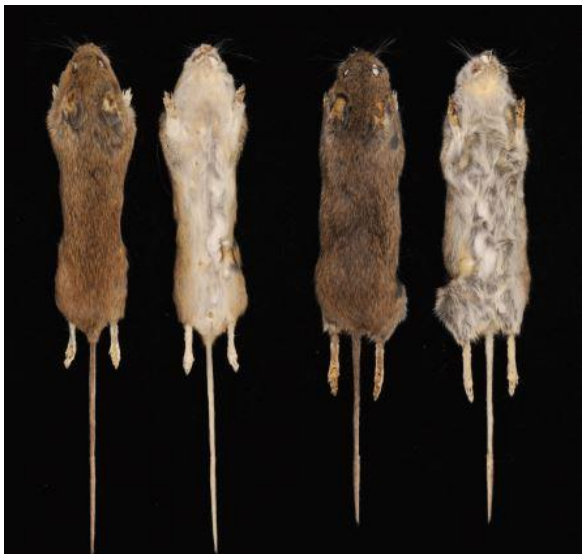


Figure 1 – Skins of *Mus cypriacus* from Paramytha, Limassol, in dorsal and ventral view (Kryštufek *et al.*, 2009)

Nowadays, *M. cypriacus* has mostly been found in areas with vast cultivation at moderate altitudes, 300-900 metres up to 1605 meters (Macholán *et al.,* 2007). However, it has also been found in areas near rivers, where lives sympatrically with the *M. m. domesticus* (IUCN 2019). It is almost absent in urbans areas or in area with a massive anthropogenic pressure (Cucchi *et al.,* 2006; Kryštufek *et al.,* 2009). The species has been listed as Least Concern on the IUCN list (Amori, 2017). However, the IUCN report has identified that there is a need to increase research in its population size, distribution trends, life history, ecology, and threats.

1.5 Invasive Mus species on Cyprus

Cyprus is considered the first island of the Mediterranean colonised by the house mouse, *M. m. domesticus* (Cucchi *et al.* 2002). The house mouse originated in a geographic area which encompassed west central Asia and the northern Indian subcontinent (Hardouin *et al.* 2015, Boursot *et al.* 1993). The house mouse ancestry split into three different sub-species (*M. m. domesticus, M. m. musculus, M. m. castaneus*) around 0.9 million years ago (Boursot *et al.* 1993). *M. m. castaneus* is found in East Asia, *M. m. musculus* in Asia and Eastern Europe and *M. m. domesticus* in Western Europe (Boursot *et al.* 1993). *M. m. domesticus* and *M. cypriacus* co-occur in Cyprus since the Neolithic and have been found sharing the same habitat (Cucchi *et al.,* 2006; Kryštufek *et al.,* 2009; IUCN 2019).

1.6 Hybridisation between Mus species

In Europe there are 6 *Mus* groups: *M. spretus* (endemic to the Mediterranean area an also present in South West Europe and Northern Africa), *M. macedonicus* (in south Balkans, Asia Minor, the Caucasus and in the Middle East), *M. spicilegus* (seen in Slovakia, Hungary, Serbia, Bulgaria, Moldova, and Ukraine), *M. m. musculus* (Eastern Europe) , *M. m. domesticus* (Western Europe) and *Mus cypriacus* (endemic to Cyprus) (Cucchi *et al.,* 2005).

Hybridisation incompatibilities have been well reported between *M. m. domesticus* and *M. m. musculus* (e.g. Turner and Harr 2014), yet those two species are not entirely genetically isolated and can hybridise. In Europe, there is a hybrid zone (located between Bulgaria to Denmark) between *M. m. domesticus* and *M. m. musculus* (Sage *et al*., 1986; Boursot *et al.,* 1993). In Japan, *M. m. musculus* and *M. m. castaneus* have hybridised producing a new species of the house mouse: *M. m. molossinus* (Yonekawa *et al.*, 1998). Furthermore, *Mus spretus*, *Mus spicilegus* and *Mus macedonicus*, even if considered sympatric with the *M. musculus domesticus* subspecies, have been reported to produce hybrids in natural environment (Guenet & Bonhomme, 2003).

Hybridisation between *M. m. domesticus* and *M. spretus* has also been reported mainly for the immunity to the pesticide warfarin. It has been discovered that *M. m. domesticus* from Spain and Germany have the whole or partial *vkorc1* gene of *M. spretus* providing them protection to warfarin (Song *et al.* 2011).

1.7 Conservation genetics

Conservation efforts must always encompass the genetic health of populations (Deem *et al.,* 2001). Molecular markers have extensively been used in conservation genetics (Schlötterer, 2004; Schwartz *et al.,* 2007; Beebee, 2018). Different markers provide different information on the genetic history and health of the studied populations. The molecular markers used in this study are reviewed below:

- **Mitochondrial DNA**

Mitochondrial DNA (mtDNA) is a widely used genetic marker thanks to its maternal inheritance and high abundance in cells, which make it easy to DNA extract and amplify

(Heggenes *et al*., 2016). MtDNA is a haploid molecule (Baker 2000), has vast intragenomic variability, and, depending on the region that is used, high substitution rates (Heggenes *et al.,* 2016). It has a high mutation rate, compare to nuclear markers, because of lack of repairing mechanisms, and only some regions of the D-Loop accumulate free mutations (Wanga *et al.,* 2015). Thanks to those factors, it is considered a suitable maker to study the origin and the evolution of species and it is also used as a marker for phylogenetic analysis, genetic variation and relatedness among species (Silva *et al.,* 2009; Lakra *et al.,* 2010; Borrel *et al.,* 2012). Mitochondrial control regions or D-loop (from the name "displacement loop") is known as a non-coding control region; and its structure is formed when a DNA double helix is invaded by a single-stranded DNA or RNA molecule, which creates a region of base pairing with one of the polynucleotides of the helix (Reyes *et al.*, 2004; Gupta *et al.,* 2015). The D-loop is particularly essential due to the high presence of mutations at a nearly neutral rate; furthermore, it contains transcription and replication elements which act as a detector for cellular DNA damage (Greider, 1999). The other two widely used mtDNA markers are the protein-coding cytochrome b (*cytb*) and cytochrome c oxidase subunit 1 (COI) regions (Ursenbacher *et al.* 2006; Kvie *et al.* 2013).

During the early 1960s, the molecular clock hypothesis was firstly introduced; with the assumption that substitutions (which are those mutations that do not undergo through the repair processes and results in permanent changes in a DNA sequence) happen at a constant rate (Brown, 2002; Bromham, 2009). In mammals, the molecular clock for mtDNA is faster than that for the nuclear DNA because of the nucleus DNA repair mechanism which are absent in mitochondria (Birky *et al.,*1989; Bromham and Penny, 2003). Therefore, mtDNA has a rapid evolutionary rate and can be used to detect events that happened at a longer time scale (Nabholz *et al.,* 2008).

- **Restriction site associated DNA (RAD) Sequencing**

Next-Generation Sequencing (NGS) has revolutionised molecular biology (Ekblom and Galindo 2011). In particular, NGS has the ability to produce gigabases of genetic data easily and at reasonably low cost (Hudson, 2008). Restriction site associated DNA (RAD) sequencing is one of the NGS based approach (Davey *et al.* 2013).

Restriction site associated DNA sequencing (RADseq) is a reduced representation genome sequencing strategy, created to examine anywhere from 0.1 to 10% of a selected genome. RADseq works by first fragmenting the target genome using restriction enzyme (Arnold *et al.,* 2013). After digestion, a series of molecular processing steps modify the DNA into a fragment library proper for sequencing on an NGS platform. The use of restriction enzyme to cut the DNA into fragments and the use of molecular identifiers to link sequence reads to particular individuals (Baird *et al.,* 2008). RADseq generates thousands of single nucleotide polymorphisms (SNPs) from any organisms (Hoffman *et al.,* 2014). Mutation rate for SNP markers is significantly lower than for microsatellites (Kronholm *et al.,* 2010), and they evolve in a manner described by mutation models, such as the infinite sites model (Vignal *et al.*, 2002; Morin *et al.,* 2004). RADseq can be used to carry out studies in conservation biology, phylogenetic and phylogeography (McCormack *et al.* 2013). RADseq usually generates thousands to tens of thousands of single nucleotide polymorphisms (SNPs) makers (Davey *et al.,* 2013), which allow to identify genetic signatures with better outcome compare to

microsatellites markers and mitochondrial genes (Emerson *et al.,* 2010). Even if RADseq is considered an ideal strategy for conservation genetics studies, the markers that are generated during the analysis must be carefully treated, in order to separate high quality markers from the possibly biased (Davey *et al.,* 2012). Allele dropout and null alleles can be caused by the presence of polymorphisms in restriction sites (Arnold *et al.,* 2013; Gautier *et al*., 2013); PCR duplicates, that can cause genotyping errors, deviating allele frequency and lead to false positive alleles (Andrews *et al.,* 2014); errors in the sequence and/or shorter length fragment can cause fewer loci (Andrews *et al*., 2016); and lastly, study of expected and observed heterozygosity can be truly complicated when a low coverage and a high percentage of missing values is identified (Hodel *et al*., 2017). However, many of those flaws can be controlled or filtered with bioinformatics pipelines (Davey *et al.,* 2012). In conclusion, for this research, RADseq has been used to test the level of genome-wide heterozygosity for *M. cypriacus* as well as F_{ST}, linkage disequilibrium, and population genetics and structure of *M. cypriacus* and *M. m. domesticus.*

- **Microsatellites**

Microsatellites markers, simple sequence repeats (SSRs) or short tandem repeats (STRs), are widely used for testing genetic diversity and population genetic structure (Bhargava & Fuentes, 2010; Guichoux *et al.,* 2011; Putman & Carbone, 2014). Microsatellites are regions of noncoding DNA with many simple identifiable alleles. The alleles are distinguished by the number of times a short sequence of nucleotides is repeated. Short tandem DNA repeats units are usually 2-6 base pairs length that are randomly distributed in the nuclear genome (Bhargava & Fuentes, 2010). The numbers of repeats evolve over time (Guichoux *et al.,* 2011).

Lots of user-friendly software are available for investigating population genetic analysis using microsatellites; furthermore, microsatellites are easy and low-cost to implement thanks to the vast abundance of primers available from previous studies (Sunnucks, 2000; Väli *et al.,* 2008; DeFaveri *et al*., 2013).

Microsatellites are considered selectively neutral as they do not influence phenotypic expression, therefore, they are ideal markers for population analysis; meaning that they can give genetic signatures without being influenced by natural selection (Silvertown and Charlesworth, 2001). Microsatellites can be highly polymorphic even in small populations, because of the numerous amounts of mutation due to slippage during DNA replication (Schlötterer, 2000; Navascués and Emerson, 2005). Microsatellites markers are commonly used to detect population genetic structure or parentage analyses (Fischer *et al.,* 2000; Ouborg *et al.,* 2010), as well as hybridisation and introgression (Randi, 2008; Trigo *et al.,* 2013; McIntosh *et al.,* 2014), (Lexer *et al*., 2007), parentage analysis (Jones and Ardren, 2003), and population demography (Sakaguchi *et al*.,2013). However, there are some possible issues using microsatellites (Hoffman & Amos, 2005), such as large allele dropout (Miller *et al.,* 2002; Johnson and Haydon, 2006), null alleles (Callen *et al*., 1993; Pemberton *et al.,* 1995; Dakin and Avise, 2004), homoplasy (Grimaldi & Crouau-Roy, 1997; Estoup *et al.*, 2002) and unclear mutational mechanisms (Ellegren, 2004; Selkoe and Toonen, 2006). Both nuclear microsatellite loci and mitochondrial DNA sequences represent rapidly evolving DNA sequences that are informational for answer questions relative to population

level (Vignal *et al.*, 2002). Despite this, the high information content, produced by high mutation rate, can cause several limitations on subsequent data analysis (Morin, *et al*., 2004). The much higher mutation rate of microsatellites, estimated to be as high as 1 × 10−5 Kruglyak *et al.,* 1998) when compared to the 1 × 10−9 for SNPs (Martinez-Arias *et al.*, 2001) can be a cause of concern, particularly when studying for linkage disequilibrium and association (Vignal *et al.*, 2002). For this research, microsatellites have been used to test the level of heterozygosity for *M. cypriacus* as well as F_{ST}, linkage disequilibrium, and population genetics and structure of *M. cypriacus* and *M. m. domesticus,* in order to compare them with the results obtained from the RASseq. And lastly, microsatellites have also been used to detect genetic signature of recent bottlenecks in *M. cypriacus.*

2. Aims, objectives, and study design

2.1 Aims and objectives

Little is known regarding the environmental threats faced by *Mus cypriacus.* However, as an endemic species, it most probably faced with habitat destruction, invasive species, pollution, population size and overexploitation.

The aim of this study was to characterise the genetic diversity of *M. cypriacus* in Cyprus in order to investigate the genetic health of its populations, determine population connectivity and investigate potential hybridisation with the invasive house mouse *M. m. domesticus* using three different markers. To achieve this, two specific objectives were addressed.

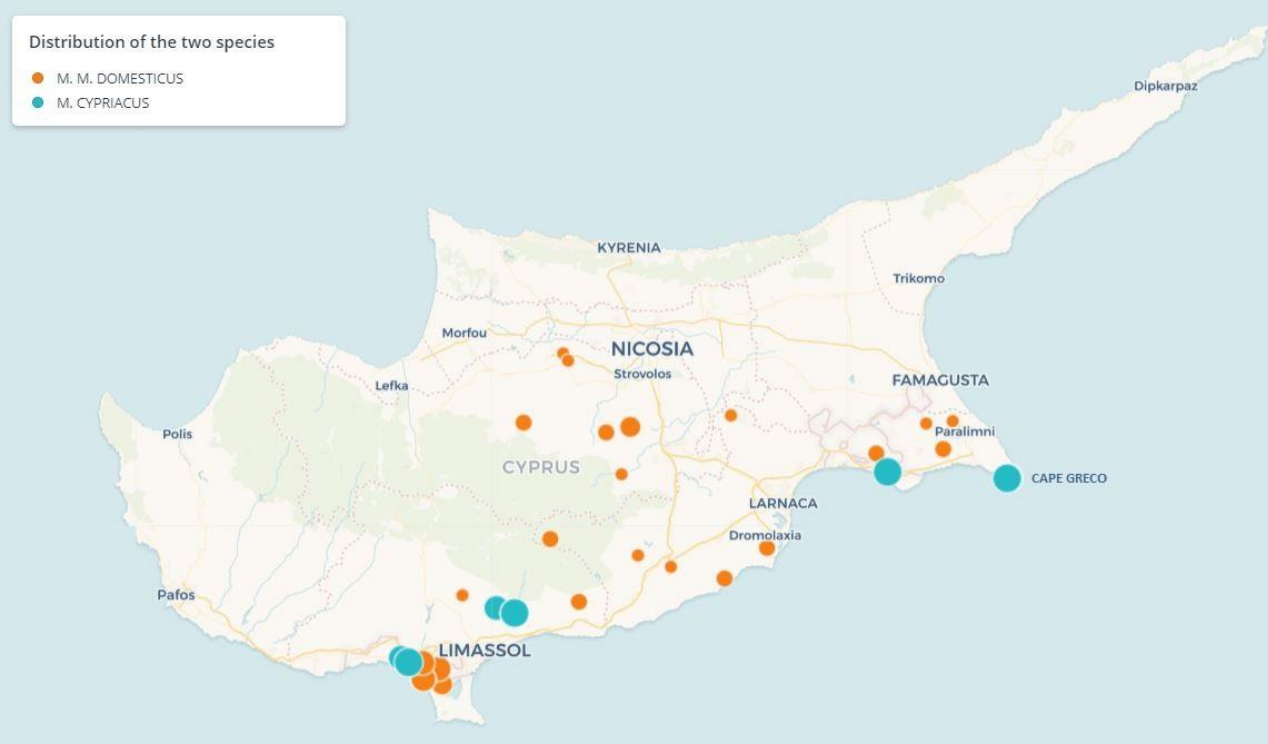


Figure 2 – ***Mus musculus domesticus* and *Mus cypriacus* in Cyprus** – the map above shows the distribution of the samples collected for this research in Cyprus. The *M. m. domesticus* are represented by the orange circles and the turquoise circles markers represent *M. cypriacus*

O1: investigate the genetic diversity and structure of *M. cypriacus* in Cyprus using the mitochondrial D-loop, RADseq and microsatellites.

O2: Infer possible hybridisation between *M. m. domesticus* and *M. cypriacus*, using microsatellites and RADseq, as it can be considered as a threat for endemic species.

2.2 Study Design

Three different molecular markers were used to investigate the population genetics and conservation of the endemic *M. cypriacus (figure 3).*

(1) The phylogeny of the Cypriot mouse was inferred using mitochondrial D-loop data. An analysis of phylogenetic inference and phylogenetic network was conducted

(2) The population structure and a possible hybridization were analysed with both RADseq and microsatellites. The results obtained from the two markers were compared.

(3) Lastly, the population demography of *M. cypriacus* was analysed using mitochondrial D-loop, to run the mismatch distribution, and microsatellite loci to run the bottleneck analysis.

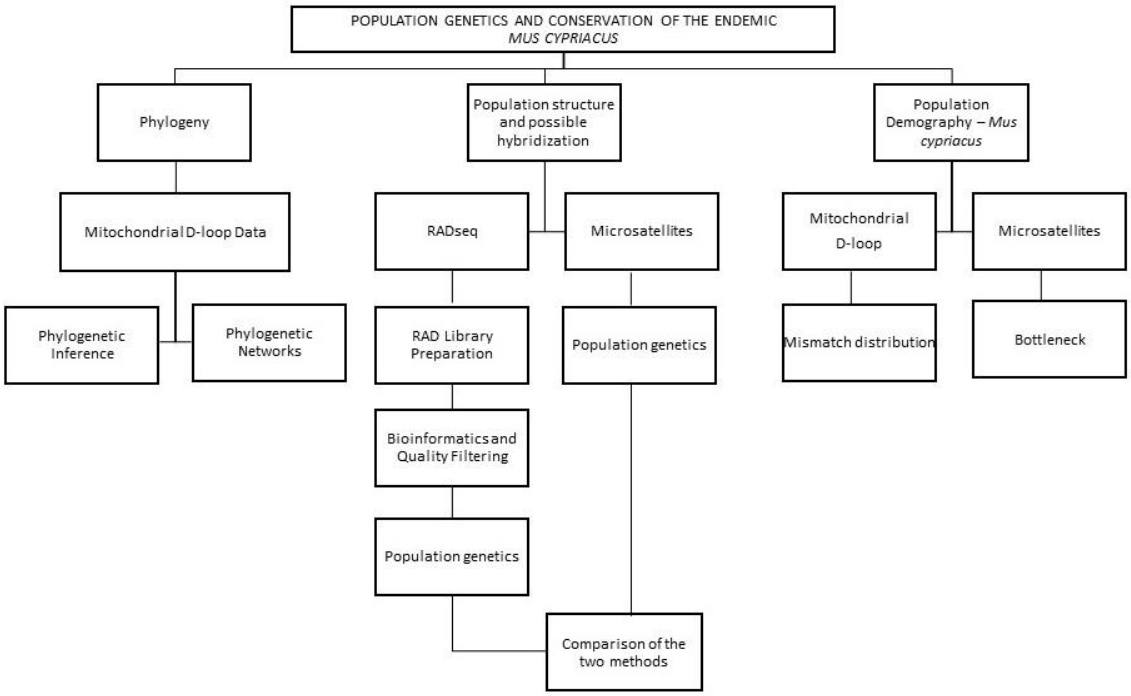


Figure 3 – ***Study design*** – The above flow chart illustrates all the procedures that have been used to investigate the genetic health of *M. cypriacus* populations, determine population connectivity and investigate potential hybridisation with *M. m. domesticus* using three different markers.

3. Materials and Methods

3.1 Sampling:

A total of 13 samples of *Mus cypriacus* were collected for the study as by caught while sampling for *Acomys nesiotes in* 2015 from 3 locations across the island (Figure 4). *M. cypriacus* samples were compared to 41 samples of *Mus musculus domesticus* (Figure 4). Overall, a total of 54 samples were collected from 39 localities across Cyprus (Figure 4), sampling scheme is described in Garcia-Rodriguez *et al.* 2018. All the samples were collected following local regulations for field collection of small mammals.

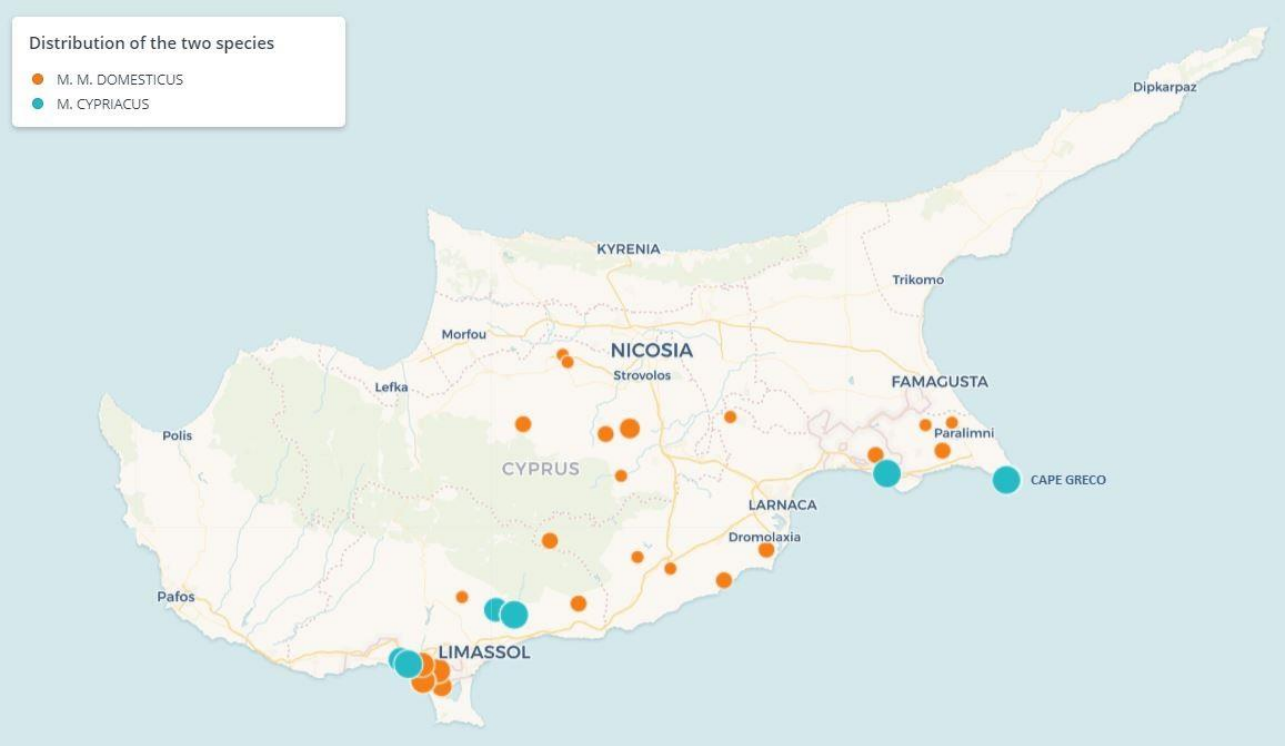


Figure 4 – ***Mus musculus domesticus* and *Mus cypriacus* in Cyprus** – the map above shows the distribution of the two species in Cyprus. The *M. m. domesticus* are represented by the orange circles and the turquoise markers represent *M. cypriacus.* The size of the circles represents the number of samples caught in a specific locus. The smallest circles show only one sample, and they increase in sizes if more than one sample has been found.

3.2 Mitochondrial D-loop – Phylogenetic analysis

A phylogenetic reconstruction was performed using a total of 54 Mitochondrial D-loop (= control region). Three *Mus musculus domesticus* from Cyprus (García-Rodríguez *et al.,* 2018), 13 *Mus cypriacus*, 2 *Mus spretus* (GenBank: MK089345, MK089344), 2 *Mus musculus castaneus* (GenBank: AB649628, AB649629), 2 *Mus musculus musculus* (GenBank: KR866365, KR866364), 2 *Mus macedonicus* (GenBank: AF506193, AF506192), and 27 sequences retrieved from Genbank of *Mus cypriacus* (EU106194- EU106281). One sequences of *Rattus rattus* (Genebank: HQ334447) and two sequences of *Rattus norvegicus* (Genebank: X04733, X04734) were used as outgroup. Overall, 54 sequences were firstly trimmed to the same size (809 bp) after visual inspection, using BioEdit v.7.0.4 software (Hall, 1999).

Out of the 54 sequenced analyses, 41 haplotypes were found and therefore, were used to construct the final phylogenetic tree. Phylogenetic analysis was performed using MrBayes 3.2 (Ronquist *et al.,* 2012) with MCMC = 2 000 000. The first 25% trees were rejected as burn-in, with the remaining trees being used to create the consensus tree. The Bayesian inference uses the Markov chain Monte Carlo (MCMC) algorithm, which forms a posterior distribution [an accumulation of approximately 1000 phylogenetic trees that illustrates the unsureness regarding the evolutionary relationships within a set of sequences (Lanfear *et al.*, 2016)]. The MCMC algorithm examines the space of all the plausible phylogenetic trees, systematically registering the trees it comes across (Aberer *et al.* 2014; Bouckaert *et al.* 2014)

Phylogenetic networks were created using the software PopART (Population Analysis with Reticulate Trees), evaluated with the median-joining option (Leigh and Bryant, 2015).

3.3 Mitochondrial D-loop – Population Demography

Out of the 40 sequences analysed (13 samples of *Mus cypriacus* and 27 sequences of *Mus cypriacus* retrieved from Macholán *et al.,* 2007), 29 haplotypes were found. Genetic diversity indexes (haplotype number and estimation of nucleotide polymorphism) were calculated using DNAsp (Librado and Rozas 2009). Two standard neutrality tests, Tajima”s D (Tajima, 1989) and Fu”s FS (Fu, 1997), were also tested using DNAsp to test for potential deviation from selection neutrality and/or recent population expansion or decline (Librado and Rozas 2009). Tajima”s (1989) D test compute the discrdance between the estimate of theta from various segregatin sites and from avarage pair-wise sequence diverence. The negative value can be interprete as a signal of purifying selection or as demographic expansion. Fu’s (1997) calculates the possibility of observing a certain number of haplotypes, given particular value of theta. The test works by evalueting the discordance in values of theta derived from number of haplotypes and average pair-wise sequence divergence. Same for the Tajima’s D test, negative value can be interprete as a signal of purifying selection or as demographic expansion. Differences in theta summary statistics, based on different population genetic analysis, will detect demographic changes. Demographic changes, in fact, can be identify thanks to the distribution of the allelic frequencies. The mismatch distribution is a frequency graph of pair-wise differences between haplotypes.

Mismatch distribution was calculated with DNAsp (Librado and Rozas 2009). The observed values were compered against the expected from the population expansion model with parameters estimated using the generalized nonlinear least-squares approach of Schneider & Excoffier (1999) using Arlequin software v. 3.5.2.2 (Excoffier and Lischer *et al.,* 2010). The population growth-decline analysis is based on three parameters: Theta Initial θ_0 (theta before the population Growth or Decline), Theta Final θ_1 (theta after the population Growth or Decline), and τ (Tau) is the date of the Growth or Decline measured in units of mutational time (Rogers and Harpending 1992).

3.4 RAD Library Preparation

Two sets of run of RADseq libraries for the 54 samples (13 samples of *M. cypriacus* and 41 samples of *M. m. domesticus,* Figure 5) were prepared following the protocol of Etter *et al.* (2011). This protocol is to make reduced complexity genomic libraries that are individually labelled and pooled for sequencing on an Illumina MiSeq based on modified ddRAD protocols (Peterson *et al.* 2012). The library construction is based on an efficient combined restriction digest/adaptor ligation. In this case, the restriction enzymes *Csp6I* (which cleaves 5”- G^TAC -3”sites) and *PstI* (which cleaves 5”- CTGCA^G -3” sites) were chosen to digest genomic DNA.

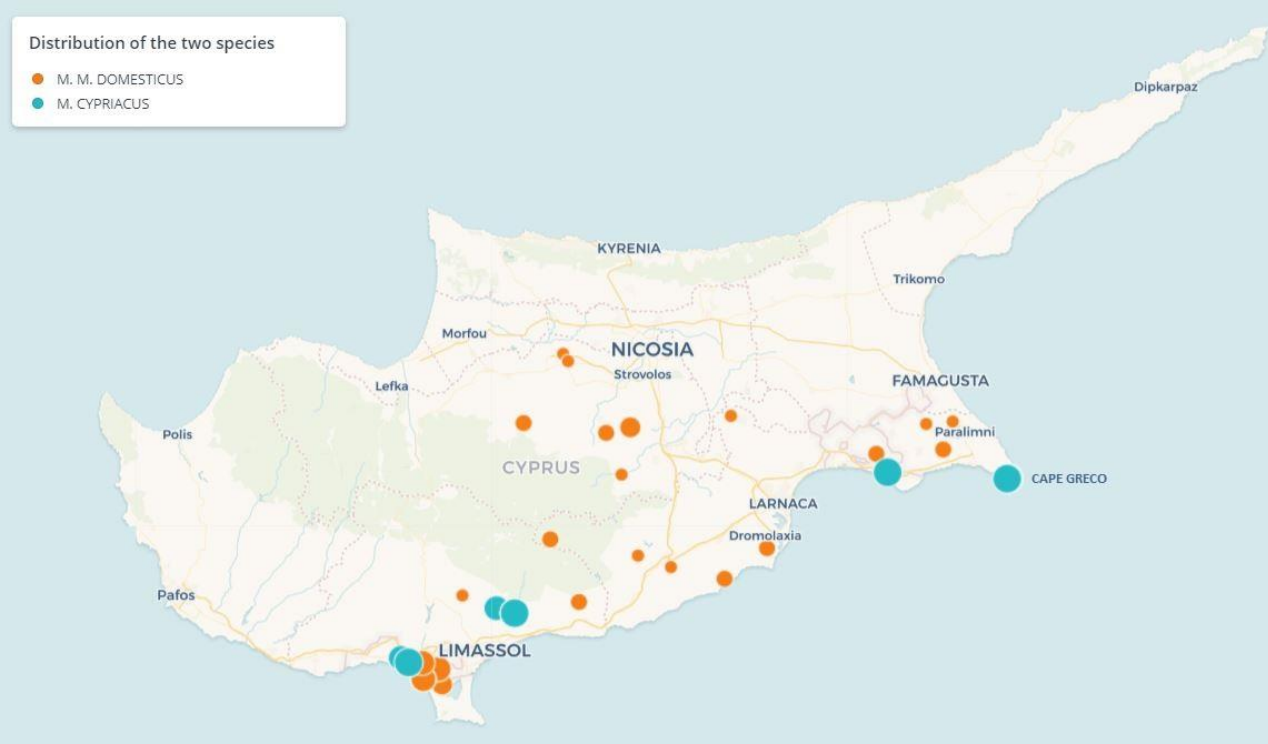


Figure 5 – ***Mus musculus domesticus* and *Mus cypriacus* in Cyprus RADseq**– the map above shows the distribution of the two species in Cyprus. The *M. m. domesticus* are represented by the orange circles and the turquoise markers represent *M. cypriacus.* The size of the circles represents the number of samples caught in a specific locus. The smallest circles show only one sample, and they increase in sizes if more than one sample has been found.

3.5 Bioinformatics and Quality Filtering

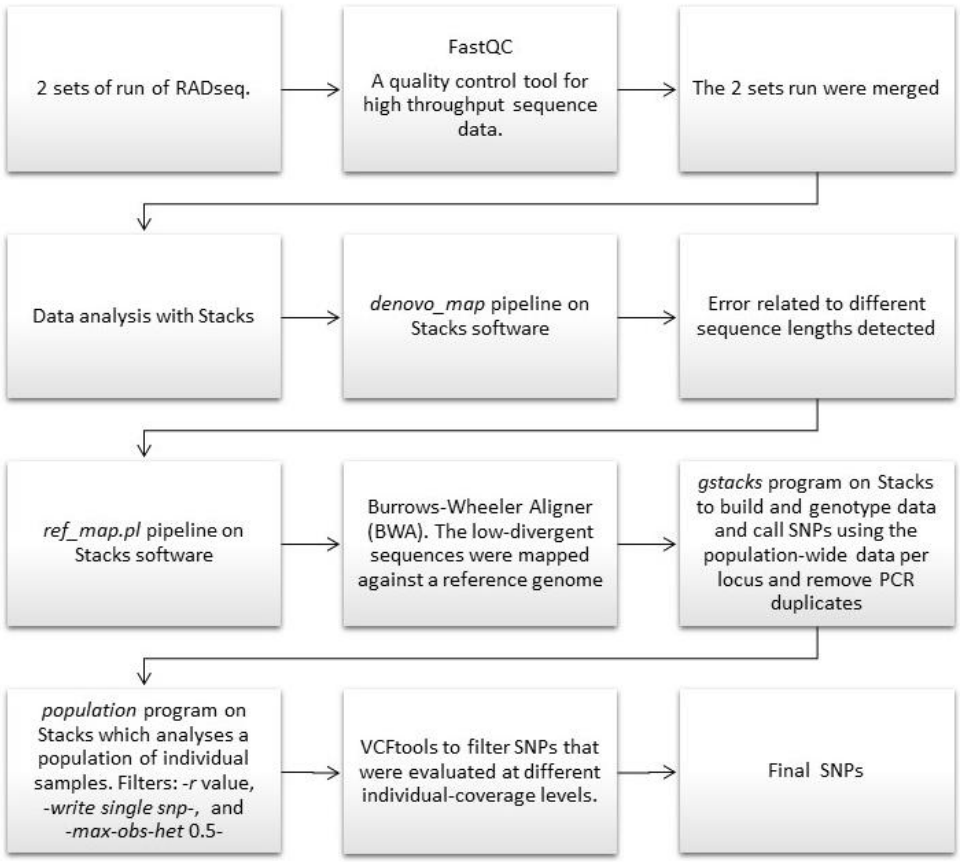


Figure 6 – **Bioinformatics and quality filtering pipeline**– A flowchart of the pipeline followed for the bioinformatics and quality filtering. First, reads quality was checked with FASTQC. Then, the two sets of runs were merged. The next stages were done using Stacks pipelines. First *denovo_map* pipeline was considered and tested. However, due to an error, *ref_map.pl* pipeline was used to conduct the analysis. A standard alignment program that incorporates Burrows-Wheeler algorithm, BWA, was used to align the sequences against a reference genome. In the next stage, the genotypes program was executed, *gstacks* programme, to generate loci by combining single- or paired- end reads that have been aligned against the reference genome and sorted. Then, the populations program (*population*) tabulates the state of loci within and among populations, calculates population genetics statistics and exports to a number of additional, useful formats. Last, VCFtools was used to filter SNPs that were evaluated at different individual-coverage levels.

RADseq data quality was checked using FASTQC version 0.11.8 (www.bioinformatics.babraham.ac.uk/projects/fastqc/). Data analysis was conducted using Stacks software (Catchen *et al*., 2013 b). During the first analysis, only one set of run of RADseq was considered, to investigate the data and the various parameters (Figure 6).

After, a secondary sequencing was run, during which the two sets of run of RADseq were merged, with the final aim to obtain a larger number of SNPs and robust loci (Figure 6).

At the beginning of the analysis, the *denovo_map* pipeline was taken in consideration to execute the Stack pipeline. The program works by executing the Stacks pipeline by running each of the Stacks components individually (Catchen *et al.* 2013). Different parameters were

used for this analysis: "*-M 4*" [number of mismatches allowed between stacks within individuals (for ustacks)], "*-n 2*" [number of mismatches allowed between stacks between individuals (for cstacks)], "*-T 15*" (the number of threads/CPUs to use) and "*-X*" [additional options for specific pipeline components (in this case *"ustacks : -m 3"*)]. As mention previously, the *denovo_map* programme performs several stages: *ustacks, cstacks, sstacks, tsv2ban, gstacks* and *populations*.

Due to an error revealed during the first analysis (*denovo_map* pipeline) related to different sequence lengths detected, data analysis was conducted using *ref_map.pl* pipeline on Stacks software (Catchen *et al*., 2013 b). During this analysis, the samples were aligned against a reference genome of the *Mus musculus domesticus* (GenBank: KV417259), using a standard alignment program known as Burrows-Wheeler Aligner, (BWA) [(Figure 6) Carver *et al.,* 2010]. BWA is a software package for mapping short low-divergent sequences against a reference genome and consists of three algorithms: BWA-backtrack, BWA-SW and BWA-MEM. It works by first constructing the FM-index for the reference genome and then, the chosen aligned algorithm is invoked with a -sub-command. For this research, the BWA-MEM algorithm was chosen. It first seeds alignments with maximal exact matches (MEMs) and then extending seeds with the affine-gap Smith-Waterman algorithm (SW).

Once the RADseq were aligned against the reference genome, using the *ref_map* pipeline and a defined population map, the *gstacks* module was used (Figure 6). The program generates loci by combining single- or paired- end reads that have been aligned against the reference genome and sorted (Catchen *et al.* 2013 b.). When *ref_map* analysis is run, the *gstacks* is the first program executed and will generate loci by combining single- or paired-end reads that have been aligned against the reference genome and sorted (Catchen *et al.* 2013 b.).

After, the *population* module was used to call genotypes, calculate population statistics, F-statistics (Figure 6). During this analysis, different parameters were tested to investigate changes in the number of SNPs. 1: the *-r* value-, which consists of the minimum percentage of individuals in a population required to process a locus for that population. 2: the use -*write_single_snp*- restrict data analysis to only the first SNP per locus, to avoid linkage between markers. 3: the presence of a population map (which consists of the prefix of each sample in the analysis in the first column, followed by an integer or string in the second column indicating the population). Furthermore, the *population* program can export data directly for a vast variety of analysis program (Catchen *et al*. 2013 b.). During this research, the script produced genotype output in multiple formats, i.e., STRUCTURE-format file and GENEPOP-format file (Catchen *et al.,* 2013 a.; Larson *et al.*, 2014; Munshi-South *et al*., 2016; Stobie *et al.,* 2019).

Various combinations of *-r* (0, 0.25,0.50, 0.66, 0.75) parameters were tested along with the presence or absence of a defined population map, to investigate changes in the number of SNPs obtained, and in the percentage of missing values among samples (Table 1). When 25% of individuals in the population were required to process a locus for that population (*r* =0.25), 8889 SNPs were reported, with a total of 80.17% of missing data. After, 50% of individuals in the population were required to process a locus for that population (*r* =0.50), showing 443 SNPs, with a total of 62.16 % of missing data. Lastly, 65% of individuals in the population were required to process a locus for that population (*r* =0.65), reporting a total of 158 SNPs with a total of 43.71% of missing data.

During the final analysis of the *population* module, the *-max-obs-het* 0.5- specifies a maximum observed heterozygosity required processing a nucleotide site and locus was added as a parameter, along with *-r* value equals to 0.25, reporting a total of 5323 SNPs.

"r" value	0	0.25	0.50	0.65	0.75
WITHOUT DEF. POP.					
Loci	605012	9820	458	187	132
Variant sites remained after filtration	16382	534	127	76	56
WITH DEF. POP.MAP. (1dataset)					
Loci	N/A	23953	2541	N/A	N/A
Variant sites remained after filtration	N/A	3747	228	N/A	N/A
WITH DEF. POP.MAP. (2datasets)					
Loci	N/A	49072	7652	1642	N/A
Variant sites remained after filtration	N/A	8890	443	158	N/A

Table 1– Analysis of «r» value, which help filtering data- corresponds to the minimum percentage of individuals in a population required to process a locus for that population. The data were analysed without a defined population map and with a defined population map, considering 0%, 25%, 50%, 65% and 75% of individuals at each site. The variant sites remained after filtrations correspond to the SNPs.

VCFtools (http://vcftoools.sourceforge.net/) software was then used to filter SNPs that were evaluated at different individual-coverage levels (Figure 6) with "*--max-missing"* as a parameter, for both species together. For this study, missing data tested weret 85% ("*--max-missing 0.85")*, 80% ("*--max-missing 0.80"),* and 75% ("*--max-missing 0.75")*

A set of specific *M. cypriacus* SNPs was also obtained excluding individuals with more than 75% of missing data ("*--max-missing 0.75"* parameter).

3.6 RADseq analysis

The heterozygosity and the mean number of alleles per locus were calculated using GENETIX 4.03 (Belkhir *et al.,* 2004). Populations pairwise F_{ST} and linkage disequilibrium were calculated with Arlequin software v. 3.5.2.2 (Excoffier and Lischer, 2010), for the both *M. cypriacus* and *M. m. domesticus* together and for *M. cypriacus* individually.

The average number of alleles and absolute number of private alleles were calculated with a rarefaction method for each population, using the HP-Rare 1.1 software (Kalinowski, 2005).

Population genetic structure was investigated using STRUCTURE (Pritchard *et al.,* 2000). STRUCTURE is a software package for using multi-locus genotype data. Its functionality consists of the investigation of the presence of different populations, allocating individuals to populations, studying hybrid zones and estimating population allele frequencies in circumstances where individuals are admixed (Pritchard *et al.,* 2000; Rosenberg, 2004; Earl & VonHoldt, 2012). The tested K values ranging from 1 to 7, based on a previous study done by García-Rodríguez *et al.* (2018), and each K was run 10 times. The number of burn-in steps was set to 10,000 and Markov Chain Monte-Carlo (MCMC) was set to 100,000. ΔK was estimated using the Evanno method (Evanno, *et al.* 2005) implemented in STRUCTURE HARVESTER (Earl and vonHoldt, 2012) to obtain the most likely value of K. All structure results were joined together among replicates using CLUMPP 1.1.2 (Jakobsson & Rosenberg, 2007) and summarized graphically using DISTRUCT1.1 (Rosenberg 2004).

Multivariate analysis type FCA (factorial correspondence analysis) was performed using the function AFC-3D in Genetix (Belkhir *et al.,* 2004).

Population structure of *M. cypriacus* only was also tested using K from 1 to 4. Ten STRUCTURE runs per K value were executed, with a length of burn-in steps to 10,000 and MCMC steps was set to 100,000.

ΔK was estimated using the Evanno method (Evanno, *et al*. 2005) implemented in STRUCTURE HARVESTER (Earl and vonHoldt, 2012) to obtain the most likely value of K. All structure results were joined together among replicates using CLUMPP 1.1.2 (Jakobsson & Rosenberg, 2007) and summarized graphically using DISTRUCT1.1 (Rosenberg 2004).

Population genetic structure was then further exanimated using Discriminant Analyses of Principal Component (DAPC) which was performed using the R-package "*adegenet*" (Jombart *et al.,* 2010). The function *find.clusters* was used to identify the optimal number of clusters (K) that maximises the variation between groups (Jombart *et al*., 2010). The BIC (Bayesian Information Criterion) scores were analysed to determine the optimal number of clusters. Principal components (PC) were used as predictors for existing clusters, for discriminant analysis in the individuals studied.

3.7 Microsatellite Analysis

15 microsatellites loci from 36 samples of *M. m. domesticus* (García-Rodríguez *et al.* 2018) and 13 samples of *M. cypriacus* were used to compare with RAD sequencing results ,

Structure analysis was performed for the 15 microsatellites using STRUCTURE (Pritchard *et al.,* 2000). K values were tested from 1 to 7, with 10 replicates for each of several values of K, with a length of the burn-in steps of 10,000 and MCMC steps was set to 100,000.

Structure analysis was also conducted looking only at *M. cypriacus* (using 16 microsatellites specific to *M. cypriacus* only), testing K values from 1 to 4, and executed 10 STRUCTURE runs per K value, with a burn-in steps of 10,000 and MCMC steps was set to 100,000.

To identify the more probable K, ΔK was estimated using the Evanno method (Evanno, *et al*. 2005) from STRUCTURE HARVESTER (Earl and vonHoldt, 2012). All structure results were joined together among replicates using CLUMPP 1.1.2 (Jakobsson & Rosenberg, 2007) and summarised graphically using DISTRUCT1.1 (Rosenberg 2004).

Also, in this case, Discriminant Analyses of Principal Component (DAPC) was performed using the R-package "*adegenet*" (Jombart *et al.,* 2010), to identify the optimal number of clusters (K) that maximises the variation between groups.

The heterozygosity and the mean number of alleles per locus were calculated using GENETIX 4.03 (Belkhir *et al.,* 2004). Populations pairwise F_{ST} and linkage disequilibrium was calculated with Arlequin software v. 3.5.2.2 (Excoffier and Lischer, 2010), for the two species and for *M. cypriacus*.

3.8 Microsatellite Analysis - Population Demography

The average number of alleles and absolute number of private alleles were calculated with a rarefaction method for each population, using the HP-Rare 1.1 software (Kalinowski, 2005). Furthermore, to detect genetic signature of recent bottlenecks in *M. cypriacus*, the software BOTTLENECK v.1.2.02 was used (Piry *et al*., 1995). Two models were used during this analysis: the stepwise mutation model (SMM) and the two-phase model (TPM), with 95% of the mutation single-step and variance of 12 (Piry *et al.*, 1995).

4. Results

4.1 Phylogenetic Inference & Population Demography

A final alignment of 41 haplogroups (809 bp) of the 54 sequences [3 *Mus musculus domesticus* from Cyprus (García-Rodríguez *et al.,* 2018), 13 *Mus cypriacus*, 2 *Mus spretus* (GenBank: MK089345, MK089344), 2 *Mus musculus castaneus* (GenBank: AB649628, AB649629), 2 *Mus musculus musculus* (GenBank: KR866365, KR866364), 2 *Mus macedonicus* (GenBank: AF506193, AF506192), and 27 sequences retrieved from Genbank of *Mus cypriacus* (EU106194- EU106281). One sequences of *Rattus rattus* (Genebank: HQ334447) and two sequences of *Rattus norvegicus* (Genebank: X04733, X04734) were used as outgroup] was used to calculate a bayesian tree (Figure 3).

Our samples did cluster with previous published *M. cypriacus* sequences from Macholan *et al* (2007). As expected, *M. cypriacus* was found to be closely related to *M. macedonicus*. (Figure 7). Interestingly, no phylogeographic pattern was found. In fact, the pattern revealed by the Bayesian tree (Figure 7) appears random, as the haplotypes from different regions of Cyprus (Figure 8) are mixed.

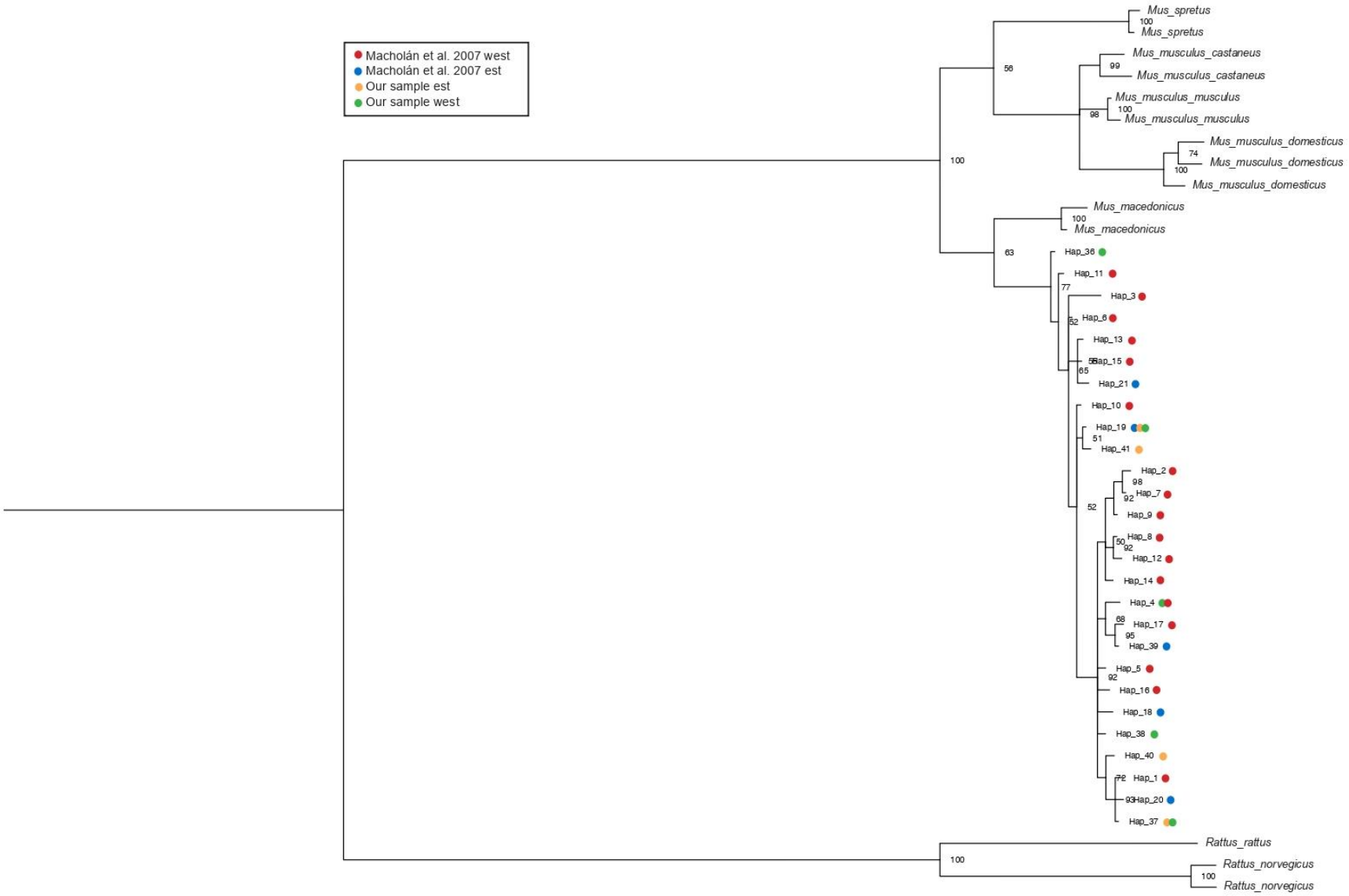


Figure 7 – **The Bayesian tree** -the D-loop of *Mus musculus domesticus* from Cyprus, *Mus cypriacus*, *Mus spretus*, *Mus musculus castaneus*, *Mus musculus musculus*, *Mus macedonicus*, were used and sequences of *Rattus rattus* and *Rattus norvegicus* were used as outgroups

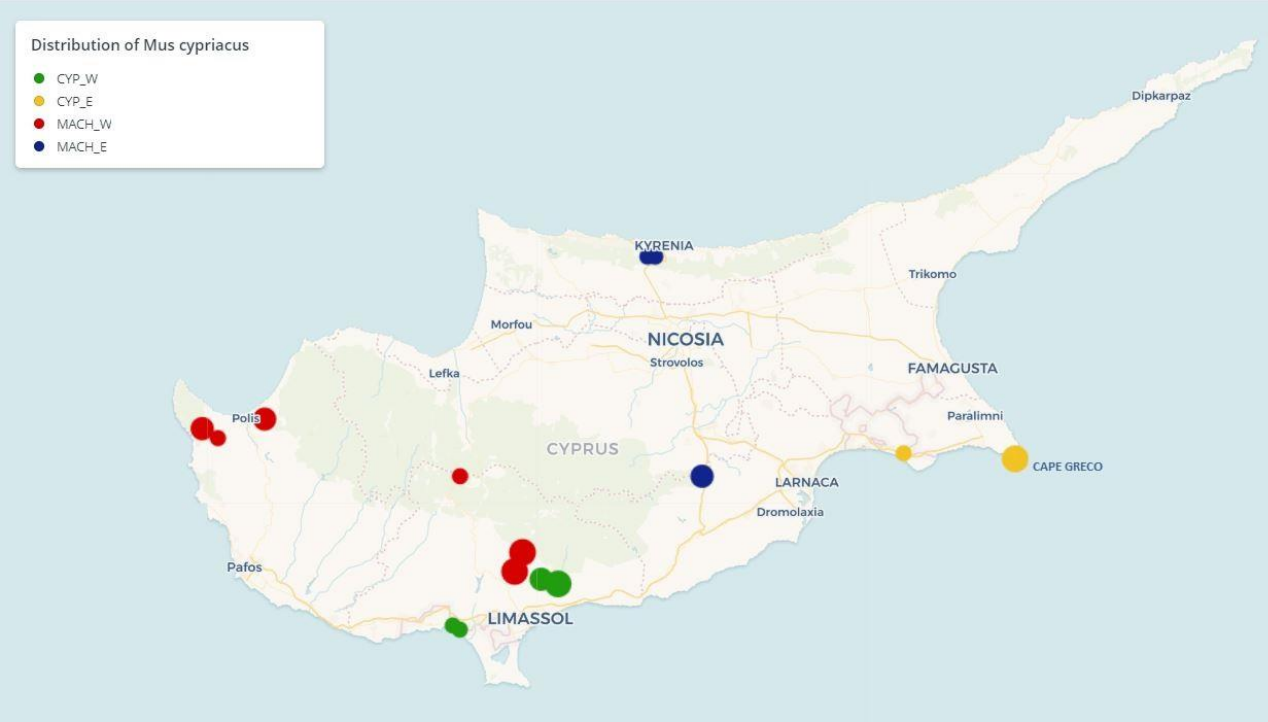


Figure 8 – **Distribution of *Mus cypriacus* in Cyprus –**The green icons correspond to the samples caught in the West part of the island (this study), the red icons are the samples caught in the West part of the island from Macholán *et al.* (2007), the yellow are the samples caught from the East part of the island (this study) and the blue are the samples caught in the East part of the island from Macholán *et al.* (2007)

A phylogenetic network was drawn using *M. cypriacus* samples from the present study and from Macholán *et al.* (2007) – see Figure 9). A total of 40 samples representing 29 haplotypes were used (Figure 9 and 10) and no phylogeography signal was found.

A second network was also calculated using only samples from this study (Figure 11). Out of 13 samples, 9 haplotypes were found. No phylogeography signal was found which is unexpected as samples from Cape Greco (west) – see Figure 12) and Limassol (east of the island) were used.

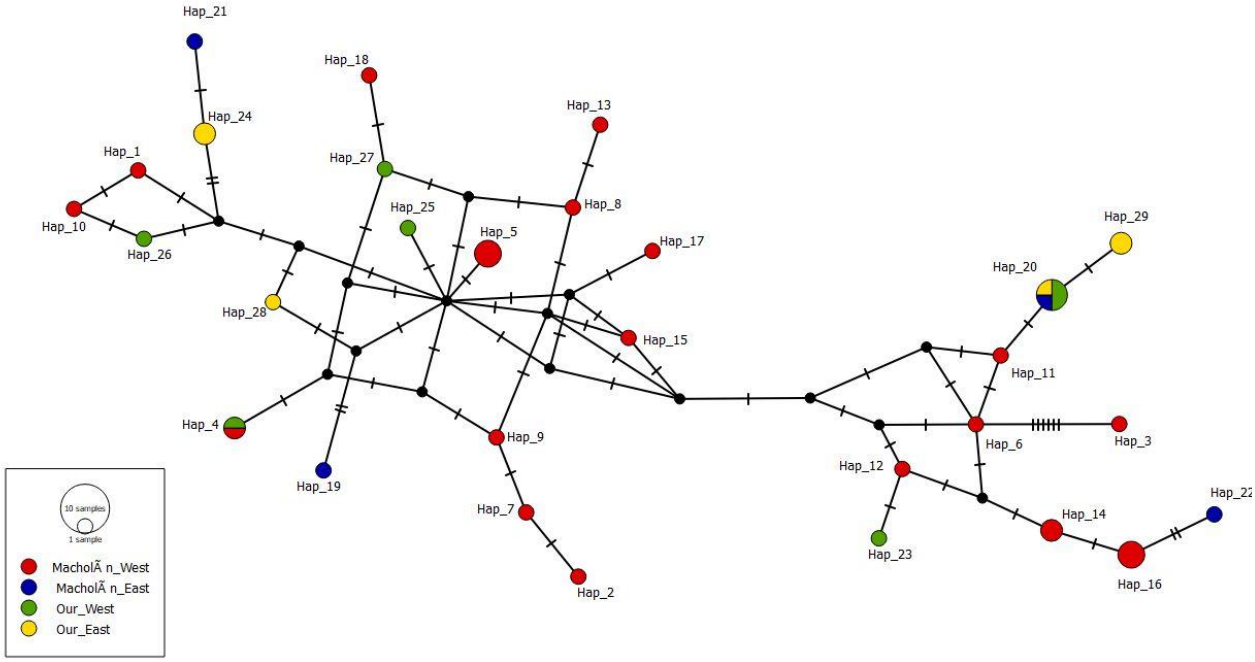


Figure 9 – **Phylogenetic network of *Mus cypriacus* samples from this study and Macholán et al. (2007).** The size of the circles corresponds to the number of samples.
The green circles correspond to the samples caught in the West part of the island for this research, the red circles are the samples caught in the West part of the island from Macholán *et al*. (2007), the yellow are the samples caught from the East part of the island for this research and the blue are the samples caught in the East part of the island from Macholán *et al*. (2007) (see Figure 10)

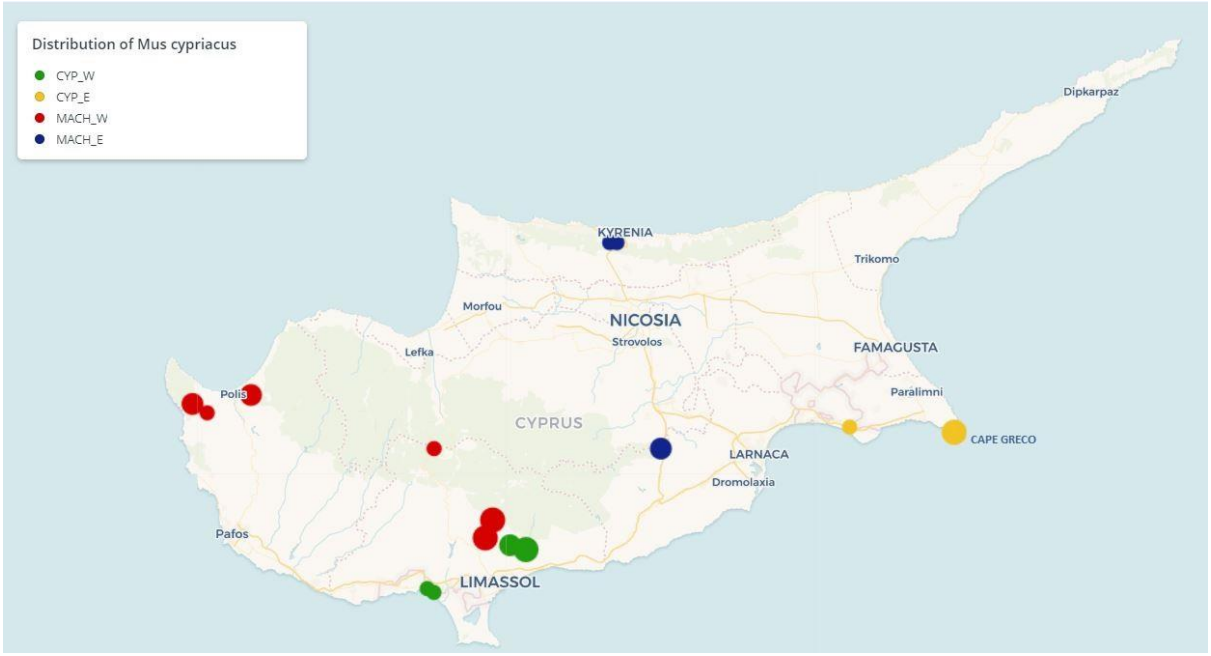


Figure 10 – **Distribution of *Mus cypriacus* in Cyprus –**The green icons correspond to the samples caught in the West part of the island (this study), the red icons are the samples caught in the West part of the island from Macholán *et al.* (2007), the yellow are the samples caught from the East part of the island (this study) and the blue are the samples caught in the East part of the island from Macholán *et al.* (2007)

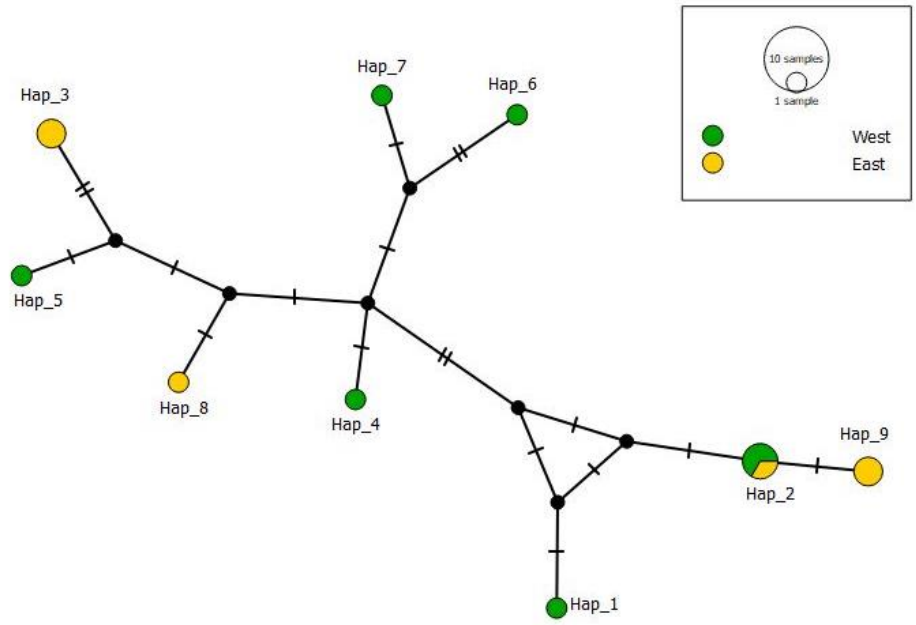


Figure 11 – Phylogenetic network of *Mus cypriacus* samples from this study.
The size of the circles corresponds to the number of samples.
The green circles correspond to the samples caught in the West part of the island, while the yellow are the samples caught from the East part of the island (see Figure 10)

Figure 12 – **Distribution of *Mus cypriacus* in Cyprus from this study only –**The green icons correspond to the samples caught in the West part of the island. The yellow are the samples caught from the East part of the island

The overall haplotype diversity (h) and nucleotide (π) diversity was found to be 0.979 (h) and 0.007 (π). Tajima's D and Fu's Fs neutrality tests were negative (D= - 0.768;P > 0.10 and Fs= - 18.966; P > 0.10) indicating absence of selection. The mismatch distribution of the D-loop sequences showed an unimodal distribution (Figure 13), a bell-shaped distribution of substitution differences between pairs of haplotypes, which is an indication of population expension (Rogers and Hamperding, 1992) or through an expantion with high levels of migration (Excoffier, 2004). The confidence of intervals around all the three variable was esistemated using Arlequin software v. 3.5.2.2 (Excoffier and Lischer *et al.,* 2010) using a parametric bootstrap with 100 or 1000 replicates. Approximate times of population expansion τ (in 1/2 u units, where u is the mutation rate for the whole sequence) was 7.196 and population sizes before the expansion (θ_0) and at present (θ_1) were found to be 2.184 (θ_0) and 89.128 (θ_1).

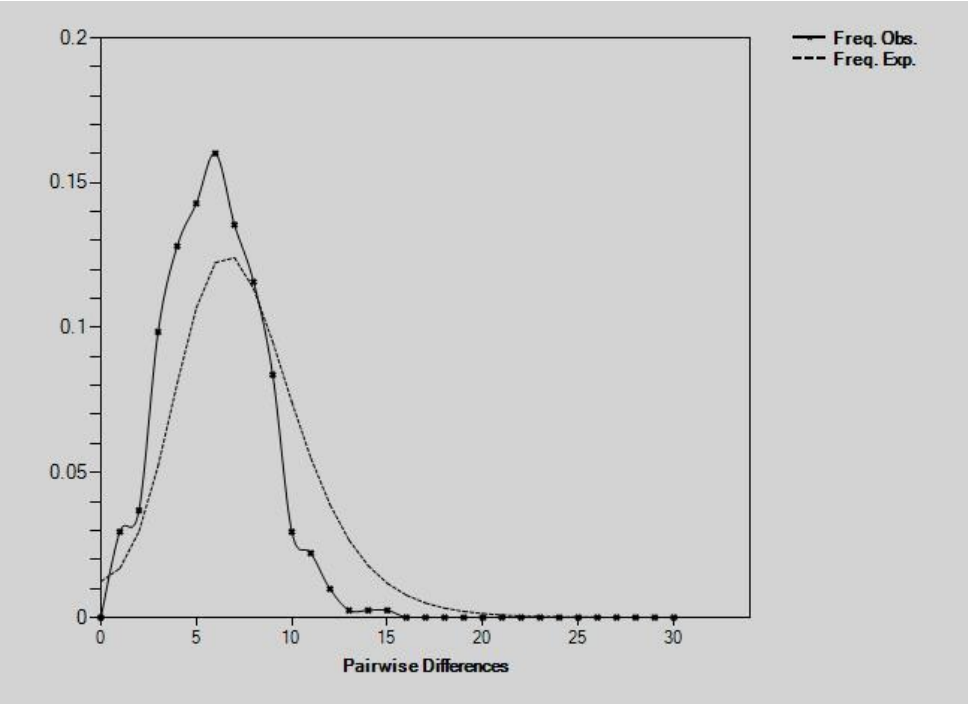


Figure 13 – **Mismatch distribution**- for the 40 samples of *Mus cypriacus*.
The expected distribution under a model of population expansion is given as a continuous line, and the observed distribution is given as a dashed line based on the population expansion function with parameters estimated using a generalized nonlinear least-squares approach.

4.2 SNP discovery and filtration

During the initial stage of the analysis (the *denovo_map* pipeline), the programme revealed two major warning: 1. Difference sequence lengths detected, this will interfere with Stacks algorithms, and 2. Input reads contained 140 uncalled nucleotides. Because of the first error mention above, *gstacks* aborted. Therefore, samples were aligned against a reference genome of the *Mus musculus domesticus* (GenBank: KV417259), using a standard alignment program known as Burrows-Wheeler Aligner, BWA. It generated a total of 33013948 BAM (Binary Alignment/ Map) records. When *ref_map* analysis is run, the *gstacks* is the first program execute. It kept 8811676 primary alignments (27.6%), of which 3807298 reverse reads, it skipped 2207709 primary alignments with insufficient mapping qualities (6.9%), it skipped 16493486 excessively soft-clipped primary alignments (51.6%), it skipped 4430525 unmapped reads (13.9%), and it also skipped some suboptimal (secondary/supplementary) alignment records. Overall, per sample, read 600253.6 records/sample (180374-2854755), it kept 10.0%-43.0% of these. The programme built 992687 loci, comprising 5004378 forward reads and 2535276 matching paired-end reads; the mean insert length was 238.6 (sd: 102.3). It removed 2469102 unpaired (forward) reads (49.3%); and kept 2535276 read pairs in 687276 loci. It then removed 152205 read pairs whose insert length had already been seen in the same sample as putative PCR duplicates (6.0%); and kept 2383071 read pairs. A total of 687276 genotyped loci were left at the end of *gstacks* analysis; with an effective per-sample coverage of mean=1.0x, stdev=0.0x, min=1.0x, max=1.1x, and a mean number of sites per locus: 244.3. After, the *population* programme was run, it removed 638204 loci that did not pass sample/population constraints from 687276 loci. It kept 49072 loci, composed of 9481355 sites; 35545 of those sites were filtered, and a total of 5325 variant sites remained. Overall, 9454366 genomic sites, of which 20951 were covered by multiple loci (0.2%). The mean genotyped sites per locus: 192.07bp (stderr 0.17).

Population summary statistics were as follow:

1. A “13.908” samples per locus for *M. m. domesticus*; pi: 0.25512; all/variant/polymorphic sites: 4811577/4833/3499; private alleles: 853
2. A “5.0605” samples per locus for *M. cypriacus*; pi: 0.11427; all/variant/polymorphic sites: 8077699/2775/736; private alleles: 148

Population pair divergence statistics were between 1-2: mean Fst: 0.69117; mean Phi_st: 0.76146; mean Fst': 0.75418

A final number of 5323 SNPs were then filtered with VCFtools allowing 25% of missing data. 46 common loci were found in *M. m. domesticus* and *M. cypriacus*. The same procedure was then done also within species and 71 loci out of 5323 SNPs were kept for *M. cypriacus*, with 25% of missing values.

4.3 Population genetics of M. cypriacus and M. m. domesticus - RADseq

A total of 46 SNPs were analysed using STRUCTURE (Pritchard *et al.,* 2000) for the 54 samples. The 46 loci were tested for linkage disequilibrium and none was detected (Figure 1.a supplementary material). Overall, for 41 *M. m. domesticus* the mean of expected and observed heterozygosity for 46 loci were 0.52 and 1.00 respectively (Table 1.a

supplementary material), and the average number of alleles per locus was 2.844; while, for 13 *M. cypriacus* the mean of expected and observed heterozygosity for 46 loci were respectively 0.533 and 1.000 (Table 1.a supplementary material), and the average number of alleles per locus was 2.522.

Using the Evanno method (Evanno, *et al*. 2005) – Figure 2.a supplementary material) and investigating the convergence of the run using CLUMPP (Jakobsson & Rosenberg, 2007), K=2 was found to be the best model. The 54 individuals were correctly assigned to their species, identifying 13 samples of *M. cypriacus* and 41 *M. m. domesticus* (Figure 14). No admixture between species was found.

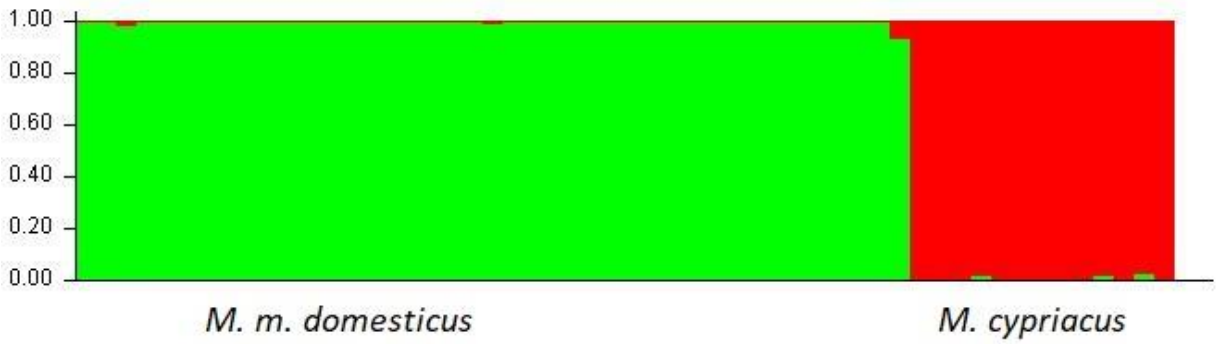


Figure 14 – **STRUCTURE analysis of RADseq** - STRUCTURE analysis of K=2 for a total of 46 SNPs in common for the two species. The red bars are for 13 samples of *M. cypriacus* and the green bars represent 41 samples of *M. m. domesticus.*

The Correspondence Analysis (CA) was performed using Genetix (Belkhir *et al*., 2004; (Figure 15). Axis 1 described 20.15% of the variation and is explained by the species membership (*M. m. domesticus* in yellow in Figure 15 and *M. cypriacus* in blue in Figure 15).

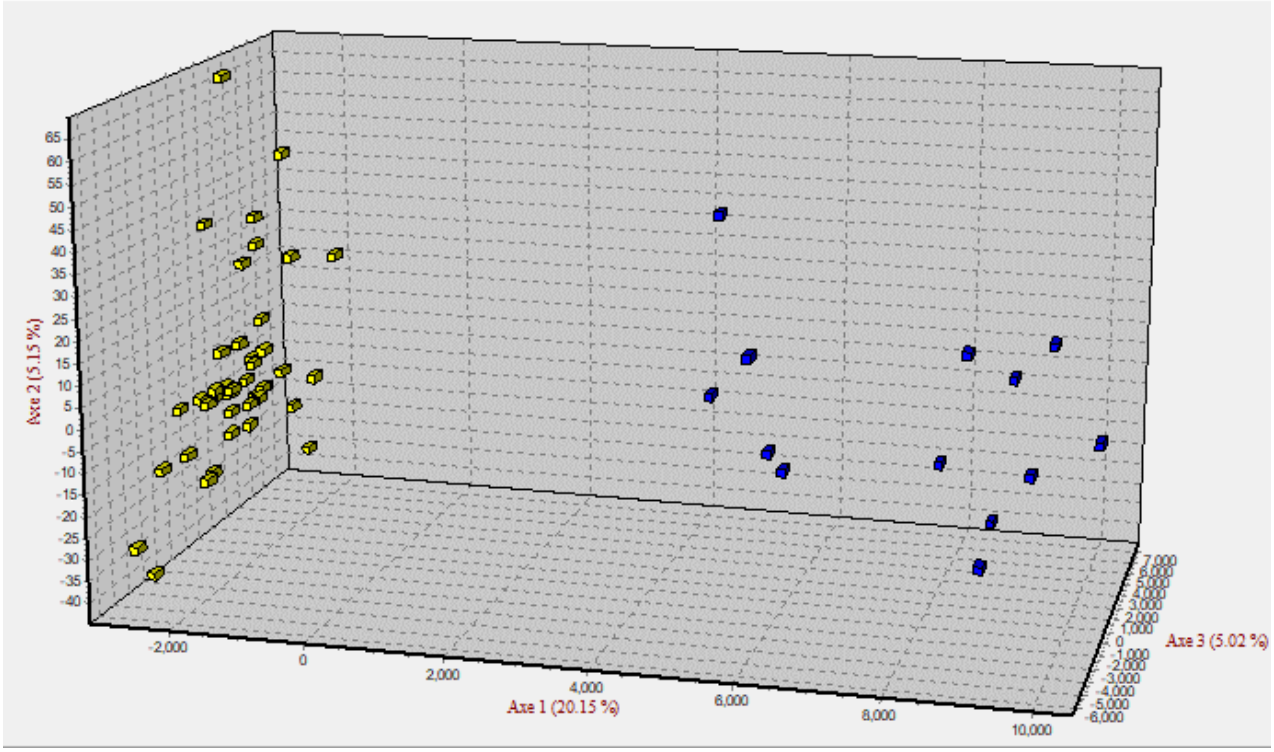


Figure 15 – **Correspondence Analysis RADseq -** Spatial representation of the tridimensional factorial correspondence analysis carried out with GENETIX (Belkhir *et al.,* 2004), every square representing an individual. The blue squares represent the 13 *Mus cypriacus*, while the yellow squares are the *Mus musculus domesticus*

4.4 Population genetics of M. cypriacus and M. m domesticus – Microsatellites.

A total of 15 microsatellites were analysed using STRUCTURE (Pritchard *et al.*, 2000) for the 49 samples. The 15 loci were tested for linkage disequilibrium (LD; Figure 3a supplementary material) and no locus was found to be linked, *M. m. domesticus* expected and observed heterozygosity was calculated to be 0.826 and 0.756 respectively, and the average number of alleles per locus was 11.733. *M. cypriacus* the mean of expected and observed heterozygosity for 15 loci were respectively 0.828 and 0.705, and the average number of alleles per locus was 10.266.

Using the Evanno method (Evanno, *et al*. 2005) – Figure 4.a supplementary material) and investigating the convergence of the run using CLUMPP (Jakobsson & Rosenberg, 2007), K=2 was selected as the best model. The 49 individuals were correctly assigned to their species, identifying 13 samples of *M. cypriacus* and 36 *M. m. domesticus*. No admixture between species was found (Figure 16).

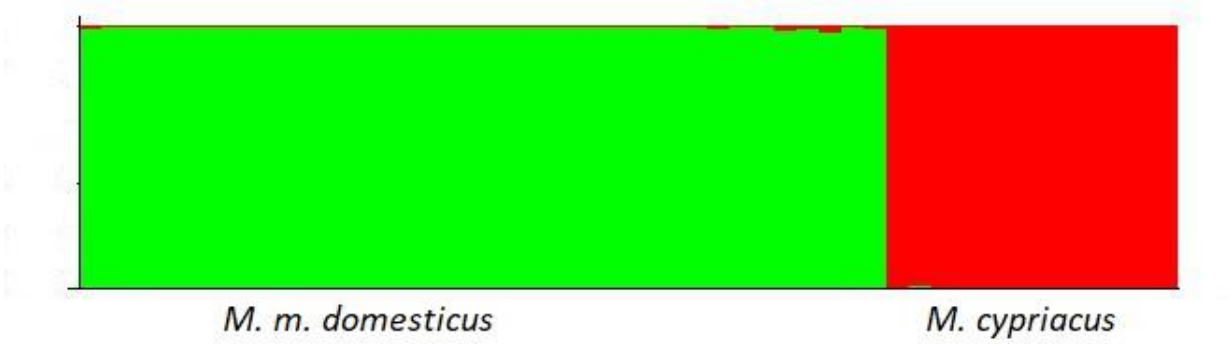


Figure 16 – **STRUCTURE analysis of msat** -The STRUCTURE analysis is K=2 for a total of 15 microsatellites in common for the two species. It is represented by 2 different colours, where the green bars are for the 36 *M. m. domesticus* samples and the other are for the 13 *M. cypriacus* samples.

The Correspondence Analysis (CA) illustrates the position of individual genotypes projected onto a 3D space (Figure 17). It is possible to observe the two species, *M. m. domesticus* and *M. cypriacus*, well separated; with the first axis explaining the variation at 6,99%, 4,84% of the variation is explain by axis 2 and 4,54 % of the variation is explain by axis 3. Overall, no admixture between the two species was found.

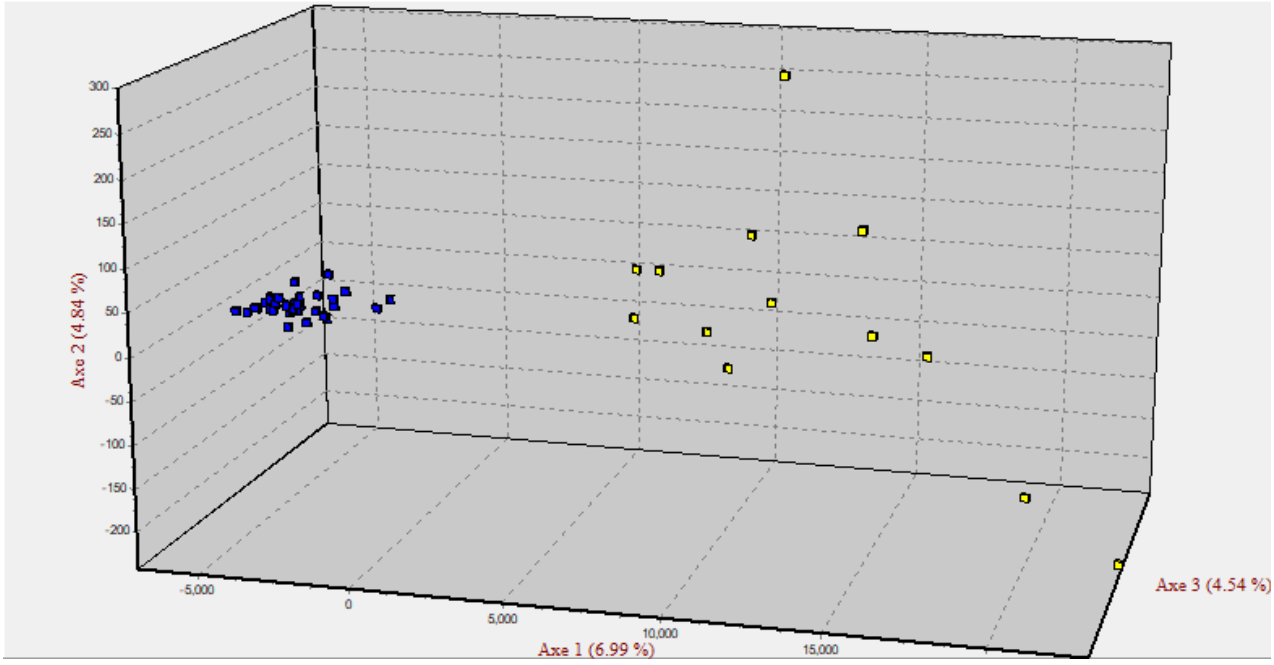


Figure 17 – **Correspondence Analysis msat -** Spatial representation of the tridimensional factorial correspondence analysis carried out with GENETIX every square representing an individual. The yellow squares represent the 13 *M. cypriacus*, while the blue squares are the 36 *M. m. domesticus*

4.5 Population Structure – M. cypriacus

Population analysis was performed for *Mus cypriacus* only using 71 SNPs obtained from the RADseq and 16 microsatellites.

Heterozygosity as well as mean numbers of alleles (Table 2) were calculated among *M. cypriacus*, respectively for SNPs and microsatellites. The samples were divided into two populations Cape Greco (East) and Limassol (West), for both markers (Figure 18).

Markers	**N**	**H_{exp}**	**H_{obs}**	**N. of loci**	**Average n. of alleles per locus**
SNPs- Limassol	7	0.542	1.000	71	2.493
SNPs- Cape Greco	6	0.532	1.000	71	2.366
Msat- Limassol	7	0.803	0.706	16	7.250
Msat- Cape Greco	6	0.754	0.741	16	6.375

Table 2 – Table reporting the two types of markers (Markers), number of samples (N), expected and observed heterozygosity based on the type of markers (H_{exp} and H_{obs} respectively), number of loci per makers (N. of loci),and the average number of alleles per locus (Average n. of alleles per locus).

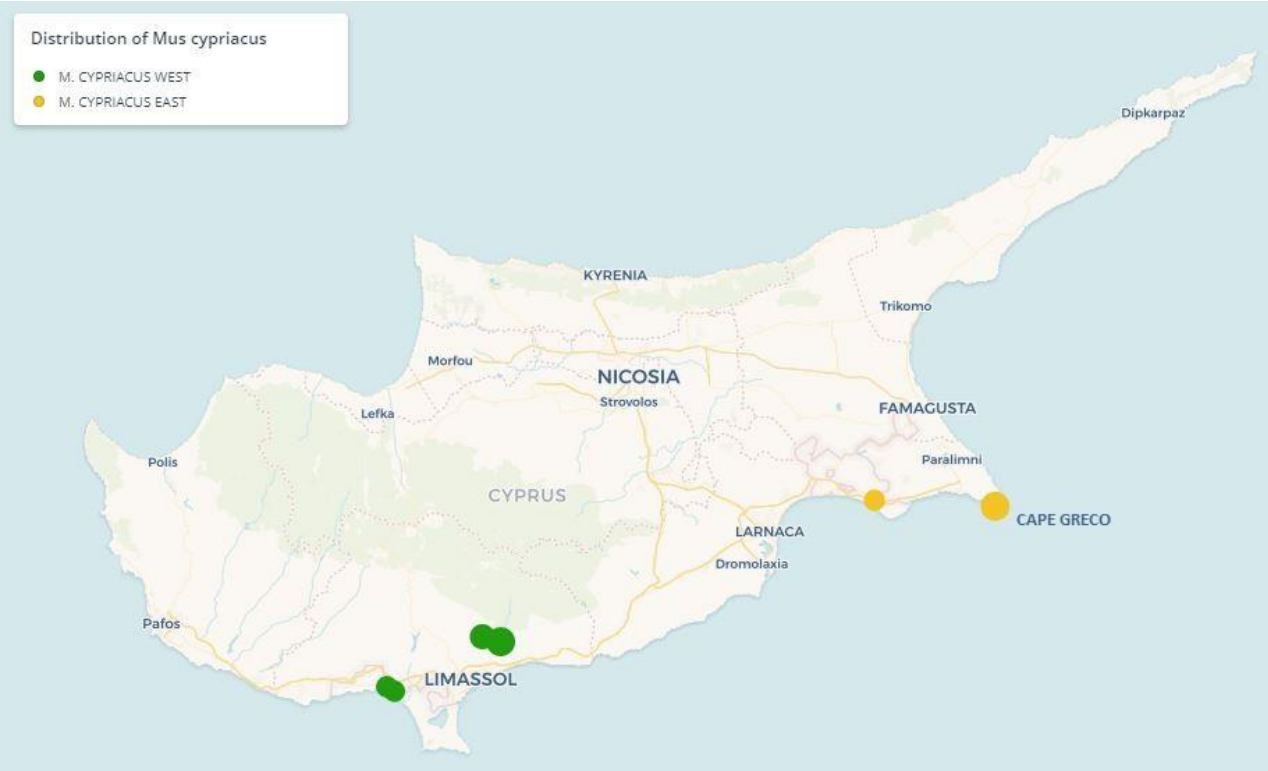


Figure 18 – **Distribution of *Mus cypriacus* in Cyprus from this study only –**The green icons correspond to the samples caught in the West part of the island. The yellow are the samples caught from the East part of the island

Population genetic structure of 13 *M. cypriacus* samples was investigated using STRUCTURE (Pritchard *et al.,* 2000) – Figure 5.a and Figure 6.a supplementary material) and DAPC analysis (Jombart *et al.,* 2010) in order to identify a possible population structure. No population subdivisions were found when using both analyses for both markers.

Populations pairwise F_{ST} between Limassol (West) and Cape Greco (East) is low (F_{ST} = 0.022) for the SNPs; while, F_{ST} within *M. cypriacus* samples from West and East for the microsatellites was 0.032 and the inbreeding coefficient for all the samples was 0.157 (Table 2.a supplementary material).

The population demography was investigated further using BOTTLENECK (Piry *et al.* 1999). Results can be found in Figure 19. No recent genetic population bottleneck has been detected related to heterozygote excess; most probably due to a population expansion or introduction of rare alleles (Luikart and Cornuet, 1998).

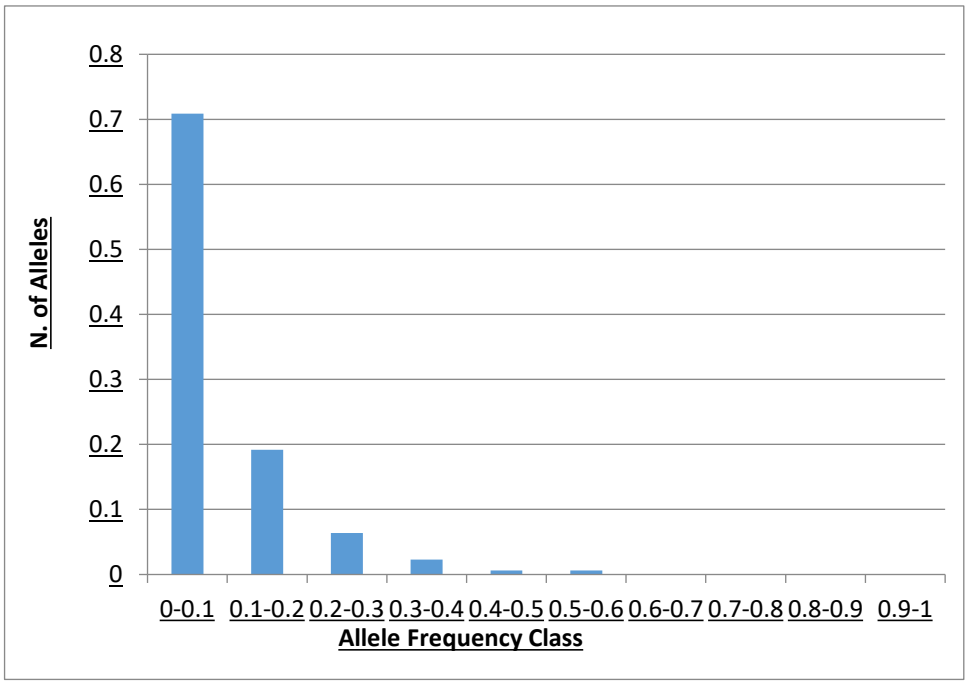


Figure 19 – **Bottleneck** - Distribution of allele frequencies expected for loci for *Mus cypriacus*. Blue bars represent the proportion of alleles expected in each of 10 allele frequency classes. The mean heterozygosity expected for random sample of loci having the illustrated distribution is 0.80

5. Discussion

This study has investigated the genetic population structure of the Cypriot mouse *Mus cypriacus* and potential hybridisation with the house mouse *Mus musculus domesticus* using 3 types of genetic markers. No population genetic structure was evident in this study for *M. cypriacus* even though three different genetic makers (Mitochondrial D-loop, SNP and microsatellites) were used. There was no evidence of a recent genetic bottleneck for the species. No genetic signature of hybridisation between *M. m. domesticus* and *M. cypriacus* was found.

5.1 Genetic diversity

Interestingly, high genetic diversity (when compared to *M. m. domesticus*) was detected for *M. cypriacus* with all the markers studied. This result was surprising as island populations often have a lower genetic diversity compared to their mainland counterparts mostly due to founder effects (Jones *et al.,* 2004; Miller *et al.,* 2011; Hardouin *et al.,* 2010; Hardouin *et al.*, 2018). Indeed, the level of genetic diversity of *M. cypriacus* was comparable to the one found in *M. m. domesticus* (e.g. average number of alleles per loci found to be 2.8 for *M. m. domesticus* and 2.5 for *M. cypriacus*). This result is unexpected as it has been shown the colonization pattern of house mouse on Cyprus is complex and the result of several human introductions (Garcia- Rodriguez *et al.* 2018). The level of genetic diversity of *M. cypriacus* were found high with both SNPs data and microsatellites, H_O = 1.000 and n_a = 2.522 for the RADseq and, H_O = 0.828 and n_a = 10.266 for the microsatellite loci. Other rodent species have shown high or similar level of genetic and allelic diversity on islands when compared to the mainland populations. For example, the Coues' rice rat (*Oryzomys couesi cozumelae*) from Cozumel Island, Mexico was found to have a genetic diversity similar to the mainland species (*O. couesi*) (Vega *et al.*, 2007). Same was observed for the insular *Oryzomys argentatus* from the Florida Keys (USA) and mainland *O. palustris natator* from the Everglades (USA) - Wang *et al.,* 2005). According Frankham, (1997) the levels of genetic and allelic diversity has been associated to island sizes, as well as other factors. In fact, larger islands can normally host and support higher population size; and, on the other side, events like inbreeding, genetic bottlenecks and higher extinction rates are normally associated to smaller islands with smaller population size. Endemic island species normally have low levels of genetic diversity (Frankham 1997), which compromises their adaptability and evolutionary potential, making island species more vulnerable to extinction (Frankham 1998). Cyprus is the third largest island of oceanic origin located in the Eastern part of the Mediterranean basin, with an area of 9,251 km^2 (Kryštufek and Vohralík, 2001). It was connected to mainland only during the Messinian salinity crisis, and therefore, it has been isolated for more than 5.3 million years (Kryštufek and Vohralík, 2001). *M. cypriacus* diverged c. 430 000–610 000 years ago (coalescent ≈ 490 000 years ago) from *Mus macedonicus* and 830 000–1.2 million years ago (coalescent ≈ 780 000 years ago) from *Mus spicilegus*. Hence, this split dates earlier than the beginning of the first glacial period. The hypothesis that a southward and westward expansion of this ancestor which then colonized Cyprus, accidentally crossed a deep marine strait that separated the island even during the minimum sea levels. A subsequent divergence between the island and mainland

populations giving rise to *M. cypriacus* and *M. macedonicus* 400 000–600 000 years ago. *Mus cypriacus* fossil remains, dating back to the Pleistocene (Reese, 1999; Kryštufek and Vohralík, 2001). Those factors could encourage adequately high population size of *M. cypriacus* as to avoid substantial levels of inbreeding or genetic drift; however, they might allow high level of heterozygosity and the presence of rare alleles.

5.2 Population structure and possible hybridisation

No population structure in *M. cypriacus* was found on the island even though samples were collected 121 km apart (distance between Cape Greco and Limassol). The geographical distributions of small rodents are influenced by phylogenetic affinities of species, interactions and environmental factors (Vazquez *et al.*, 2000). Unfortunately, little information is available on the dispersal ability and transport mechanisms of *M. cypriacus*. According to Eble *et al*. (2009), the recognition of endemic species subpopulations can be counterintuitive and can also lead to an increase of range-wide panmixia. Furthermore, it has been proved that endemic species can manifest more genetic diversity within a limited geographical area compared to their mainland counterparts, even if they exhibit lower dispersal ability (Bohonak 1999; Siegel *et al.* 2003; Shanks *et al.* 2003). Our population structure analysis, in fact, indicates an absence of population structure. According to Macholán *et al*., (2007), the lack of geographical structure and the absence of connection between geographical and genetic distances could have been generated by sufficiently high gene flow among populations within the island. Most probably, *M. cypriacus* started an exponential population growth approximately 100000 years ago (Macholán et al., 2007). Indeed, according to García-Rodríguez *et al.* (2018), the house mouse in Cyprus also revealed little population structure on the island, potentially due to the high levels of transportation, and mice, within farms and agricultural settings on Cyprus. However, we did not find *M. cypriacus* in farms, therefore it is not possible to assume any conclusions in regard to that.

The mismatch distribution for *M. cypriacus* suggested a recent population expansion. Indeed, the coalescence analysis rejects the null hypothesis of a stable population, which is in agreement with the results obtained by Macholán *et al*., (2007). Both Fu's (1997), Fs tests and Tajima's (1989), D were not significant across all *M. cypriacus*. Generally, statistics based on haplotype frequency (e.g., Fs) are more powerful at detecting recent and moderate bottlenecks, whereas tests that rely on frequency spectrum of mutations (e.g., D) are best at detecting old and severe bottlenecks (Depaulis *et al.* 2003). Invasive species, predators, and competitors, as well as anthropogenic events, can affect the number of individuals within a population, leading in reductions of population size and genetic bottleneck (Frankham, 1997, Frankham, 1998). Cyprus is a hotspot area, and all the before mentioned factors are present in the island, being threats for the Cypriote mouse. However, no recent genetic population bottleneck has been detected related to heterozygote excess; most probably due to a population expansion (Luikart and Cornuet, 1998).

Nevertheless, invasive species usually pose a threat to native island endemics (Mellink *et al.,* 2002, Vázquez-Domínguez *et al.,* 2004). On Cyprus, *M. m. domesticus* and the domestic cat *Felis silvestris* arrived c. 8,000 years B.C (Vigne *et al.*, 2004; Vigne *et al.,* 2012; García-Rodríguez *et al.,* 2018), as *Acomys nesiotes* (Barome *et al.,* 2001). The black rat *R. rattus* arrived in the island during the roman period (McCormick, 2003), and the Norway rat *Rattus*

norvegicus arrived later (Musser and Carleton, 2005). The genus *Rattus* are considered strong competitors of indigenous species (Harper and Cabrera, 2010), they are also known to kill mice (Karli, 1956; Bridgman *et al.,* 2013) and other animals, such as seabirds (Stapp, 2002). Furthermore, Frynta *et al*., (2006) showed that *M. cypriacus* tend to avoid the smell of domestic cats, but it does not recognize the smell of the *R. norvegicus* as competitor. Those introduced species might influence *M. cypriacus* populations, however, it seems this has not been shown on its genetic diversity.

Introgressive hybridisation between wild and domestic mouse species have been described in Northern Africa (Song *et al*. 2011) however on Cyprus no sign of introgression between *M. m. domesticus* and *M. cypriacus* was found even though they were found to share the same habitat in one of our sampling sites in Xylophagou (Figure 2). The *M. cypriacus* specimen as well as the other seven *M. m. domesticus* species were found in an abandoned quarry close to houses and fields and thus contact between the two species is possible, at least for this site.

5.3 Units for conservation

The results collected in the present study suggested that *M. cypriacus* population is panmictic on Cyprus. This result is unexpected *as M. cypriacus* is mostly found in cultivation terraces with vineyard, grassy fields, and bushes (Cucchi *et al.,* 2006) and absent in areas with intense anthropogenic pressure, such as farms or humans' abodes, where instead the *M. m. domesticus* has been found to be abundant (Cucchi 2005 and present study). During the last few decades, the number of tourists increased on Cyprus (Saveriades, 2000) with high development efforts in hotel building, luxury apartments, villas and golf clubs, abolishing or altering natural environment mostly near the sea (Hadjimitsis, 2010; Zachariadis, 2012; Welz, 2015). The results of our genetic analysis suggest that this increase in urbanisation has not affected the dispersion of *M. cypriacus*. The Cypriot mouse maintains high levels of genetic and allelic diversity. Furthermore, the results obtained by the population structure of the Cypriot mouse indicate that individuals are not completely isolated. However, those factors of environment alterations might intensify isolation of groups and potentially lead to extinction (Neuwald, 2010).

5.4 Limitations for the analysis of M. cypriacus

The main limitation of our study is the low sample size which was 13 individuals. In order to have a better understanding of the population and its potential conservation, more populations need to be sampled across the island. Due to this low sample size, the initial results indicating a genetically healthy population must be considered with caution especially in light of the high development experiences in Cyprus that has led to various conservation issues, such as habitat perturbation, urbanisation and introduction of exotic species. A large number of samples or a large number of loci are suggested for the calculation of genetic statistic, mainly when diploid markers, such as microsatellites are used (Toro *et al*. 2002, Kalinowski 2005). However, when studying threatened and endangered species, obtaining a large number of samples could result in a real challenge, because of their dispersal capacity or found in remote areas (Pruett and Winker, 2008). According to Smith and Wang

(2014), different population studies can still be obtained without biases in small samples (above 10 and 20 individuals in well-differentiated and poorly differentiated populations), such as measures of expected heterozygosity, differentiation and population structure. Furthermore, Pruett and Winker (2008) stated that when the nuclear genetic diversity of a species is not known, the samples size should include a minimum of 20 individuals, ideally 30.

Species distribution is crucial for monitoring threatened and endangered species (Gaston, 1996; Kumar and Stohlgren, 2009). However, the distribution data available for those vulnerable species are often insufficient, making it extremely difficult to analyses habitat modelling (Ferrier *et al.,* 2002; Engler *et al*., 2004). There are a variety of species distribution modelling methods, that allow scientist to predict species distribution (Guisan and Zimmermann, 2000; Guisan and Thuiller, 2005; Elith *et al.,* 2006; Wisz *et al.,* 2008). However, most of those methods are sensitive to the small sample size, and the outcome is an inaccurate prediction of habitat distribution of the species studied (Wisz *et al.,* 2008). The Cypriote mouse has only been described recently (Cucchi *et al*. 2006; Bonhomme *et al*., 2004), and only little is known on the population size, population trends and distribution of the species. *Mus cypriacus* was found mainly in the Troodos region between 300 and 900 meters a.s.l. It is mostly found in habitat comprises abandoned cultivation terraces with vineyard, grassy fields, and bushes such as Mastic trees, Terebinths Thorny Broom and Thorny Gorse (Cucchi *et al*. 2006). Further studies with a brother number of samples are essential to understand better the distribution and the dynamics of the Cypriote mouse.

Lastly but not least, another significant limitation for this study was related to low library quality. The genomic libraries constructed are individually labelled and pooled for sequencing on an Illumina MiSeq based on modified ddRAD protocols (Peterson *et al.* 2012). The library construction was based on an efficient combined restriction digest/adaptor ligation. Two restriction enzymes were used digest genomic DNA: Csp6I (which cleaves 5'-G^TAC -3'sites) and PstI (which cleaves 5'- CTGCA^G -3' sites). The reaction conditions permit that sticky end adapters and T4 ligase are added to the reaction such that adaptors are ligated to the restriction sites. Importantly, the adaptors do not reconstitute the restriction sites.

SNP markers generated must be carefully treated, in order to separate high-quality markers from the possibly biased (Davey *et al.,* 2012). There are several potential sources of error that could affect RADseq generated data such as PCR duplicates and allele dropout (Kimberly *et al.,* 2016). Furthermore, high levels of DNA degradation have been proved to decrease the potential SNPs data drastically and ultimately eliminate the usefulness of the ddRADseq approach (Graham *et al.,* 2015); "the potential SNPs available for population decreased on average by approximately 96.5%per individual in 96-h treatment." However, due to short lengths methods for analysing RADseq, sometimes require mapping sequencing reads to a whole sequence genome from the same or the closer species (Li *et al.,* 2008). For this analysis, a final number of 5323 SNPs were filtered with VCFtools, leaving the number of SNPs called at 75% of individuals, keeping 46 robust loci in common for *M. cypriacus* and *M. m. domesticus*. The same procedure was then done also within species, and 71 loci out of 5323 SNPs were kept for *M. cypriacus*, with 25% of missing values. There was a considerable reduction of potential SNPs after filtration, and a small number of loci

were left available for this analysis. However, the results obtained with the available microsatellites loci were the same compared to the results obtained from the SNPs.

6. Conclusion and future work

The Cypriot mouse is one the three surviving palaeoendemic mammal species found on Mediterranean islands and therefore of conservation interest (Gippoliti and Amori 2006). The species has only been described recently (Cucchi *et al*. 2006; Bonhomme *et al*., 2004) which might explain the lack of knowledge on *M. cypriacus*. This study is the first to date to investigate the population structure of the Cypriot mouse using nuclear markers. The data suggests that the species is comprised of a single and demographically stable panmictic population. Due to the small sample size, however, more research is needed to confirm these results. Furthermore, only little is known on the population size, population trends and the distribution of this species. Future work needs to estimate population sizes, provide a detailed distribution map and be complemented with mark-release-recapture work to better understand the dispersal of the species. This work will allow a determination of the factors that are contributing to the apparent genetic health of the population as well as identify any potential threats to it.

References

Aberer, A. J., Stamatakis, A., & Ronquist, F., 2016. “An efficient independence sampler for updating branches in Bayesian Markov chain Monte Carlo sampling of phylogenetic trees.” *Systematic biology*, 65(1), pp. 161-176.

Abbott R., Albach D., Ansell S., Arntzen J., Baird S., Bierne N., Boughman J., Brelsford A., Buerkle C., and Buggs R. 2013. “Hybridisation and speciation”. *Journal Of Evolutionary Biology*, 2, pp. 229-246.

Abdelkrim, J., Pascal, M., Samadi, S., 2005. “Island colonization and founder effects: The invasion of the Guadeloupe islands by ship rats (Rattus rattus)”. *Molecular Ecology*, 14, pp. 2923-2931.

Allendorf, F. W., Leary, R. F., Spruell, P., and Wenburg, J. K., 2001. “The problems with hybrids: setting conservation guidelines.” *Trends in ecology and evolution*, 16(11), pp. 613-622.

Amori, G. 2017. “*Mus cypriacus*." The IUCN Red List of Threatened Species 2017: pp. e.T136641A22406364.

Amori, G., Gippoliti, S. and Helgen, K.M., 2008. “Diversity, distribution, and conservation of endemic island rodents.” *Quaternary International*, vol. 182, no. 1, pp. 6–15.

Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G. and Hohenlohe, P.A., 2016. “Harnessing the power of RADseq for ecological and evolutionary genomics.” *Nature Reviews Genetics,* no. 2, pp. 81.

Andrews, K.R., Hohenlohe, P.A., Miller, M.R., Hand, B.K., Seeb, J.E. and Luikart, G., 2014. “Trade-offs and utility of alternative RADseq methods: Reply to Puritz et al. [Puritz JB, 2014]” *Molecular Ecology*, no. 24, pp. 5943.

Arnold, B., Corbett-Detig, R.B., Hartl, D. and Bomblies, K., 2013. “RADseq underestimates diversity and introduces genealogical biases due to non-random haplotype sampling.” *Molecular Ecology,* no. 11, pp. 3179.

Auffray, J.C., Vanlerberghe, F. and Britton-Davidian, J.A.N.I.C.E., 1990. “The house mouse progression in Eurasia: a palaeontological and archaeozoological approach.” *Biological Journal of the Linnean Society*, 41(1-3), pp.13-25.

Auffray, J. C., & Britton-Davidian, J. A. N. I. C. E., 2012. “*The house mouse and its relatives: systematics and taxonomy. Evolution of the house mouse*.” Cambridge studies in morphology and molecules: new paradigms in evolutionary biology. Cambridge University Press, Cambridge, pp. 1-34.

Baier, F., Sparrow, D. J., and Wiedl, H. J., 2009. “*The amphibians and reptiles of Cyprus*.” Frankfurt am Main: Edition Chimaira.

Baillie, J. E., Hilton-Taylor, C., and Stuart, S. N., 2004. “2004 IUCN red list of threatened species.” IUCN.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. and Johnson, E.A., 2008. “Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers.” *PLoS ONE,* vol. 3, no. 10, pp. 1–7.

Baker, A.J., 2000. “*Molecular Methods in Ecology*.” Blackwell Science ltd, London

Bate D.M.A., 1903. “On an extinct species of genet (*Genetta plesictoides*, sp.n.) from the Pleistocene of Cyprus.” *Proc. Zool. Soc. Lon*. 2: pp.121–124.

Beebee, T.J.C., 2018. “Genetic contributions to herpetofauna conservation in the British Isles.” *Herpetological Journal*, vol. 28, no. 2, pp. 51–62.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and Bonhomme, F., 2004. “GENETIX 4.05, logiciel sous Windows pour la génétique des populations“. Montpellier, France: Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II.

Bennett, E.L., Blencowe, E., Brandon, K., Brown, D., Burn, R.W., Cowlishaw, G., Davies, G., Dublin, H., Fa, J.E. and Milner-Gulland, E.J., 2007. “Hunting for Consensus: Reconciling Bushmeat Harvest, Conservation, and Development Policy in West and Central Africa”. *CONSERVATION BIOLOGY -BOSTON MASSACHUSETTS*-, no. 3, pp. 884.

Bhargava, A. and Fuentes, F.F., 2010. “Mutational Dynamics of Microsatellites.” *Molecular Biotechnology*, no. 3, pp. 250-66.

Birky, C. W., Fuerst, P., & Maruyama, T., 1989. “Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes.” *Genetics*, 121(3), pp. 613-627.

Bonhomme, F., 1986. “Evolutionary relationships in the genus *Mus*.” *In the wild mouse in immunology.* Springer, Berlin, Heidelberg, pp. 19-34.

Bonhomme, F., Orth, A., Cucchi, T., Hadjisterkotis, E., Vigne, J.-D., Auffray, J.-C., 2004. “De´couverte d”une nouvelle espe`ce de souris sur l”ıˆle de Chypre.” *C. R. Biol.* 327, pp. 501–507.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., and Drummond, A.J., 2014. “BEAST 2: a software platform for Bayesian evolutionary analysis.” *PLoS computational biology*, vol. 10, no. 4, p. e1003537.

Borrel Y.J., Pinera J.P., Sanchez P., Blanco G., 2012. “Mitochondrial DNA and microsatellite genetic differentiation in the European anchovy *Engraulis encrasicolus* L.” *ICES. J. Mar. Sci. 69*: pp. 1357-1371.

Bohonak, A.J., 1999. “Dispersal, gene flow and population structure.” *Q Rev Biol* 72, pp.21–45.

Bouchemousse S., Liautard-Haag C., Bierne N., and Viard, F., 2016. “Distinguishing contemporary hybridisation from past introgression with postgenomic ancestry-informative SNPs in strongly differentiated Ciona species”. *Molecular Ecology*, 25, 21, pp. 5527-5542.

Boursot, P., Auffray, J. C., Britton-Davidian, J., and Bonhomme, F., 1993. “The evolution of house mice.” *Annual review of ecology and systematics*, 24(1), pp. 119-152.

Brennan A., Woodward G., Seehausen O., Muñoz-Fuentes V., Moritz C., Guelmami A., Abbott R.J. and Edelaar P., 2014. “Hybridisation due to changing species distributions: adding problems or solutions to conservation of biodiversity during global change?” *Evolutionary ecology research*, no. 6, pp. 475.

Bridgman LJ, Innes J, Gillies C, Fitzgerald NB, Miller S., and King, C.M., 2013. “Do ship rats display predatory behaviour towards house mice?” *Animal Behaviour* 86, pp. 257–268.

Bromham, L., 2009. “Why do species vary in their rate of molecular evolution?” *Biology letters*, 5(3), pp. 401-404.

Bromham, L., and Penny, D., 2003. “The modern molecular clock.” *Nature Reviews Genetics*, 4(3), 216-224.

Brooks, T. M., Mittermeier, R. A., Fonseca, G. A. B., Gerlach, J., Hoffmann, M., Lamoreux, J. F., Mittermeier, C. G., Pilgrim, J. D. and Rodrigues A. S. L., 2006. “Global Biodiversity Conservation Priorities.” *Science*, vol. 313, no. 5783, pp. 58.

Brown, T. A., 2002. “*Molecular phylogenetics*”. In Genomes. 2nd edition. Wiley-Liss.

Burlakova, L. E., Karatayev, A. Y., Karatayev, V. A., May, M. E., Bennett, D. L., and Cook, M. J., 2011. “Endemic species: contribution to community uniqueness, effect of habitat alteration, and conservation priorities.” *Biological Conservation*, 144(1), pp. 155-165.

Callen DF, Thompson AD, Shen Y, Phillips HA, Richards RI, Mulley JC, Sutherland GR., 1993. “Incidence and origin of null alleles in the (AC)n microsatellite markers.” *American Journal of Human Genetics* 52, pp. 922–927.

Carver, T., Böhme, U., Otto, T.D., Parkhill, J. and Berriman, M., 2010. “BamView: viewing mapped read alignment data in the context of the reference sequence.” *Bioinformatics* (Oxford, England), vol. 26, no. 5, pp. 676–677.

Catchen, J., Bassham, S., Wilson, T., Currey, M., O”Brien, C., Yeates, Q., and Cresko, W. A., 2013, a. “The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing”. *Molecular ecology,* 22(11), pp. 2864-2883.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A. and Cresko, W.A., 2013, b. “Stacks: an analysis tool set for population genomics”. *Molecular Ecology*, no. 11, pp. 3124.

Ceballos, G., Ehrlich, P.R., Barnosky, A.D., García, A., Pringle, R.M. and Palmer, T.M., 2015. “Accelerated modern human–induced species losses: Entering the sixth mass extinction”. *Science advances,* 1(5), pp. e1400253.

Cheylan, G., 1991. “Patterns of Pleistocene turnover, current distribution and speciation among Mediterranean mammals.” *Biogeography of Mediterranean invasions*, pp.227-262.

Chen, X. Y., and He, F., 2009. “Speciation and endemism under the model of island biogeography.” *Ecology*, 90(1), pp. 39-45.

Ciesla, W.M., 2004. “Forests and forest protection in Cyprus”, *Forestry Chronicle,* no. 1, pp. 107.

Crutzen, P. J., 2002. “Geology of manking.” *Nature*, 415, 23.

Cucchi T., Vigne J.D., Auffray J.C., Croft P., Peltenburg E., 2002. “Passive transport of the house mouse *Mus musculus domesticus* to Cyprus at the Early Preceramic Neolitic (late 9th and 8th millenia cal. BC).” *C. R. Palevol*. 1, pp. 235–241.

Cucchi, T., Auffray, J. C., and Vigne, J. D., 2012. “Synanthropy and dispersal in the Near East and Europe: zooarchaeological review and perspectives.” *Evolution of the House mouse*, 3, pp. 65.

Cucchi, T., Vigne, J. D., and Auffray, J. C., 2005. “First occurrence of the house mouse (*Mus musculus domesticus* Schwarz and Schwarz, 1943) in the Western Mediterranean: a zooarchaeological revision of subfossil occurrences.” *Biological Journal of the Linnean Society,* 84(3), pp. 429-445.

Cucchi, T., Orth, A., Auffray, J.C., Renaud, S., Fabre, L., Catalan, J., Hadjisterkotis, E., Bonhomme, F. and Vigne, J.D., 2006. “A new endemic species of the subgenus *Mus* (*Rodentia*, Mammalia) on the Island of Cyprus”. *Zootaxa*, (1241), pp.1-36.

Dakin EE, Avise JC., 2004. “Microsatellite null alleles in parentage analysis.” *Heredity* 93, pp. 504–509.

Davey, J. W., Cezard, T, Fuentes-Utrilla, P, Eland, C, Gharbi, K and Blaxter, ML., 2013. “Special features of RAD Sequencing data: implications for genotyping.” *Molecular Ecology*, pp. 3151.

Deem, SL, Karesh, WB and Weisman, W 2001, “Putting Theory into Practice: Wildlife Health in Conservation.” *Conservation Biology -Boston Massachusetts-*, no. 5, pp. 1224-1233.

DeFaveri, J., Viitaniemi, H., Leder, E. and Merilä, J., 2013. “Characterizing genic and nongenic molecular markers: comparison of microsatellites and SNPs.” *Molecular Ecology Resources*, vol. 13, no. 3, pp. 377–392.

Depaulis, F., Mousset, S., Veuille, M., 2003. “Power of neutrality tests to detect bottlenecks and hitchhiking.” *J Mol Evol.* 57(Suppl 1): pp. 190–S200.

Dierking, J., Phelps, L., Praebel, K., Ramm, G., Prigge, E., Borcherding, J., Brunke, M., and Eizaguirre, C., 2014. “Anthropogenic hybridisation between endangered migratory and commercially harvested stationary whitefish taxa (*Coregonus* spp.).” *Evolutionary Applications*, no. 9, pp. 1068.

Diethart M., Ingo B., Wiebke M. and Teja T., 2004. “Population Size and the Risk of Local Extinction: Empirical Evidence from Rare Plants”. *Oikos*, vol. 105, no. 3, pp. 481.

Earl, D.A. and VonHoldt, B.M., 2012. “STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method.” *Conservation Genetics Resources*, no. 2, pp. 359.

Eble, J. A., Toonen, R. J. and Bowen, B. W., 2009. “Endemism and dispersal: comparative phylogeography of three surgeonfishes across the Hawaiian Archipelago.” *Marine Biology - Berlin- Springer Verlag-*, pp. 689.

Ekblom, R., and Galindo, J., 2011. “Applications of next generation sequencing in molecular ecology of non-model organisms.” *Heredity*, 107(1), pp. 1-15.

Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, ScachettiPereira R, Schapire RE, Soberon J, Williams S, Wisz MS, Zimmermann NE, 2006. “Novel methods improve prediction of species' distributions from occurrence data.” *Ecography* 29, pp. 129-151.

Ellegren H. 2004. “Microsatellites: simple sequences with complex evolution.” *Nature Reviews Genetics* 5, pp. 435–445.

Emerson, K. J., Merz, C. R., Catchen, J. M., Hohenlohe, P. A., Cresko, W. A., Bradshaw, W. E., and Holzapfel, C. M., 2010. “Resolving postglacial phylogeography using high-throughput sequencing.” *Proceedings of the national academy of sciences*, 107(37), pp. 16196-16200.

Engler R, Guisan A, Rechsteiner L, 2004. “An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data.” *J. Appl. Ecol.* 41, pp. 263-274.

Estoup A, Jarne P, Cornuet JM., 2002. “Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis.” *Molecular Ecology* 11, pp. 1591–1604.

Etter P.D., Preston J.L., Bassham S., Cresko W.A., Johnson E.A., 2011*. PLoS One*, 6(4):e18561.

Evanno G., Regnaut S., and, Goudet J., 2005. “Detecting the number of clusters of individuals using the software structure: a simulation study.” *Molecular ecology* 14(8), pp. 2611–2620.

Excoffier, L., 2004. “Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model.” *Molecular Ecology,* 13(4), pp. 853-864.

Excoffier, L. and 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: pp. 564-567.

Fenchel, T., and Finlay, B.J., 2004. “The Ubiquity of Small Species: Patterns of Local and Global Diversity.” *BioScience*, vol. 54, no. 8, pp. 777.

Ferrier S, Watson G, Pearce J, Drielsma M, 2002. “Extended statistical approaches to modeling spatial pattern in biodiversity: the north-east New South Wales experience. I. Species-level modeling.” *Biodiversity and Conservation* 11, pp. 2275–2307.

Fischer, M., Husi, R., Prati, D., Peintinger, M., Kleunen, M. van and Schmid, B. ,2000. “RAPD variation among and within small and large populations of the rare clonal plant *Ranunculus reptans* (*Ranunculaceae*).” *American Journal of Botany*, 87 (8), pp.1128–1137.

Frankham, R., 1997. “Do island populations have less genetic variation than mainland populations?”. *Heredity* 78, pp. 311–327.

Frankham, R., 1998. “Inbreeding and extinction: island populations”. *Cons Biol* 12, pp. 665–675.

Frankham, R., Ballou, J. D., Briscoe, D. A., and McInnes, K.H., 2010. "*Introduction to Conservation Genetics*." Cambridge University Press.

Frynta, D., Baladová, M., Eliášová, B., Lišková, S. and Landová, E., 2015. "Why not to avoid the smell of danger? Unexpected behavior of the Cypriot mouse surviving on the island invaded by black rats." *Current Zoology*, 61(4), pp.781-791.

Fu, Y.X., 1997. "Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection." *Genetics*, 147(2), pp.915-925.

García-Rodríguez, O., Andreou, D., Schutkowski, H., Stafford, R., Stewart, J.R., Hardouin, E.A., Herman, J.S., Mitsainas, G.P., Searle, J.B., Bonhomme, F., and Hadjisterkotis, E., 2018. "Cyprus as an ancient hub for house mice and humans." *Journal of Biogeography*, vol. 45, no. 12, pp. 2619.

Gaston KJ, 1996. "Species richness: measure and measurement. Biodiversity: a biology of numbers and difference (ed. K.J. Gaston)." *Blackwell Science*, Oxford, pp. 77–113.

Gautier, M., Gharbi, K., Cezard, T., Foucaud, J., Kerdelhu, C., Pudlo, P., Cornuet, J.M. and Estoup, A., 2013. "The effect of RAD allele dropout on the estimation of genetic variation within and between populations." *Molecular Ecology*, no. 11, pp. 3165.

Gillespie, R.G, Claridge, E.M. and Roderick, G.K., 2007. "Biodiversity dynamics in isolated island communities: interaction between natural and human-mediated processes". *Molecular Ecology*, no. 1, pp. 45.

Gippoliti, S. and Amori, G., 2004. "Mediterranean Island mammals: are they a priority for biodiversity conservation?". *Biogeographia–The Journal of Integrative Biogeography*, 25(1).

Graham, CF, Glenn, TC, McArthur, AG, Boreham, DR, Kieran, T, Lance, S, Manzon, RG, Martino, JA, Pierson, T, Rogers, SM, Wilson, JY & Somers, CM, 2015. "Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq)." *Molecular Ecology Resources*, vol. 15, no. 6, pp. 1304–1315.

Greider C.W., 1999. "telomeres do D-loop-T-loop". *Cell* 97, pp. 419-422.

Grimaldi MC, Crouau-Roy B., 1997. "Microsatellite allelic homoplasy due to variable flanking sequences." *Journal of Molecular Evolution* 44, pp. 336–340.

Gröning, J. and Hochkirch, A., 2008. "Reproductive interference between animal species". *The Quarterly Review of Biology,* 83(3), pp.257-282.

Guenet, J.L. and Bonhomme, F., 2003. "Wild mice: an ever-increasing contribution to a popular mammalian model." *Trends In Genetics*, no. 1, pp. 24.

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., Lepoittevin, C., Malausa, T., Revardel, E., Salin, F. and Petit, R.J., 2011. "Current trends in microsatellite genotyping." *Molecular ecology resources*, 11(4), pp.591-611.

Guisan A, Thuiller W, 2005. "Predicting species distribution: offering more than simple habitat models." *Ecol. Lett.* 8, pp. 993–1009.

Guisan A, Zimmermann NE, 2000. "Predictive habitat distribution models in ecology." *Ecol. Modell.* 135, pp. 147–186.

Günther, A., 1879. “Notice of a Collection of Mammals and Reptiles from Cyprus.” *In: Proceedings of the Zoological Society of London.* Oxford, UK: Blackwell Publishing Ltd, Vol. 47, No. 1, pp. 741-741.

Gupta, A., Bhardwaj, A., Sharma, P., and Pal, Y., 2015. “Mitochondrial DNA-a tool for phylogenetic and biodiversity search in equines.” *Journal of Biodiversity and Endangered Species*, S1: S1.006.

Hadjikyriakou, G., and Hadjisterkotis, E., 2002. “The adventive plants of Cyprus with new records of invasive species.” *Zeitschrift für Jagdwissenschaft,* 48(1), pp. 59-71.

Hadjimitsis, D.G., 2010. Brief communication" Determination of urban growth in catchment areas in Cyprus using multi-temporal remotely sensed data: risk assessment study*". Natural Hazards and Earth System Sciences*, 10(11), pp.2235-2240.

Hadjinicolaou, P., Giannakopoulos, C., Zerefos, C., Lange, M. A., Pashiardis, S., and Lelieveld, J., 2011. “Mid-21st century climate and weather extremes in Cyprus as projected by six regional climate models.” *Regional Environmental Change,* 11(3), 441-457.

Hadjisterkotis, E., 2001. “The Cyprus mouflon, a threatened species in a biodiversity “hotspot” area.” *Proceedings of the International Mouflon Symposium*, Sopron, pp. 71-81.

Hadjisterkotis, E., Masala, B., and Reese, D. S., 2000. “The origin and extinction of the large endemic Pleistocene mammals of Cyprus.” *Biogeographia–The Journal of Integrative Biogeography*, 21(1).

Hall, A. V., and Veldhuis, H. A., 1985. *“South African red data book: plants-Fynbos and Karoo biomes.”* National Scientific Programmes Unit: CSIR.

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

Hardouin, E.A., Andreou, D., Zhao, Y., Chevret, P., Fletcher, D.H., Britton, J.R. and Gozlan, R.E., 2018. “Reconciling the biogeography of an invader through recent and historic genetic patterns: the case of topmouth gudgeon *Pseudorasbora parva*.” *Biological invasions*, 20(8), pp.2157-2171.

Hardouin, E.A., Chapuis, J.L., Stevens, M.I., Van Vuuren, J.B., Quillfeldt, P., Scavetta, R.J., Teschke, M. and Tautz, D., 2010. “House mouse colonization patterns on the sub-Antarctic Kerguelen Archipelago suggest singular primary invasions and resilience against re-invasion.” *BMC Evolutionary Biology,* 10(1), pp.325.

Hardouin, E. A., Orth, A., Teschke, M., Darvish, J., Tautz, D., and Bonhomme, F., 2015. “Eurasian house mouse (*Mus musculus L.*) differentiation at microsatellite loci identifies the Iranian plateau as a phylogeographic hotspot.” *BMC Evolutionary Biology*,15, 26.

Hardouin, E.A., Andreou, D., Zhao, Y., Chevret, P., Fletcher, D.H., Britton, J.R. and Gozlan, R.E., 2018. “Reconciling the biogeography of an invader through recent and historic genetic patterns: the case of topmouth gudgeon *Pseudorasbora parva*”. *Biological Invasions*, 20 (8), pp. 2157-2171.

Harper GA, Cabrera LF, 2010. “Response of mice Mus musculus to the removal of black rats *Rattus rattus* in arid forest on Santa Cruz Island, Galapagos.” *Biol. Invasions* 12: 1449–1452.

Harrison, D. L. and Bates, P. J. J. 1991. “*The Mammals of Arabia*.” - Harrison Zool. Mus. Publ., Seuenoaks, England.

Harrison, R., and Larson, E., 2014. “Hybridisation, introgression, and the nature of species boundaries”. *The Journal Of Heredity*, 105 Suppl 1, pp. 795-809.

Heggenes, J., Kvie, K. S. and Røed, K. H., 2016. “Merging and comparing three mitochondrial markers for phylogenetic studies of Eurasian reindeer (*Rangifer tarandus*)” *Ecology and Evolution* (20457758), 6(13), pp. 4347–4358.

Hinten, G., Harriss, F., Rossetto, M., & Braverstock, P. R., 2003. “Genetic variation and island biogreography: microsatellite and mitochondrial DNA variation in island populations of the Australian bush rat, *Rattus fuscipes greyii*.” *Conservation Genetics*, 4(6), pp. 759-778.

Hodel, R.G.J., Chen, S., Payton, A.C., McDaniel, S.F., Soltis, P. and Soltis, D.E., 2017. “Adding loci improves phylogeographic resolution in red mangroves despite increased missing data: comparing microsatellites and RAD-Seq and investigating loci filtering.” *Scientific Reports*, vol. 7, no. 1, pp. 17598.

Hoffman JI, Amos W., 2005. “Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion.” *Molecular Ecology* 14, pp. 599–612.

Hoffman, J.I., Simpson, F., David, P., Rijks, J.M., Kuiken, T., Thorne, M.A., Lacy, R.C. and Dasmahapatra, K.K., 2014. “High-throughput sequencing reveals inbreeding depression in a natural population.” *Proceedings of the National Academy of Sciences*, 111(10), pp.3775-3780.

Hudson, M.E., 2008. “Sequencing breakthroughs for genomic ecology and evolutionary biology”. *Molecular ecology resources*, 8, pp. 3-17.

Irl, S.D.H., Schweiger, A.H., Medina, F.M., Fernández, P.J.M., Harter, D.E.V., Jentsch, A., Provenzale, A., Steinbauer, M.J., Beierkuhnlein, C. and Kueffer, C., 2017. “An island view of endemic rarity-Environmental drivers and consequences for nature conservation.” *Diversity and Distributions*, vol. 23, no. 10, pp. 1132–1142.

IŞIK, K., 2011. “Rare and endemic species: why are they prone to extinction?” *Turkish Journal of Botany*, vol. 35, no. 4, pp. 411–417.

IUCN 2019. The IUCN Red List of Threatened Species. Version 2019-2. https://www.iucnredlist.org

Jakobsson, M. and Rosenberg, N.A., 2007. “CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure.” *Bioinformatics -Oxford-,* no. 14, pp. 1801.

Johnson PCD, and Haydon DT., 2006. “Maximum-likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data.” *Genetics* 175, pp. 827–842.

Jombart, T., Devillard, S. and Balloux, F., 2010. “Discriminant analysis of principal components: a new method for the analysis of genetically structured populations.” *BMC Genetics*, vol. 11, pp. 94.

Jones AG, Ardren WR. 2003. "Methods of parentage analysis in natural populations.” Molecular Ecology 12, pp. 2511–2523.

Jones, M.E., Paetkau, D., Geffen, E.L.I. and Moritz, C., 2004. “Genetic diversity and population structure of Tasmanian devils, the largest marsupial carnivore”. *Molecular Ecology,* 13(8), pp. 2197-2209.

Kalinowski, S.T, 2005. “HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness”. *Molecular Ecology Notes* 5(1), pp. 187 – 189.

Kalinowski, S. T., 2005. “Do polymorphic loci require large sample sizes to estimate genetic distances?” *Heredity* 94, pp. 33-36.

Karli P, 1956. “The Norway rat’s killing response to the white mouse: An experimental analysis.” *Behaviour* 10 (1/2), pp. 81–103.

Koh, L.P., Dunn, R.R., Sodhi, N.S., Colwell, R.K., Proctor, H.C., and Smith, V.S., 2004. “Species Coextinctions and the Biodiversity Crisis”. *Science*, vol. 305, no. 5690, pp. 1632–1634.

Kronholm, I., Loudet, O., & de Meaux, J., 2010. “Influence of mutation rate on estimators of genetic differentiation-lessons from Arabidopsis thaliana”. *BMC genetics,* 11(1), pp. 33.

Kruglyak S., Durrett R.T., Schug M.D., Aquadro C.F., 1998. “Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations”. *Proc. Natl. Acad. Sci. USA* 95, pp. 10774–10778.

Kryštufek, B., and Vohralík, V., 2001. “*Mammals of Turkey and Cyprus.*” Zgodovinsko društvo za južno Primorsko.

Kryštufek, B., Vohralík, V., and Janžekovič, F., 2009. “*Mammals of Turkey and Cyprus: Rodentia II: Cricetinae, Muridae, Spalacidae, Calomyscidae, Capromydae, Hystricidae, Castoridae*.” Univerza na Primorskem, Znanstveno-raziskovalno središče, Založba Annales.

Kumar, S., & Stohlgren, T. J., 2009. “Maxent modeling for predicting suitable habitat for threatened and endangered tree *Canacomyrica monticola* in New Caledonia.” *Journal of Ecology and natural Environment*, 1(4), pp. 94-98.

Kvie K. S., Hogner S., Aarvik L., Lifjeld J. T., and Johnsen A., 2013. “Deep sympatric mtDNA divergence in the autumnal moth (*Epirrita autumnata*).” *Ecol. Evol*. 3, pp. 126–144.

Lakra W.S., Goswami M., Gopalakrishnan A., Singh D.P., Singh A., Nagpure N.S., 2010. “Genetic relatedness among fish species of Genus *Channa* using mitochondrial DNA genes.” *Biochem. Syst. Ecol*. 38, pp. 1212-1219.

Lanfear, R., Hua, X., & Warren, D. L., 2016. “Estimating the effective sample size of tree topologies from Bayesian phylogenetic analyses.” *Genome biology and evolution*, 8(8), pp. 2319-2332.

Larson, W.A., Seeb, L.W., Everett, M.V., Waples, R.K., Templin, W.D. and Seeb, J.E., 2014. “Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*)”. *Evolutionary Applications*, 7(3), pp.355-369.

Leigh, J.W. & Bryant, D., 2015. “POPART: full-feature software for haplotype network construction”. *Methods Ecol. Evol.* **6**, pp. 1110–1116.

Lexer C, Buerkle CA, Joseph JA, Heinze B, Fay MF., 2007. “Admixture in European Populus hybrid zones makes feasible the mapping of loci that contribute to reproductive isolation and trait differences.” *Heredity* 98, pp. 74–84.

Li H, Ruan J, Durbin R., 2008, “Mapping short DNA sequencing reads and calling variants using mapping quality scores.” *Genome Res.,* 18 (11), pp. 1851-8.

Librado, P., and Rozas J., 2009. “DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.” *Bioinformatics* 25, pp. 1451–1452.

Luikart, G. and Cornuet, J.M., 1998. “Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data”. *Conservation biology*, 12(1), pp.228-237.

Macholán, M., Vyskočilo vá, M., Bonhomme, F., Kryštufek, B., Orth, A., Vohralík, V., 2007. “Genetic variation and phylogeography of free-living mouse species (genus *Mus*) in the Balkans and the Middle East.” *Mol. Ecol*., 16, pp. 4774-4788.

Malcolm, J. R., Liu, C., Neilson, R. P., Hansen, L., and Hannah, L. E. E., 2006. “Global warming and extinctions of endemic species from biodiversity hotspots.” *Conservation biology*, 20(2), pp. 538-548.

Marra, A.C., 2005. “Pleistocene mammals of Mediterranean islands”, *Quaternary International*, vol. 129, no. 1, pp. 5–14.

Martinez-Arias R., Calafell F., Mateu E., Comas D., Andres A., Bertranpetit J., 2001. “Sequence variability of a human pseudogene”. *Genome Res*. 11, pp. 1071–1085.

McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C. and Brumfield, R.T., 2013. “Applications of next-generation sequencing to phylogeography and phylogenetics.” *Molecular Phylogenetics And Evolution*, vol. 66, no. 2, pp. 526–538.

McGowan, A., Broderick, A. C., Deeming, J., Godley, B. J., and Hancock, E. G., 2001. “Dipteran infestation of loggerhead (*Caretta caretta*) and green (*Chelonia mydas*) sea turtle nests in Northern Cyprus.” *Journal of Natural History*, 35(4), pp. 573-581.

McIntosh EJ, Rossetto M, Weston PH, Wardle GM., 2014. “Maintenance of strong morphological differentiation despite ongoing natural hybridization between sympatric species of Lomatia (*Proteaceae*).” *Annals of Botany* 113, pp. 861–872.

Mellink, E., Ceballos, G., Luévano, J., 2002. “Population demise and extinction threat of the Angel de la Guarda deer mouse (*Peromyscus guardia*)”. *Biological Conservation*, 180, pp. 107-111.

Miller CR, Joyce P, Waits LP., 2002. “Assessing allelic dropout and genotyping reliability using maximum likelihood.” *Genetics* 160, pp. 357–366.

Miller, W., Hayes, V.M., Ratan, A., Petersen, D.C., Wittekindt, N.E., Miller, J., Walenz, B., Knight, J., Qi, J., Zhao, F. and Wang, Q., 2011. “Genetic diversity and population structure of the endangered marsupial *Sarcophilus harrisii* (Tasmanian devil).” *Proceedings of the National Academy of Sciences*, 108(30), pp.12348-12353.

Morin, P. A., Luikart, G., & Wayne, R. K., 2004. “SNPs in ecology, evolution and conservation”. *Trends in ecology & evolution*, 19(4), pp. 208-216.

Muhlfeld, C.C., Kalinowski, S.T., McMahon, T.E., Taper, M.L., Painter S., Leary R.F., Allendorf F.W., 2009. “Hybridisation rapidly reduces fitness of a native trout in the wild.” *Biology letters*, 5, pp. 328-331.

Munshi-South, J., Zolnik, C. P. and Harris, S. E., 2016. “Population genomics of the Anthropocene: urbanization is negatively associated with genome-wide variation in white-footed mouse populations”. *Evolutionary Application*s, 9(4), pp. 546–564.

Musser G.G. and Carleton M.D., 2005. “Superfamily Muroidea. *In: Wilson DE, Reeder DM ed. Mammal Species of the World: A Taxonomic and Geographic Reference*.” Vol 2. 3rd edn. Baltimore: The Johns Hopkins University Press, pp. 894–1531.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., and Kent, J., 2000. “Biodiversity hotspots for conservation priorities.” *Nature*, 403(6772), pp. 853.

Nabholz, B., Glémin, S., and Galtier, N., 2008. “Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis.” *Molecular biology and evolution*, 25(1), pp. 120-130.

Navascués, M. and Emerson, B. C., 2005. “Chloroplast microsatellites: Measures of genetic diversity and the effect of homoplasy.” *Molecular Ecology*, 14 (5), pp.1333–1341.

Neuwald, J.L., 2010. “Population isolation exacerbates conservation genetic concerns in the endangered Amargosa vole, *Microtus californicus scirpensis*.” *Biol Conserv* 143, pp. 2028–2038.

Orsini P., Bonhomme F., Britton-Davidian J., 1983. Le complexe d”espèces du genre Mus en Europe Centrale et Orientale. II Critères d”identification, répartition et caractéristiques écologiques.” *Zeitschrift für Säugetierkunde,* 48, pp. 86–95

Ouborg, N. J., Pertoldi, C., Loeschcke, V., Bijlsma, R. and Hedrick, P. W., 2010. “Conservation genetics in transition to conservation genomics.” *Trends in Genetics*, 26 (4), pp.177–187.

Pastorini, J., Zaramody, A., Curtis, D.J., Nievergel, C.M. and Mundy, N.I., 2009. “Genetic analysis of hybridisation and introgression between wild mongoose and brown lemurs.” *BMC Evolutionary Biology*, vol. 9, pp. 1–13.

Pemberton JM, Slate J, Bancroft DR, Barrett JA., 1995. “Nonamplifying alleles at microsatellite loci - a caution for parentage and population studies.” *Molecular Ecology* 4, pp. 249–252.

Perry, W.,L., Lodge, D.,M., Feder, J.,L., 2002. “Importance of hybridisation between indigenous and nonindigenous freshwater species: an overlooked threat to North American biodiversity”. *Systematic biology*, 51, pp. 255-275.

Peterson B.K., Weber J.N., Kay E.H., Fisher H.S., and Hoekstra H.E., 2012. ”Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species.” *PLoS ONE* 7(5) pp. e37135.

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts, C.M. and Sexton, J.O., 2014. “The biodiversity of species and their rates of extinction, distribution, and protection.” *Science* (New York, N.Y.), vol. 344, no. 6187, pp. 1246752.

Piry, S., Luikart, G., Cornuet, J.M., 1995. “Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data.” *Heredity* 90(4).

Pritchard, J.K., Stephens, M. & Donnelly, P., 2000. ‘Inference of population structure using multilocus genotype data’. *Genetics -USA-,* no. 2, pp. 945

Primack, R.B., 2014. “*Essentials of conservation biology*.” Sunderland: Sinauer Associates, Sixth edition.

Pritchard, J. K., Stephens, M. and Donnelly, P., 2000. “Inference of population structure using multilocus genotype data”. *Genetics*, 155(2), pp. 945-959.

Pruett, C., and Winker, K., 2008. “The effects of sample size on population genetic diversity estimates in song sparrows *Melospiza melodia*.” *Journal of Avian Biology,* 39(2), pp. 252-256.

Purvis, A., Gittleman, J. L., Cowlishaw, G., and Mace, G. M., 2000. “Predicting extinction risk in declining species.” *Proceedings of the royal society of London. Series B: Biological Sciences,* 267(1456), pp. 1947-1952.

Putman, A.I. and Carbone, I., 2014. “Challenges in analysis and interpretation of microsatellite data for population genetic studies”. *Ecology and Evolution* (20457758), vol. 4, no. 22, pp. 4399–4428.

Pysek, P., Jarosík, V., Hulme, P.E., Kühn, I., Wild, J., Arianoutsou, M., Bacher, S., Chiron, F., Didziulis, V., Essl, F., Genovesi, P., Gherardi, F., Hejda, M., Kark, S., Lambdon, P.W., Desprez-Loustau, M.-L., Nentwig, W., Pergl, J., Poboljsaj, K., Rabitsch, W., Roques, A., Roy, D.B., Shirley, S., Solarz, W., Vilà, M. and Winter, M., 2010. ““Disentangling the role of environmental and human pressures on biological invasions across Europe.” *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 107, no. 27, pp. 12157–12162.

Randi E., 2008. “Detecting hybridization between wild species and their domesticated relatives.” *Molecular Ecology* 17, pp. 285–293.

Rebelo, A. G., and Siegfried, W. R., 1990. “Protection of fynbos vegetation: ideal and real-world options.” *Biological Conservation*, 54(1), pp. 15-31.

Reese, D., 1999. “*Mouse. In: Simons, A.H. (ed.) Faunal extinction in the island society: Pygmy hippopotamus hunters on Cyprus*.” Kluwer Academic Press, New York, pp: 169-170.

Reyes, A., Gissi, C., Catzeflis, F., Nevo, E., Pesole, G. and Saccone, C., 2004. “Congruent mammalian trees from mitochondrial and nuclear genes using Bayesian methods.” *Molecular biology and evolution*, 21(2), pp.397-403.

Robertson, B.C., Gemmell, N.J., 2004. “Defining eradication units to control invasive pests.” *Journal of Applied Ecology*, 41, pp. 1042-1048.

Rogers, A.R., and H. Harpending., H., 1992. “Population growth makes waves in the distribution of pairwise genetic differences.” *Mol. Biol. Evol*. 9: pp. 552-569.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., and Huelsenbeck, J., P., 2012. “MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space.” *Systematic Biology*,61, 539–542.

Rosenberg, N.A., 2004. “distruct: a program for the graphical display of population structure.” *Molecular Ecology Notes*, no. 1, pp. 137.

Sage, R. D., Heyneman, D., Lim, K. C., and Wilson, A. C., 1986. “Wormy mice in a hybrid zone.” *Nature*, 324(6092), 60.

Sakaguchi S, Bowman DMJS, Prior LD, Crisp MD, Linde CC, Tsumura Y, Isagi Y, 2013. “Climate, not Aboriginal landscape burning, controlled the historical demography and distribution of fire-sensitive conifer populations across Australia.” *Proceedings of the Royal Society B-Biological Sciences* 280, pp. 20132182.

Saveriades, A., 2000. “Establishing the social tourism carrying capacity for the tourist resorts of the east coast of the Republic of Cyprus.” *Tourism management*, 21(2), pp.147-156.

Schlötterer, C., 2000. “Evolutionary dynamics of microsatellite DNA.” *Chromosoma*, 109 (6), pp. 365–371.

Schlötterer, C., 2004. “The evolution of molecular markers—just a matter of fashion?”. *Nature reviews genetics*, 5(1), pp.63.

Schneider S, and Excoffier L., 1999. “Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA.” *Genetics*, 102, pp. 1079–1089.

Schwartz, Michael K., Gordon Luikart, and Robin S. Waples, 2007. “Genetic Monitoring as a Promising Tool for Conservation and Management.” *Trends in Ecology and Evolution* 22 (1) pp. 25–33.

Selkoe K.A., Toonen R.J., 2006. “Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers.” *Ecol Lett.,* 9 pp.615–29.

Shanks AL, Grantham BA, Carr MH, 2003. “Propagule dispersal distance and the size and spacing of marine reserves.” *Ecol Appl* 13: pp.159–S169.

Siegel DA, Kinlan BP, Gaylord B, Gaines SD, 2003. “Lagrangian descriptions of marine larval dispersion.” *Mar Ecol Prog Ser* 260, pp.83–96.

Silva P, Guan X, Ho-Shing O, Jones J, Xu J, Hui D, Notter D, Smith E, 2009. “Mitochondrial DNA-based analysis of genetic variation and relatedness among Sri Lankan indigenous chickens and the Ceylon Junglefowl (*Gallus lafayetti*).” *Anim. Genet*. 40, pp. 1-9.

Silvertown, J. and Charlesworth, D., 2001. “*Introduction to plant population biology*.” 4th ed. London, UK: Blackwell Science.

Smith, O., & Wang, J., 2014. “When can noninvasive samples provide sufficient information in conservation genetics studies?”. *Molecular Ecology Resources,* 14(5), pp. 1011-1023.

Song Y., Endepols S., Klemann N., Richter D., Matuschka F., Shih C., Nachman M., and Kohn M., 2011. “Adaptive Introgression of Anticoagulant Rodent Poison Resistance by Hybridisation between Old World Mice.” *Current Biology*, 15, pp. 1296- 1301.

Steadman, D.W., 2006. “*Extinction and biogeography of tropical Pacific birds*.” University of Chicago Press.

Stebbins, G. L., 1942. “The genetic approach to problems of rare and endemic species.” *Madrono*, 6(8), pp. 241-258.

Stobie, C.S., Cunningham, M.J., Oosthuizen, C.J. and Bloomer, P., 2019. “Finding stories in noise: Mitochondrial portraits from RAD data”. *Molecular Ecology Resources*, vol. 19, no. 1, pp. 191–205.

Strassburg, B. B., Rodrigues, A. S., Gusti, M., Balmford, A., Fritz, S., Obersteiner, M., Kerry Turner, R., and Brooks, T. M., 2012. “Impacts of incentives to reduce emissions from deforestation on global species extinctions.” *Nature Climate Change*, 2(5), pp. 350.

Sunnucks, P., 2000. “Efficient genetic markers for population biology.” *Trends in Ecology and Evolution,* 15, pp. 199–203.

Taylor S., Larson E., and Harrison R., 2015. “Hybrid zones: windows on climate change”. *Trends In Ecology and Evolution*, 30, 7, p. 398-406.

Tajima, F., 1989. “Statistical method for testing the neutral mutation hypothesis by DNA polymorphism.” *Genetics*, 123(3), pp.585-595.

Thomas, C. D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N, de Siqueira, M.F., Grainger, A. and Hannah, L., 2004. “Extinction risk from climate change.” *Nature -London-*, no. 6970, pp. 145.

Todesco, M., Pascual, M.A., Owens, G.L., Ostevik, K.L., Moyers, B.T., Hübner, S., Heredia, S.M., Hahn, M.A., Caseys, C., Bock, D.G. and Rieseberg, L.H., 2016. “Hybridisation and extinction”, *Evolutionary Applications*, vol. 9, no. 7, pp. 892.

Toro, M., Barraga´n, C,O´vilo, C., Rodrigan˜ez, J., Rodriguez, C.and Silio´, L., 2002. “Estimation of coancestry in Iberian pigs using molecular markers.” *Conserv. Genet.* 3, pp. 309-320.

Trigo TC, Schneider A, de Oliveira TG, Lehugeur LM, Silveira L, Freitas TRO, Eizirik E., 2013. “Molecular data reveal complex hybridization and a cryptic species of Neotropical wild cat.” *Current Biology* 23, pp. 2528–2533.

Tsangari, H., Paschalidou, A. K., Kassomenos, A. P., Vardoulakis, S., Heaviside, C., Georgiou, K. E., and Yamasaki, E. N., 2016. “Extreme weather and air pollution effects on cardiovascular and respiratory hospital admissions in Cyprus.” *Science of the Total Environment*, 542, pp. 247-253.

Tsintides, T.C. and L. Kourtellarides, 1998. "*The endemic plants of Cyprus*." Bank of Cyprus Group and Cyprus Association of Professional Foresters, Nicosia. pp. 123.

Turner, L. M., and Harr, B., 2014. "Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions." *Elife*, 3, pp. e02504.

Unger, F., and Kotschy, T., 1865. "*Die insel Cypern: ihrer physischen und organischen natur nach mit rücksicht auf ihre früherre geschichte, geschildert.*" Рипол Классик.

Ursenbacher S., Carlsson M., Helfner V., Tegelstrom H., and Fumagalli L., 2006. "Phylogeography and Pleistocene refugia of the adder (*Vipera berus*) as inferred from mitochondrial DNA sequence data." *Mol. Ecol.* 15, pp. 3425–3437.

Väli, Ü., Einarsson, A., Waits, L., and Ellegren, H., 2008. "To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations?" *Molecular ecology*, 17(17), pp. 3808-3817.

Vázquez-Domínguez, E., Ceballos, G., Cruzado, J., 2004. "Extirpation of an insular subspecies by a single introduced cat: The case of the endemic deer mouse Peromyscus guardia on Estanque Island, Mexico". *Oryx*, 38, pp. 347-350.

Vences, M., Wollenberg, K. C., Vieites, D. R., and Lees, D. C., 2009. "Madagascar as a model region of species diversification." *Trends in ecology and evolution*, 24(8), pp. 456-465.

Vega, R., Vázquez-Domínguez, E., Mejía-Puente, A., & Cuarón, A. D., 2007. "Unexpected high levels of genetic variability and the population structure of an island endemic rodent (Oryzomys couesi cozumelae)". *Biological Conservation*, 137(2), pp. 210-222.

Vié, J.-C., Hilton-Taylor, C., Pollock, C., Ragle, J., Smart, J., Stuart, S. and Tong, R., 2008. "The IUCN Red List: A key conservation tool, Gland, International Union for Conservation of Nature."

Vignal, A., Milan, D., SanCristobal, M., & Eggen, A., 2002. "A review on SNP and other types of molecular markers and their use in animal genetics". *Genetics selection evolution*, 34(3), pp. 275-305.

Violaris, M., Vasquez, M. I., Samanidou, A., Wirth, M. C., and Hadjivassilis, A., 2009. "The mosquito fauna of the Republic of Cyprus: a revised list." *Journal of the American Mosquito Control Association,* 25(2), pp. 199-203.

Wang, Y.Q, Williams, D.A., Gaines, M.S., 2005. "Evidence for a recent genetic bottleneck in the endangered Florida Keys silver rice rat (Oryzomys argentatus) revealed by microsatellite DNA analyses". *Conservation Genetics*, 6, pp. 575-585.

Wanga, T. Y., Hsua, H. H., Chena, Y. M., Chuac, C. S., Kuoc, H. C., and Chena, T. Y., 2015. "Mitochondrial D-loop as Marker for Genetic Diversity Study of Giant Grouper (*Epinephelus lanceolatus*)" Broodstocks in Taiwan. *J. Fish. Soc. Taiwan,* 42(2), pp. 87-94.

Welz, G., 2015. "*European products: Making and unmaking heritage in Cyprus*." Berghahn Books.

Whaley, D. J., and Dawes, J. C., 2003. "Cyprus breeding birds atlas." *Bird Census*, 63.

Wilson, J. R., Dormontt, E. E., Prentis, P. J., Lowe, A. J., and Richardson, D. M., 2009. “Something in the way you move: dispersal pathways affect invasion success.” *Trends in ecology and evolution*, 24(3), 136-144.

Wilson, K.A., McBride, M.F., Bode, M. and Possingham, H.P. 2006. “Prioritizing global conservation efforts.” *Nature -London-*, no. 7082, pp. 337.

Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, NCEAS Predicting Species Distributions Working Group, 2008. “Effects of sample size on the performance of species distribution models.” *Divers. Distrib.* 14, pp. 763-773.

Worm, B., Barbier, E.B., Beaumont, N., Duffy, J.E., Folke, C., Halpern, B.S., Jackson, J.B., Lotze, H.K., Micheli, F., Palumbi, S.R. and Sala, E., 2006. “Impacts of biodiversity loss on ocean ecosystem services”. *Science*, 314(5800), pp.787-790.

Yonekawa, Y., Harada, A., Okada, Y., Funakoshi, T., Kanai, Y., Takei, Y., Terada, S., Noda, T., and Hirokawa, N., 1998. “Defect in synaptic vesicle precursor transport and neuronal cell death in KIF1A motor protein–deficient mice.” *The Journal of cell biology*, 141(2), pp. 431-441.

Zachariadis, T., 2012. “Climate Change in Cyprus: Impacts and Adaptation Policies”. *Cyprus Economic Policy Review*, Vol.6, No. 1, pp. 21-37.

Zalasiewicz, J., Williams, M., Fortey, R., Smith, A., Barry, T. L., Coe, A. L., Bown, P. R., Rawson, P. F., Gale, A., Gibbard, P., Gregory, F. J., Hounslow, M. W., Kerr, A. C., Pearson, P., Knox, R., Powell, J., Waters, C., Marshall, J., Oates, M. and Stone, P., 2011. “Stratigraphy of the Anthropocene”. *Philosophical Transactions of the Royal Society* A, 369, pp. 1036-1055.

Appendices

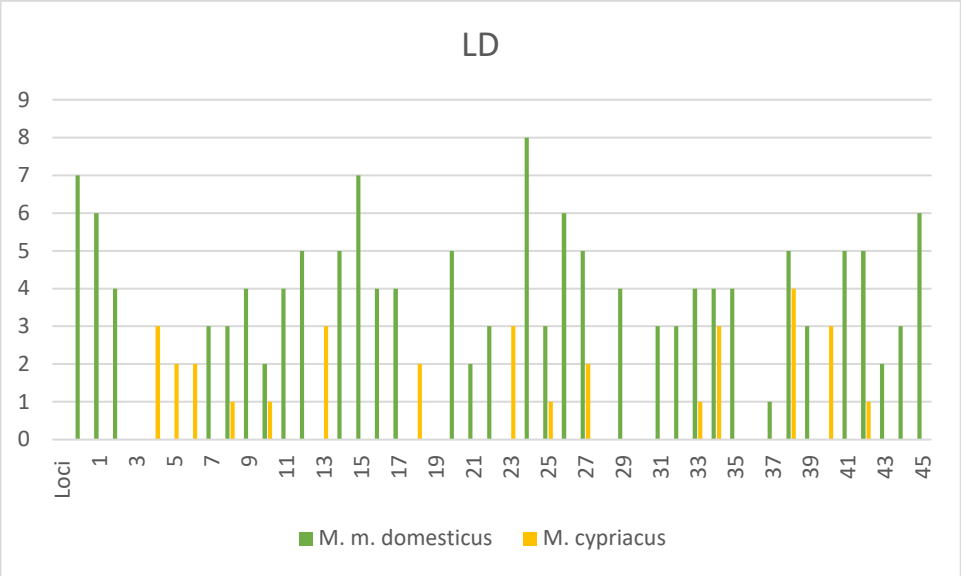


Figure 1.a – **Linkage Disequilibrium RADseq –** a total of 46 loci in common between *M. m. domesticus* and *M. cypriacus* were used to look at the linkage disequilibrium. The number of linked loci (x axis) are reported in the graph above respectively for each species at specific locus (y axis). The green bars are for *M. m. domesticus* and the orange bars are for the *M. cypriacus*

LOCUS	**POPULATION**	
	M. m. domesticus	*M. cypriacus*
18510:49		
H exp.	0.5277	0.5000
H obs.	1	1
19331:6:		
H exp.	0.5293	0.5000
H obs.	1	1
40473:3:		
H exp.	0.5135	0.5000
H obs.	1	1
57594:48		
H exp.	0.5000	0.5000
H obs.	1	1
76131:6:		
H exp.	0.5000	0.5744
H obs.	1	1
87856:3:		

H exp.	0.5000	0.5612
H obs.	1	1
87894:22		
H exp.	0.5000	0.7041
H obs.	1	1
110544:9		
H exp.	0.5413	0.5355
H obs.	1	1
121222:2		
H exp.	0.5256	0.5382
H obs.	1	1
189487:4		
H exp.	0.5256	0.5000
H obs.	1	1
192571:1		
H exp.	0.5460	0.5382
H obs.	1	1
215754:9		
H exp.	0.5143	0.5000
H obs.	1	1
341731:1		
H exp.	0.5278	0.5000
H obs.	1	1
354424:5		
H exp.	0.5000	0.6488
H obs.	1	1
363626:3		
H exp.	0.5382	0.5000
H obs.	1	1
381065:4		
H exp.	0.6169	0.5000
H obs.	1	1
446302:1		
H exp.	0.5571	0.5000
H obs.	1	1

548063:9		
H exp.	0.5243	0.5355
H obs.	1	1
570989:9		
H exp.	0.5000	0.5744
H obs.	1	1
586085:3		
H exp.	0.5000	0.5000
H obs.	1	1
609303:4		
H exp.	0.5256	0.5000
H obs.	1	1
621561:6		
H exp.	0.5382	0.5000
H obs.	1	1
621587:6		
H exp.	0.5269	0.5000
H obs.	1	1
688196:1		
H exp.	0.5000	0.6600
H obs.	1	1
852088:6		
H exp.	0.5135	0.5000
H obs.	1	1
925079:4		
H exp.	0.5464	0.5710
H obs.	1	1
925341:1		
H exp.	0.5249	0.5000
H obs.	1	1
986650:6		
H exp.	0.5238	0.5864
H obs.	1	1
987240:5		
H exp.	0.5000	0.5000
H obs.	1	1

987396:7		
H exp.	0.5547	0.5000
H obs.	1	1
987400:3		
H exp.	0.5000	0.5355
H obs.	1	1
987450:1		
H exp.	0.5128	0.5000
H obs.	1	1
987469:1		
H exp.	0.5506	0.5000
H obs.	1	1
987494:1		
H exp.	0.6050	0.5694
H obs.	1	1
987496:2		
H exp.	0.5172	0.5355
H obs.	1	1
987520:1		
H exp.	0.5238	0.5000
H obs.	1	1
988653:1		
H exp.	0.5000	0.5000
H obs.	1	1
988901:3		
H exp.	0.5262	0.5000
H obs.	1	1
989624:1		
H exp.	0.6012	0.6094
H obs.	1	1
990511:3		
H exp.	0.5143	0.5000
H obs.	1	1
991164:8		
H exp.	0.5000	0.6172

H obs.	1	1
991167:2		
H exp.	0.6275	0.5000
H obs.	1	1
991450:1		
H exp.	0.5256	0.5651
H obs.	1	1
991934:3		
H exp.	0.5119	------
H obs.	1	0.0000
992004:6		
H exp.	0.5166	0.5000
H obs.	1	1
992155:2		
H exp.	0.6071	------
H obs.	1	0.0000

Table 1.a – Table reporting expected and observed heterozygosity for each locus, respectively for 41 *M. m. domesticus*

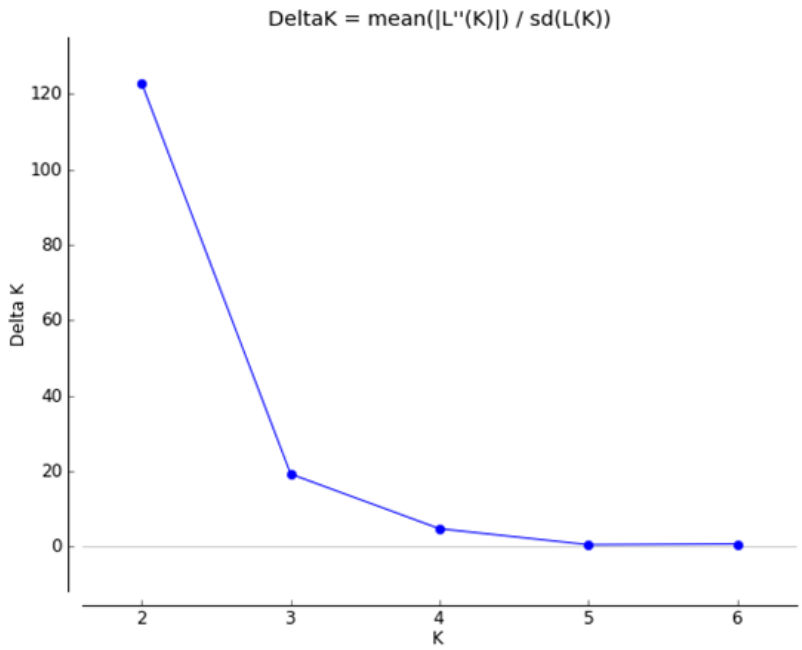


Figure 2.a – **Delta K RADseq all**- shows only the uppermost clustering level, not necessarily the actual number of subpopulations. In this case, K= 2 is the most recommended

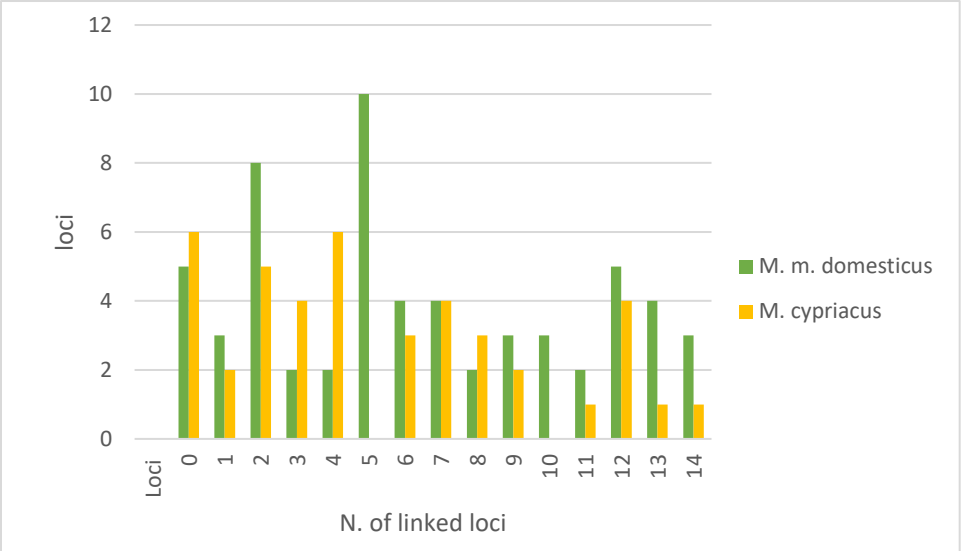


Figure 3.a – **Linkage Disequilibrium msat –** a total of 15 microsatellite loci in common between *M. m. domesticus* and *M. cypriacus* were used to look at the linkage disequilibrium. The number of linked loci (x axis) are reported in the graph above respectively for each species at specific locus (y axis). The green bars are for *M. m. domesticus* and the orange bars are for the *M. cypriacus*

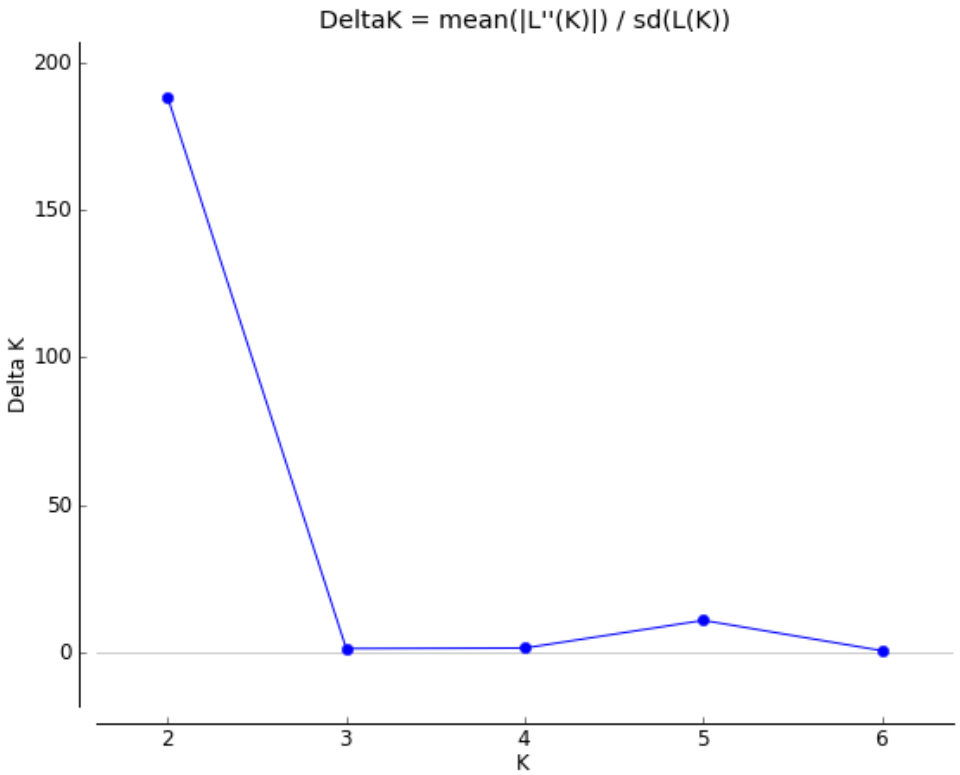
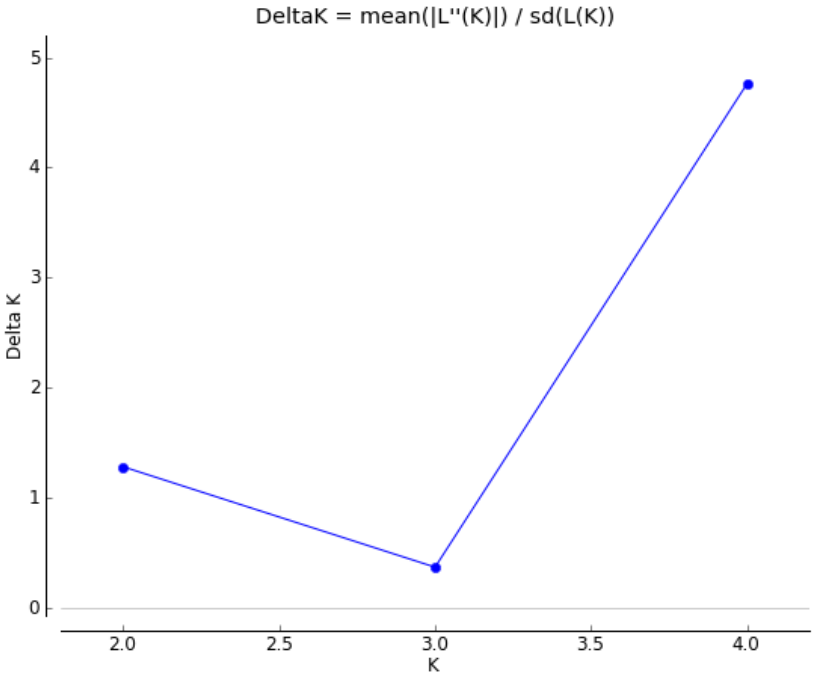


Figure 4.a – **Delta K msat all**- shows only the uppermost clustering level, not necessarily the actual number of subpopulations. In this case, K= 2 is the most recommended



clustering level, not necessarily the actual number of subpopulations

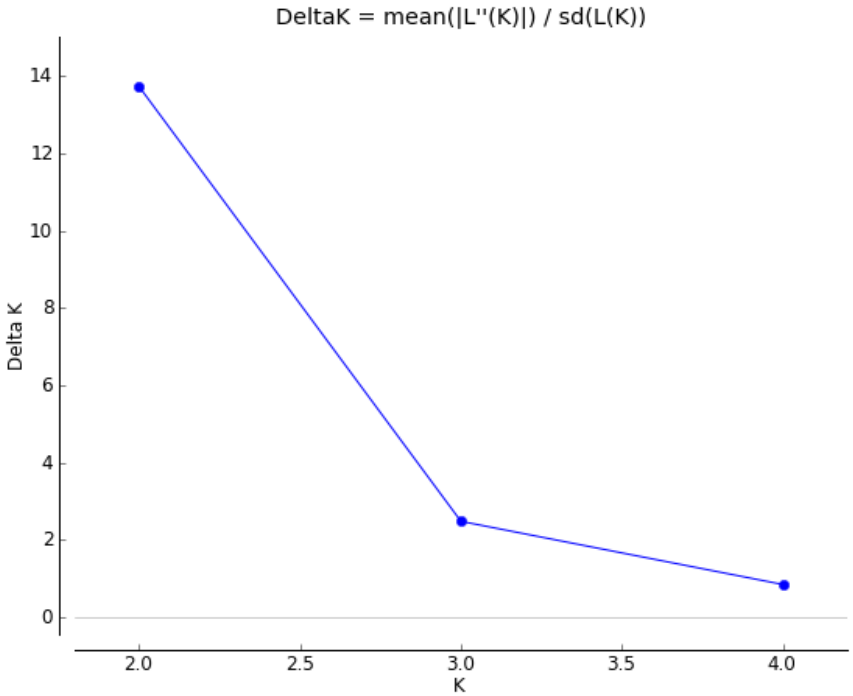


Figure 6.a – **Delta K msat *M. cypriacus***- shows only the uppermost clustering level, not necessarily the actual number of subpopulations

LocName	Ho	Hs	Ht	Dst	Dst"	Ht"	Fst	Fst"	Fis
Chr02_	0.690	0.920	0.939	0.019	0.038	0.958	0.020	0.040	0.250
Chr09_	0.548	0.895	0.918	0.023	0.045	0.940	0.025	0.048	0.388
Chr11_	0.583	0.925	0.928	0.003	0.006	0.931	0.003	0.006	0.369
Chr08_	0.690	0.898	0.937	0.039	0.078	0.976	0.042	0.080	0.231
Chr16_	0.443	0.857	0.857	0.000	0.000	0.857	0.000	0.000	0.483
Chr18_	0.929	0.893	0.896	0.003	0.006	0.899	0.003	0.006	-0.040
Chr19_	0.833	0.862	0.898	0.036	0.073	0.935	0.040	0.078	0.033
Chr04_	0.667	0.917	0.910	-0.007	-0.014	0.903	-0.008	-0.015	0.273
Chr13_	0.833	0.900	0.908	0.008	0.017	0.917	0.009	0.018	0.074
Chr17_	0.417	0.675	0.671	-0.004	-0.008	0.667	-0.006	-0.013	0.383
Chr03_	0.762	0.776	0.781	0.005	0.010	0.786	0.006	0.013	0.018
Chr05_	0.857	0.668	0.679	0.011	0.022	0.690	0.016	0.032	-0.283
Chr05_	0.657	0.829	0.890	0.060	0.121	0.950	0.068	0.127	0.207
Chr07_	0.845	0.865	0.870	0.005	0.010	0.875	0.006	0.011	0.023
Chr14_	0.833	0.926	0.915	-0.011	-0.021	0.905	-0.012	-0.023	0.100
Chr15_	1.000	0.933	0.948	0.016	0.032	0.964	0.017	0.033	-0.072
Overall	0.724	0.859	0.872	0.013	0.026	0.885	0.015	0.029	0.157

Table 2.a – Table reporting expected and observed heterozygosity based on the type of markers, Populations pairwise F_{ST} and the inbreeding coefficient for the microsatellites loci of *M. cypriacus.*