UNIVERSITÀ POLIT'ECNICA DELLE MARCHE

DIPARTIMENTO SCIENZE DELLA VITA E DELL'AMBIENTE

Corso di Laurea Magistrale

BIOLOGIA MOLECOLARE E APPLICATA

Genetic variation by means mtDNA and nDNA in a *Podarcis* population

Variazione genetica mediante mtDNA e nDNA in una popolazione di *Podarcis*

Relatore:
Prof.ssa
Oliana Carnevali

Tesi di Laurea
Magistrale di:
Vincenza Fiorentino

Correlatori:
Prof.ssa
Cori Ramon

Dott.ssa
Marta Bassitta

Sessione: Straordinaria di Febbraio

Anno Accademico: 2019/2020

A la isla de Mallorca,
una experiencia formidable que attesoraré
en mi corazòn.

Index

Summary

Riassunto

1. Introduction

1.1 Mitochondrial and Nuclear DNA

1.1.1 Structure

1.1.2 Functions

1.1.3 Applications

1.2 Study Model: Podarcis hispanicus

1.2.1 Description

1.2.2 Geographic Distribution Area

1.2.3 Genus, species and subspecies: a complex path

1.3 Objective

2. Material and Methods

2.1 Molecular Analysis

2.1.1 Extraction of DNA from tissue

2.1.2 Target Gene Amplification

2.1.3 Agarose Gel

2.1.4 DNA Purification

2.1.5 Sequencing Reaction

2.1.6 DNA Precipitation

2.1.7 Sequencing with ABI 6130 Sequencer

2.2 Bioinformatic Analysis

2.2.1 Codon Code Aligner

2.2.2 BioEdit v7.2.5.

2.2.3 Mega X

2.2.4 DNAsp v6.

3. Results

3.1 Phylogenetic trees

3.2 Genetic diversity

4. Conclusions

5. Acknowledgment

6. Bibliography

Summary

This thesis work is focused on the genetic analysis of two mitochondrial genes (Region Control and two fragments of Cytochrome b) and a nuclear gene (MC1R) to assess population structure and genetic diversity in a population of Podarcis, located in the southeast of the Iberian Peninsula, in a place called Galera. Were implemented laboratory protocols of DNA extraction using the chloroform/phenol method, PCR, electrophoretic stroke in agarosio and use of the Gel Capture program to visualize DNA bands, purification using the INVITEK MSB kit® Spin PCRapace, preparation for the sequencing reaction by inserting the Dye Terminator, ethanol-purified DNA precipitation and sequencing with the Applied Biosystem's ABI 3130 sequencer, to obtain the nucleotide sequences of the genes examined. Subsequently, computational analysis procedures were used to build phylogenetic trees (with Mega X software) and to study the genetic diversity (with DNAsp software) of the population under consideration and other known *Podarcis* species. From the cladistic based on molecular data relating to the specimens of interest we can highlight the evolutionary relationships of the various taxa, as well as the degree of kinship; in particular, samples from the geographical area studied were phylogenetically distant from other Iberian species of *Podarcis* (*lusitanicus, liolepis, guadarramae, virescens and*

muralis), the latter used as an outgroup. Therefore the results show a high degree of divergence, both of mtDNA and nDNA, at an intraspecific and interspecific level, suggesting a possible taxonomic reassessment of the genus *Podarcis* hispanicus, as the specimens sampled in the south-eastern area of the Iberian Peninsula, called Galera, are characterized by an obvious genetic diversity compared to the species of *Podarcis* already known. At the level of phylogenetic sistematics, therefore, they could constitute an autonomous taxon. As regards the future prospects, it is necessary to expand the number of populations under consideration, to expand the geographical sampling interval and to use individuals from other species in order to carry out a more detailed intespecific counter-type.

Riassunto

Questo lavoro di tesi è incentrato sull'analisi genetica di due geni mitocondriali (Regione Controllo e due frammenti di Citocromo b) e un gene nucleare (MC1R) per valutare la struttura della popolazione e la diversità genetica in una popolazione di *Podarcis*, localizzata al sud-est della penisola Iberica, in una località chiamata Galera. Sono stati messi in atto protocolli laboratoristici di estrazione di DNA con il metodo cloroformio/fenolo, PCR, corsa elettroforetica in agarosio e utilizzo del programma Gel Capture per visualizzare le bande di DNA, purificazione impiegando il kit INVITEK

MSB® Spin PCRapace, preparazione alla reazione di sequenziamento inserendo il Dye Terminator, precipitazione DNA purificato in etanolo e sequenziamento con il sequenziatore ABI 3130 dell'Applied Biosystem, per ottenere le sequenze nucleotidiche dei geni esaminati. In seguito, sono state impiegate procedure di analisi computazionale per la costruzione di alberi filogenetici (con il software Mega X) e per lo studio della diversità genetica (con il software DNAsp) della popolazione in esame e di altre specie di *Podarcis* già conosciute. Dalla cladistica basata su dati molecolari relativi agli esemplari di interesse si possono evidenziare le relazioni evoluzionistiche dei vari taxa, nonché il grado di parentela; in particolare i campioni provenienti dall'area geografica oggetto di studio risultano essere filogeneticamente distanti dalle altre specie Iberiche di *Podarcis* (*lusitanicus, liolepis, guadarramae, virescens e muralis*), quest'ultima utilizzata come outgroup. Pertanto i risultati ottenuti mostrano un elevato grado di divergenza, sia del mtDNA che del nDNA, a livello intraspecifico e interspecifico, suggerendo una possibile rivalutazione tassonomica del genere *Podarcis hispanicus*, in quanto gli esemplari campionati nell'area sud-orientale della penisola Iberica, denominata Galera, sono caratterizzati da un'evidente diversita' genetica rispetto le specie di *Podarcis* gia' note. A livello di sistematica filogenetica, quindi, potrebbero costituire un taxon autonomo. Per uquanto riguarda le

prospettive future è necessario ampliare il numero delle popolazioni in esame, espandere l'intervallo di campionamento geografico e utilizzare piu' individui di altre specie per poter effettuare un confronto interspecifico più dettagliato.

1. INTRODUCTION

Fig.1 : Podarcis hispanica (Galera) Lorc, junio 2009

1.1 – Mitochondrial and nuclear DNA

DNA is the depository molecule of genetic information. The most important function of DNA is therefore to transmit the hereditary characteristics from one individual to another. There are two different types of DNA: nuclear and mitochondrial.

Mitochondrial DNA (mtDNA) is a widely used tool for evolutionary studies, analysis of population structure, studies aimed at identifying species and resolving phylogeological problems or inherent in phyletic reconstructions of species and populations.

Also the nuclear DNA (nDNA) in the field of phylogenesis is useful to study the origin of some lizard's lineages, gene flow and genetic, morphological

and criptic diversification of a species, in order to carry out an accurate phylogeographic scheme. Both nuclear DNA and mitochondrial DNA are extremely resistant to the passage of time, even if the mitochondrial is more protected from oxidative damage. Therefore the phylogenomic analysis, using both mitochondrial and nuclear DNA, allows the resolution of complex philogeographic patterns typical of areas that constitute an evident biodiversity hotspot. [1,2,3,4]

1.1.2 - Structure

The mtDNA is much smaller and simpler than that contained in the cell nucleus. It is a double-stranded circular DNA molecule, resides inside the mitochondria, that is, power stations that carry out the cellular process of oxidative phosphorylation. The general structure of a mitochondrial DNA molecule recalls the general structure of nuclear DNA, in fact is analogoment to it is a biopolymer, consisting of two long nucleotide filaments.

Beyond the aforementioned analogies, mitochondrial DNA has some structural peculiarities, which distinguish it considerably from nuclear DNA. First of all, it is a circular molecule, while nuclear DNA is a long polymeric chain organized in chromosomes and complexed with histone proteins to form chromatin. Furthermore, mtDNA is a molecule containing a markedly lower

number of genes and the mitochondrial genetic code differs from the standard nuclear genetic code.

The nDNA is in diploid form inherited from two haploid parental gametes, one paternal and one maternal (According to the Mendelian inheritance), instead the mtDNA is haploid and us of purely maternal origin. The mutation frequency of nDNA and mtDNA is the same, but mtDNA repair less efficiently and the populational size is minor, thus the evolutionary rate is major. So the low fidelity of mtDNA polymerase, lack of proof-reading, and the apparent lack of mtDNA repair mechanisms have led to a higher rate of mutation in the mtGenome as compared to the nuclear genome. [5,6,7]

1.1.2 - Functions

From a functional point of view, mitochondrial DNA has some peculiar characteristics that clearly distinguish it from nuclear DNA. It is semi-independent, in the sense that it needs the intervention of some proteins synthesized starting from the nuclear DNA in order to properly carry out its functions. By contrast, nuclear DNA is completely autonomous and produces everything it needs to perform its tasks properly. mtDNA has a genetic code slightly different from that of nuclear DNA.This leads to a series of differences in the realization of proteins: if a certain sequence of nucleotides

in the nuclear DNA leads to the creation of a certain protein, the same sequence in the mtDNA leads to the formation of a slightly different protein. It has very few non-coding nucleotide sequences, that is, they do not produce any protein, tRNA or rRNA. The mitochondrial genome has 37 genes coding for two ribosomial RNA, 22 transport RNA and 13 proteins that are part of enzymatic complexes responsable for oxidative phosphorylation. In percentage terms, only 3% of mtDNA is not coding. In contrast, nDNA is coding for only 7%, so it contains a large number of non-coding nucleotide sequences.[8]

1.1.3 - Applications

Mitochondrial DNA has several properties that make it a highly employed and unique molecular marker:

- Contained in mitochondria and it is present in numerous copies in cells (100/1000 copies for cell)
- Circular structure then closed
- Generally matrilineal inheritance (transmission of uniparental genetic material, even if there are examples of paternal mtDNA that infiltrates into the embryo)
- Easily insulated
- It does not undergo recombination

- High mutation rate

Nuclear DNA's properties are:

- Localized in the nucleus of the eukaryotic cells and at most it can be found replicated in two copies
- Linear structure with open ends
- Genetic information inherited from both parents (Mendelian trasmission)
- Low mutation frequency

These peculiarities, high mutation and passage of genetic material only for maternal line with very low recombination, make mtDNA a powerful tool to trace matrilinearity and has been used to study many species up to generations of hundreds of years ago.

In summary, the various fields of application of lizard's mtDNA and nDNA are:

- Evolutionary species studies (molecular differences studied on the basis oh haplotypes, that is nucleotide portions inherited for many generations without being mutated)
- Phylogenetic reconstructions and historical biogeography (trace the evolutionary history of a set of species illustrated by phylogenetic trees). [9]
- Studies on microevolutionary process and genesis of intraspecific biodiversity hotspots

- Spatial segregation studies linked to different morphotypes
- Analysis of the distribution of haplotypes [10]
- Studies on speciation, morphological evolution and adaptation to extreme environment [11,14]
- Studies on chromatic polymorphism, food diet and behavioral aspects [12]
- Potential toxicity of a harmful compound on lizard's survival
- Studies on lizard's termoregulation [13]

1.2 – Study model : Podarcis hispanicus

Podarcis hispanicus Wagler, 1830 is a genus of small reptiles of the family of the Lacertidae of the suborder of the Sauri, commonly known as lizards.[15] For a long time it was considered a simple subgenus of Lacertidae, from which it is distinguished substantially by the ventral scales of a subrectangular shape rather than trapezoidal with oblique lateral margins. Even so the taxonomic determination has not concluded and *Podarcis hispanicus* is considered more than as a species, as a paraphyletic group and it would not be strange that in a short time some of the different forms that are recognized in Iberian Peninsula ''Lusitanica'', ''Virescens'', ''Liolepis'', ''Galera'' can give rise to another new species. Iberian Peninsula provide many of the

world's biodiversity hotspots because continously generate new species. These high speciation rates are facilitated by specific island and iberian peninsula characteristics that lead to a wealth of ecological, biogeographic and evolutionary processes. These include: natural fragmentation, long term isolation, high altitudes creating habitat heterogeneity, complex intra-island landscapes and island emergence due to volcanism and eustatic sea-level changes. The lizard genus *Podarcis* encompasses about 19 species and is widespread across the Mediterranean. *Podarcis hispanicus* is endemic to Iberian Peninsula and it exhibits a great intraspecific variability in morphological traits, especially colour pattern and melanism.[16]

1.2.1 – Description

The Iberian lizard varies greatly in size and color throughout its range.[17] The anatomical design of the Iberian lizard is clearly determined by its rupicolous habits. Effectively the main characteristics that draw attention when looking at a specimen close up are the robust body although of small size and easily collapsible, husiform design, extremely long fingers and tail that are focused on guaranteeing quick horizontal and vertical movements on walls and stones. They also allow to take refuge easily within small fissures and cracks or under a rock. Its head is not very large, rather flattened and little differentiated from the rest of the body. The snout is sharp and the eyes,

slightly prominent, have black circular pupils with yellowish iris. The presence of occipital scale is also characteristic. The body appears covered by a multitude of small granular scales, The legs are slender and have five long fingers, especially the posterior ones, endowed with equally long nails. The tail is also very long, and accounts for almost 2/3 of the total length of the body. The background color of the back is extremely variable even in specimens belonging to the same population. Although brown specimens abound, there are also others with a nice greenish coloration. It is frequent the presence of small spots of black color that sometimes form longitudinal bands on the sides. The throat is usually whitish or yellowish, free of spots and the belly is dirty white or grayish dotted with small black specks. The tail has the same tones as the rest of the body, except in newborns and juveniles, which are endowed with a spectacurarly colored appendage of green or turquoise blue. Cases of albino and melanic specimens have been documented. The males generally reach larger sizes, more robust head as well as longer legs and tail. The females have more muted tones and are notable for the presence of several lateral bands (alternately light and dark) on the sides. [18,19]

1.2.2- Geographic distribution area

Regarding the geographical distribution area, *P. hispanicus* is widespread in almost the entire Iberian Peninsula [21] and in the southern areas of France.

Its range extends to the Columbretes and Chafarinas Islands. The habitat of *P.hispanicus* is represented by rocky or rich in vegetation areas. It can also be found in stone walls or in cracks in country houses. Reported from sea level up to 3,481 meters high. The lizard *Podarcis hispanicus* is the protagonist in the setting of the Iberian Peninsula, as we can see in the figure below (Fig.2).

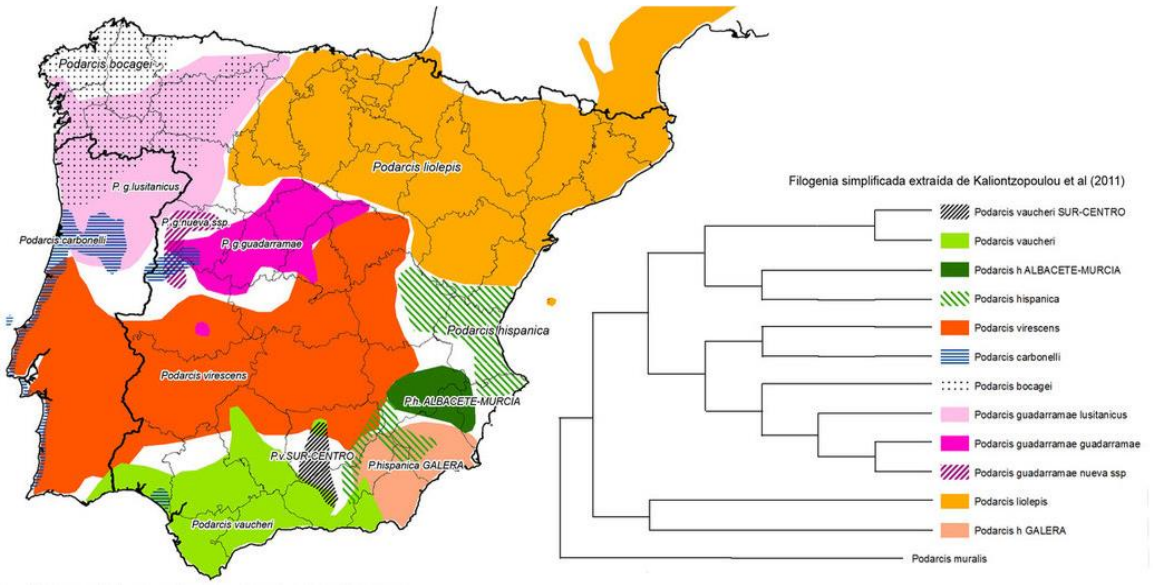

Fig.2 : The following map shows the geographical distribution of the different Podarcis lineages present in the Iberian peninsula, as well as a very simplified phylogeny in which the relationships between them are shown.

For simplicity it can be differentiated into 4 large groups depending on the geographical location [20]:

1. North-East group *(Podarcis liolepis)* (Fig.3)

It is the oldest group and is distributed mainly by the northeast. It is found in the communities of Catalonia, Aragòn, Castilla y Leon, Pais Vasco, Navarra and Comunidad Valenciana, although its limits are not clearly known.

Fig.3 : Podarcis liolepis. Female (above) and male (below). Penyagolosa (Catellòn). V.Sancho. 2006

2. South-East *group* *(Podarcis vaucheri-hispanic*a) (Fig.4-5)

It is distributed to the South and East of the peninsula. *P.vaucheri* occupies from the southern tip of Portugal to Almeria, although some populations of Granada have enough differences to be considered as separate form (perhaps a subspecies). Regarding *P.hispanica* the picture is still very confusing and some populations could show a mixture or introgression with *P.liolepis*. *P.Galera*, present in Murcia, Almeria and Granada, has not yet been

described as a new species and remains a form of *P.hispanica*. In any case there are enough genetic differences to be considered at least as a subspecie.

Fig.4: Podarcis vaucheri. Seville Sebas Gomez 2012

Fig.5: Podarcis hispanica. Santa Pols (Alicante) BPO Natura, 2011

3. Western group *(Podarcis carbonelli-virescens)*(Fig.6)

This group includes Podarcis carbonelli, formerly considere a subspecies of *P.bocagei*. It has a common ancestor with *P.virescens*, new species broken off from the old *P.hispanica*, which occupies Castilla-La Mancha, Extremadura and southern of Portugal.

Fig.6: Podarcis virescens Andujar (Jaen) Toni Alcocer, 2012

4. North-West group *(Podarcis bocagei-guadarramae)*

This group includes *P. bocagei*, which occupies the northwest of the Iberian peninsula and northern areas of the portugal, and *P.guadarramae*, a new species separated from the ancient *P.hispanica*.

Genetic studies have revealed that it is a complex that is still young from the evolutionary point of view, formed by species and subspecies (or morphotypes) very similar to each other and until a few years ago all gathered under the name of *Podarcis hispanicus*.

Fig.7 : P.guadarramae. Female (above) and male (below). Cional (Zamora) 2014

1.2.3 – Genus, species and subspecies : a complex path

The interrelations between the morphological variability and the phylogenesis of the species are still the subject of a wide debate among scholars, concerning the investigation of the evolutionary processes that affect the

differentiation characters. However the strong variability of morphological characters has made it very difficult to draw an exhaustive picture of the evolution of the genus Podarcis. Indeed, the attempts made to reconstruct the phylogenetic relationships within the taxon are very different from each other, making them still today unclear even within the Lacertidae family. [22]

1.3 – Objective

"The time will come I believe, thug I shall not live to see it, when we shall have very fairly true genealogical trees of each great Kingdom of nature" Darwin, (1857) Letter to Thomas Henry Huxley

This study aims to provide the most detailed analysis of genetic variation of Podarcis of the Galera's region, using both mitochondrial and nuclear genes. With this research work we intend to classify the genus Podarcis of the Galera region as a univocal species, respect to the other phylogenetically related species, through the molecular analysis carried out on different mitochondrial (RC, C450, C600) and nuclear (MC1R) markers, and subsequent allignment of the corresponding nucleotide sequences to evaluate the degree of genetic diversity. Purpose of the thesis:

- Evaluation of intraspecific genetic diversity among different populations of endemic lacertids of the Galera region and the city of Don Fadrique (Granada), in the south-east of Spain.
- Evaluation of interspecific genetic divergence of these populations and other *Podarcis* species distributed in the Iberian Peninsula. (sisters species)

Considering genetic variation could reveal the degree of connection between populations of Podarcis lyzard of the Galera and of the other Spanish places and generate hypotheses about their recent evolutionary history and lead to the recognition of this 'cryptic species', so far known as *Podarcis* of the Galera's region.

2. MATERIALS AND METHODS

The project focuses on the molecular analysis of eight samples (GALE6, GALE7, GALE8, GALE9, GALE10, GALE11, Df4, Df5) of Lacertids of the genus *Podarcis* located in the geographical area of Galera, in the south-east of the Iberian peninsula (Fig.8). The *Podarcis hispanicus* complex forms a very heterogeneous group with complicated phylogenetic relationships. Objective of the project is the identification of inter and intra population phylogenetic relationships, between the Podarcis samples examined coming from Galera and the already classified species. The laboratory analyzes were conducted in order to obtain the sequence of different mitochondrial and nuclear genes (RC, C450, C600, MC1R), fondamental for the construction of phylogenetic trees and the consequent distinction and classification of the species examined [24,25].

Fig.8: Sampling map of specimens studied and analysed.

2.1 – Molecular analysis

Following the sampling activity, the findings collected in the field using non-invasive methods, were useful for the implementation of laboratory procedures that involved the performance of different analysis methods:

- Phenol/Chloroform DNA extraction
- Amplification of target regions
- Agarose gel
- Amplified DNA purification
- Sequencing reaction
- Ethanol precipitation
- Bioinformatic analysis

2.1.1 – Extraction of DNA from tissue

A standard phenol-chloroform protocol was used for DNA extraction. The first phase of the extraction involves the lysis of the cells with suitable buffers containing detergents (SDS, Triton X-100) and with Proteinase K, in order to solubilize the membranes, complex proteins and extract nucleic acids. Proteinase K has the specific function of digesting proteins. It is a protease of

the saprophytic fungus Tritirachium album and is particularly suitable for digestions that occur quickly.

The DNA was extracted from the tissue (tail tips) of lyzards belonging to the Galera populations with manual method involving the use of Phenol/Chloroform-Proteinase K. Subsequently the DNA was evaluated quantitatively by Nanodrop.

As reagents have been used:

- Lysis Buffer
- SDS 10%
- Proteinase K (powder)
- NaCl 5 M
- Phenol
- Isoamyl chloroform (24:1)
- Absolute ethanol (-20°C)
- Te, (ph=7,5)

The preliminary steps of the extraction procedure consisted in cutting the lizard's tail, discarding the first piece to eliminate any contamination and inserting the pieces inside a 15 ml tube and add 1 ml of lysis buffer. Homogenization until the tissue was well broken and insertion in a

thermostatic bath overnight at 37°C. The consecutive step was carried out under a hood and consisted of the addition of 5 ml of phenol, centrifuge at 3.000 rpm for 10 min and supernatant recovery (3.5 ml aprox.). Addiotion of 4 ml of isoamyl chloroform and centrifuge at 3.000 rpm for 10 min and supernatant recovery, eliminating the pellet.

The final step included precipitation by adding 100% absolute ethanol, which is indispensable for concentrating DNA solutions and eliminating phenol and chloroform residues and storage in an overnight freezer. The following morning a centrifuge of 20 min was performed at 4.000 rpm liquid decantation and subsequent addition of 70% ethanol which allows the removal of most of the monovalent cations that can interfere in the activity of enzymes used in the subsequent molecular analyzes. Drying with open tubes at room temperature.

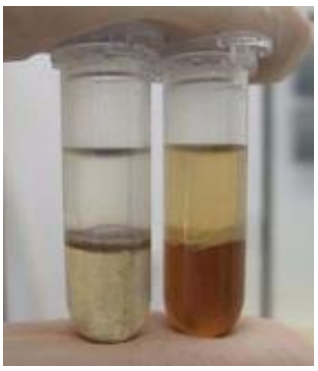

Fig.9: Stratification pellet (organic phase containing proteins), interphase (denaturated proteins) and surnatant (aqueous phase containing nucleic acids).

2.1.2 – Target gene amplification

For the amplification of the gene regions of interest (profile utilized shown in table 2) was performed the PCR protocol described in Table1. We have utilized different primers depending on the type of gene to be amplified (shown in table 3). PCR was conducted in a final volume of 25 µL.

Reagents	Volume
MgCl2	1,75 µL
T10x	2,5 µL
dNTPs	1 µL
F/R	0,5/0,5 µL
BioTaq	0,05 µL
H2O	variable

Table 1: Polimerase Chain Reaction reagents and corresponding volumes.

Genes	T°C annealing
RC	50°C
C450	49°C
C600	50°C
MC1R	50°c

Table 2: Amplification profile used for the different genes

Genes	Primers and Sequences
RC	F:L15022 5'-TACCCTTGCTCATAGCATAACTG-3' H:00292 5'-GTCTTGTTGACTGTAATTAACCGATA-3'
C450	F:L14724 5'-TGACTTGAARAACAYCGTTG-3' R:H15175 5'-CCCTCAGAATGATATTTGTCCTCA-3'
C600	F:L14724 5'-TGACTTGAARAACAYCGTTG-3' R:H15175 5'-CCCTCAGAATGATATTTGTCCTCA-3'
MC1R	**MC1R-PF** 5'-GGCNGCCATYGTCAANAACCGGAACC-3' **MC1R-PR** 5'-CTCCGRAAGGCRTAAATNATGGGGTCCAC-3'

Table 3: List of primers used depending on the gene to be amplified

Amplification conditions involved an initial denaturation step of 2 min at 94 °C; 35 cycles of 10 s at 95 °C, 20 s at 50 °C, 90 s at 72 °C, and a final extension step of 7 min at 72 °C.

2.1.3 – Agarose gel

The gel was prepared by heating 0,5g of agarose dissolved in 50 mL of TAE buffer solution and subsequent addition of ethidium bromide. Waiting time

for gel solidification of 30 min and observation of the electrophoretic run result through the Gel Capture program.

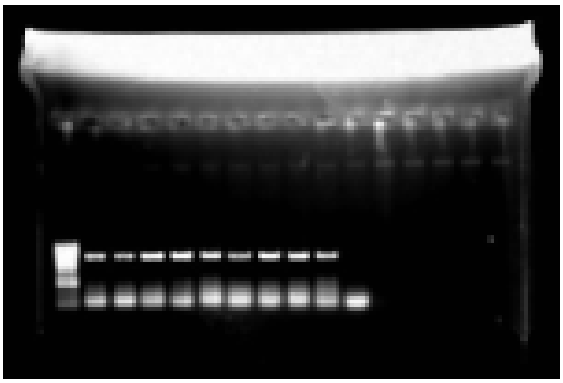

Fig.10 : Exemplary image of the electrophoretic migration of the MC1R gene on agarose gel.

2.1.4 – Gene purification

Following the amplification reaction, the PCR products were purified from any primer dimers, unincormporated dNTPs, supercoated DNA and protein residues using the INVITEK MSB® Spin PCRapace kit. The MSB Spin PCRapace is the fastest system for DNA fragment purification. Use it for DNA product cleanup from amplification reaction mixtures. Ready in only 7 minuts. Two step format; only binding and elution buffers. Purification ultra fast of Nucleid acid with spin column technology. + no washing needed.

At the end of the purification process the DNA was quantified with Nanodrop nanoVue Plus for a quantitative evaluation. The spectrophotometric analysis allows to determine the quantity of ultraviolet light absorbed by the sample under examination by photodetector and consequently to evaluate the concentration of nucleic acid (The Lambert-Beer Law: quantification of the light absorbed by a substance in solution is directly proportional to its concentration in the solution). The degree of purity in instead estimated OD260/OD280. Optimal purity around 1,8. High values reveal an excessive presence of RNA. On the other hand, if protein contamination or phenol residues are present, the A260/A280 will have a significantly lower value.

2.1.5 – Sequencing reaction

As a preliminary step to the sequencing reaction was prepared a mix containing primers conjugated to the Dye Terminator for fluorescent marking in the next sequencing procedure.

PCR sequencing mixture:

- Amplified DNA (1 μL for [20]DNA)
- Primer F/R to [2] in two separate Mixes
- Buffer solution

- DNA polymerase
- Nucleotides
- Terminator nucleotides

-	Reagents	-	Volume
-	**T5x**	-	1,2µL
-	**Primers**	-	1 µL
	Pre-Mix	-	0,6 µL
-	**H20**	-	variable

Table4: component Mix of the Terminator reaction.

2.1.6 – Ethanol precipitation

To eliminate any contaminants that might interfere with the subsequent analysis, I performed ethanol precipitation. The precipitation of deoxyribonucleic acid was carried out by inserting 52 µL of Mix containing 50 µl of cold EtOH 100% + 2 µl of NaAC (Na-Acetate pH 5.6) in the tubes containing purified DNA. Shake well, leave the samples 15 minutes at T room. Centrifuge for 30 minutes at maximum speed, supernatant removal with particular attention not to touch the invisible pellet deposited on the wall of the tube (Remove alcohol without removing the Pellet). After centrifugation the supernatant solution is removed, leaving a pellet of crude DNA. Wash with 100 µl of 70% cold EtOH, shake and leave at room temperature for 10 minutes. This removes some of the salts present in the leftover supernatant and bound to DNA pellet making the final DNA cleaner. Later centrifufate at max rpm for 10 minutes. Drying under paper overnight (leave to air-dry).

2.1.7 – Sequencing with ABI 3130 Sequencer

To sequencing these gene in question, I took my samples to another lab where the sequence was made. At the end of the sequencing process, they sent me the results in the form of a elletrophoferogram (base calling) and I corrected any reading errors. Sequencing of both strands of PCR products was carried out on an automated ABI 3130 Sequencer (Applied Biosystems, Foster City, CA, USA) using a BigDye® Terminator v. 3.1 (Fig.11).

Fig.11: Applied Biosystems® 3130 Genetic Analyzer is the latest generation of 4-capillary electrophoresis instruments for the low to medium throughput laboratories. The system offers industry-leading performance, plus sophisticated automation capabilities allowing you to save time, reduce costs and increase productivity.

Sequencing is performer on ABI 3130 Automatic Genetic Analyzer Sequencer (Applied Biosystems) equipe with 4 capillaries and is based on a

capillary electrophoresis process fully automated. Sequence reactions are carried out with Big chemistry Dye Terminator in a “cycle sequencing” PCR reaction in which during the extension of the amplified product the dideoxynucleotiditrifosfati ddNTPs are incorporated (terminators) cojugated to 4 different fluorescent molecole. The detection takes place through laser-induced fluorescence in 4 spectral channels. After the sequence reaction we proceed with the run electrophoretic. Electrophoresis takes place on the capillary: the DNA is electrokinetically loaded into the electrophoretic support, exploiting the negative charge of the DNA and the application of a difference of potential at the ends of the support. The migration of the fragments is follone by detecting the emissions in fluorescence at different wavelenghts of different fluorochromes after the excitation caused by laser. The emissions are collected and analyzed by a photosensitive camere CCD (charge coupled device), which processes the different fluorescence signals with high sensitività. The sequence of DNA bands marked by the four fluorochromes is displayed in a graph called electropherogram, characterized by a succession of peaks of four different colors, which correspond to the fluorescent emissions of different fluorochromes, as the various fragments of different lenghts. Nucleotide reaches, during the electrophotetic run, the position of the detector. The raw data obtained from the sequenze are

assembled, edited, analyzed by means of dedicated software (Sequencing Analysis v2.0) in order to obtain the nucleotide sequence.

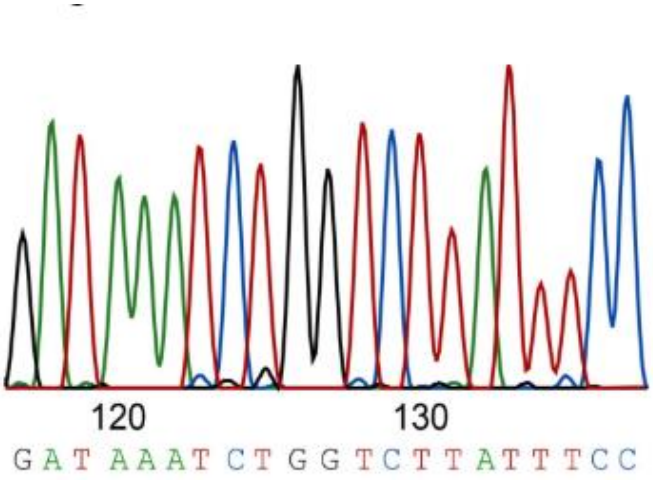

Fig.12: Electroferogram, example of a portion of a sequence of MC1R gene obtained with Genetic Analyzer Sequencer and elaborated with Sequencing Analysis v2.0 software.

2.2 – Bioinformatic analysis

Regarding the bioinformatic profile, three different programs were used depending on the analysis to be performed:

- CODON CODE ALIGNER to clean nucleotide sequences and to align the Forward and Reverse sequences
- BIOEDIT v7.2.5 for the alignments of the sequences of different populations of the same species and for the alignments with sequences of other species

- MEGA – X for the graphic representation of phylogenetic trees
- DNAsp v6. for the study of genetic diversity

2.2.1 – Codon Code Aligner

It is the CodonCode Aligner is an easy-to-use program for base calling, sequence assembly, contig editing, and mutation detection. Features include chromatogram editing, end clipping, and vector trimming, sequence assembly and contig editing, aligning cDNA against genomic templates, sequence alignment and editing, alignment of contigs to each other with ClustalW, MUSCLE, or built-in algorithms, mutation detection, including detection of heterozygous single-nucleotide polymorphism, analysis of heterozygous insertions and deletions [26].

2.2.2 – BioEdit v7.2.5

BioEdit is a biological sequence alignment editor and sequence analysis program which offers a variety of useful features. We have used this program to align all the gene sequences, already cleaned and aligned in pairs (F and R) with CodonCode Aligner, trough ClustalW multialignment and then create a consensus nucleotide sequence for each sequence in which both forward and reverse are present. Subsequently the aligned sequences of the genes C450,

C600, RC, MC1R were inserted and re-aligned with the same gene sequences of other podarcis species [27].

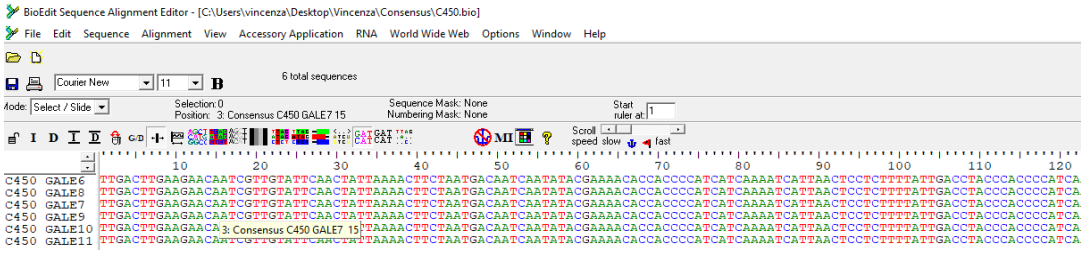

Fig.13: Consensus sequences of the C450 mitocondrial gene

2.2.3 – Mega X

At this point we have used computational phylogenetics (software Mega X) for the construction of phylogenetic trees , in order to have en evolutionary view of the species under examination. Over the past decade, genome sequencing has become an efficient and potent means to investigate a broad array of biological systems, from large-scale studies of biological diversity to tracking the evolution. The numerous steps required to glean interpretable and actionable results from raw sequence data invariably require comparative analysis of molecular sequences to discover functional and adaptive genome differences. Molecular Evolutionary Genetics Analysis (Mega) software

provides tools to conduct such analyses. Mega includes a large repertoire of programs for assembling sequence alignments, inferring evolutionary trees, estimating genetic distances and diversities, inferring ancestral sequences, computing timetrees, and testing selection. The tree was designed using the Neighbor joining and boostraping procedure, to assess the correctness of each branch. [31, 32]

- Neighbor Joining: Software for phylogenetic reconstruction based on evolutionary distance (requires in fact a distance matrix). Performs a clustering for the creation of phylogenetic trees (distance calculation and clustering method). It requires knowledge of the distance of each taxa pair and takes as input a distance matrix that specifies the distance between each taxa pair. The algorithm starts with a completely unresolved tree, with star topology, and proceeds with the next steps until all the lengths of the branches are known. Method for the reconstructing phylogenetic trees from evolutionary distance data. The principle of this method is to find pairs of operational taxonomic units that minimize the total branch lenght at each stage of clustering starting with a starlike tree. The branch lenghts as well as the topology of a parsimonious tree can quickly be obtained by using this method. NJ principle: the program looks for a pair of known X and Y such that the XY distance is the smallest possible, thus creating a new node in the

tree that joins the two closest nodes, calculates the distance of each of the nodes of the pair from their ancestral node and all the knots outside of this pair from their ancestral node. You start the algorithm again by replacing the neighboring taxa pair and using the distances calculated in the previous step. This process is repeated until all nodes are fixed.

- Bootstrapping: method to determine the significante of a result resulting from a complex analysis. It is a random resampling of sequences to which are applied to real sequences. In this way we have two values, one effective that of the sequences analyzed and one that represents the value i would obtain with a random sample. From their comparison it is possible to obtain an estimate of the truthfulness of the result. In general greater is the number of bootstrapping operation, greater is the reliability of values. Higher is the percentage of bootstrapping results that match with the true result, greater is the precision of the estimate of the true value.

How distance method were used: No. of differences, p-distance and Jukes-Cantor Model.

- No. of differences: this distance is the number of sites in which twoo sequences compared differ.
- p-distance: this distance is the proportion of nucleotide sites at which two sequences being compared are different. It is obtained by dividing

the number of nucleotide differences by the total number of nucleotides compared.

- Jukes and Cantor model: the rate of nucleotide substitution is the same for all pairs of the four nucleotides.As well, it assume that the different sites are independent.

For each tree the muralis species has been inserted as an outgroup (tree root). Were obtained a total of 6 phylogenetic trees, three using mtDNA(C450,RC,C600) and three with the nDNA (MC1R) [28].

2.2.4 – DNAsp v6

To obtain estimates of intraspecific and interspecific biodiversity was used the DNAsp software v6 (DNA Sequencing Polymorphism Analysis for Large Data Sets), which analyze large data sets such as those generated by high throughput sequencing technologies. DnaSP is a software package that performs extensive population genetics analyses using DNA sequence data from a multiple sequence aligned data, so DNAsp allowing DNA polymorphism analyses on multiple data files and on large datasets. Altogether, the present version of DnaSP has the appropriate features for exhaustive exploratory analyses using high-throughput DNA polymorphism data. DnaSP estimates several measures of molecular evolutionary due to

DNA sequence variation within and between populations.The observed parameters are: number of variable sites, haplotype diversity and nucleotide diversity. To assess nuclear diversity, however, haplotype phases for MC1R were resolved for heterozigotic individuals using DNAsp software which implements an algorithm from the program phase. The analysis of DNA polymorphisms is a powerful approach to understand the evolutionary process and to establish the functional significance of particular genomic regions [29,30].

3. RESULTS

3.1 Phylogenetics trees

The following cladograms (Fig.14-15) highlight the kinship relations between the lizard's populations analyzed coming from Galera (samples GALE6, GALE9, GALE10, GALE11) and other Iberian Podarcis species (lusitanicus, guadarramae, virescens, liolepis, muralis). The other samples analysed were not used in the construction of the phylogenetic tree, although repeated attempts at PCR failed to amplify the template of interest. Only one phylogenetic mitochondrial tree (p-distance method) and on enuclear tree (no of differences method) have been reported, due to the high similarità of the results. As shown in the respective trees it can be noted that the populations studied represent a monophyletic group with the other populations coming from the same region (ph21, AB8, ph22, AB9, Ro1, Ro2, GALE3, GALE5, GALE1, Df1). More in detail, the populations GALE6 and GALE11 are genetically closer than the GALE9 and GALE10, which in turn are phylogenically closer to the remaining lacertid population of Galera. This taxon visibly differs from other *Podarcis* species, therefore the interspecific diversity is remarkable. While the clade *P.virescens* is more closely related to the species lusitanicus and guadarramae. Another markedly differentiated

taxon is *liolepis* [34, 36, 37]. The phylogenetic model observed within the species coincides with the actual taxonomic division into subspecies.

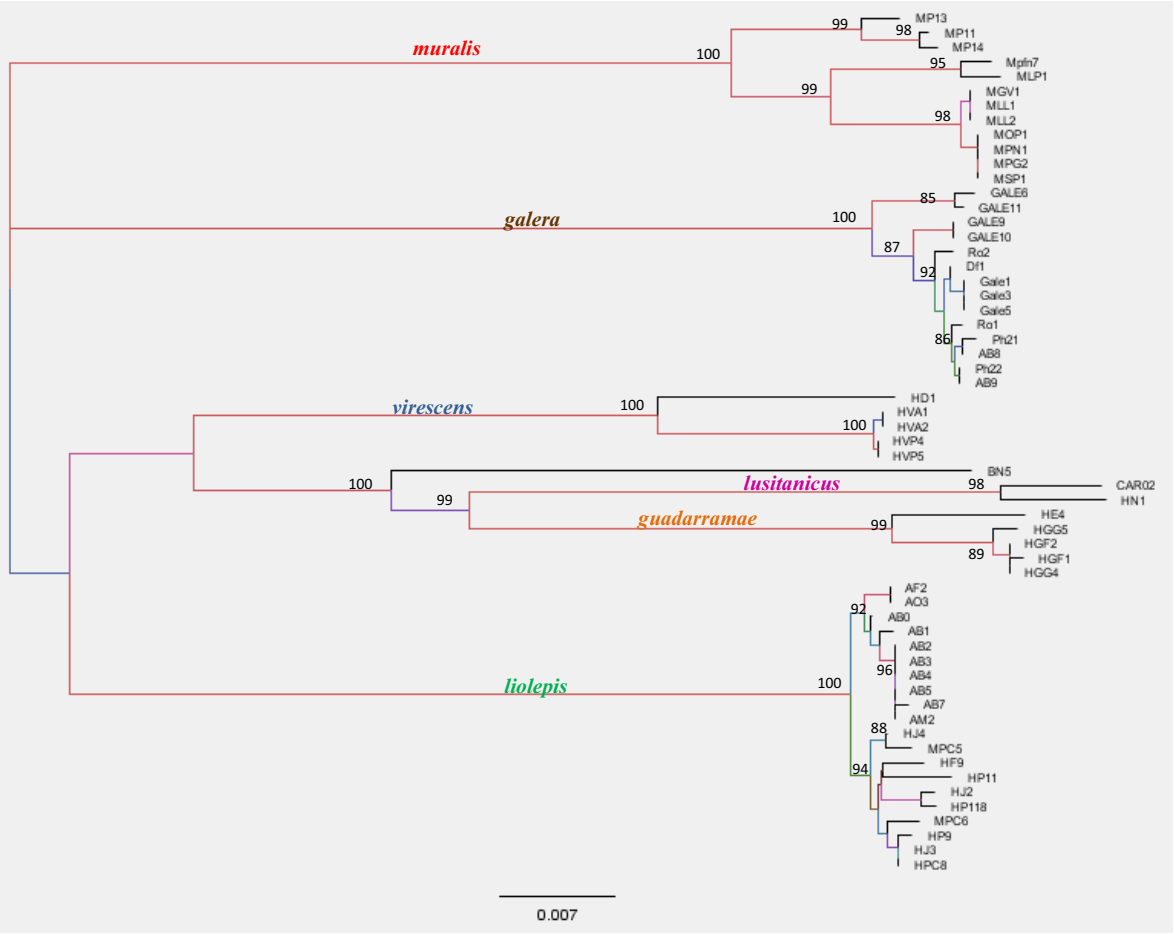

Fig.14 : phylogenetic tree built using the sequences of the mitochondrial genes examined (C450,C600,RC) using the Mega software. To visualize the tree we used the software FigTree. Statistical method Neighbor-Joining, p-distance. As an outgroup (monophyletic group taken as a reference) was used *P.muralis.*

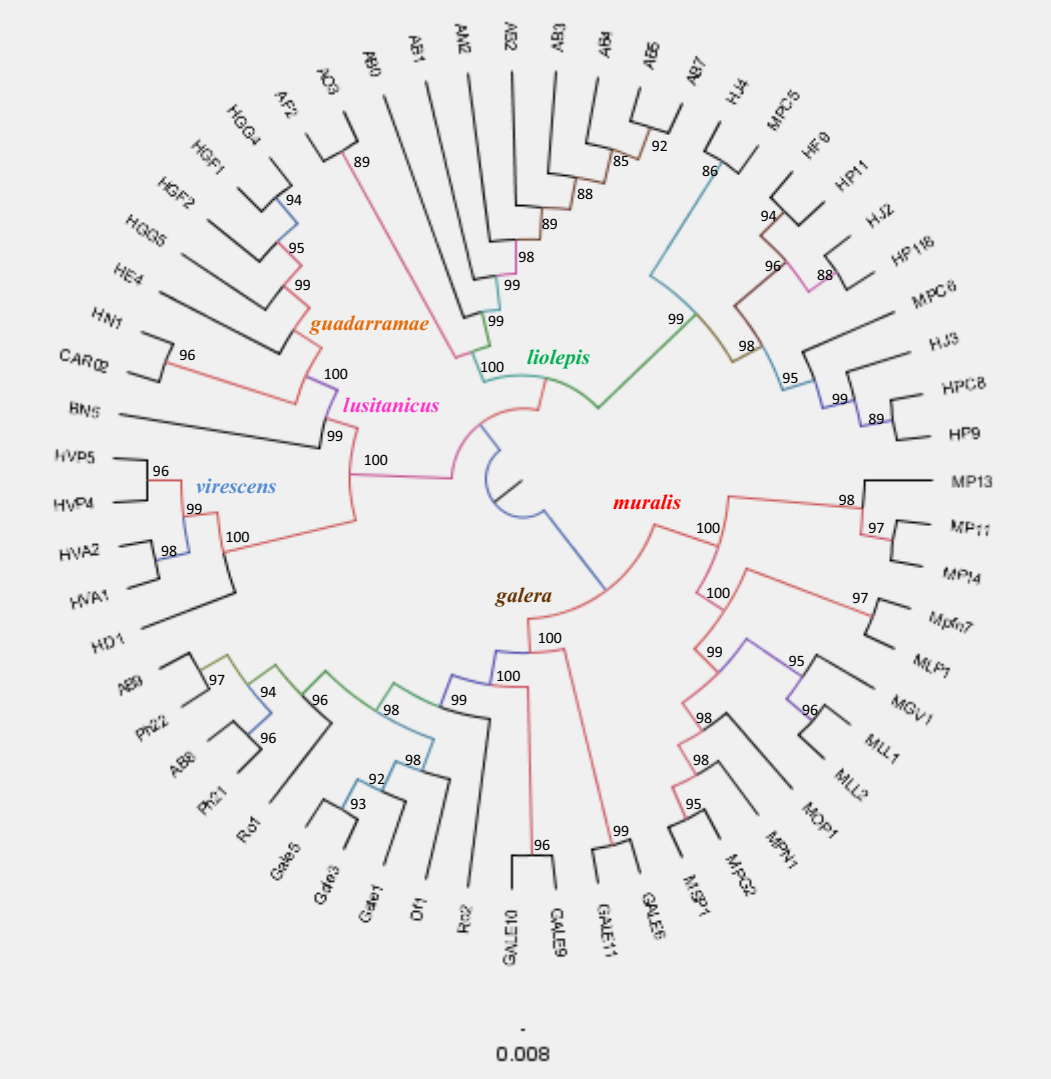

Fig.15 : phylogenetic tree built using the sequences of the nuclear gene MC1R, drawn with Mega using the statistical method Neighbor-Joining, Jukes-Cantor. As an outgroup was used *P.muralis*.

3.2 Genetic diversity

Intraspecific genetic diversity (software DNAsp) of differents populations of the same lizard's species (Table5-6). The interspecific genetic diversity was calculated observing the genetic diversity value of *Podarcis* of *Galera*'s region compared to other *Podarcis* species. [33, 35]

Species	N=n° of sequences	S=n° of variable sites	h=number of haplotypes	Hd=haplotype diversity	K=average number of nucleotide differences	Pi=nucleotide diversity (for site)
Galera	14	21	10	0,945±0,045	6,12088	0,0051±0,0013
Lusitanicus	2	15	2	1,000 ±0,500	15,00000	0,01147±0,00573
Liolepis	20	28	14	0,937±0,043	6,04211	0,00465±0,00043
Muralis	12	54	7	0,864±0,079	19,40909	0,01492±0,00293
Guadarramae	5	22	4	0,9005±0,161	9,00000	0,00688±0,00305
Virescens	5	36	3	0,800±0,164	14,80000	0,01131±0,00614

Table 5 : Diversity parameters for the concatenated mtDNA dataset (1199 bp)- in different lineages of Iberian *Podarcis*.

Species	N=number of sequences	S=number of variable sites	H=number of haplotypes	Hd=haplotype diversity	K=average number of nucleotide differences	Pi=nucleotide diversity (for site)
Galera	28	3	3	0,540±0,086	0,85714	0,00195±0,00044
Lusitanicus	4	3	2	0,667±0,204	2,00000	0,00407±0,00124
Liolepis	40	4	5	0,636±0,051	0,80000	0,00163±0,00022
Muralis	24	7	7	0,826±0,050	2,36232	0,00480±0,00072
Guadarramae	10	1	2	0,533±0,095	0,53333	0,00108±0,00019
Virescens	10	3	2	0,356±0,159	1,06667	0,00217±0,00097

Table 6 : Diversity parameters for the nDNA dataset (MC1R) in different lineages of Iberian *Podarcis*.

4. CONCLUSIONS

Recently, several works have focused on the lacertid lizards of the genus Podarcis, revealing cases of hidden diversity and offering evidence that suggests the revision of the extant taxonomical arrangements within the genus. In summary, I analyzed four mtDNA (RC, C450, C600) and one nuclear marker (MC1R) in populations of *Podarcis* sampled in the southeast of Spain and I found a hight genetic variability both intra- and inter-population as can be seen from the table above, corresponding to deep evolutionary divergence that characterized the genus *Podarcis* endemic to the Iberian peninsula.This study will motivate and provide evidence of taxonomic revaluation of these populations located in the Galera region compared to other Iberian forms, due to a correlation between genetic diversity of *Podarcis* of the Galera's regions populations and paleogeography of the regions they inhabit is discussed [38,39, 40]. In the future, by expanding the populations of lizards examined, expanding the geographical sampling range and adding new Iberian species for interspecific comparison, it will be possible to classify individuals from the region Galera as an autonomous taxonomic category, given the high degree of genetic diversity.

5. Acknowledgment

Finally the fateful day has arrived: writing these phrases of thanks is the final touch of my thesis. it was a time of deep learning, not only at the scientific level, but above all personal.

A heartfelt thank goes to my university which gave me the opportunity to do my thesis work abroad, allowing me to live one of the most significant experiences of my life.

A special thank goes to my supervisors, Oliana Carnevali of the Polytechnic University of Marche and Cori Ramon of the University of the Balearic Islands; for their precious advices and teachings without which I could not have completed this brilliant achievement. I’m extremely grateful to my copresenter Marta Bassitta for her help with laboratory work.

Another great thank goes to my mother and father, who with their sweet and tireless support, both moral and economic, have given me the opportunity to embark on this university path and have been a guide and stimulus for my professional training and my personal growth. Thanks also to my brother who is a great pizza maker and has always encouraged me in the studio because jokingly he always told me that I am "the smart sister".

I would also like to thank my grandparents: Gioconda who for me is an ajoint mother and delights me with her pearls of wisdom; Adelina and Antonio who unfortunately will not be able to rejoice with me but protect me and watch me from up there and have been my greatest strength to always give the best and never give up; and my grandfather Vincenzo that I never knew but I know he was an exceptional person so I guess he's proud of me right now. Another member of my family that I would like to thank is my four-legged friend, my bodyguard, my good giant, 50 kg of pure sweetness, my favorite stress reliever, who is always there waiting for me wailing and ready to play.

A warm thank also goes to my colleagues working in the Termoli's chemistry industry "Momentive", in particular to Dr. Cristina Domitriu and Gabriele Casolino who gave me the wonderful opportunity to undertake my first experience working in the laboratory world, and the analysts who welcomed me as a daughter and taught me, between coffee and a chat, the tricks of the trade.

I would also like to thank all my friends and classmates, as well as adventures, who have always supported me for good and bad, from moments of despondency and melancholy to moments of joy and carefreeness.

The greatest thanks, however, go to a special person who has been close to me throughout the period of writing this thesis, supporting me, raising my mood in the darkest moments and enduring me in moments of madness. I will always cherish this unforgettable experience in my heart.

6. BIBLIOGRAPHY

1. Multigenic phylogeny, phylogeny and population structure of the species group *Podarcis* cretensis in the southern Balkans.Spilani L, Bougiouri K, Antoniou A, Psonis N, Poursanidis D, Lymberakis P, Poulakakis N.Mol Phylogenet Evol . Set 2019.

2. Senczuk G, Havenstein K, Milana V, Ripa C, De Simone E, Tiedemann R, Castiglia R.Rep . Sci . Spotlight on the islands: the origin and diversification of a new lineage of the Italian mural lizard *Podarcis siculus* in the Western Pontine islands. 11 ottobre 2018.

3. Psonis N, Antoniou A, Karameta E, Leaché AD, Kotsakiozi P, Darriba D, Kozlov A, Stamatakis A, Poursanidis D, Kukushkin O, Jablonski D, Crnobrnja-Isailović J, Gherghel I, Lymberakis P, Poulakakis N. Resolve complex phylogeographic patterns in the Balkan peninsula by using closely related wall lizard species as a model species. 21 Marzo 2018.

4. Psonis N, Antoniou A, Kukushkin O, Jablonski D, Petrov B, Crnobrnja-Isailović J, Sotiropoulos K, Gherghel I, Lymberakis P, Poulakakis. Hidden diversity in the subgroup of species *Podarcis tauricus* in the light of multilocus phylogeny and species demarcation. 2017 gennaio.

5. Watson J.D. and Crick F.H.C. A Structure for Deoxyribose Nucleic Acid. Nature febbraio 2007.

6. Emmanuel D. Ladoukakise Eleftherios Zouros. Evolution and inheritance of animal mitochondrial DNA: rules and exceptions. Dicembre, 2017.

7. Brown WM, George M Jr., Wilson AC, Rapid evolution of mitochondrial DNA. 1971

8. Sung Ryul Lee, Jin Han. Mitochondrial nucleus: shield and switch of the mitochondrial genome. Giugno 2017

9. DJ Harris, C Pinho, MA Carretero, C Corti, W Bohme . Determination of the genetic diversity whithin the insular lizard Podarcis tiliguerta using mtDNA sequence data, with a reassessment of the phylogeny of Podarcis. 2005

10. Terrasa B., Brown R.P, Castro J.A., Giménez P., Hoskison P., Pérez-Mellado V., Hernández-Estevez JA, Picornell A., Ramon M.M.. Genetic analysis suggests that the Balearic lizard and *P.pityusensis* have evolved separately for five milion years. (2005).

11. Harris DJ, et al.Mol. Rapid speciation, morphological evolution and adaptation to estreme environments in South African sand lizards as revealed by mitochondrial gene sequences. 2004

12. Buades JM[1], Rodríguez V, Terrasa B, Pérez-Mellado V, Brown RP, Castro JA, Picornell A, Ramon MM. Variability of the mc1r gene in melanic and non-melanic Podarcis lilfordi and Podarcis pityusensis from the Balearic archipelago. 2013

13. Effective thermoregulation in a newly established population of Podarcis siculus in Greece: a possible advantage for a successful invader Grigoris Kapsalas1, Ioanna Gavriilidi1 , Chloe Adamopoulou2 ,JohannesFoufopoulos Panayiotis Pafilis1, 2016.

14. Aleksandar Urošević Katarina Ljubisavljević Dušan Jelić AnaIvanović. Variation in the cranium shape of wall lizards (Podarcis spp.): effects of phylogenetic constraints, allometric constraints and ecology. 2012

15. Jose Antonio Mateo Miras, Marc Cheylan, M. Saïd Nouira, Ulrich Joger, Paulo Sá-Sousa, Valentin Pérez-Mellado, Iñigo Martínez-Solano. Iberian Wall Lizard *''Podarcis hispanicus''*, 2009

16. Marco Oliverio*, Marco A. Bologna and Paolo Mariottini. Molecular biogeography of the Mediterranean lizards Podarcis Wagler, 1830 and Teira Gray, 1838 (Reptilia, Lacertidae). 2016

17. Jesús Ortega Giménez. Determinant of phenotypic variation in the Iberian wall lizard species complex (*Podarcis hispanicus)*. 2015

18. Eduardo Crespo, Luís Vicente and Paulo Sá-Sousa. Morphological variability of Podarcis hispanica (Sauria: Lacertidae) in Portugal. 2002

19. DJ HnnnISI, P Sá-Sousa - Herpetological Journal. Species distinction and relationships of the western Iberian Podarcis lizards (Reptilia, Lacertidae) based on morphology and mitochondrial DNA sequences. 2001

20. Philippe Geniez1 , Alexandre Cluchier , Alexandre Cluchier1, Paulo Sá-Sousa, Claude P. Guillaume1 & Pierre-André Crochet Pierre-André Crochet,. Systematics of the Podarcis hispanicus-complex (Sauria, Lacertidae): Redefinition, morphology and distribution of the nominotypical taxon. 2007

21. José Maria Lopez Sanchez. Map of *Podarcis hispanica's* distribution. 2015.

22. D.James Harris, Paulo Sa-Sousa. Molecular phylogenetics of Iberian wall lizard: is *Podarcis hispanica* a species complex? 2002.

23. Philippe Geniez, Alexander Cluchier, Paulo Sa Sousa, Pierre Andre Uncinetto. Systematics of the Podarcis hispanicus-complex (Sauria,Lacertidae): Redefinition, morphology and distribution of the nominotypical taxon. Aprile 2007.

24. V.Rodriguez, J.M. Buades, R.P. Brown, B.Terrasa, Perez-Mellado, C.Corti, M.Delaugerre, A.Picornell, and M.M. Ramon. Evolutionary history of *Podarcis tiliguerta* on Corsica and Sardinia, 2017.

25. Mol Biol Evol. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. 2006.

26. Genetics in Medicine. Codon Code Aligner version 1.2 Released. 2004

27. Tom Hall, Ibis Biosciences, Carlsband. BioEdit: An important software for molecular biology. 2007

28. Sudhir Khumar, Glen Stecher, Michael Li, Christina Knyar, Koichiro Tamura. Mega X: Molecular Evolutionary Genetics Analysis across computing platforms. 2018

29. P.Librado and J.Rozas. DNAsp a software for comprehensive analysis of DNA polimorphysm data. 2009

30. Rozas J, Ferrer-Mata A, Sanchez-Del Barro JC, Sancher-GraciaA, Librado P, Guirao-Rico S. DNAsp 6: DNA sequence polimorphysm analysis of large data sets. 2017

31. A general species delimitation method with applications to phylogenetic placements. Bioinformatics. Zhang J, Kapli P, Pavlidis P, Stamatakis A. 2013.

32. A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. Iq-tree, 2015.

33. Wiens JJ. Species delimitation: new approaches for discovering diversity. 2007.

34. Reexamination of the Iberian and North African Podarcis (Squamata: Lacertidae) phylogeny based on increased mitochondrial DNA sequencing. Pinho C, Ferrand N, Harris DJ. 2006.

35. Molecular phylogeny and biogeography of the wall-lizard *Podarcis erhardii* (Squamata: Lacertidae). P.Lymberakis A.Antoniou D.Chalkia E.Zouros M.Mylonas E.Valakos,2003.

36. Phylogenetic relationships and biogeography of *Podarcis* species from the Balkan Peninsula, by bayesian and maximum likelihood analyses of

mitochondrial DNA sequences. P.Lymberakis E.Valakos E.Zouros[c]M.Mylonas, 2005.

37. A. Bellati, D.Pellitter Rosa, R.Sacchi, A.Nistri, A.Galimberti, M.Casiraghi, M. Fasola, P. Galeotti. Morphological variability and molecular discentation: a study of Podarcis muralis of the Tuscan Archipelago.2011.

38. Martina Podnar, Werner Mayer, Nikola Tvrtknovic. Mitochondrial phylogeography of the Dalmatian wall lizard, *Podarcis melisellensis* (Lacertidae). 2004

39. P. Lymberakis , N. Poulakakis , A. Kaliontzopoulou , E. Valakos & . Two new species of *Podarcis* (Squamata; Lacertidae) from Greece. 2010

40. P. Bloor, V. Rodríguez, B. Terrasa, R. P. Brown, V. Pérez-Mellado, J. A. Castro, A. Picornell & M. M. Ramon. Polymorphic microsatellite loci for the Balearic Island Lizard *Podarcis lilfordi* (Squamata: Lacertidae). 2010