



UNIVERSITAT DE
BARCELONA

**Population genetics of the endemic lizard
Podarcis lilfordi (Günther, 1874)
from the Balearic Islands**

Katherin E. A. Otálora Acevedo



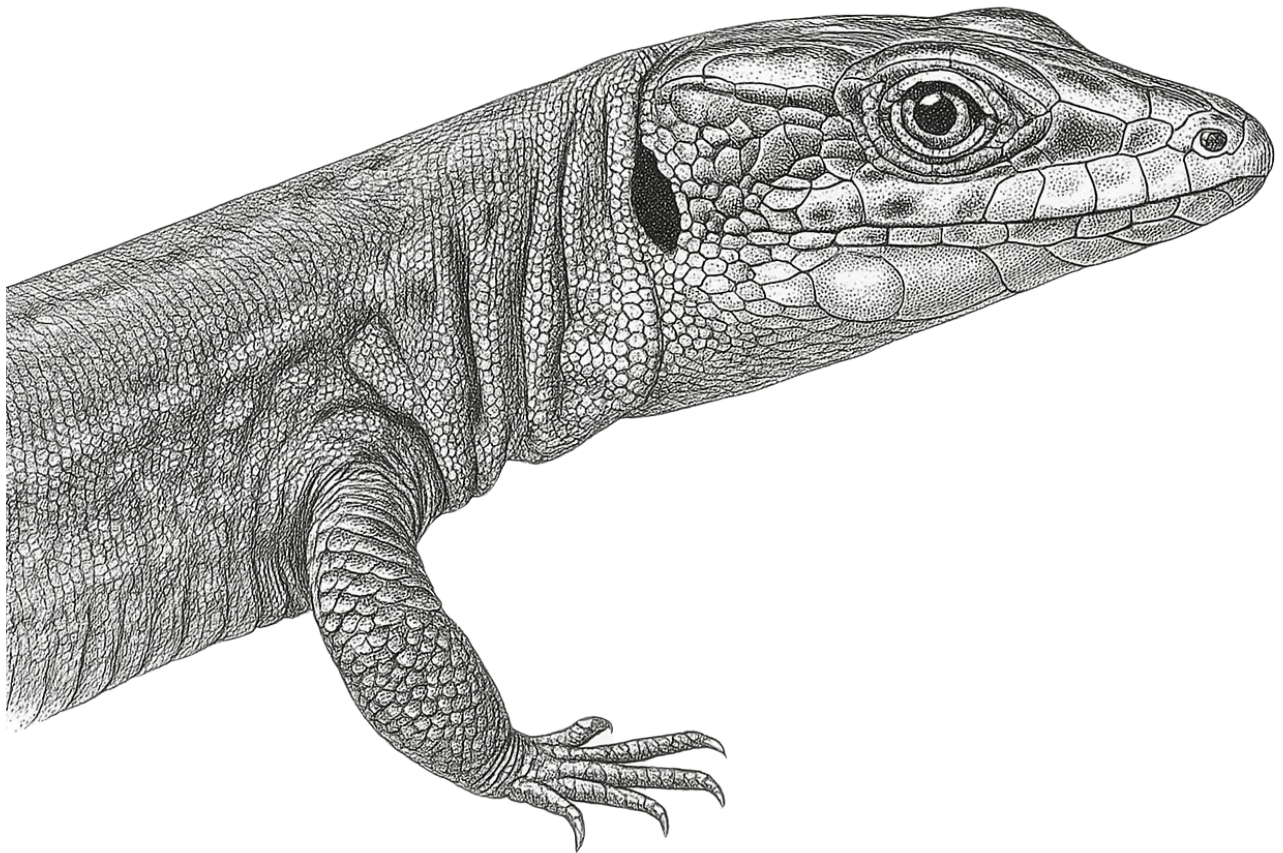
Aquesta tesi doctoral està subjecta a la llicència **Reconeixement 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution 4.0. Spain License.**

POPULATION GENETICS OF THE ENDEMIC LIZARD *Podarcis lilfordi* (Günther, 1874) FROM THE BALEARIC ISLANDS

Katherin E. A. Otálora Acevedo
Thesis 2025





UNIVERSITAT_{DE} BARCELONA

POPULATION GENETICS OF THE ENDEMIC LIZARD
Podarcis lilfordi (Günther, 1874) FROM THE BALEARIC
ISLANDS

GENÉTICA POBLACIONAL DE LA LAGARTIJA
ENDÉMICA *Podarcis lilfordi* (Günther, 1874) EN LAS
ISLAS BALEARES

GENÈTICA POBLACIONAL DE LA SARGANTANA
ENDÈMICA *Podarcis lilfordi* (Günther, 1874) A LES ILLES
BALEARS

Otálora, K. Population genetics of the endemic lizard Podarcis lilfordi (Günther, 1874) from the Balearic Islands. Doctoral thesis. Universitat de Barcelona, Barcelona, Spain.

All diagrams and schemes by Katherin Eliana Alejandra Otálora Acevedo.

Cover photographs generated by ChatGPT (OpenAI, 2023) on chatgpt.com, based on web-available photos.



UNIVERSITAT_{DE}
BARCELONA

Programa de Doctorat en Biodiversitat (UB)
Departament de Biologia Evolutiva, Ecologia i Ciències
Ambientals de la Universitat de Barcelona
(BEECA)

POPULATION GENETICS OF THE ENDEMIC LIZARD
Podarcis lilfordi (Günther, 1874) FROM THE BALEARIC
ISLANDS

GENÉTICA POBLACIONAL DE LA LAGARTIJA ENDÉMICA *Podarcis lilfordi* (Günther, 1874) EN LAS ISLAS BALEARES.

Dissertation presented by Katherin Eliana Alejandra Otálora Acevedo to apply for the
doctoral degree by the Universitat de Barcelona

Katherin Eliana Alejandra Otálora Acevedo
Barcelona, 2025

Approval of the director and tutor of the doctoral thesis:

Dr. Laura Baldo
Associate professor
Director

Dr. Raül Ramos
Associate professor
Tutor

A mi padre, madre, hermanos, abuelita,
sobrinito, familia, amigas y red de mujeres.

ACKNOWLEDGMENTS

Quiero expresar mi profunda gratitud a todas las personas que han contribuido de manera significativa a la realización de esta tesis doctoral, ha sido un proceso muy enriquecedor lleno de aprendizajes personales y profesionales.

En primer lugar, agradezco sinceramente a mi directora de tesis, Dra. Laura Baldo por su orientación, paciencia y apoyo constante a lo largo de este arduo proceso. Sus consejos y dedicación fueron fundamentales para dar forma a esta investigación. Al Dr. Raül Ramos por su colaboración constante, escucha y apoyo para sacar adelante la tesis en la última fase.

Asimismo, quiero extender mi agradecimiento al equipo de investigación EVOK, Dr. Giacomo Tavecchia, Dr. Andreu Rotger, Dra. Jessica Gomez-Garrido, Dr. Joan Lluís Riera, Dr. Carles Galia Camps a los miembros de mi comisión de seguimiento Dr. Matha Pascual y Dr. Salvador Carranza, por sus valiosas sugerencias y comentarios que enriquecieron este trabajo.

Agradezco también a la Universitat de Barcelona y a la Beca Doctoral del Ministerio de Ciencias de Colombia por proporcionar los recursos necesarios y el ambiente propicio para llevar a cabo esta investigación.

A mi familia, cuyo amor ha sido el combustible que ha impulsado mi determinación y dedicación. A mi padre, José Otálora, mi razón de ser, quien me motiva día a día para alcanzar mis objetivos y me apoya en cada paso que doy, a mi hermana Evelyn y a mi sobrinito Thomas, por su amor, sabiduría y fortaleza. Su permanencia y confianza en mí son la razón principal por la que he alcanzado este logro. A mi madre Olga Acevedo, a mis hermanos Sebastián y Diego, por su cariño y apoyo. A mi abuelita Ana Elvia Espitia, por su incondicional amor. A mis grandes amigas y amigos Laura Escarmena, Alba Camacho, Guillem Rius, Fernanda Gonzales, Ángel Rain, Mario Hugo, Fernanda Mejía, Aida Viza, Julián Vega, Laia Pérez-Sorribes, Santiago Hernández, Andrea Ramirez, Andrea Arias, por acompañarme, apoyarme, entenderme, celebrarme, escucharme en momentos de duda y crear hermosos lazos de bienestar durante este proceso.

A todas las personas que, con un gesto, una palabra o un abrazo, han sido parte de este proceso, gracias por acompañar esta valiosa experiencia.

Katherin Eliana Alejandra Otálora Acevedo received a Doctoral Scholarship, Colombian Ministry of Science, Technology, and Innovation (MINCIENCIAS885/2020). Additionally, this research was supported by the Govern de les Illes Balears, Consellera d'Educació, Universitat i Recerca (PRD2018/25), by the Institute for Catalan Studies (PRO2020-S02-BALDO) and by MCIN/AEI/10.13039/501100011033/ and by FEDER A way to do Europe (Project PID2022-141578NB-C22).

The research presented in this dissertation, entitled **“Population Genetics of the Endemic Lizard *Podarcis lilfordi* (Günther, 1874) from the Balearic Islands”**, is structured around three core scientific chapters, each corresponding to a research article: two already published and one currently in preparation for submission. In addition, a methodological article was co-authored and published, providing foundational support for the development of the first chapter.

The methodological study, **“Chromosome-level genome assembly of Lilford’s wall lizard, *Podarcis lilfordi* (Günther, 1874) from the Balearic Islands (Spain)”**, was published in DNA Research (Vol. 30, Issue 3, 2023, DOI: <https://doi.org/10.1093/dnares/dsad008>), with a Journal Impact Factor of 3.9 (SCI, 2023). This journal ranks in the first quartile (Q1) within the “Genetics & Heredity” category, holding the 8th position out of 180 indexed journals, with an SCIE score of 91.2.

The first scientific chapter corresponds to the article **“Population genetics and phylogeographic history of the insular lizard *Podarcis lilfordi* (Günther, 1874) from the Balearic Islands based on genome-wide polymorphic data”**, published in Ecology and Evolution (Vol. 14, 2024, DOI: <https://doi.org/10.1002/ece3.11407>). This study provides a detailed analysis of the genetic diversity and phylogeographic structure of the species. The journal has an Impact Factor of 2.3 (SCI, 2023) and is ranked in the first quartile (Q1) in the Ecology category according to the Scimago Journal Rank (SJR), where it occupies the 41st position out of 175 journals, with an SJR score of 0.858 (2024).

The second chapter is currently being finalized and is provisionally titled **“Historical demographic inference in island populations: the case of the Lilford’s wall lizard, *Podarcis lilfordi* (Günther, 1874)”**. This chapter focuses on reconstructing the historical demography of the species, offering novel insights into island vertebrate evolution and informing conservation strategies related to population resilience.

The third chapter investigates the immunogenomic repertoire of *P. lilfordi* and corresponds to the article **“Antimicrobial peptides (AMPs) of lizards: a first comprehensive characterization of beta-defensins, ovo-defensins and cathelicidins from the Balearic lizard *Podarcis lilfordi* and closely related Lacertidae species”**, published in BMC Genomics (Vol. 26, 2025, DOI: <https://doi.org/10.1186/s12864-025-11797-2>). The journal reports an Impact Factor of 3.7 (SCI, 2023) and is ranked in the first quartile (Q1) according to the Scimago Journal Rank (SJR), where it holds an SJR score of 1.003 (2024) and ranks 12th out of 96 journals in the Genetics category.

Together, these chapters constitute a comprehensive scientific contribution to the understanding of the genomic architecture, demographic history, and immunological diversity of an insular endemic species, with broader implications for evolutionary biology and conservation management.

ABSTRACT

Island ecosystems offer unique windows into microevolutionary processes, particularly when recurrent climatic oscillations and anthropogenic pressures shape the genetic architecture of endemic taxa. This thesis investigates the population genetics, historical demography, and innate immune gene diversity of the Balearic lizard *Podarcis lilfordi*—an endangered lacertid lizard restricted to 43 islets across the Mallorca, Menorca, and Cabrera archipelagos—with the overarching goal of informing conservation strategies for highly fragmented insular populations.

Chapter 1 explores the phylogeographic structure and genome-wide patterns of genetic diversity within the species. Using a newly assembled reference genome and >300,000 SNPs from 191 individuals across 16 islets, the analyses reveal two deeply divergent lineages corresponding to the Mallorca–Cabrera and Menorca archipelagos (mean global $F_{ST} = 0.247 \pm 0.09$). Despite severe fragmentation, islet populations retain moderate levels of nucleotide diversity ($\pi = 0.144 \pm 0.021$) and low inbreeding coefficients ($F_{IS} < 0.10$), suggesting that historical connectivity, past population sizes, and possibly selection may have contributed to the retention of genetic diversity, although the precise mechanisms remain uncertain. Allelic richness increases significantly with islet surface area, indicating that larger islands may better sustain genetic diversity. Maximum-likelihood phylogenies based on concatenated SNPs support a Mallorcan origin for *P. lilfordi*, followed by colonization events toward Menorca and Cabrera during Pleistocene glacial lowstands, likely facilitated by temporary land connections or reduced sea barriers during periods of lowered sea levels.

Chapter 2 aims at reconstructing major historical demographic events within *P. lilfordi* looking at three hierarchical levels: (i) within the species, (ii) within archipelagos, and (iii) within individual populations/islands, focusing on three islands under current demographic study. For this purpose, site frequency spectra models (assuming a four-year generation time and a mutation rate of 1×10^{-8}) reveal strong population contractions during the Riss (~300–130 ka), Würm (~115–11.7 ka), and LGM (~33–19 ka), likely driven by habitat loss and climatic instability. Subsequent interglacial sea-level rises intensified fragmentation, although rare rafting events or short-lived land bridges likely permitted limited gene flow. A sharp decline in N_e around 4.3–4.5 ka coincides with early

human settlement, followed by habitat transformation and predator introduction which caused the extirpation of *P. lilfordi* from Mallorca and Menorca by ~2 ka and severe bottlenecks on remnant islets.

Chapter 3 explores the genetic potential of *P. lilfordi* to respond to infection threats by investigating its innate immune gene repertoire. Specifically, the chapter presents the first genome-wide characterization of AMPs in this species and compares them with close relatives in the Lacertidae family. Three AMP families were identified— beta-defensins, ovo-defensins (including three proline-rich variants), and cathelicidins—totaling 77 genes. Most beta- and ovo-defensins are located on chromosome 3, and cathelicidins on chromosome 12, flanked by conserved genomic markers. Comparative analyses revealed 58 orthologs or paralogs shared with *P. muralis*, *P. raffonei*, and *Zootoca vivipara*, including cathelicidins and ovo-defensins-PrAMPs, most of which were one-to-one orthologs between species, as well as recent gene expansions in beta- and ovo-defensins. Phylogenetic reconstructions suggest potential convergent evolution of AMP sequences across distantly related squamates. Despite historical bottlenecks and absence of gene flow, *P. lilfordi* retains high AMP gene diversity, potentially contributing to population resilience in isolated environments.

In summary, this thesis demonstrates that *P. lilfordi* has withstood both Pleistocene climatic shifts and recent human impacts while preserving significant genetic and functional variation. The identification of distinct genetic lineages and high immunogenomic diversity highlights the importance of managing each islet population as a distinct conservation unit. These findings offer a strong genomic framework for biodiversity conservation strategies targeting this key insular species, with broader implications for all isolated vertebrate populations in dynamic island ecosystems.

CATALÀ

Els ecosistemes insulars ofereixen finestres úniques als processos microevolutius, especialment quan les oscil·lacions climàtiques recurrents i les pressions antropogèniques configuren l'arquitectura genètica dels tàxons endèmics. Aquesta tesi investiga la genètica de poblacions, la demografia històrica i la diversitat gènica immunitària innata del sargantana balear *Podarcis lilfordi* —un lacèrtid en perill d'extinció restringit a 43 illots dels arxipèlags de Mallorca, Menorca i Cabrera— amb l'objectiu

general d'informar les estratègies de conservació per a poblacions insulars altament fragmentades.

El Capítol 1 explora l'estructura filogeogràfica i els patrons de diversitat genètica a nivell del genoma dins de l'espècie. Utilitzant un genoma de referència recentment assemblet i >300.000 SNP de 191 individus de 16 illots, les anàlisis revelen dos llinatges profundament divergents corresponents als arxipèlags de Mallorca-Cabrera i Menorca (F_{ST} global mitjana = $0,247 \pm 0,09$). Malgrat la fragmentació severa, les poblacions d'illots mantenen nivells moderats de diversitat de nucleòtids ($\pi = 0,144 \pm 0,021$) i coeficients de consanguinitat baixos ($F_{IS} < 0,10$), cosa que suggereix que la connectivitat històrica, les mides de les poblacions passades i possiblement la selecció poden haver contribuït a la retenció de la diversitat genètica, tot i que els mecanismes precisos continuen sent incerts. La riquesa al·lèlica augmenta significativament amb la superfície dels illots, cosa que indica que les illes més grans poden sostenir millor la diversitat genètica. Les filogènies de màxima probabilitat basades en SNP concatenats donen suport a un origen mallorquí de *P. lilfordi*, seguit d'esdeveniments de colonització cap a Menorca i Cabrera durant les zones glacials baixes del Plistocè, probablement facilitats per connexions terrestres temporals o barreres marítimes reduïdes durant períodes de baixada del nivell del mar.

El Capítol 2 té com a objectiu reconstruir els principals esdeveniments demogràfics històrics dins de *P. lilfordi* observant tres nivells jeràrquics: (i) dins de l'espècie, (ii) dins dels arxipèlags i (iii) dins de poblacions/illes individuals, centrant-se en tres illes que s'estan estudiant actualment. Per a aquest propòsit, els models d'espectres de freqüència de llocs (assumint un temps de generació de quatre anys i una taxa de mutació d' 1×10^{-8}) revelen fortes contraccions poblacionals durant el Riss (~300–130 ka), el Würm (~115–11,7 ka) i l'Últim Màxim Glacial (~33–19 ka), probablement impulsades per la pèrdua d'hàbitat i la inestabilitat climàtica. Els augments posteriors del nivell del mar interglacial van intensificar la fragmentació, tot i que els rars esdeveniments de ràfting o els ponts terrestres de curta durada probablement van permetre un flux genètic limitat. Una forta disminució de N_e al voltant de 4,3–4,5 ka coincideix amb els primers assentaments humans, seguida de la transformació de l'hàbitat i la introducció de depredadors que van causar l'extirpació de *P. lilfordi* de Mallorca i Menorca en ~2 ka i greus colls d'ampolla als illots restants.

El Capítol 3 explora el potencial genètic de *P. lilfordi* per respondre a les amenaces d'infecció investigant el seu repertori de gens immunitaris innats. Concretament, el capítol presenta la primera caracterització a nivell de genoma dels pèptids antimicrobians (AMP) d'aquesta espècie i els compara amb parents propers de la família Lacertidae. Es van identificar tres famílies d'AMP: beta-defensines, ovo-defensines (incloses tres variants riques en prolina) i catelicidines, que sumen un total de 77 gens. La majoria de les beta- i ovo-defensines es troben al cromosoma 3, i les catelicidines al cromosoma 12, flanquejades per marcadors genòmics conservats. Les anàlisis comparatives van revelar 58 ortòlegs o paràlegs compartits amb *P. muralis*, *P. raffonei* i *Zootoca vivipara*, incloent-hi catelicidines i ovo-defensines-PrAMP, la majoria de les quals eren ortòlegs un a un entre espècies, així com expansions gèniques recents en beta- i ovo-defensines. Les reconstruccions filogenètiques suggereixen una possible evolució convergent de les seqüències d'AMP a través d'esquamats distants. Malgrat els colls d'ampolla històrics i l'absència de flux genètic, *P. lilfordi* conserva una alta diversitat gènica d'AMP, cosa que pot contribuir a la resiliència de la població en entorns aïllats.

En resum, aquesta tesi demostra que *P. lilfordi* ha resistit tant els canvis climàtics del Plistocè com els impactes humans recents, alhora que ha preservat una variació genètica i funcional significativa. La identificació de diferents llinatges genètics i una alta diversitat immunogenòmica destaca la importància de gestionar cada població d'illots com una unitat de conservació diferent. Aquestes troballes ofereixen un marc genòmic sòlid per a estratègies de conservació de la biodiversitat dirigides a aquesta espècie insular clau, amb implicacions més àmplies per a totes les poblacions de vertebrats aïllats en ecosistemes insulars dinàmics.

CASTELLANO

Los ecosistemas insulares ofrecen una perspectiva única sobre los procesos microevolutivos, especialmente cuando las oscilaciones climáticas recurrentes y las presiones antropogénicas configuran la arquitectura genética de los taxones endémicos. Esta tesis investiga la genética de poblaciones, la demografía histórica y la diversidad de genes inmunitarios innatos de la lagartija balear *Podarcis lilfordi*, una especie en peligro de extinción que se encuentra restringida a 43 islotes de los archipiélagos de Mallorca,

Menorca y Cabrera, con el objetivo general de informar sobre las estrategias de conservación de las poblaciones insulares altamente fragmentadas.

El Capítulo 1 explora la estructura filogeográfica y los patrones genómicos de diversidad genética dentro de la especie. Utilizando un genoma de referencia recientemente ensamblado y más de 300 000 SNP de 191 individuos de 16 islotes, los análisis revelan dos linajes profundamente divergentes correspondientes a los archipiélagos de Mallorca-Cabrera y Menorca (F_{ST} global medio = $0,247 \pm 0,09$). A pesar de la grave fragmentación, las poblaciones de los islotes conservan niveles moderados de diversidad nucleotídica ($\pi = 0,144 \pm 0,021$) y bajos coeficientes de endogamia ($F_{IS} < 0,10$), lo que sugiere que la conectividad histórica, el tamaño de las poblaciones en el pasado y, posiblemente, la selección pueden haber contribuido a la conservación de la diversidad genética, aunque los mecanismos precisos siguen siendo inciertos. La riqueza alélica aumenta significativamente con la superficie de los islotes, lo que indica que las islas más grandes pueden mantener mejor la diversidad genética. Las filogenias de máxima verosimilitud basadas en SNP concatenados respaldan un origen mallorquín para *P. lilfordi*, seguido de eventos de colonización hacia Menorca y Cabrera durante los glaciares del Pleistoceno, probablemente facilitados por conexiones terrestres temporales o barreras marítimas reducidas durante los períodos de descenso del nivel del mar.

El Capítulo 2 tiene como objetivo reconstruir los principales acontecimientos demográficos históricos dentro de *P. lilfordi*, analizando tres niveles jerárquicos: (i) dentro de la especie, (ii) dentro de los archipiélagos y (iii) dentro de poblaciones/islas individuales, centrándose en tres islas que actualmente son objeto de estudio demográfico. Para ello, los modelos de espectros de frecuencia de sitios (suponiendo un tiempo de generación de cuatro años y una tasa de mutación de 1×10^{-8}) revelan fuertes contracciones de la población durante el Riss (~300-130 ka), el Würm (~115-11,7 ka) y el Último Máximo Glacial (~33-19 ka), probablemente impulsadas por la pérdida de hábitat y la inestabilidad climática. Las subidas del nivel del mar durante los interglaciales posteriores intensificaron la fragmentación, aunque es probable que algunos casos aislados de rafting o puentes terrestres de corta duración permitieran un flujo genético limitado. El fuerte descenso de N_e alrededor de 4,3-4,5 ka coincide con los primeros asentamientos humanos, seguidos de la transformación del hábitat y la introducción de

depredadores, lo que provocó la extinción de *P. lilfordi* en Mallorca y Menorca hace ~2 ka y graves cuellos de botella en los islotes restantes.

El Capítulo 3 explora el potencial genético de *P. lilfordi* para responder a las amenazas de infección mediante la investigación de su repertorio de genes inmunitarios innatos. En concreto, el capítulo presenta la primera caracterización genómica completa de los péptidos antimicrobianos (AMP) en esta especie y los compara con parientes cercanos de la familia Lacertidae. Se identificaron tres familias de AMP: beta-defensinas, ovo-defensinas (incluidas tres variantes ricas en prolina) y catelicidinas, con un total de 77 genes. La mayoría de las beta- y ovo-defensinas se encuentran en el cromosoma 3, y las catelicidinas en el cromosoma 12, flanqueadas por marcadores genómicos conservados. Los análisis comparativos revelaron 58 ortólogos o parálogos compartidos con *P. muralis*, *P. raffonei* y *Zootoca vivipara*, incluidas las catelicidinas y las ovo-defensinas-PrAMP, la mayoría de los cuales eran ortólogos uno a uno entre especies, así como expansiones genéticas recientes en las beta- y ovo-defensinas. Las reconstrucciones filogenéticas sugieren una posible evolución convergente de las secuencias de AMP en escamosos lejanamente emparentados. A pesar de los cuellos de botella históricos y la ausencia de flujo génico, *P. lilfordi* conserva una alta diversidad de genes AMP, lo que podría contribuir a la resiliencia de la población en entornos aislados.

En resumen, esta tesis demuestra que *P. lilfordi* ha resistido tanto los cambios climáticos del Pleistoceno como los impactos humanos recientes, al tiempo que ha conservado una importante variación genética y funcional. La identificación de linajes genéticos distintos y una alta diversidad inmunogenómica pone de relieve la importancia de gestionar cada población de islotes como una unidad de conservación distinta. Estos hallazgos ofrecen un sólido marco genómico para las estrategias de conservación de la biodiversidad dirigidas a esta especie insular clave, con implicaciones más amplias para todas las poblaciones de vertebrados aisladas en ecosistemas insulares dinámicos.

LIST OF ABBREVIATIONS

Abbreviation	Meaning
AccNos	Accession Numbers
ADP3	Antimicrobial Peptide Database
ALT	Alternative Allele
AMP	Antimicrobial Peptides
Ar	Allelic Richness
BAM	Binary Alignment Map
BD	Beta-Defensin
BLAST	Basic Local Alignment Search Tool
bp	Base Pair
BWA	Burrows-Wheeler Aligner
CATH	Cathelicidin
CDS	Coding DNA Sequence
CLD	Cathelin-Like Domain
CNAG	Centro Nacional de Análisis Genómico
CTSB	Cathepsin B
DNA	Deoxyribonucleic Acid
EN	Endangered
ESU	Evolutionarily Significant Unit
FASTK	Fas-activated serine/threonine kinase
FIS	Inbreeding Coefficient
FST	Fixation Index
GATK	Genome Analysis Toolkit

GBS	Genotyping-by-Sequencing
GO	Gene Ontology
GTR+ASC	General Time Reversible Model + Ascertainment Bias Correction
Ho	Observed Heterozygosity
IBD	Isolation by Distance
IUCN	International Union for Conservation of Nature
Ka	Thousand years ago
KLHL18	Kelch-Like Protein 18
LD	Linkage Disequilibrium
LGM	Last Glacial Maximum
MF	Molecular Functions
ML	Maximum Likelihood
MSC	Messinian Salinity Crisis
MTMR	Myotubularin-Related Protein
MUSCLE	Multiple Sequence Comparison by Log-Expectation
Mya	Millions of years ago
NCBI	National Center for Biotechnology Information
Ne	Effective Population Size
NJ	Neighbor Joining
ORF	Open Reading Frame
OVOD	Ovo-Defensins
PA	Private Allele
PO	Prior Odds
PrAMP	Proline-Rich Antimicrobial Peptide
RADseq	Restriction Site-Associated DNA Sequencing

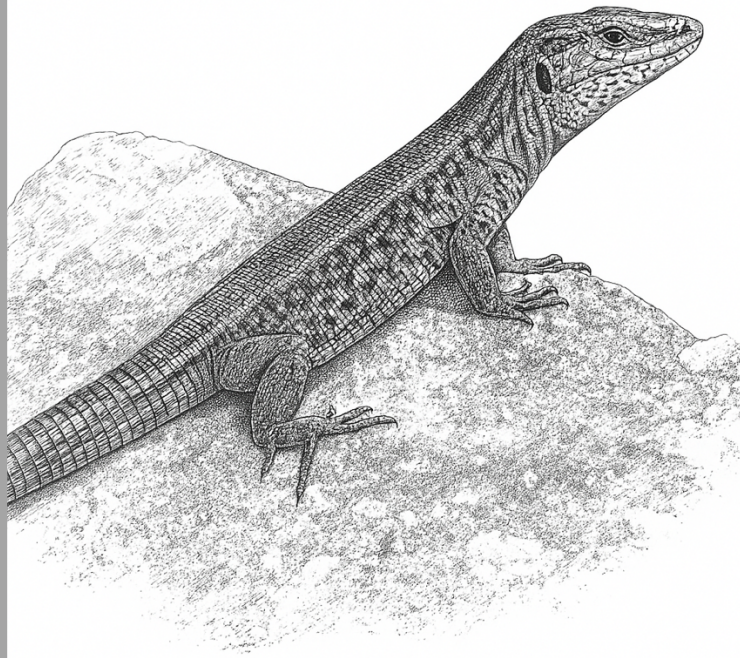
REF	Reference Position
SFS	Site Frequency Spectrum
SNPs	Single Nucleotide Polymorphisms
SVL	Snout-Vent Length
TBLASTN	Translated nucleotide databases using a protein query
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
VCF	Variant Call Format
XKR6	XK, Kell blood group complex subunit-related family member 6
XPO1	Exportin-1 Protein

CONTENTS

I. ACKNOWLEDGMENTS	7
II. ABSTRACT (ENG/CAT/CAST)	10
III. LIST OF ABBREVIATIONS	17
1. GENERAL INTRODUCTION	24
1.1. Evolutionary context and characterization of <i>P. lilfordi</i> .	24
1.1.1. Introduction to the study of insular species	24
1.1.1.1. Evolutionary processes on islands: genetic drift and selection	24
1.1.2. The Gymnesian islands	26
1.1.2.1. Historical and environmental background of the Balearic islands	26
1.1.3. Lizards from the genus <i>Podarcis</i>	28
1.1.3.1. The genus <i>Podarcis</i>	28
1.1.3.2. The Balearic lizard <i>P. lilfordi</i> : History, phenotypic differentiation, ecology and conservation considerations	29
1.2. Population genetics/ genomics and phylogeography	31
1.2.1. Genetic and genomic tools	31
1.2.1.1. Single nucleotide polymorphisms (SNPs) and their application to population studies	31
1.2.1.2. Demographic inference	33
1.2.2. Genomic foundations for <i>P. lilfordi</i>	34
1.2.2.1. Population genetics of <i>P. lilfordi</i> : current knowledge and gaps	34
1.2.2.2. Sequenced and assembled genome	38
1.3. Antimicrobial peptides (AMPs) in reptiles	40
1.3.1. Reptile innate immunity	40
1.3.2. Introduction to AMPs	41
1.3.2.1. Types and characteristics: beta-defensins, cathelicidins, and ovo-defensins	43
	20

1.3.3. AMPs in reptiles and squamata	45
1.3.3.1. Diversity and characteristics in reptiles	45
1.3.3.2. <i>P. lilfordi</i> and Lacertidae innate immunity: gap of knowledge	46
2. RESEARCH OBJECTIVES	49
3. CHAPTER I	52
Population genetics and phylogeographic history of the insular lizard <i>Podarcis lilfordi</i> (Günther, 1874) from the Balearic Islands based on genome-wide polymorphic data	52
4. CHAPTER II	99
Historical demographic inference in island populations: the case of the Balearic lizard, <i>Podarcis lilfordi</i> (Günther, 1874)	100
5. CHAPTER III	128
Antimicrobial peptides (AMPs) of lizards: a first comprehensive characterization of beta-defensins, ovo-defensins and cathelicidins from the Balearic lizard <i>Podarcis lilfordi</i> and closely related Lacertidae species	129
6. GENERAL DISCUSSION	166
6.1. Introduction to the discussion	166
6.2. Evolution in insular environments	168
6.3. Adaptive immune response in island populations: relationship between genetic diversity and AMP diversification as a defense against local pathogens	172
6.4. Conservation of <i>P. lilfordi</i> : the need for integrative studies to assess the species adaptive potential in the face of future threats	176
6.5. Limitations and future directions	179
6.5.1. Genomic resolution: moving towards high-density genomics	179
6.5.2. Identification of adaptive loci: integrating GWAS and outlier analyses	179
6.5.3. Demographic inference: challenges in sampling and model complexity	179
6.5.4. Antimicrobial peptides (AMPs) characterization: intraspecific diversity underexplored	180
6.5.5. Functional data integration: towards an integral evolutionary genomics framework	180
7. CONCLUSIONS	183
	21

8. REFERENCES	186
9. ANNEX	210
10. ORIGINAL PUBLICATIONS	216



INTRODUCTION

1. GENERAL INTRODUCTION

1.1. Evolutionary context and characterization of *P. lilfordi*

1.1.1. Introduction to the study of insular species

1.1.1.1. *Evolutionary processes on islands: genetic drift and selection*

Islands are natural evolutionary laboratories, providing isolated and controlled environments where processes like genetic drift and natural selection play dominant roles in shaping biodiversity (Gillespie, 2016; Losos & Ricklefs, 2009). The unique combination of biotic, abiotic, and stochastic factors ensures that each island population has a distinct demographic history (MacArthur & Wilson, 1967). The isolation and typically small population sizes characteristic of islands create ideal conditions for studying genetic drift, a process involving the random fluctuation of allele frequencies within populations over time (Frankham, 1997). On islands, genetic drift is amplified due to the limited gene pool, often resulting in the fixation or loss of alleles independently of their adaptive value (Wright, 1931). Stochastic events, such as environmental disturbances or natural disasters (Banks et al., 2013), can cause a dramatic reduction in population size, leading to a significant reduction in genetic diversity and causing populations to diverge genetically from their mainland relatives or neighboring islands. Genetic drift decreases genetic variation in a manner that is inversely related to the size of the population (Crow & Kimura, 1970). As a result, drift is anticipated to be particularly significant in islands with small population sizes (Woolfit, & Bromham, 2005). Genetic drift on islands can also occur due to a founder effect, where a few individuals colonize a new environment and establish a population with limited genetic variation (Mayr, 1942). Both stochasticity and founder effects critically shaped island populations, influencing evolutionary trajectories in ways distinct from those seen on continents (Clegg et al., 2002).

In addition to genetic drift, natural selection is a powerful evolutionary force on islands, often operating under unique conditions compared to the mainland. Stabilizing selection

is frequently predominant, favoring intermediate traits optimized for the resource-limited environments of many islands, which enhances survival and maintains genetic stability within populations (Grant & Grant, 1995; Russo et al., 2009; Losos & Ricklefs, 2009; Estes & Arnold, 2007; Schluter, 2000). Nonetheless, islands also provide fertile ground for directional selection, particularly when environmental shifts or novel resources drive populations toward new adaptive peaks, as seen in the specialized beak evolution of Galápagos finches (Grant & Grant, 2002; Clegg & Owens, 2002). This is further illustrated in behavioral traits, such as altered call frequencies in island populations of the amphibian *Boana albomarginata* (Rebouças et al., 2020). While less common, disruptive selection may occur when distinct ecological niches within the same island favor extreme over intermediate phenotypes, potentially driving intraspecific divergence (Schluter, 2000; Lack, 1947; Clegg et al., 2008). A classic example involves *Geospiza fortis* on Santa Cruz Island, where bimodal beak sizes persisted through strong disruptive selection during droughts, despite constraints against the most extreme morphs due to interspecific competition or resource scarcity (Hendry et al., 2008; Grant & Grant, 2006). Although rare, disruptive selection on islands may promote ecological divergence and even speciation under sympatric conditions (Hendry et al., 2008), with *G. fortis* illustrating how assortative mating and fluctuating resources influence adaptive radiation (Martin & Wainwright, 2013).

Overall, the interplay between drift and selection on islands fosters high levels of endemism and specialization, with limited gene pools and ecological isolation accelerating evolutionary processes beyond what is typically observed on the mainland (Frankham, 1997; Grant & Grant, 2008). This dynamic is encapsulated in the concept of “island syndrome,” where life-history traits, morphology, and behavior evolve to match insular constraints (Baeckens & Van Damme, 2020; Novosolov et al., 2013), making islands essential natural laboratories for studying the mechanisms driving biodiversity and speciation worldwide.

1.1.2. The Gymnesian islands

1.1.2.1. *Historical and environmental background of the Balearic islands*

The Balearic Islands, comprising the Pityusic Islands (Ibiza and Formentera) and the Gymnesic Islands (Mallorca, Menorca, Cabrera, and surrounding islets), represent a geologically diverse archipelago in the western Mediterranean Sea. Their formation is closely connected to important geological events during the Miocene epoch (23–5.3 Mya). At that time, the Alpine Orogeny, part of the Baetic System, shaped the islands into a typical horst-and-graben landscape (a system of uplifted blocks and depressions). Mallorca, which belongs to the NE–SW oriented Balearic Promontory, shows a particularly complex relief because the Earth's crust there was affected by both extensional (stretching) and compressional (squeezing) processes during the Cenozoic. This geological history, together with wider Mediterranean tectonic dynamics such as backarc extension and slab rollback, has left a legacy of great structural diversity and remains an active subject of scientific research (Cuerda, 1989; Issar, 2004).

The isolation of the Balearic Islands became more pronounced following the MSC (Messinian Salinity Crisis) (around 5.3 Mya, which is regarded as the point at which the archipelago became more distinctly separated from mainland Europe and North Africa (Clauzon et al., 1996; F. Gautier et al., 1994; Krijgsman et al., 1999). During the Pleistocene epoch (approximately 0.2 Mya), the fluctuating Mediterranean Sea levels caused the islands to experience cycles of submergence and reemergence, shaping their distinct geological and ecological composition (Goy et al., 1997). These climatic fluctuations during the Pleistocene, including events like the Würm Glaciation (115,000–11,700 years ago), had significant effects on the islands' biota, with glaciations altering the paleogeography of the Mediterranean region and influencing the evolutionary processes of species (Terrasa et al., 2009). The Mediterranean region was shaped by these glaciations, and the changing climate over the past 2.5 million years contributed to the evolutionary development of species living on the islands.

Human colonization of the Balearic Islands did not occur until around 4420–4220 cal years before present (BP), which allows for a clear understanding of the natural evolution of the islands' landscapes throughout much of the Holocene (11.7 ka to present) (Burjachs et al., 2017; Fernandes et al., 2020). Prior to human settlement, climatic oscillations during

the Early and Middle Holocene had already begun transforming the landscape, with significant changes in vegetation occurring before human arrival. Vegetation shifts on Menorca occurred between approximately 5.8 and 4.7 Ka, while Majorca exhibited floristic changes over a broader time span, from around 7.2 to 2.5 Ka. In contrast, Ibiza—located within the southern Pityusic Islands—experienced more gradual vegetation transitions, which appear to have coincided with the 4.2 Ka climatic anomaly, a period that also corresponds with the earliest evidence of human presence on the island (Burjachs et al., 2017). This highlights the dominant influence of climate in shaping the landscape, with human impact becoming more significant during the Late Holocene.

The Gymnesian Islands, present more varied topography compared to the smaller, flatter Pityusic Islands of Ibiza and Formentera (Mayol et al., 2020). These islands, with their distinctive landscapes, became biodiversity hotspots due to their geographical isolation and the diverse Mediterranean climates that prevail on the islands. Hot, dry summers and mild, wet winters create a variety of ecological niches, which in turn contribute to the development of numerous endemic species, particularly among reptiles and invertebrates (Aurelle et al., 2022; Barbosa et al., 2012; Terrasa et al., 2009; Thompson, 2020). The unique evolutionary processes that have occurred on the islands are driven by the Mediterranean Sea acting as a barrier, allowing for the differentiation of species and subspecies between island populations and their mainland relatives (Aurelle et al., 2022; Terrasa et al., 2009).

The islands status as continental islands, rather than oceanic islands, further explains their unique evolutionary trajectories. Unlike oceanic islands, whose biota are shaped largely by long-distance dispersal, continental islands like the Balearics share a historical connection with the mainland. Vicariance, the geographic isolation of species driven by ancient climatic and geological shifts, has played a central role in the evolution of the Balearic Islands' biota (Gillespie & Roderick, 2002). Additionally, processes such as "relaxation" (the loss of species due to reduced area) and "relictualization" (the development of endemics after the extinction of mainland relatives) are more common on continental islands (Gillespie & Roderick, 2002). The evolution of species on the Gymnesian Islands, including the emergence of unique endemic species, has been significantly influenced by these processes, providing a rich field for understanding the mechanisms of speciation and the impact of historic climatic fluctuations and geological transformations on island ecosystems.

These historical dynamics are closely linked to the present-day ecological diversity of the Balearic archipelago, where insular isolation has promoted the emergence of unique habitats and species (Terrasa et al., 2009). Notable examples of this insular biota include the extinct Balearic cave goat *Myotragus balearicus*, a case of extreme insular adaptation (Alcover et al., 2000), and the Mallorcan midwife toad *Alytes muletensis*, which survives only in restricted mountain habitats (Moore et al., 2004). The endemic lizard *Podarcis lilfordi*, now confined to the Gymnesic islets, depends on structurally complex habitats for shelter and reproduction (Pons & del Valle, 2020; Llorens & Gil, 2017). Altogether, these cases highlight how the combination of environmental heterogeneity and isolation has turned the Balearic Islands into a natural laboratory for understanding speciation and biodiversity under insular conditions (Barberá et al., 2012; Pavón et al., 2019; Larrucea et al., 2016).

1.1.3. Lizards from the genus *Podarcis*

1.1.3.1. The genus *Podarcis*

The Balearic Islands are home to several endemisms, including two lizard species from the genus *Podarcis*: *P. lilfordi* (Günther, 1874), and *P. pityusensis* (Boscá, 1883). The genus *Podarcis* represents one of the most diverse and abundant reptile groups in Southern Europe and the Mediterranean region, notable for its remarkable diversity and adaptability. First described by Wrangel in 1830, *Podarcis* includes the well-known wall lizards found across Europe and North-West Africa. The genus currently comprises approximately 26 recognized species (Poulakakis et al., 2005; Yang et al., 2021) and nearly 300 subspecies, including several inhabiting islands (Böhme et al., 1985). This extensive diversity has led to numerous taxonomic revisions, driven by the discovery of new phylogenetic relationships (Poulakakis et al., 2005; Yang et al., 2021).

Over the past 20 million years, the diversification of lizards in the genus *Podarcis* has been largely influenced by recurrent climatic shifts and tectonic transformations within the Mediterranean basin (Terrasa et al., 2009; Yang et al., 2021). Phylogenetic reconstructions suggest an early separation—dated to over 11 million years ago—between two principal continental lineages, associated with the Iberian and Balkan Peninsulas, alongside two distinct groups composed of island-endemic taxa. A major phase of diversification is inferred to have occurred between 6.5 and 4 million years ago,

coinciding with the MSC. Throughout this evolutionary trajectory, gene flow among divergent lineages has played a continuous role, leading to the emergence of genomes composed of mixed ancestries. This history of introgression has contributed to the appearance of hybrid forms, some of which have given rise to insular endemic species.

Overall, the extensive genomic diversity of this genus and current widespread distribution underscores its ability to diversify and adapt to a wide range of environments, making it the most successful reptile genus in Southern Europe (Harris & Arnold, 1999). Given the complex geological history and ecological heterogeneity of these islands, it is conceptually appropriate to focus on *Podarcis* as a taxonomic framework through which broader questions of island biogeography and evolutionary diversification can be addressed (Poulakakis et al., 2005). This genus has been repeatedly highlighted in the literature as a valuable system for investigating the evolutionary consequences of isolation, habitat structuring, and restricted gene flow in island and continental contexts (Yang et al., 2021). Moreover, its repeated use in phylogeographic and evolutionary studies makes it particularly well-suited for integrating genomic, ecological, and spatial data to disentangle the drivers of insular divergence. In this sense, the genus *Podarcis* offers an ideal entry point for examining how landscape-scale features of the Balearics influence lineage diversification in vertebrate taxa.

1.1.3.2. *The Balearic lizard P. lilfordi: History, phenotypic differentiation, ecology and conservation considerations*

P. lilfordi (Günther, 1874), commonly known as Lilford's wall lizard, is an emblematic lizard of the Balearic Islands, currently restricted to a fragmented network of 43 small offshore islets surrounding the main islands of Menorca, Mallorca, and Cabrera. It is a medium-sized reptile with an omnivorous diet, and exhibits marked phenotypic diversity among populations (Rotger et al., 2016), with substantial variation in body size, scalation patterns, and dorsal coloration, including the recurrent evolution of melanistic morphs (Pérez-Cembranos et al., 2020; Bassitta et al., 2020). Morphological divergence correlates strongly with climatic gradients across the archipelagos, with lizards from more arid islets generally possessing fewer and larger scales, an adaptation likely linked to thermoregulation and water retention (Pérez-Cembranos et al., 2020). In terms of morphology, this lizard is relatively sturdy, with males attaining snout–vent lengths of up to 81 mm and females up to 75 mm (Pérez-Mellado, 1998). Body size variation follows an

inverse Bergmann's rule, with larger individuals predominating in lower-latitude, warmer habitats such as those in Mallorca and Cabrera (Pérez-Cembranos et al., 2020). Nonetheless, multiple factors may also drive variation in body size, including changes in diet, levels of intraspecific competition, and reductions in predation pressure (Uller et al., 2019). These phenotypic and ecological differences are further shaped by historical isolation, limited gene flow, and local selective pressures, resulting in distinct evolutionary trajectories among insular populations (Pérez-Cembranos et al., 2020). Approximately 2,000 years ago during the Holocene, the Balearic lizard vanished from Mallorca and Menorca, most likely due to the introduction of terrestrial predators by the Romans (Alcover, 2000; Terrasa et al., 2009). All these aspects make this species a model system to study evolutionary adaptation and genetic differentiation processes in the Balearic Islands (Pérez-Cembranos et al., 2021; Terrasa et al., 2009).

P. lilfordi is classified as an endangered species, with an EN category by the IUCN (Pérez-Mellado & Martínez-Solano, 2009a). Several factors represent a current threat to the long-term persistence of these populations. Many islets that host these lizards have very small populations, rendering them highly susceptible to disturbances. Human activities can disrupt local populations and undermine genetic processes in the populations (Rotger et al., 2021). Increasing populations of yellow-legged gulls (*Larus cachinnans*) pose a predation threat, and the use of poisoned bait in campaigns to eradicate gulls or rats exacerbates the risk to the lizards. On the other hand, the introduction of goats to some islands has led to vegetation destruction, which is critical for the lizards' habitat and survival (Mayol et al., 2020; Pérez-Mellado & Pérez-Cembranos, 2025). At present, the horseshoe snake (*Hemorrhois hippocrepis*), a predator specialized on lizards that has been accidentally introduced in the Balearic Islands, represents the most important threat to the persistence of *P. lilfordi* populations (Mencía et al., 2017).

1.2. Population genetics/ genomics and phylogeography

1.2.1. Genetic and genomic tools

1.2.1.1. *Single nucleotide polymorphisms (SNPs) and their application to population studies*

Advances in molecular genetics, particularly the application of single nucleotide polymorphisms (SNPs), have revolutionized population genetic studies by offering insights into the organism evolutionary history, adaptation, and conservation (Bassitta et al., 2021; Prates et al., 2018; Yang et al., 2021). SNPs are highly abundant, reproducible, and provide broad genome-wide coverage, making them a powerful alternative to traditional markers such as microsatellites and mitochondrial DNA (Nielsen et al., 2020). While early concerns such as ascertainment bias were raised (Morin et al., 2004), advances in sequencing and computational tools have made SNPs cost-effective and widely accessible. These advantages enable researchers to investigate genetic diversity, population history—including the reconstruction of past demographic events—and to detect signatures of selection with greater precision (Fischer et al., 2017; Morin et al., 2004; Nielsen et al., 2020; Puritz et al., 2012). In conservation contexts, SNPs are used to monitor genetic health, detect inbreeding, and identify populations at risk of genetic erosion, supporting the design of targeted management strategies (Morin et al., 2004; Nielsen et al., 2020).

In addition to their utility in population-level monitoring, SNPs provide a powerful framework for understanding local adaptation and evolutionary divergence (Morin et al., 2004; Nielsen et al., 2020). Outlier SNPs, which deviate from neutral expectations, are particularly informative for detecting loci under selective pressure (Ahrens et al., 2018; Excoffier et al., 2009; Narum & Hess, 2011). These loci often correspond to traits that influence fitness, including responses to environmental stressors, predation, or habitat heterogeneity (Ahrens et al., 2018; Bassitta et al., 2021; Prates et al., 2018). Identifying

such genomic signatures enhances our understanding of the molecular mechanisms driving evolution in natural populations.

Research in lizards and marine vertebrates has demonstrated how SNP-based approaches can reveal the genetic basis of local adaptation and population divergence across diverse ecological contexts. In *Anolis* lizards, Prates et al. (2020) used genome–environment association methods to identify genetic differentiation driven by temperature and climatic gradients, illustrating how environmental selection shapes genomic variation. In marine species such as the turtle *Chelonia mydas*, SNPs have been used to detect loci associated with thermoregulation, hypoxia, melanism, morphogenesis, osmoregulation, diet, and reproduction (Álvarez-Varas et al., 2021). These loci have been linked to ecological specialization and behaviors such as natal homing, providing insight into how environmental and life-history factors drive genetic divergence. In lizards, outlier loci have similarly been associated with selective pressures such as beach–inland transitions and altitudinal gradients (Brown et al., 2023).

Genotyping-by-sequencing (GBS) has emerged as a groundbreaking method for SNP discovery and genotyping, widely adopted across diverse taxa due to its simplicity, cost-effectiveness, and scalability. Initially developed by (Elshire et al., 2011), GBS employs restriction enzymes to reduce genomic complexity, followed by the addition of barcoded adapters, enabling the simultaneous analysis of large populations. This makes it particularly advantageous for applications in genetic diversity, genome-wide association studies, and marker-assisted selection, even in species with complex or non-model genomes (Ruperao et al., 2023; Torkamaneh et al., 2016).

A key strength of GBS lies in its streamlined workflow, which minimizes sample preparation steps compared to techniques like RAD-seq and ddRAD-seq. This efficiency not only reduces costs but also enhances the throughput of sequencing, especially on platforms like Illumina, which are optimized for processing large datasets (De Donato et al., 2013). Furthermore, GBS provides reliable SNP detection even in genetically diverse or highly heterozygous populations, making it suitable for studies in agriculture, forestry, and conservation genomics (Torkamaneh et al., 2016; Ulaszewski et al., 2021). Its capacity to balance reduced genome representation with high-resolution genotyping ensures GBS remains a cornerstone in modern population genomics (Ruperao et al., 2023; Ulaszewski et al., 2021).

1.2.1.2. *Demographic inference*

Demographic inference using SNPs has become essential for reconstructing population history and dynamics, especially in non-model organisms. Historical and contemporary effective population size (N_e) estimates provide insight into genetic drift, migration, and bottleneck events (Marchi et al., 2021; Noskova et al., 2023). The integration of SNP frequency spectra (histogram that shows the distribution of different allele frequencies for a collection of genetic variations in a population sample) further enhances these analyses, allowing for precise modeling of allele frequency changes over time and the detection of past demographic shifts (Liu & Fu, 2020). Additionally, usually the methods like Sequentially Markovian Coalescent (SMC) frameworks, are normally utilized with dense SNP datasets to infer N_e dynamics, revealing population responses to environmental and anthropogenic pressures (Li & Durbin, 2011; Schiffels & Durbin, 2014; Sellinger et al., 2021; Terhorst et al., 2016).

The utility of SNPs in estimating N_e spans across temporal scales. Contemporary estimates are critical for conservation, as they assess genetic diversity and inbreeding risks in endangered populations. Historical N_e calculations, however, offer a deeper understanding of past population expansions or contractions, informing on population/species ability to persist through climatic or geological changes and overall guiding strategies for species conservation (Dedato et al., 2022; Hohenlohe et al., 2020).

By linking genomic variation to effective population sizes, demographic inference illuminates the interplay between genetic diversity and demographic events. SNP analyses coupled with environmental and ecological data have further refined models, identifying how specific pressures shape genetic structure, guiding strategies for species conservation (Dedato et al., 2022; Hohenlohe et al., 2020).

1.2.2. Genomic foundations for *P. lilfordi*

1.2.2.1. Population genetics of *P. lilfordi*: current knowledge and gaps

Genetic data suggest that *P. lilfordi* and its closely related species *Podarcis pityusensis* began to diverge after the Mediterranean basin was re-flooded, at the end of the Messinian salinity event (5.3 Mya) (Brown et al., 2008; Terrasa et al., 2009). Mitochondrial DNA studies suggest that these two species form a basal group within the genus *Podarcis*, with *P. sicula* being their closest relative (Oliverio et al., 2000). Cytochrome b sequence analyses estimate that the divergence between *P. lilfordi* and *P. pityusensis* occurred approximately 4.95 million years ago, which aligns with the geological separation of the Gymnesian Islands (Mallorca, Menorca) from the Pityusic Islands about 5 Mya (Terrasa et al., 2004). During the onset of the Quaternary period around 2.6 Mya, the cladogenesis of the Balearic lizard of Menorca and Mallorca took place (Brown et al., 2008; Terrasa et al., 2009). Concurrently, during the late Pliocene epoch, sea levels exceeded those of the present day (Siddal et al. 2003), thereby obstructing any terrestrial connection between the two islands (Pérez-Cembranos et al., 2021).

Population genetics of *P. lilfordi* has been long limited to mitochondrial DNA and a few nuclear markers, including microsatellites, which constrained the resolution of evolutionary and demographic inferences (Terrasa et al., 2009; Rotger et al., 2021). According to these partial genetic data, the phylogeographic structure of *P. lilfordi* reveals four main clades, reflecting deep insular divergence shaped by historical and geological processes. Menorcan populations diverged ~2.6 Mya, showing a south-to-north haplotype distribution (Brown et al., 2008; Terrasa et al., 2009), while a distinct western Mallorcan clade (Dragonera, Malgrats, Toro) separated ~2 Mya due to Pleistocene allopatry. The remaining clades encompass central/northern Mallorca and Cabrera, with the latter isolated ~1.2 Mya, suggesting a northward gene flow (Brown et al., 2008; Terrasa et al., 2009) (**Figure 1**). Haplotype networks show stronger structuring and diversity in Mallorca-Cabrera populations versus the star-like pattern in Menorca, indicative of recent expansion (Terrasa et al., 2009). Complementing these insights, Rotger et al. (2021) analyzed microsatellites in three islets on the South of Mallorca (Na Moltana, Na Guardia, and En Curt), revealing a correlation between genetic diversity and island size. En Curt showed marked genetic differentiation likely due to a founder effect,

while Moltona and Na Guardia were more similar, though the latter displayed signs of bottlenecking, possibly from human settlement. Anthropogenic impacts, including translocations and quarrying, further shaped gene flow and demography, with En Curt hosting the densest population and lizards exhibiting slower growth despite larger sizes—emphasizing how insular ecological dynamics interact with evolutionary trajectories. However, these advances still leave major questions unanswered, requiring higher-resolution genomic data to clarify patterns of variation and historical processes.

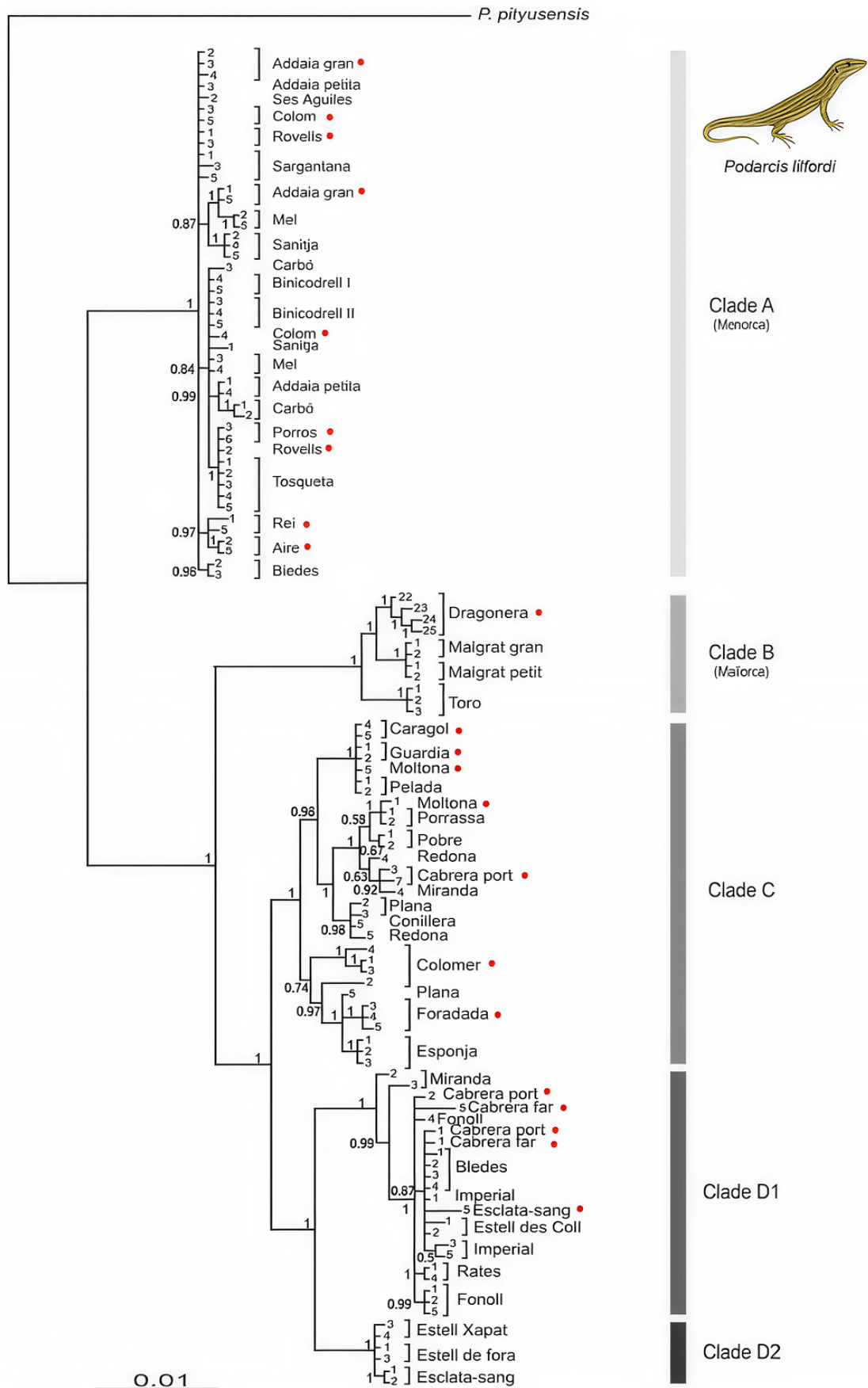


Figure 1. Bayesian-inference phylogeny illustrating the relationships among taxa, with principal clades (A–D2) labeled. Support values above each branch denote posterior probability estimates derived from the topology’s posterior distribution. The red dot indicates the populations included in this study. (Adapted from Terrasa et al., 2009).

A comprehensive understanding of the genetic diversity and population structure of *P. lilfordi* across the Balearic archipelago remains incomplete. Despite the insights gained from mitochondrial and microsatellite data (Terrasa et al., 2009; Rotger et al., 2021), current knowledge does not yet clarify how genetic diversity is distributed across the archipelago—particularly whether patterns of variation align with geographical features such as islet size and isolation. Similarly, the phylogeographic relationships among island populations are only partially resolved, with major questions still open regarding the directionality of colonization events, the root of the species’ phylogeny, and the historical role of geographic barriers. The absence of genome-wide data continues to hinder efforts to test hypotheses on hierarchical clustering, dispersal routes, and the emergence of independent evolutionary lineages—key aspects for understanding the historical biogeography and insular diversification of this species.

In addition to the limited phylogeographic resolution, a major gap in the population genomic knowledge of *P. lilfordi* concerns the near-complete absence of studies addressing adaptive genetic variation, as most investigations have focused on neutral markers, overlooking regions potentially under selection (Terrasa et al., 2009; Rotger et al., 2021). Consequently, the influence of environmental heterogeneity, ecological pressures, and geographic isolation on local adaptation remains largely unexplored, with no genome-wide scans to detect outlier loci or selective sweeps. At the onset of this research (2021), a ddRAD-seq dataset for *P. lilfordi* had recently become available (Bassitta et al., 2021), providing the first genome-wide SNP resource for the species and demonstrating the utility of such data to resolve population differentiation. Together with previously published mitochondrial sequences (Terrasa et al., 2009), these resources established a valuable framework to contextualize further genomic analyses. Building upon this foundation, the present study develops a high-quality reference genome and expands population-level sampling through Genotyping-by-Sequencing (GBS), thereby advancing the understanding of both neutral and adaptive processes in *P. lilfordi* and informing evidence-based conservation strategies. This genome-wide perspective

begins with the development of a high-quality reference genome, which is described in the following section.

1.2.2.2. *Sequenced and assembled genome*

A high-contiguity genome assembly for *P. lilfordi* was generated during this PhD thesis to obtain a reference for SNPs mapping (Gomez-Garrido et al., 2023). The assembly was obtained integrating advanced sequencing approaches, including 10X Genomics linked reads, Oxford Nanopore long-read technology, and Hi-C scaffolding, and transcriptomic data from both Illumina and PacBio platforms. This integrated approach yielded a contiguous genome of approximately 1.5 Gb with 99% of sequences mapped to chromosomes and an annotation of 25,663 protein-coding genes (Gomez-Garrido et al., 2023). The genome architecture comprises 18 autosomes along with a pair of sex chromosomes (Z and W), assembled with high resolution through the use of Hi-C data and detailed manual refinement (Gomez-Garrido et al., 2023). In addition, a fully annotated mitogenome of 17,251 bp highlights conserved vertebrate gene structures (Figure 2).

Comparisons with *P. muralis* demonstrate strong genome collinearity, affirming evolutionary stability over 18–20 Mya. The assembly achieved a scaffold N50 of 90 Mb and included comprehensive annotation of repetitive elements (39%), aligning with closely related species. This genomic resource established a new benchmark for reptilian genomics, offering unprecedented opportunities to explore molecular underpinnings of phenotypic diversity and inform conservation genomics, especially as *P. lilfordi* faces growing environmental and anthropogenic challenges (Gomez-Garrido et al., 2023).

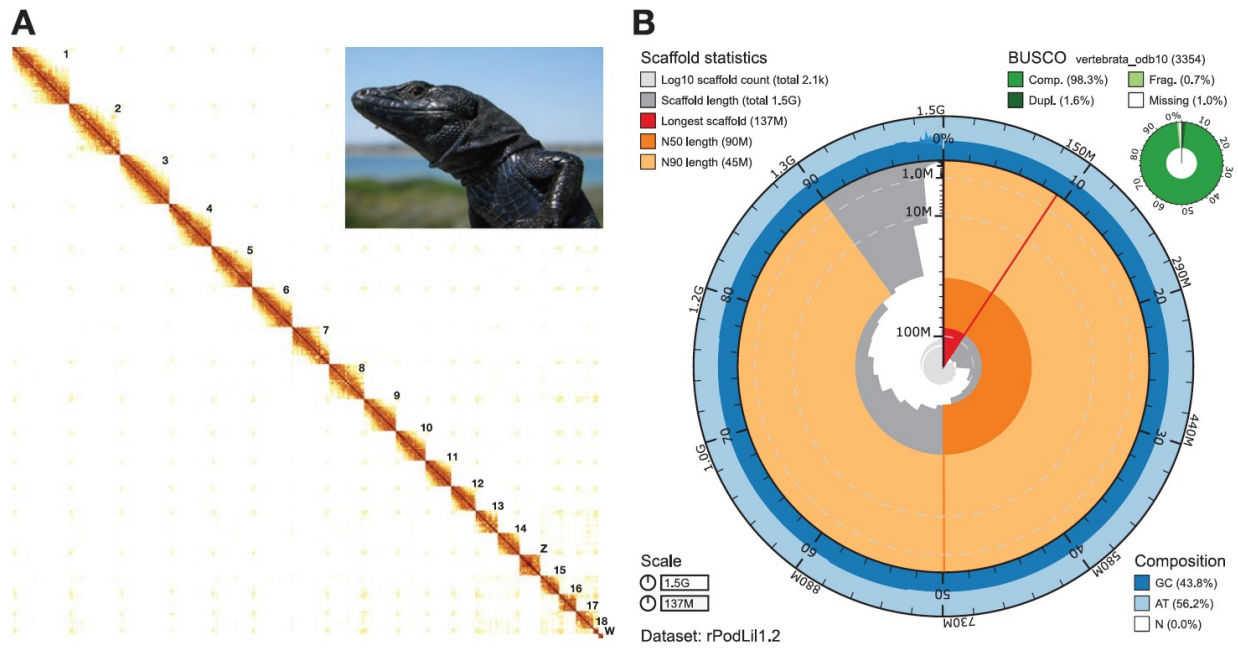


Figure 2. Visual summary of the rPodLil1.2 assembly. (A) Hi-C contact map of the genome assembly visualized in PretextView. The map shows 20 superscaffolds, ordered from longest to shortest, corresponding to the 18 autosomes and the two sex chromosomes (Z and W). Additionally, there are a total of 44 short unlocalized scaffolds within the main chromosomes, and 2,084 short unplaced scaffolds (not visible at the resolution shown). (B) Snailplot summarising assembly metrics, including scaffold statistics, BUSCO completeness, total size and base composition. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 1.460.085.851 bp assembly. The distribution of record lengths is shown in dark grey with the plot radius scaled to the longest record present in the assembly (137 Mb, shown in red). Orange and pale-orange arcs show the N50 and N90 record lengths (90 Mb and 45 Mb), respectively. The pale grey spiral shows the cumulative record count on a log scale with white scale lines showing successive orders of magnitude. The blue and pale-blue area around the outside of the plot shows the distribution of GC, AT and N percentages in the same bins as the inner plot. A summary of complete, fragmented, duplicated and missing BUSCO genes in the vertebrata_odb10 set is shown in the top right.

1.3. Antimicrobial peptides (AMPs) in reptiles

1.3.1. Reptile innate immunity

Reptiles represent valuable models for studying innate immunity, as they retain both primitive recognition systems and adaptive components, and their immune responses are shaped by a long evolutionary history under varied environmental pressures (Field et al., 2022; Zimmerman, 2020). Unlike mammals, which have been extensively studied, the immune systems of reptiles—particularly within Crocodylia, Testudines, Squamata, and Rhynchocephalia—have only recently begun to be characterized in detail (Zimmerman et al., 2010; Nash & Ryan, 2023; van Hoek et al., 2014). In this context, it is important to note that immune defense constitutes a fundamental physiological mechanism by which organisms confront infections caused by parasites or pathogens (Lochmiller, 1996), playing a pivotal role in determining survival (Lochmiller et al., 2000). Immunity is classically categorized into innate and adaptive branches. Innate immunity relies on pattern recognition receptors (PRRs) capable of detecting pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs), which signal the presence of infection or tissue damage, respectively (Zimmerman, 2020). Once triggered by PAMPs or DAMPs, PRRs activate the innate immune cascade (Kumar et al., 2011), which in turn primes the adaptive system, enabling it to generate a specific response and immunological memory against subsequent exposures (Cooper et al., 2006).

In all multicellular organisms, the innate immune arm is conserved (Zimmerman et al., 2010), though its functional prominence varies among lineages. In reptiles, which are ectotherms, the immune response is skewed toward innate mechanisms due to their lower metabolic requirements compared to endothermic organisms, as they do not internally regulate body temperature (Pough, 1980). The innate immune system provides rapid, generalized protection and initiates downstream adaptive responses (Iwasaki et

al., 2015). Key components of this system include the complement cascade, phagocytic leukocytes, inflammation, behavioral fever, and notably, AMPs.

Due to their strong reliance on innate responses and minimal adaptive interference, reptiles offer a distinct model for studying innate immunity, though research remains limited (Zimmerman et al., 2010; French et al., 2007a, 2007b, 2008; Graham et al., 2011; Neuman-Lee et al., 2015, 2017). The effectiveness of innate immunity in reptiles is strongly shaped by ecological and physiological conditions. Seasonal changes, nutritional state, and stress have all been shown to influence immune function across multiple species, highlighting the significant role of environmental context in modulating immunocompetence (Nash & Ryan et al., 2023). These findings underscore the importance of considering dynamic ecological variables when assessing immune function in reptiles. Within this framework, AMPs emerge as particularly relevant immune effectors, as their structural diversity, tissue-specific expression, and deep evolutionary conservation in reptiles warrant focused efforts toward their description, annotation, and characterization to contribute to the process of ecoimmunology and One Health.

1.3.2. Introduction to AMPs

AMPs represent a heterogeneous group of small, naturally occurring proteins or peptides that play a key role in the homeostasis of the innate immune system. Widely distributed across diverse organisms, AMPs exhibit potent broad-spectrum activity against bacteria, fungi, viruses, and protozoan parasites (Lei et al., 2019; Mohammed et al., 2017; Zhang et al., 2021). In vertebrates, families such as cathelicidins and defensins participate in immune signaling, inflammation regulation, and the orchestration of responses that link innate and adaptive immunity (Huan et al., 2020). They target various microbial structures, such as membranes and intracellular components, with three types of membrane actions (barrel-stave, toroidal-pore, carpet models), exploiting their cationic and amphipathic nature to selectively interact with anionic bacterial membranes—a mechanism that allows them to act against antibiotic-resistant microorganisms (Li et al., 2022) (**Figure 3**). In contrast, the largely neutral composition of host cell membranes reduces their susceptibility to this mechanism of action (van Hoek et al., 2019). Over 3,700 AMPs (APD3 data base) have been identified across multiple biological kingdoms,

including animals, plants, fungi, and microorganisms, underlining their universal importance (Huan et al., 2020; Zhang et al., 2021; Wang et al., 2015).

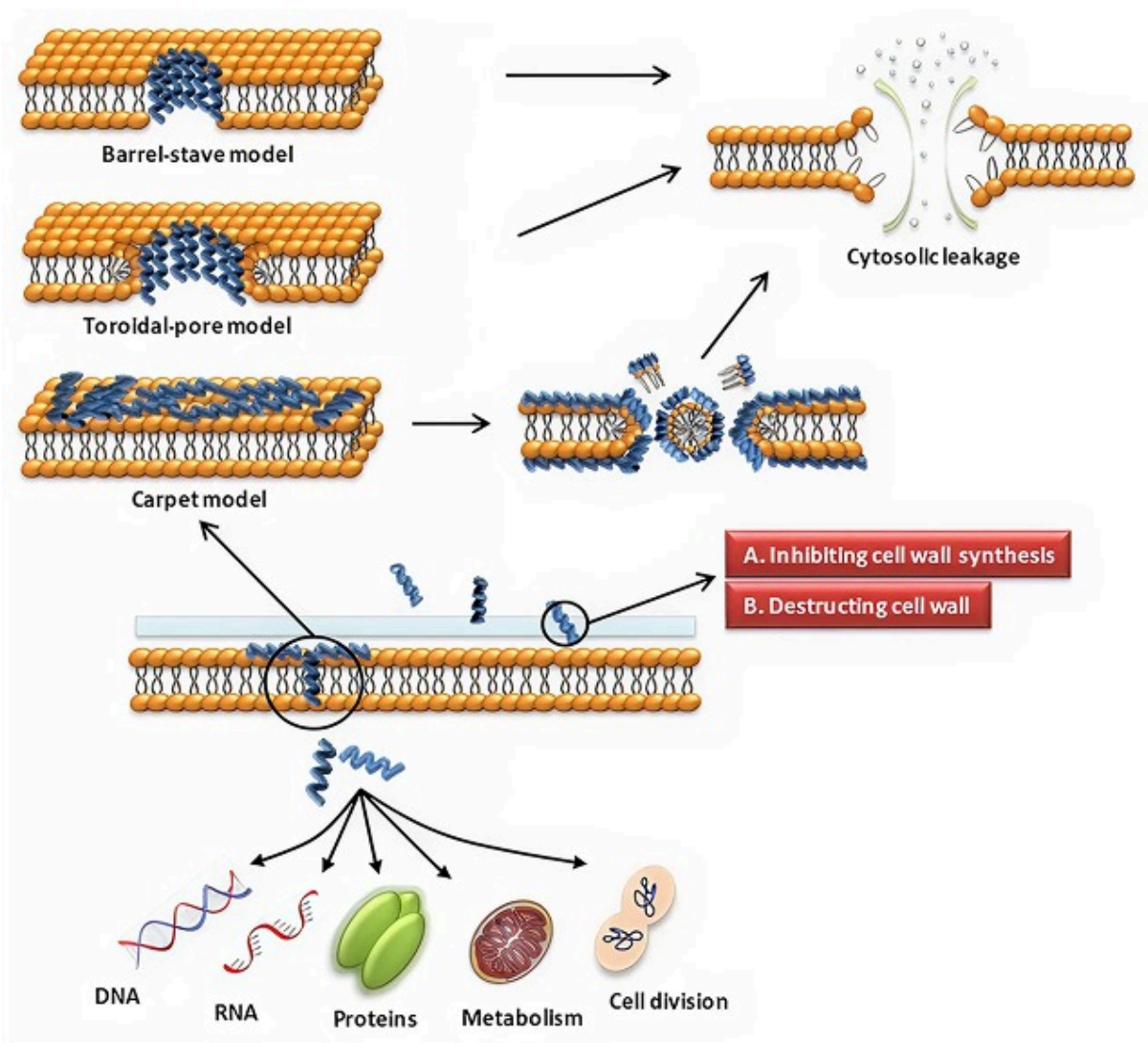


Figure 3. Schematic Overview of Antimicrobial Peptide (AMP) bactericidal mechanisms. Some AMPs compromise the cytoplasmic membrane via distinct mechanisms—barrel-stave pore formation, toroidal-pore generation and carpet-like disruption—thereby elevating membrane permeability, provoking cytosolic leakage and ultimately causing cell lysis. Other AMPs target the cell wall by inhibiting peptidoglycan synthesis and disrupting its structural integrity. A further subset of AMPs translocates into the cell—either by direct penetration or endocytic uptake—and exerts antimicrobial effects on intracellular components, including nucleic acids, organelles and, in fungi, species-specific proteins (Source: Li et al., 2022).

The adaptation of AMPs in different species highlights their evolutionary significance and versatility (Torres et al., 2019). In particular, the continuous discovery of new AMPs and their diverse mechanisms of action is facilitated by advances in genomic mining and machine learning techniques. These technologies enable the identification and development of AMPs with enhanced properties, offering promising therapeutic options to combat the increasing global health threat of antimicrobial resistance (Madhavan et al., 2023; Rai et al., 2016; Rojas-Pirela et al., 2023; Zhang et al., 2021).

Beyond their antimicrobial properties, AMPs are involved in a range of biological processes, including immune modulation, angiogenesis, wound healing, and antitumor effects (Ma et al., 2020; Wu et al., 2014). These additional functions make them a promising alternative to address the growing challenge of antibiotic resistance in modern medicine (Huan et al., 2020). Despite their potential, natural AMPs face limitations, such as poor stability and cytotoxicity, which has led to extensive research focused on optimizing these peptides for therapeutic use (Huan et al., 2020; Madhavan et al., 2023; Rai et al., 2016; Rojas-Pirela et al., 2023; Zhang et al., 2021). The AMPs in crocodilians have been identified as promising candidates for therapeutic use due to their potent antimicrobial properties (Bishop et al., 2015; Santana et al., 2021).

1.3.2.1. Types and characteristics: beta-defensins, cathelicidins, and ovo-defensins

AMPs are classified into various subgroups based on their amino acid sequence, structure, and properties. The three major vertebrate families include defensins (α -, β -, θ -, and ovo-defensins), histatins and cathelicidins (Findlay et al., 2016; Haney et al., 2019; Wang, 2021). AMPs are categorized into four groups based on their secondary structures: α (helical structures), β (beta-strands), $\alpha\beta$ (combined helical and beta-strands, such as beta-defensins), and non- $\alpha\beta$ (lacking both structures, e.g., indolicidin). This classification highlights the structural diversity of AMPs and their varied mechanisms of action (Wang, 2021). On the other hand, the vertebrate defensin family encompasses alpha-, beta-, theta-, and ovo-defensins, with alpha- and theta-defensins found exclusively in mammals, while ovo-defensins are specific to birds and reptiles (van Hoek et al., 2019).

Beta-defensins, characterized by their β -sheet structure, typically contain 2–8 cysteine residues that form disulfide bonds essential for their antimicrobial activity (Koehbach &

Craik, 2019). These peptides are widely found in mammals, including humans, where they play a critical role in the immune response by disrupting microbial membranes through electrostatic interactions and hydrophobic effects (Koehbach & Craik, 2019; Zhang et al., 2021; Zhu & Gao, 2013). In addition to mammals, beta-defensins have been identified in reptiles, amphibians, and birds, where they serve similar protective functions, particularly in defense against environmental pathogens (Santana et al., 2021; van Hoek et al., 2019). In reptiles, β -defensins display notable structural variability. Although the canonical architecture comprises 2 to 3 exons—typically encoding a signal sequence, a protein with six cysteine residues forming the β -defensin core, and occasionally an additional short exon—there is considerable interspecific variation. These include differences in exon number and length, the number of β -defensin domains per gene, unconventional cysteine arrangements, and deviations in the total number of cysteines (van Hoek et al., 2019). Ovo-defensins represent a unique subfamily of AMPs, predominantly found in egg whites and reproductive tissues of birds and reptiles. They share a β -sheet structure similar to beta-defensins but have distinctive properties that enhance their antimicrobial efficacy. Ovo-defensins can bind metal ions, such as zinc, increasing their activity against bacterial strains like *Escherichia coli* and *Staphylococcus aureus*. They play a crucial role in protecting the egg from pathogens, and their structural diversity offers opportunities for designing novel AMPs for therapeutic purposes (Whenham et al., 2015).

Cathelicidins, another well-known family, are typically cationic α -helical AMPs that exhibit amphipathic properties, unique to vertebrates (Bals & Wilson, 2003). They are characterized by precursor proteins containing conserved N-terminal pre-pro-cathelin domains, essential for their structure and function (van Harten et al., 2018). This subgroup is best represented by LL-37, the only human cathelicidin, which is activated from the precursor hCAP18 by serine protease in neutrophils (Bhattacharjya et al., 2024). LL-37 undergoes conformational changes in the presence of certain anions, adopting an α -helical structure that enhances its antimicrobial and immune-modulating functions. These peptides act mainly by integrating into microbial membranes, causing membrane disruption and cell death, while also modulating immune responses during infection (Zhang et al., 2021).

1.3.3. AMPs in reptiles and squamata

1.3.3.1. Diversity and characteristics in reptiles

Reptiles, as some of the oldest amniotes, possess a vast, but still largely underexplored array of AMPs (Mason-Smith, 2023; Modesto & Anderson, 2004; van Hoek, 2014). Recent advancements in proteomics and genomics have started to shed light on the genetic and functional diversity of these peptides across various reptilian genomes, indicating a rich reservoir of antimicrobial capabilities (Dalla Valle et al., 2012; Mason-Smith, 2023; van Hoek et al., 2019).

The Squamata order, which includes both snakes and lizards, carry a remarkable variety of AMPs tailored to their distinct ecological niches and challenges (van Hoek, 2014). In snakes, for instance, AMPs play crucial roles in their venom composition: proteomic analyses of snake venom have revealed a complex array of peptides with diverse antimicrobial properties, which contribute to their ability to neutralize pathogens and subdue prey (Aird et al., 2013; Oguiura et al., 2023; Zhao et al., 2008).

Lizards, the major group within Squamata, also exhibit a similar diversity of AMPs, according to the few studies available for the Komodo dragon (*Varanus komodoensis*) and the green anole (*Anolis carolinensis*) (van Hoek et al., 2014; Dalla Valle et al., 2012). The Komodo dragon genome encodes 66 β -defensins, six ovo-defensins and three cathelicidins, many arranged in clusters, reflecting strong innate defenses against its pathogen-rich environment (van Hoek et al., 2019). Similarly, in *A. carolinensis*, combining the reference genome with EST (Expressed sequence tag) libraries, identified 32 β -defensin-like genes with conserved cysteine motifs, variable tissue expression and alternative splicing, while phylogenetic evidence links some of these peptides to venom proteins, suggesting a shared ancestral origin (Dalla Valle et al., 2012). Genomic approaches in lizards, have identified several AMP genes, which are essential for their survival in varied environments, highlighting the evolutionary flexibility of their immune systems to cope with diverse microbial challenges (Alibardi, 2013; Dalla Valle et al., 2012; van Hoek et al., 2019).

These studies of snakes and lizards reveal both shared and unique features in their AMP repertoires, offering valuable insights into their evolutionary development (Dalla Valle et al., 2012; van Hoek et al., 2019). Future research on AMPs in Squamata holds significant potential to advance our understanding on their remarkable resistance of reptiles to infections as well as on their ability to thrive in the different environments. It will also contribute to the discovery of novel toxins and the development of novel therapeutics to combat microbial infections.

1.3.3.2. *P. lilfordi* and Lacertidae innate immunity: gap of knowledge

To date there are no published studies characterizing AMPs in *P. lilfordi*, nor within the family Lacertidae, which includes genera such as *Podarcis* and *Zootoca*. In the case of *P. lilfordi*, a species subject to strong ecological constraints and genetic isolation, the exploration of AMPs offers an opportunity to investigate potential immunological adaptations associated with insular living. The characterization of AMP diversity in this context is critical for advancing our knowledge of eco-immunology in insular vertebrates. This perspective complements genomic and demographic analyses, contributing to a more integrated understanding of mechanisms underpinning persistence and resilience in small, fragmented populations. Additionally, AMP analyses directly connect to population demography and phylogeographic structure, providing novel insights into how immune function may shape long-term survival in endangered island reptiles.

At the same time, a comparative analysis across distinct lizard species might inform on the evolutionary context of AMPs diversification in *P. lilfordi*. Although genomic data exist for other Lacertidae—such as *Podarcis muralis*, *Podarcis raffonei*, and *Zootoca vivipara*—there has been no specific genomic or functional characterization of AMPs in any of these closely related species. As a result, key aspects remain unaddressed, including the complete repertoire of AMP gene families, their genomic organization and chromosomal clustering, the conservation or divergence of cysteine motifs, the extent of gene duplication or loss and their pattern of molecular evolution. Moreover, the phylogenetic relationships of *P. lilfordi* AMPs to orthologs in other Lacertidae, as well as potential lineage-specific innovations associated with insular living, remain unexplored. Addressing these gaps will be essential to clarify the evolutionary trajectories and adaptive significance of AMPs in insular reptiles.



RESEARCH OBJECTIVES

2. RESEARCH OBJECTIVES

2.1 Main goal

The main objective of this thesis is to study microevolutionary processes in allopatric populations of the endemic lizard *P. lilfordi*, distributed across the islets of the Menorca, Mallorca, and Cabrera archipelagos in the Balearic Islands. Using genomic data (SNPs and a newly assembled genome), the aim is to understand how geographic fragmentation, climatic shifts, and anthropogenic pressures have shaped the genetic diversity, phylogeographic structure, demographic history, and innate immune repertoire of this threatened endemic species.

2.2 Specific goals

2.2.1. Assess the genetic diversity and population structure of *P. lilfordi* using SNP data

Hypothesis: A genetic structure is expected that mirrors the geographical layout of the archipelagos, wherein larger and less isolated islets retain higher levels of genetic diversity, while smaller, more remote islets exhibit stronger effects of genetic drift and potential inbreeding.

2.2.2. Reconstruct the phylogeographic relationships among populations and identify signals of selective pressures in *P. lilfordi*

Hypothesis: The genetic diversification of *P. lilfordi* should reflect an initial colonization event from Mallorca or Menorca, followed by geographic isolation and local population adaptation. These dynamics are presumed to have left detectable signatures in both phylogeographic structure and selected functional loci, including associations with specific environmental and geographic features of each islet.

2.2.3. Identify signals of demographic changes in *P. lilfordi* populations.

Hypothesis: The demographic history of *P. lilfordi* has been strongly shaped by Pleistocene glaciations and, more recently, by habitat fragmentation driven by anthropogenic activities. These pressures are expected to have caused severe and

heterogeneous population bottlenecks, with the most pronounced effects in smaller, more isolated populations.

2.2.4. Identify and describe antimicrobial peptide (AMP) using the recently assembled *P. lilfordi* genome.

Hypothesis: The AMP repertoire of *P. lilfordi* should display a diversity in number and composition comparable to that described for the closest related species with available AMP data, such as the Komodo dragon, due to the *P. lilfordi* geographic isolation and low gene flow. We further expect this repertoire to include clear orthologues in other Lacertidae species and to exhibit conserved structural cysteine motifs characteristic of each AMP family — β -defensins, ovo-defensins, and cathelicidins— consistent with a monophyletic origin of reptilian antimicrobial defense. In line with the rapid evolutionary dynamics of immune genes, we anticipate low phylogenetic resolution across AMP sequences, even among lineages that diverged millions of years ago—a pattern likely driven by functional convergence and the structural constraints imposed on antimicrobial activity.



3. CHAPTER I

Population genetics and
phylogeographic history of the
insular lizard *Podarcis lilfordi*
(Günther, 1874) from the
Balearic Islands based on
genome-wide polymorphic
data

Population genetics and phylogeographic history of the insular lizard *Podarcis lilfordi* (Günther, 1874) from the Balearic Islands based on genome-wide polymorphic data

Katherin Otalora^{1,5}, Joan Lluís Riera¹, Giacomo Tavecchia², Andreu Rotger², José Manuel Igual², Jean-Remi Paul Trotta³, Laura Baldo^{1,4*}

¹Department of Evolutionary Biology, Ecology and Environmental Sciences, Biology Faculty, University of Barcelona (UB), Barcelona, Av. Diagonal 645, Barcelona 08028, Spain

²Animal Demography and Ecology Unit, (GEDA-IMEDEA, CSIC -UIB), c. M. Marques, 07190, Esporles, Spain.

³CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), 08028 Barcelona, Spain

⁴Institute for Research on Biodiversity (IRBio), University of Barcelona (UB), Barcelona, Spain

⁵Fundación FUNMAJO, EBA, RAIEC, Biodiversity Branch, Tunja, Boyacá, Colombia.

*Corresponding author:

Laura Baldo

baldo.laura@ub.edu

DOI: <https://doi.org/10.1002/ece3.11407>

Abstract

Islands provide a great system to explore the processes that maintain genetic diversity and promote local adaptation. We explored the genomic diversity of the Balearic lizard *Podarcis lilfordi*, an endemic species characterized by numerous small insular populations with large phenotypic diversity. Using the newly available genome for this species, we characterized more than 300,000 SNPs, merging Genotype by Sequencing (GBS) data with previously published Restriction-site associated DNA sequencing (RADSeq) data, providing a dataset of 16 island populations (191 individuals) across the range of species distribution (Menorca, Mallorca, and Cabrera). Results indicate that each islet hosts a well-differentiated population ($F_{st}=0.247\pm0.09$), with no recent immigration/translocation events. Contrary to expectations, most populations harbor a considerable genetic diversity (mean nucleotide diversity, $\pi=0.144\pm0.021$), characterized by overall low inbreeding values ($F_{is}<0.1$). While the genetic diversity significantly decreased with decreasing islet surface, maintenance of substantial genetic diversity even in tiny islets suggests variable selection or other mechanisms that buffer genetic drift. Maximum-likelihood tree based on concatenated SNP data confirmed the existence of the two major independent lineages of Menorca and Mallorca/Cabrera. Multiple lines of evidence, including admixture and root testing, robustly placed the origin of the species in the Mallorca Island, rather than in Menorca. Outlier analysis mainly retrieved a strong signature of genome differentiation between the two major archipelagos, especially in the sexual chromosome Z. A set of proteins were target of multiple outliers and primarily associated with binding and catalytic activity, providing interesting candidates for future selection studies. This study provides the framework to explore crucial aspects of the genetic basis of phenotypic divergence and insular adaptation.

1. Introduction

Island populations represent biological scenarios of greatest evolutionary interest due to their geographic isolation, which accelerates the process of organismal diversification (Warren et al., 2015; Whittaker et al., 2017), and their restricted scale of study, that allow close monitoring (Drake et al., 2002; Losos and Ricklefs, 2009). They also typically host a high rate of endemism, making them a global conservation priority (Manes et al., 2021; Sivaperuman et al., 2008). The Balearic Islands, encompassing the Pityusic (Ibiza and Formentera) and Gimnesic (Mallorca, Menorca, Cabrera and associated islets) islands, comprise a large number of islands that originated following a complex geological and climatic history related to the Mediterranean sea level variations during the Pleistocene (0.2- Mya) (Cuerda, 1989; Goy et al., 1997). Among the several endemic species, they host the two species of Balearic lizard (i.e., *P. lilfordi* from the Gimnesic islands and *P. pityusensis* from the Pityusic islands). They are among the most extensively studied vertebrates of the archipelago, due to their remarkable phenotypic diversity (Pérez-Cembranos et al., 2020), high population densities and reduction of anti-predatory behavior associated to the “island syndrome” (Cooper and Pérez-Mellado, 2012; Hawlena et al., 2009; Novosolov et al., 2013; Rotger et al. 2023).

P. lilfordi (Günther, A., 1874) presently occurs on 43 off-the-coast islets of Mallorca and Menorca, as well as in the Cabrera archipelago. Current distribution is the result of a vicariance process during the Mediterranean sea desiccation and subsequent reflooding (5.33 Mya) (Krijgsman et al., 1999; Terrasa et al., 2009). The disappearance of *P. lilfordi* from the main islands of Menorca and Mallorca dates back more than 2000 years, following the introduction of predators by humans (Alcover, 2000). The species is now confined on islets that provide an effective refuge for to date populations. Within these islets, the species evolved in a context of limited resources and absence of major predators and competitors, greatly diversifying in morphology, pigmentation (Pérez-Cembranos et al., 2020; Rotger et al., 2021), behavior and life history traits (Salvador, A., 2009; Rotger et al., 2023). Major forces driving this diversification can be ascribed to the high intrapopulation competition (density-dependent factors) (Grant and Benton, 2000; Le Galliard et al., 2010; Massot et al., 1992), as well as intense genetic drift resulting from recurrent bottlenecks, primarily affecting the smallest islets (Bassitta et al., 2021; Charlesworth et al., 2003; Rotger et al., 2021; Terrasa et al., 2009). This context of rapid phenotypic differentiation among insular populations makes *P. lilfordi* an interesting

vertebrate model for ecological and evolutionary studies (Camargo et al., 2010), particularly those addressing the genomic basis of local adaptation and persistence of small populations (Bassitta et al., 2021; Yang et al., 2022).

Genetic analyses of this species have been long limited to mtDNA, few nuclear genes (Brown et al., 2008; Terrasa et al., 2009) and microsatellites for few populations (Bloor et al., 2011; Rotger et al., 2021). A recent study based on single nucleotide polymorphisms (SNPs) data by Restriction Site Associated DNA Sequencing (RADSeq) (Bassitta et al., 2021) has expanded on previous knowledge on genetic diversity and phylogeographic pattern of this species, confirming the existence of two well-discriminated genetic lineages separating populations from the major archipelagos of Menorca and Mallorca/Cabrera (Brown et al., 2008; Terrasa et al., 2009). Recently, the *P. lilfordi* genome has been sequenced (Gomez-Garrido et al., 2023) providing the framework for a comprehensive exploration of population-level genomics of this species.

Here, we used for the first time the newly available reference genome of *P. lilfordi* to map and annotate newly generated genome-wide polymorphic markers obtained by two independent sequencing methods, Genotype by Sequencing (GBS) and RADSeq. We specifically characterized SNPs of eight new populations (100 individuals) by GBS, and integrated RADSeq data of additional 10 populations (91 individuals from Bassitta et al. (2021), for a comprehensive analysis of 16 populations (two were common to both studies), spanning the main range of distribution of this species. Individual GBS and RADSeq datasets were independently analyzed as well as merged to obtain a representative set of common SNPs loci across all populations, which allowed for a robust comparative genomics within *P. lilfordi*.

Our main objectives were to: a) assess the level of intraspecific genetic diversity and structuring of the different populations of *P. lilfordi*; b) reconstruct the species main phylogeographic scenario of colonization of the Balearic Islands; c) identify potential signatures of genome diversification and loci underpinning the phenotypic diversity and insular adaptation of this species.

2. Materials and Methods

2.1. Sampling and DNA extractions

Tails tissues of 100 specimens of *P. lilfordi* were collected between 2015 and 2016 from eight islets in the Menorca and Mallorca archipelagos (Balearic Islands) (10-22 samples per islet) (Figure 1 and Table 1). Islets encompassed a representative subsample of the current species geographic distribution and vary in several geographic and ecological aspects, including surface area, maximum altitude, vegetation cover and type, presence of human settlements and other vertebrates, as well as lizard demographic traits (Mayol et al., 2020; Pérez-Mellado et al., 2008; Rotger et al., 2021). Specimens were captured using pitfall traps placed along paths and vegetation edges, sexed according to visual examination and morphology (Rotger et al., 2016), weighted, and body size measured as snout to vent length (SVL) (see Table S1 for specimen information). Tails were preserved in 100% ethanol and kept at -80°C until processed. Permits for sampling were provided by Conselleria d'Agricultura i Medi Ambient i Territori, Govern de les Illes Balears (CEP 31/2015 to LB and CEP 6/14 to GT). Genomic DNA was extracted with DNAeasy Tissue and Blood kit (Qiagen) with RNase treatment. DNA concentration was measured with Qubit 2 fluorometer (Invitrogen, Waltham). About 1 ug of genomic DNA per sample was sent to *Centro Nacional de Analisis Genómico* (CNAG, Spain) for sequencing. DNA integrity and concentration was further measured on a fragment analyzer (Agilent Bioanalyzer) and high-quality samples with at least 10 ng/μl were selected for sequencing.

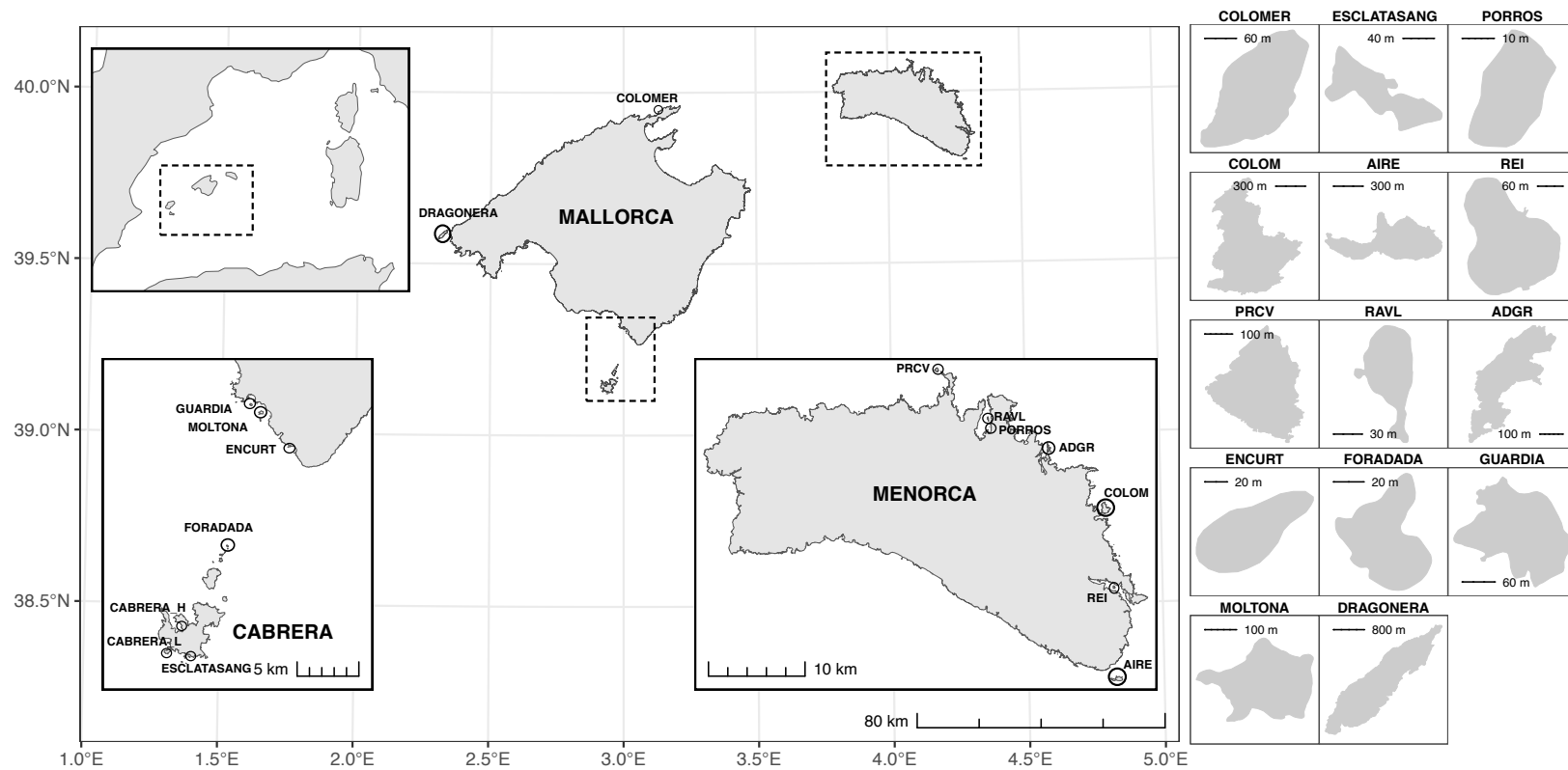


Figure 1. Geographic map of the 16 localities (15 islets) under study, sampled in the archipelagos of Mallorca, Cabrera, and Menorca.

2.2. Sequencing

A restriction site-associated DNA library was generated by GBS method (Elshire et al., 2011). To select for the best enzyme with reproducible genomic fragments across samples (a critical criterion for GBS experiments), we first generated a pilot library (four samples) with Illumina MiSeq (150 bp), using two common restriction enzymes, PstI and ApeKI. PstI provided the best reproducibility and number of variants and was therefore chosen to generate the final library.

Genomic DNA was digested with the PstI, fragments were tagged with individual barcodes, PCR-amplified, multiplexed, and sequenced on dual lanes on an Illumina HiSeq 2500 platform (2x100 bp). Raw read sequences were demultiplexed and quality checked using FastQC v0.10.1 (Wingett and Andrews, 2018); available at <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.

Additionally, raw RADseq genomic data of *P. lilfordi* generated on a previous study (Bassitta et al., 2021) were downloaded from the online repository (PRJNA645796, 91 samples from 10 populations). The GBS dataset (this study) and the RADseq dataset (Bassitta et al., 2021) shared the restriction enzyme PstI, which allowed data integration for comparative purposes (see below). Two localities of Colom and Aire were sampled and sequenced in both studies (different individuals) (Table 1 and S1).

Table 1. Sampling design and environmental variables

Major Archipelagos	Archipelago	Type	Official name ^a (locality)	Short name	No. of Individuals			Geographic variables ^a					
					GBS	RADSeq ^b	Total	X-coord	Y-coord	Surface area (ha)	Maximum Altitude (msnm)	SDI	Lizard Density ^c (ind/ha)
Menorca	Menorca	Islet	Illa Gran d'Addaia	ADGR	10		10	603284,8	4430301,1	7,5	21,7	2,251	1062
Menorca	Menorca	Islet	Illa de l'Aire	AIRE	11	12	23	610443,5	4406455,4	30,5	16,1	1,9461	4098
Mallorca/Cabrera	Cabrera	Inland	Cabrera (Harbour)	CABRERA_H		12	12	494459,1	4333606,0	1142,0	171,4	3,5241	45059,0
Mallorca/Cabrera	Cabrera	Inland	Cabrera (Lighthouse)	CABRERA_L		8	8	493279,9	4331172,0	1142,0	171,4	3,5241	45059,0
Menorca	Menorca	Islet	Illa d'en Colom	COLOM	10	10	20	609129,9	4424151,2	59,3	44,0	2,1169	1616
Mallorca/Cabrera	Mallorca	Islet	el Colomer	COLOMER		4	4	511211,5	4421664,5	2,8	105,4	1,1866	3339
Mallorca/Cabrera	Mallorca	Islet	sa Dragonera	DRAGONERA		4	4	441589,9	4381854,5	272,7	352,8	2,0490	729
Mallorca/Cabrera	Mallorca	Islet	Illot d'En Curt	EN CURT	23		23	503125,0	4347776,2	0,4	3,1	1,1943	1799
Mallorca/Cabrera	Cabrera	Islet	Esclatasang	ESCLTS		10	10	495084,6	4330730,1	0,6	39,2	1,7303	3118
Mallorca/Cabrera	Cabrera	Islet	Illot de na Foradada	FORADADA		10	10	498200,0	4339516,6	0,3	5,7	1,2941	2595
Mallorca/Cabrera	Mallorca	Islet	Na Guardia	GUARDIA	15		15	500100,3	4351217,2	2,5	8,6	1,3915	1784
Mallorca/Cabrera	Mallorca	Islet	Na Moltona	MOLTONA	11		11	500926,8	4350573,0	5,7	7,9	1,3946	2093
Menorca	Menorca	Islet	Porros de Fornells	PORROS		11	11	597159,9	4432654,7	0,1	2,2	1,1277	1676
Menorca	Menorca	Islet	Illa des Porros	PRCV	10		10	591657,0	4438581,0	8,6	18,4	1,6212	1100
Menorca	Menorca	Islet	Illa des Revells	RAVL	10		10	597032,1	4433305,5	0,4	8,0	1,5184	NA
Menorca	Menorca	Islet	Illa del Rei	REI		10	10	610060,7	4415982,1	4,3	15,2	1,2214	403

^a Vector geographic information was extracted from Base Topogràfica de les Illes Balears (https://ideib.caib.es/geoserveis/rest/services/public/GOIB_BTIB_IB/MapServer/27), publicly available from Institut Cartogràfic i Geogràfic de les Illes Balears (ICGIB). Elevations were obtained from the 2-m resolution MDT02 Digital Elevation Model (DEM) publicly available from the Centro Nacional de Información Geográfica (<https://centrodedescargas.cnig.es>).

^b From Bassitta et al. 2021

^cBased on Pérez-Mellado et al., 2008 and Bassitta et al., 2021

2.3. SNPs calling and filtering

Raw sequence data from GBS and RADSeq were independently processed as well as merged (COMBINED dataset). Thereafter, the three generated datasets (GBS, RADSeq and COMBINED) were processed through the same pipeline.

All barcoded raw reads were mapped to the *P. lilfordi* genome (rPodLil1.1, available at <https://denovo.cnag.cat/podarcis> (Gomez-Garrido et al., 2023) using the Burrows-Wheeler Aligner (BWA) software (v2.1) to generate individual Binary Alignment Map (BAM) files. The genome corresponds to a single female individual sampled in the Aire population (the population was included in our dataset). Variant calling per sample was performed with Genome Analysis Toolkit HaplotypeCaller from GATK v. 3.7 (Van der Auwera GA & O'Connor BD., 2020) to produce gVCF files (minimum-mapping-quality 20). gVCF files were finally genotyped for variant calling (i.e., SNPs).

Prior to data filtering, SNP and sample statistics were explored using VCFTools v. 0.1.16 (Danecek et al., 2011; De la Cruz and Raska, 2014). SNPs were then filtered using the following settings: --max-alleles 2, --remove-indels, --min-meanDP 10, --max-meanDP 50, --minQ 30, --maf 0.05, --max-missing 0.75 (individual datasets) or 0.90 (combined).

To retrieve common loci between GBS and RADSeq datasets, the max-missing parameter was set to 0.9, thus retaining loci shared by at least 90% of the specimens and effectively removing all loci unique to either GBS or RADSeq.

2.4. Population genetic diversity

We used the *populations* module in STACKS v.2.61 (Catchen et al., 2013) to calculate diversity statistics per population, including the observed heterozygosity (H_o), expected heterozygosity (H_e), nucleotide diversity (π), inbreeding coefficient (F_{is}) and number of private alleles (PAs). Allelic richness (A_r) was calculated with the 'hierfstat' package in R (Goudet et al., 2015) using four random individuals per population, i.e. the minimum number of samples per population (minimum number of alleles, min.all=8). Significant associations between genetic diversity indexes (H_e , π , A_r , F_{is}) and geographic variables (islet surface area, maximum altitude, and the shoreline development index, SDI, i.e., a measure of island perimeter complexity, all log-transformed) were assessed with simple and multiple linear regression in R.

2.5. Population genetic differentiation and clustering

For differentiation and structural analyses, loci under Linkage Disequilibrium (LD) were removed using *plink* v2.00a3.3 (Purcell et al., 2007) (`--indep-pairwise: 50 ['kb'] 5 0.20`) (Table 2). Pairwise *F_{st}* values (Weir & Cockerham, 1984) were estimated with the function *stampFst* from 'StAMPP' R package (Pembleton et al., 2013), and significant differences for each population pair were evaluated with confidence limits at 95% after 1000 bootstrap iterations. Heatmaps based on *F_{st}* were created using *pheatmap* in R (Raivo Kolde, 2019).

Isolation by distance (IBD) was evaluated by correlating the *F_{st}* genetic distances with linear geographic distances among populations, using a Mantel test (Mantel, 1967) with function *mantel.randtest* in the 'ade4' package in R (Dray and Dufour, 2007) with 1000 permutations. The test was performed for Menorca and Mallorca/Cabrera archipelago separately to control for distance effect between major archipelagos.

To reconstruct evolutionary relationships among island populations, we first used a Discriminant Analysis of Principal Components (DAPC), integrated in the 'Adegenet' R package (Jombart, 2008). In this multivariate statistical approach, sample genetic variance is partitioned into groups to maximize discrimination, without prior knowledge on genetic structure. The number of meaningful PCs retained was chosen by a cross-validation with the function *xvalDapc* (Jombart & Collins, 2015) with default parameters (1000 replicates).

Population structure, including contemporary and historic admixture among populations and their relationships, were further investigated using ADMIXTURE v1.3.0 (Alexander et al., 2009), setting the numbers of co-ancestry clusters (*K*) from 2 to 16 (maximum number of populations). Cross-validation was performed with the `-cv` option (10-fold) to retain the best value of *K* clusters (lowest cv error). Membership probabilities were calculated according to the retained discriminant functions and specimens showing above 90% of assignment probability to an island different from their source were considered as recently translocated, or potential mislabeled and removed (none detected).

2.6. Phylogeographic pattern

TreeMix v. 1.13 (Pickrell and Pritchard, 2012) was used to infer population splitting and mixture patterns. The method employs allele frequencies to build a graph-based model of the population network (as opposed to a bifurcating tree) by first building a Maximum Likelihood (ML) tree and then searching for migration events that increase the composite likelihood (Flesch et al., 2020; Pickrell and Pritchard, 2012). The program utilizes a Gaussian approximation to model genetic drift (drift parameter) along each population (Flesch et al., 2020; Pickrell and Pritchard, 2012). Using the combined data set filtered for LD, a ML tree was built with a window size (k) of 500 SNPs, evaluating from 0 to 16 migrations edges (m), 10 iterations per edge, using the “-noss” option to prevent overcorrection of sample size. The optimal number of significant migration edges was then inferred from the second-order rate of change in likelihood (Δm) across incremental values of m with the *OptM* package in R (Fitak, 2021).

In addition to network analyses and as a complementary result, a Maximum Likelihood (ML) tree was built using the concatenated SNP dataset (2867 bp). Prior to concatenation, the combined VCF file was filtered using BCFtools (Li, 2011) to remove individuals with more than 10% of missing data ($N=8$) and candidate SNP outliers ($N=18$, see section 2.7). The filtered VCF file was then converted to PHYLIP format using the *vcf2phylip.py* script (Ortiz, 2019). Consensus sequences for each population were estimated with the function *consensusString* in the R package ‘Biostrings’. The two Cabrera localities were considered as a single population ($F_{st}=0.03$) and data merged. A ML tree was then built on the consensus alignment with IQ-TREE 2 v2.2.0.8 (Minh et al., 2020) using variable sites only and applying an ascertainment bias correction for SNP data (model GTR+ASC) (Lewis, 2001), with 10000 pseudo-replicates.

To identify the most ancestral population in our dataset we used IQ-TREE 2 with non-reversible substitution models (model 12.12) (Naser-Khdour et al., 2022) with 1,000 ultrafast bootstrap replicates, using both the consensus alignment as well as one random specimen per population (all sites or only variables). The program performs a bootstrap analysis to obtain several ML rooted bootstrap trees; it then computes *rootstrap* support values for each branch in the tree, as the proportion of rooted bootstrap trees that have the root on that branch. A root testing was then performed with option --root-test to compare the log-likelihoods of the trees being rooted on every branch of the ML tree.

The resulting trees were visualized and edited with Figtree v1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

2.7. Identification and functional annotation of SNPs outliers

To detect footprint of genomic diversification and potential selection, outlier loci were identified using two independent approaches: PCAdapt software v 4.3.5 (Luu et al., 2017) and BayeScan v 2.1 (Foll and Gaggiotti, 2008). The PCAdapt package implemented in R software detects outlier loci based on PCA by assuming that markers excessively related to population structure are candidates for potential adaptation. The number of PCs to retain was chosen after checking score plots for population structuring, setting $maf=0.01$ and distance to 'mahalanobis'. The distribution of loadings (SNP contribution to each PC) was uniform, indicating no relevant LD effect. P-values were corrected for false discovery rate (FDR) using a cut-off $q < 0.01$ for outlier retention. BayeScan is designed to detect potential genetic loci under selection by analyzing variations in allele frequencies among specific groups with a Multinomial-Dirichlet model. The prior odd (PO) for neutrality indicates the ratio of selected:neutral sites (e.g., 1:1000) and provides a measure of uncertainty on the likelihood of the neutral model compared to the selection model (Lotterhos and Whitlock, 2014). The sensitivity of the analysis to the PO was evaluated using alternative values (1:100, 1:10000). The final MCMC chain was run for 20 short pilot runs with 5000 integrations, 50000 burn-in, thinning interval of 10, and PO set to 100. Loci were filtered for $q\text{-value} < 0.01$.

PCAdat and Bayescan results were finally intersected, retaining outliers with $F_{st} > 0.8$, thus providing a conservative set of candidate loci. The above analysis was performed separately for each dataset (GBS, RADseq and COMBINED data).

The candidate SNP outliers were annotated by cross-referencing the SNP position against the GFF file of the *rPodLil1.1* genome assembly (Gomez-Garrido et al., 2023) for Gene ID association. For outliers falling in protein-coding region, a Gene Ontology (GO) annotation was performed, followed by a functional enrichment analysis with gGOST in g:Profiler [<https://biit.cs.ut.ee/gprofiler/gost>] on individual datasets (GBS and RADSeq). The *Podarcis muralis* genome was used as a reference to determine the functional categories (Biological processes (BP), molecular functions (MF) and cellular components (CC)) that were significantly enriched ($FDR < 0.05$). For the subset of outliers falling within coding regions (CDS), we annotated the codon position and assessed whether the

alternative allele (ALT) translated into a synonymous or nonsynonymous substitution with respect to the reference position in the genome (REF).

2.8. Outlier association with environmental variables

A potential association between outliers and geographic/environmental variables (Major Archipelagos, Islet surface, Maximum altitude and SDI) was assessed through a permutational multivariate analysis of variance (PERMANOVA) with the function *adonis2* in the 'vegan' R package (Oksanen et al., 2022). A Euclidean distance matrix among individuals was built on allele frequencies calculated with *gl.percent.freq* in the 'dartR' package (Gruber et al., 2018). Previously, missing genotype values were imputed using the function 'gl.impute' in the dartR package. Model selection was performed by first fitting a model with all two-way interactions and all main terms, and then refitting the model after excluding non-significant interactions. Significance of model terms was assessed using marginal tests (1000 permutations). SNPs that significantly discriminated between major archipelagos were identified with logistic regression models using $p < 0.01$ after adjusting p-values to minimize false discovery rate (FDR). A genotype matrix coded as 0/1/2 was obtained using function *extract.gt* in the 'vcfR' R package (Knaus and Grünwald, 2017, 2016) and used to visualize geographic clustering of individual SNP outliers by a Principal Components Analysis (PCA).

3. Results

We generated a total of 15 billion paired-end GBS raw reads for 100 specimens from eight distinct localities/islands (99% passing the filters, Table S2). Newly generated data and previous RADSeq data from Bassitta et al. (2021) were individually processed through the same pipeline and combined for an integrative analysis of all 16 populations. After genome mapping, variant calling and filtering, we obtained a total of 143272 biallelic SNPs for the GBS dataset (100 individuals, 8 localities), 166943 for the RADseq dataset (91 individuals, 10 localities) and 5523 for the COMBINED dataset (191 individuals, 16 localities, including only common loci between GBS and RADSeq datasets) (Table 2). As expected, multiallelic variants represented a minor fraction of the total polymorphic sites (2.1% for GBS and 2.5% for RADSeq) and were excluded from the study.

Table 2. Summary of raw and filtered SNP count per dataset.

Dataset	Study	N ind	N pop	Raw	SNP count	
					Filtered ^a	LD pruned ^b
GBS	This study	100	8	1.888.392	143272	34065
RADSeq	Bassitta <i>et al.</i> 2021	91	10	4.851.070	166943	58207
COMBINED	Both studies	191	16	6.394.354	5523	2876

^a Filters: --max-alleles 2, --remove-indels, --min-meanDP 10, --max-meanDP 50, --minQ 30, --maf 0.05, --max-missing 0.75 (individual datasets) or 0.90 (combined)

^b After Linkage disequilibrium (LD): --indep-pairwise 50 ['kb'] 5 0.20

Both raw and filtered SNPs were widely distributed across the 18 autosomal and two sexual chromosomes (Z and W), with SNP frequencies being proportional to chromosome size (Figure 2; chromosome W representation strictly depends on the number of males in the dataset, see Table S1). Both sequencing methods covered the entire genome diversity, with chromosome representation being highly congruent among the three datasets (Figure 2).

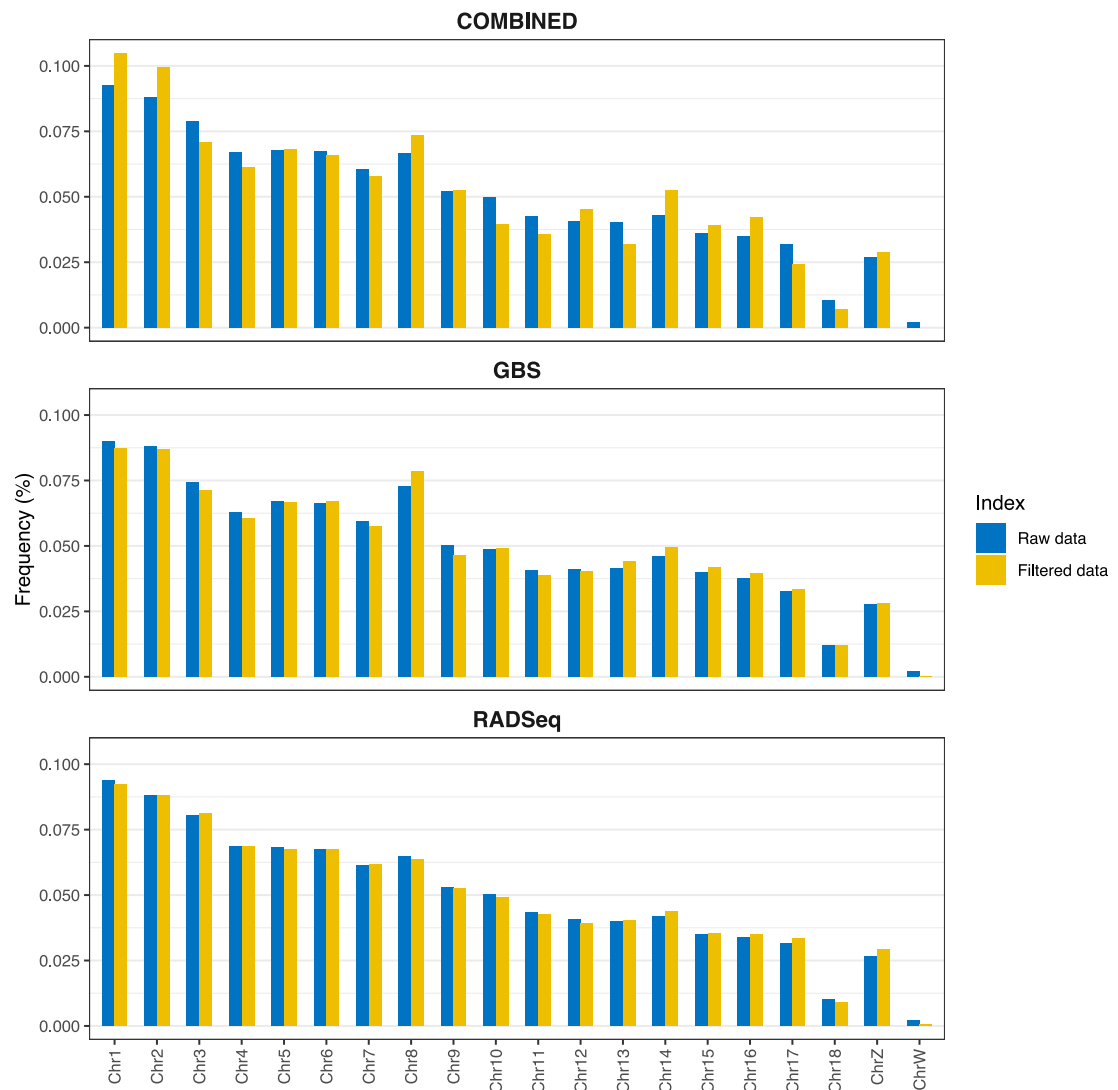


Figure 2. Chromosome distribution of raw and filtered SNP data for each of the three datasets (COMBINED, GBS and RADSeq).

3.1. Population/Islet genetic diversity

Genetic diversity based on the COMBINED dataset is summarized in Figure 3. While diversity statistics for the individual datasets (GBS and RADSeq) was slightly higher in the absolute estimates per locality (Table S3), the relative pattern found among localities was highly preserved, indicating that the reduced COMBINED dataset is overall representative of the relative genetic diversity within this species.

Population nucleotide diversity (P_i) ranged between 0.104 and 0.174, with the three archipelagos showing similar ranges: 0.123- 0.156 (Cabrera), 0.109-0.174 (Mallorca) and 0.104-0.158 (Menorca) (Figure 3). The observed heterozygosity (H_o) showed comparable values as P_i , varying between 0.095 and 0.140. Most diverse populations/islets per archipelago were the two Cabrera localities (Cabrera), Na Guardia and Na Moltón (Mallorca), and Aire and Colom (Menorca); the least diverse populations were the small islets of Porros (Menorca), Colomer (Mallorca) and Esclatasang (Cabrera) (Figure 3). F_{is} values were positive for all populations and range from 0.009 (Porros) to 0.107 (Moltón), indicating low inbreeding effect despite the small size of some islands. Individual datasets (GBS and RADSeq) based on a more extensive number of SNPs confirmed low F_{is} values and major trend among populations (Table S3). The Allelic richness (A_r) was also comparable across localities, ranging between 1.245 (Porros) and 1.438 (Dragonera). The number of private alleles (PA) estimated on individual datasets (Table S3) vary greatly, with highest values found in the four smallest islets of Porros, En Curt, Esclatasang and Foradada ($n > 1000$, island area < 1.50 ha), with the small islet of En Curt presenting the largest number ($n=5316$) (Table S3).

Of the genetic diversity indexes, A_r , P_i and F_{is} showed positive coefficients with the islet surface and negative with maximum altitude in multiple regression models (A_r : $p < 0.001$ only for Area, P_i : $p = 0.001$ for Area and $p < 0.005$ for Max altitude, F_{is} : $p < 0.05$ for Area and Max altitude), indicating that the smallest islets typically host a more reduced genetic diversity, although no evident inbreeding depression.

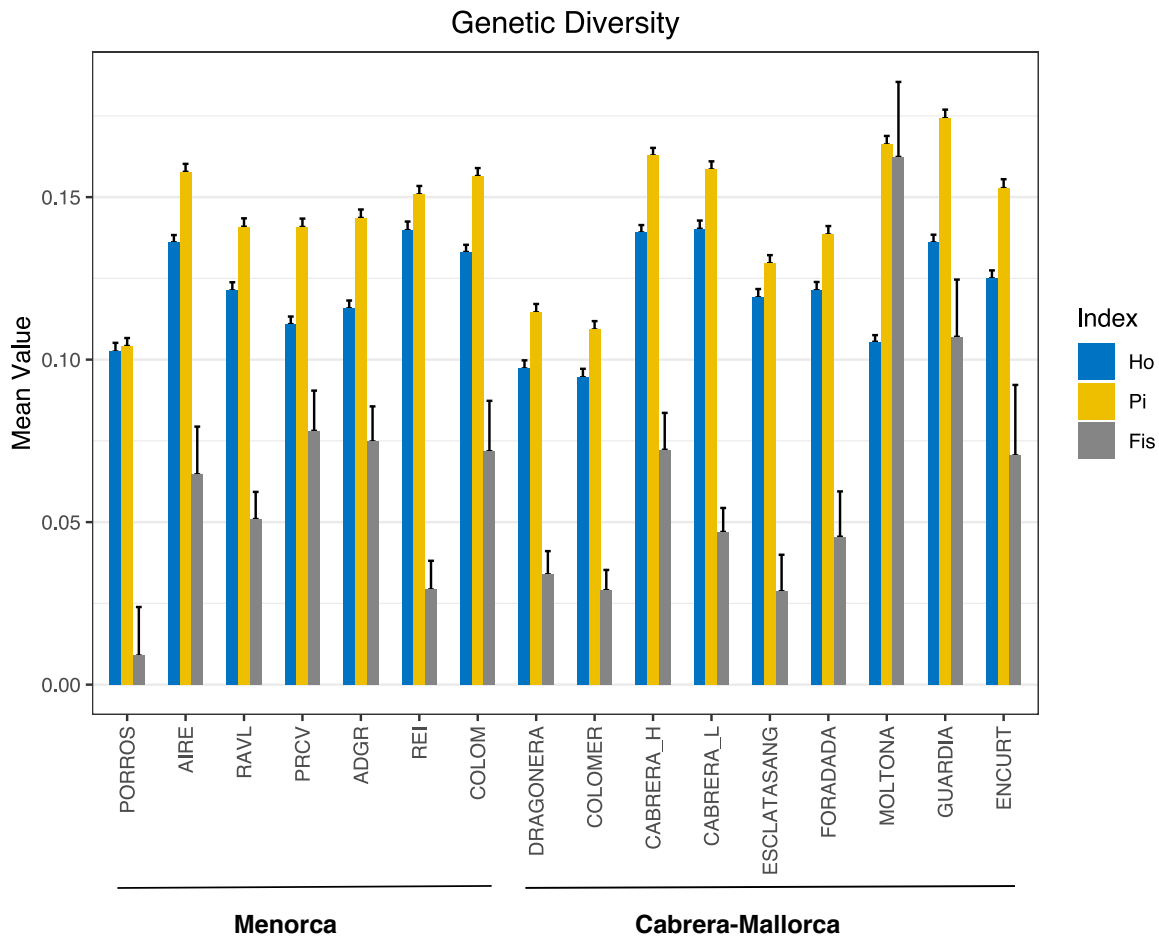


Figure 3. Genetic diversity estimates per population (Pi: nucleotide diversity, Ho: Observed heterozygosity, and Fis: inbreeding coefficient) with corresponding error bars, based on 5523 SNPs (COMBINED dataset).

3.2. Population genetic differentiation and structure

Pairwise F_{st} values (Weir & Cockerham, 1984) estimated for the combined dataset (2876 SNPs after LD pruning, Table 2) indicated that all populations were significantly differentiated ($p < 0.01$ for all pair-wises) (Figure 4). Global levels of differentiation among islet populations were generally high, with a mean $F_{st} = 0.247$ (ranging between 0.030 and 0.473). These values were highly comparable to those estimated on individual datasets, indicating that this reduced SNP dataset is also representative of the genetic differentiation among the studied populations.

The most discriminated population was the small islet of En Curt (Mallorca), showing an average F_{st} of 0.338 with respect to all other islands (minimum value with Moltona, F_{st} =0.24, Figure 4). The least differentiated populations were Aire and Colom (0.047, Menorca), Na Guardia and Na Moltona (0.064, Mallorca), and the two Cabrera localities (0.030). In this latter case, F_{st} value <0.05 supports the existence of a unique panmictic population within the major Cabrera Island.

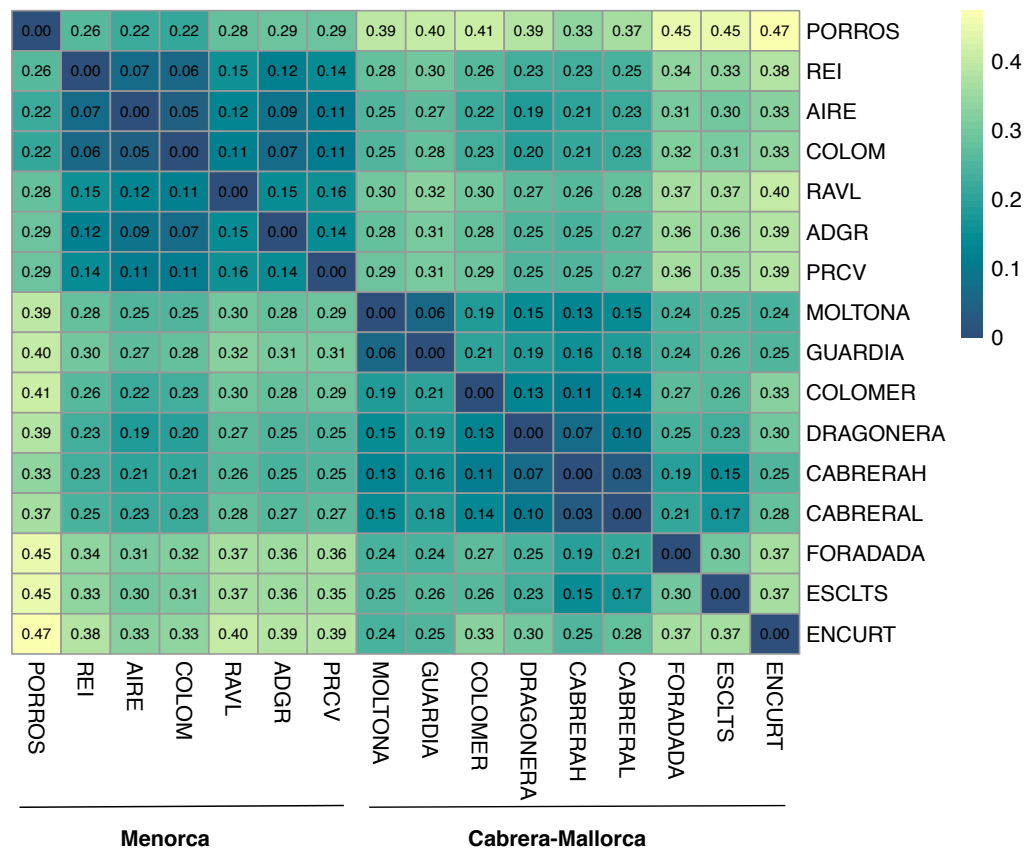
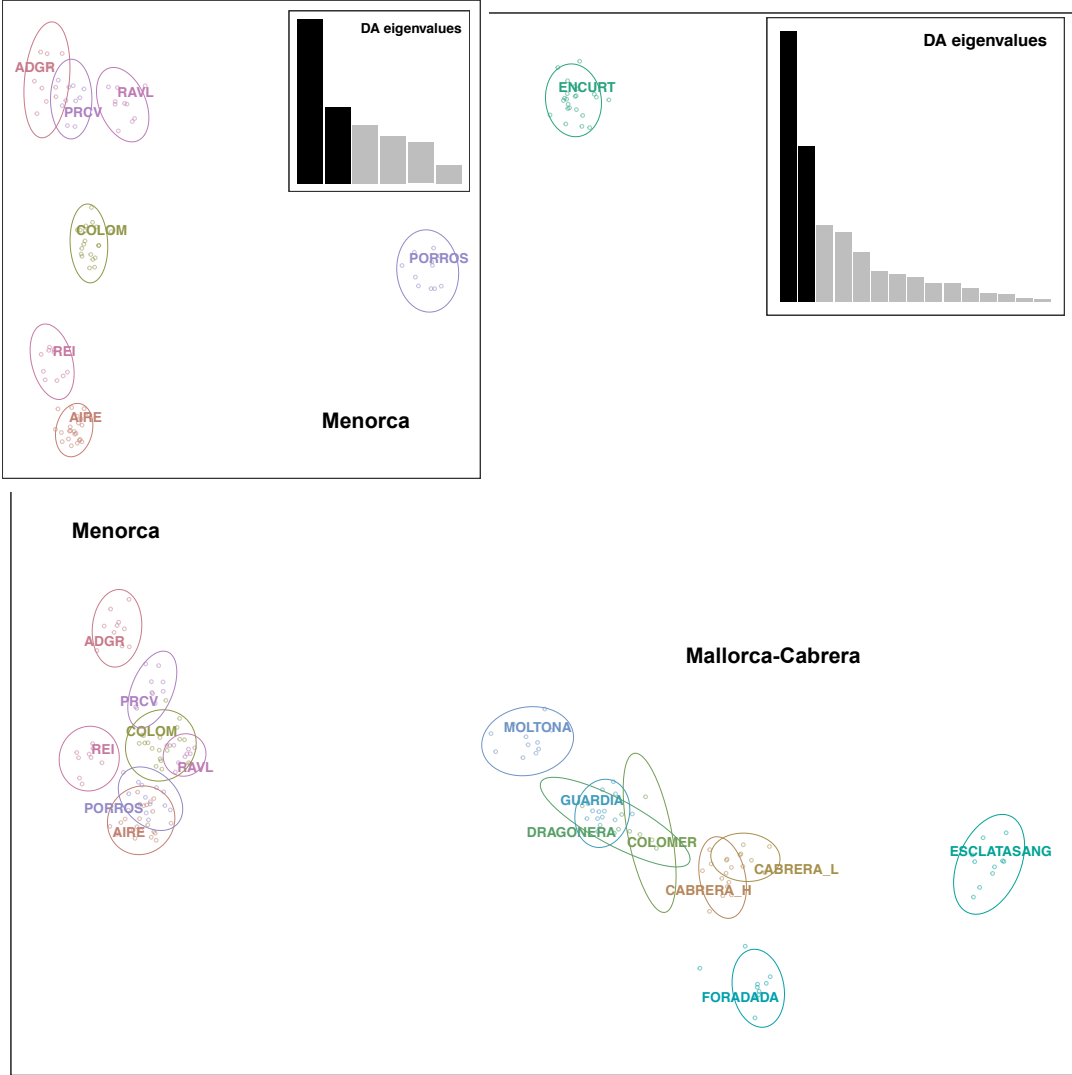


Figure 4. Heatmap of pairwise genetic differentiation (F_{st}) between the 16 populations of *P. lilfordi* (COMBINED dataset, 2876 SNPs). All F_{st} estimates were significant ($p < 0.01$ after permutations).

Within major archipelagos, no significant correlation was found between genetic (F_{st}) and geographic distances ($p > 0.1$), excluding a scenario of isolation-by-distance.

Population genetic clustering with DAPC (Figure 5A) was consistent with F_{st} -based distances (Figure 4) and the geographic distribution of these populations (Figure 1). Membership probabilities according to the retained discriminant function (64 PCs) correctly assigned all individuals to their source population/islet, excluding recent translocations or mislabeling. According to the first Discriminant Function (DF) (35% of variance), three major genetic clusters were identified (Figure 5A): I) the group of Menorca islands, highly homogeneous and with poor internal discrimination, II) the group of Cabrera and Mallorca islands, characterized by a stronger genetic differentiation among populations, and III) the islet of En Curt, forming an independent lineage. DAPC on the subset of Menorcan populations only (30 PCs retained) slightly increased the population resolution, highlighting a clear divergence of the islet of Porros (Figure 5A, top left panel).

A)



B)

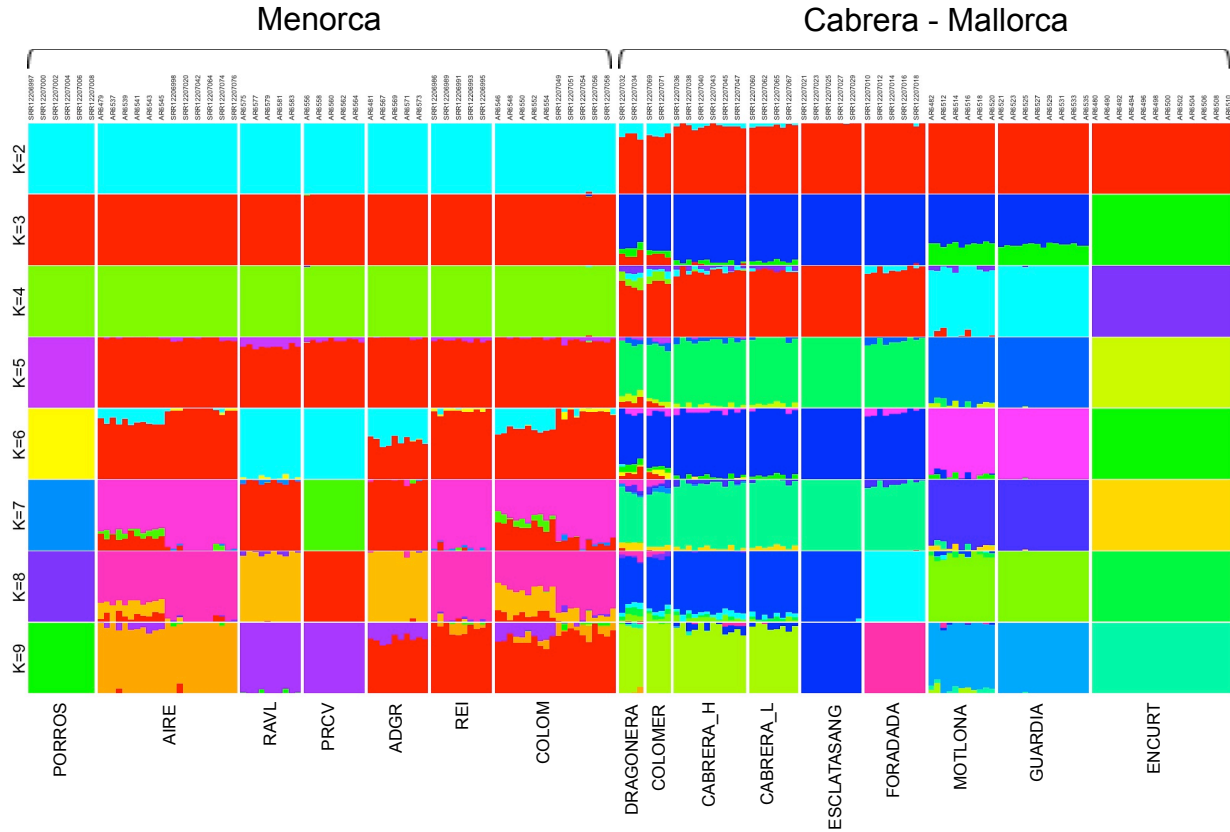


Figure 5. Population genetic structure based on 2876 SNPs (COMBINED dataset, after LD). A) DAPC retaining the first 64 PCs (65% of the explained variance). B) Unsupervised clustering by ADMIXTURE depicting K =2 to K=9 (best value) clusters.

ADMIXTURE analyses were used to further explore population structure, as well as shared ancestry (Figure 5B). Genetic clustering clearly evidenced a major discrimination between Menorca and Cabrera/Mallorca populations (K=2), followed by progressive substructuring within Cabrera/Mallorca (from K=3), and ultimately within Menorca (from K=5). The best-supported number of present genetic clusters is 9 (CV=0.386), corresponding to four genetic clusters in Menorca and five in Mallorca/Cabrera, in line with the major structure recovered by DAPC analysis (Figure 5A). The proportion of shared ancestry among populations varies from zero (independent lineages of the islets

of En Curt, Porros, Foradada and Esclatasang) up to 40% (the highly admixed population of Colom). Cabrera island and distant islands from North-West of Mallorca (Dragonera and Colomer) formed a unique genetic cluster, with a small proportion of shared ancestries also with the small islets of Foradada and Esclatasang (North and South of Cabrera, respectively) (see map in Figure 1). According to the optimal current grouping ($K=9$), no shared ancestry was detected between the two major archipelagos of Menorca and Mallorca/Cabrera, indicating a strong genetic differentiation. However, lower substructuring, from $K=2$ to $K=9$ retrieved a clear signature of shared ancestry between the Northern Mallorca populations of Dragonera and Colomer with the Menorcan populations of Colom, Rei and Aire.

3.3. Phylogeographic scenario

Prior to phylogenetic reconstruction, we assessed past migration events among populations using allele frequencies on Treemix and OptM analysis to support the best number of edges. OptM analysis indicated a single migration edge explaining 98.9% of the variance. The event with the greatest weight occurred from Colomer into Cabrera-Escatlasang (mean migration weight: 0.454032, mean p-value: $3.22478e-11$), followed by a migration event from Dragonera into Escatlasang-Cabrera (mean migration weight: 0.380029, mean p-value: $6.71e-05$) (Figure S1).

A midpoint rooted ML tree was then built using consensus sequences for each population on the concatenated SNP dataset (Figure 6). The phylogenetic tree supports the major grouping obtained by the ADMIXTURE, Fst and DAPC analyses, largely in line with the population geographic distribution (Figure 1). Menorcan and Cabrera/Mallorca populations formed clearly distinct phylogenetic clades (bootstrap, bs =100%). Within the Mallorca/Cabrera clade, further substructuring indicated three well differentiated clades: i) the southern Mallorca islets of En Curt, Na Moltana and Na Guardia (bs= 99%), ii) Cabrera islets of Foradada and Esclatasang (bs=76%), with Cabrera as a basal lineage, and iii) the two northern Mallorcan islands of Colomer and Dragonera (bs=80%) (Figure 1 and Figure 6). Within the Menorca clade, substantial support is given only to the grouping of Porros de Cavalleria (PRCV) and Revells (RAVL) (bs=73%), while the small islet of Porros de Fornells (PORROS) represents an independent lineage, highly differentiated from to the rest of Menorcan populations (bs=89%). The remaining

Menorcan populations were poorly resolved, consistently with their lower F_{st} values (Figure 4) and higher admixture levels (Figure 5B).

Tree rooting using non-reversible models on IQ-tree consistently identified the population of Colomer (Mallorca) as the most likely ancestral lineage. The island showed the highest rootstrap value (86,13%, after 10K bootstrapping on consensus sequences, all sites; values <2% for all other branches) (Figure S2). This result was robust to the use of consensus sequences with variant sites only and to subsampling to one random specimen per population, all consistently recovering the highest rootstrap value for Colomer (71.35 and 69.54, respectively). Root testing provided the maximum p-value of the approximately-unbiased (AU) test for the branch leading to Colomer ($p=0.85$). All branches within the Menorca clade showed a $p<0.05$, significantly excluding Menorca as a potential source of the ancestral population. Rootstrap values and root testing considering only the archipelago of Menorca did not provide any clear root support within Menorca.

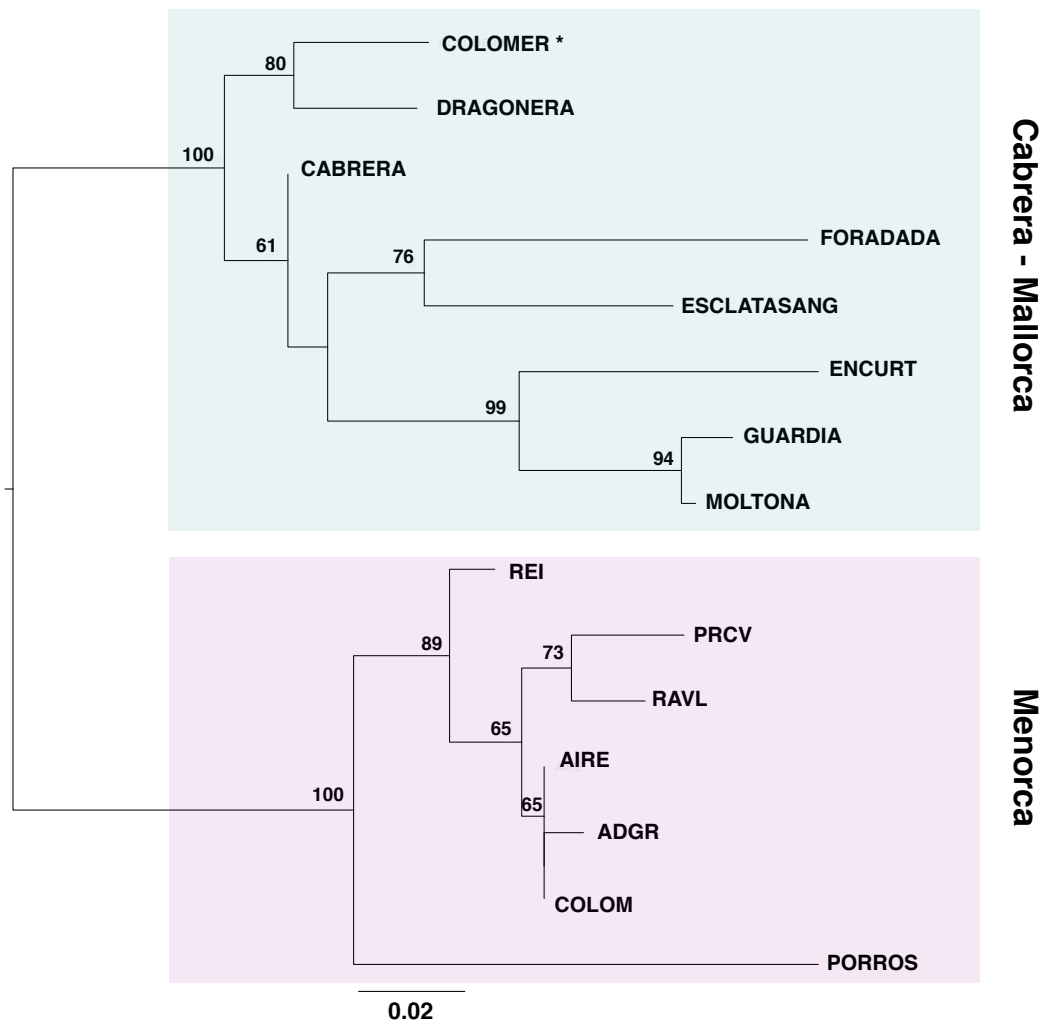


Figure 6. Midpoint-rooted ML tree based on consensus sequences per population after SNP concatenation (2876 SNPs, only variable sites, model GTR+ASC, 10,000 bootstrap replicates). Node numbers indicate bootstrap support. The asterisk indicates the most probable ancestral population according to IQ-TREE 2 root analysis.

3.4. Outlier identification

We explored potential selective pressures driving population genomic differentiation by identification of outlier SNPs. Intersection of outputs from two independent methods (PCAdapt and Bayescan) provided a total of 18 candidate outliers for the COMBINED dataset, 222 for the GBS and 190 for the RADSeq dataset ($n_{\text{total}} = 446$, $q < 0.01$ and $F_{\text{ST}} > 0.8$, see Table S4 for a complete list).

Pattern of distribution of outliers per chromosome was highly concordant among the three datasets, with most outliers falling within the sexual chromosome Z (>16% for all datasets), followed by a large representation on chromosome 2 (>9%) and 15 (12 and 8%, respectively) (Figure 7A). Outliers in chromosomes 3 and 4 were detected only by RADSeq.

About 60% of the outliers fell within protein-coding genes (44% for the COMBINED) corresponding to a total of 182 proteins (Figure 7A, pie charts). Only a minor proportion fell within CDS ($n = 14$ outliers, corresponding to 11 annotated proteins). Five of these SNPs caused a nonsynonymous substitution (ALT) with respect to the reference genome (REF) and were detected in the glypican-3, B-cell lymphoma 6 protein, the cytochrome P450 2C19-like, the pore complex Nup214 genes and an uncharacterized protein (see details in Table S4).

Functional enrichment analysis of annotated proteins with outliers indicated a comparable profile for GBS and RADSeq, showing a significant overrepresentation of binding proteins and catalytic activity (MF category), and metabolic process, growth, and cellular homeostasis (BP category) (BH, $p < 0.001$, Figure S3).

Of the overall 182 proteins with outliers, a small subset ($n = 13$) was consistently retrieved by more than one dataset and most ($n = 9$) were target of multiple outliers per dataset (up to seven) (Figure 7B). Three of these proteins were ion channels: the calcium channels inositol 1,4,5-trisphosphate receptor type 2 and the voltage-dependent L-type calcium channel subunit alpha-1C-like, and the potassium voltage-gated channel subfamily KQT member 1-like. Two proteins presented an outlier position identified by more than one dataset: the cullin-1 protein (SNP Position POS:937545) detected by all three datasets, and the voltage-dependent L-type calcium channel subunit alpha-1C-like (POS:55226525) detected by both GBS and COMBINED (Table S4).

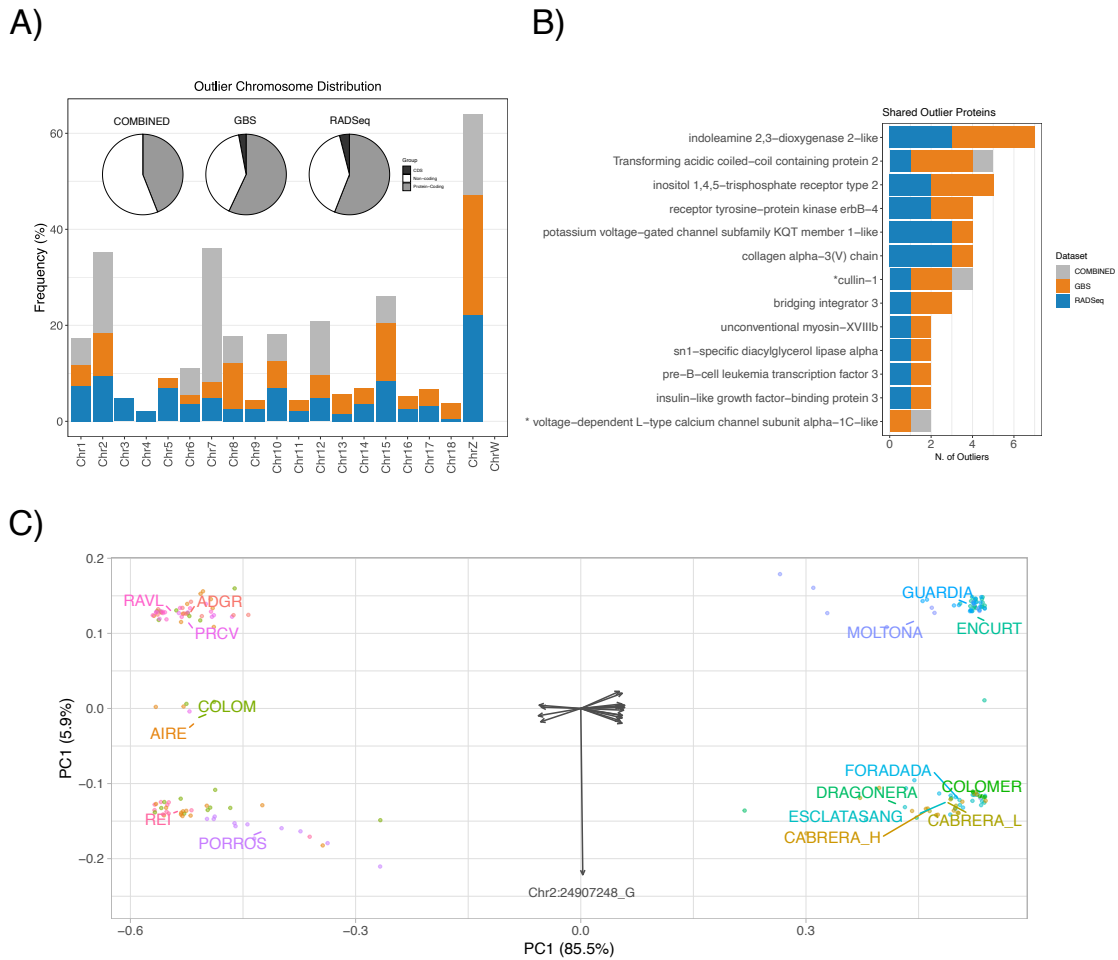


Figure 7. Candidate SNP outliers identified for GBS (n= 222), RADSeq (n=190), and COMBINED (n=18). A) Outliers chromosome distribution (barplot) and proportion per gene category (coding, CDS and non-coding regions) (pie chart). B) Proteins with outliers detected by multiple datasets and corresponding number of outliers. The asterisk indicates proteins with shared SNP outliers among datasets. C) Distance biplot of a Principal Component Analysis (PCA) based on the outlier genotype distance matrix for the COMBINED dataset (18 SNPs). Circles are specimens plot to optimally reproduce their Euclidean distances; arrows are unscaled SNPs scores showing their contribution to each principal component.

3.5. Major Archipelagos-segregating SNPs

Finally, we explored environmental and geographic variables driving outlier distribution. PERMANOVA analyses indicated that Major Archipelago (Menorca and Mallorca/Cabrera, Table 1) was the only explanatory variable for all three datasets ($p < 0.001$, 1000 permutations), with all individual outliers, except one, significantly contributing to geographic between these two major phylogenetic lineages (Figure 6) (binomial test $p < 0.01$). PCA based on outlier genotype matrix clearly evidenced this major clustering (85.5% of variance explained for the COMBINED dataset, Figure 7C, see also Figure S4 for individual outlier genotype clustering). Only one SNP departed from this trend (SNP Chr2_24907248), recovering a clustering of South Mallorca with a subset of the Menorcan islands (Figure 7C and Figure S4). The SNP falls within the sidekick2 protein (Table S4).

A high percentage of outliers had allelic variants that perfectly segregated between major archipelagos, i.e., had null frequency in one archipelago while being present in the other (n total = 154; 11% for COMBINED, 42% for GBS and 31% for RADSeq) (Table S5). These fixed allelic differences can be considered as specific markers of the two major phylogenetic clades (Figure 6).

4. Discussion

4.1. Genetic diversity of *P. lilfordi*: evolution in islands and population conservation

A fundamental goal in population ecology and evolution is to understand the processes that maintain genetic diversity, and those that drive intraspecific/interpopulation divergence across geographic space and time (Avise, 2000). In small insular populations, free of predators and with no immigrants, genetic diversity is primarily driven by genetic drift (due to environmental and demographic stochasticity, both particularly important in tiny islets) and density-dependent selection due to competition for the limited local resources (Hoffmann et al., 2021; Hunt et al., 2022; Travis et al., 2023). Founder effect and high inbreeding levels are also expected to further reduce the population genetic variance and accelerate the process of divergence from the original source population (Keller and Waller, 2002). Each *P. lilfordi* island clearly hosted a distinct population (all F_{st} distances were significant), with no contemporary gene flow (no recent translocations detected). Few small islets showed a remarkably high differentiation from all other islets (i.e. Porros and En Curt, $F_{st} > 0.2$ for all pairs, see Figure 4), indicating either a strong founder effect and/or intense local drift and selection due to their reduced island size (Sendell-Price et al., 2021). Despite being effectively closed, *P. lilfordi* populations harbored relatively high genetic diversity ($H_o > 0.09$ and $P_i > 0.1$ for all populations), even higher than other continental and insular species of *Podarcis* (Sabolić and Štambuk, 2021). Diversity was overall largely comparable among all 16 populations studied and across their range of distribution (Figure 1 and Figure 3); nonetheless, association analyses supported the expected pattern of decreasing genetic diversity with decreasing island size (Furlan et al., 2012). Smaller islands also presented the highest number of private alleles, indicative of an ongoing process of differentiation. Mallorca/Cabrera archipelago typically hosted a higher genetic diversity than Menorca (Figure 3), with the islands of Cabrera, Na Guardia and Na Moltona (South of Mallorca) currently harboring the largest genetic diversity present in this species.

As previously reported by Bassitta et al. (2021), we also observed low levels of inbreeding for most populations ($F_{is} < 0.08$), even for tiny islets such as Esclatasang, Foradada, Porros, Revells and En Curt (all with Area < 0.7 ha, $F_{is} < 0.5$, Table 1 and Figure 3). While small populations are expected to show a reduced heterozygosity (Keller and Waller, 2002), F_{is} values in *P. lilfordi* indicated a slight trend of positive increase with increasing

islet surface, with the highest estimate found for the island of Moltona ($F_{is}=0.16$). A small inbreeding coefficient, negatively associated with island size, was also found in a previous study based on microsatellites from the three neighboring populations of Moltona, Na Guardia and En Curt (Rotger et al., 2021). A potential explanation is the existence of a population spatial substructure within larger islands, which would cause a non-random mating of individuals effectively reducing the observed heterozygosity. Nonetheless, we invite caution in interpretation of these results as SNP-based estimation of the inbreeding depression in wild populations is notoriously challenging (Schmidt et al., 2021) and would require pedigree-studies for a reliable quantification (Kardos et al., 2016).

Both the relatively high genetic diversity and low F_{is} values observed in *P. lilfordi* would suggest the existence of potential mechanisms for buffering inbreeding depression in these lizard populations, increasing their chances of persistence. In line with this, recent studies support the ability of insular lizards to counteract genetic depletion even in presence of a strong founder effect (see Sherpa et al., 2023).

4.2. A phylogeographic scenario of *P. lilfordi* colonization of the Balearic Islands

P. lilfordi genetic diversity was largely structured according to the main geographic distribution, with all analyses supporting the existence of two major genetic lineages separating the large archipelagos of Menorca from Mallorca/Cabrera (Figure 5 and 6). The tree topology was largely consistent with the ones previously reported based on mtDNA and SNPs data (Bassitta et al., 2021; Brown et al., 2008; Pérez-Cembranos et al., 2020; Terrasa et al., 2009). While Mallorca/Cabrera lineage is further sub-structured in four well-supported clades ($bs>80\%$, Figure 6), phylogenetic relationships among Menorcan populations remain largely unresolved, as also indicated by their low degree of differentiation (Figure 4, F_{st} values < 0.1). Menorcan islands were once largely panmictic (Brown et al., 2008; Terrasa et al., 2009) and a signature of past admixture can still be observed among most populations of this archipelago (Figure 5B), with the exclusion of the small islet of Porros, now representing an independent lineage (Figure 6).

Admixture analysis ($K=2$, Figure 5B), root inference by IQ-tree (Figure S2) and root testing all strongly suggest that the origin of the species occurred in the main island of Mallorca, with most ancestral populations in our dataset being identified in the North of Mallorca. The highest rootstrap support (86.13%) specifically points to the small island of Colomer as the population retaining nowadays most ancestral polymorphisms, in line with

previous proposals (Bassitta et al., 2021; Terrasa et al., 2009). This root placement is consistent with the geographic history of both Dragonera and Colomer islands which were once connected to the Serra of Tramuntana, a mountain chain running up North of Mallorca, before their detachment from the main Mallorca Island about 2.3 Mya (Brown et al., 2008; Terrasa et al., 2009). The Serra of Tramuntana is considered to be the last refuge of *P. lilfordi* (Bailón, 2004) before its extinction in the main Mallorca Island, as testified by fossil records dating around 2000 years old found in the Muleta cave, in the central part of the Tramuntana (Alcover, 2000; Kotsakis, 1981). The Colomer islet is characterized by high lizard density (Pérez-Mellado et al., 2008) and limited accessibility (due to high altitude and coastal shape). Although introduction of lizards from other islands cannot be ruled out, the islet has reduced chances of external gene flow, compatible with the retention of ancestral polymorphisms (Bassitta et al., 2021; Terrasa et al., 2009).

The following process of colonization within the archipelago of Mallorca/Cabrera did not follow an isolation by distance model (Mantel test, $p > 0.1$). According to admixture, phylogenetic analyses and Treemix reconstructions of past migration events, North of Mallorca was then the source of colonization of main Cabrera Island, with which they form a unique genetic clade (see Admixture results, Figure 5B). This closed genetic relationship was also reported by Bassitta et al., 2021, although it partly differs from mtDNA results (Terrasa et al., 2009).

Following this scenario, species expansion within the archipelago of Cabrera proceeded with the colonization of small surrounding islets (Esclatasang and Foradada), as supported by the observed coancestry with Cabrera Island (see Figure 5B). This suggests a founder effect in these small islands, followed by independent evolution leading to the nowadays well-differentiated genetic clade of Esclatasang and Foradada (Figure 5 and 6). Similarly, admixture and ML tree analyses indicated that islets on North of Cabrera could have been a source of colonization for Southern Mallorca islets, specifically the islet of Moltona, with which they still share a small proportion of coancestry (Figure 5B and 6), while the geographically closed sister population of Na Guardia has currently lost all signature of coancestry with Northern Cabrera (Figure 5B). The observed clustering of Cabrera with South Mallorca islets confirms previous haplotype grouping based on mtDNA data, which also suggested a directionality of gene flow from Cabrera to South Mallorca (Network III; Terrasa et al., 2009). Moltona would have then seeded the small

islet of En Curt, with which they form a unique clade (bs=99%, Figure 6), as evidenced by their shared coancestry (Figure 5B). The tiny islet of En Curt (0,44 ha) would have followed a process of accelerated differentiation presumably driven by a founder effect, extensive genetic drift, and intense density-dependent selection with more than 1500 ind/ha (Ruiz De Infante Anton et al., 2013). As an alternative scenario, the three islets on the South of Mallorca directly derived from the ancestral mainland Mallorca population, now extinct, a scenario that we presently cannot exclude.

According to bathymetric, geological and genetic data subsequent colonization of Menorca occurred around ~ 2.8 Mya (Brown et al., 2008; Terrasa et al., 2009). We were not able to assess the root within this archipelago due to high levels of admixture. Menorcan populations were highly panmictic during a large period of time and the colonization and subsequent isolation of most Menorcan islands was gradual (Pretus et al., 2004), following a vicariance type of colonization, with an asymmetric flow from south to north of the distribution of the species (Terrasa et al., 2009).

4.3. Signature of genomic diversification of *P. lilfordi* populations

Local adaptation in effectively closed populations is a trade-off between genetic drift and selective pressures (Savolainen et al., 2013). The two processes are intrinsically associated with the island surface and its exposure to open sea, both affecting the impact of environmental stochasticity (stronger in smaller islets) and resource availability (reduced in smaller islands). Particularly, the vegetation cover, highly correlated to the island size, provides a strong cascade effect on lizard resource availability, including pollen and fruits production, insect visitor frequency and diversity, and even seabirds presence, all potential sources of dietary items (Pérez-Mellado et al., 2008; Ruiz De Infante Anton et al., 2013; Salvador, 2009; Santamaría et al., 2020). Along with stochasticity, this gives rise to a heterogeneous landscape across islands, which is expected to drive independent processes of lizard phenotypic diversification.

We explored the signature of genomic diversification that might underpin this lizard phenotypic diversity through outlier analysis. Outlier distribution showed a clearly skewed chromosome representation, with most outliers falling within the sexual chromosome Z, consistently for all datasets analyzed (Figure 7A). Recent studies in birds and lizards showed that closely related species often present high differentiation on the Z chromosome (Kulikova et al., 2022; Rovatsos et al., 2019), a pattern that is typically explained as faster Z evolution and lower recombination rates (Irwin, 2018; Lima, 2014;

Mank et al., 2010; Wright et al., 2015). The existence of potential "islands of differentiation" within the sexual chromosomes has also been hypothesized (Lavretsky et al., 2019). At present, detection of such "islands" in *P. lilfordi* Z chromosome would require a higher sequencing coverage than the one currently provided by GBS and RADseq (i.e., genome resequencing).

Most outliers fell within protein-coding genes, although the large majority were intronic (Figure 7A). Substantial evidence supports the notion that introns have a crucial and evolutionarily conserved function in controlling gene expression in eukaryotes (Kumari et al., 2022; Rose, 2019), providing a particularly rapid mechanism for increasing variation in proteome variance by production of a diverse array of alternative splicing variants (AS) (Reixachs-Solé and Eyras, 2022; Wang et al., 2015). Moreover, most genes with outliers were associated with protein binding and catalytic activities and were involved in metabolic and growth processes (Figure S3), critical molecular functions for gene expression regulation and phenotypic diversification (Van Nostrand et al., 2020). In small, isolated populations, extensive modulation of gene expression conferring phenotypic plasticity could represent a major mechanism to counteract the loss of genetic diversity (Fulgione et al., 2023; Sherpa et al., 2023), a hypothesis that needs to be validated by whole transcriptome and epigenome data (Chapelle and Silvestre, 2022; Fulgione et al., 2023).

We found a subset of protein-coding genes targets of multiple outliers according to both independent sequencing methods, GBS and RADSeq (Figure 7B), which might represent interesting candidates for further exploration of population genomic diversification in *P. lilfordi*. Of these proteins, few represented potassium and calcium channels, which could hint to an important regulatory role of osmotic pressure in lizards (Dantzler and Bradshaw, 2008). We also highlight the cullin-1 protein, with a shared outlier SNP detected by all datasets. The protein is known to be involved in ubiquitination and subsequent proteasomal degradation of target proteins (Duan et al., 2011; Gao et al., 2011; Scott et al., 2016). Recent evidence indicates a critical role of this protein in modulation of the transcription factor c-MYC, a major regulator of gene expression and cell proliferation (Sweeney et al., 2020). While there are no specific studies on cullin-1 protein in lizards, c-MYC was linked to the cellular regenerative response after tail amputation (Alibardi, 2017). The gradual loss of tail autotomy ability in insular lizards is a hallmark of their

reduced antipredator response following insular adaptation (part of the island syndrome) (Cooper et al., 2004; Pafilis et al., 2009; Perez-Mellado, 1997).

Exploration of the major variables driving outlier genotypes primarily recovered a signature of past geographic separation between Menorca and Cabrera/Mallorca, with a large proportion of loci presenting fixed allelic differences between archipelagos (Figure 7C). These SNPs are putatively derived from a past founder effect that occurred during initial colonization of the Menorca archipelago (Brown et al., 2008; Terrasa et al., 2009) or the result of unclear selective forces, including potential selective sweeps (Brown et al., 2023; Campagna et al., 2022; Stephan, 2019). In all cases, they support no recent secondary contact between Archipelagos, in line with previous mtDNA-based phylogeographic reconstructions (Terrasa et al., 2009). We note that given the major confounding effect of archipelago, the significance/impact of other environmental variables in driving genome diversification cannot be reliably assessed with the current sampling design (Bassitta et al. 2021). Additional sampling, along with phenotypic data are required to clarify the putative adaptive role of these genomic changes.

Conclusions

The Balearic lizard *P. lilfordi* has recently entered the genomic era, offering a new interesting insular system to explore trends in vertebrate genome evolution. Our study presents a first use of the *P. lilfordi* genome to reliably mapped extensive SNP data. Comparative analysis using independent methods and datasets provided robustness to our findings. We found support for the origin of this species colonization in the Mallorca Island, confirming previous proposals and providing a framework for future studies of the evolutionary trajectories in genome diversification for this insular species. The substantial genetic diversity observed in these effectively closed populations suggest that they have potential mechanisms to partially counteract genetic drift and inbreeding depression that are worth further investigation by genome resequencing and inclusion of additional populations. Furthermore, integration of population demographic data (ongoing) and collection of detailed phenotypic data, still scarce for this species, will be critical to evaluate the genomic plasticity and ability to persist of these insular populations.

Data availability

The raw data with the individual sequences are available at the Sequence Read Archive (SRA) (BioProject ID: PRJNA1070579).

Acknowledgments

This study was supported by the CAIB - Government of the Balearic Islands (project PRD2018/25 to G.T.), the Institut d'Estudis Catalans under the Catalan Biogenome Project initiative (PRO2020-S02 to L.B.), the MCIN/AEI/10.13039/501100011033/FEDER, UE (PID2022-141578NB-C22 to L.B.) and the Doctoral Scholarship, Colombian Ministry of Science, Technology and Innovation (MINCIENCIAS885/2020 to K.O)

References

1. Alcover, J.A., 2000. Vertebrate Evolution and Extinction on Western and Central Mediterranean Islands. *Tropics* 10, 103–123. <https://doi.org/10.3759/tropics.10.103>
2. Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. <https://doi.org/10.1101/gr.094052.109>
3. Alibardi, L., 2017. Immunolocalization of c-myc-positive cells in lizard tail after amputation suggests cell activation and proliferation for tail regeneration. *Acta Zool.* 98, 114–124. <https://doi.org/10.1111/azo.12153>
4. Avise, J.C., 2000. *Phylogeography: The History and Formation of Species*. Harvard University Press. <https://doi.org/10.2307/j.ctv1nzfjg7>
5. Bailón, S., 2004. Fossil records of Lacertidae in mediterranean islands: the state of the art. Pp. 37-62. En: Pérez-mellado, V., Riera, N., Perera, A. (Eds.). *The biology of Lacertid lizards. Evolutionary and Ecological Perspectives*. Institut Menorquí d'Estudis. Recerca, 8. Maó.
6. Bassitta, M., Brown, R.P., Pérez-Cembranos, A., Pérez-Mellado, V., Castro, J.A., Picornell, A., Ramon, C., 2021. Genomic signatures of drift and selection driven by predation and human pressure in an insular lizard. *Sci. Rep.* 11, 6136. <https://doi.org/10.1038/s41598-021-85591-x>
7. Bloor, P., Rodríguez, V., Terrasa, B., Brown, R.P., Pérez-Mellado, V., Castro, J.A., Picornell, A., Ramon, M.M., 2011. Polymorphic microsatellite loci for the Balearic Island Lizard *Podarcis lilfordi* (Squamata: Lacertidae). *Conserv. Genet. Resour.* 3, 323–325. <https://doi.org/10.1007/s12686-010-9354-7>
8. Brown, R.P., Sun, H., Jin, Y., Meloro, C., 2023. Habitat-associated Genomic Variation in a Wall Lizard from an Oceanic Island. *Genome Biol. Evol.* 15, evad193. <https://doi.org/10.1093/gbe/evad193>
9. Brown, R.P., Terrasa, B., Pérez-Mellado, V., Castro, J.A., Hoskisson, P.A., Picornell, A., Ramon, M.M., 2008. Bayesian estimation of post-Messinian divergence times in Balearic Island lizards. *Mol. Phylogenet. Evol.* 48, 350–358. <https://doi.org/10.1016/j.ympev.2008.04.013>
10. Camargo, A., Sinervo, B., Sites, J.W., 2010. Lizards as model organisms for linking phylogeographic and speciation studies: LIZARD PHYLOGEOGRAPHY AND

- SPECIATION. *Mol. Ecol.* 19, 3250–3270. <https://doi.org/10.1111/j.1365-294X.2010.04722.x>
11. Campagna, L., Mo, Z., Siepel, A., Uy, J.A.C., 2022. Selective sweeps on different pigmentation genes mediate convergent evolution of island melanism in two incipient bird species. *PLOS Genet.* 18, e1010474. <https://doi.org/10.1371/journal.pgen.1010474>
 12. Chapelle, V., Silvestre, F., 2022. Population Epigenetics: The Extent of DNA Methylation Variation in Wild Animal Populations. *Epigenomes* 6, 31. <https://doi.org/10.3390/epigenomes6040031>
 13. Charlesworth, B., Charlesworth, D., Barton, N.H., 2003. The Effects of Genetic and Geographic Structure on Neutral Variation. *Annu. Rev. Ecol. Evol. Syst.* 34, 99–125. <https://doi.org/10.1146/annurev.ecolsys.34.011802.132359>
 14. Cooper, W.E., Pérez-Mellado, V., 2012. Historical influence of predation pressure on escape by Podarcis lizards in the Balearic Islands: PREDATION PRESSURE AND ESCAPE. *Biol. J. Linn. Soc.* 107, 254–268. <https://doi.org/10.1111/j.1095-8312.2012.01933.x>
 15. Cooper, W.E., Pérez-Mellado, V., Vitt, L.J., 2004. Ease and effectiveness of costly autotomy vary with predation intensity among lizard populations. *J. Zool.* 262, 243–255. <https://doi.org/10.1017/S095283690300462X>
 16. Cuerda, J.B., 1989. Los tiempos cuaternarios en Baleares. Conselleria de Cultura, Educació i Esports, Direcció General de Cultura, Govern Balear.
 17. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., 1000 Genomes Project Analysis Group, 2011. The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
 18. Dantzler, W. H, Bradshaw, S. D., 2008. 10 Osmotic and Ionic Regulation in Reptiles. *Osmotic and ionic regulation: cells and animals*, 443.
 19. De la Cruz, O., Raska, P., 2014. Population structure at different minor allele frequency levels. *BMC Proc.* 8, S55. <https://doi.org/10.1186/1753-6561-8-S1-S55>
 20. Drake, D.R., Mulder, C.P.H., Towns, D.R., Daugherty, C.H., 2002. The biology of insularity: an introduction. *J. Biogeogr.* 29, 563–569. <https://doi.org/10.1046/j.1365-2699.2002.00706.x>
 21. Dray, S., Dufour, A.-B., 2007. The ade4 Package: Implementing the Duality Diagram for Ecologists. *J. Stat. Softw.* 22. <https://doi.org/10.18637/jss.v022.i04>

22. Duan, S., Skaar, J.R., Kuchay, S., Toschi, A., Kanarek, N., Ben-Neriah, Y., Pagano, M., 2011. mTOR Generates an Auto-Amplification Loop by Triggering the β TrCP- and CK1 α -Dependent Degradation of DEPTOR. *Mol. Cell* 44, 317–324. <https://doi.org/10.1016/j.molcel.2011.09.005>
23. Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E., 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLOS ONE* 6, e19379. <https://doi.org/10.1371/journal.pone.0019379>
24. Fitak, R.R., 2021. *OptM*: estimating the optimal number of migration edges on population trees using *Treemix*. *Biol. Methods Protoc.* 6, bpab017. <https://doi.org/10.1093/biomethods/bpab017>
25. Flesch, E.P., Graves, T.A., Thomson, J.M., Proffitt, K.M., White, P.J., Stephenson, T.R., Garrott, R.A., 2020. Evaluating wildlife translocations using genomics: A bighorn sheep case study. *Ecol. Evol.* 10, 13687–13704. <https://doi.org/10.1002/ece3.6942>
26. Foll, M., Gaggiotti, O., 2008. A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics* 180, 977–993. <https://doi.org/10.1534/genetics.108.092221>
27. Fulgione, D., Maselli, V., Riviuccio, E., Aceto, S., Salvemini, M., Buglione, M., 2023. Evolutionary Plasticity in Insular Lizard, Adapting over Reproduction, Metabolism, and Color Variation. *Biology* 12, 1478. <https://doi.org/10.3390/biology12121478>
28. Furlan, E., Stoklosa, J., Griffiths, J., Gust, N., Ellis, R., Huggins, R.M., Weeks, A.R., 2012. Small population size and extremely low levels of genetic diversity in island populations of the platypus, *Ornithorhynchus anatinus*: Low Genetic Diversity in Island Platypuses. *Ecol. Evol.* 2, 844–857. <https://doi.org/10.1002/ece3.195>
29. Gao, D., Inuzuka, H., Tan, M.-K.M., Fukushima, H., Locasale, J.W., Liu, P., Wan, L., Zhai, B., Chin, Y.R., Shaik, S., Lyssiotis, C.A., Gygi, S.P., Toker, A., Cantley, L.C., Asara, J.M., Harper, J.W., Wei, W., 2011. mTOR Drives Its Own Activation via SCF β TrCP-Dependent Degradation of the mTOR Inhibitor DEPTOR. *Mol. Cell* 44, 290–303. <https://doi.org/10.1016/j.molcel.2011.08.030>
30. Gomez-Garrido, J., Cruz, F., Alioto, T.S., Feiner, N., Uller, T., Gut, M., Sanchez Escudero, I., Tavecchia, G., Rotger, A., Otalora Acevedo, K.E., Baldo, L., 2023. Chromosome-level genome assembly of Lilford's wall lizard, *Podarcis lilfordi* (Günther, 1874) from the Balearic Islands (Spain) (preprint). *Genomics*. <https://doi.org/10.1101/2023.02.15.528647>

31. Goudet J, Jombart, T., Kamvar, Z., Archer, E., Hardy, O., 2015. hierfstat: estimation and tests of hierarchical F-statistics. R Package Version 0.04-22. Available <https://CRAN.R-project.org/package=hierfstat>.
32. Goy, J.L.G., Zazo, C., Barceló, J.C., Goy, J.L.G., Zazo, C., Barceló, J.C., 1997. Evolución de las áreas margino-litorales de la costa de Mallorca (I. Baleares) durante el Ultimo y Presente Ingerglacial: Nivel del mar Holoceno y Clima. *Bol. Geológico Min.* 108, 127–136.
33. Gràcia Lladó, F., Vicens Xamena, D., 1998. Aspectes geomorfològics quaternaris del litoral de Mallorca, in: *Aspectes geològics de les Balears: Mallorca, Menorca i Cabrera*, 1998, ISBN 84-7632-368-9, págs. 307-329. Presented at the *Aspectes geològics de les Balears: Mallorca, Menorca i Cabrera, Servei de Publicacions i Intercanvi Científic*, pp. 307–329.
34. Grant, A., Benton, T.G., 2000. ELASTICITY ANALYSIS FOR DENSITY-DEPENDENT POPULATIONS IN STOCHASTIC ENVIRONMENTS. *Ecology* 81, 680–693. [https://doi.org/10.1890/0012-9658\(2000\)081\[0680:EAFDDP\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[0680:EAFDDP]2.0.CO;2)
35. Gruber, B., Unmack, P.J., Berry, O.F., Georges, A., 2018. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol. Ecol. Resour.* 18, 691–699. <https://doi.org/10.1111/1755-0998.12745>
36. Günther, A., 1874. Description of a new European species of Zootoca. *Ann. Mag. Nat. Hist.* 14, 158–159.
37. Hawlena, D., Pérez-Mellado, V., Cooper, J., 2009. Morphological traits affect escape behaviour of the Balearic lizards (*Podarcis lilfordi*). *Amphib.-Reptil.* 30, 587–592. <https://doi.org/10.1163/156853809789647167>
38. Hoffmann, A.A., Miller, A.D., Weeks, A.R., 2021. Genetic mixing for population management: From genetic rescue to provenancing. *Evol. Appl.* 14, 634–652. <https://doi.org/10.1111/eva.13154>
39. Hunt, D.A.G.A., DiBattista, J.D., Hendry, A.P., 2022. Effects of insularity on genetic diversity within and among natural populations. *Ecol. Evol.* 12, e8887. <https://doi.org/10.1002/ece3.8887>
40. Irwin, D.E., 2018. Sex chromosomes and speciation in birds and other ZW systems. *Mol. Ecol.* 27, 3831–3851. <https://doi.org/10.1111/mec.14537>

41. Jombart, T., 2008. *adeigenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
42. Jombart, T., & Collins, C., 2015. A tutorial for discriminant analysis of principal components (DAPC) using *adeigenet* 2.0.0. *London Imperial College London MRC Cent. Outbreak Anal. Model.*
43. Kardos, M., Taylor, H.R., Ellegren, H., Luikart, G., Allendorf, F.W., 2016. Genomics advances the study of inbreeding depression in the wild. *Evol. Appl.* 9, 1205–1218. <https://doi.org/10.1111/eva.12414>
44. Keller and Waller, L., 2002. Inbreeding effects in wild populations. *Trends Ecol. Evol.* 17, 230–241. [https://doi.org/10.1016/S0169-5347\(02\)02489-8](https://doi.org/10.1016/S0169-5347(02)02489-8)
45. Knaus, B.J., Grünwald, N.J., 2017. *vcfr*: a package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* 17, 44–53. <https://doi.org/10.1111/1755-0998.12549>
46. Knaus, B.J., Grünwald, N.J., 2016. *Vcfr*: a package to manipulate and visualize VCF format data in R (preprint). *Bioinformatics*. <https://doi.org/10.1101/041277>
47. Kotsakis, T., 1981. Le Lucertole (Lacertidae, Squamata) del Pliocene, Pleistocene e Olocene delle Baleari. *Bolletí de la Societat d'Historia Natural de les Balears*, 25: 135–150.
48. Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., Wilson, D.S., 1999. Chronology, causes and progression of the MSC. *Nature* 400, 652–655. <https://doi.org/10.1038/23231>
49. Kulikova, I.V., Shedko, S.V., Zhuravlev, Y.N., Lavretsky, P., Peters, J.L., 2022. Z-chromosome outliers as diagnostic markers to discriminate Mallard and Chinese Spot-billed Duck (Anatidae). *Zool. Scr.* 51, 401–414. <https://doi.org/10.1111/zsc.12540>
50. Kumari, A., Sedehizadeh, S., Brook, J.D., Kozlowski, P., Wojciechowska, M., 2022. Differential fates of introns in gene expression due to global alternative splicing. *Hum. Genet.* 141, 31–47. <https://doi.org/10.1007/s00439-021-02409-6>
51. Lavretsky, P., DaCosta, J.M., Sorenson, M.D., McCracken, K.G., Peters, J.L., 2019. ddRAD-seq data reveal significant genome-wide population structure and divergent genomic regions that distinguish the mallard and close relatives in North America. *Mol. Ecol.* 28, 2594–2609. <https://doi.org/10.1111/mec.15091>

52. Le Galliard, J.F., Marquis, O., Massot, M., 2010. Cohort variation, climate effects and population dynamics in a short-lived lizard. *J. Anim. Ecol.* 79, 1296–1307. <https://doi.org/10.1111/j.1365-2656.2010.01732.x>
53. Lewis, P.O., 2001. A Likelihood Approach to Estimating Phylogeny from Discrete Morphological Character Data. *Syst. Biol.* 50, 913–925. <https://doi.org/10.1080/106351501753462876>
54. Li, H., 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>
55. Lima, T.G., 2014. Higher levels of sex chromosome heteromorphism are associated with markedly stronger reproductive isolation. *Nat. Commun.* 5, 4743. <https://doi.org/10.1038/ncomms5743>
56. Losos, J.B., Ricklefs, R.E., 2009. Adaptation and diversification on islands. *Nature* 457, 830–836. <https://doi.org/10.1038/nature07893>
57. Lotterhos, K.E., Whitlock, M.C., 2014. Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Mol. Ecol.* 23, 2178–2192. <https://doi.org/10.1111/mec.12725>
58. Luu, K., Bazin, E., Blum, M.G.B., 2017. *pcadapt*: an R package to perform genome scans for selection based on principal component analysis. *Mol. Ecol. Resour.* 17, 67–77. <https://doi.org/10.1111/1755-0998.12592>
59. Manes, S., Costello, M.J., Beckett, H., Debnath, A., Devenish-Nelson, E., Grey, K.-A., Jenkins, R., Khan, T.M., Kiessling, W., Krause, C., Maharaj, S.S., Midgley, G.F., Price, J., Talukdar, G., Vale, M.M., 2021. Endemism increases species' climate change risk in areas of global biodiversity importance. *Biol. Conserv.* 257, 109070. <https://doi.org/10.1016/j.biocon.2021.109070>
60. Mank, J.E., Vicoso, B., Berlin, S., Charlesworth, B., 2010. EFFECTIVE POPULATION SIZE AND THE FASTER-X EFFECT: EMPIRICAL RESULTS AND THEIR INTERPRETATION. *Evolution* 64, 663–674. <https://doi.org/10.1111/j.1558-5646.2009.00853.x>
61. Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27, 209–220.
62. Massot, M., Clobert, J., Pilorge, T., Lecomte, J., Barbault, R., 1992. Density Dependence in the Common Lizard: Demographic Consequences of a Density Manipulation. *Ecology* 73, 1742–1756. <https://doi.org/10.2307/1940026>

63. Mayol, J., Bibiloni, G., Pons, G., Ballesteros, E., Díaz, D., 2020. Descripció General de les Petites illes i els illots de les Balears.
64. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* 37, 1530–1534. <https://doi.org/10.1093/molbev/msaa015>
65. Naser-Khdour, S., Quang Minh, B., Lanfear, R., 2022. Assessing Confidence in Root Placement on Phylogenies: An Empirical Study Using Nonreversible Models for Mammals. *Syst. Biol.* 71, 959–972. <https://doi.org/10.1093/sysbio/syab067>
66. Novosolov, M., Raia, P., Meiri, S., 2013. The island syndrome in lizards. *Glob. Ecol. Biogeogr.* 22, 184–191. <https://doi.org/10.1111/j.1466-8238.2012.00791.x>
67. Oksanen, J. et al., 2022. vegan: Community Ecology Package. R Package Version 45-3 <https://CRAN.R-project.org/package=vegan>
68. Ortiz, 2019. vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis. <https://zenodo.org/record/2540861>.
69. Pafilis, P., Foufopoulos, J., Poulakakis, N., Lymberakis, P., Valakos, E.D., 2009. TAIL SHEDDING IN ISLAND LIZARDS [LACERTIDAE, REPTILIA]: DECLINE OF ANTIPREDATOR DEFENSES IN RELAXED PREDATION ENVIRONMENTS. *Evolution* 63, 1262–1278. <https://doi.org/10.1111/j.1558-5646.2009.00635.x>
70. Pembleton, L.W., Cogan, N.O.I., Forster, J.W., 2013. St AMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol. Ecol. Resour.* 13, 946–952. <https://doi.org/10.1111/1755-0998.12129>
71. Pérez-Cembranos, A., Pérez-Mellado, V., Alemany, I., Bassitta, M., Terrasa, B., Picornell, A., Castro, J.A., Brown, R.P., Ramon, C., 2020. Morphological and genetic diversity of the Balearic lizard, *Podarcis lilfordi* (Günther, 1874): Is it relevant to its conservation? *Divers. Distrib.* 26, 1122–1141. <https://doi.org/10.1111/ddi.13107>
72. Perez-Mellado, V., C. Corti, and P. Lo Cascio., 1997. Tail autotomy and extinction in Mediterranean lizards. A preliminary study of continental and insular populations. *J. Zool.* 243:553–541. *J Zool* 243553–541.
73. Pérez-Mellado, V., Martín-Vallejo, J., Brown, R., Picornell, A., Castro, J., Ramón, M.M., Terrassa, B., García-Díez, T., Hernández-Estévez, J.Á., 2008. Population density in *Podarcis lilfordi* (Squamata, Lacertidae), a lizard species endemic to small islets in the Balearic Islands (Spain). *Amphib.-Reptil.* 29, 49–60. <https://doi.org/10.1163/156853808783431587>

74. Pickrell, J.K., Pritchard, J.K., 2012. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLOS Genet.* 8, e1002967. <https://doi.org/10.1371/journal.pgen.1002967>
75. Pretus, J., Marques, R, Pérez-Mellado, V., 2004. Holocene sea level rise and range fragmentation of *Podarcis lilfordi* on Minorcan islets: a vicariance scenario reviewed through a mtDNA tree. *Pérez-Mellado V, Riera N, Perera A, (eds). The Biology of lacertid lizards: Evolutionary and Ecological Perspectives. Institut Menorquí d'Estudis, 8, 279-291.*
76. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., De Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* 81, 559–575. <https://doi.org/10.1086/519795>
77. Raivo Kolde, 2019. pheatmap: Pretty Heatmaps. R package Version 1.0.12.
78. Reixachs-Solé, M., Eyra, E., 2022. Uncovering the impacts of alternative splicing on the proteome with current omics techniques. *WIREs RNA* 13, e1707. <https://doi.org/10.1002/wrna.1707>
79. Rose, A.B., 2019. Introns as Gene Regulators: A Brick on the Accelerator. *Front. Genet.* 9, 672. <https://doi.org/10.3389/fgene.2018.00672>
80. Rotger, A., Igual, J.M., Genovart, M., Rodríguez, V., Ramon, C., Pérez-Mellado, V., Bibiloni, G., Rita, J., Tavecchia, G., 2021. Contrasting Adult Body-Size in Sister Populations of the Balearic Lizard, *Podarcis lilfordi* (Günther 1874) Suggests Anthropogenic Selective Pressures. *Herpetological Monographs*, 35(1), 53-64. <https://doi.org/10.1655/HERPMONOGRAPHS-D-19-00005>
81. Rotger, A., Igual, J.M., Smith, J.J., Tavecchia, G., 2016. Relative role of population density and climatic factors in shaping the body growth rate of Lilford's Wall Lizard (*Podarcis lilfordi*). *Can. J. Zool.* 94, 207–215. <https://doi.org/10.1139/cjz-2015-0188>
82. Rotger, A., Tenan, S., Igual, J., Bonner, S., Tavecchia, G., 2023. Life span, growth, senescence and island syndrome: Accounting for imperfect detection and continuous growth. *J. Anim. Ecol.* 92, 183–194. <https://doi.org/10.1111/1365-2656.13842>
83. Rovatsos, M., Reháč, I., Velenský, P., Kratochvíl, L., 2019. Shared Ancient Sex Chromosomes in Varanids, Beaded Lizards, and Alligator Lizards. *Mol. Biol. Evol.* 36, 1113–1120. <https://doi.org/10.1093/molbev/msz024>
84. Ruiz De Infante Anton, J., Rotger, A., Igual, J.M., Tavecchia, G., 2013. Estimating lizard population density: an empirical comparison between line-transect and capture–recapture methods. *Wildl. Res.* 40, 552. <https://doi.org/10.1071/WR13127>

85. Sabolić, Štambuk, 2021. Genomic patterns of phenotypic adaptation in Italian wall lizard *Podarcis siculus* (Rafinesque-Schmaltz, 1810). Dr. Diss. Univ. Zagreb Fac. Sci. Dep. Biol. 2021.
86. Salvador, A., 2009. Lagartija balear – *Podarcis lilfordi*. Encicl. Virtual Los Vertebr. Esp. Museo Nacional de Ciencias Naturales, Madrid. *Enciclopedia Virtual de los Vertebrados Españoles*.1, 3-22.
87. Santamaría, S., Enoksen, C.A., Olesen, J.M., Tavecchia, G., Rotger, A., Igual, J.M., Traveset, A., 2020. Diet composition of the lizard *Podarcis lilfordi* (Lacertidae) on 2 small islands: an individual-resource network approach. *Curr. Zool.* 66, 39–49. <https://doi.org/10.1093/cz/zoz028>
88. Savolainen, O., Lascoux, M., Merilä, J., 2013. Ecological genomics of local adaptation. *Nat. Rev. Genet.* 14, 807–820. <https://doi.org/10.1038/nrg3522>
89. Schmidt, T.L., Jasper, M., Weeks, A.R., Hoffmann, A.A., 2021. Unbiased population heterozygosity estimates from genome-wide sequence data. *Methods Ecol. Evol.* 12, 1888–1898. <https://doi.org/10.1111/2041-210X.13659>
90. Scott, D.C., Rhee, D.Y., Duda, D.M., Kelsall, I.R., Olszewski, J.L., Paulo, J.A., De Jong, A., Ovaa, H., Alpi, A.F., Harper, J.W., Schulman, B.A., 2016. Two Distinct Types of E3 Ligases Work in Unison to Regulate Substrate Ubiquitylation. *Cell* 166, 1198-1214.e24. <https://doi.org/10.1016/j.cell.2016.07.027>
91. Sendell-Price, A.T., Ruegg, K.C., Robertson, B.C., Clegg, S.M., 2021. An island-hopping bird reveals how founder events shape genome-wide divergence. *Mol. Ecol.* 30, 2495–2510. <https://doi.org/10.1111/mec.15898>
92. Sherpa, S., Paris, J.R., Silva-Rocha, I., Di Canio, V., Carretero, M.A., Ficetola, G.F., Salvi, D., 2023. Genetic depletion does not prevent rapid evolution in island-introduced lizards. *Ecol. Evol.* 13, e10721. <https://doi.org/10.1002/ece3.10721>
93. Sivaperuman, C., Jaisankar, I., Velmurugan, A., Sharma, T.V.R.S., 2008. Tropical Islands: Ecosystem and Endemism, in: *Biodiversity and Climate Change Adaptation in Tropical Islands*. Elsevier, pp. 31–52. <https://doi.org/10.1016/B978-0-12-813064-3.00002-8>
94. Stephan, W., 2019. Selective Sweeps. *Genetics* 211, 5–13. <https://doi.org/10.1534/genetics.118.301319>
95. Sweeney, M.A., Iakova, P., Maneix, L., Shih, F.-Y., Cho, H.E., Sahin, E., Catic, A., 2020. The ubiquitin ligase Cullin-1 associates with chromatin and regulates transcription of

- specific c-MYC target genes. *Sci. Rep.* 10, 13942. <https://doi.org/10.1038/s41598-020-70610-0>
96. Terrasa, B., Pérez-Mellado, V., Brown, R.P., Picornell, A., Castro, J.A., Ramon, M.M., 2009. Foundations for conservation of intraspecific genetic diversity revealed by analysis of phylogeographical structure in the endangered endemic lizard *Podarcis lilfordi*. *Divers. Distrib.* 15, 207–221. <https://doi.org/10.1111/j.1472-4642.2008.00520.x>
 97. Travis, J., Bassar, R.D., Coulson, T., Reznick, D., Walsh, M., 2023. Density-Dependent Selection. *Annu. Rev. Ecol. Evol. Syst.* 54, 85–105. <https://doi.org/10.1146/annurev-ecolsys-110321-055345>
 98. Van der Auwera GA & O'Connor BD., 2020. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra* (1st Edition). O'Reilly Media.
 99. Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R., Blue, S.M., Chen, J.-Y., Cody, N.A.L., Dominguez, D., Olson, S., Sundararaman, B., Zhan, L., Bazile, C., Bouvrette, L.P.B., Bergalet, J., Duff, M.O., Garcia, K.E., Gelboin-Burkhart, C., Hochman, M., Lambert, N.J., Li, H., McGurk, M.P., Nguyen, T.B., Palden, T., Rabano, I., Sathe, S., Stanton, R., Su, A., Wang, R., Yee, B.A., Zhou, B., Louie, A.L., Aigner, S., Fu, X.-D., Lécuyer, E., Burge, C.B., Graveley, B.R., Yeo, G.W., 2020. A large-scale binding and functional map of human RNA-binding proteins. *Nature* 583, 711–719. <https://doi.org/10.1038/s41586-020-2077-3>
 100. Wang, Y., Liu, J., Huang, B., Xu, Y.-M., Li, J., Huang, L.-F., Lin, J., Zhang, J., Min, Q.-H., Yang, W.-M., Wang, X.-Z., 2015. Mechanism of alternative splicing and its regulation. *Biomed. Rep.* 3, 152–158. <https://doi.org/10.3892/br.2014.407>
 101. Warren, B.H., Simberloff, D., Ricklefs, R.E., Aguilée, R., Condamine, F.L., Gravel, D., Morlon, H., Mouquet, N., Rosindell, J., Casquet, J., Conti, E., Cornuault, J., Fernández-Palacios, J.M., Hengl, T., Norder, S.J., Rijsdijk, K.F., Sanmartín, I., Strasberg, D., Triantis, K.A., Valente, L.M., Whittaker, R.J., Gillespie, R.G., Emerson, B.C., Thébaud, C., 2015. Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecol. Lett.* 18, 200–217. <https://doi.org/10.1111/ele.12398>
 102. Weir & Cockerham, 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, 38, 1358-1370.

103. Whittaker, R.J., Fernández-Palacios, J.M., Matthews, T.J., Borregaard, M.K., Triantis, K.A., 2017. Island biogeography: Taking the long view of nature's laboratories. *Science* 357, eaam8326. <https://doi.org/10.1126/science.aam8326>
104. Wingett, S.W., Andrews, S., 2018. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research* 7, 1338. <https://doi.org/10.12688/f1000research.15931.2>
105. Wright, A.E., Harrison, P.W., Zimmer, F., Montgomery, S.H., Pointer, M.A., Mank, J.E., 2015. Variation in promiscuity and sexual selection drives avian rate of Faster-Z evolution. *Mol. Ecol.* 24, 1218–1235. <https://doi.org/10.1111/mec.13113>
106. Yang, W., Feiner, N., Pinho, C., While, G.M., Kaliontzopoulou, A., Harris, D.J., Salvi, D., Uller, T., 2022. Extensive introgression and mosaic genomes of Mediterranean endemic lizards. *Nat. Commun.* 12, 2762. <https://doi.org/10.1038/s41467-021-22949-9>

Supplementary Material

Table S1: Individual sample metadata

Table S2: GBS read summary

Table S3: Genetic diversity estimates for individual datasets (GBS and RADSeq)

Table S4: List of outliers identified by both Bayescan and PCADAPT ($Q < 0.01$, $F_{st} > 0.8$) for the three datasets: Combined ($N=18$), GBS (222) and RADSEQ (190) and corresponding annotation

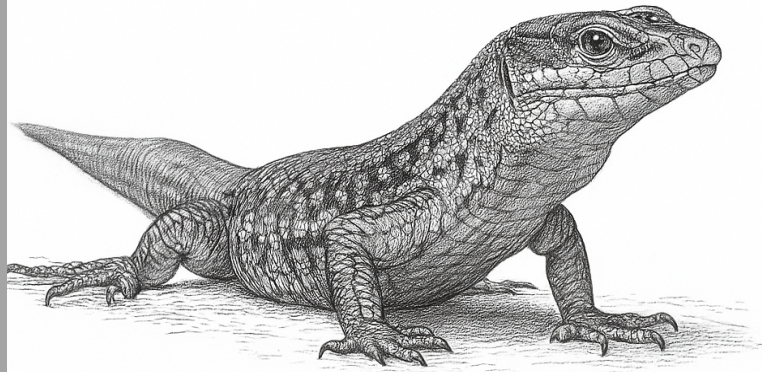
Table S5: Archipelago-segregating SNPs (fixed alleles)

Figure S1: The ML-graph inferred by TreeMix with 16 migration events based on the COMBINED dataset. The arrow indicates the two significant migration events with the highest migration weight. Drift parameter is shown on the x-axis.

Figure S2: ML tree with rootstrap support values for root inference. Values were estimated with IQ-TREE 2 on consensus sequences per population (2867 bp), using a non-reversible model (model 12.12, all sites, 10000 bootstrap replicates).

Figure S3. Functional enrichment of GOs for proteins with outliers, according to A) GBS and B) RADSeq dataset.

Figure S4. PCA of the 18 outliers detected in the COMBINED dataset. All but one indicated a clear clustering by archipelago. Black dots correspond to the most frequent allele, while Red dots indicate the heterozygotic genotype.



4. CHAPTER II

Historical demographic
inference in island populations:
the case of the Balearic lizard,
Podarcis lilfordi (Günther, 1874)

Historical demographic inference in island populations: the case of the Balearic lizard, *Podarcis lilfordi* (Günther, 1874)

Katherin Eliana A. Otalora Acevedo^{1,2,7*}, Carles Galià-Camps^{3,4,6}, Giacomo Tavecchia⁵, Andreu Rotger Vallespir⁵, Laura Baldo^{1,6}

¹Department of Evolutionary Biology, Ecology and Environmental Sciences, Biology Faculty, University of Barcelona (UB), Barcelona, Av. Diagonal 645, Barcelona 08028, Spain

² UNAD, BIOTICS, ECOGEN, Tunja, Boyacá, Colombia

³ Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Avda. Diagonal, 643, 08028 Barcelona, Spain

⁴ Centre d'Estudis Avançats de Blanes (CEAB, CSIC), Accés Cala St. Francesc 14, 17300 Blanes, Girona, Spain

⁵ Animal Demography and Ecology Unit, (GEDA-IMEDEA, CSIC -UIB), c. M. Marques, 07190, Esporles, Spain.

⁶ Institute for Research on Biodiversity (IRBio), University of Barcelona (UB), Barcelona, Spain

⁷ Fundación FUNMAJO, EBA, RAIEC, Biodiversity Branch, Cucaita, Boyacá, Colombia.

* Corresponding author:

Katherin Eliana A. Otalora Acevedo

katherinhotalora1@gmail.com

Abstract

The Balearic wall lizard (*Podarcis lilfordi*), endemic to 46 across the Mallorca, Menorca, and Cabrera archipelagos, persists in a highly fragmented environment shaped by paleoclimatic shifts, sea-level changes, and long-standing human impact. With a restricted and patchy distribution, *P. lilfordi* serves as a model for understanding how environmental changes—both ancient and recent—have influenced population persistence, genetic diversity, and evolutionary trajectories. Reconstructing its demographic history is therefore essential to reveal the underlying processes shaping its current genetic structure and to inform effective conservation strategies for island species. This study aimed to reconstruct the major demographic historical events of *P. lilfordi* using genomic data derived from single nucleotide polymorphisms (SNPs). We tackled three temporal scales: I) the general demographic reconstruction of the species, II) the early separation of the species in the two main archipelagos of Mallorca-Cabrera and Menorca, and III) the subsequent demographic events within the archipelago of Mallorca and within specific islets populations. At the islet level, we focused on three close populations under current demographic study (Na Guardia, Na Moltona, and En Curt). SNP data were analyzed using stairway plots to infer historical changes in effective population size (N_e).

Results revealed that the integration of previously described mitochondrial data with the genomic dataset from this study supported the hypothesis that divergence among the Mallorca, Menorca, and Cabrera populations began during the First Donau Glaciation (~2.3–1.8 million years ago), with subsequent cladogenesis events occurring around ~1.2 million years ago. The Riss (~300,000–130,000 years ago), Würm (~115,000–11,700 years ago), and the LGM (Last Glacial Maximum) (~33,000–19,000 years ago; the coldest stage within the last glaciation) led to substantial declines in N_e , likely due to habitat contraction and harsher environmental conditions. During interglacial periods, rising sea levels further fragmented habitats, reinforcing genetic differentiation among populations. However, transmarine dispersal events via rafting may have occasionally facilitated sporadic contact between islands. The analysis at the islet level of Na Guardia, Na Moltona and En Curt showed a marked demographic decline associated with the glacial events described above at the species level, as well as anthropogenic factors. The arrival of humans to the Balearic archipelago approximately 4,000 years ago intensified the decline of populations due to habitat alteration, the introduction of non-native

species and direct human pressure. These impacts ultimately led to the extinction of *P. lilfordi* on Mallorca and Menorca around 2,000 years ago, as well as significant demographic reductions in the three small islets analyzed in the final stages of the demographic reconstruction. Overall, these findings largely support available geological data and previous proposed scenarios for the *P. lilfordi* diversification, highlighting the intricate interplay between past climatic events and recent human-induced pressures in shaping the evolutionary trajectory of insular species. They also underscore the urgent need for conservation strategies that recognize the islets as independent genetic units to ensure the long-term viability of *P. lilfordi*.

Keywords: *Podarcis lilfordi*, Effective population size (N_e), Balearic Islands, Stairway Plot, conservation biology, Pleistocene glaciations, Transmarine dispersal.

1. Introduction

The demographic changes of lizard species of the genus *Podarcis* in the Balearic Islands represent an interesting field of study to understand how paleoclimatic events have shaped biogeography and genetic diversity in this island environment. Balearic Island lizards, *P. lilfordi*, *P. pityuensis*, as well as *P. sicula*, have been subjected to a range of evolutionary pressures. Among these pressures, outstands the significant changes in sea level during events such as the MSC 5,33 Mya (Terrasa et al., 2009; Bassita et al., 2021; Podnar et al., 2005) and several glaciations' events (i.e. Donau Glaciation (~2.3–1.8 million years ago), which have potentially influenced the distribution, genetic structure, and persistence of their populations over time (Terrasa et al., 2009; Hewitt, 2004).

The Balearic wall lizard *P. lilfordi* (Günther, A., 1874), also known as the Balearic wall lizard, is currently distributed across 43 islets in the Mallorca, Menorca, and Cabrera archipelagos (Terrasa et al., 2009). However, it is classified as an endangered species by the IUCN due to severe fragmentation of its populations and a continuous decline in its distribution area and habitat quality (Pérez-Mellado, 2004; Salvador, 2015). Several anthropogenic factors had further exacerbated the vulnerability of these populations, including the translocation of individuals by visitors between islets, illegal capture for trade, destruction of vegetation by introduced goats, and currently the introduction of predators such as the Iberian snakes *Hemorrhois hippocrepis* and feral cats (Álvarez et al., 2010; Carretero and Pinya, 2011). These pressures have also contributed to reduced

population numbers on many islets, highlighting the urgent need for conservation efforts (Pérez-Mellado, 2004; Mayol, 2004; Pérez-Cembranos et al. 2020).

Analyses of the historical and contemporary demography of *P. lilfordi* are crucial to understand its conservation status and the effects of paleoclimatic and contemporary fluctuations on the species. It is proposed that *P. lilfordi* populations may have experienced a demographic history marked by cycles of population expansion and contraction in response to changes in sea level and habitat availability. During glacial periods, such as the Riss (ca. 300,000 years ago) and the Würm (ca. 115,000 years ago), the drop in sea level may have facilitated the reconnection of some islands and the formation of saline corridors. These corridors could have enabled secondary contact between previously isolated populations, potentially leading to genetic introgression (Yang et al., 2021). Alternatively, excessive salinity in these areas might have further promoted lineage diversification (Terrasa et al., 2009). On the other hand, interglacial periods, characterized by rising sea level, likely caused habitat fragmentation and reduced connectivity between populations, fostering the formation of population refugia and genetic bottlenecks that have left a significant mark on the current genetic blueprint of the species (Hewitt, 2004). In contemporary times, the disappearance of *P. lilfordi* from the main islands of Mallorca and Menorca about 2,000 years ago (Alcover, 2000) may have had a considerable impact on the species' effective population size, potentially exacerbating the effects of bottlenecks and affecting the species' genetic dynamics (Bouzat et al., 2010; Broquet et al., 2010). Altogether these historical events, along with more recent bottlenecks experienced by present islets (Rotger et al., 2021), could have influenced the species' ability to adapt to environmental changes and its current conservation status.

Despite this geological and climatic framework, current knowledge on major demographic events occurred in the species remain patchy and largely speculative. Here, we aim to disentangle the demographic changes experienced by *P. lilfordi* populations across its entire range, with a focus on small islets characterized by limited area, and reduced population size. For this purpose, we gathered published genomic data from 191 individuals inhabiting 16 representative islets of the species' current distribution (Bassita et al., 2021 and Otalora et al., 2024) using neutral markers (SNPs) to identify both ancient and more recent demographic trends. We have carried out specific demographic stages for the entire species at three levels: at species (16 populations across the species'

distribution), archipelago (Mallorca-Cabrera and Menorca) and islets for three close populations from southern Mallorca —Na Guardia, Na Moltona, and En Curt—that belong to the same genetic clade (Terrasa et al., 2009) and are considered the same subspecies (i.e. *P. lilfordi jordansi*; Salvador 2015). Despite their geographic proximity, these populations show clear differences in body size, growth, and coloration (Rotger et al., 2020 and 2021). The islets differ in surface area and vegetation, being Na Moltona the largest (~5 ha.) and most vegetated islet, and En Curt the smallest (~0.3 ha.) and least vegetated. Na Guardia (~ 2.5 ha.) shares a similar vegetation assembly with Moltona. These contrasting patterns highlight a complex interplay of ecological constraints, demographic dynamics, and historical human disturbance. These three islets are the focus of an ongoing demographic study based on capture-recapture data (Rotger et al., 2016; Rotger et al., 2023).

Understanding the demographic history of *P. lilfordi* can offer valuable insights not only applicable to the conservation management plans for this species, but also for other species facing similar challenges in fragmented habitats. These insights are essential to develop effective conservation strategies that consider both the preservation of genetic diversity and the maintenance of connectivity between isolated populations, ensuring the long-term survival of the Balearic lizard in its increasingly threatened insular environment.

2. Methodology

2.1. Sampling and genotyping

We used previously published SNPs from 191 individuals of *P. lilfordi* collected from 16 localities across its current distribution (Bassita et al., 2021; Otalora et al., 2024) (Figure 1). The procedures for capture, tissue sampling, sequencing, and SNP discovery of all individuals have been described in Otalora et al. (2024). Three types of data were used for the demographic analyses in this study: the genotyping-by-sequencing (GBS) dataset (Otalora et al., 2024), the restriction-site associated DNA (RAD-Seq) dataset (Bassita et al., 2021), and the combined dataset (COMBINED, GBS+RAD-Seq, merged) (Otalora et al., 2024). The calling and filtering process was performed with VCFTools v. 0.1.16 (Danecek et al., 2011; De la Cruz and Raska, 2014); using the following parameters: --max-allels 2, --remove-indels, --min-meanDP 10, --max-meanDP 50, --max-meanDP 50, -

-max-meanDP 50, --max-meanDP 50, - 173 minQ 30, --maf 0.05, --max-missing 0.75 (individual datasets) or 0.90 (combined); and linkage disequilibrium (LD) using plink v2.00a3.3 (Purcell et al., 2007) (--indep-pairwise: 50 ["kb"] 5 0.20). The outlier loci were identified using two independent approaches: PCAdapt software v 4.3.5 (Luu et al., [2017](#)) and BayeScan v 2.1 (18 SNPs) (Foll & Gaggiotti, [2008](#)).

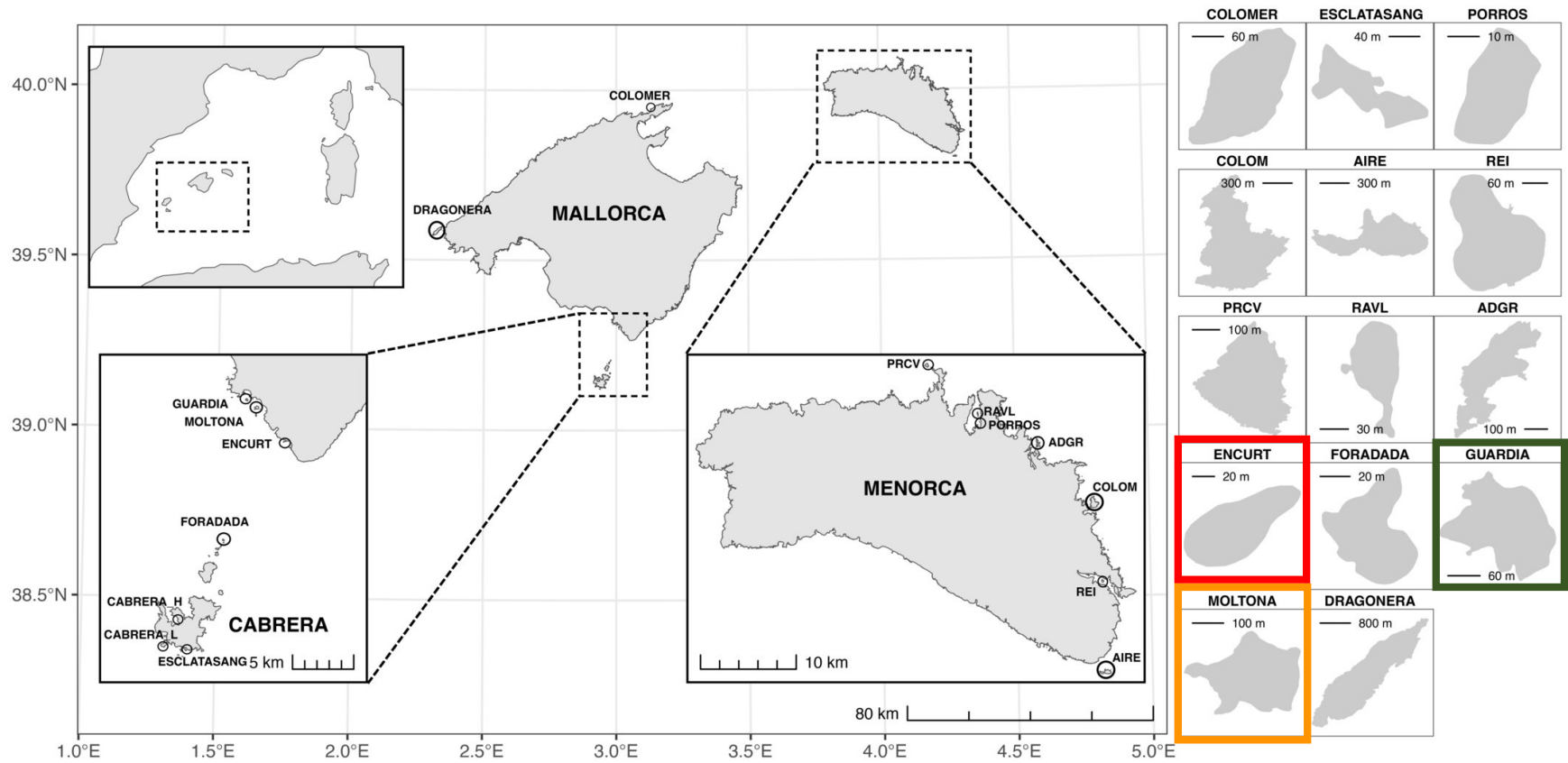


Figure 1. Geographic map of the 16 localities under study, sampled in the archipelagos of Mallorca, Cabrera, and Menorca. The three islets under demographic study are highlighted in the red, green, and yellow boxes.

We generated three stages to evaluate demographic events at different biological and geographic scales: species, archipelago, and islet levels. Each dataset was designed to retrieve distinct aspects of the demographic history of *P. lilfordi*. First, for the species-level analysis, we constructed a comprehensive dataset including all 16 populations across the species' distribution to reconstruct major historical demographic events affecting the species as a whole (Figure 1, Table 1). Second, to explore archipelago-level demographic events, we created two subsets: one including individuals from the Mallorca/Cabrera group, and another comprising individuals from Menorca (Otalora et al., 2024). Both the species-level and archipelago-level analyses were based on the COMBINED dataset (GBS + RAD-Seq), comprising 2,867 SNPs. Third, for the islet-level analysis, we focused on three populations of particular interest located in the south of Mallorca: En Curt, Na Guardia, and Na Moltona. For each of these populations, we performed genotyping using the GBS dataset exclusively. SNP filtering was carried out as previously described (see Methods), but applied independently for each population, allowing us to recover a larger number of markers per islet. The resulting islet-specific datasets contained 6,393 SNPs for Na Guardia, 3,522 SNPs for Na Moltona, and 2,936 SNPs for En Curt. This approach allowed us to treat each islet as an independent genetic unit and to reconstruct their recent demographic history individually (Table 1).

Table 1. Sampling design and sequencing data available from Otalora et al., 2024.

Archipelago	Islet	No. of Individuals		Total No. of Individuals	Latitude	Longitude	SNPs (Raw / Filtered/LD / Outliers) – Species & Archipelago	SNPs (Raw / Filtered / LD/ Outliers) – Islets
		This study	Bassitta et al 2021					
Menorca	Aire	11	12	23	39°48'3"	4°17'29	63,943,45/5523/2876/2867	N/A
	Rei	N/A	10	10	39°53'11"	4°17'14		
	Addaya Gran	10	N/A	10	40°1'	4°12'35		
	Colom	10	10	20	39°57'34"	4°16'40		
	Rovells	10	N/A	10	40°2'40"	4°8'15		
	Porros de Fornell	N/A	11	11	40°2'19"	4°8'20		
	Porros de Cavalleria	10	N/A	10	40°5'32"	4°4'31		
Cabrera	Foradada	N/A	10	10	39°12'25"	2°58'44		
	Cabrera	N/A	20	20	39°8'30"	2°56'30		
	Esclatasang	N/A	10	10	39°7'32"	2°56'36		
Mallorca	Colomer	N/A	4	4	39° 55' 52"	3° 6' 45"		
	Dragonera	N/A	4	4	39°58' 5"	2°31'96		
	Na Guardia	15	N/A	15	39°18'37"	3°0'4		
	Na Moltona	11	N/A	11	39°18'17"	3°0'39		
	En Curt	23	N/A	23	39°16'45"	3°2'10		
Total		100	91	191				

Filters: --min-alleles 2, --max-alleles 2, --min-meanDP10, --max-meanDP50, --minQ30, --maf 0.05, --max-missing 0.75 (combined) or 0.90 (individual datasets).

LD: After linkage disequilibrium (LD): --indep-pair-wise 50 ["kb"] 5 0.20. Candidate outliers: 18 SNPs.

2.2. Demographic history models

Each genomic dataset was transformed into the site frequency spectrum (SFS) format using the easySFS.py script (<https://github.com/isaacovercast/easySFS>). We included all SNPs using the (-a) flag and used the (--preview) option to inspect how the number of segregating sites varied across different projection values (i.e., sample sizes). Based on this output, we selected the projection that best balanced the retention of individuals (number of alleles) and the number of segregating sites. This balance point, hereafter referred to as TRADE (*Trade-off between Retained Alleles and Data Efficiency*), corresponds to the inflection point where a noticeable drop in the number of segregating sites occurs with only minimal gain in retained alleles. TRADE was determined by calculating the differences in the number of segregating sites between successive projections and identifying the point of steepest decline, representing the optimal compromise between missing data and information retention. This procedure was applied independently for each dataset. Finally, we re-ran easySFS.py using the selected projection value with the (--proj) flag.

To infer population size changes in *P. lilfordi*, we used STAIRWAY PLOT v2.1.1 (Liu & Fu, 2020), a model-flexible and nonparametric method that estimates changes in effective population size (N_e) over time from the SFS using a composite likelihood approach. This method does not require the specification of a predefined demographic model. The software can handle folded SFSs, which is especially suitable for non-model organisms, as it does not require the inference of ancestral allele states. The previously generated SFS files were used as input. For each analysis, we generated 200 input replicates using the program's default SFS subsampling procedure. We excluded singletons, as their reliability is lower compared to higher-frequency variants, and their inclusion may bias demographic inference. We used an average mutation rate per site per generation of 1.0×10^{-8} —a midpoint estimate based on the reported range of 0.6 to 1.4×10^{-8} (Bergeron et al., 2023). The generation time for *P. lilfordi* was defined as four years, reflecting the estimated median reproductive age of females (Rotger et al., 2023). The results were visualized using ggplot2 in R (TRADE), and major geological events were overlaid to contextualize the inferred demographic changes. Confidence intervals were represented as shaded ribbons, with the darker purple band indicating the 75% confidence interval (N_e 12.5%– N_e 87.5%) and the lighter purple band representing the 95% confidence

interval (Ne 2.5%–Ne 97.5%), both derived from 200 replicates. This allowed us to reconstruct past demographic events in *P. lilfordi* through time. STAIRWAY PLOT assumes neutral evolution and independence among SNPs, as expected in SFS-based demographic models. These assumptions are standard and generally reasonable given our SNP filtering and dataset characteristics.

Ne estimates obtained with STAIRWAY PLOT 2 do not represent the real or census population size (N). Instead, they provide a genetic estimate of major changes or fluctuations in past effective population size, defined as the number of individuals that effectively contribute genes to the next generation (Wright, 1931; Crow & Kimura, 1970). These values are derived under standard coalescent assumptions, including panmixia, neutrality, absence of population structure, non-overlapping generations, and constant Ne within each modeled time interval (Fedorca et al., 2024). As explained by Liu & Fu (2020), " $\theta_k = 4N_k\mu$, where N_k is the effective size of the population during time block k , and μ is the mutation rate per bp per generation," indicating that both mutation rate and generation time directly influence the absolute scale of Ne estimates.

3. Results

Species and archipelago-level events

Demographic analyses identified key fluctuations in the Ne of *P. lilfordi* over time. Furthermore, the evolutionary history of *P. lilfordi* appears to be strongly influenced by large-scale climatic events, such as the MSC (~5.3 Mya), which may have shaped the initial configuration of the insular ecosystems, and the species' formation ~4.9 Mya, marking the beginning of its diversification (Figure 1). Over time, the interplay between geographical isolation, climatic fluctuations, and changes in habitat availability has shaped the population structure of *P. lilfordi*, with recurrent episodes of demographic expansion and contraction. Notably, the Riss and Würm Glaciations appear to have been critical events in reducing effective population size. The demographic reconstruction of major events using the whole dataset, i.e. at species-level (Figure 2), shows an important initial bottleneck associated with the Würm glaciation occurring ~115,000 years ago (ya), after which a period of stability in Ne is reflected (flat line in the plot, indicating no change in effective size), from 74,000 to 21,500 ya. However, this was followed by a second population decline coinciding with the onset of the LGM, occurring from ~19,000 to

~33,000 ya. Finally, a third and drastic reduction of N_e is observed, beginning around 4,500 ya and becoming evident by approximately 4,000 ya, coinciding with the initial arrival of humans to the Balearic archipelago (Figure 2). This early anthropogenic impact predates the later historical colonization events, namely the Phoenician settlement around the 7th century BCE (Before Common Era, equivalent to a. C.) and the subsequent Roman conquest of the Balearic Islands in 123 BCE (~2,000 years before present; Vogiatzakis et al., 2008). Finally, the arrival of humans to the islands represented an additional pressure factor, possibly contributing to the extinction of some populations and the fragmentation of others.

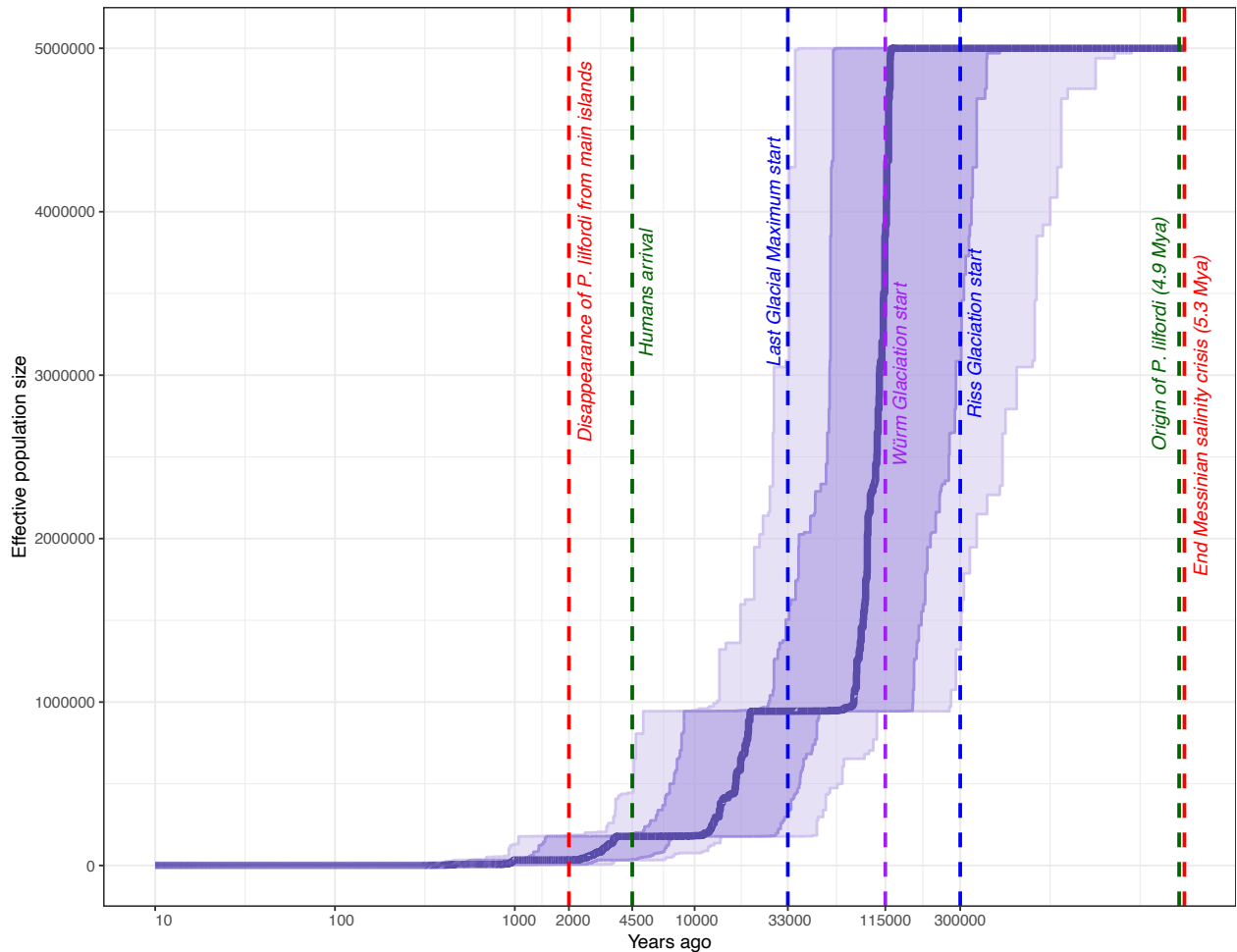


Figure 2. Demographic history of *P. lilfordi* trade inferred from COMBINED data using StairwayPlot2. The bold dashed vertical line marks major events, which coincide with the onset of a significant decline in effective population size.

Demographic reconstruction of the Mallorca-Cabrera Archipelago (Figure 3) shows a marked reduction in effective population size around 1.2 million years ago (Mya) (First Cladogenesis event Mallorca-Cabrera DNAmT). Nevertheless, the 75% confidence interval indicates that this event could have progressively occurred starting from First Donau Glaciation (~2.3 Mya-1.8 Mya). After the decline in N_e , there is a plateau spanning from ~1Mya to 500,000 ya, after which there is a recovery of the N_e .

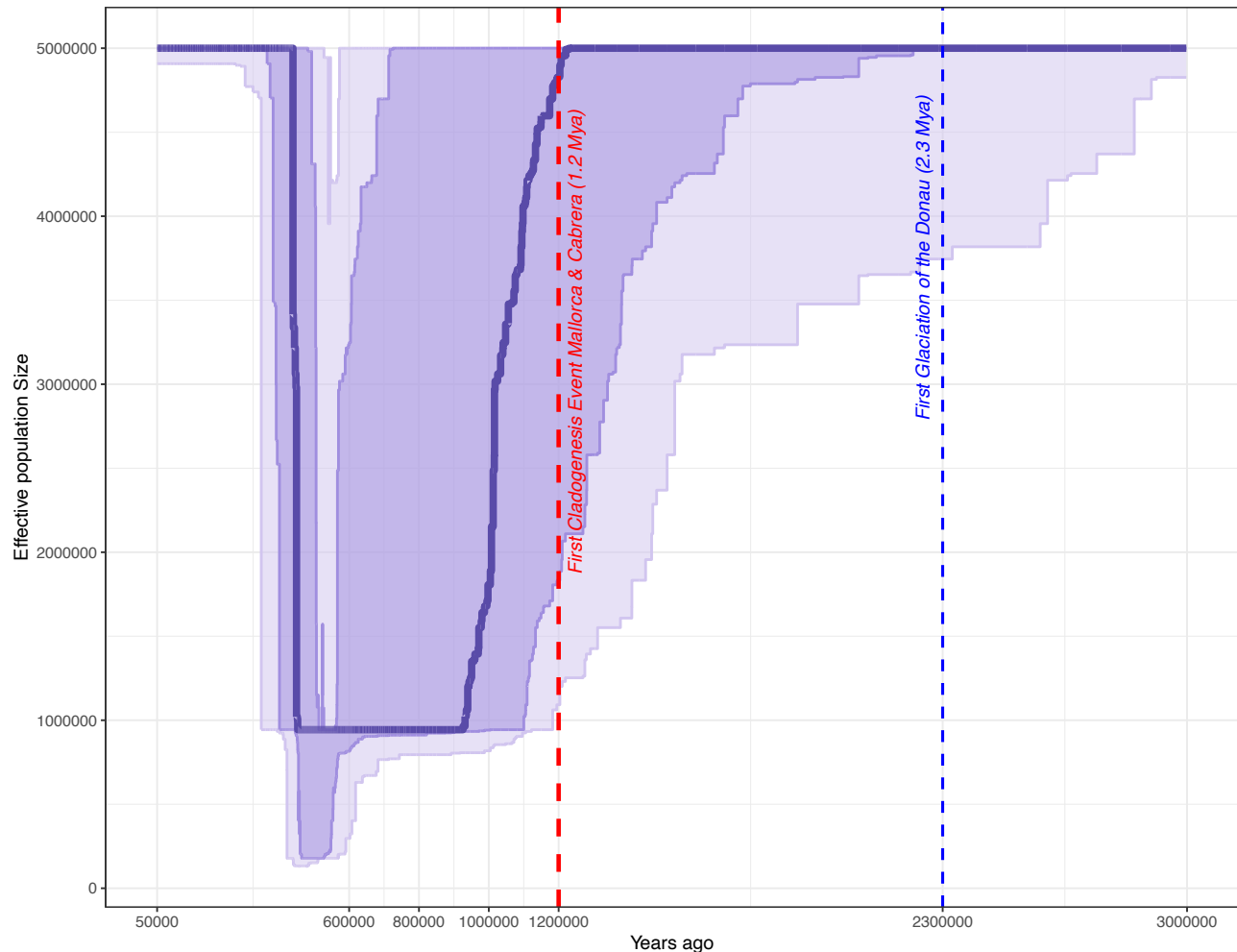


Figure 3. Demographic history of *P. lilfordi* Trade Archipelago Mallorca-Cabrera inferred from COMBINED data using StairwayPlot2. The bold dashed vertical line marks the major events, coinciding with the onset of a significant bottleneck in effective population size.

Similarly, the demographic reconstruction of Menorca (Figure 4) shows an initial period of stability, followed by a significant reduction around ~2.1 Mya within the 75% confidence interval. This decline approximately coincides with the First Donau Glaciation

(~2.3 Mya) and may also be temporally close to the cladogenesis event proposed for Menorca (~2.6 Mya DNAmT) (Terrasa et al., 2009; Brown et al., 2008). Following this reduction in effective population size, the results suggest a prolonged period of demographic stability, followed by a recovery of N_e around ~700,000 years ago, which closely resembles the trajectory inferred for the Mallorca/Cabrera archipelago.

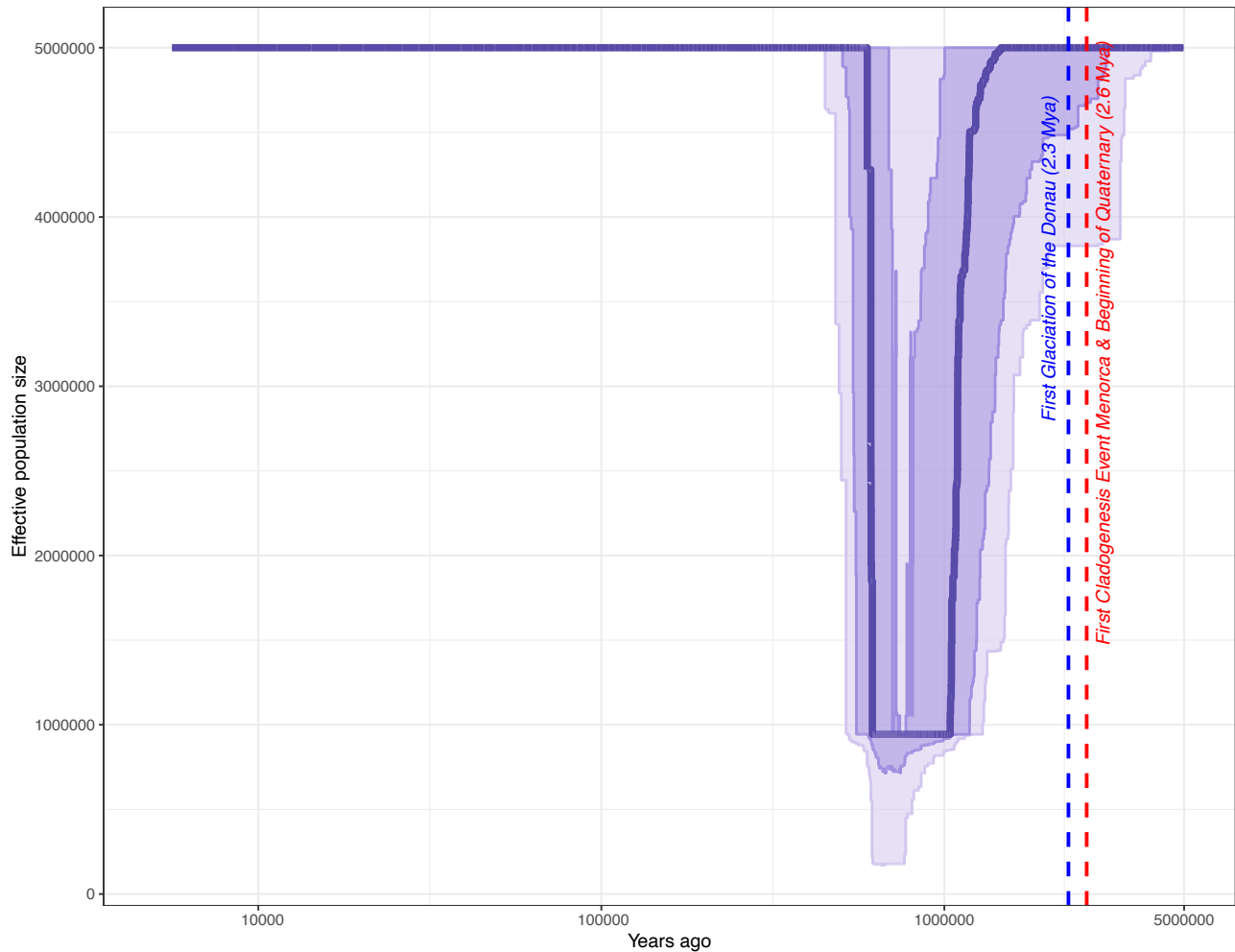


Figure 4. Demographic history of *P. lilfordi* Trade Archipelago Menorca inferred from COMBINED data using StairwayPlot2. The bold dashed vertical line marks the major events, coinciding with the onset of a significant bottleneck in effective population size.

Population-level demography

Regarding the islet approach, the demographic reconstruction of En Curt Islet (Figure 5) reveals an initial reduction in effective population size coinciding with the Riss Glaciation

around (~300,000–130,000 years ago), followed by a prolonged period of stability. Nevertheless, climatic changes associated with the Würm Glaciation triggered a sharp decline in population size, marking the second reduction in N_e . Subsequently, a third population decline is evidenced that could be a consequence of the LGM, suggesting that these glacial events played a decisive role in shaping the demographic structure of this population. A final and pronounced demographic collapse is observed around 4,500 years ago, coinciding with the arrival of humans to the archipelago.

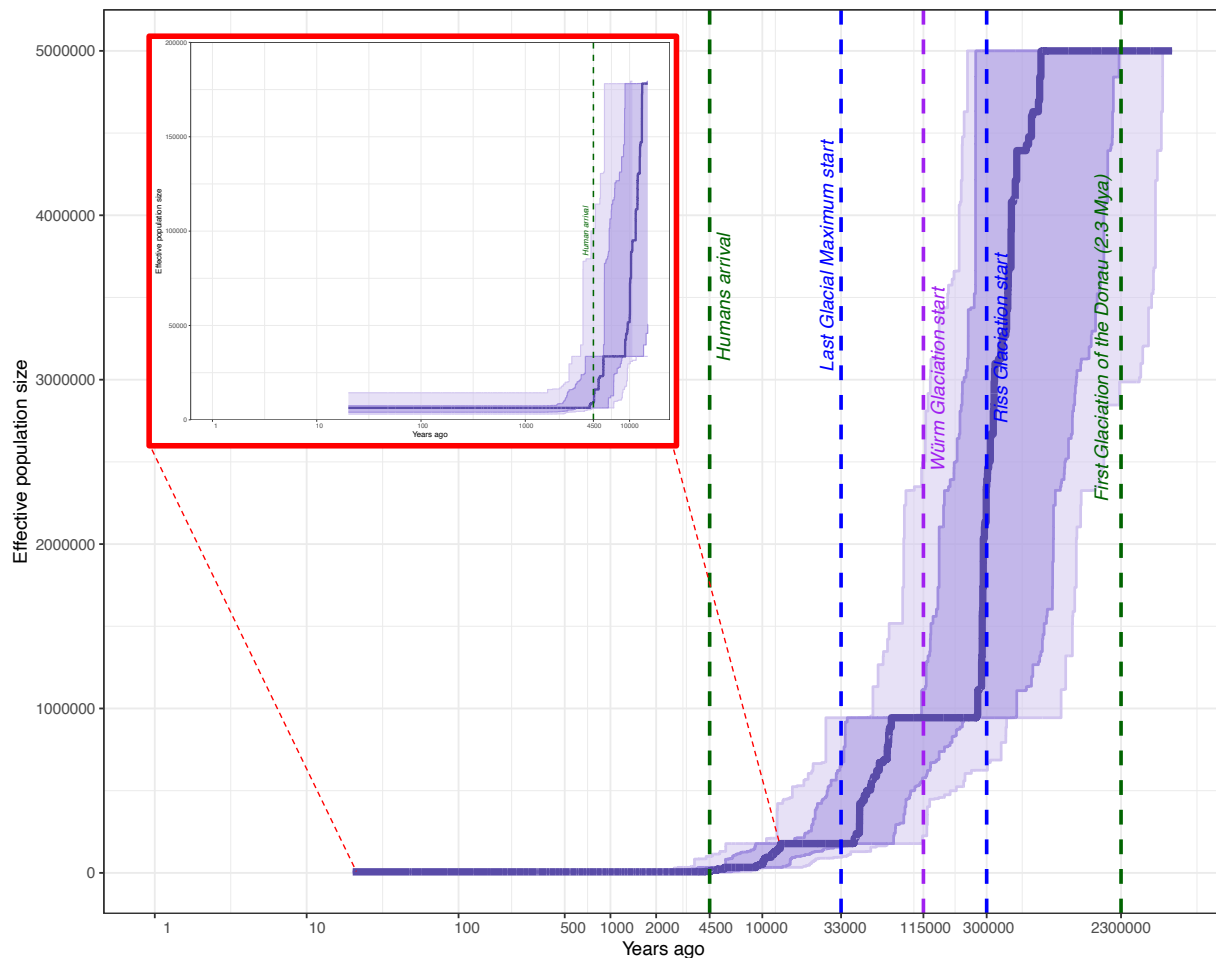


Figure 5. Demographic history of *P. lilfordi* from Islet En Curt inferred from GBS data using StairwayPlot2. The bold dashed vertical line marks major events that coincide with the onset of a significant decline in effective population size. A red inset shows a zoom from 0 to 15 000 years ago, providing higher resolution of the more recent islet history.

On the other hand, the demographic reconstruction of Na Guardia Islet (Figure 6) shows a similar pattern to that of En Curt Islet, indicating that both populations, due to their geographic proximity, share a comparable demographic history. In both cases, the Riss Glaciation initiates a first Ne decline, followed by stability, and then successive declines associated with the Würm Glaciation and the LGM. However, while Na Guardia and also Na Moltona (Figure 7) show signs of a final bottleneck coinciding with human arrival, their demographic trajectories differ notably from En Curt. Rather than a rapid and abrupt decline, both Na Guardia and Na Moltona exhibit a more gradual decrease in Ne, marked by intermittent plateaus over a period of approximately 2,800 years, suggesting a prolonged and possibly buffered demographic contraction. These differences highlight varying degrees of demographic resilience or exposure among these closely situated islet populations.

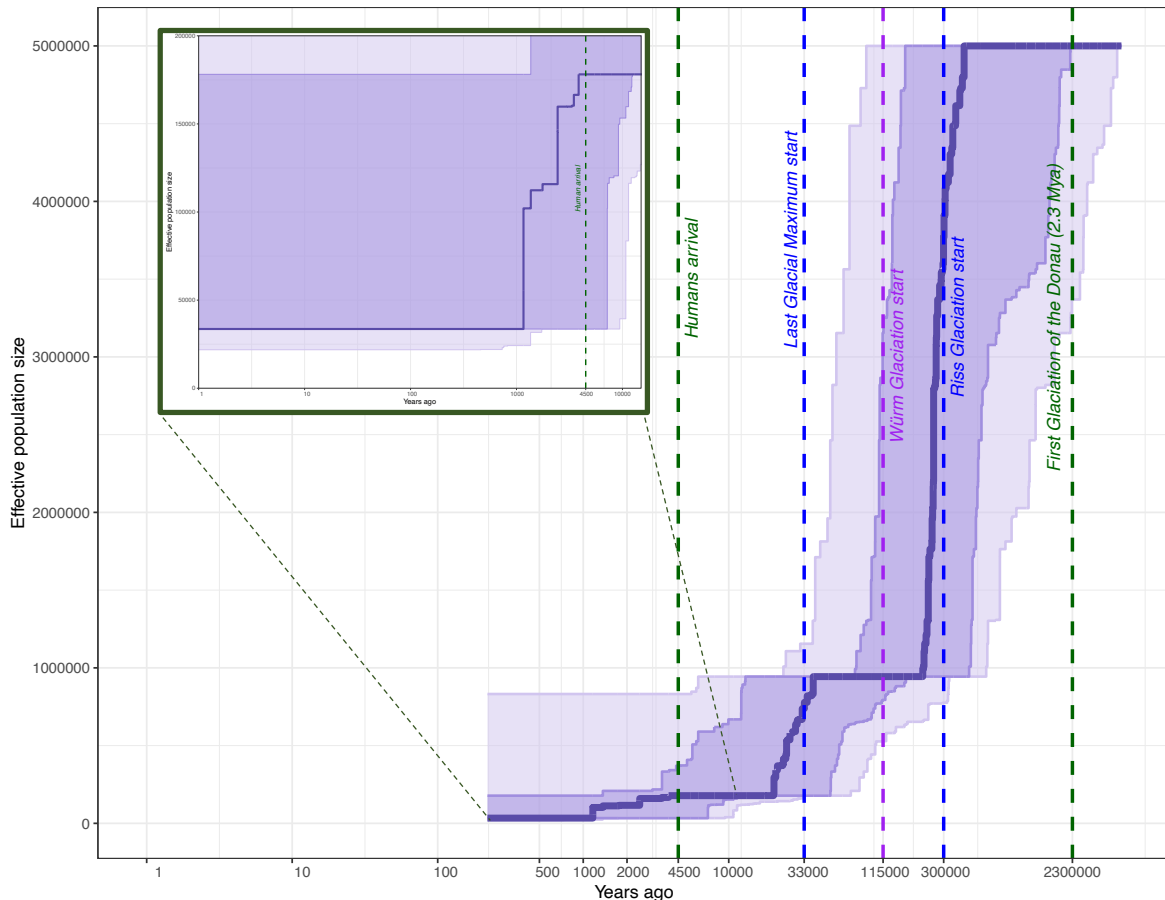


Figure 6. Demographic history of *P. lilfordi* trade Islet Na Guardia inferred from GBS data using StairwayPlot2. The bold dashed vertical line marks major events that coincide with

the onset of a significant decline in effective population size. A green inset shows a zoom from 0 to 15 000 years ago, providing higher resolution of the more recent islet history.

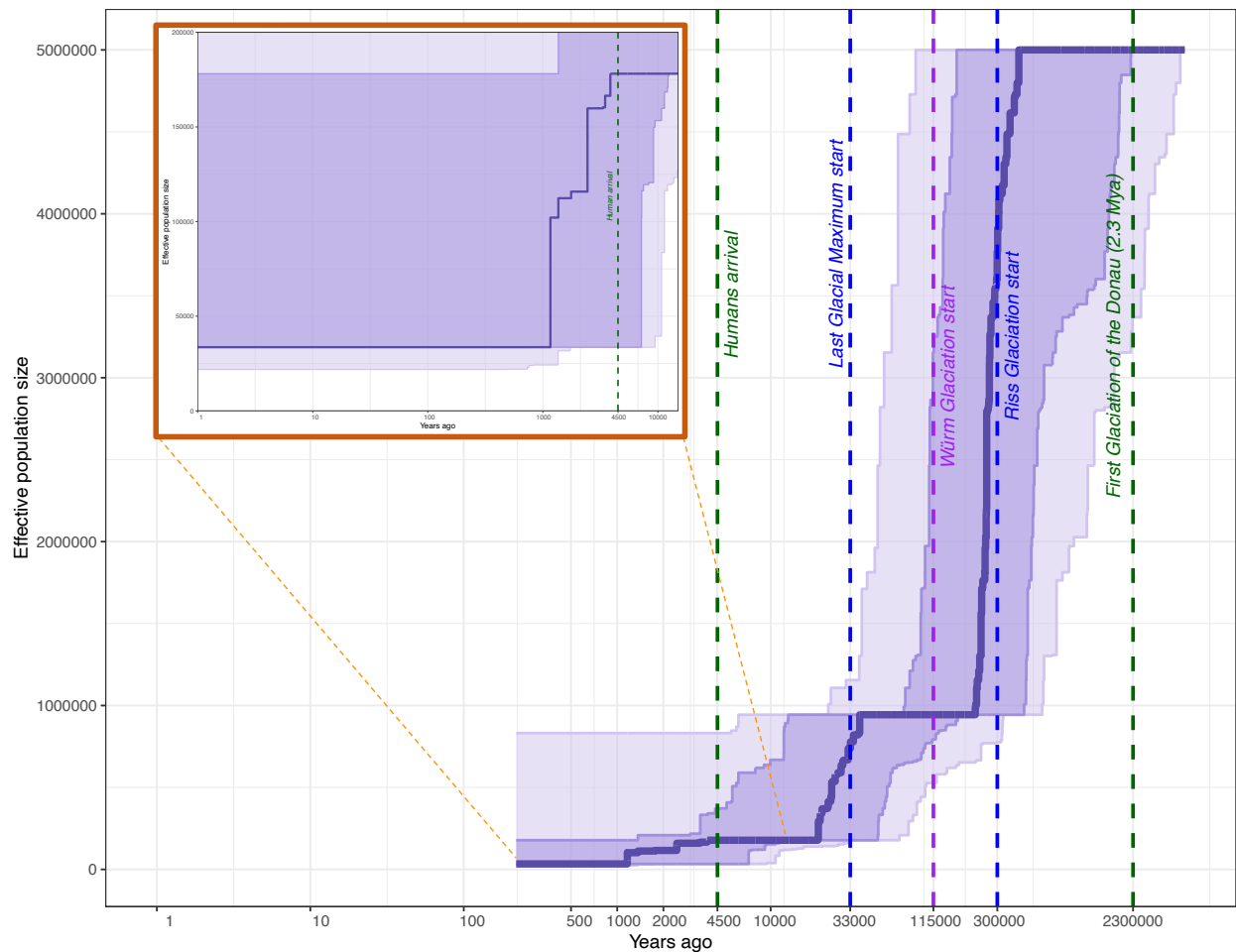


Figure 7. Demographic history of *P. lilfordi* trade Islet Na Moltona inferred from GBS data using StairwayPlot2. The bold dashed vertical line marks major events that coincide with the onset of a significant decline in effective population size. An orange inset shows a zoom from 0 to 15 000 years ago, providing higher resolution of the more recent islet.

4. Discussion

4.1. Historical and demographic events in the species and main archipelagos

The evolutionary history of *P. lilfordi* is linked to climatic and geological events that have shaped the structure of island populations and their genetic diversification (Perez-Cembranos et al., 2020; Otalora et al., 2024; Bassita et al., 2021; Terrassa et al., 2009; Bisconti et al., 2011). A pivotal episode was the MSC (MSC; ~5.96–5.33 Ma), which likely facilitated the formation and diversification of *P. lilfordi* after the desiccation of the Mediterranean. This event led to the extinction of aquatic fauna but allowed the adaptation and diversification of species resistant to extreme conditions, such as reptiles (Brown et al., 2008; Terrassa et al., 2009; Agiadi et al., 2024). During the MSC, climatic and sea-level fluctuations generated temporary land bridges across the western Mediterranean (Brown et al., 2008), enabling population contact, range expansion, and colonization of novel ecological niches. Following this phase, the Zanclean flood (~5.33 Ma) refilled the basin, submerging these corridors and isolating emerging island populations (Garcia-Castellanos et al., 2020). These alternating periods of connection and isolation were instrumental in driving the adaptation of *P. lilfordi* to the developing archipelago environments.

At the whole-species scale (Figure 1), N_e fluctuations align with major climatic events such as the Würm glaciation and the LGM, when sea-level and temperature changes fragmented habitats, creating refugia and driving periods of decline and isolation (Terrasa et al., 2009). This pattern is consistent with findings in other Balearic taxa, such as the small red scorpionfish *Scorpaena notata*, whose demographic reconstructions also point to Late Pleistocene expansions likely driven by interglacial sea-level changes and climatic oscillations (Riera et al., 2025). Comparable palaeoenvironmental shifts influenced the demography of the Balearic (*Puffinus mauretanicus*) and Yelkouan (*Puffinus yelkouan*) shearwaters, with phylogenies and demographic models indicating a split associated with Pleistocene glacial cycles, followed by secondary contact during interglacial periods and high interspecific gene flow that buffered inbreeding and preserved genome-wide diversity (Izquierdo-Arànega et al., 2024). Together, these examples highlight that Late Pleistocene sea-level dynamics not only drove demographic contractions and expansions but also created repeated opportunities for secondary contact and genetic exchange among insular populations.

Building on these demographic patterns, geological and oceanographic conditions during interglacial periods likely shaped the colonization routes and historical connectivity among islands. These conditions may have promoted episodes of transmarine dispersal or island colonization via temporary land bridges, followed by vicariance as sea levels rose in *P. lilfordi*. Since the mid-late Pliocene, sea levels have generally remained higher than at present and have not fallen enough to expose the 70 m channel between Mallorca and Menorca (Emig and Geistdoerfer, 2004), suggesting transmarine dispersal. This is consistent with patterns observed in other Mediterranean reptiles, where colonization often occurred without land bridges, implying, potentially, rafting as the dispersal mechanism (Carranza et al., 2006). In previous genomic studies (Otalora et al., 2024; Bassita et al., 2021), no evidence of recent translocation or admixture was detected among insular populations of *P. lilfordi*, although TreeMix analyses indicated potential historical migration events between some populations of this species (Otalora et al., 2024), suggesting that gene flow may have occurred in the past, albeit requiring further evidence to confirm its extent.

On the other hand, the geographic isolation between the Mallorca/Cabrera and Menorca populations, dated in our analyses to approximately 2.6 Mya, constitutes a new estimate based on genomic data. This finding is consistent with mitochondrial-based evidence (Brown et al., 2008; Terrasa et al., 2009) suggesting an early Pleistocene divergence, coinciding with the Donau glaciation (~2.3–1.8 Mya). This period was characterized by significant global cooling and the expansion of large ice sheets in the Northern Hemisphere (Webb and Bartlein, 1992). Around 1.2 Mya—at the transition from the Donau glacial to the Donau-Günz interglacial—a cladogenesis event separated the Mallorca and Cabrera populations, likely via temporary land bridges or rafting-mediated dispersal (Carranza et al., 2006). This isolation led to a reduction in N_e and a bottleneck (Figure 3), followed by a recovery around 500,000 years ago, possibly related to shifts in climate and insular connectivity.

Regarding Menorca, the demographic reconstruction also shows a similar trend, with an initial population stabilization followed by a significant reduction in N_e around 2.1 Mya, coinciding with the end of the Donau glaciation (1.8 Mya) and the onset of the interglacial period. This event appears to have affected the viability of the Menorca population, possibly due to extreme climatic conditions. The decrease in N_e in Menorca coincides with the first cladogenesis event that separated the populations between Menorca and

Mallorca/Cabrera (~2.6 Mya), leaving persistent genetic traces in current populations (Terrasa et al., 2009; Gracia and Vicens, 1998) (Figure 3). Later, a recovery of N_e was identified around 700,000 years ago, largely coinciding with the recovery observed in the Mallorca/Cabrera populations.

Moreover, ecological conditions may also have contributed to the observed demographic patterns in *P. lilfordi*. The population growth of *P. lilfordi* in these islands was probably facilitated by the availability of new ecological niches (both space and dietary resources), and by the lack of predators and competitors (Hutchinson, 1957). The subsequent introduction of alien species, such as cats, martens, stoats and rodents since the archipelago was settled, about 4.300-4.200 years ago (Bover and Alcover, 2008; Traveset et al., 2009) has had a negative impact on *P. lilfordi* populations, especially on the main islands of Mallorca and Menorca, where the species went locally extinct approximately 2,000 years ago (Alcover, 2000). The absence of natural predators on smaller islands has led to high population densities, but also a decrease in anti-predator behaviors, increasing the vulnerability of these populations (Blumstein et al., 2005; Cooper & Pérez-Mellado, 2011; Pérez-Cembranos et al., 2013).

4.2. Historical and demographic events on the three islets (En Curt, Na Moltona and Na Guardia)

The populations of En Curt, Na Moltona, and Na Guardia show demographic patterns influenced by the Riss and Würm glaciations and the LGM. During the Riss glaciation (~300,000 years ago), the populations experienced a decline in N_e , which was followed by a temporary stabilization. The Würm glaciation (~115,000–11,700 years ago) and the LGM (~20,000 years ago) caused further reductions, with limited recovery due to small population sizes and isolation. These conditions amplified the effects of genetic drift, reducing genetic variability and increasing extinction risk, aligns with the “extinction vortex” theory (Nordstrom et al., 2023).

Human arrival (~4,300–4,500 years ago) (Fernandes et al., 2020) exacerbated this process. Anthropogenic pressures and the introduction of non-native species contributed to a further reduction in genetic diversity and increased the vulnerability of these populations. According to the evolutionary rescue models, small populations face greater difficulties in adapting to abrupt environmental changes due to the loss of genetic diversity (Nordstrom et al., 2023). Human intervention would have aggravated

this genetic loss, intensifying the likelihood of extinction and hindering adaptation. Furthermore, Rotger et al. (2021) indicated that in Na Guardia, the bottleneck associated with human arrival was an additional factor, while in En Curt, another bottleneck was caused by human activities, such as quarrying.

Given this historical context, our demographic reconstructions based on SNPs offer valuable insights into long-term demographic changes, and it is essential to interpret these results within the appropriate geological and evolutionary framework. The coalescent-based approach implemented in Stairway Plot reflects deep genealogical histories and may capture demographic signals that predate the physical emergence of the current islets (En Curt, Na Moltona, and Na Guardia). The precise geological formation dates of the southern islets of Mallorca remain uncertain. It has been proposed that these landforms emerged due to Holocene sea-level rise and erosive processes following the separation of the main islands of Mallorca-Cabrera and Menorca. According to Van den Berg (2015), estimated dates of isolation due to sea-level rise are approximately <2,000 years BP for all three islets, suggesting that they may have been connected to Mallorca until relatively recently. Nevertheless, phylogeographic analyses based on mitochondrial DNA offer an additional perspective: Terrasa et al. (2009) and Brown et al. (2008) identified a cladogenesis event in the western islets of Mallorca—such as Dragonera, Malgrats, and Toro—dated to ~2.3 Mya, indicating that mitochondrial divergence in these insular populations may have been linked to early geological separation from the main island. This raises the possibility that a comparable process may have occurred in the southern islets of Mallorca, including En Curt, Na Moltona, and Na Guardia, where mitochondrial lineage divergence might also correspond to older geological episodes. Thus, the demographic signals detected in these populations may reflect deep evolutionary histories shaped not only by Holocene sea-level changes, but also by earlier geological or ecological barriers that promoted isolation and divergence over longer timescales.

5. Conclusions

The demographic reconstruction of *P. lilfordi*, based on SNPs across all 16 populations, reveals a complex history shaped by major paleoclimatic events, geographical isolation, and anthropogenic pressures. At the species level, three marked declines in N_e were detected: during the Würm glaciation (~115 kya), at the onset of the LGM (~19–33 kya),

and a drastic collapse ~4 kya coinciding with human arrival. In the Mallorca/Cabrera archipelago, SNPs data indicate a divergence from Menorca ~2.6 Mya, consistent with mitochondrial evidence, with a major bottleneck around 1.2 Mya and later recovery (~500 kya). Menorca shows a similar trajectory, with a reduction in N_e around 2.1 Mya (Donau glaciation) and recovery ~700 kya. The islets of En Curt, Na Moltana, and Na Guardia exhibit demographic contractions linked to the Riss, Würm, and LGM glaciations, followed by final declines associated with human impact; however, Na Guardia and Na Moltana show a more gradual contraction than En Curt. The absence of exposure of the 70 m channel between Mallorca and Menorca since the mid-late Pliocene suggests that historical gene flow likely involved transmarine dispersal, potentially via rafting (Carranza et al., 2006). While the geological formation dates of the southern islets remain uncertain (<2,000 years BP; Van den Berg, 2015), phylogeographic signals indicate possible deeper divergences (~2.3 Mya; Terrasa et al., 2009; Brown et al., 2008). These findings support that large-scale climatic events—including the MSC, the Riss and Würm glaciations, LGM and anthropogenic pressures—have been important in shaping the genetic structure and demographic trajectories of *P. lilfordi*, and that integrating SNPs reconstructions with geological context provides a valuable approximation of the evolutionary history of this insular reptile. Importantly, these demographic insights provide the first evidence-based framework to elucidate the species' demographic history, laying the groundwork for more robust evolutionary studies and guiding evidence-based conservation strategies that integrate long-term historical processes into the conservation of *P. lilfordi*. Taken together, our results show that climatic oscillations and human activity have been decisive in shaping *P. lilfordi* demography and provide a species-wide demographic baseline from genome-scale SNP data. A key limitation is the reliance on SNP data alone; integrating ecological and functional genomic evidence will be essential. Future research should combine demographic reconstructions with adaptive variation and habitat modeling to better evaluate resilience under ongoing climate change.

Fundings

This study was supported by the MCIN/AEI/10.13039/501100011033/FEDER, UE (PID2022-141578NB-C22 to L.B.), and the Doctoral Scholarship, Colombian Ministry of Science, Technology and Innovation (MINCIENCIAS885/2020 to K.O.). Institutional support to CNAG was from the Spanish Ministry of Science, Innovation and Universities, Fondo de Investigaciones Sanitarias cofunded with ERDF funds (PI19/01772), the 2014-2020 Smart Growth Operating Program, and the Generalitat de Catalunya through the Departament de Recerca i Universitats and Departament de Salut.

Author contributions

L.B, K.O. and C.G.C conceived the study. K.O. and C.G.C analyzed the results. K.O. wrote the initial draft of the article. All authors reviewed and contributed to the final version of the article. Correspondence to K.O.

References

1. Agiadi, K., Hohmann, N., Gliozzi, E., Thivaïou, D., Bosellini, F., Taviani, M., ... & García-Castellanos, D. (2024). A revised marine fossil record of the Mediterranean before and after the MSC. *Earth System Science Data Discussions*, 2024, 1-11
2. Álvarez, C., Mateo, J. A., Oliver, J., & Mayol, J. (2010). Los ofidios ibéricos de introducción reciente en las Islas Baleares. *Boletín de la Asociación Herpetológica Española*, 21, 126-131.
3. Alcover, J.A. (2000) Vertebrate evolution and extinction on western and central Mediterranean islands. *Tropics*, 10, 103–123.
4. Bassitta, M., Brown, R. P., Pérez-Cembranos, A., Pérez-Mellado, V., Castro, J. A., Picornell, A., & Ramon, C. (2021). Genomic signatures of drift and selection driven by predation and human pressure in an insular lizard. *Scientific Reports*, 11(1), 1–13.
5. Bergeron, L. A., Besenbacher, S., Zheng, J., Li, P., Bertelsen, M. F., Quintard, B., ... & Zhang, G. (2023). Evolution of the germline mutation rate across vertebrates. *Nature*, 615(7951), 285-291
6. Brown, R. P., Terrasa, B., Pérez-Mellado, V., Castro, J. A., Hoskisson, P. A., Picornell, A., & Ramon, M. M. (2008). Bayesian estimation of post-Messinian divergence times in Balearic Island lizards. *Molecular Phylogenetics and Evolution*, 48(1), 350-358.
7. Broquet, T., Angelone, S., Jaquière, J., Joly, P., LENA, J. P., Lengagne, T., ... & Perrin, N. (2010). Genetic bottlenecks driven by population disconnection. *Conservation Biology*, 24(6), 1596-1605.
8. Bisconti, R., Canestrelli, D., Colangelo, P., & Nascetti, G. (2011). Multiple lines of evidence for demographic and range expansion of a temperate species (*Hyla sarda*) during the last glaciation. *Molecular Ecology*, 20(24), 5313-5327.
9. Blumstein, D. T., & Daniel, J. C. (2005). The loss of anti-predator behaviour following isolation on islands. *Proceedings of the Royal Society B: Biological Sciences*, 272(1573), 1663-1668.
10. Bouzat, J. L. (2010). Conservation genetics of population bottlenecks: the role of chance, selection, and history. *Conservation Genetics*, 11(2), 463-478.
11. Carranza, S., Harris, D.J., Arnold, E.N., Batista, V., Gonzalez, J.P., 2006. Phylogeography of the lacertid lizard, *Psammodromus algirus*, in Iberia and across the Strait of Gibraltar. *Journal of Biogeography* 33, 1279–1288.

12. Carretero, M. A., & Pinya, S. (2011). The Balearic herpetofauna: a species update and a review on the evidence. *Acta herpetologica*: 6, 1, 2011, 59-80.
13. Cooper, W., & Pérez-Mellado, V. (2011). Escape by the Balearic Lizard (*Podarcis lilfordi*) is affected by elevation of an approaching predator, but not by some other potential predation risk factors. *Acta Herpetologica*, 6(2), 247-259.
14. Crow, J. F., & Kimura, M. (1970). *An Introduction to Population Genetics Theory*. Harper & Row.
15. Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27, 2156– 2158. [https:// doi. org/ 10. 1093/ bioin forma tics/ btr330](https://doi.org/10.1093/bioinformatics/btr330)
16. De la Cruz, O., & Raska, P. (2014). Population structure at different minor allele frequency levels. *BMC Proceedings*, 8, S55. [https:// doi. org/ 10. 1186/ 1753-6561-8- S1- S55](https://doi.org/10.1186/1753-6561-8-S1-S55)
17. Emig, C., & Geistdoerfer, P. (2005). The Mediterranean deep-sea fauna: historical evolution, bathymetric variations and geographical changes. arXiv preprint [q-bio/0507003](https://arxiv.org/abs/0507003).
18. Fedorca, A., Mergeay, J., Akinyele, A. O., Albayrak, T., Biebach, I., Brambilla, A., ... & Kopatz, A. (2024). Dealing with the complexity of effective population size in conservation practice.
19. Fernandes, D. M., Mittnik, A., Olalde, I., Lazaridis, I., Cheronet, O., Rohland, N., ... & Reich, D. (2020). The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nature Ecology & Evolution*, 4(3), 334-345.
20. Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180, 977–993. [https:// doi. org/ 10. 1534/ genet ics. 108. 092221](https://doi.org/10.1534/genetics.108.092221)
21. Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: a review. *Genetics Research*, 66(2), 95-107.
22. Garcia-Castellanos, D., Micallef, A., Estrada, F., Camerlenghi, A., Ercilla, G., Periañez, R., & Abril, J. M. (2020). The Zanclean megaflood of the Mediterranean–Searching for independent evidence. *Earth-Science Reviews*, 201, 103061.
23. Gracia, F. & Vicens, D. (1998) Aspectes geomorfològics quaternaris del litoral de Mallorca. Aspectes geològics de les Balears (ed. by J.J. Fornós). Universitat de les Illes Balears, Palma de Mallorca, Spain.

24. Günther, A. (1874). Description of a new European species of *Zootoca*. *Annals and Magazine of Natural History*, 14, 158–159
25. Hewitt, G. M. (2004). A climate for colonization. *Heredity*, 92, 1-2
26. Hutchinson, G.E. 1957. Concluding remarks. *population studies: animal ecology and demography*. Cold Spring Harbor Symposia on Quantitative Biology 22:415 – 427.
27. Izquierdo-Aranega, G., Cuevas-Caballe, C., Giannelli, F., Paris, J. R., Bourgeois, K., Trucchi, E., ... & Rozas, J. (2024). Pervasive and recurrent hybridisation prevents inbreeding depression in Europe's most threatened seabird. *bioRxiv*, 2024-10
28. Liu, X., & Fu, Y. X. (2020). Stairway Plot 2: demographic history inference with folded SNP frequency spectra. *Genome biology*, 21(1), 280.
29. Luu, K., Bazin, E., & Blum, M. G. B. (2017). Pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17, 67–77. <https://doi.org/10.1111/1755-0998.12592>
30. Nordstrom, S. W., Hufbauer, R. A., Olazcuaga, L., Durkee, L. F., & Melbourne, B. A. (2023). How density dependence, genetic erosion and the extinction vortex impact evolutionary rescue. *Proceedings of the Royal Society B*, 290(2011), 20231228.
31. Otalora, K., Riera, J. L., Tavecchia, G., Rotger, A., Igual, J. M., Trotta, J.-R. P., & Baldo, L. (2024). Population genetics and phylogeographic history of the insular lizard (Günther, 1874) from the Balearic Islands based on genome-wide polymorphic data. *Ecology and Evolution*, 14(5), e11407.
32. Podnar, M., Mayer, W. & Tvrtkovic, N. (2005) Phylogeography of the Italian wall lizard, *Podarcis sicula*, as revealed by mitochondrial DNA sequences. *Molecular Ecology*, 14, 575–588.
33. Pérez-Cembranos, A., Pérez-Mellado, V., & Cooper, W. E. (2013). Predation risk and opportunity cost of fleeing while foraging on plants influence escape decisions of an insular lizard. *Ethology*, 119(6), 522-530.
34. Pérez-Mellado, V., & Martínez-Solano, I. (2009a). *Podarcis lilfordi*. The IUCN Red List of Threatened Species 2009: e.T17795A7481971. <https://doi.org/10.2305/IUCN.UK.2009.RLTS.T17795A7481971.en>.
35. Pérez-Cembranos, A., Pérez-Mellado, V., Alemany, I., Bassitta, M., Terrasa, B., Picornell, A., ... & Ramon, C. (2020). Morphological and genetic diversity of the Balearic lizard, *Podarcis lilfordi* (Günther, 1874): Is it relevant to its conservation?. *Diversity and Distributions*, 26(9), 1122-1141.

36. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81, 559–575. [https:// doi. org/ 10. 1086/ 519795](https://doi.org/10.1086/519795)
37. Riera, J., Tudurí, A., Guijarro, B., Ordines, F., Picornell, A., & Ramírez-Amaro, S. (2025). Genetic Diversity, Connectivity and Demographic History of the Small Red Scorpionfish *Scorpaena notata* at a Small Scale in the Balearic Islands (Western Mediterranean). *Diversity*, 17(6), 405.
38. Rotger, A., Igual, J. M., Genovart, M., Rodríguez, V., Ramon, C., Pérez-Mellado, V., Bibiloni, G., Rita, J., & Tavecchia, G. (2021). Contrasting Adult Body-Size in Sister Populations of the Balearic Lizard, *Podarcis lilfordi* (Günther 1874) Suggests Anthropogenic Selective Pressures. *Herpetological Monographs*, 35(1), 53–64.
39. Rotger, A., Igual, J. M., Smith, J. J., & Tavecchia, G. (2016). Relative role of population density and climatic factors in shaping the body growth rate of Lilford’s Wall Lizard (*Podarcis lilfordi*). *Canadian Journal of Zoology*, 94(3), 207-215.
40. Rotger, A., Tenan, S., Igual, J. M., Bonner, S., & Tavecchia, G. (2023). Life span, growth, senescence and island syndrome: Accounting for imperfect detection and continuous growth. *Journal of Animal Ecology*, 92(1), 183-194.
41. Ryman, N., Laikre, L., & Hössjer, O. (2019). Do estimates of contemporary effective population size tell us what we want to know?. *Molecular Ecology*, 28(8), 1904-1918.
42. Salvador, A. (2015). *Lagartija balear – Podarcis lilfordi*. In A. Salvador & A. Marco (Eds.), *Enciclopedia virtual de los vertebrados Españoles*. Museo Nacional de Ciencias Naturales. [http:// www.vertebradosibericos.org/](http://www.vertebradosibericos.org/)
43. Terrasa, B., Pérez-Mellado, V., Brown, R. P., Picornell, A., Castro, J. A., & Ramon, M. M. (2009). Foundations for conservation of intraspecific genetic diversity revealed by analysis of phylogeographical structure in the endangered endemic lizard *Podarcis lilfordi*. *Diversity and Distributions*, 15(2), 207-221.
44. Van den Berg, M. (2015). *Endemism and diversity of Lacertidae on the Balearic Islands* [Unpublished manuscript]. Naturalis Biodiversity Center.
45. Vogiatzakis, I. N., Pungetti, G., & Mannion, A. M. (Eds.). (2008). *Mediterranean island landscapes: Natural and cultural approaches*. Springer Science & Business Media. <https://doi.org/10.1007/978-1-4020-5064-0>.

46. Webb III, T., & Bartlein, P. J. (1992). Global changes during the last 3 million years: climatic controls and biotic responses. *Annual review of Ecology and Systematics*, 141-173.
47. Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16(2), 97–159.
48. Yang, W., Feiner, N., Pinho, C., While, G. M., Kaliontzopoulou, A., Harris, D. J., ... & Uller, T. (2021). Extensive introgression and mosaic genomes of Mediterranean endemic lizards. *Nature Communications*, 12(1), 2762.



5. CHAPTER III

Antimicrobial peptides (AMPs) of lizards: a first comprehensive characterization of beta-defensins, ovo-defensins and cathelicidins from the Balearic lizard *Podarcis lilfordi* and closely related Lacertidae species

Antimicrobial peptides (AMPs) of lizards: a first
comprehensive characterization of beta-defensins, ovo-
defensins and cathelicidins from the Balearic lizard
Podarcis lilfordi and closely related Lacertidae species

Katherin Otalora^{1,2}, Jessica Gomez-Garrido^{3,4}, Laura Baldo^{1,5*}

¹ Department of Evolutionary Biology, Ecology and Environmental Sciences, Biology Faculty, University of Barcelona (UB), Barcelona, Av. Diagonal 645, Barcelona 08028, Spain

² Fundación FUNMAJO, EBA, RAIEC, Biodiversity Branch, Cucaita, Boyacá, Colombia

³ Centro Nacional de Análisis Genómico (CNAG), Baldiri Reixac 4, 08028 Barcelona, Spain

⁴ Universitat de Barcelona (UB), Barcelona, Spain

⁵ Institute for Research on Biodiversity (IRBio), University of Barcelona (UB), Barcelona, Spain

*Corresponding author:

Laura Baldo

E-mail: baldo.laura@ub.edu

DOI: <https://doi.org/10.1186/s12864-025-11797-2>

Abstract

Background: Reptiles exhibit remarkable resistance to infections, making them a critical model for studying the molecular basis of innate immunity. Key components of their innate immune response are the antimicrobial peptides (AMPs), small, rapidly evolving molecules with a broad spectrum of antimicrobial activity. Despite their importance, AMPs in reptiles remain largely understudied, with no current studies available on one of the most successful reptilian lineages, lizards from the Lacertidae family. In this study, we performed an extensive mining of published Lacertidae genomes and present the first comprehensive characterization of their major AMP families —beta-defensins (BDs), ovoidensins (OVODs), and cathelicidins (CATHs) — with the aim of understanding their genomic diversity and evolution.

Results: We first identified a nearly complete catalogue of antimicrobial proteins from *P. lilfordi*, an endemic lizard of the Balearic Islands (Spain): this included 63 BDs, eight OVODs, three of which are proline-rich proteins (OVOD-PrAMPs), and four CATHs. These AMPs are arranged in chromosome clusters of mostly contiguous peptides, flanked by highly conserved marker proteins. Using this fine-scale annotation, we mined corresponding orthologues and closed paralogues from *P. muralis*, *P. raffonei* and *Zootoca vivipara*, identifying a total of 58 AMPs. All AMPs consistently locate on chromosome 3 (BDs and OVODs) and chromosome 12 (CATHs), supporting a monophyletic origin for the reptilian antimicrobial defense. All Lacertidae AMPs show a multiple exon structure (two to four exons) and a characteristic cysteine motif (six cysteines in BDs, eight in OVODs, and four in CATHs), consistently with previous findings for vertebrates. Comparative sequence analyses indicate pervasive duplication events in tandem of both BDs and OVODs, with clear evidence of recent gene expansion. In contrast, OVOD-PrAMPs and CATHs mostly present a one-to-one ortholog across species. Despite this high level of intra-genomic diversity, we also identified multiple instances in which different species from distantly related Squamata families (Lacertidae, Varanidae and Dactyloidae) share identical or nearly identical peptides, suggesting convergent evolution.

Conclusions: Overall, our findings support a remarkable diversity of AMPs in lizards, driven by an ongoing process of gene expansion, primarily via duplication in tandem, and rapid diversification. The observed instances of peptide identity or similarity, along with the putative maintenance of specific polymorphisms across species, provide a basis

for investigating the potential functional equivalence of these peptides among lizards and underscore the need for population-level studies of allele intraspecific diversity.

1. Introduction

AMPs are a large class of small peptides (15–150 amino acids) that are found abundantly in nature and play a crucial role in the innate immune systems of all organisms, from bacteria to vertebrates (Lazzaro et al., 2020; Huan et al., 2020; Chakraborty et al., 2022; Erdem Büyükkiraz & Kesmen, 2022). They exhibit an ample spectrum of antimicrobial effects, primarily targeting bacteria, viruses, fungi and parasites, as well as potent immunomodulatory activity (Huan et al., 2020; Chakraborty et al., 2022). In addition to serving as a primary defense against pathogens, AMPs act as key modulators of host–microbe interactions (Cardoso et al., 2022).

AMPs are typically derived from post-translational modifications involving the proteolytic cleavage of large precursor proteins (Zhang & Gallo, 2016). Most active peptides present a net positive charge and an amphipathic nature, enabling direct membrane binding and disruption in target microbes (Chakraborty et al., 2022; Erdem Büyükkiraz & Kesmen, 2022; Zhang & Gallo, 2016). Anionic peptides have been also reported; however, their mechanisms of action remain mostly unclear (Harris et al., 2009).

Coevolution of AMPs and pathogen resistance has led to a highly functionally and structurally diverse repertoire of AMPs (Huan et al., 2020; Erdem Büyükkiraz & Kesmen, 2022), resulting from extensive gene duplication, pseudogenization, deletion as well as strong positive selection (Lazzaro et al., 2020; Semple et al., 2005). Nonetheless, evidence of polymorphism maintenance by balancing selection and convergent evolution has also been documented (Lazzaro et al., 2020; Chakraborty et al., 2022), indicating an overall complex pattern of molecular evolution. Hence, despite their ample spectrum of microbial targets, recent studies have shown that AMPs response can be remarkably specific, with single polymorphisms significantly influencing resistance to infections (Lazzaro et al., 2020; Hanson, 2024).

Among the plethora of AMP-producing proteins found in vertebrates, the main classes are represented by defensins and cathelicidins (with the notable exception of caecilian amphibians that appear to lack defensins) (Huan et al., 2020; Benítez-Prián et al., 2024). Both are characterized by the presence of cysteine residues (ranging from four to eight) arranged in conserved motifs and forming disulfide bonds that contribute to their

structural stability. The number of cysteine residues and spacing of cysteine-pair alignment currently define the known subclasses (Semple et al., 2006; Semple & Dorin, 2012). Beta-defensins are the most abundant and diverse subclass found in all vertebrates, known for their immunomodulatory and toxic activities, as well as their involvement in developmental processes, fertility, wound healing and cancer, among others (Semple & Dorin, 2012). They produce small active peptides (20–50 residues) with a characteristic six-cysteine motif (Semple & Dorin, 2012), and are typically arranged in large, discrete clusters within a chromosome (Zhu & Gao, 2013; Tu et al., 2015). In the same syntenic block of the beta-defensins are the ovo-defensins, a small subclass of AMP-producing proteins structurally similar to beta-defensins (containing a six- to eight-cysteine motif). Ovo-defensins have been described so far only in avian and reptile species, and are named for their expression in the oviduct and putative role in egg protection (Whenham et al., 2015; Gong et al., 2010).

Cathelicidins are a small class of largely studied, vertebrate-specific proteins with multifunctional activities marked by potent antimicrobial and immunomodulatory properties (Kościuczuk et al., 2012; Bals & Wilson, 2003). They contain a typical four-cysteine motif and are synthesized as large precursor protein comprising a conserved cathelin-like domain (CLD) and a functional, highly variable antimicrobial domain (Kościuczuk et al., 2012; Tomasinsig & Zanetti, 2005).

While AMPs have been extensively studied in humans, other mammals and amphibians (particularly frogs), a comprehensive genomic characterization of AMP diversity is still lacking in most other vertebrate groups (Wang et al., 2016). Reptiles are an increasingly attractive model for antimicrobial defense studies, as they represent one of the earliest branches to evolve an innate immune response among extant vertebrates, and are known for their remarkable resistance to infections (Field et al., 2022; Buchmann, 2014). Yet, AMP studies in reptiles are still at their beginnings (Santana et al., 2021; Siddiqui et al., 2022).

Over the past two decades, research has identified genes associated with major classes of AMPs in a limited number of representative species across the four major reptilian orders (i.e. Squamata, Testudines, Rhynchocephalia and Crocodylia) (Santana et al., 2021). Within Squamata (snakes and lizards), AMP diversity has been documented in species such as the Komodo dragon (*Varanus komodoensis*) (van Hoek et al., 2019), the green anole lizard (*Anolis carolinensis*) (Dalla Valle et al., 2012), and few snake species,

including the king cobra (*Ophiophagus hannah*) (Zhao et al., 2008) and the Burmese python (*Python bivittatus*) (Kim et al., 2017). No formal AMP characterization is yet available for any members of the most successful lineage of the Squamata order: the family Lacertidae.

Lacertids are a highly species-diverse group of true lizards, primarily encompassing wall lizards that are widely distributed throughout the Mediterranean region. Few Lacertidae genomes have been recently published, including the common wall lizard, *Podarcis muralis* (Andrade et al., 2019), the Aeolian wall lizard, *P. raffonei* (Gabrielli et al., 2023), and the viviparous lizard, *Zootoca vivipara* (Yurchenko et al., 2020), thus providing the genomic resources for a first comprehensive overview of the AMP diversity within this reptilian family. In addition, our research group has recently produced a high-quality chromosome-level genome assembly and annotation for Lilford's wall lizard *P. lilfordi* (Gomez-Garrido et al., 2023), an endemic species of the Balearic Islands where it inhabits multiple islets surrounding the islands of Menorca, Mallorca and Cabrera. Due to the insularity of this species and the long-term monitoring of several of these populations (Rotger et al., 2016), *P. lilfordi* is a growing natural model system for integrative studies of host and associated microbial symbionts (Otalora et al., 2024; Baldo et al., 2018; Baldo et al., 2023), as well as pathogen resistance (unpublished) in a context of a rapidly changing environment.

Here we present the first and comprehensive identification of major classes of AMPs (defensins and cathelicidins) in the *P. lilfordi* genome by means of extensive genomic search. To further understand AMP diversity and evolution, we additionally mined AMPs from all currently published Lacertidae genomes (*P. muralis*, *P. raffonei* and *Z. vivipara*). Comparative analyses of these newly identified AMPs, along with previously published data from the Komodo dragon (family Varanidae) (van Hoek et al., 2019) and the green anole lizard (family Dactyloidae) (Dalla Valle et al., 2012), provided a first overview of the antimicrobial defense repertoire for this reptilian family, while also shedding light on major patterns of AMP evolution within Squamata.

2. Methods

2.1. Identification of candidate defensin and cathelicidin genes in *P. lilfordi*

We employed a multiple genomic approach to identify candidate defensin and cathelicidin genes within the *P. lilfordi* genome, including genome-wide TBLASTN screening of reference AMPs, filtering by characteristic cysteine motifs, genomic cluster delineation via conserved flanking markers, Open Reading Frame (ORF) identification, transcriptomic data validation, and physical properties and structural homology prediction. Each step was manually curated to ensure accurate and robust annotation.

Specifically, we first used the protein definitions file (PODLIA.protein_definitions.txt) available at <https://denovo.cnag.cat/podarcis> to retrieve annotated AMPs from the recently published genome assembly (Gomez-Garrido et al., 2023). This dataset included four putative beta-defensins, one cathelicidin, and no ovo-defensins.

Second, we queried all available AMPs from the closest reptile species, the *V. komodoensis* (Komodo dragon) (van Hoek et al., 2019), and the *A. carolinensis* (the green anole) (Dalla Valle et al., 2012), in a TBLASTN search (Altschul et al., 1990) against the *P. lilfordi* genome. The Komodo dragon includes 65 described beta-defensin peptides, six ovo-defensins (only mature peptides), and three cathelicidins (complete genes) (van Hoek et al., 2019); the green anole lizard includes 32 beta-defensins (complete mRNA), three cathelicidins (complete genes) and no ovo-defensins (Dalla Valle et al., 2012). For each protein family, we applied relaxed TBLASTN search parameters (e-value = 1, gap open/extension = 8/2) to maximize output retrieval. The BLAST outputs were then filtered according to the presence of conserved cysteine (C) motifs for each AMP family: six cysteines for beta-defensins, with a relaxed cysteine spacing from van Hoek et al. (2019); a terminal "CC" motif and between six to eight total cysteine residues for ovo-defensins (van Hoek et al., 2019; Dalla Valle et al., 2012); at least four cysteine residues for cathelicidins (Tomasinsig & Zanetti, 2005).

Third, we identified the chromosome and boundaries of defensin and cathelicidin clusters, by searching for conserved flanking marker genes previously documented in literature (van Hoek et al., 2019; Dalla Valle et al., 2012; Cheng et al., 2015; van Hoek, 2014). For beta and ovo-defensins we targeted the exportin-1 protein (XPO1), the cathepsin B protein (CTSB), the XK, Kell blood group complex subunit-related family member 6 (XKR6) and myotubularin phosphatase domain-containing protein 9 (MTMR9) (van Hoek et al., 2019; Zhang et al., 2019). For cathelicidins we targeted the fas-activated

serine/threonine kinase (FASTK) and kelch-like family member 18 (KLHL18) (van Hoek et al., 2019; Cheng et al., 2015).

Throughout this systematic approach, we verified and corrected the orientation of all sequences, prioritizing non-overlapping sequences with the highest bit score. All analyses were performed with custom scripts in R software (2023).

2.2. Gene annotation

To obtain the complete ORFs for the candidate AMPs that passed the above filters, we queried all available full mRNA and proteins from *A. carolinensis* (32 beta-defensins mRNA and four cathelicidin complete proteins) and *V. komodoensis* (only three cathelicidins proteins) using the program *exonerate* v2.4.0 (Slater & Birney, 2005) in *protein2genome* mode against chr-3 (defensins) and chr-12 (for cathelicidins) of the *P. lilfordi* genome. For ovo-defensins, as no complete proteins were available, partial ovo-defensin proteins newly identified from *P. lilfordi* were searched against the NCBI nucleotide (nt) database: best RNA hits were translated to proteins and searched back against the *P. lilfordi* genome.

For each identified AMP gene, we used *P. lilfordi* Illumina and PacBio RNA-seq data from multiple individuals and tissues (Bioproject PRJNA897120) to accurately determine exon–intron structures, using this transcriptional evidence to guide the annotation process. This approach enabled us to validate most of the genes previously identified using partial sequences as well as to annotate novel genes.

Finally, we further validated all candidate peptides/proteins by their physical properties using `EMBOSS` tool `pepstats` [https://www.ebi.ac.uk/jdispatcher/seqstats/emboss_pepstats], and structural homology searches by protein prediction of secondary and tertiary structure in *Phyre2* (Kelley et al., 2015).

We named newly identified proteins and peptides in *P. lilfordi* using a standard nomenclature applied in other species consisting of the genus and species initials as a prefix, followed by a protein code and a number indicating their genomic order (van Hoek et al., 2019; Dalla Valle et al., 2012). Newly identified candidate proteins from Lacertidae were similarly named and numbered according to the *P. lilfordi* closest sequence, when available.

AMP chromosome location was visualized with the R function `GC_chart` in `geneviewer` R library. For the UPGMA (unweighted pair group method with arithmetic mean) dendrogram, a protein distance matrix was estimated with the R function `dist.alignment` (matrix = "identity", gap = T) in `seqinr` library, and a dendrogram generated accordingly.

2.3. Comparative AMP analysis across reptile species

To explore patterns of diversity and evolution of AMPs in Lacertidae, we queried all *P. lilfordi* AMP proteins in a BlastP and TBLASTN search against all published Lacertidae genomes to retrieve corresponding orthologs and paralogs. Available Lacertidae genomes included the Aeolian wall lizard, *P. raffonei* (Gabrielli et al., 2023), the common wall lizard, *P. muralis* (Andrade et al., 2019), and the viviparous lizard, *Z. vivipara* (Yurchenko et al., 2020). Additionally, we used available genome annotation and transcriptome data to guide our search. For comparative purposes, we also integrated available sequences from the Komodo dragon and the green lizard. Amino acid sequences were aligned using MUSCLE vs. 5.0 (Edgar, 2004), visualized and manually curated in Jalview vs. 2.11.3.3 (Waterhouse et al., 2009). Due to the high rate of amino acid substitutions and extensive gene duplications, homology inference within the same protein family is limited. Therefore, alignment curation was performed according to the following guidelines: a) we first aligned C-residues within the conserved motif, b) we maximized homology among highly similar sequences, c) we optimized similarity within the conserved regions among all sequences, and d) for highly variable regions (i.e. functional peptide), we minimized indels among divergent sequences.

To identify clusters of similar peptides, as well as potential orthologs and paralogs within and across species, a Neighbor Joining (NJ) tree (Saitou & Nei, 1987) was built with MEGA 4.0 (Tamura et al., 2007) with 1000 bootstrap values. The resulting trees were visualized using the `ggtree` R package (Nava et al., 2009). The use of distance-based methods (NJ and UPGMA), rather than model-based phylogenetic inference approaches such as Maximum Likelihood (ML), was guided by the complex nature of the AMP dataset. This complexity includes extensive gene paralogy, low sequence homology beyond the conserved cysteine residues, and incomplete gene recovery across the studied species. Distance-based methods represent a conservative and widely used approach in AMP research (Tu et al., 2015), as they minimize the risk of false inferences related to substitution model selection and do not assume true aminoacidic homology.

While this approach limits formal phylogenetic interpretation, it allows for the identification of clusters of similar peptides and broad patterns of sequence similarity without overinterpreting evolutionary relationships.

3. Results

3.1. AMP genomic clusters in *P. lilfordi*

Genome-wide screening of *P. lilfordi* enabled the identification of two major AMP gene clusters, located on chromosomes 3 and 12 (Figure 1).

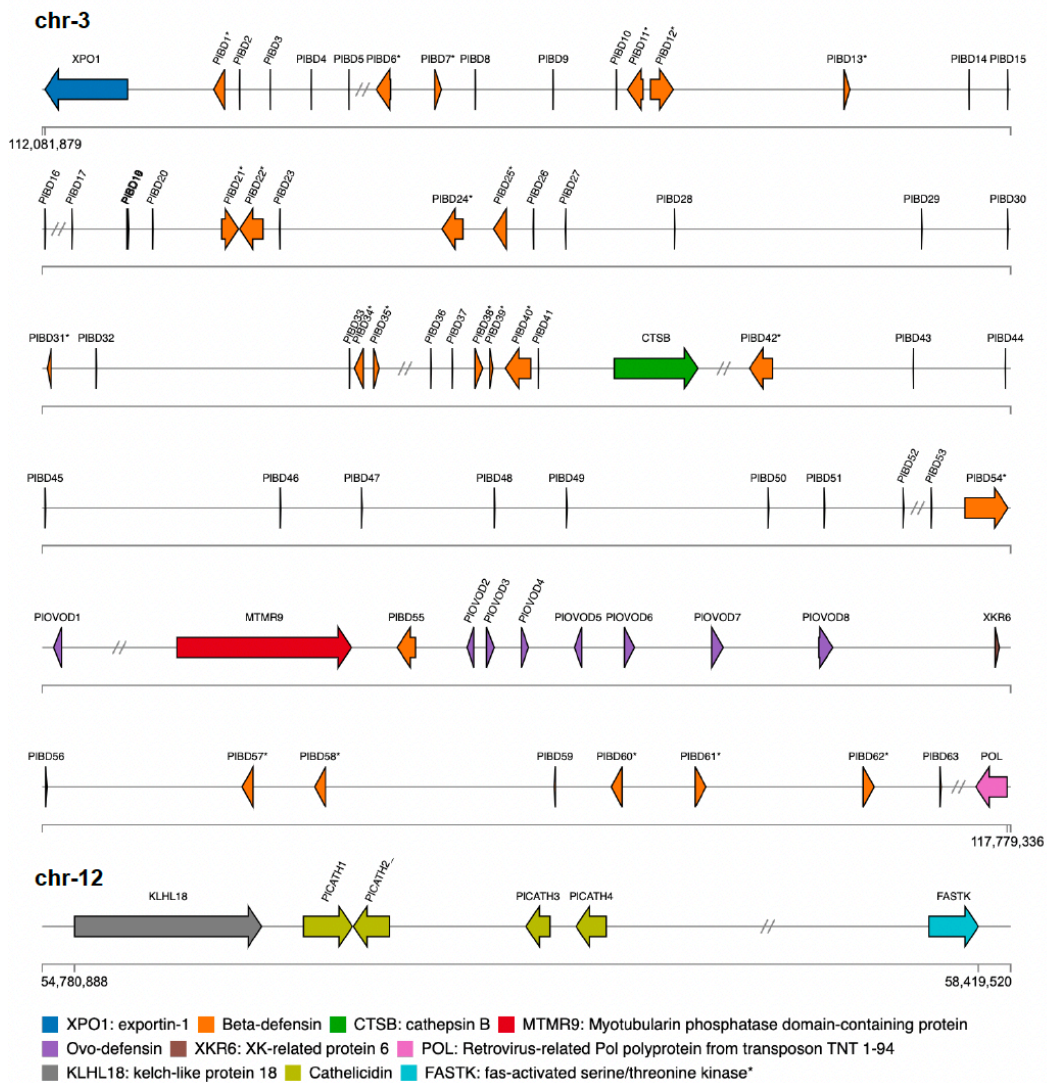


Figure 1. *P. lilfordi* genome localization of beta-defensins (chr-3), ovo-defensins (chr-3) and cathelicidins (chr-12). Beta-defensins include both partial and complete ORFs (marked with an asterisk). Marker genes for major clusters are also shown. Long intergenic regions are shortened and marked by a double //.

Chr-3 cluster (112164644-117711852 bp, about 5.5 Mb) includes 71 putative defensins: 63 beta-defensins (BDs) and eight ovo-defensins (OVODs). The defensin block is flanked by the exportin-1 protein (XPO1) and the retrovirus-related Pol polyprotein from transposon TNT 1-94 (POL), with the cathepsin B protein (CTSB) falling within the block.

Ovo-defensins are found within the beta-defensin block, flanked by the myotubularin phosphatase domain-containing protein 9 (MTMR9) and the XK, Kell blood group complex subunit-related family, member 6 protein (XKR6).

Chr-12 cluster (54832594-58419520 bp, about 3.5 Mb) includes four cathelicidins (CATHs) arranged in a single block, flanked by the kelch-like family member 18 protein (KLHL18) and the fas-activated serine/threonine kinase protein (FASTK) (Figure 1).

3.2. Characterization of *P. lilfordi* beta-defensins

All 63 BDs are arranged into two distinct clusters within chr-3, separated by ovo-defensins: one encompassing PIBD1-PIBD55 and the other PIBD56-63, separated by more than 2,000,000 bases (Figure 1 and Supplemental Table S1). All BDs identified present a six-cysteine motif (29 and 35 amino acid (AA)) with the following conserved spacing: C-X(6-7)-C-X(3-5)-C-X(8-14)-C-X(5-7)-CC (Figure 2). In addition, most sequences show two highly conserved glycine (G) residues at positions 7 and 28, and an isoleucine (I) at position 27 of the alignment (Figure 2). Both PIBD12 and PIBD22 carry a double cysteine motif in tandem at the C-terminal, summing a total of 65 AMPs.

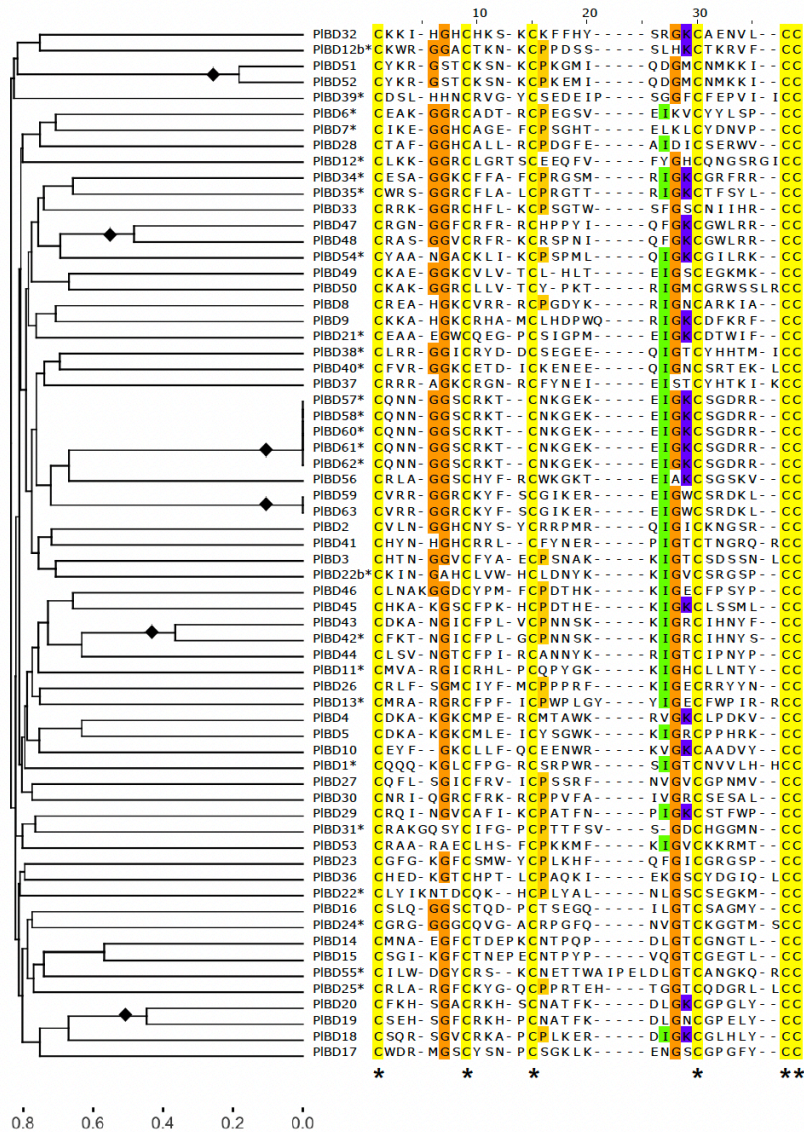


Figure 2. Alignment and UPGMA dendrogram of *P. lilfordi* putative beta-defensin peptides (N=65) showing the six-cysteine domain (C-X(6-7)-C-X(3-5)-C-X(8-14)-C-X(5-7)-CC). Alignment colors show residues above 30% identity threshold. Shared cysteine residues are marked with an asterisk. Dendrogram height corresponds to the square root of sequence identity distances. Branch squares highlight closely related peptides (cutoff at 0.5 height, corresponding to <25% of dissimilarity). An asterisk in sequence IDs correspond to peptides with complete ORFs. PIBD12 and PIBD22 carried a double motif, and the second motif was labelled with “b”.

Most peptides present a cationic charge, with isoelectric points ranging between 6.8 to 10.7, with only 13 out of the 65 peptides showing a negative or no charge (Supplemental Table S1). According to protein homology searches, all peptides (at least 30 AA) have their best hits to an antimicrobial protein/beta-defensin, with confidence >80% and sequence coverage >70%, except for PIBD39 (Supplemental Table S2). Prediction of secondary structure indicates that most cysteine domains contain two to three beta-sheets.

The UPGMA (unweighted pair group method with arithmetic mean) dendrogram identified six clusters of similar sequences (<25% dissimilarity based on an arbitrary cutoff), typically grouping contiguous or spatially proximal peptides within the chromosome (e.g., PIBD51-52, PIBD47-48, PIBD19-20, see also Figure 1). Two of these clusters encompass multiple BDs with identical motifs (PIBD57-58,60-62 and PIBD59,63), suggesting recent duplications in tandem.

Of the 65 beta-defensin motifs identified, we were able to detect the complete open reading frame (ORF) for 24 of them (Figure 3). According to published *P. lilfordi* transcriptomic data from multiple individuals and tissues (Bioproject: PRJNA897120, but also see genome browser https://denovo.cnag.cat/genomes/podarcis/browse/rPodLil1.2_annot/), fourteen of these BDs were transcribed, guiding our gene annotation (Supplemental Table S1).

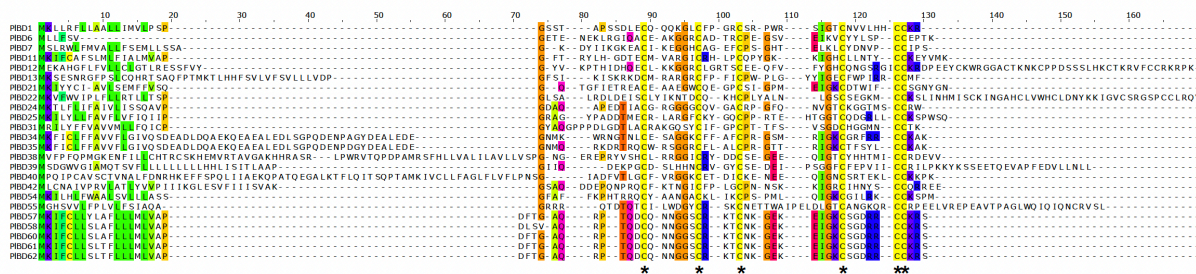


Figure 3. Alignment of *P. lilfordi* complete beta-defensin proteins (N=24). Proteins encompass two to four exons (see Supplemental Table S1 for intron breaks). The six-cysteine residues are marked with an asterisk. Alignment colors show residues above 30% identity threshold. PIBD12 and 22 carry a double cysteine motif (the second motif was not aligned).

Complete protein size ranges from 53 to 118 AA arranged from two to four exons, with the majority presenting two-three exons, with large exon size variation (exon-1: 6-55 AA, exon-2: 21-69 AA, exon-3: 0-49 AA, exon-4: 0-45 AA) (see Supplemental Table S1 for intron position). They contain a N-terminal with a putative signal peptide (first exon and part of the second exon), followed by the cysteine motif at the C-terminal encoding for the AMP. Structural homology searches based on complete proteins confirm strong similarity to previously identified beta-defensins, with confidence >90%, although with generally low sequence identity (35-65%) (Supplemental Table S2). Again, only PIBD39 shows poor confidence (<80%).

Notably, the BD cluster encompassing identical peptides (PIBD58-62, Figure 2) also displays near AA identity at full protein length (Figure 3), supporting recent whole-gene duplication.

3.3. Comparative analysis of beta-defensins across reptiles

To investigate patterns of beta-defensin diversity and evolution within the Lacertidae, we performed extensive BLAST searches against the genomes of closely related species to *P. lilfordi* (hereafter named Pl), including *P. raffonei* (hereafter Pr), *P. muralis* (Pm) and *Z. vivipara* (Zv). This analysis yielded several complete beta-defensin proteins: 11 from Pr (eight newly identified), seven from Pm (six new) and nine from Zv (six new). Notably, many of these newly characterized beta-defensins were previously annotated as uncharacterized non-coding RNAs or mRNAs. Detailed AccNos for these sequences are provided in Supplemental Table S3.

BD peptides from all reptile species share the following six-cysteine motif: C-X(6-7)-C-X(3-7)-C-X(8-14)-C-X(4-7)-CC (see Supplemental Figure S1 for alignment).

The Neighbor-Joining (NJ) tree (Figure 4A) identified fourteen multispecies clades (bootstrap support >70) and only one clade of *P. lilfordi*-specific sequences (clade 34).

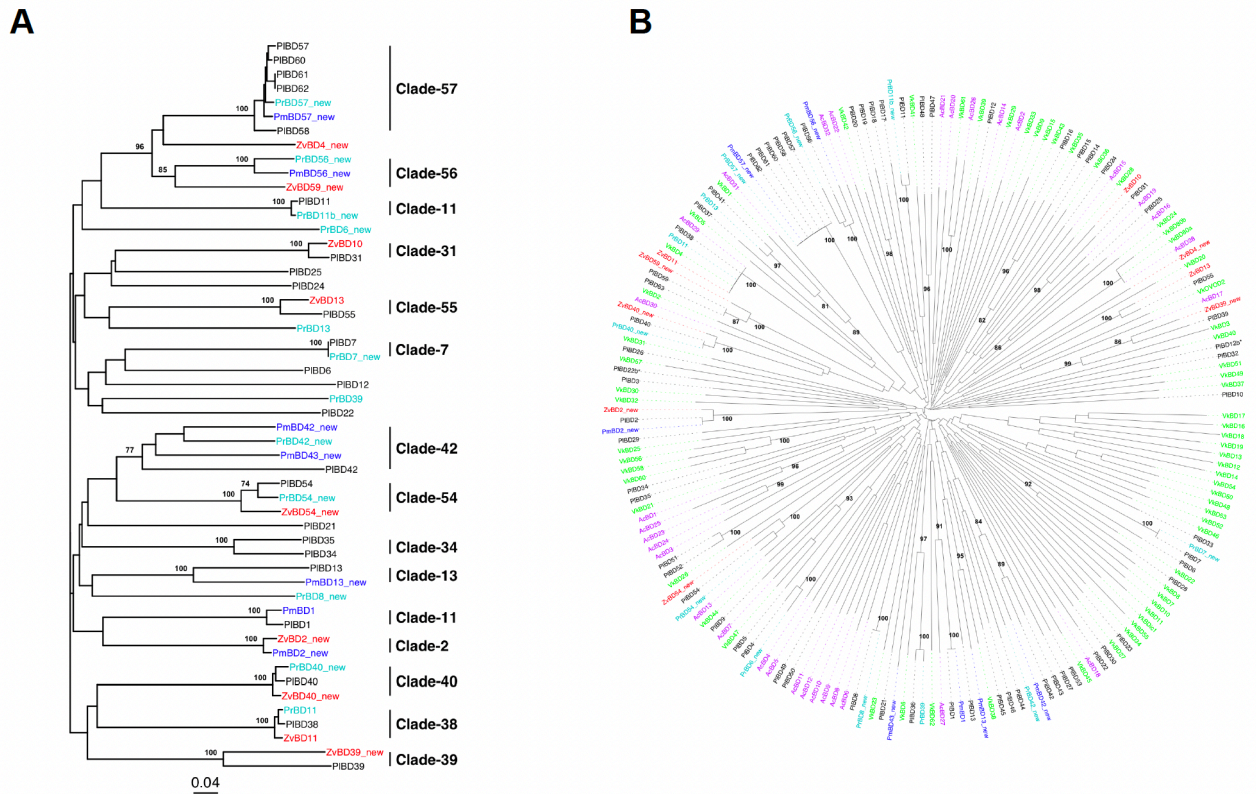


Figure 4. Beta-defensin comparative analysis across multiple related reptile species. A) NJ tree based on complete proteins (N=47). Major clusters (bootstrap support>70) are highlighted and named according to the PIBD reference within the clade. B) NJ tree based on the shared cysteine domain from all five species (C-X(6-7)-C-X(3-7)-C-X(8-14)-C-X(4-7)-CC) (N=190 sequences), highlighting major clusters (bootstrap support>80). Proteins were named according to their public annotation; if uncharacterized, they were numbered according to their closest *P. lilfordi* sequence and marked with “new” (see Supplemental Table S3 for species AccNos). *Pl*: *P. lilfordi*; *Pm*: *P. muralis*; *Pr*: *P. raffonei*; *Zv*: *Z. viviparus*; *Vk*: *V. komodoensis*; *Ac*: *A. carolinensis*.

Clades 54, 56 and 57 reflect species relative distances consistent with their taxonomic grouping: all *Podarcis* spp. cluster together, with *Zv* positioned at a basal node, supporting gene orthology. Most other clades do not provide species or genus resolution: these include clade-38 and 40 (grouping *Podarcis* spp. and *Zv*), clade-1 (*Pm*-*Pl*), clade-2 (*Pm*-*Zv*) and clade-7 and 11 (*Pl*-*Pr*) (Figure 4A). Interestingly, clade-7 shows full identity of BD7 between *Pr* and *Pl*, suggesting putative convergent evolution (but see Discussion).

Due to a potential incomplete sampling of the full repertoire of beta-defensin full proteins from *Podarcis* spp. and *Zv*, we further explored pattern of BD evolution in reptiles based on cysteine motifs only, using a larger dataset that encompassed also the Komodo dragon and the green anole (N=190 sequences: Pl=65, Pm = 7, Pr= 12, Zv = 9, *A. carolinensis* (Ac) = 32, and *V. komodoensis* (Vk)= 65) (Supplemental Table S3). The NJ tree (Figure 4B) highlights three main aspects: within species recent duplications, species-specific lineages, and across species peptide convergence. Aside from instances of duplication in Pl, no other recent intraspecific duplication was observed for the any other species, except for the previously reported duplication in Vk (i.e. VkBD80a and b). Moreover, most Ac and Vk lineages appear to be species-specific. Nonetheless, both patterns could also result from pervasive gene loss and/or incomplete sampling of beta-defensin diversity in all six species under study, unclear at present. Notably, we identified eight clusters showing motif identity across distinct species (Figure 4B). Such instances of motif identity are found for all species pairwise, although more frequently within the *Podarcis* genus (Pl-Pr: five clades; Pm-Pl: two clades; Ac-Vk: two clades, Pl-Pr-Zv-Ac-Vk: one clade). Remarkably, the clade encompassing AcBD29, PIBD38, VkBD4, ZvBD11 and PrBD11 (corresponding to clade-38 in Figure 4A) shows motif identity among all species but Pm, thus encompassing all three reptilian families under study (Lacertidae, Varanidae and Dactyloidae).

3.4. Characterization of *P. lilfordi* ovo-defensins

We detected a total of eight complete ovo-defensin proteins in chr-3 of *P. lilfordi*, tightly clustered within or near the region flanked by MTMR9 and XKR6 proteins (Figure 1 and Supplemental Table S4). Protein size ranges from 64 to 239 AA arranged in two exons: the first exon putatively encodes for the signal peptide (about 21 AA), whereas the second exon includes the putative AMP (Figure 5).

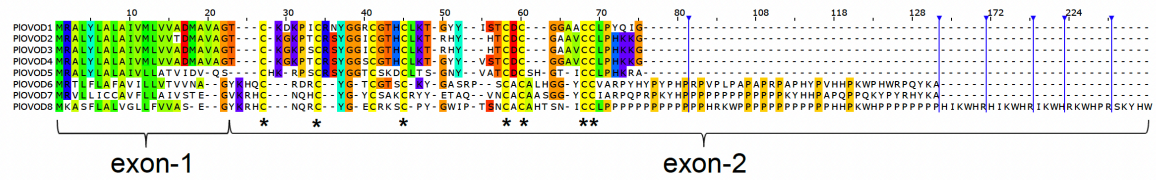


Figure 5. Alignment of *P. lilfordi* complete ovo-defensin proteins (N=8). All proteins share a two-exons structure with a common intron position and the eight-cysteine motif C-X(3-5)-C-X(3-6)-C-X(3-6)-C-X(8-9)-C-X1-C-X(4-6)-CC, with a positive net charge. Blue lines mark the start of long stretches of prolines (>10) at OVOD7 and OVOD8. Shared cysteine residues are highlighted with an asterisk.

Intron size is relatively small, ranging from 1.3 kb and 2.3 kb, except for OVOD1, which has substantially larger intron (3.5 kb). All OVODs show an eight-cysteine motif with a highly conserved spacing: C-X(3-5)-C-X(3-6)-C-X3-C-X(8-9)-C-X1-C-X(4-6)-CC. PIOVOD7 contains two additional CC in the signal peptide. All peptides show a conserved Tyrosine (Y) and a Glycine (G) at position 37 and 38, respectively (Figure 5). PIOVOD6-8 are proline-rich proteins (PrAMPs), with PIOVOD7-8 displaying long stretches of up to 31 consecutive prolines (see Figure 5 for position of proline repeats and Supplemental Table S4 for complete sequences). In PIOVOD7-8, proline repeats are interspersed with repetitive lysine-tryptophan-histidine (KWH) and lysine-tyrosine-histidine (KYH) motifs. Transcriptome data indicate that two of the OVODs (PIOVOD1 and PIOVOD4) are expressed (see Supplemental Table S4).

All OVODs exhibit a highly cationic charge (3.5 to 20) at exon 2 (i.e. the putative AMP), with isoelectric points ranging from 8.1 and 10.5 (Supplemental Table S4). Protein homology search revealed low-confidence assignment (<71%) and coverage (<58%) for most peptides (Supplemental Table S2). The best hit was given for PIOVOD5 to a hydrolase/hydrolase inhibitor (carboxypeptidase inhibitor) (96.8% of confidence and 57% of coverage), and for PIOVOD1 which matched an antimicrobial peptide (82% and 52%) (Supplemental Table S2). The remaining proteins did show any significant matches (>75% of confidence), likely due to the limited classification currently available for this group of defensins and their intrinsically disordered structures.

PIOVOD2-4 share large similarity (90-97% of AA identity), with PIOVOD2 and PIOVOD3 differing by a single AA (Figure 5) likely due to a recent duplication event. OVODs that are located proximally on the chromosome tend to share higher protein similarity (Figure

1), consistent with a duplication in tandem, as previously observed for BDs. PLOVOD1-5 show little homology with PLOVOD6-8 (i.e., OVOD-PrAMPs), with only 7-24% of AA identity between them. Apart from a conserved eight-cysteine residue motif, a two-exon structure, and a closed proximity on chr-3 (Figure 1), these two groups of OVODs may represent distinct protein families or subfamilies.

3.5. Comparative analysis of ovo-defensins across reptiles

Through extensive BLAST searches and genome mining, we successfully retrieved several complete PLOVOD orthologs/paralogs from closely related species, none of which has been previously identified: five from *Zv*, four from *Pm*, and eight from *Pr* (Supplemental Table S3). Additionally, we included three previously described partial OVODs from *Vk* that met the minimum requirement of eight cysteine residues (i.e., *VkOVOD3*, 4 and 6). Due to the high fragmentation of current genome draft for *Varanus*, we were unable to retrieve the full protein sequences for this species.

The eight-cysteine motif was highly conserved across all reptile species, with a consistent spacing between the third and fourth cysteine, and between the fifth and the sixth cysteine: C-X(3-6)-C-X(3-6)-C-X3-C-X(8-12)-C-X1-C-X(4-6)-CC (Supplemental Figure S2). Notably, OVOD6-8 are proline-rich proteins in all Lacertidae, with OVOD7-8 showing long stretches of prolines interspaced by the conserved motif repeats KW(H) and KYH in all species (for *Pm*, only in OVOD8) (Supplemental Table S3 and Figure S2).

Based on the NJ tree constructed from complete proteins, we identified five main clades (A-D, bootstrap support>80) (Figure 6).

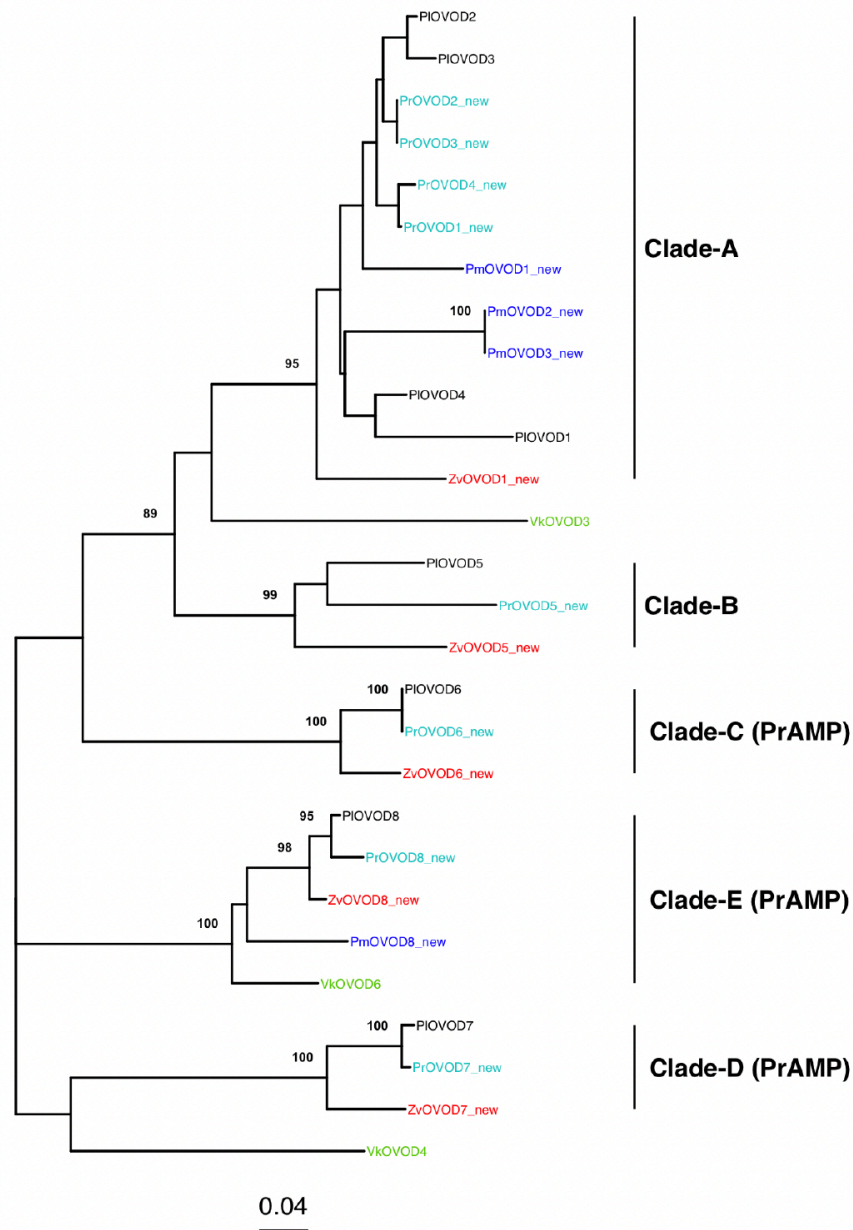


Figure 6. NJ tree based on complete ovo-defensin proteins (partial only for Vk) from multiple related reptile species (N=17). All species shared the following cysteine motif: C-X(3-6)-C-X(3-6)-C-X3-C-X(8-12)-C-X1-C-X(4-6)-CC. Major clades (bootstrap support>80) are labelled as A–E. Recent gene duplications are framed. New proteins were marked with “new” and numbered according to their closest PI sequence for Zv or to their sequential order in the chromosome for Pr and Pm (see Supplemental Table S3). *Pl*: *P. lilfordi*; *Pm*: *P. muralis*; *Pr*: *P. raffonei*; *Zv*: *Z. viviparus*; *Vk*: *V. komodoensis*.

Within Clade A, we observe several instances of recent intraspecific duplication within all *Podarcis* spp., including Pl (OVOD2-3, with one AA difference, and OVOD1-4), Pr (OVOD2-3, OVOD1-4, each pair sharing identical cysteine motifs) and Pm (OVOD2-3 with full protein identity). Clade A and B group together with high support, suggesting a common origin from a past lineage duplication. The remaining three clades (C-E) include the proline-rich proteins OVOD6-8, all of which have a one-to-one ortholog in all Lacertidae, with no recent duplications. OVOD8 shows cysteine motif identity (though not full gene identity) across Pl, Pr, Pm and Zv, and OVOD6 shows identity between Pl and Pr (Supplemental Figure S2).

3.6. Characterization of *P. lilfordi* cathelicidins

We identified four complete cathelicidin proteins on chr-12 of *P. lilfordi* (PICATH1-4), flanked by KLHL18 and FASTK proteins (Figure 1 and Supplemental Table S5). PICATH2 was labelled as PICATH2-OH following the annotation of a closed related protein in Zv, named as ZvCATH2-OH to maintain the previous annotation from the king cobra (*Ophiophagus hannah*), where this protein was first described [27]. PICATH1 and PICATH2-OH are separated by only 266 bp, although located on opposite strands (Supplemental Table S5).

Protein size of the four CATHs range from 161 to 171 AA and are all organized into four exons (Figure 7). All four CATHs were expressed according to transcriptome data. They exhibit a conserved four-cysteines domain with a highly conserved spacing: C-X10-C-X10-C-X(16-19)-C.

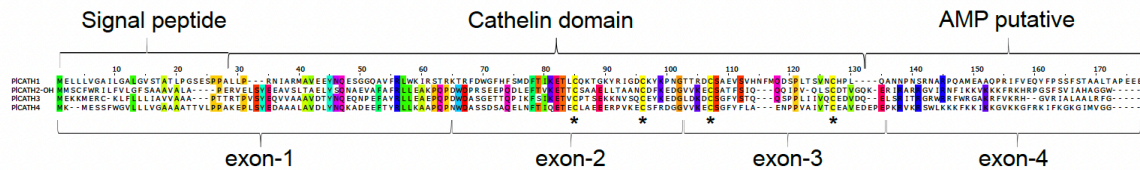


Figure 7. Alignment of *P. lilfordi* complete cathelicidin proteins (inactive precursor) (N=4). All proteins share a four-exons structure and the four-cysteine motif C-X10-C-X10-C-X(6-19)-C. Main domain structure (signal peptide, cathelin domain and putative AMP) was predicted by homology searches. Shared cysteine residues are highlighted with an asterisk.

Homology searches returned best matches to the porcine CATHs, with 99.9% of confidence and coverage above 50% for all four proteins (Supplemental Table S2). Folding predictions indicate that all four PICATHs represent inactive precursors, structured into three domains: 1) a variable N-terminal domain with a signal peptide of slightly varying length (first 23-27 AA), 2) a conserved "cathelin-like domain" (CLD) containing the cysteine motif (100-101 AA long), with one alpha helix and four beta-sheets, and 3) a highly variable C-terminal domain containing the putative functional AMP (36-42 AA), with proteolytic cleavage predicted at position 130 (Figure 7).

Except for PICATH1, all other AMP domains present a highly cationic charge (11.5-15) due to multiple Lys (K) and Arg (R) residues, with isoelectric points ranging between 12 to 13 (Supplemental Table S5). The AMP domain from PICATH1 largely differs in AA content from all other CATHs, lacking K and G residues and showing a net negative charge (-2).

Overall protein similarity among the four CATHs is low (<40%), indicating no recent duplications.

3.7. Comparative analysis of cathelicidins across reptiles

For comparative purposes, we retrieved all available CATHs from the Lacertidae under study and additionally identified six novel ones: five from Pm (one newly identified), four from Pr (three new), and four from Zv (two new). We also included the three previously described CATHs from Vk and Ac. Most of the novel CATHs were previously annotated as uncharacterized mRNA (see Supplemental Table S3 for AccNos).

During sequence alignment, we found that a single CATH-coding mRNA sequence from Pm, Pr and Zv contained two consecutive cathelicidin proteins: the first protein showed high similarity to PICATH4, followed by a second protein with high similarity to PICATH3. Coexpression of CATH4-CATH3 was not observed in Pl, where the two proteins are separated by 5901 bp (Figure 1 and Supplemental Table S5). For alignment purposes, we manually split the mRNA to recover the individual proteins and aligned them along with the other sequences.

The multispecies alignment indicates that all reptiles under study share a highly conserved four-cysteine motif: C-X(9-10)-C-X10-C-X(17-21)-C (Supplemental Figure S3). The Valine-130 (following the last cysteine residue), predicted to be the site of proteolytic

cleavage for the active peptide in PI, is also conserved across all reptiles in CATH2-4, but not in CATH1 (Supplemental Figure S3).

The NJ tree based on full proteins (Figure 8) clearly depicts four highly supported clusters of one-to-one orthologous proteins (Clade 1-4), largely reflecting the main phylogenetic position of the two Lacertidae genera: all *Podarcis* species together (with no further resolution at species level), while *Zv* locates at a basal node.

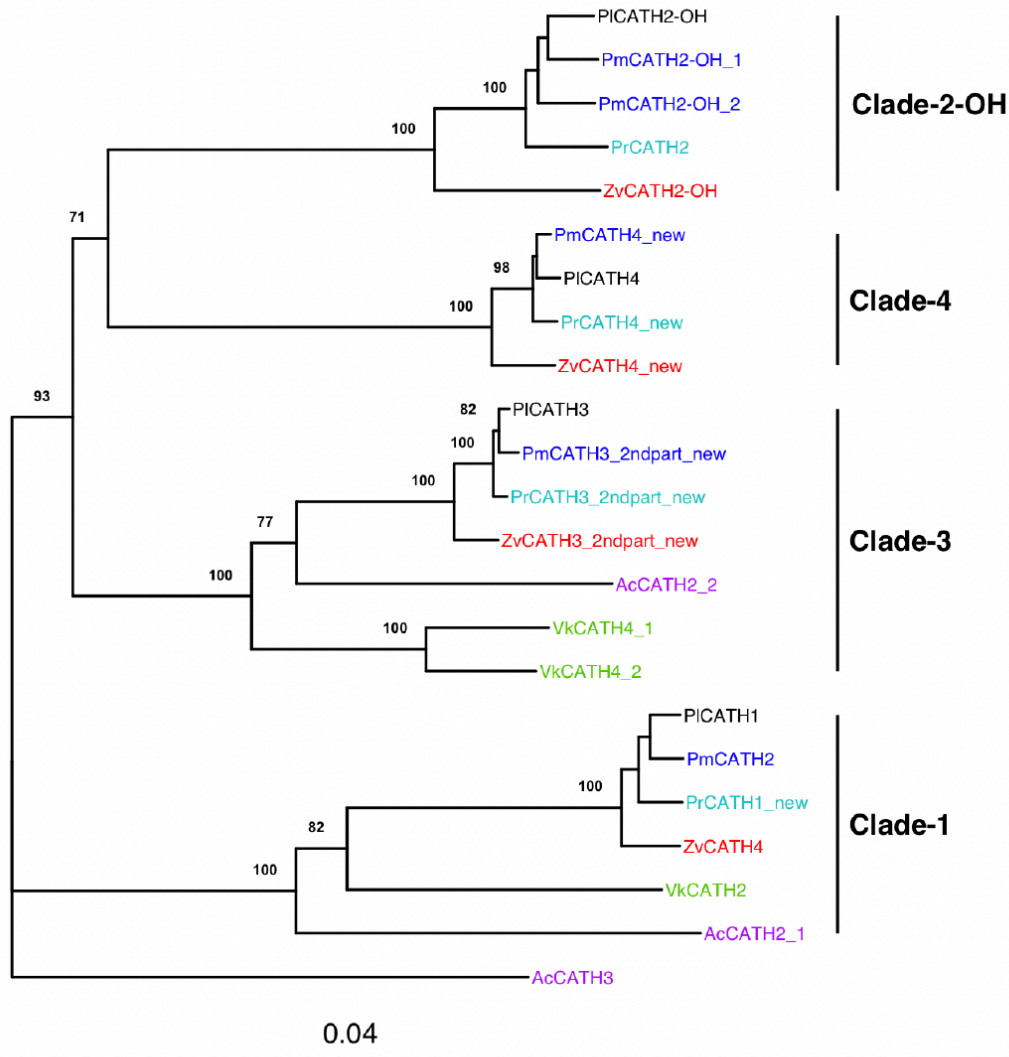


Figure 8. NJ tree based on complete cathelicidin proteins from multiple related reptile species (N=23). All reptiles shared the four-cysteine motif C-X(9-10)-C-X10-C-X(17-21)-C. Major clades (bootstrap support>70) are labelled based on the *P. lilfordi* reference protein.

Moreover, the species Ac and Vk are basal to both clade 1 and 3, with Vk presenting two closely related paralogs in CATH3. Both species lack orthologues to clade 2-OH or clade 4. However, given the incomplete genome assembly of *V. komodoensis*, this absence may also be attributed to a failure in sequence identification for this species. A recent CATH duplication is observed only in Pm, with two proteins falling within clade 2-OH.

4. Discussion

The innate immune system of lizards remains a largely understudied area of research (Field et al., 2022; van Hoek, 2014). In this study we provided the first comprehensive genomic characterization of major AMP classes in *P. lilfordi* and closely related species within the family Lacertidae, with a comparative analysis extended to members of other Squamata families, the Komodo dragon and the green anole.

For all Squamata species under study, AMPs consistently clustered in two chromosomes: chr-3 for beta- and ovo-defensins, and chr-12 for cathelicidins, making these regions important repositories for the reptile antimicrobial defense. This genomic organization has been previously reported in crocodiles and birds, but not in mammals (Santana et al., 2021), supporting an avian and reptilian monophyletic origin. Within these regions, *P. lilfordi* flanking proteins of AMP clusters (i.e., XPO1 and CTSB for beta-defensins, XKR6 and MTMR9 for ovo-defensins, and FASTK and KHLL18 for cathelicidins) were consistent with those previously described in other reptiles, but also mammals (Santana et al., 2021; Dalla Valle et al., 2012; van Hoek, 2014), underscoring their importance as critical reference markers for AMP cluster identification in vertebrates.

4.1. Beta-defensins diversity and evolution in reptiles

Beta-defensins represent the most expanded and diverse AMP family in vertebrates (Tu et al., 2015). Currently, no studies are available for Lacertidae, with Komodo dragon and the green anole being the closest species with a comprehensive characterization of beta-defensins (van Hoek et al., 2019; Dalla Valle et al., 2012).

In *P. lilfordi*, the diversity of beta-defensins (65 peptides, 24 with complete ORFs) largely aligned with that previously described in the green anole (N = 32) (Dalla Valle et al., 2012) and the Komodo dragon (N = 65) (van Hoek et al., 2019). All proteins presented substantial diversity in amino acid sequence, length, molecular weight, and isoelectric points. However, they all shared the canonical six-cysteine motif and a predominantly cationic charge, as reported for all vertebrates (Semple & Dorin, 2012; Tu et al., 2015;

Nava et al., 2009). Moreover, complete beta-defensins exhibited broad structural variation, with two to four exons (two being the most recurrent), consistent with observations in the green anole (Dalla Valle et al., 2012) and the Komodo dragon (van Hoek et al., 2019).

In *P. lilfordi*, we observed that sequence similarity was typically higher among chromosomally adjacent peptides or full proteins, suggesting a process of paralogous expansion via duplication in tandem. Instances of identity or near identity among pairs of contiguous peptides further indicate that this process is still ongoing. These observations align with previous research highlighting a major role of gene duplication in the increasing diversity of this gene family (Lazzaro et al., 2020; Semple & Dorin, 2012; Tu et al., 2015). Notably, we also identified instances of beta-defensins (PIBD12 and PIBD22) carrying two consecutive cysteine motifs at the C-terminal, a structural organization that was previously observed in other reptiles, but not in mammals (Guyot et al., 2022), and that suggests the concomitant expression of multiple AMPs. Given the low sequence similarity between the two consecutive domains, these double motifs are unlikely to result from intra-protein duplication in tandem; rather they most likely arose from sequence translocation or other chromosomal rearrangements that remain to be clarified.

In line with this complex pattern of intra-genomic diversification, comparative beta-defensin analyses across reptiles recovered a generally low phylogenetic signal due to pervasive gene duplication, as well as putative gene loss and rapid diversification at the active peptide (Tennessen, 2005). While the extensive AMP genome screening of *P. lilfordi* provided a nearly complete catalogue of beta-defensins for this species, the present AMP characterization in the other Lacertidae genomes may be incomplete, limiting our ability to infer the relative role of gene loss and species-specific gene duplication in the observed diversity. Nonetheless, we identified several clades of putative true orthologs across all Lacertidae (Figure 4A), supporting a substantial level of evolutionary conservation for some beta-defensins, as previously observed in mammals (Lazzaro et al., 2020; Nava et al., 2009). These clades showed variable branch lengths, suggesting heterogeneous rates of protein changes along species-specific lineages, likely driven by protein niche functional specialization in host–pathogen interactions (Chakraborty et al., 2022; Semple et al., 2006). However, identification of true orthologues is needed to properly address rates of evolution and selection.

Interestingly, we observed several instances of cysteine motif identity or high similarity between members of the genus *Podarcis*, particularly between the two closely related species *P. lilfordi* and *P. raffonei* (Yang et al., 2022). Instances of identical peptides were also observed among members of distantly related genera and families, for example, a beta-defensin peptide from clade-38 was shared across all three studied families: Lacertidae, Dactyloidae, and Varanidae (Figure 4B). In the green anole, this anionic peptide (i.e. AcBD29) is expressed in all tissues, although its function remains unknown (Dalla Valle et al., 2012). Similarly, the beta-defensin peptide AcBD27, which presents high similarity to clade-1 peptides from *P. lilfordi* (PIBD1) and *P. muralis* (PmBD1), is known to be highly expressed in wounded lizards, particularly following tail loss, where it plays a role in preventing infection in the absence of inflammation, a key factor in successful regeneration (Lai et al., 2019; Nguyen et al., 2011). These findings underscore the need for species-specific studies to clarify whether identical or highly similar beta-defensin peptides or specific polymorphisms arise in parallel and serve functionally equivalent roles across different species.

Convergent evolution of AMPs has been previously reported in both vertebrates and invertebrates, with multiple studies documenting convergence of critical residues under similar antimicrobial pressures across diverse taxa (Lazzaro et al., 2020; Unckless & Lazzaro, 2016). In anurans, for instance, equivalent signal motifs are thought to have arisen independently in three distinct evolutionary lineages (König & Bininda-Emonds, 2011). While the instances of identical peptides observed across our studied reptiles may suggest convergent evolution, establishing a framework of true one-to-one orthologs is essential to rule out alternative explanations such as ancestral polymorphism. Population-level studies are also needed to assess allele diversity and investigate the potential role of balancing selection in polymorphism maintenance (Lazzaro et al., 2020). Finally, phenotypic analyses targeting specific residues or motifs will be critical to clarify selective pressures and the molecular mechanisms underlying AMP evolution.

4.2. Ovo-defensins and PrAMPs diversity and evolution in Lacertidae

Unlike beta-defensins, ovo-defensins represent a smaller subfamily of still largely understudied defensins, currently described only in birds and a few reptiles (but none in Lacertidae), where they appear to be specifically expressed in the oviduct and contribute to egg immune defense (Whenham et al., 2015; van Hoek et al., 2019; Zhang et al., 2019).

The eight OVODs identified in *P. lilfordi* were all located within the same chr-3 cluster region, flanked by conserved marker genes, MTMR9 and XKR6, although PIOVOD1 was located immediately outside the gene cluster. This chromosomal arrangement was previously described in other vertebrates (Zhang et al., 2019), supporting a monophyletic origin for this subfamily. We confirmed the two-exon structure and the higher divergence at exon-2, as previously reported for avian and reptilian ovo-defensins (Whenham et al., 2015). In contrast to beta-defensins, OVODs present two extra cysteines (eight in total), which contrasts with previous reports of a six-cysteine motif for OVODs in reptiles (Whenham et al., 2015). Notably, this eight-cys motif and two-exon structure was observed in all reptiles under study (both Lacertidae species and the Komodo dragon), strongly suggesting that these features represent specific structural characteristics of this protein subfamily. Differences in the number and arrangement of cysteines between beta- and ovo-defensins may underscore fundamental functional differences that are currently unclear (Erdem Büyükkiraz & Kesmen, 2022).

Despite structural commonalities and a similar chromosomal arrangement, the eight PIOVODs encompassed a remarkable amino acid diversity, suggesting the potential for further categorization into distinct defensin subfamilies. While ovo-defensins from clade A appear to have undergone an extensive process of gene expansion through duplication in tandem in all *Podarcis* species (Figure 6), the remaining clades (B–E) show a single ortholog for each species, indicating relative evolutionary stability. We note that presence of OVODs in the viviparous species *Z. vivipara* suggests that protein expression may not be restricted to eggs.

Interestingly, these clades encompass PrAMPs (Lai et al., 2019), containing either multiple non-adjacent prolines (P) (OVOD6) or proline repeats at the C-terminal, typically interspersed with motifs rich in lysine (K), tryptophan (W), histidine (H), arginine (R) and tyrosine (Y) (OVOD7 and OVOD8) (Figure 5). Their structure and lack of any significant protein folding suggest they may represent “extended AMPs,” a class of AMPs known for containing high proportions of proline (P) or tryptophan (W) residues and lacking regular secondary structure elements (Nguyen et al., 2011). Notably, both P- and W-rich peptides are known for their potent antimicrobial activity due to their ability to permeate the outer membrane and additionally target intracellular pathways (Lai et al., 2019; Mishra et al., 2018; Hernandez-Gordillo et al., 2013). For instance, recent studies have shown that the polyproline host defense peptide bactenecin 7 (Bac7) can bind to the wall

polysaccharide biofilm of hypermucoviscous bacteria, such as *Klebsiella pneumoniae*, actively disrupting the bacterial cell wall (De Los Santos et al., 2024; Fleeman & Davies, 2022). Additionally, KW repeats (observed here in OVOD8) have been shown to enhance AMP bactericidal activity (Gopal et al., 2013), while KY residues (seen in OVOD7) cooperate with H to facilitate DNA cleavage in DNase proteins (Korn et al., 2002).

Interestingly, we found several instances of cys-motif identity in OVOD-PrAMPs across species, with OVOD8 carrying the same motif in all Lacertidae species. As observed for beta-defensins, this suggests a putative role of convergent evolution in shaping PrAMPs diversity, although alternative scenarios remain to be explored.

4.3. Cathelicidins diversity and evolution in Lacertidae

Cathelicidins are a small and ancient family of potent AMPs, known for their functional plasticity and broad spectrum of activity against a wide range of microorganisms (Wang, 2014; van Harten et al., 2018; Whang et al., 2022). To date, they have been described in several reptile species within the order Squamata (Whang et al., 2022), including the Komodo dragon (van Hoek et al., 2019), but no formal description exists for any Lacertidae species, apart from gene annotation of few representative cathelicidin proteins in public genomes.

In *P. lilfordi*, we identified four cathelicidin proteins (CATH1–4), all of which are tightly clustered on chr-12, flanked by the conserved genes FASTK and KLHL18, previously reported in birds and reptiles (van Hoek et al., 2019; Whang et al., 2022; Zhang et al., 2019). These proteins all exhibit the classical four-exon precursor structure common to vertebrates, characterized by a variable signal peptide, a conserved four-cysteine motif, and a highly divergent C-terminal that carries the active peptide (Chang et al., 2006; Xiao et al., 2006; Tomasinsig & Zanetti, 2005). Of the four cathelicidins, three present mature peptides with a highly cationic net charge (>10), in line with the average positive charge found for reptiles, including the Komodo dragon (van Harten et al., 2018). Interestingly, CATH1 is the only protein presenting an anionic AMP across all reptiles studied, with a net charge of -1 to -2 and a significantly distinct AA composition, suggesting it may have a rather different mechanism of action compared to the other cathelicidins. To date, a single anionic cathelicidin has been formally described and functionally characterized in salamanders, where it is produced by the skin and exhibits potent anti-inflammatory and wound-healing activities, although not direct antimicrobial effects (Zhao et al., 2009). Functional studies are required to assess whether this cathelicidin plays a similar role in

lizard wound healing following injuries or tail autotomy, a conserved anti-predatory strategy observed in all lizard species (Bateman & Fleming, 2009).

All Lacertidae species studied presented four cathelicidins, except for *P. muralis* (with five in total), all located on chr-12, with a one-to-one ortholog in all species, supporting a monophyletic origin and phylogenetic conservation within this taxonomic group. Three of these orthologous clades (CATH2–4) clustered into a major, well-supported clade, including also representatives of the Komodo dragon and the green anole, and pointing to a common origin by a past duplication within Squamata. A fifth CATH was identified in *P. muralis*, resulting from a recent duplication in Clade-2-OH. This event indicates that the process of gene duplication in cathelicidins is still ongoing, although it is less pervasive than in defensins (Whang et al., 2022). Interestingly, we found that CATH3 and CATH4 were coexpressed as a single large RNA in all Lacertidae, except in *P. lilfordi*, suggesting a further post-translational process to concomitantly generate the two distinct active peptides. We found no previous reports in literature describing such an RNA arrangement for cathelicidins. However, we observed similar instances of a multidomain structure in beta-defensins (see above), suggesting it might provide some functional gain, unclear at present (Mookherjee et al., 2020). In line with this, recent AMP coexpression models in experimental systems have shown enhanced antimicrobial activity and immunity (Li et al., 2023; He et al., 2021), pointing to potential cooperative interactions among naturally coexpressed AMPs, which warrants further investigation.

5. Conclusions

Our findings substantially expand current knowledge on the diversity of AMPs in reptiles and their mechanisms of evolution, characterized by gene duplications and rapid evolution.

For the Lilford's wall lizard, the identification of key players in innate immunity sets the basis for exploring the role of AMPs in their eco-immunology and the potential responses of these insular populations to rising infection threats associated with increasing anthropogenic impacts on these islands. Healthy *P. lilfordi* populations play significant roles in their ecosystems and a fine understanding of their genetic potential in immune response can offer new insights into resilience mechanisms, guiding species-specific conservation efforts for these isolated populations.

The comprehensive characterization of major AMP classes in *P. lilfordi* will also offer a valuable resource for fine-scale annotation of these peptides in other reptile species,

facilitating comparative studies across Squamata and guiding a comprehensive understanding of lineage-specific patterns of gene duplication/loss. Particularly, the identified AMP cysteine motif identity and whole-protein similarity across species lay the groundwork for exploring conserved mechanisms of antimicrobial action in reptiles and assessing their value as candidates for novel therapeutic applications.

Nonetheless, we also emphasize that relying on a single reference genome per studied species currently limits our ability to fully capture intraspecific AMP diversity. This limitation constrains our conclusions regarding the extent of gene duplications and losses, the impact of convergent evolution and the role of balancing-selection in maintaining polymorphisms. Population-level studies will be critical to address these aspects. In this context, future integrative analyses combining population genomics and expression data from multiple island populations of Lilford's wall lizard offer an exciting opportunity to uncover key evolutionary trajectories and functional roles of AMPs in lizards.

Declarations

Availability of data and materials

All sequence data analysed in this study are included in the Supplementary Tables, along with the AccNos for published sequences. All complete nucleotide and protein sequences newly described for *P. lilfordi* have been uploaded at European Nucleotide Archive (ENA) under AccNo PRJEB89980. All complete and partial *P. lilfordi* AMPs are also available in the additional file.

Funding

This study was funded by MICIU/AEI/ 10.13039/501100011033 and by “ERDF A way of making Europe”, grant PID2022-141578NB-C22 to LB. Doctoral support to KO was from the Doctoral Scholarship, Colombian Ministry of Science, Technology and Innovation (MINCIENCIAS885/2020 to K.O.).

Authors' contributions

LB conceived the study. LB, KO and JG analyzed the results. LB and KO wrote the paper. All authors read and approved the final manuscript.

Acknowledgements

Institutional support to CNAG was given from the Spanish Ministry of Science, Innovation and Universities, Fondo de Investigaciones Sanitarias cofunded with ERDF funds (PI19/01772), the 2014-2020 Smart Growth Operating Program, and the Generalitat de Catalunya through the Departament de Recerca i Universitats and Departament de Salut.

References

1. Lazzaro BP, Zasloff M, Rolff J. Antimicrobial peptides: Application informed by evolution. *Science*. 2020;368:eaau5480.
2. Huan Y, Kong Q, Mou H, Yi H. Antimicrobial peptides: Classification, design, application and research progress in multiple fields. *Front Microbiol*. 2020;11:582779.
3. Chakraborty S, Chatterjee R, Chakravorty D. Evolving and assembling to pierce through: Evolutionary and structural aspects of antimicrobial peptides. *Comput Struct Biotechnol J*. 2022;20:2247–58.
4. Erdem Büyükkiraz M, Kesmen Z. Antimicrobial peptides (AMPs): A promising class of antimicrobial compounds. *J Appl Microbiol*. 2022;132:1573–96.
5. Cardoso MH, Meneguetti BT, Oliveira-Júnior NG, Macedo MLR, Franco OL. Antimicrobial peptide production in response to gut microbiota imbalance. *Peptides*. 2022;157:170865.
6. Zhang LJ, Gallo RL. Antimicrobial peptides. *Current biology : CB*. 2016;26.
7. Harris F, Dennison SR, Phoenix DA. Anionic antimicrobial peptides from eukaryotic organisms. *Curr Protein Pept Sci*. 2009;10:585–606.
8. Semple CAM, Maxwell A, Gautier P, Kilanowski FM, Eastwood H, Barran PE, et al. The complexity of selection at the major primate beta-defensin locus. *BMC Evol Biol*. 2005;5:32.
9. Hanson MA. When the microbiome shapes the host: immune evolution implications for infectious disease. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2024;379:20230061.
10. Benítez-Prián M, Lorente-Martínez H, Agorreta A, Gower DJ, Wilkinson M, Roelants K, et al. Diversity and Molecular Evolution of Antimicrobial Peptides in Caecilian Amphibians. *Toxins*. 2024;16:150.
11. Semple CA, Gautier P, Taylor K, Dorin JR. The changing of the guard: Molecular diversity and rapid evolution of beta-defensins. *Mol Divers*. 2006;10:575–84.
12. Semple F, Dorin JR. β -Defensins: multifunctional modulators of infection, inflammation and more? *J Innate Immun*. 2012;4:337–48.
13. Zhu & Gao. Evolutionary origin of β -defensins. *Developmental & Comparative Immunology*. 2013;39:79–84.

14. Tu J, Li D, Li Q, Zhang L, Zhu Q, Gaur U, et al. Molecular Evolutionary Analysis of β -Defensin Peptides in Vertebrates. *Evolutionary Bioinformatics Online*. 2015;11:105.
15. Whenham N, Lu TC, Maidin MBM, Wilson PW, Bain MM, Stevenson ML, et al. Ovo-defensins, an oviduct-specific antimicrobial gene family, have evolved in birds and reptiles to protect the egg by both sequence and intra-six-cysteine sequence motif spacing. *Biol Reprod*. 2015;92:154.
16. Gong D, Wilson PW, Bain MM, McDade K, Kalina J, Hervé-Grépinet V, et al. Gallin; an antimicrobial peptide member of a new avian defensin family, the ovo-defensins, has been subject to recent gene duplication. *BMC Immunology*. 2010;11:1–15.
17. Kościuczuk EM, Lisowski P, Jarczak J, Strzałkowska N, Jóźwik A, Horbańczuk J, et al. Cathelicidins: family of antimicrobial peptides. A review. *Mol Biol Rep*. 2012;39:10957–70.
18. Bals R, Wilson JM. Cathelicidins - a family of multifunctional antimicrobial peptides. *Cellular and Molecular Life Sciences CMLS*. 2003;60:711–20.
19. Tomasinsig L, Zanetti M. The cathelicidins--structure, function and evolution. *Current protein & peptide science*. 2005;6.
20. Wang G, Li X, Wang Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res*. 2016;44:D1087–93.
21. Field EK, Hartzheim A, Terry J, Dawson G, Haydt N, Neuman-Lee LA. Reptilian Innate Immunology and Ecoimmunology: What Do We Know and Where Are We Going? *Integrative and comparative biology*. 2022;62.
22. Buchmann K. Evolution of Innate Immunity: Clues from Invertebrates via Fish to Mammals. *Front Immunol*. 2014;5:108394.
23. Santana FL, Estrada K, Ortiz E, Corzo G. Reptilian β -defensins: Expanding the repertoire of known crocodylian peptides. *Peptides*. 2021;136.
24. Siddiqui R, Maciver SK, Khan NA. Gut microbiome-immune system interaction in reptiles. *Journal of applied microbiology*. 2022;132.
25. van Hoek ML, Prickett MD, Settlege RE, Kang L, Michalak P, Vliet KA, et al. The Komodo dragon (*Varanus komodoensis*) genome and identification of innate immunity genes and clusters. *BMC Genomics*. 2019;20:1–18.
26. Dalla Valle L, Benato F, Maistro S, Quinzani S, & Alibardi L. Bioinformatic and molecular characterization of beta-defensins-like peptides isolated from the green lizard *Anolis carolinensis*. *Developmental & Comparative Immunology*. 2012;36:222–9.

27. Zhao H, Gan TX, Liu XD, Jin Y, Lee WH, Shen JH, et al. Identification and characterization of novel reptile cathelicidins from elapid snakes. *Peptides*. 2008;29.
28. Kim D, Soundrarajan N, Lee J, Cho HS, Choi M, Cha SY, et al. Genomewide Analysis of the Antimicrobial Peptides in *Python bivittatus* and Characterization of Cathelicidins with Potent Antimicrobial Activity and Low Cytotoxicity. *Antimicrobial agents and chemotherapy*. 2017;61.
29. Andrade P, Pinho C, Pérez I de Lanuza G, Afonso S, Brejcha J, Rubin C-J, et al. Regulatory changes in pterin and carotenoid genes underlie balanced color polymorphisms in the wall lizard. *Proc Natl Acad Sci U S A*. 2019;116:5633–42.
30. Gabrielli M, Benazzo A, Biello R, Ancona L, Fuselli S, Iannucci A, et al. A high-quality reference genome for the critically endangered Aeolian wall lizard, *Podarcis raffonei*. *The Journal of heredity*. 2023;114.
31. Yurchenko AA, Recknagel H, Elmer KR. Chromosome-Level Assembly of the Common Lizard (*Zootoca vivipara*) Genome. *Genome Biol Evol*. 2020;12:1953–60.
32. Gomez-Garrido J, Cruz F, Alioto TS, Feiner N, Uller T, Gut M, et al. Chromosome-level genome assembly of Lilford's wall lizard, *Podarcis lilfordi* (Günther, 1874) from the Balearic Islands (Spain). *DNA Res*. 2023;30.
33. Rotger A, Igual JM, Smith JJ, Tavecchia G. Relative role of population density and climatic factors in shaping the body growth rate of Lilford's Wall Lizard (*Podarcis lilfordi*). *Can J Zool*. 2016;94:207–15.
34. Otalora K, Riera JL, Tavecchia G, Rotger A, Igual JM, Trotta J-RP, et al. Population genetics and phylogeographic history of the insular lizard (Gunther, 1874) from the Balearic Islands based on genome-wide polymorphic data. *Ecol Evol*. 2024;14:e11407.
35. Baldo L, Riera JL, Mitsi K, Pretus JL. Processes shaping gut microbiota diversity in allopatric populations of the endemic lizard *Podarcis lilfordi* from Menorcan islets (Balearic Islands). *FEMS microbiology ecology*. 2018;94.
36. Baldo L, Tavecchia G, Rotger A, Igual JM, Riera JL. Insular holobionts: persistence and seasonal plasticity of the Balearic wall lizard (*Podarcis lilfordi*) gut microbiota. *PeerJ*. 2023;11.
37. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of molecular biology*. 1990;215.
38. Cheng Y, Prickett MD, Gutowska W, Kuo R, Belov K, Burt DW. Evolution of the avian β -defensin and cathelicidin genes. *BMC Evol Biol*. 2015;15:188.

39. van Hoek ML. Antimicrobial peptides in reptiles. *Pharmaceuticals (Basel)*. 2014;7:723–53.
40. Zhang L, Chen D, Yu L, Wei Y, Li J, Zhou C. Genome-wide analysis of the ovo-defensin gene family: Monophyletic origin, independent gene duplication and presence of different selection patterns. *Infect Genet Evol*. 2019;68:265–72.
41. Slater GSC, Birney E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*. 2005;6:1–11.
42. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*. 2015;10.
43. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*. 2004;32.
44. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics (Oxford, England)*. 2009;25.
45. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;;406–25.
46. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular biology and evolution*. 2007;24.
47. Nava GM, Escorcía M, Castañeda MP. Molecular diversity of the antimicrobial domain of beta-defensin 3 and homologous peptides. *Comp Funct Genomics*. 2009;2009:983636.
48. Guyot N, Landon C, Monget P. The Two Domains of the Avian Double- β -Defensin AvBD11 Have Different Ancestors, Common with Potential Monodomain Crocodile and Turtle Defensins. *Biology*. 2022;11.
49. Tennesen JA. Molecular evolution of animal antimicrobial peptides: widespread moderate positive selection. *J of Evolutionary Biology*. 2005;18:1387–94.
50. Yang W, Feiner N, Salvi D, Laakkonen H, Jablonski D, Pinho C, et al. Population Genomics of Wall Lizards Reflects the Dynamic History of the Mediterranean Basin. *Mol Biol Evol*. 2022;39.
51. Unckless RL, Lazzaro BP. The potential for adaptive maintenance of diversity in insect antimicrobial peptides. *Philos Trans R Soc Lond B Biol Sci*. 2016;371:20150291.
52. König E, Bininda-Emonds ORP. Evidence for convergent evolution in the antimicrobial peptide system in anuran amphibians. *Peptides*. 2011;32:20–5.

53. Lai P-K, Tresnak DT, Hackel BJ. Identification and elucidation of proline-rich antimicrobial peptides with enhanced potency and delivery. *Biotechnol Bioeng*. 2019;116:2439–50.
54. Nguyen LT, Haney EF, Vogel HJ. The expanding scope of antimicrobial peptide structures and their modes of action. *Trends Biotechnol*. 2011;29:464–72.
55. Mishra AK, Choi J, Moon E, Baek K-H. Tryptophan-Rich and Proline-Rich Antimicrobial Peptides. *Molecules*. 2018;23.
56. Hernandez-Gordillo V, Geisler I, Chmielewski J. Dimeric unnatural polyproline-rich peptides with enhanced antibacterial activity. *Bioorganic & medicinal chemistry letters*. 2013;24:556.
57. De Los Santos L, Beckman RL, DeBarro C, Keener JE, Torres MDT, de la Fuente-Nunez C, et al. Polyproline peptide targets *Klebsiella pneumoniae* polysaccharides to collapse biofilms. *Cell reports Physical science*. 2024;5.
58. Fleeman RM, Davies BW. Polyproline Peptide Aggregation with *Klebsiella pneumoniae* Extracellular Polysaccharides Exposes Biofilm Associated Bacteria. *Microbiology spectrum*. 2022;10.
59. Gopal R, Seo CH, Song PI, Park Y. Effect of repetitive lysine-tryptophan motifs on the bactericidal activity of antimicrobial peptides. *Amino Acids*. 2013;44:645–60.
60. Korn C, Scholz SR, Gimadutdinow O, Pingoud A, Meiss G. Involvement of conserved histidine, lysine and tyrosine residues in the mechanism of DNA cleavage by the caspase-3 activated DNase CAD. *Nucleic acids research*. 2002;30.
61. Alford MA, Baquir B, Santana FL, Haney EF, Hancock REW. Cathelicidin Host Defense Peptides and Inflammatory Signaling: Striking a Balance. *Frontiers in microbiology*. 2020;11.
62. Tossi A, Gerdol M, Caporale A, Pacor S, Mardirossian M, Scocchi M, et al. Cathelicidins—a rich seam of antimicrobial peptides waiting for exploitation. *Front Drug Discov (Lausanne)*. 2024;4.
63. Izadpanah A, Gallo RL. Antimicrobial peptides. *J Am Acad Dermatol*. 2005;52 3 Pt 1:381–90; quiz 391–2.
64. Gennaro R, Zanetti M. Structural features and biological activities of the cathelicidin-derived antimicrobial peptides. *Biopolymers*. 2000;55.
65. Zhong L, Liu J, Teng S, Xie Z. Identification of a Novel Cathelicidin from the Genome with Antibacterial Activity by Multiple Mechanisms. *Toxins (Basel)*. 2020;12.

66. Luo X, Ouyang J, Wang Y, Zhang M, Fu L, Xiao N, et al. A novel anionic cathelicidin lacking direct antimicrobial activity but with potent anti-inflammatory and wound healing activities from the salamander *Tylostotriton kweichowensis*. *Biochimie*. 2021;191.
67. Alibardi L. Ultrastructural immunolocalization of cathelicidin-like peptides in granulocytes of normal and regenerating lizard tissues. *Acta Histochemica*. 2014;116:363–71.
68. Chen J, Peng J, Ma C, Zhang L, Wu X, Wei H, et al. Co-Expression of Pig IL-2 and Fusion Bovine Cathelicidin Gene by Recombinant Plasmids in Yeast and Their Promotion of Mouse Antibacterial Defense. *Biology (Basel)*. 2022;11.
69. Kuddus MR, Yamano M, Rumi F, Kikukawa T, Demura M, Aizawa T. Enhanced expression of cysteine-rich antimicrobial peptide snakin-1 in *Escherichia coli* using an aggregation-prone protein coexpression system. *Biotechnol Prog*. 2017;33:1520–8.

Supplementary Material

Table S1: Beta-defensin peptides and complete proteins from *P. lilfordi*.

Table S2: Summary of beta-defensins, ovo-defensins and cathelicidin protein homology results by Phyre2.

Table S3: List of AccNos of complete defensin and cathelicidin proteins from published reptile genomes. Newly identified proteins are indicated with 'new'.

Table S4: Ovo-defensin complete proteins from *P. lilfordi*.

Table S5: Cathelicidin proteins from *P. lilfordi*.

Figure S1: Multispecies alignment of beta-defensins proteins (partial) from the Lacertidae species under study (Pm, Pr, Pl and Zv).

Figure S2: Multispecies alignment of ovo-defensins proteins (complete) from the Lacertidae species under study and the Komodo dragon (Pm, Pr, Pl, Zv and Vk).

Figure S3: Multispecies alignment of cathelicidins proteins (partial) from Lacertidae species, the Komodo dragon and the green anole lizard (Pm, Pr, Pl, Zv, Vk and Ac).



GENERAL DISCUSSION

6. GENERAL DISCUSSION

6.1. Introduction to the discussion

This thesis examines how island environments act as natural laboratories for evolutionary processes by exploring the complex interplay between population genomics, historical demography, and antimicrobial resistance in *P. lilfordi*. Unlike approaches that analyze these factors in isolation, a holistic perspective reveals their interdependence and reciprocal influence on the species' evolutionary trajectory. Geographic fragmentation, drift and selective pressures acting on these islands, imposed by specific ecological niches and historical demographic events have synergistically interacted, shaping the species genetic diversity and adaptation within this insular system.

The present discussion is structured around three main axes, which seek to integrate genetic, demographic, and immunological approaches to the evolutionary study of *P. lilfordi*. Each axis addresses key questions aimed at deepening our understanding of how adaptations are shaped in insular contexts, how immune defense mechanisms diversify in response to local selective pressures, and the broader implications of these findings for the conservation of the species:

AXIS 1. Evolution in insular environments

*How does the integration of population genetics and historical demographic patterns provide a deeper understanding of the evolutionary adaptation of *P. lilfordi* within the insular ecosystems of the Balearic Islands, compared to approaches that examine these factors in isolation? What roles do these components play in shaping specific adaptive traits, and how do they interact to contribute to the evolutionary distinctiveness of the species in this insular context?*

AXIS 2. Adaptive immune response in island populations: relationship between genetic diversity and AMP diversification as a defense against local pathogens

*How is the genetic diversity of *P. lilfordi* populations linked to AMP diversification under the ecological and pathogenic pressures of the Balearic Islands, and which factors may have driven the evolution of these peptides as specific defense mechanisms in comparison with other Squamata species in insular contexts?*

AXIS 3. Conservation of *P. lilfordi*: the need for integrative studies to assess the species adaptive potential in the face of future threats
How can population genomics, phylogeography, demographic analyses, and an immunogenomic perspective inform our understanding of the species' adaptive capacity in response to future threats, such as climate change or emerging diseases? Furthermore, how essential is this comparative framework for designing effective and sustainable conservation strategies?

6.2. Evolution in insular environments

The evolution of *P. lilfordi* within the insular ecosystems of the Balearic Islands serves as an exemplary case of how evolutionary mechanisms interact in non-linear ways to produce unique trajectories of local specialization. Beyond the well-documented genetic differentiation among islands, it is essential to focus on the finer-scale processes that explain how this divergence occurs, particularly under extreme isolation, ecological specificity, and severe demographic constraints. In this scenario, it becomes crucial to first obtain a clear understanding of the species phylogeographic history and genetic structure, and to evaluate the potential for adaptation at loci under selection.

A key contribution of the first chapter of this thesis was the genetic structure analysis of *P. lilfordi* (Chapter I), which revealed strong differentiation among island populations ($F_{ST} = 0.247 \pm 0.09$), underscoring the impact of geographic isolation on evolutionary divergence (Otalora et al., 2024; Terrasa et al., 2009). Two clearly differentiated genetic lineages (Menorca and Mallorca/Cabrera) were identified (Otalora et al., 2024) which supports evidence from previous studies (Bassitta et al., 2021; Terrasa et al., 2009), and multiple lines of evidence—including root phylogenetic analyses—point to Mallorca as the species' origin, ruling out Menorca (Otalora et al., 2024). Despite the small size of some populations, genetic diversity remained relatively high ($\pi = 0.144 \pm 0.021$), suggesting that selection, mutation, and historical demography have counteracted expected genetic erosion (Otalora et al., 2024). Geographic fragmentation and islet area (A_r : $p < .001$) have driven marked differentiation among small islands such as Porros and En Curt, where unique mutations and alleles have accumulated under local selective pressures and absence of gene flow (Hurston et al., 2009; Sendell-Price et al., 2021; Warren et al., 2015; Whittaker et al., 2017).

Unlike many insular species that experience reduced genetic diversity from founder effects or drift, *P. lilfordi* retains high heterozygosity ($H_o > 0.09$) (Otalora et al., 2024), exceeding values in other *Podarcis* species, both insular and continental (Sabolić & Štambuk, 2021). This may reflect balanced selection that maintains allelic diversity, along with gene flow between islands potentially facilitated in the past by sea level fluctuations. (Bassitta et al., 2021; Fisher, 1930; Funk et al., 2016; Grant, 1998; Kahilainen et al., 2014; Wright, 1931, 1951; Brown et al., 2008; Terrasa et al., 2009). Island evolution is not solely defined by genetic loss but can also foster adaptive divergence under specific

evolutionary forces—particularly genetic drift and natural selection—where drift accelerates allele fixation in small, isolated populations, enhancing differentiation even without strong selective pressures (Funk et al., 2016; Bassitta et al., 2021). This resilience contrasts with other insular reptiles that, due to founder effects or drift, show reduced variability and greater vulnerability to environmental change and invasive species (Buckland et al., 2014; Templeton et al., 2001). Nonetheless, it highlights the need to protect these genetic and evolutionary units, which are vital for the species' biodiversity and ecological stability.

The analysis of outlier SNPs revealed genomic regions with high differentiation in *P. lilfordi*, concentrated on chromosomes Z, 2, and 15—potential “islands of differentiation” that may reflect the action of selection, mutational hotspots, or other genomic processes (Foll & Gaggiotti, 2008; Gautier et al., 2010; Lavretsky et al., 2019; Luu et al., 2017). Nearly 60% of outliers occurred in coding genes, many linked to protein binding and catalytic functions essential to metabolism and growth, indicating functional adaptation to distinct island environments (Van Nostrand et al., 2020). The resulting genetic configuration, shaped by factors beyond drift or isolation, reflects historical divergence between Menorca and Cabrera/Mallorca, with numerous fixed allelic differences (Otalora et al., 2024), supported by phylogeographic evidence of long-term isolation (Terrasa et al., 2009). Beyond geography, ecological barriers and habitat heterogeneity have contributed to differentiation, with small islands acting as independent evolutionary units (ESUs) shaped by local adaptation (Otalora et al., 2024; Terrasa et al., 2009).

The demographic history of *P. lilfordi* (Chapter II) reflects how successive environmental pressures—glaciations, interglacial fragmentation, and human arrival—have shaped its genetic architecture through the interaction of various evolutionary forces (mutation, genetic drift, natural selection, and gene flow). Three major demographic declines were detected: during the Würm glaciation, the LGM, and following human colonization ~4,000 years ago (Fernandes et al., 2020). These events reduced N_e , decreased habitat connectivity and dispersal, eroded standing variation while enabling drift-mediated allele fixation and selection on locally advantageous variants, and increased extinction risk, especially on smaller islands. The introduction of exotic predators and habitat alteration further exacerbated these trends during the Holocene (Orueta, 2003; Salvador, 2009). The current genetic structuring is consistent with these contractions, long-term isolation and restricted gene flow (Otalora et al., 2024). The disappearance of the species from

Mallorca and Menorca (~2,000 years ago) and its persistence on surrounding islets illustrate vulnerability to combined climatic and anthropogenic forces (Alcover et al., 1999).

Far from simply eroding genetic variation, these demographic contractions contributed to the shaping of the species' adaptive architecture. At regional and local scales, reconstructions link island geography to lineage divergence and evolutionary processes. In the Mallorca–Cabrera system, a contraction ~1.2 Ma aligns with a divergence event (DNAmt), plausibly associated with interglacial sea-level rise and reduced connectivity enhancing drift and limiting gene flow (Brown et al., 2008; Terrasa et al., 2009). Menorca shows an earlier decline (~2.1 Ma), coincident with climatic instability during the Donau glaciation (Lladó & Xamena, 1998). On En Curt, Na Moltona and Na Guardia, Pleistocene events (Riss, Würm) progressively reduced N_e , and a Holocene collapse—coincident with early human presence—signals localized anthropogenic pressure (Fernandes et al., 2020; Rotger et al., 2021). Additional disturbance between the 17th–19th centuries (e.g., limestone quarrying) likely intensified demographic instability (Salvà Matas, 2013), with greater impact in Na Guardia due to a sustained human footprint—harbor proximity, lighthouse, ruderal flora, sporadic rats, and archaeological evidence of long-term settlement (Guerrero, 2000; Rotger et al., 2021). Overall, the demographic record of *P. lilfordi* reflects how historical climatic fluctuations and human disturbance have shaped its genetic structure, promoting population divergence and local adaptation under the influence of evolutionary forces in dynamic island environments.

In small and geographically isolated populations, demographic constraints strongly influence the dynamics of genetic variation. Beneficial mutations can more readily reach fixation, promoting specialized traits, as observed in islets such as En Curt and Porros, where pronounced divergence is reflected in the high number of private alleles—even within areas smaller than one hectare (Otalora et al., 2024). Reduced effective population sizes increase the role of genetic drift, which may lead to the accumulation of deleterious mutations; nevertheless, beneficial variants are more likely to become fixed due to the limited genetic variability (García-Dorado, 2012; Whitlock, 2000). Thus, demographic bottlenecks not only purge deleterious alleles but also facilitate the emergence of adaptive traits. Importantly, historical demography functions not merely as a backdrop but as a selective filter, narrowing the range of viable adaptive solutions and shaping the direction and pace of evolutionary responses (Balick et al., 2015; Grossen et al., 2020).

Understanding such processes requires multidisciplinary perspectives, since genetic differentiation among island populations emerges from the interplay of geographic isolation and evolutionary forces (Cerca et al., 2023; Gillespie, 2016; Stankowski et al., 2024; Templeton, 2006; Wolf et al., 2010).

Although direct evidence of population-level convergence within *P. lilfordi* is currently lacking, it is plausible that similar selective pressures—such as low population density, resource stress, competitive pressures, or microbial exposure—may lead to functionally analogous responses between geographically isolated populations (Barthe et al., 2022). These hypotheses warrant further investigation using population-level functional genomic and transcriptomic analyses. This type of convergence, in the absence of gene flow, illustrates the repeated action of natural selection and genetic drift on conserved functional pathways across ecologically analogous environments. Integrating genomic and demographic data is crucial to understand how functional diversity is generated and maintained. The case of *P. lilfordi* shows how isolation, bottlenecks, and local adaptation interact to drive evolutionary innovation. As an insular model, this species highlights the value of small populations as evolutionary laboratories, sustaining adaptive potential through finely tuned immune responses.

6.3. Adaptive immune response in island populations: relationship between genetic diversity and AMP diversification as a defense against local pathogens

The extraordinary genetic diversity observed in *P. lilfordi* populations offers a compelling framework to investigate how insular environments may shape immune defense mechanisms. Within this context, it is plausible to consider a coevolutionary scenario in which islands, functioning as ecologically closed systems, tend to harbor persistent and specialized pathogen assemblages. The reduced influx of external microbial strains can stabilize these pathogen communities over time, thereby imposing strong, continuous, and host-specific selective pressures that may drive accelerated diversification of AMPs in their immune systems (Barrat-Charlaix & Neher, 2024; Lazzaro et al., 2020; Rahnamaeian, 2011). These host-pathogen interactions could be central not only to immunogenomic adaptation but also to the generation and maintenance of intraspecific diversity within isolated populations (Djahoui et al., 2025; Hulse et al., 2023). This possible pattern is supported by the remarkable expansion of genes encoding beta-defensins (n=65), ovo-defensins (n=8)—including three PrAMPs—and cathelicidins (n=4) in *P. lilfordi* (Otalora et al., 2025), these exhibit typical vertebrate AMP features, including multi-exon structures, conserved cysteine motifs, and many of them the product of recent tandem duplications, organized in conserved chromosomal blocks and flanked by highly stable marker proteins (Otalora et al., 2025; Semple & Dorin, 2012).

Based on the available evidence, the diversification of AMPs in *P. lilfordi* might not merely be a byproduct of relaxed selection, but rather the outcome of potential selective processes driven by a mosaic of pathogen-specific (e.g. *Pantoea* spp., *Citrobacter* spp., *Morganella* spp., *Pseudomonas*, *Enterobacter* spp., *Staphylococcus* spp. and *Escherichia coli*; and parasites, e.g. *Haemogregarines*, *Ophionyssus natricis*, coccidia, *Dicrocoelidae*) pressures on each islet (Armstrong ,2018; Gilroy et al., 2016; Dipineto et al., 2018; Garrido & Mellado et al., 2013). While only a single genome is currently available for *P. lilfordi*, it is conceivable that distinct populations might exhibit differentiated AMP profiles as a consequence of variability in the composition, abundance, and virulence of local microbial communities (Macke et al., 2017). In the RNA data obtained by pooling the five individuals sampled in Aire for the genome assembly, we observed that expression among AMPs was heterogeneous, with some peptides being expressed and

others not. Nevertheless, because data were pooled, we do not have expression patterns at the individual or population level. This could be consistent with the diversity in differential expression patterns of certain AMPs in specific tissues and the presence of peptide variants with structural and functional divergence, such as PrAMPs, like PIOVOD7 and PIOVOD8, which show specialization in disrupting bacterial biofilms (Otalora et al., 2025; Lai et al., 2019). Currently, since we lack population-level data, we cannot clearly understand the diversity and functional profiles of the AMPs present in each islet. This comparative approach will allow us to assess whether populations have diverged immunologically in response to specific pathogenic pressures and will validate the existence of patterns of adaptive differentiation in the innate immune system of *P. lilfordi*.

Regarding the evolutionary forces underlying this pattern, it is crucial to consider how insular conditions shape the functioning of the immune system. In ecosystems characterized by low diversity of exogenous pathogens, minimal predation pressure, and relative ecological stability (Losos & Ricklefs, 2009), selective pressures on generalist immune responses may be relaxed, and demographic effects and evolutionary forces—such as genetic drift—tend to amplify the erosion of immune effectiveness observed in many insular taxa (Barthe et al., 2022). Additionally, it has been proposed that the decrease in exposure to diverse pathogens constitutes a hallmark of the so-called island syndrome, directly linked to the low parasitic pressure characteristic of these environments (Lobato et al., 2017; Matson & Beadell, 2010; Wikelski et al., 2004).

In this context, the immune responses of island vertebrates are often conditioned by the characteristics of the system itself and its environmental limitations, which favor rapid and specific defense mechanisms over more complex and multiphase immune responses (van Hoek, 2014). Innate immunity, which serves as the first line of defense, allows a rapid and broad response and includes phagocytes, macrophages, and dendritic cells, responsible for the nonspecific recognition of pathogens (Akira & Takeda, 2004; Alberts et al., 2002). In this system, Toll-like receptors (TLRs)—essential transmembrane proteins—activate signaling cascades that culminate in the production of various effector molecules, including beta-defensins (BDs) (Velová et al., 2018). Due to their central and highly conserved function, both TLRs and BDs evolve under purifying selection, a regime that eliminates harmful mutations to preserve the structural and functional integrity of these molecules (Barthe et al., 2022; Grueber et al., 2014; van Dijk et al., 2008). However, in island populations with reduced effective sizes, an increase in nonsynonymous variation

in these genes has been documented (Barthe et al., 2022; Gonzalez-Quevedo et al., 2014, 2015). While these gene-level patterns highlight the dynamics of innate immune genes within insular populations, their broader evolutionary meaning becomes clearer when integrated into a macroevolutionary perspective.

From this broader perspective, distance-based reconstructions of the main AMP families in Squamata of this thesis, indicate distinct evolutionary trajectories: β -defensins form well-supported multispecies clades, some of which align with established taxonomic relationships; ovo-defensins display highly conserved cysteine motifs and one-to-one orthology across Lacertidae, with limited evidence of recent duplication; and cathelicidins, despite overall low sequence similarity among variants, show no signs of recent intraspecific duplication. Nevertheless, despite their effectiveness, these innate mechanisms are not exempt from stochastic exposures to emerging pathogens, introduced, for example, by migratory birds or vectors associated with human activity. This suggests that latent genetic diversity could act as a functional reservoir, activated in the face of new infectious threats. From a macroevolutionary perspective, this evolutionary tension between functional constraint and adaptability is exemplified by the presence of structurally conserved AMPs in phylogenetically distant reptilian lineages, including species from the families Lacertidae, Varanidae, and Dactyloidae (*P. lilfordi*, *P. muralis*, *P. raffonei*, *Zootoca vivipara*, *Anolis carolinensis*, *Varanus komodoensis*). This molecular homoplasy provides potential evidence of convergent evolution, likely driven by similar selective pressures in ecologically analogous habitats that demand immunological solutions functionally equivalent, despite phylogenetic divergence (Otalora et al., 2025).

Under this framework, the evolutionary trajectories of AMPs in *P. lilfordi* align with foundational concepts of island evolutionary theory, including founder effects, genetic drift, intensified local selection, the island syndrome, and phylogenetic constraint (Barthe et al., 2022; Otalora et al., 2025). Beyond their canonical role in host defense, AMPs have also been implicated in regulating host–microbiota interactions (Cardoso et al., 2022). Increasing evidence suggests that these molecules contribute to the maintenance of microbial homeostasis by modulating both pathogen clearance and symbiont tolerance (Muniz et al., 2012). This functional versatility may be particularly relevant in insular systems, where microbial communities and host immune responses are shaped by

prolonged coevolution under spatial and ecological constraints (Otalora et al., 2025; Baldo et al., 2023).

The integration of these findings highlights the utility of *P. lilfordi* as a model system for investigating the molecular architecture of immune diversification in isolated vertebrate lineages. The comprehensive annotation of its AMP repertoire provides a robust framework for evaluating immunogenetic resilience in the face of pathogen emergence, anthropogenic disturbance, and climate-driven environmental change. Future research incorporating environmental transcriptomics, comparative immunogenomics, and ecological modeling will be instrumental in assessing the adaptive potential of these relict island populations and elucidating broader principles governing immune system evolution under ecological and spatial isolation.

6.4. Conservation of *P. lilfordi* : the need for integrative studies to assess the species adaptive potential in the face of future threats

Developing effective and sustainable conservation strategies for vulnerable insular species like *P. lilfordi* has become imperative within a global context characterized by accelerated climate change and the emergence of novel pathogens. Projections indicate that the sustained increase in average temperatures across the Mediterranean basin will exacerbate thermal stress on isolated populations, adversely affecting their physiology, immune competence, and long-term viability (IPCC, 2021). Such environmental perturbations are also predicted to modulate the composition and functional stability of host-associated microbiota potentially enhancing the prevalence and transmission of infectious diseases, thus compounding the extinction risk for species with restricted ranges and low interpopulation connectivity (Alemany et al., 2022; Baldo et al., 2023; Uwishema et al., 2023).

In this context, integrating genomic, phylogeographic, demographic, and eco-immunological approaches becomes essential to anticipate adaptive responses and design evidence-based conservation plans (Sutherland et al., 2004). A recent genomic study has provided high-quality chromosome-level assembly for *P. lilfordi*, facilitating the assessment of genetic diversity and adaptive potential across its fragmented populations (Gomez-Garrido et al., 2023). Phylogeographic analyses have revealed significant genetic differentiation among isolated populations, underscoring the importance of preserving intraspecific diversity (Otalora et al., 2024; Terrasa et al., 2009; Bassita et al., 2021). Moreover, investigations into gut microbiota have demonstrated its role in host ecology and evolution, highlighting the need to consider microbial communities in conservation strategies (Alemany et al., 2022; Baldo et al., 2023).

P. lilfordi, is highly susceptible to extinction (Pérez-Mellado & Martínez-Solano., 2009b). The species' small and isolated populations face an elevated risk of genetic diversity loss, leading to inbreeding and the fixation of deleterious alleles. Over the long term, these processes can severely compromise the species' adaptive capacity in the face of environmental change, potentially initiating a trajectory toward the "extinction vortex" (Blomqvist et al., 2010; Fagan & Holmes, 2006; Gilpin & Soule, 1986; Godwin et

al., 2020). Consequently, identifying appropriate management units—such as (Evolutionarily Significant Unit) ESUs—is essential to ensure a conservation framework that maintains both the species' viability and its evolutionary potential.

ESUs serve as valuable tools to ensure the conservation of species and populations that possess unique genetic and ecological characteristics (Funk et al., 2012; Hoelzel, 2023; Reilly et al., 2012). Approaches used to identify ESUs are often based on reciprocal monophyly and significant divergence of allele frequencies (Fraser & Bernatchez, 2001; Moritz, 1994). These analytical methods allow us to evaluate genetic differences between populations and make species conservation decisions (Rodríguez et al., 2013). In the case of *P. lilfordi*, population genomic analyses presented in the first chapter of this thesis revealed an exceptionally high pattern of genetic differentiation; indicative of an extremely fragmented population structure. Some populations, such as En Curt and Foradada, exhibit divergence levels exceeding 0.4, along with the presence of thousands of private alleles (Otalora et al., 2024).

Demographic reconstructions based on SFS inferences indicate that *P. lilfordi* populations have undergone recurrent bottlenecks associated with paleoclimatic events and more recent anthropogenic pressures (Otalora et al., 2024). Despite these fluctuations, several extant populations maintain relatively high nucleotide diversity and low inbreeding coefficients, suggesting the influence of historical demographic rescue or sporadic inter-island dispersal events, such as rafting. This genetic resilience is further supported by a correlational analysis, which reveals a strong and statistically significant positive relationship between allelic richness (A_r) and island area (log-transformed; $p < .001$), consistent with predictions from island biogeography theory and the equilibrium model (Costanzi & Steifetten, 2019; MacArthur & Wilson, 1967). These findings highlight the need to consider each islet (both large and small) as an ESU in conservation planning, lending support to proposals such as that of Pérez-Cembranos et al. (2021). This breaks away from the traditional view that small populations are necessarily genetically depauperate, instead promoting a more dynamic framework that also values the potential adaptive capacity and evolutionary resilience of these units.

Within this framework, a robust evaluation of resilience demands the integration of population genomics, historical demography, local ecology, and the functional characterization of immune responses. The capacity of certain populations to retain

functional immune diversity despite historical N_e collapses highlights a latent form of resilience that may prove critical in the face of future stressors such as emerging infectious diseases or novel climatic regimes. This integrative perspective—one that prioritizes evolutionary functionality within its historical context—must become the cornerstone of any effective conservation strategy for *P. lilfordi*.

From a conservation standpoint, this reality necessitates a paradigm shift. *P. lilfordi* should not be managed simply as a fragmented species, but rather as a collection of distinct evolutionary and functional units (Bassitta et al., 2021; Pérez-Cembranos et al., 2021; Terrasa et al., 2009). Immunological profiles, genomic configurations, and demographic histories differ across populations—each representing a unique reservoir of adaptive potential (Otalora et al., 2024). Neglecting this specificity risks eroding critical components of the species' long-term adaptive potential and compromising its evolutionary trajectory.

Ultimately, the case of *P. lilfordi* reinforces an insular conservation model in which evolutionary history, demographic dynamics, genomic architecture, and immunogenomic diversity are viewed as interdependent dimensions of population viability. Effective conservation must move beyond the protection of organisms or habitats to encompass the preservation of populations as independent evolutionary genetic units. Each ESU embodies an irreplaceable expression of the species' adaptive potential. In the face of accelerating environmental change and the rise of emerging biological threats, conservation strategies must anticipate threats—rather than merely react to loss.

6.5. Limitations and future directions

While this thesis has significantly advanced the study of the Balearic lizard, additional research directions could complement these findings and enrich our understanding of the evolutionary and adaptive processes that have shaped this insular species.

6.5.1. *Genomic resolution: moving towards high-density genomics*

SNP-based approaches, particularly those utilizing GBS and RAD-seq technologies, have provided robust insights into population-level patterns. Nevertheless, the inherently partial genomic coverage of these methods constrains the resolution of fine-scale genetic structure, local adaptation, and evolutionary dynamics. Overcoming these limitations will require comprehensive whole-genome resequencing (WGR) of multiple individuals across populations, enabling the detection of a broader spectrum of genomic variation, including structural variants and regulatory elements often overlooked by reduced-representation methods.

6.5.2. *Identification of adaptive loci: integrating GWAS and outlier analyses*

The systematic application of genome-wide association studies (GWAS) and outlier detection analyses remains largely unexplored in *P. lilfordi*. Implementing these approaches is critical for identifying loci under positive selection, particularly those involved in responses to extreme environmental pressures and emerging pathogens. To maximize the power of GWAS in this context, it will be essential to prioritize the collection of high-quality phenotypic data that can be directly linked to genomic variation. Given the species' small effective population sizes and the profound effects of isolation, it is essential to employ multivariate analytical frameworks capable of disentangling authentic signals of selection from stochastic patterns generated by intense genetic drift.

6.5.3. *Demographic inference: challenges in sampling and model complexity*

Recent reconstructions using Stairway Plot analyses have revealed major demographic events associated with climatic oscillations and human-mediated impacts. Nevertheless, the resolution of historical demographic trajectories remains limited by the relatively small number of loci and individuals sampled per population. Achieving a finer resolution will necessitate denser genomic datasets and increased sample size combined with more

sophisticated coalescent-based modeling approaches, incorporating ancestral migration, secondary colonization events, and population size fluctuations at microgeographic scales.

6.5.4. *Antimicrobial peptides (AMPs) characterization: intraspecific diversity underexplored*

The pioneering identification of AMPs in *P. lilfordi* of this thesis uncovered a diversified set of β -defensins, ovo-defensins, and cathelicidins. However, this effort was largely based on a single reference genome, limiting the exploration of intraspecific variability within these critical immune effectors. Considering the evidence for recent gene duplications, balancing selection, and convergent evolution, expanding functional and population-level investigations across different islands is imperative to elucidate the adaptive significance of AMP polymorphisms.

6.5.5. *Functional data integration: towards an integral evolutionary genomics framework*

Although the AMP repertoire has been structurally characterized, a functional validation is required. Addressing this gap demands the development of antimicrobial assays targeting pathogenic species relevant to insular ecosystems, (e.g. *Pantoea* spp., *Citrobacter* spp., *Morganella* spp., *Pseudomonas*, *Enterobacter* spp., *Staphylococcus* spp. and *Escherichia coli*; and parasites, e.g. *Haemogregarines*, *Ophionyssus natricis*, *coccidia*, *Dicrocoelidae*) coupled with expression studies under controlled environmental conditions. Furthermore, a multi-omics approach is essential to unravel the complex regulatory and evolutionary mechanisms underlying AMP diversification. In particular, epigenetic profiling, particularly DNA methylation analyses, together with microbiome characterization, will offer unprecedented insights into host-microbiota interactions and their role in shaping immune resilience under biotic and abiotic stressors.



CONCLUSIONS

7. CONCLUSIONS

This thesis provides an integrated perspective on the phylogeography, historical demography, and immune genomic diversity of *P. lilfordi*, offering crucial insights into the evolutionary processes that shape insular populations and informing future conservation strategies. The key findings from each chapter are summarised as follows:

1. Genome-wide analyses revealed two major phylogeographic lineages in *P. lilfordi*, separating Mallorca–Cabrera from Menorca populations, and identified Mallorca as the ancestral origin of the species, thereby challenging previous colonization hypotheses. The pronounced genetic differentiation among all islet populations indicates long-term isolation and underscores the distinctiveness of each, justifying their designation as separate Evolutionarily Significant Units (ESUs), reinforcing the need for fine-scale conservation strategies tailored to individual islet populations.
2. Despite their small sizes and fragmentation, most populations retained considerable genetic diversity and low inbreeding levels, suggesting the action of buffering forces such as balancing selection. This hypothesis warrants further investigation through integrative approaches combining population genomics, functional assays, and environmental association analyses. Comprehensive whole-genome resequencing across multiple populations will be essential to capture the full spectrum of genomic variation, including structural variants (SVs) often overlooked by reduced-representation methods.
3. Outlier analyses identified highly differentiated loci between major lineages, particularly on the Z chromosome, supporting hypotheses of "genomic islands of divergence". Candidate loci were enriched in binding and catalytic functions, suggesting roles in local adaptation during insular diversification. Functional studies are required to validate these findings and to explore their ecological significance, and the systematic application of GWAS and outlier detection analyses to identify loci under positive selection while distinguishing adaptive signals.

4. Demographic analyses demonstrate that *P. lilfordi* has undergone recurrent bottlenecks since the Quaternary, driven by glaciations, sea-level rise, and anthropogenic pressures. Potential rafting dispersal events, together with the marked demographic variation observed in microinsular populations (e.g., Na Moltana, Na Guardia, En Curt), underscore the complex interplay between isolation, and occasional connectivity. These findings provide genomic evidence linking past climatic events with recent human impacts. Future research should integrate genomic and demographic approaches to evaluate connectivity and resilience under environmental change, and apply species distribution models (SDMs) to anticipate range shifts and guide conservation strategies for *P. lilfordi*.
5. Genomic analyses identified a diverse AMP repertoire (65-beta-defensins, 8 ovo-defensins, 4 cathelicidins) with distinct architectures and evolutionary dynamics. Tandem duplications in beta/ovo-defensins contrast with conserved cathelicidins, underscoring heterogeneous selective trajectories. Comparative analyses across squamates revealed conserved AMP motifs and novel OVOD PrAMPs in *P. lilfordi*, suggesting convergent evolution and functional innovation under pathogen pressures. Functional assays across different islets will be critical to establish the link between AMP genomic diversity and effective immune performance, clarifying how such diversity contributes to pathogen defense in fragmented insular populations.
6. The integration of genomic variation and immune resilience is fundamental to anticipate species responses to climatic and ecological change. Future research should integrate genomic surveys, demographic reconstructions, and experimental assays to map both variation and adaptive responses to environmental change. This knowledge will be essential for identifying critical evolutionary units and guiding conservation strategies that ensure the long-term viability of *P. lilfordi* and other insular systems.

8. REFERENCES

- Ahrens, C. W., Rymer, P. D., Stow, A., Bragg, J., Dillon, S., Umbers, K. D. L., & Dudaniec, R. Y. (2018). The search for loci under selection: Trends, biases and progress. *Molecular Ecology*, 27(6), 1342–1356. <https://doi.org/10.1111/mec.14549>
- Aird, S. D., Watanabe, Y., Villar-Briones, A., Roy, M. C., Terada, K., & Mikheyev, A. S. (2013). Quantitative high-throughput profiling of snake venom gland transcriptomes and proteomes (*Ovophis okinavensis* and *Protobothrops flavoviridis*). *BMC Genomics*, 14(1), 1–27. <https://doi.org/10.1186/1471-2164-14-790>
- Akira, S., & Takeda, K. (2004). Toll-like receptor signalling. *Nature Reviews Immunology*, 4(7), 499–511. <https://doi.org/10.3389/fimmu.2014.00461>
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Innate immunity. En *Molecular biology of the cell* (4.^a ed.). Garland Science. <https://www.ncbi.nlm.nih.gov/books/NBK21054/>
- Alcover, J. A. (2000). Vertebrate evolution and extinction on western and central Mediterranean islands. *Tropics*, 10(1), 103–123. <https://doi.org/10.3759/tropics.10.103>
- Alcover, J. A., & Dover, P. (2000). Extreme insular evolution in *Myotragus balearicus* Bate 1909 (Artiodactyla, Caprinae). *Tropics*, 10(1), 189–201. <https://doi.org/10.3759/tropicos.10.189>
- Alcover, J. A., Seguí, B., & Bover, P. (1999). Extinctions and local disappearances of vertebrates in the Western Mediterranean Islands. En *Extinctions in Near Time* (pp. 165–188). https://doi.org/10.1007/978-1-4757-5202-1_8
- Aleman, I., Pérez-Cembranos, A., Pérez-Mellado, V., Castro, J. A., Picornell, A., Ramon, C., & Jurado-Rivera, J. A. (2022). Faecal microbiota divergence in allopatric populations of *Podarcis lilfordi* and *P. pityusensis*, two lizard species endemic to the Balearic Islands. *Microbial Ecology*, 85(4), 1564–1577. <https://doi.org/10.1007/s00248-022-02019-3>
- Alibardi, L. (2013). Ultrastructural immunolocalization of beta-defensin-27 in granulocytes of the dermis and wound epidermis of lizard suggests they contribute to the anti-microbial skin barrier. *Anatomy & Cell Biology*, 46(4), 246. <https://doi.org/10.5115/acb.2013.46.4.246>

Álvarez-Varas, R., Rojas-Hernández, N., Heidemeyer, M., Riginos, C., Benítez, H. A., Araya-Donoso, R., Reséndiz, E., Lara-Uc, M., Godoy, D. A., Muñoz-Pérez, J. P., Alarcón-Ruales, D. E., Alfaro-Shigueto, J., Ortiz-Alvarez, C., Mangel, J. C., Vianna, J. A., & Véliz, D. (2021). Green, yellow or black? Genetic differentiation and adaptation signatures in a highly migratory marine turtle. *Proceedings of the Royal Society B: Biological Sciences*. <https://doi.org/10.1098/rspb.2021.0754>

Armstrong, C. (2018). *The role of pathogens in shaping genetic variation within and among populations in an island bird* [Doctoral dissertation, University of East Anglia]. <https://ueaeprints.uea.ac.uk/id/eprint/69970/1/2018ArmstrongCPhD.pdf>

Aurelle, D., Thomas, S., Albert, C., Bally, M., Bondeau, A., Boudouresque, C.-F., ... Fady, B. (2022). Biodiversity, climate change, and adaptation in the Mediterranean. *Ecosphere*, 13(4), e3915. <https://doi.org/10.1002/ecs2.3915>

Baeckens, S., & Van Damme, R. (2020). The island syndrome. *Current Biology*, 30(8), R338–R339. <https://doi.org/10.1016/j.cub.2020.03.029>

Banks, S. C., Cary, G. J., Smith, A. L., Davies, I. D., Driscoll, D. A., Gill, A. M., ... Peakall, R. (2013). How does ecological disturbance influence genetic diversity? *Trends in Ecology & Evolution*, 28(11), 670–679. <https://doi.org/10.1016/j.tree.2013.08.005>

Baldo, L., Tavecchia, G., Rotger, A., Igual, J. M., & Riera, J. L. (2023). Insular holobionts: Persistence and seasonal plasticity of the Balearic wall lizard (*Podarcis lilfordi*) gut microbiota. *PeerJ*, 11, e14511. <https://doi.org/10.7717/peerj.14511>

Balick, D. J., Do, R., Cassa, C. A., Reich, D., & Sunyaev, S. R. (2015). Dominance of deleterious alleles controls the response to a population bottleneck. *PLOS Genetics*, 11(8), e1005436. <https://doi.org/10.1371/journal.pgen.1005436>

Bals, R., & Wilson, J. M. (2003). Cathelicidins: A family of multifunctional antimicrobial peptides. *Cellular and Molecular Life Sciences*, 60(4), 711–720. <https://doi.org/10.1007/s00018-003-2186-9>

Barberá, C., Moranta, J., Ordines, F., Ramón, M., de Mesa, A., Díaz-Valdés, M., Grau, A. M., & Massutí, E. (2012). Biodiversity and habitat mapping of Menorca Channel (western Mediterranean): Implications for conservation. *Biodiversity and Conservation*, 21(3), 701–728. <https://doi.org/10.1007/s10531-011-0210-1>

Barbosa, A., Sillero, N., Martínez-Freiría, F., & Real, R. (2012). Ecological niche models in Mediterranean herpetology: Past, present, and future. *Unpublished manuscript*. <https://doi.org/10.13140/2.1.3746.6560>

Barrat-Charlaix, P., & Neher, R. A. (2024). Eco-evolutionary dynamics of adapting pathogens and host immunity. *eLife*, 13, e97350. <https://doi.org/10.7554/eLife.97350>

Barthe, M., Doutrelant, C., Covas, R., Melo, M., Illera, J. C., Tilak, M.-K., ... Nabholz, B. (2022). Evolution of immune genes in island birds: Reduction in population sizes can explain island syndrome. *Peer Community Journal*, 2, e186. <https://doi.org/10.24072/pcjournal.186>

Bassitta, M., Brown, R. P., Pérez-Cembranos, A., Pérez-Mellado, V., Castro, J. A., Picornell, A., & Ramon, C. (2021). Genomic signatures of drift and selection driven by predation and human pressure in an insular lizard. *Scientific Reports*, 11(1), 1–13. <https://doi.org/10.1038/s41598-021-85591-x>

Bhattacharjya, S., Zhang, Z., & Ramamoorthy, A. (2024). LL-37: Structures, antimicrobial activity, and influence on amyloid-related diseases. *Biomolecules*, 14(3), 320. <https://doi.org/10.3390/biom14030320>

Bishop, B. M., Juba, M. L., Devine, M. C., Barksdale, S. M., Rodriguez, C. A., Chung, M. C., Russo, P. S., Vliet, K. A., Schnur, J. M., & Van Hoek, M. L. (2015). Bioprospecting the American alligator (*Alligator mississippiensis*) host defense peptidome. *PLOS ONE*, 10, e0132819–e0132819. <https://doi.org/10.1371/journal.pone.0117394>

Blomqvist, D., Pauliny, A., Larsson, M., & Flodin, L.-Å. (2010). Trapped in the extinction vortex? Strong genetic effects in a declining vertebrate population. *BMC Evolutionary Biology*, 10(1), 1–9. <https://doi.org/10.1186/1471-2148-10-33>

Boscá, E. (1883). Exploración herpetológica de la isla de Ibiza. *Anales de la Sociedad Española de Historia Natural*, 12, 241–250.

Böhme, W., Hutterer, R., & Bings, W. (1985). *Die Stimme der Lacertidae, speziell der Kanareneidechsen* (Reptilia: Sauria). https://www.researchgate.net/publication/284297455_Die_Stimme_der_Lacertidae_speziell_der_Kanareneidechsen_Reptilia_Sauria

Brown, R. P., Sun, H., Jin, Y., & Meloro, C. (2023). Habitat-associated genomic variation in a wall lizard from an oceanic island. *Genome Biology and Evolution*, 15(11), evad193. <https://doi.org/10.1093/gbe/evad193>

Brown, R. P., Terrasa, B., Pérez-Mellado, V., Castro, J. A., Hoskisson, P. A., Picornell, A., & Ramon, M. M. (2008). Bayesian estimation of post-Messinian divergence times in Balearic Island lizards. *Molecular Phylogenetics and Evolution*, 48(1), 350–358. <https://doi.org/10.1016/j.ympev.2008.04.013>

Buckland, S., Cole, N. C., Groombridge, J. J., Küpper, C., Burke, T., Dawson, D. A., Gallagher, L. E., & Harris, S. (2014). High risks of losing genetic diversity in an endemic Mauritian gecko: Implications for conservation. *PLOS ONE*, 9(6), e93387. <https://doi.org/10.1371/journal.pone.0093387>

Burjachs, F., Pérez-Obiol, R., Picornell-Gelabert, L., Revelles, J., Servera-Vives, G., Expósito, I., & Yll, E. I. (2017). Overview of environmental changes and human colonization in the Balearic Islands (Western Mediterranean) and their impacts on vegetation composition during the Holocene. *Journal of Archaeological Science: Reports*, 12, 845–859. <https://doi.org/10.1016/j.jasrep.2016.09.018>

Cardoso, M. H., Meneguetti, B. T., Oliveira-Júnior, N. G., Macedo, M. L. R., & Franco, O. L. (2022). Antimicrobial peptide production in response to gut microbiota imbalance. *Peptides*, 157, 170865. <https://doi.org/10.1016/j.peptides.2022.170865>

Cerca, J., Cotoras, D. D., Bieker, V. C., De-Kayne, R., Vargas, P., Fernández-Mazuecos, M., ... Martin, M. D. (2023). Evolutionary genomics of oceanic island radiations. *Trends in Ecology & Evolution*, 38(7), 631–642. <https://doi.org/10.1016/j.tree.2023.02.003>

Clauzon, G., Suc, J.-P., Gautier, F., Berger, A., & Loutre, M.-F. (1996). Alternate interpretation of the MSC: Controversy resolved? *Geology*. [https://doi.org/10.1130/0091-7613\(1996\)024<0363:AIOTMS>2.3.CO;2](https://doi.org/10.1130/0091-7613(1996)024<0363:AIOTMS>2.3.CO;2)

Clegg, S. M., Degnan, S. M., Kikkawa, J., Moritz, C., Estoup, A., & Owens, I. P. (2002). Genetic consequences of sequential founder events by an island-colonizing bird. *Proceedings of the National Academy of Sciences*, 99(12), 8127–8132. <https://doi.org/10.1073/pnas.102583399>

Clegg, S. M., Frentiu, F. D., Kikkawa, J., Tavecchia, G., & Owens, I. P. F. (2008). 4000 years of phenotypic change in an island bird: Heterogeneity of selection over three

microevolutionary timescales. *Evolution*, 62(9), 2393–2410.
<https://doi.org/10.1111/j.1558-5646.2008.00437.x>

Clegg, S. M., & Owens, I. P. F. (2002). The “island rule” in birds: Medium body size and its ecological explanation. *Proceedings of the Royal Society B: Biological Sciences*, 269(1498), 1359–1365. <https://doi.org/10.1098/rspb.2002.2024>

Cooper, M. D., & Alder, M. N. (2006). The evolution of adaptive immune systems. *Cell*, 124(4), 815–822. <https://doi.org/10.1016/j.cell.2006.02.001>

Costanzi, J. M., & Steifetten, Ø. (2019). Island biogeography theory explains the genetic diversity of a fragmented rock ptarmigan (*Lagopus muta*) population. *Ecology and Evolution*, 9(7), 3937–3949. <https://doi.org/10.1002/ece3.5007>

Crow, J., & Kimura, M. (1970). *An introduction to population genetics theory*. Harper & Row.

Cuerda, J. (1989). *Geological history of the Balearic Islands*.

Dalla Valle, L., Benato, F., Maistro, S., Quinzani, S., & Alibardi, L. (2012). Bioinformatic and molecular characterization of beta-defensins-like peptides isolated from the green lizard *Anolis carolinensis*. *Developmental and Comparative Immunology*, 36(1), 222–229. <https://doi.org/10.1016/j.dci.2011.05.004>

Dedato, M. N., Robert, C., Taillon, J., Shafer, A. B. A., & Côté, S. D. (2022). Demographic history and conservation genomics of caribou (*Rangifer tarandus*) in Québec. *Evolutionary Applications*, 15(12), 2043–2053. <https://doi.org/10.1111/eva.13495>

De Donato, M., Peters, S. O., Mitchell, S. E., Hussain, T., & Imumorin, I. G. (2013). Genotyping-by-sequencing (GBS): A novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing. *PLOS ONE*, 8(5), e62137. <https://doi.org/10.1371/journal.pone.0062137>

Dipineto, L., Raia, P., Varriale, L., Borrelli, L., Botta, V., Serio, C., ... Rinaldi, L. (2018). Bacteria and parasites in *Podarcis sicula* and *P. sicula klemmerii*. *BMC Veterinary Research*, 14(1), 392. <https://doi.org/10.1186/s12917-018-1708-5>

Djahoui, E., Loeuille, N., & Rohr, R. P. (2025). The role of pathogens in creating and maintaining host polymorphism. *bioRxiv*, 2025.02.13.638051. <https://doi.org/10.1101/2025.02.13.638051>

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLOS ONE*, 6(5), e19379. <https://doi.org/10.1371/journal.pone.0019379>

Excoffier, L., Hofer, T., & Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity*, 103(4), 285–298. <https://doi.org/10.1038/hdy.2009.74>

Fagan, W. F., & Holmes, E. E. (2006). Quantifying the extinction vortex. *Ecology Letters*, 9(1), 51–60. <https://doi.org/10.1111/j.1461-0248.2005.00845.x>

Fernandes, D. M., Mittnik, A., Olalde, I., Lazaridis, I., Cheronet, O., Rohland, N., ... Reich, D. (2020). The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nature Ecology & Evolution*, 4(3), 334–345. <https://doi.org/10.1038/s41559-020-1102-0>

Field, E. K., Hartzheim, A., Terry, J., Dawson, G., Haydt, N., & Neuman-Lee, L. A. (2022). Reptilian innate immunology and ecoimmunology: What do we know and where are we going? *Integrative and Comparative Biology*, 62(6), 1557–1571. <https://doi.org/10.1093/icb/icac116>

Findlay, F., Proudfoot, L., Stevens, C., & Barlow, P. G. (2016). Cationic host defense peptides: Novel antimicrobial therapeutics against Category A pathogens and emerging infections. *Pathogens and Global Health*, 110(4–5), 137–147. <https://doi.org/10.1080/20477724.2016.1195036>

Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., Holderegger, R., & Widmer, A. (2017). Estimating genomic diversity and population differentiation—An empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *BMC Genomics*, 18, 69. <https://doi.org/10.1186/s12864-016-3459-7>

Fisher, R. A. (1930). *The genetical theory of natural selection*. Clarendon Press.

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180(2), 977–993. <https://doi.org/10.1534/genetics.108.092221>

Frankham, R. (1997). Do island populations have less genetic variation than mainland populations? *Heredity*, 78(Pt 3), 311–327. <https://doi.org/10.1038/hdy.1997.46>

French, S. S., DeNardo, D. F., & Moore, M. C. (2007a). Trade-offs between the reproductive and immune systems: Facultative responses to resources or obligate responses to reproduction? *American Naturalist*, 170, 79–89. <https://doi.org/10.1086/518569>

French, S. S., Johnston, G. I. H., & Moore, M. C. (2007b). Immune activity suppresses reproduction in food-limited female tree lizards *Urosaurus ornatus*. *Functional Ecology*, 21, 1115–1122. <https://doi.org/10.1111/j.1365-2435.2007.01311.x>

French, S. S., Fokidis, H. B., & Moore, M. C. (2008). Variation in stress and innate immunity in the tree lizard (*Urosaurus ornatus*) across an urban-rural gradient. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 178, 997–1005. <https://doi.org/10.1007/s00360-008-0290-8>

Fraser, D. J., & Bernatchez, L. (2001). Adaptive evolutionary conservation: Towards a unified concept for defining conservation units. *Molecular Ecology*, 10(12), 2741–2752. <https://pubmed.ncbi.nlm.nih.gov/11903888/>

Funk, W. C., Lovich, R. E., Hohenlohe, P. A., Hofman, C. A., Morrison, S. A., Sillett, T. S., Ghalambor, C. K., Maldonado, J. E., Rick, T. C., Day, M. D., Polato, N. R., Fitzpatrick, S. W., Coonan, T. J., Crooks, K. R., Dillon, A., Garcelon, D. K., King, J. L., Boser, C. L., Gould, N., & Andelt, W. F. (2016). Adaptive divergence despite strong genetic drift: Genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). *Molecular Ecology*, 25(10), 2176–2194. <https://doi.org/10.1111/mec.13605>

Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution*, 27(9), 489–496. <https://doi.org/10.1016/j.tree.2012.05.012>

García-Dorado, A. (2012). Understanding and predicting the fitness decline of shrunk populations: Inbreeding, purging, mutation, and standard selection. *Genetics*, 190(4), 1461–1476. <https://doi.org/10.1534/genetics.111.135541>

Gautier, F., Clauzon, G., Suc, J. P., Cravatte, J., & Violanti, D. (1994). Age et durée de la crise de salinité messinienne. *Manuscrito*, https://www.researchgate.net/publication/279621556_Age_and_duration_of_the_Messinian_salinity_crisisAge_et_duree_de_la_crise_de_salinite_messinienne

Gautier, M., Hocking, T. D., & Foulley, J.-L. (2010). A Bayesian outlier criterion to detect SNPs under selection in large data sets. *PLOS ONE*, 5(8), e11913. <https://doi.org/10.1371/journal.pone.0011913>

Garrido, M., & Pérez-Mellado, V. (2013). Prevalence and intensity of blood parasites in insular lizards. *Zoologischer Anzeiger*, 252(4), 588–592.

Gillespie, R. G. (2016). Island time and the interplay between ecology and evolution in species diversification. *Evolutionary Applications*, 9(1), 53–73.

Gillespie, R. G., & Roderick, G. K. (2002). Arthropods on islands: Colonization, speciation, and conservation. *Annual Review of Entomology*, 47. <https://doi.org/10.1146/annurev.ento.47.091201.145244>

Gilpin, M. E., & Soule, M. E. (1986). Minimum viable populations: Processes of species extinction. *Biological Conservation*, 61(2), 77–79. https://scispace.com/papers/minimum-viable-populations-processes-of-species-extinction-35qd9qebwf?citations_page=71

Gilroy, D., van Oosterhout, C., Komdeur, J., & Richardson, D. S. (2016). Avian β -defensin variation in bottlenecked populations: The Seychelles warbler and other congeners. *Conservation Genetics*, 17(3), 661–674. <https://doi.org/10.1007/s10592-016-0813-x>

Godwin, J. L., Lumley, A. J., Michalczyk, Ł., Martin, O. Y., & Gage, M. J. G. (2020). Mating patterns influence vulnerability to the extinction vortex. *Global Change Biology*, 26(8), 4226–4239. <https://doi.org/10.1111/gcb.15186>

Gomez-Garrido, J., Cruz, F., Alioto, T. S., Feiner, N., Uller, T., Gut, M., Sanchez Escudero, I., Tavecchia, G., Rotger, A., Otalora Acevedo, K. E., & Baldo, L. (2023). Chromosome-level genome assembly of Lilford's wall lizard, *Podarcis lilfordi* (Günther, 1874) from the Balearic Islands (Spain). *DNA Research*, 30(3). <https://doi.org/10.1093/dnares/dsac008>

Gonzalez-Quevedo, C., Phillips, K. P., Spurgin, L. G., & Richardson, D. S. (2014). 454 screening of individual MHC variation in an endemic island passerine. *Immunogenetics*, 67(3), 149–162. <https://doi.org/10.1007/s00251-014-0822-1>

Gonzalez-Quevedo, C., Spurgin, L. G., Illera, J. C., & Richardson, D. S. (2015). Drift, not selection, shapes toll-like receptor variation among oceanic island populations. *Molecular Ecology*, 24(23), 5852–5863. <https://doi.org/10.1111/mec.13437>

Graham, S. P., Earley, R. L., Guyer, C., & Mendonça, M. T. (2011). Innate immune performance and steroid hormone profiles of pregnant versus nonpregnant cottonmouth snakes (*Agkistrodon piscivorus*). *General and Comparative Endocrinology*, 174, 348–353. <https://doi.org/10.1016/j.ygcen.2011.09.015>

Grant, P. R. (1998). *Evolution on islands*. Oxford University Press. <https://global.oup.com/academic/product/evolution-on-islands-9780198501718>

Grant, P. R., & Grant, B. R. (1995). Predicting microevolutionary responses to directional selection on heritable variation. *Evolution*, 49(2), 241–251. <https://doi.org/10.1111/j.1558-5646.1995.tb02236.x>

Grant, P. R., & Grant, B. R. (2002). Unpredictable evolution in a 30-year study of Darwin's finches. *Science*, 296(5568). <https://doi.org/10.1126/science.1070315>

Grant, P. R., & Grant, B. R. (2006). Evolution of character displacement in Darwin's finches. *Science*, 313(5784), 224–226. <https://doi.org/10.1126/science.1128374>

Grant, P. R., & Grant, B. R. (2008). *How and why species multiply: The radiation of Darwin's finches*. Princeton University Press. <https://doi.org/10.1525/auk.2008.41008.3>

Grossen, C., Guillaume, F., Keller, L. F., & Croll, D. (2020). Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. *Nature Communications*, 11(1), 1–12. <https://doi.org/10.1038/s41467-020-14803-1>

Grueber, C. E., Wallis, G. P., & Jamieson, I. G. (2014). Episodic positive selection in the evolution of avian Toll-like receptor innate immunity genes. *PLOS ONE*, 9(3), e89632. <https://doi.org/10.1371/journal.pone.0089632>

Guerrero, V. (2000). *Colonización humana en ambientes insulares: Interacción con el medio y adaptación cultural*. Servei de Publicacions i Intercanvi Científic. <https://dialnet.unirioja.es/servlet/libro?codigo=11204>

Haney, E. F., Straus, S. K., & Hancock, R. E. W. (2019). Reassessing the host defense peptide landscape. *Frontiers in Chemistry*, 7, 435645. <https://doi.org/10.3389/fchem.2019.00043>

Harris, D., & Arnold, E. (1999). Relationships of wall lizards, *Podarcis* (Reptilia: Lacertidae) based on mitochondrial DNA sequences. <https://www.jstor.org/stable/1447608>

Hendry, A. P., Huber, S. K., De León, L. F., Herrel, A., & Podos, J. (2008). Disruptive selection in a bimodal population of Darwin's finches. *Proceedings of the Royal Society B: Biological Sciences*. <https://doi.org/10.1098/rspb.2008.1321>

Hoelzel, A. R. (2023). Where to now with the evolutionarily significant unit? *Trends in Ecology & Evolution*, 38(12), 1134–1142. <https://doi.org/10.1016/j.tree.2023.07.005>

Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2020). Population genomics for wildlife conservation and management. *Molecular Ecology*, 30(1), 62. <https://doi.org/10.1111/mec.15720>

Huan, Y., Kong, Q., Mou, H., & Yi, H. (2020). Antimicrobial peptides: Classification, design, application and research progress in multiple fields. *Frontiers in Microbiology*, 11, 582779. <https://doi.org/10.3389/fmicb.2020.582779>

Hulse, S. V., Antonovics, J., Hood, M. E., & Bruns, E. L. (2023). Host–pathogen coevolution promotes the evolution of general, broad-spectrum resistance and reduces foreign pathogen spillover risk. *Evolution Letters*, 7(6), 467. <https://doi.org/10.1093/evlett/grad051>

Hurston, H., Voith, L., Bonanno, J., Foufopoulos, J., Pafilis, P., Valakos, E., & Anthony, N. (2009). Effects of fragmentation on genetic diversity in island populations of the Aegean wall lizard *Podarcis erhardii* (Lacertidae, Reptilia). *Molecular Phylogenetics and Evolution*, 52(2). <https://doi.org/10.1016/j.ympev.2009.03.028>

IPCC. (2021). *Climate change 2021: The physical science basis*. <https://www.ipcc.ch/report/ar6/wg1/>

Issar, A. S. (2004). *Climate changes during the Holocene and their impact on hydrological systems*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511535703>

Iwasaki, A., & Medzhitov, R. (2015). Control of adaptive immunity by the innate immune system. *Nature Immunology*, 16, 343–353. <https://doi.org/10.1038/ni.3123>

Kahilainen, A., Puurtinen, M., & Kotiaho, J. S. (2014). Conservation implications of species–genetic diversity correlations. *Global Ecology and Conservation*, 2, 315–323. <https://doi.org/10.1016/j.gecco.2014.10.013>

Koehbach, J., & Craik, D. J. (2019). The vast structural diversity of antimicrobial peptides. *Trends in Pharmacological Sciences*, 40(7), 517–528. <https://doi.org/10.1016/j.tips.2019.04.012>

Krijgsman, W., Hilgen, F. J., Raffi, I., Sierro, F. J., & Wilson, D. S. (1999). Chronology, causes and progression of the MSC. *Nature*, 400(6745), 652–655. <https://doi.org/10.1038/23231>

Kumar, H., Kawai, T., & Akira, S. (2011). Pathogen recognition by the innate immune system. *International Reviews of Immunology*, 30, 16–34. <https://doi.org/10.3109/08830185.2010.529976>

Lack, D. (1947). *Darwin's finches*. Cambridge University Press. <https://www.cambridge.org/us/universitypress/subjects/life-sciences/darwin/darwins-finches?format=PB&isbn=9780521272421>

Lai, P.-K., Tresnak, D. T., & Hackel, B. J. (2019). Identification and elucidation of proline-rich antimicrobial peptides with enhanced potency and delivery. *Biotechnology and Bioengineering*, 116(10), 2439–2450. <https://doi.org/10.1002/bit.27092>

Larrucea, J. R., Ballesteros, E., Gràcia, Á. G., Grivé, M. M., & Mellado, V. P. (2016). Tejiendo naturaleza: El archipiélago de Cabrera, lugar de encuentro y armonía entre la gea, la fauna y la flora. En *El Parque Nacional marítimo-terrestre del Archipiélago de Cabrera: Un paisaje entre la tierra y el mar*. <https://www.researchgate.net/publication/311203179>

Lavretsky, P., DaCosta, J. M., Sorenson, M. D., McCracken, K. G., & Peters, J. L. (2019). ddRAD-seq data reveal significant genome-wide population structure and divergent genomic regions that distinguish the mallard and close relatives in North America. *Molecular Ecology*, 28(10), 2594–2609. <https://doi.org/10.1111/mec.15091>

Lazzaro, B. P., Zasloff, M., & Rolff, J. (2020). Antimicrobial peptides: Application informed by evolution. *Science*. <https://doi.org/10.1126/science.aau5480>

Lei, J., Sun, L., Huang, S., Zhu, C., Li, P., He, J., Mackey, V., Coy, D. H., & He, Q. (2019). The antimicrobial peptides and their potential clinical applications. *American Journal of Translational Research*, 11(7). <https://pubmed.ncbi.nlm.nih.gov/31396309/>

Li, H., & Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature*, 475(7357), 493–496. <https://doi.org/10.1038/nature10231>

Li, X., Zuo, S., Wang, B., Zhang, K., & Wang, Y. (2022). Antimicrobial mechanisms and clinical application prospects of antimicrobial peptides. *Molecules*, 27(9), 2675. <https://doi.org/10.3390/molecules27092675>

Liu, X., & Fu, Y.-X. (2020). Stairway Plot 2: Demographic history inference with folded SNP frequency spectra. *Genome Biology*, 21(1), 1–9. <https://doi.org/10.1186/s13059-020-02243-5>

Lladó, F. G., & Xamena, D. V. (1998). Aspectes geomorfològics quaternaris del litoral de Mallorca. En *Aspectes geològics de les Balears: Mallorca, Menorca i Cabrera* (pp. 307–329). <https://dialnet.unirioja.es/servlet/libro?codigo=767549>

Llorens, L., & Gil, L. (2017). The Balearic Islands. En *The vegetation of the Iberian Peninsula* (pp. 3–33). <https://doi.org/10.1007/978-3-319-54867-8>

Lobato, E., Doutrelant, C., Melo, M., Reis, S., & Covas, R. (2017). Insularity effects on bird immune parameters: A comparison between island and mainland populations in West Africa. *Ecology and Evolution*, 7(11), 3645–3656. <https://doi.org/10.1002/ece3.2788>

Lochmiller, R. L. (1996). Immunocompetence and animal population regulation. *Oikos*, 76, 594–602. <https://www.jstor.org/stable/3546356>

Lochmiller, R. L., & Deerenberg, C. (2000). Trade-offs in evolutionary immunology: Just what is the cost of immunity? *Oikos*, 88, 87–98. <https://doi.org/10.1034/j.1600-0706.2000.880110.x>

Losos, J. B., & Ricklefs, R. E. (2009). Adaptation and diversification on islands. *Nature*, 457(7231), 830–836. <https://doi.org/10.1038/nature07893>

Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17(1), 67–77. <https://doi.org/10.1111/1755-0998.12592>

MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography* (Vol. 1). Princeton University Press. https://books.google.com.co/books?id=a10cdkywhVgC&printsec=frontcover&hl=es&source=gbs_ge_summary_r&cad=0

Macke, E., Tasiemski, A., Massol, F., Callens, M., & Decaestecker, E. (2017). Life history and eco-evolutionary dynamics in light of the gut microbiota. <https://doi.org/10.1111/oik.03900>

Madhavan, M., Sumodan, P. K., Dhanya, C. R., Mary, A. S., & Mustafa, S. (2023). Antimicrobial peptides (AMPs): Current state and future prospects for the treatment of human parasitic diseases. En *Natural product based drug discovery against human*

parasites (pp. 203–228). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-9605-4_11

Marchi, N., Schlichta, F., & Excoffier, L. (2021). Demographic inference. *Current Biology*, 31(6), R276–R279. <https://doi.org/10.1016/j.cub.2021.01.053>

Martin, C. H., & Wainwright, P. C. (2013). Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science*. <https://doi.org/10.1126/science.1227710>

Ma, R., Wong, S. W., Ge, L., Shaw, C., Siu, S. W. I., & Kwok, H. F. (2020). and MD simulation study to explore physicochemical parameters for antibacterial peptide to become potent anticancer peptide. *Molecular Therapy – Oncolytics*, 16, 7–19. <https://doi.org/10.1016/j.omto.2019.12.001>

Mason-Smith, P. (2023). *Reptilian beta-defensins* [Doctoral dissertation, Birmingham City University]. https://www.open-access.bcu.ac.uk/15144/1/Paul%20Mason-Smith%20PhD%20Thesis%20published_Final%20version_Submitted%20Jun%202023_Final%20Award%20Nov%202023%20.pdf

Matson, K. D., & Beadell, J. S. (2010). Infection, immunity, and island adaptation in birds. https://www.researchgate.net/profile/Kevin-Matson-2/publication/220049600_Infection_immunity_and_island_adaptation_in_birds/links/0deec5213617a9f35c000000/Infection-immunity-and-island-adaptation-in-birds.pdf

Mayol, J., Bibiloni, G., Pons, G., Ballesteros, E., & Díaz, D. (2020). *Descripció general de les petites illes i els illots de les Balears*. http://ibdigital.uib.es/greenstone/sites/localsite/collect/monografiesHistoriaNatural/index/assoc/MonografiesSHNB_2020vol0/29.dir/MonografiesSHNB_2020vol029.pdf

Mayr, E. (1942). *Systematics and the origin of species from the viewpoint of a zoologist*. Harvard University Press. <https://www.hup.harvard.edu/books/9780674862500>

Mencía, A., Ortega, Z., & Pérez-Mellado, V. (2017). From tameness to wariness: Chemical recognition of snake predators by lizards in a Mediterranean island. *PeerJ*, 5, e2828. <https://doi.org/10.7717/peerj.2828>

Modesto, S. P., & Anderson, J. S. (2004). The phylogenetic definition of Reptilia. *Systematic Biology*, 53(5), 815–821. <https://doi.org/10.1080/10635150490503026>

- Mohammed, I., Said, D. G., & Dua, H. S. (2017). Human antimicrobial peptides in ocular surface defense. *Progress in Retinal and Eye Research*, 61, 1–22. <https://doi.org/10.1016/j.preteyeres.2017.03.004>
- Moore, R. D., Griffiths, R. A., & Román, A. (2004). Distribution of the Mallorcan midwife toad (*Alytes muletensis*) in relation to landscape topography and introduced predators. *Biological Conservation*, 116(3), 327–332. [https://doi.org/10.1016/S0006-3207\(03\)00202-7](https://doi.org/10.1016/S0006-3207(03)00202-7)
- Morin, P. A., Luikart, G., Wayne, R. K., & SNP Workshop Group. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, 19(4), 208–216. <https://doi.org/10.1016/j.tree.2004.01.009>
- Moritz, C. (1994). Defining “Evolutionarily Significant Units” for conservation. *Trends in Ecology & Evolution*, 9(10), 373–375. [https://doi.org/10.1016/0169-5347\(94\)90057-4](https://doi.org/10.1016/0169-5347(94)90057-4)
- Muniz, L. R., Knosp, C., & Yeretssian, G. (2012). Intestinal antimicrobial peptides during homeostasis, infection, and disease. *Frontiers in Immunology*, 3, 310. <https://doi.org/10.3389/fimmu.2012.00310>
- Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, 11, 184–194. <https://doi.org/10.1111/j.1755-0998.2011.02987.x>
- Nash, A., & Ryan, E. J. (2023). Immunity in sea turtles: Review of a host–pathogen arms race millions of years in the running. *Animals*, 13(4), 556. <https://doi.org/10.3390/ani13040556>
- Neuman-Lee, L. A., Fokidis, H. B., Spence, A. R., Van der Walt, M., Smith, G. D., Durham, S., & French, S. S. (2015). Food restriction and chronic stress alter energy use and affect immunity in an infrequent feeder. *Functional Ecology*, 29, 1453–1462. <https://doi.org/10.5061/dryad.dk166>
- Neuman-Lee, L. A., & French, S. S. (2017). Endocrine-reproductive-immune interactions in female and male Galápagos marine iguanas. *Hormones and Behavior*, 88, 60–69. <https://doi.org/10.1016/j.yhbeh.2016.10.017>
- Nielsen, E. S., Beger, M., Henriques, R., & von der Heyden, S. (2020). A comparison of genetic and genomic approaches to represent evolutionary potential in conservation

planning. *Biological Conservation*, 251, 108770.
<https://doi.org/10.1016/j.biocon.2020.108770>

Noskova, E., Abramov, N., Iliutkin, S., Sidorin, A., Dobrynin, P., & Ulyantsev, V. I. (2023). GADMA2: More efficient and flexible demographic inference from genetic data. *GigaScience*, 12, giad059. <https://doi.org/10.1093/gigascience/giad059>

Novosolov, M., Raia, P., & Meiri, S. (2013). The island syndrome in lizards. *Global Ecology and Biogeography*, 22(2), 184–191. <https://doi.org/10.1111/j.1466-8238.2012.00791.x>

Oguiura, N., Sanches, L., Duarte, P. V., Sulca-López, M. A., & Machini, M. T. (2023). Past, present, and future of naturally occurring antimicrobials related to snake venoms. *Animals*, 13(4), 744. <https://doi.org/10.3390/ani13040744>

Oliverio, M., Bologna, M. A., & Mariottini, P. (2000). Molecular biogeography of the Mediterranean lizards *Podarcis* Wagler, 1830 and *Teira* Gray, 1838 (Reptilia, Lacertidae). *Journal of Biogeography*, 27(6), 1403–1420. <https://doi.org/10.1046/j.1365-2699.2000.00517.x>

Orueta, J. F. (2003). *Manual práctico para el manejo de vertebrados invasores en islas de España y Portugal*.
https://www.researchgate.net/publication/236629147_Manual_de_manejo_de Vertebrados_invasores_en_islas

Otalora, K., Riera, J. L., Tavecchia, G., Rotger, A., Igual, J. M., Trotta, J.-R. P., & Baldo, L. (2024). Population genetics and phylogeographic history of the insular lizard (Günther, 1874) from the Balearic Islands based on genome-wide polymorphic data. *Ecology and Evolution*, 14(5), e11407. <https://doi.org/10.1002/ece3.11407>

Otalora, K., Gómez-Garrido, J., & Baldo, L. (2025). Antimicrobial peptides (AMPs) of lizards: A first comprehensive characterization of beta-defensins, ovo-defensins and cathelicidins from *Podarcis lilfordi* and closely related Lacertidae species. *BMC Genomics*, 26(1), 677. <https://doi.org/10.1186/s12864-025-11797-2>

Pavón, I., De Arcas, G., Sigcha, L., San Millán-Castillo, R., Suárez, E., & Hermida, L. F. (2019). Data collection methodology evolution for soundscape assessing: A case of study in Menorca Island (Spain).
<https://www.ingentaconnect.com/content/ince/incec/2019/00000259/00000003/art00029>

- Pérez-Cembranos, A., Pérez-Mellado, V., Alemany, I., Bassitta, M., León, A., Terrasa, B., Picornell, A., Castro, J. A., Brown, R. P., & Ramon, C. (2021). Diversidad y evolución de la lagartija balear, *Podarcis lilfordi* (Günther, 1874). *Revista de Menorca*, 100, 113–142. <https://produccioncientifica.usal.es/documentos/6381692318a84b178feaa674>
- Pérez-Mellado, V. (1998). *Podarcis lilfordi* (Günther, 1874). En A. Salvador (Coord.), *Reptiles* (pp. 272–282). En M. A. Ramos et al. (Eds.), *Fauna Ibérica*. Museo Nacional de Ciencias Naturales.
- Pérez-Mellado, V., & Martínez-Solano, I. (2009a). *Podarcis lilfordi*. *The IUCN Red List of Threatened Species* 2009: e.T17795A7481971. <https://doi.org/10.2305/IUCN.UK.2009.RLTS.T17795A7481971.en>
- Pérez-Mellado, V., & Martínez-Solano, I. (2009b). *Podarcis lilfordi*. *The IUCN Red List of Threatened Species* 2009: e.T17795A7481971. <http://dx.doi.org/10.2305/IUCN.UK.2009.RLTS.T17795A7481971.en>
- Pérez-Mellado, V., & Pérez-Cembranos, A. (2025). Effects of the introduction of an herbivore on an endangered lizard. *European Journal of Wildlife Research*, 71(3). <https://doi.org/10.1007/s10344-025-01925-4>
- Pons, G. X., & del Valle, L. (2020). Bionomic cartography of the Balearic Sea: Evolution of coastal marine habitats of Mallorca. *Journal of Coastal Research*, 95(sp1), 153–157. <https://doi.org/10.2112/SI95-030.1>
- Pough, H. F. (1980). The advantages of ectothermy for tetrapods. *The American Naturalist*, 115, 92–112. <https://www.journals.uchicago.edu/doi/abs/10.1086/283547>
- Poulakakis, N., Lymberakis, P., Valakos, E., Zouros, E., & Mylonas, M. (2005). Phylogenetic relationships and biogeography of *Podarcis* species from the Balkan Peninsula, by Bayesian and maximum likelihood analyses of mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, 37(3), 845–857. <https://doi.org/10.1016/j.ympev.2005.06.005>
- Prates, I., Penna, A., Rodrigues, M. T., & Carnaval, A. C. (2018). Local adaptation in mainland anole lizards: Integrating population history and genome–environment associations. *Ecology and Evolution*, 8(23), 11932–11944. <https://doi.org/10.1002/ece3.4650>

Puritz, J. B., Addison, J. A., & Toonen, R. J. (2012). Next-generation phylogeography: A targeted approach for multilocus sequencing of non-model organisms. *PLOS ONE*, 7(3), e34241. <https://doi.org/10.1371/journal.pone.0034241>

Rahnamaeian, M. (2011). Antimicrobial peptides: Modes of mechanism, modulation of defense responses. *Plant Signaling & Behavior*, 6(9), 1325. <https://doi.org/10.4161/psb.6.9.16319>

Rai, M., Deshmukh, S. D., Ingle, A. P., Gupta, I. R., Galdiero, M., & Galdiero, S. (2016). Metal nanoparticles: The protective nanoshield against virus infection. *Critical Reviews in Microbiology*, 42(1), 46–56. <https://doi.org/10.3109/1040841X.2013.879849>

Rebouças, R., Augusto-Alves, G., & Toledo, L. F. (2020). Evolution of treefrogs' calls in tropical islands might be under directional selection. *Journal of Zoology*, 312(1), 43–52. <https://doi.org/10.1111/jzo.12792>

Reilly, S. B., Marks, S. B., & Jennings, W. B. (2012). Defining evolutionary boundaries across parapatric ecomorphs of Black Salamanders (*Aneides flavipunctatus*) with conservation implications. *Molecular Ecology*, 21(23). <https://doi.org/10.1111/mec.12068>

Rojas-Pirela, M., Kemmerling, U., Quiñones, W., Michels, P. A. M., & Rojas, V. (2023). Antimicrobial peptides (AMPs): Potential therapeutic strategy against trypanosomiasis? *Biomolecules*, 13(4), 599. <https://doi.org/10.3390/biom13040599>

Rodríguez, V., Brown, R. P., Terrasa, B., Pérez-Mellado, V., Castro, J. A., Picornell, A., & Ramon, M. M. (2013). Multilocus genetic diversity and historical biogeography of the endemic wall lizard from Ibiza and Formentera, *Podarcis pityusensis* (Squamata: Lacertidae). *Molecular Ecology*, 22(19), 4829–4841. <https://doi.org/10.1111/mec.12443>

Rotger, A., Igual, J. M., Genovart, M., Rodríguez, V., Ramon, C., Pérez-Mellado, V., Bibiloni, G., Rita, J., & Tavecchia, G. (2021). Contrasting adult body-size in sister populations of the Balearic lizard, *Podarcis lilfordi* (Günther 1874) suggests anthropogenic selective pressures. *Herpetological Monographs*, 35(1), 53–64. <https://produccioncientifica.usal.es/documentos/618608828a2f583ff09af545?lang=en>

Ruperao, P., Bajaj, P., Subramani, R., Yadav, R., Lachagari, V. B. R., Lekkala, S. P., Rathore, A., Archak, S., Angadi, U. B., Singh, R., Singh, K., Mayes, S., & Rangan, P. (2023). A pilot-scale comparison between single and double-digest RAD markers generated using GBS

strategy in sesame (*Sesamum indicum* L.). *PLOS ONE*, 18(6), e0286599. <https://doi.org/10.1371/journal.pone.0286599>

Russo, D., Teixeira, S., Cistrone, L., Jesus, J., Teixeira, D., Freitas, T., & Jones, G. (2009). Social calls are subject to stabilizing selection in insular bats. *Journal of Biogeography*, 36(12), 2212–2221. <https://doi.org/10.1111/j.1365-2699.2009.02158.x>

Sabolić, I., & Štambuk, A. (2021). *Genomic patterns of phenotypic adaptation in Italian wall lizard Podarcis siculus (Rafinesque-Schmaltz, 1810)* [Doctoral dissertation]. <https://repozitorij.pmf.unizg.hr/islandora/object/pmf:9358>

Salvador, A. (2009). Lagartija balear – *Podarcis lilfordi* (Günther, 1874). https://digital.csic.es/bitstream/10261/109156/1/podlil_v3.pdf

Salvà Matas, C. (2013). La memòria d'un paisatge gravat: Les pedreres de marès, empremta territorial del paisatge identitari de Mallorca. V *Seminario Internacional de Investigación en Urbanismo* (Barcelona-Buenos Aires), 1733–1749. <https://doi.org/10.5821/siiu.5931>

Santana, F. L., Estrada, K., Ortiz, E., & Corzo, G. (2021). Reptilian β -defensins: Expanding the repertoire of known crocodylian peptides. *Peptides*, 136, 170473. <https://doi.org/10.1016/j.peptides.2020.170473>

Santana, F. L., Arenas, I., Haney, E. F., Estrada, K., Hancock, R. E. W., & Corzo, G. (2021). Identification of a crocodylian β -defensin variant from *Alligator mississippiensis* with antimicrobial and antibiofilm activity. *Peptides*, 140, 170549. <https://doi.org/10.1016/j.peptides.2021.170549>

Schiffels, S., & Durbin, R. (2014). Inferring human population size and separation history from multiple genome sequences. *Nature Genetics*, 46(8), 919–925. <https://doi.org/10.1101/005348>

Schluter, D. (2000). *The ecology of adaptive radiation*. Oxford University Press. <https://global.oup.com/academic/product/the-ecology-of-adaptive-radiation-9780198505228>

Sellinger, T. P. P., Abu-Awad, D., & Tellier, A. (2021). Limits and convergence properties of the sequentially Markovian coalescent. *Molecular Ecology Resources*, 21(7), 2231–2248. <https://doi.org/10.1111/1755-0998.13416>

Semple, F., & Dorin, J. R. (2012). β -Defensins: Multifunctional modulators of infection, inflammation and more? *Journal of Innate Immunity*, 4(4), 337–348. <https://doi.org/10.1159/000336619>

Sendell-Price, A. T., Ruegg, K. C., Robertson, B. C., & Clegg, S. M. (2021). An island-hopping bird reveals how founder events shape genome-wide divergence. *Molecular Ecology*, 30(11), 2495–2510. <https://doi.org/10.1111/mec.15898>

Siddall, M., Rohling, E. J., Almogi-Labin, A., Hemleben, C., Meischner, D., Schmelzer, I., & Smeed, D. A. (2003). Sea-level fluctuations during the last glacial cycle. *Nature*, 423(6942), 853–858. <https://doi.org/10.1038/nature01690>

Stankowski, S., Cutter, A. D., Satokangas, I., Lerch, B. A., Rolland, J., Smadja, C. M., Segami Marzal, J. C., Cooney, C. R., Feulner, P. G. D., Domingos, F. M. C. B., North, H. L., Yamaguchi, R., Butlin, R. K., Wolf, J. B. W., Coughlan, J., Heidbreder, P., Hernández-Gutiérrez, R., Barnard-Kubow, K. B., Peede, D., ... Kulmuni, J. (2024). Toward the integration of speciation research. *Evolutionary Journal of the Linnean Society*, 3(1), kzae001. <https://doi.org/10.1093/evolinnean/kzae001>

Sutherland, W. J., Pullin, A. S., Dolman, P. M., & Knight, T. M. (2004). The need for evidence-based conservation. *Trends in Ecology & Evolution*, 19(6), 305–308. <https://doi.org/10.1016/j.tree.2004.03.018>

Templeton, A. R. (2006). *Population genetics and microevolutionary theory*. <https://doi.org/10.1002/0470047356>

Templeton, A. R., Robertson, R. J., Brisson, J., & Strasburg, J. (2001). Disrupting evolutionary processes: The effect of habitat fragmentation on collared lizards in the Missouri Ozarks. *Proceedings of the National Academy of Sciences*, 98(10), 5426–5432. <https://doi.org/10.1073/pnas.091093098>

Terhorst, J., Kamm, J. A., & Song, Y. S. (2016). Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nature Genetics*, 49(2), 303–309. <https://doi.org/10.1038/ng.3748>

Terrasa, B., Pérez-Mellado, V., Brown, R. P., Picornell, A., Castro, J. A., & Ramon, M. M. (2009). Foundations for conservation of intraspecific genetic diversity revealed by analysis of phylogeographical structure in the endangered endemic lizard *Podarcis lilfordi*. *Diversity and Distributions*, 15(2), 207–221. <https://doi.org/10.1111/j.1472-4642.2008.00520.x>

Terrasa, B., Picornell, A., Castro, J. A., & Ramon, M. M. (2004). Genetic variation within endemic *Podarcis* lizards from the Balearic Islands inferred from partial cytochrome *b* sequences. *Amphibia-Reptilia*, 25(4), 407–414. <https://doi.org/10.1163/1568538042788960>

Thompson, J. D. (2020). *The ecology and biogeography of endemism*. Oxford University Press. <https://doi.org/10.1093/oso/9780198835141.003.0002>

Torkamaneh, D., Laroche, J., & Belzile, F. (2016). Genome-wide SNP calling from genotyping-by-sequencing (GBS) data: A comparison of seven pipelines and two sequencing technologies. *PLOS ONE*, 11(8), e0161333. <https://doi.org/10.1371/journal.pone.0161333>

Torres, M. D., Sothiselvam, S., Lu, T. K., & de la Fuente-Nunez, C. (2019). Peptide design principles for antimicrobial applications. *Journal of Molecular Biology*, 431(18), 3547–3567. <https://doi.org/10.1016/j.jmb.2018.12.015>

Traveset, A., Nogales, M., Alcover, J. A., Delgado, J. D., López-Darias, M., Godoy, D., ... Bover, P. (2009). A review on the effects of alien rodents in the Balearic (Western Mediterranean Sea) and Canary Islands (Eastern Atlantic Ocean). *Biological Invasions*, 11(7), 1653–1670. <https://doi.org/10.1007/s10530-008-9395-y>

Ulaszewski, B., Meger, J., & Burczyk, J. (2021). Comparative analysis of SNP discovery and genotyping in *Fagus sylvatica* L. and *Quercus robur* L. using RADseq, GBS, and ddRAD methods. *Forests*, 12(2), 222. <https://doi.org/10.3390/f12020222>

Uller, T., Laakkonen, H., Michaelides, S., While, G. M., Coulon, A., & Aubret, F. (2019). Genetic differentiation predicts body size divergence between island and mainland populations of common wall lizards (*Podarcis muralis*). *Biological Journal of the Linnean Society*, 200, 1–16. <https://doi.org/10.1093/biolinnean/blz052>

Uwishema, O., Masunga, D. S., Naisikye, K. M., Bhanji, F. G., Rapheal, A. J., Mbwana, R., ... Wellington, J. (2023). Impacts of environmental and climatic changes on future infectious diseases. *International Journal of Surgery*, 109(2), 167–170. <https://doi.org/10.1097/JS9.0000000000000160>

van Dijk, A., Veldhuizen, E. J. A., & Haagsman, H. P. (2008). Avian defensins. *Veterinary Immunology and Immunopathology*, 124(1–2), 1–18. <https://doi.org/10.1016/j.vetimm.2007.12.006>

van Harten, R. M., van Woudenberg, E., van Dijk, A., & Haagsman, H. P. (2018). Cathelicidins: Immunomodulatory antimicrobials. *Vaccines*, 6(3), 63. <https://doi.org/10.3390/vaccines6030063>

van Hoek, M. L. (2014). Antimicrobial peptides in reptiles. *Pharmaceuticals*, 7(6), 723–753. <https://doi.org/10.3390/ph7060723>

van Hoek, M. L., Prickett, M. D., Settlege, R. E., Kang, L., Michalak, P., Vliet, K. A., & Bishop, B. M. (2019). The Komodo dragon (*Varanus komodoensis*) genome and identification of innate immunity genes and clusters. *BMC Genomics*, 20(1), 382. <https://doi.org/10.1186/s12864-019-6029-y>

Van Nostrand, E. L., Freese, P., Pratt, G. A., Wang, X., Wei, X., Xiao, R., Blue, S. M., Chen, J.-Y., Cody, N. A. L., Dominguez, D., Olson, S., Sundararaman, B., Zhan, L., Bazile, C., Bouvrette, L. P. B., Bergalet, J., Duff, M. O., Garcia, K. E., Gelboin-Burkhart, C., ... Yeo, G. W. (2020). A large-scale binding and functional map of human RNA-binding proteins. *Nature*, 583(7818), 711–719. <https://doi.org/10.1038/s41586-020-2077-3>

Velová, H., Gutowska-Ding, M. W., Burt, D. W., & Vinkler, M. (2018). Toll-like receptor evolution in birds: Gene duplication, pseudogenization, and diversifying selection. *Molecular Biology and Evolution*, 35(9), 2170–2184. <https://doi.org/10.1093/molbev/msy119>

Wang, G., Li, X. y Wang, Z. (2016). APD3: la base de datos de péptidos antimicrobianos como herramienta para la investigación y la educación. *Investigación de ácidos nucleicos*, 44 (D1), D1087-D1093. <https://doi.org/10.1093/nar/gkv1278>

Wang, G. (2021). Unifying the classification of antimicrobial peptides in the antimicrobial peptide database. *Methods in Enzymology*, 663, 1. <https://doi.org/10.1016/bs.mie.2021.09.006>

Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilée, R., Condamine, F. L., Gravel, D., Morlon, H., Mouquet, N., Rosindell, J., Casquet, J., Conti, E., Cornuault, J., Fernández-Palacios, J. M., Hengl, T., Norder, S. J., Rijsdijk, K. F., Sanmartín, I., Strasberg, D., Triantis, K. A., ... Thébaud, C. (2015). Islands as model systems in ecology and evolution: Prospects fifty years after MacArthur–Wilson. *Ecology Letters*, 18(2), 200–217. <https://doi.org/10.1111/ele.12398> .

Whenham, N., Lu, T. C., Maidin, M. B. M., Wilson, P. W., Bain, M. M., Stevenson, M. L., Stevens, M. P., Bedford, M. R., & Dunn, I. C. (2015). Ovo-defensins, an oviduct-specific

antimicrobial gene family, have evolved in birds and reptiles to protect the egg by both sequence and intra-six-cysteine sequence motif spacing. *Biology of Reproduction*, 92(6), 154, 1–13. <https://doi.org/10.1095/biolreprod.114.126839>

Whitlock, M. C. (2000). Fixation of new alleles and the extinction of small populations: Drift load, beneficial alleles, and sexual selection. *Evolution*, 54(6), 1855–1861. <https://doi.org/10.1111/j.0014-3820.2000.tb01232.x>

Whittaker, R. J., Fernández-Palacios, J. M., Matthews, T. J., Borregaard, M. K., & Triantis, K. A. (2017). Island biogeography: Taking the long view of nature's laboratories. *Science*, 357(6354), eaam8326. <https://doi.org/10.1126/science.aam8326>

Wikelski, M., Foufopoulos, J., Vargas, H., & Snell, H. (2004). Galápagos birds and diseases: Invasive pathogens as threats for island species. https://www.researchgate.net/publication/42763413_Galapagos_Birds_and_Diseases_Invasive_Pathogens_as_Threats_for_Island_Species

Wolf, J. B. W., Lindell, J., & Backström, N. (2010). Speciation genetics: Current status and evolving approaches. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1547), 1717–1733. <https://doi.org/10.1098/rstb.2010.0023>

Woolfit, M., & Bromham, L. (2005). Population size and molecular evolution on islands. *Proceedings of the Royal Society B: Biological Sciences*, 272(1578), 2277–2282. <https://doi.org/10.1098/rspb.2005.3217>

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16(2), 97–159. <https://doi.org/10.1093/genetics/16.2.97>

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, 15(4), 323–354. <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>

Wu, X., Wang, Z., Li, X., Fan, Y., He, G., Wan, Y., Yu, C., Tang, J., Li, M., Zhang, X., Zhang, H., Xiang, R., Pan, Y., Liu, Y., Lu, L., & Yang, L. (2014). *In vitro* and *in vivo* activities of antimicrobial peptides developed using an amino acid-based activity prediction method. *Antimicrobial Agents and Chemotherapy*, 58(9), 5342–5349. <https://doi.org/10.1128/AAC.02823-14>

Yang, W., Feiner, N., Pinho, C., While, G. M., Kaliontzopoulou, A., Harris, D. J., Salvi, D., & Uller, T. (2021). Extensive introgression and mosaic genomes of Mediterranean

endemic lizards. *Nature Communications*, 12(1), 2762. <https://doi.org/10.1038/s41467-021-22949-9>

Zhang, Q.-Y., Yan, Z.-B., Meng, Y.-M., Hong, X.-Y., Shao, G., Ma, J.-J., Cheng, X.-R., Liu, J., Kang, J., & Fu, C.-Y. (2021). Antimicrobial peptides: Mechanism of action, activity and clinical potential. *Military Medical Research*, 8(1), 48. <https://doi.org/10.1186/s40779-021-00343-2>

Zhao, H., Gan, T.-X., Liu, X.-D., Jin, Y., Lee, W.-H., Shen, J.-H., & Zhang, Y. (2008). Identification and characterization of novel reptile cathelcidins from elapid snakes. *Peptides*, 29(10), 1685–1691. <https://doi.org/10.1016/j.peptides.2008.06.008>

Zhu, S., & Gao, B. (2013). Evolutionary origin of β -defensins. *Developmental and Comparative Immunology*, 39(1–2), 79–84. <https://doi.org/10.1016/j.dci.2012.02.011>

Zimmerman, L. M., Vogel, L. A., & Bowden, R. M. (2010). Understanding the vertebrate immune system: Insights from the reptilian perspective. *Journal of Experimental Biology*, 213, 661–671. <https://doi.org/10.1242/jeb.038315>

Zimmerman, L. M. (2020). The reptilian perspective on vertebrate immunity: 10 years of progress. *Journal of Experimental Biology*, 223(21), jeb214171. <https://doi.org/10.1242/jeb.214171>

9. ANNEX

Population genetics and phylogeographic history of the insular lizard *Podarcis lilfordi* (Günther, 1874) from the Balearic Islands based on genome-wide polymorphic data

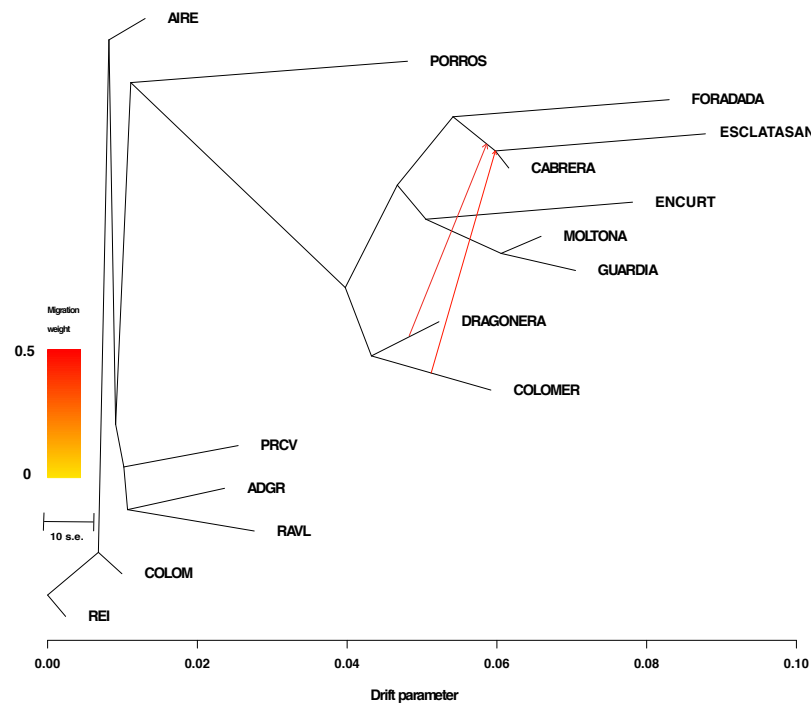


Figure S1: The ML-graph inferred by TreeMix with 16 migration events based on the COMBINED dataset. The arrow indicates the two significant migration events with the highest migration weight. Drift parameter is shown on the x-axis.

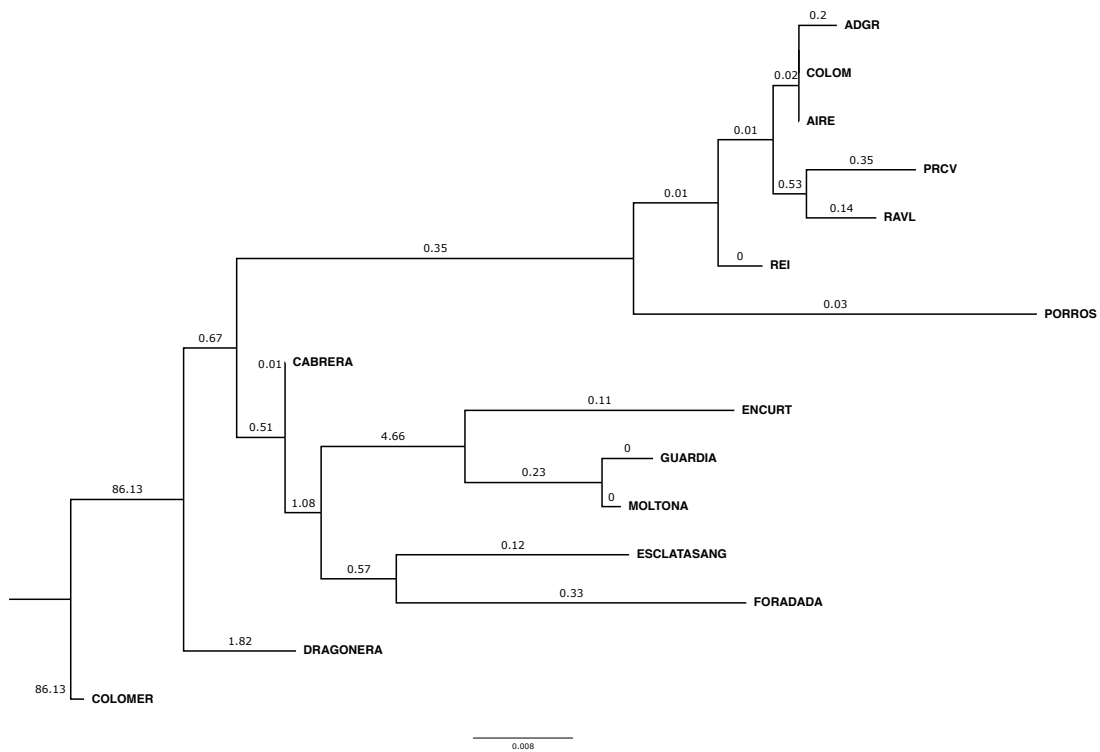


Figure S2: ML tree with rootstrap support values for root inference. Values were estimated with IQ-TREE 2 on consensus sequences per population (2867 bp), using a non-reversible model (model 12.12, all sites, 10000 bootstrap replicates).

A)



B)



Figure S3. Functional enrichment of GOs for proteins with outliers, according to A) GBS and B) RADSeq dataset.

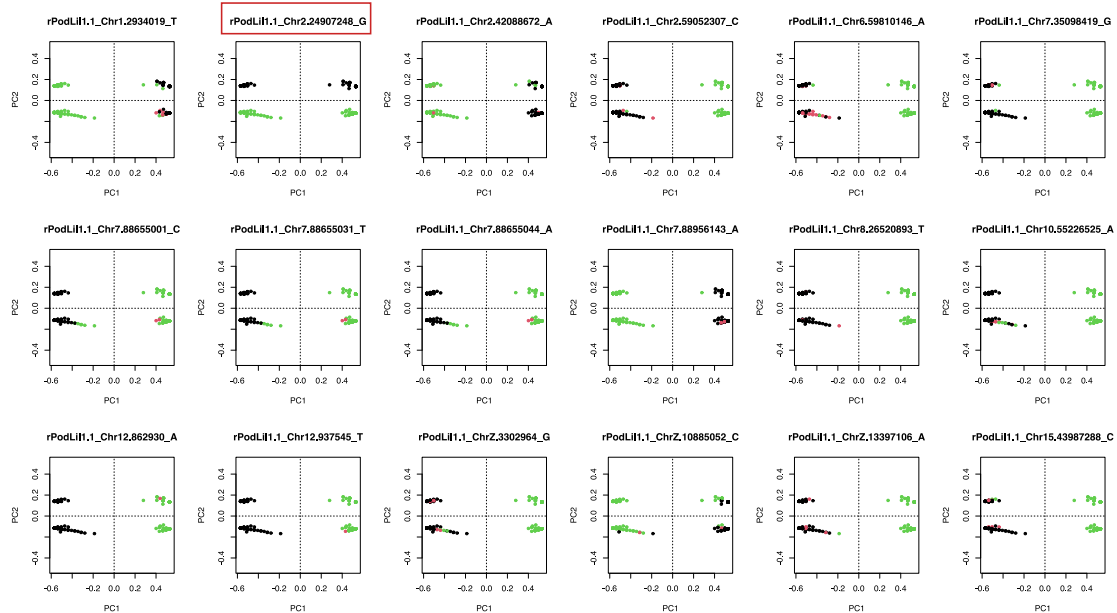


Figure S4. PCA of the 18 outliers detected in the COMBINED dataset. All but one indicated a clear clustering by archipelago. Black dots correspond to the most frequent allele, while Red dots indicate the heterozygous genotype.

Table S1-S5: <https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.11407>

Antimicrobial peptides (AMPs) of lizards: a first comprehensive characterization of beta-defensins, ovo-defensins and cathelicidins from *Podarcis lilfordi* and closely related Lacertidae species

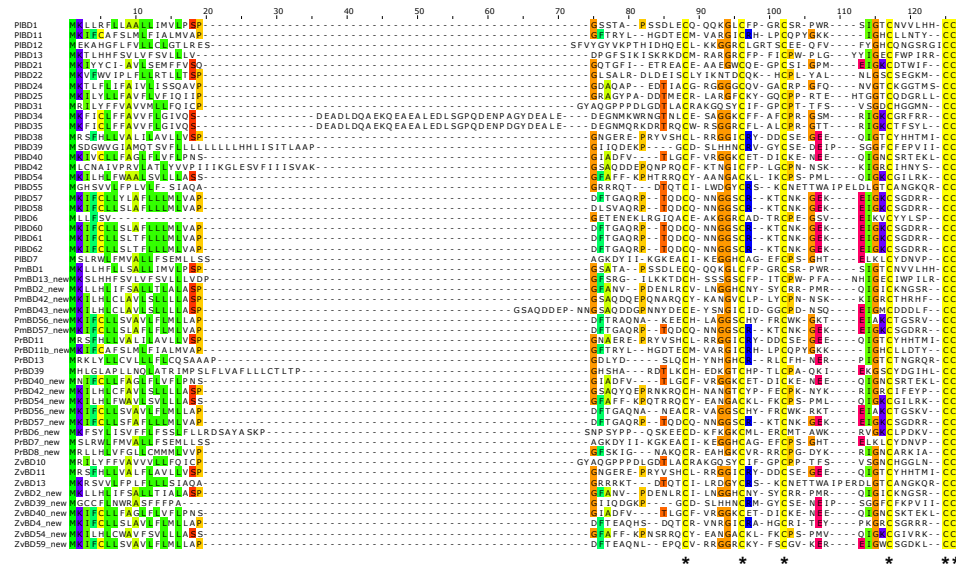


Figure S1: Multispecies alignment of beta-defensins proteins (partial) from the Lacertidae species under study (Pm, Pr, Pl and Zv).

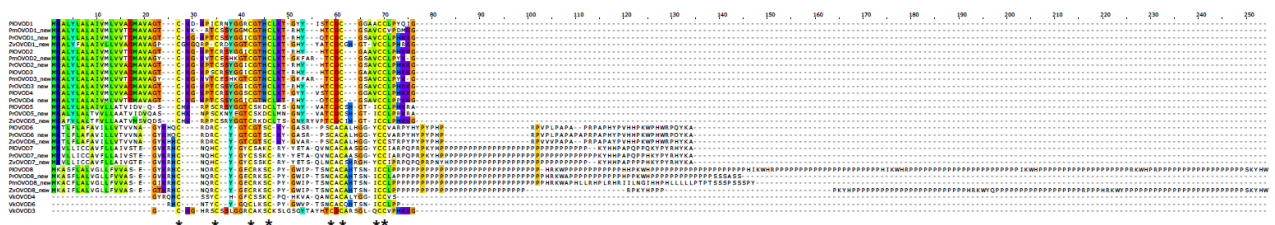


Figure S2: Multispecies alignment of ovo-defensins proteins (complete) from the Lacertidae species under study and the Komodo dragon (Pm, Pr, Pl, Zv and Vk).

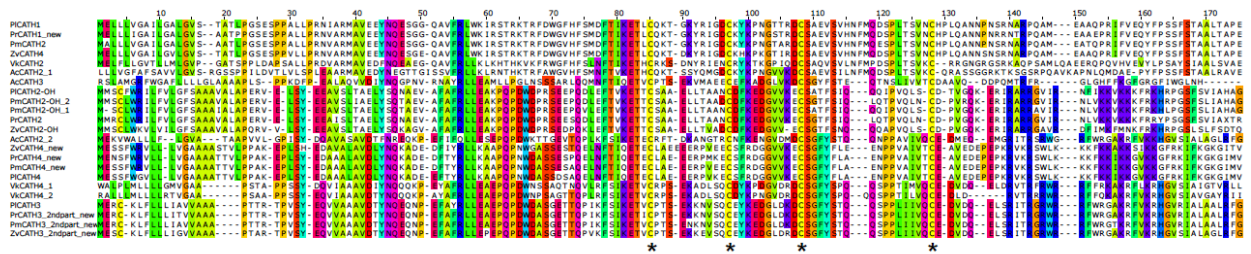


Figure S3: Multispecies alignment of cathelicidins proteins (partial) from Lacertidae species, the Komodo dragon and the green anole lizard (Pm, Pr, Pl, Zv, Vk and Ac).

Table S1-S5: <https://bmcbgenomics.biomedcentral.com/articles/10.1186/s12864-025-11797-2#Sec20>

10. ORIGINAL PUBLICATIONS

DNA Research, 2023, 30, 1–10
<https://doi.org/10.1093/dnares/dsac008>
Advance access publication 4 May 2023
Resource Article: Genomes Explored

OXFORD

Resource Article: Genomes Explored

Chromosome-level genome assembly of Lilford's wall lizard, *Podarcis lilfordi* (Günther, 1874) from the Balearic Islands (Spain)

Jessica Gomez-Garrido¹, Fernando Cruz¹, Tyler S. Alioto¹, Nathalie Feiner², Tobias Uller^{2,3}, Marta Gut^{1,3}, Ignacio Sanchez Escudero¹, Giacomo Tavecchia^{4,5}, Andreu Rotger⁴, Katherin Eliana Otalora Acevedo^{5,6}, and Laura Baldo^{5,7,*}

¹CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), 08028 Barcelona, Spain

²Department of Biology, Lund University, Lund, Sweden

³Universitat Pompeu Fabra (UPF), Barcelona, Spain

⁴Animal Demography and Ecology Unit, IMEDEA, CSIC-UIB, Esporles, Spain

⁵Department of Evolutionary Biology, Ecology and Environmental Sciences, University of Barcelona, Barcelona, Spain

⁶Fundación Motiva Inteligencia Colectiva, Biodiversity Branch, Tunja, Boyacá, Colombia

⁷Institute for Research on Biodiversity (IRBio), University of Barcelona, Barcelona, Spain

*To whom correspondence should be addressed. Tel. +34 9340 21515. Email: baldo.laura@ub.edu; g.tavecchia@imedea.uib-csic.es

Abstract

The Mediterranean lizard *Podarcis lilfordi* is an emblematic species of the Balearic Islands. The extensive phenotypic diversity among extant isolated populations makes the species a great insular model system for eco-evolutionary studies, as well as a challenging target for conservation management plans. Here we report the first high-quality chromosome-level assembly and annotation of the *P. lilfordi* genome, along with its mitogenome, based on a mixed sequencing strategy (10X Genomics linked reads, Oxford Nanopore Technologies long reads and Hi-C scaffolding) coupled with extensive transcriptomic data (Illumina and PacBio). The genome assembly (1.5 Gb) is highly contiguous (N50 = 90 Mb) and complete, with 99% of the sequence assigned to candidate chromosomal sequences and >97% gene completeness. We annotated a total of 25,663 protein-coding genes translating into 38,615 proteins. Comparison to the genome of the related species *Podarcis muralis* revealed substantial similarity in genome size, annotation metrics, repeat content, and a strong collinearity, despite their evolutionary distance (~18–20 MYA). This genome expands the repertoire of available reptilian genomes and will facilitate the exploration of the molecular and evolutionary processes underlying the extraordinary phenotypic diversity of this insular species, while providing a critical resource for conservation genomics.

Key words: Hi-C, Oxford Nanopore Technologies (ONT), RNAseq, mitochondrial genome, reptile

1. Introduction

Podarcis lilfordi (Günther, 1874), also known as Lilford's wall lizard, is an endemic lizard of the Balearic Islands (Spain), currently confined to the Cabrera archipelago and several islets surrounding Menorca and Mallorca.¹ Due to a patchy distribution (~43 isolated populations) and threats from habitat loss, the species is currently listed as endangered by the IUCN (<https://www.iucnredlist.org/species/17795/7481971>). An extensive morphological diversity has resulted in the description of a large number (currently 23) of subspecies or morphotypes, but current understanding considers each population/islet as a distinct evolutionary unit.¹ The species exhibits several insular characteristics that likely reflect the lack of natural predators, including high population densities,² high survival and low fecundity^{2–5} compared to their continental relatives. The insularity and large phenotypic diversity of *P. lilfordi* have made it a popular model for

studies of ecological adaptation and demographic dynamics of terrestrial vertebrates, as well as for testing predictions of how evolution proceeds on islands. During the past decades, substantial efforts have been directed toward understanding the species' morphological diversity,^{1,3,6} trophic ecology,⁷ life history traits and demographic resilience,^{3,8} genetics^{5,9} and associated gut microbes.^{10–12} Despite this extensive knowledge and the endemic character of *P. lilfordi*, genomic information on the species has been limited.⁹

Here we present the complete genome sequence of *P. lilfordi* (both nuclear and mitochondrial) from a single female specimen (the heterogametic sex, ZW) collected from Aire Island (Menorca), where the species was first described (Günther, 1874). We used a sequencing strategy combining long and short reads, coupled with RNAseq data from multiple tissues and specimens, to achieve a highly robust assembly and annotation. The *P. lilfordi* genome, along with

Received 9 February 2023; Revised 27 March 2023; Accepted 1 May 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of Kazusa DNA Research Institute.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from <https://academic.oup.com/dnares/article/30/3/dsac008/7151100> by guest on 25 May 2025

that of its continental relative *Podarcis muralis*,¹³ will extend current genomic resources within this highly diverse lizard genus^{14,15} and provide a genetic framework to begin understanding the processes behind the remarkable diversification of the Lilford's wall lizard. In addition, the genome will provide a critical tool for the development of a conservation plan based on population-level genomics (accounting for levels of genetic diversity, drift, and accumulation of deleterious mutations) for an accurate assessment of the species adaptive potential and resilience in face of current challenges by global change and increasing human pressure in the Balearic Islands.

2. Materials and methods

2.1. Sample collection and processing

Five adult specimens of *P. lilfordi* (subspecies *lilfordi*) were collected in April 2021 on Aire Island, to the southeast of Menorca Island (Spain). The island is one of the largest and most densely populated (surface area 34 ha, 4,098.60 lizards/ha).⁴ Specimens were caught using pitfall traps containing fruit juices placed along paths and vegetation edges. All specimens were sexed according to visual examination of femoral pores and morphology,¹⁶ weighed and body size measured as snout to vent length (see [Supplementary Table S1](#) for metadata). Individuals were sacrificed by cervical dislocation, immediately stored on dry ice, and kept at -80°C until processing. Under a sterile hood, each frozen specimen was rapidly dissected to extract all major organs, including heart, liver, lungs, kidneys, brain, testicles/ovaries, intestine and muscle tissue from the tail. All tissues were immediately stored at -80°C .

A single female adult specimen (the heterogametic sex, ZW) was chosen for genome sequencing. Tissues were sent to the Centre Nacional d'Anàlisi Genòmica (CNAG-CRG) for DNA extraction and sequencing. Samples from the remaining specimens were shipped to Lund University (Sweden) for RNA extraction. RNA sequencing was performed at the SNP&SEQ Technology Platform (for short reads), and at the Uppsala Genome Center (for long reads) (SciLifeLabs, Sweden).

Specimen sampling was granted by the Species Protection Service (Department of Agriculture, Environment and Territory, Government of the Balearic Islands), under permit CAP03/2021 (to LB). The specimen remains were deposited at the Museum of Natural History of Barcelona (Spain), under voucher name MZB 2022-5701.

2.2. Genomic DNA extractions

High-molecular-weight (HMW) gDNA was extracted from frozen liver using the Nanobind tissue kit (Circulomics) following the manufacturer's protocol. Briefly, two cryopreserved liver aliquots of 38 mg and 24 mg were homogenized under cryogenic conditions on dry ice, using a mortar and pestle. The pulverized tissue was collected into 1.5 ml tubes with lysis buffer (Circulomics). Nanobind disk (Circulomics) was used on fresh supernatant for the gDNA binding. The HMW gDNA eluate was quantified by Qubit DNA BR Assay kit (Thermo Fisher Scientific) and the DNA purity was evaluated using Nanodrop 2000 (Thermo Fisher Scientific) UV/Vis measurements. The gDNA integrity was evaluated with pulsed-field gel electrophoresis SeaKem® GOLD Agarose 1% (Lonza), using the Pippin Pulse (Sage Science). The gDNA samples were stored at 4°C .

2.3. Genome sequencing

The complete genome sequence was achieved employing a mixed sequencing strategy, combining the use of 10X linked short reads (Illumina NovaSeq 6000, 2×150 bp) for base accuracy, long reads (Oxford Nanopore Technologies, ONT) for high contiguity and repeat resolution, and Hi-C (Illumina NovaSeq 6000, 2×150 bp) for chromosome-level scaffolding ([Fig. 1](#)).

The linked reads library was prepared using the Chromium Controller instrument (10X Genomics) and Genome Reagent Kits v2 (10X Genomics) following the manufacturer's protocol. Briefly, 10 ng of HMW gDNA was portioned in GEM reactions including a unique barcode (Gemcode) after loading onto a chromium controller chip. The droplets were then recovered, isothermally incubated, fractured and the intermediate DNA library was then purified and size-selected using Silane and Solid Phase reverse immobilisation (SPRI) beads. Illumina-compatible paired-end sequencing libraries were prepared following 10X Genomics recommendations and validated on an Agilent 2100 BioAnalyzer with the DNA 7500 assay (Agilent). The library was sequenced on NovaSeq 6000 (Illumina, 2×151 bp) following the manufacturer's protocol for dual indexing. Image analysis, base calling and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (RTA v3.4.4) and followed by the generation of FASTQ files.

The ONT libraries were prepared using the 1D Sequencing kit SQK-LSK110. Briefly, 1.0 μg of the HMW gDNA was DNA-repaired and DNA-end-repaired using NEBNext FFPE DNA Repair Mix and the NEBNext UltraII End Repair/dA-Tailing Module (New England Biolabs, NEB), followed by ligation of sequencing adaptors. The library was purified with 0.4 X AMPure XP Beads and eluted in Elution Buffer. Four sequencing runs were performed on GridION Mk1 (ONT) using an R9.4.1 flow cell (FLO-MIN106D, ONT) on the MinKNOW platform version 4.2.5 for real-time monitoring. Data were collected for 110 h and base called with Guppy version 4.3.4 in high accuracy mode using the dna_r9.4.1_450bps_modbases_dam-dcm-cpg_hac.cfg model.

The Hi-C libraries were prepared using the Omni-C kit (Dovetail Genomics), following the manufacturer's protocol. Briefly, frozen muscle tissue (heart) was pulverized using a mortar and pestle immersed in a liquid nitrogen bath. Chromatin was fixed in place with formaldehyde (Sigma Aldrich), digested with DNase I and DNA extracted. DNA ends were repaired, and a biotinylated bridge adapter was ligated followed by proximity ligation of adapter-containing ends. After reverse crosslinking, the DNA was purified and followed by the preparation of Illumina-compatible paired-end sequencing libraries (omitting the fragmentation step). Biotinylated chimeric molecules were isolated using streptavidin beads before PCR enrichment of the library. The library was sequenced on NovaSeq 6000 (Illumina, 2×151 bp) following the manufacturer's protocol for dual indexing.

2.4. Transcriptome sequencing

RNA was extracted from five different tissues (heart, kidney, liver, lungs and tail) of five specimens using the RNeasy Mini Kits (Qiagen) and including an on-column DNA digestion step. RNA concentration was measured using the Qubit RNA High Sensitivity assay and RNA quality was assessed using the Eukaryote Total RNA Nano assay on the Agilent

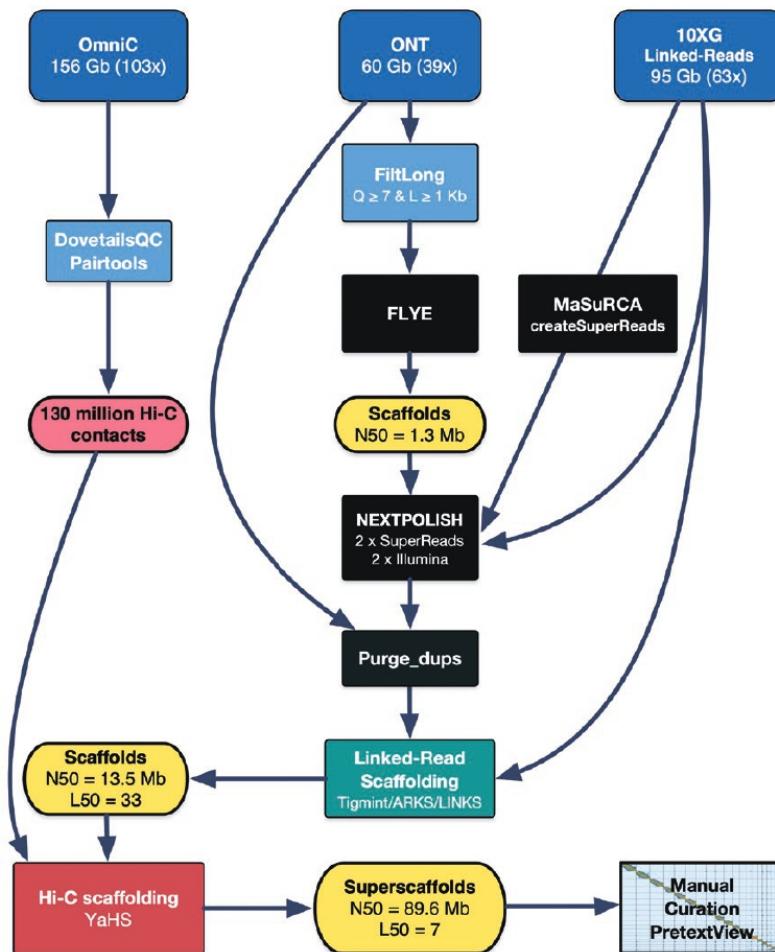


Figure 1. Assembly workflow. Summary of the main steps followed to obtain the curated chromosome-level assembly. See the Snakemake pipeline for detailed information (https://github.com/cnag-aat/assembly_pipeline).

2100 Bioanalyzer system. For short-read RNA sequencing, we pooled samples of two to four individuals per tissue in equimolar amounts (Supplementary Table S2). These five tissue-specific samples were subjected to Illumina's TruSeq Stranded mRNA library preparation kit and sequenced on NovaSeq 6000 (Illumina, 2 × 150 bp). For long-read RNA sequencing, we pooled five individual tissue samples in equimolar amounts (Supplementary Table S2) and subjected this single, mixed sample to PacBio Iso-Seq sequencing.

2.5. Nuclear genome assembly

The nuclear genome was assembled following the workflow summarized in Fig. 1. To increase reproducibility, the main steps have been included into a snakemake¹⁷ pipeline (https://github.com/cnag-aat/assembly_pipeline). A few modifications were added for this project which are detailed below, along with a summary of major steps. All intermediate and final assemblies produced during the process were evaluated as

indicated in the pipeline description, using BUSCO^{18,19} v 4.0.6 with *vertebrata_odb10*, *fasta-stats.py* (<https://github.com/galaxyproject/tools-iuc>), *Nseries.pl* and Merquy²⁰ v1.1. Note that our final assembly statistics was further evaluated with updated versions of Merquy v1.3 and BUSCO version 5.0.4 with *vertebrata_odb10*.

Before assembly, the 10X Illumina reads were preprocessed using LongRanger Basic v2.2.2 and the nanopore reads were filtered using FiltLong v0.2.0 with options '--min_length 1000 --min_mean_q 80'. The resulting set of filtered nanopore reads accounted for 50.8 Gb (about 34x coverage) with a mean Phred-scale read quality of 9.8 and a read length N50 of 31.4 Kb (Supplementary Table S3).

First, we assembled the filtered nanopore reads using Flye v2.8.3 with options '--nano-raw -i 2' (the latest performs two final rounds of polishing in the output assembly) obtaining a base assembly that comprised a total of 1.54 Gb, with an N50 of 1.36 Mb and a consensus quality (QV) of 27.19 (Fig. 1 and Supplementary Table S4).

Second, to maximize the QV of the Flye assembly, we tested several polishing tools and strategies (Supplementary Table S4). The best results were obtained using a combined strategy that exclusively used 10X Illumina reads to improve the QV. We used MaSuRCA^{21,22} version 4.0.4 using *createSuperReadsForDirectory.perl* to obtain high-quality SuperReads with a read N50 of 539 bp. The SuperReads were then aligned to the genome with Minimap2²³ v2.17 and the Illumina reads were aligned with BWA-MEM²⁴ v0.7.15. Finally, we ran NextPolish²⁵ v1.3.1 with these SuperReads, treating them as long reads, and the debarcoded 10X linked reads as paired-end Illumina for polishing the Flye assembly. In total, four polishing iterations were performed with NextPolish: two with the SuperReads and another two with the Illumina reads.

Third, to obtain a haploid reference (removing alternate haplotigs and other artificial duplications), we purged the assembly using purge_dups²⁶ v1.2.5 with cutoffs -l 5 -m 24 -u 78. This step removed 4,289 scaffolds accounting for 94,919,135 bp.

Fourth, the purged assembly was corrected and scaffolded with 10X Linked-Reads using the Faircloth's Lab pipeline (<http://protocols.faircloth-lab.org/en/latest/protocols-computer/assembly/assembly-scaffolding-with-arks-and-links.html#>). This pipeline includes Tigmint²⁷ v1.1.2, ARKS²⁸ v1.0.3 and LINKS²⁹ v1.8.5. Tigmint was used to identify and correct mis-assemblies using the linked reads in the purged assembly. The corrected assembly was then scaffolded using ARKS and LINKS, resulting in 2,783 scaffolds accounting for 1,460,390,273 bp with an N50 = 13.51 Mb (Supplementary Table S4).

Finally, the Omni-C reads were then mapped to the assembly using BWA-MEM and pre-processed using the Dovetail pipeline (https://omni-c.readthedocs.io/en/latest/fastq_to_bam.html). The filtering of the alignments was done with the default minimum mapping quality of 40. After the removal of PCR duplicates (46.34%, see Supplementary Table S5), 130,208,672 read pairs remained and were used as input to YaHS³⁰ v1.1 scaffolder with default parameters. We performed two rounds of assembly error correction and made 15 breaks, followed by ten rounds of scaffolding from higher to lower resolution (10 Mb down to 10 Kb), which produced an assembly with a span of 1,460,440,873 bp with an N50 of 89.54 Mb. The high contiguity achieved is also reflected in the scaffold L90 of 17, a value close to the number of chromosomes ($n = 19$) (Fig. 1 and Supplementary Table S4).

2.6. Manual curation

To guide manual curation of the assembly, we computed the mean Illumina coverage for all scaffolds (using BWA-MEM v0.7.15, SAMtools³¹ v1.9 and BlobTools³² v1.1) and the whole-genome alignments (WGAs) against the genome assemblies of two species belonging to family *Lacertidae*: a ZZ male of *P. muralis* (PodMur1.0) and a ZW female of *Lacerta agilis* (rLacAgi1.pri). The WGAs were produced with nucmer⁴³ and visualized with Dot (<https://github.com/MariaNattestad/dot>). This approach allowed us to 'scaffoldtype' (i.e. assign to a chromosome) the largest 20 superscaffolds (Supplementary Tables S6 and S7). Finally, the location of gaps (fasta-stats.py) and telomeres (<https://github.com/tolkitt/telomeric-identifier>), together with the Illumina coverage, were added to the contact map using PretextView (<https://github.com/wtsi-hpag/PretextView>). Manual curation was then performed using

PretextView (<https://github.com/wtsi-hpag/PretextView>). Given the high quality and contiguity of the YaHS assembly, the curation involved only 17 edits.

The Blobtoolkit³⁴ pipeline was then run on the curated assembly (Supplementary Fig. S1), using the NCBI nt database (updated on September 2022) and several BUSCO odb10 databases (sauropsida, vertebrata, metazoa, eukaryota, chlorophyta, fungi and bacteria). Removal of six contaminated scaffolds (based on GC cutoff of 0.3–0.65) (see Supplementary Table S8) resulted in the final assembly rPodLil1.2.

Finally, by comparing Illumina and ONT coverage estimates, we calculated the ratio of coverage for each sex chromosome with respect to the autosomal mean coverage (Supplementary Table S9). Whole genome alignments of the *P. lilfordi* (rPodLil1.2) against *P. muralis* (PodMur1.0) were performed with Minimap2 using the '-x asm5' option and visualized with the *pafr* R package.

2.7. Nuclear genome annotation

Gene annotation is very sensitive to both the tools and evidence-based data used. To produce an accurate gene annotation here we (i) sequenced and aligned transcriptome data to the *P. lilfordi* assembly, (ii) inferred protein content from comparative analyses against closely related annotated genomes and (iii) performed *ab initio* gene predictions (see flowchart in Supplementary Fig. S2).

As a first step, we masked all the repeat regions as they can include Open-Reading Frames (ORFs) and share certain domain homology with coding genes. For this purpose, we searched for repeats with RepeatMasker v4.1-2 (<http://www.repeatmasker.org>) using the custom RepBase repeat library available for *Podarcis*, along with a specific repeat library generated with RepeatModeler³⁵ v1.0.11 for our assembly. After excluding those repeats that were part of repetitive protein families (performing a BLAST search against Uniprot; last accessed March 2022), a final repeat annotation was produced using RepeatMasker. As this repeat annotation was produced mainly to aid in the genome annotation, low-complexity repeats were not annotated. The *P. muralis* genome was similarly processed for repeat annotation and comparative analysis.

Second, to obtain a comprehensive catalogue of the gene isoforms, we generated transcriptomic data from several tissues and multiple individuals using both regular Illumina RNA-seq and long-read PacBio Isoseq technologies (Supplementary Table S2). While regular Illumina provides an ample coverage of the full transcriptome potential, long-reads can cover an entire isoform with a single read, limiting the production of chimeric variants during transcript assembling. Short and long RNA reads were mapped to the genome assembly using STAR³⁶ v2.7.2a and Minimap2 v2.14 (with '-x splice:hq -uf' options), respectively. To infer gene models from the mappings, StringTie³⁷ v2.1.4 was run on each BAM file, setting the option '--conservative' only for the long-read libraries in order to disable the process of combining reads for transcript assembling, therefore retaining the full-length transcript sequences. Finally, all the models produced by Stringtie were combined using TACO³⁸ v0.6.3 and the resulting 'gtf' file was input into the Program to Assemble Spliced Alignments (PASA³⁹ v2.4.1) to produce PASA assemblies for annotation. High-quality junctions used during the annotation process were obtained with Portcullis⁴⁰

v1.2.0 after STAR and Minimap2 mapping. Additionally, the *TransDecoder* program (embedded in the PASA package) was run on the PASA assemblies to detect coding regions in the transcripts.

Third, we used proteomes from closely related species as a guide for predicting gene content. For this purpose, we downloaded the complete proteomes of *P. muralis*, *Pogona vitticeps* and *Pantherophis guttatus* from Uniprot in April 2022 and aligned them to the *P. lilfordi* repeat-masked assembly using Spaln⁴¹ v2.4.03.

Fourth, we performed *ab initio* gene predictions on the final assembly using three different programs: GeneID⁴² v1.4, Augustus⁴³ v3.3.4 and Genemark-ES⁴⁴ v2.3e, with and without incorporating junction evidence from the RNAseq data. The gene predictors were run with trained parameters on humans, except for Genemark, which runs in a self-trained mode.

Final genome annotation was achieved by combining all the data into consensus CDS models using EvidenceModeler-1.1.1³⁹ (EVM). Additionally, untranslated regions (UTRs) and alternative splicing forms were annotated via two rounds of PASA annotation updates (Supplementary Fig. S2).

Functional annotation was performed on the annotated proteins with Blast2GO.⁴⁵ First, a DIAMOND Blastp⁴⁶ search was made against the NCBI nr database (last accessed May 2022). Furthermore, InterProScan⁴⁷ was run to detect protein domains on the annotated proteins. All these data were combined by Blast2GO, which produced the final functional annotation.

General statistics on the genome and individual chromosomes were computed with in-house PERL scripts. The sex chromosome W presented 14 transposon-derived, short, only *ab-initio* single-copy genes. Due to the high repetitiveness of the chromosome, these genes were considered as artefacts and removed from the final annotation.

The annotation of non-coding RNAs (ncRNAs) was obtained as follows. First, the program CMsearch⁴⁸ v1.1 from the Infernal⁴⁹ package was run against the RFAM database of RNA families v12.0. Additionally, transfer RNA genes were identified by tRNAscan-SE⁵⁰ v2.08. Long non-coding RNAs (lncRNAs) were identified as those expressed transcripts (assembled by PASA) longer than 200 bp that were not included in the protein-coding annotation and not covered by a small ncRNA in more than 80% of their length. The resulting transcripts were clustered into genes (i.e. same gene assignment) using shared splice sites or significant sequence overlap.

2.8. Mitogenome assembly and annotation

To obtain the mitochondrial sequences, all ONT reads, previously filtered for whole-genome assembly with FiltLong v0.2.0 to be at least 1 Kb long and have a mean quality of 7, were mapped with Minimap2 against the *P. muralis* complete mitochondrial genome (NC_011607.1; 17,311 bp) with options: '-t \$THREADS -ax map-ont \$DATABASE \$READS'. We retained all reads with mapping quality = 12 (relatively unique) and at least 800 exact matches to the mitochondrial genome reference; these included 8,093 reads and a total of 51,692,448 bp (estimated mitochondrial coverage 2,986x).

All filtered ONT reads were assembled with Flye⁵¹ v2.9 using the options: 'flye --meta --scaffold -t 12 -i 2 -g 25k --nano-raw'. The '--meta' option is the most appropriate for uneven coverage samples and two polishing iterations were

run with the ONT reads on the final assembly with '-i 2'. Finally, the output assembly was screened for circular contigs, resulting in eight linear contigs and a single circular contig 17,112 bp long (contig_1).

As the reference mitogenome (*P. muralis*) is relatively distant (18–20 MYA), the identification of Illumina reads mapping outside the most conserved regions of the organelle is not straightforward. To overcome this issue, we mapped all the Illumina data (previously de-barcode PE 2 × 150 bp reads 10x linked-reads) to our complete Flye long-read assembly with gem-mapper, with ≤2% mismatches. Finally, a total of 326,105 read pairs were collected for further polishing (estimated coverage of the mitochondrial genome is 5,255.81x).

To further improve the sequence accuracy of the assembled mitochondrial genome, we performed two additional rounds of polishing on contig_1 with the selected Illumina reads using NextPolish v1.1.0 with Illumina PE 2 × 150 bp with options: '-paired -max_depth 5000'. The polished assembly was evaluated with Merqury v1.1 using 'k = 21' on the mitochondrial Illumina reads, *dnadiff* from MUMmer³³ package 4.0.0beta2 and fast-stats.py. The resulting circular chromosome was rotated and oriented according to the *P. muralis* reference, after detecting the appropriate Origin coordinates with the *dnadiff*.

The annotation of the mitogenome was performed using the MITOS⁵² Web Server (<http://mitos.bioinf.uni-leipzig.de/>). Manual curation was performed by checking the predicted sequences and comparing the annotation to that of the *P. muralis* mitogenome. As a result, one partial tRNA-Asp was removed from the annotation due to its absence in the *P. muralis* mitogenome annotation and to the high e-value reported by MITOS (e-value = 0.04071).

3. Results and discussion

3.1. Genome assembly

Genome sequencing yielded a total of 95 Gb of Illumina data (2 × 150 bp), 156.2 Gb of Omni-C data (2 × 150 bp) and 60 Gb of ONT data (Fig. 1). Assembly of the ONT data, followed by polishing, scaffolding, and manual curation resulted in a highly contiguous and complete assembly (rPodLil1.2) of 2,148 scaffolds accounting for 1.46 Gb, in line with the *C-value* and assembly span of *P. muralis* (1.51 Gb for PodMur1.0)¹³ (Fig. 2). It has a contig N50 of 1.48 Mb, scaffold N50 of 89.64 Mb (≥10 Mb) and QV of 40 (Supplementary Table S4), meeting the minimum quality requirement of 6.C.Q40 (megabase contig N50 and chromosomal-scale scaffold N50, with less than 1/10,000 error rate) established by the Earth Biogenome Project (EBP) for eukaryotic species with sufficient DNA and tissue.⁵³ Although the obtained contig N50 is shorter than most values found for other published chromosome-level *Lacertidae* assemblies (Supplementary Table S10), the high contig quality obtained and the use of Hi-C data allowed us to confidently assign most of the sequences (98.70%) to candidate chromosomes. In fact, the resulting assembly is consistent with the karyotype (2n = 38),⁵⁴ including 18 autosomes and two sex chromosomes (Fig. 2). Moreover, the assembly is highly complete, with 96.7% single copy complete genes, 87% k-mer completeness, and a low false duplication rate of 0.68% (Fig. 2).

Assignment to chromosomes was done via whole-genome alignments to the available genome assemblies of *P. muralis* (a

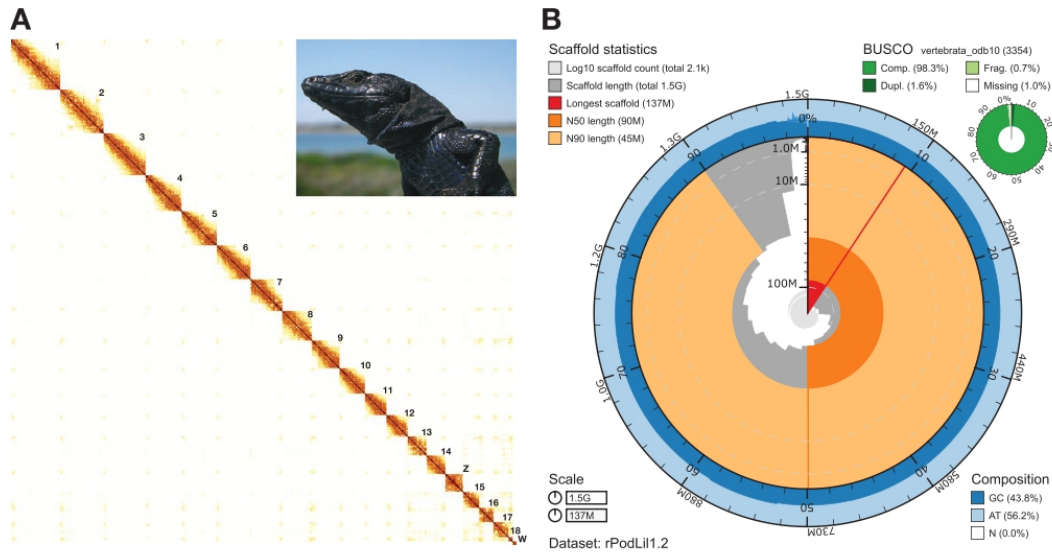


Figure 2. Visual summary of the rPodLil1.2 assembly. (A) Hi-C contact map of the genome assembly visualized in PretextView. The map shows 20 superscaffolds, ordered from longest to shortest, corresponding to the 18 autosomes and the two sex chromosomes (Z and W); the latter can be better visualized in [Supplementary Fig. S5](#). Additionally, there are a total of 44 short unlocalized scaffolds within the main chromosomes, and 2,084 short unplaced scaffolds (not visible at the resolution shown). (B) Snailplot summarising assembly metrics, including scaffold statistics, BUSCO completeness, total size and base composition. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 1,460,085,851 bp assembly. The distribution of record lengths is shown in dark grey with the plot radius scaled to the longest record present in the assembly (137 Mb, shown in red). Orange and pale-orange arcs show the N50 and N90 record lengths (90 Mb and 45 Mb), respectively. The pale grey spiral shows the cumulative record count on a log scale with white scale lines showing successive orders of magnitude. The blue and pale-blue area around the outside of the plot shows the distribution of GC, AT and N percentages in the same bins as the inner plot. A summary of complete, fragmented, duplicated and missing BUSCO genes in the vertebrata_odb10 set is shown in the top right.

male) and *L. agilis* (rLacAgi1.pri, a female). The alignments revealed a high level of collinearity to both assemblies for the autosomes as well as the Z chromosome (shown for *P. muralis* in [Supplementary Fig. S4](#)). A total of 21 inversions greater than 1 Mbp in length were observed with respect to the *P. muralis* assembly, only two of which were larger than 5 Mbp. There was no evidence of translocations or large deletions. Sex chromosome assignment was additionally supported by manual curation, showing the expected drop in sequencing coverage ([Supplementary Table S9](#)) and the typical pattern of Hi-C contacts ([Supplementary Fig. S5A](#)). Sex chromosomes in lacertids (ZZ and ZW) are known to differ in gene copy numbers, with males (ZZ) showing twice as many genes as females (ZW) due to W degeneration.⁵³ Moreover, the W is mostly heterochromatic and shows a typical high content in repetitive sequences,^{55,56} reducing the number of informative Hi-C read pairs for scaffolding in this chromosome. As expected, the Z chromosome of *P. lilfordi* was assembled into one superscaffold of 50.7 Mb in length that aligns well to both the Z chromosome of *P. muralis* and to that of *L. agilis* ([Supplementary Fig. S5B](#)), while the W chromosome was partially assembled into a 12.3 Mb superscaffold that aligns best to the superscaffold W of *L. agilis*, albeit with low identity ([Supplementary Fig. S5C](#)).

3.2. Genome annotation

While many genome assemblies have been recently released, their capability to answer certain biologically relevant questions is frequently limited by the lack of genome annotation.

Here, we provide a robust annotation for *P. lilfordi* by combining different approaches including newly generated transcriptome data, comparative proteomics, and *ab initio* gene prediction.

Overall, we annotated a total of 25,663 protein-coding genes that produce 43,578 transcripts (1.7 transcripts per gene) encoding 38,615 unique protein products ([Table 1](#)). We were able to assign functional labels to 72% (29,273) of the annotated proteins. The annotated transcripts contain 11 exons on average, with 91% of them being multi-exonic. In addition, we annotated 47,052 non-coding transcripts, including 12,785 lncRNAs and 34,267 sncRNAs ([Table 1](#)).

Running BUSCO on both the annotated protein-coding transcripts and protein sets showed a gene completeness of 97.9% and 97.1%, respectively, using the vertebrata_odb10 database. These results are in line with the genome BUSCO completeness of 98.3% and demonstrate the high accuracy of the genome annotation pipeline. Minor differences in the results obtained using the three sources of evidence are likely due to the algorithm that BUSCO uses, and the threshold established to consider a gene absent, fragmented or complete.

Comparison of *P. lilfordi* and *P. muralis* genome annotation statistics reveals only few differences ([Table 1](#)). Gene content is comparable, although we annotated ~3,000 extra protein-coding genes for *P. lilfordi*. Aligning these genes back to the PodMur1.0 genome assembly indicates that only 600 are unique to *P. lilfordi* (i.e. do not align to the *P. muralis* assembly). Most of the *P. lilfordi* proteins (96.5%) have a complete ORF (versus 83.7% in *P. muralis*) supporting our

high-quality assembly and annotation (Table 1). Moreover, the median UTR length is four times longer in *P. lilfordi* than in *P. muralis*, which is likely due to the usage of long-read transcript sequencing technologies (PacBio) for our annotation. Although most of the observed differences can be ascribed to the use of different evidence sources and annotation pipelines, additional genome resequencing data will be critical for validation.

3.3. Repetitive elements

Around 39% of the assembled genome was repetitive: 6% was annotated as repeats using the ‘podarcis’ Repbase library and an additional 33% was classified as repeats using the repeat library obtained after running Repeat Modeler (<https://www.repeatmasker.org/RepeatModeler/>) against the assembly (Table 2). These repeats were represented by transposable elements, including short, interspersed elements (SINEs), long interspersed nuclear elements (LINEs), long terminal repeat retrotransposons (LTRs) and DNA transposons. The repeat content landscape of *P. lilfordi* was highly comparable to that of *P. muralis* (38–39%), in line with their similar genome size⁵⁷: most of the repeats were classified as LINEs (~12%), followed by DNA transposons (~7%) and SINEs (4–5%) (Table 2). Percentage of LTR was slightly higher in *P. lilfordi* than *P. muralis* (2.37% against 1.36%). At present, we cannot confidently assess

whether these minor differences between species are real or a methodological bias (methods can particularly affect repeat-resolution).

3.4. Mitogenome

The resulting mitogenome assembly is a single circular contig 17,251 bp long with a QV of 44.12. Its GC content is 38.73%, very similar to the 38.55% observed in the *P. muralis* mitogenome. As expected, the average identity between both mitogenomes is 88.74%. The origin of replication was identified and used to linearize the circular chromosome by comparison to the *P. muralis* mitogenome. After annotation, it was confirmed that this was the origin of the tRNA-Phe (MT-TF) that is used as a standard origin in vertebrate mitogenomes.⁵⁸ The final annotation of the mitochondrial chromosome contained 13 protein-coding genes, 2 rRNAs and 22 tRNAs, in line with expectations for vertebrates.

4. Conclusions

Here we announce and publicly release the *Podarcis lilfordi* reference nuclear and mitochondrial genome assembly and annotation. This is the first chromosome-level reference for an endemic reptile species released within the framework of the Catalan Initiative for the Earth Biogenome Project (CBP). The *P. lilfordi* genome is the second complete genome within the

Table 1. Final genome annotation statistics for the *P. lilfordi* assembly (rPodLil1.2) and comparison to *P. muralis* (PodMur1.0)

	rPodLil1.2	PodMur1.0
Genome size (bp)	1,460,085,851	1,511,002,858
Number of protein-coding genes	25,663	22,062
Median gene length (bp)	13,456	11,808
Number of transcripts	43,578	37,240
Number of proteins	38,615	36,445
Proteins with complete ORFs	37,249	30,503
Functionally annotated proteins	29,273	25,768
Number of coding exons	232,494	219,498
Median UTR length (bp)	2,138	505
Median intron length (bp)	1,297	1,245
Exons/transcript	10.90	12.12
Transcripts/gene	1.70	1.69
Multi-exonic transcripts (proportion)	0.91	0.92
Gene density (gene/Mb)	17.58	14.60

Table 2. Repeat content of the *P. lilfordi* assembly and comparison with *P. muralis*

	rPodLil1.2 (bp)	PodMur1.0 (bp)
SINEs	58,508,447 (4.0%)	71,879,703 (4.7%)
LINEs	186,782,953 (12.8%)	182,775,149 (12.1%)
LTR elements	34,584,027 (2.4%)	20,566,821 (1.4%)
DNA transposons	105,432,906 (7.2%)	110,903,635 (7.3%)
Unclassified	180,164,025 (12.3%)	181,606,330 (12.0%)
Total interspersed repeats	565,472,358 (38.7%)	567,731,638 (37.6%)

Percentages are estimated with respect to the total genome size

highly diverse genus *Podarcis*, along with that of *P. muralis*. Despite their evolutionary divergence (18–20 MYA¹⁴), the two species exhibit substantial conservation in genome organisation and overall annotation. Given its high quality, contiguity and annotation, the *P. lilfordi* genome sequence will represent a valuable resource for evolutionary and conservation genomics studies. The resource will facilitate comparative genomics of Lacertidae and reptiles in general, and aid in the understanding of the genetic bases of vertebrate insular adaptation (i.e. the island syndrome) and demographic resilience. Additionally, the genome will represent a critical reference to explore the genetic diversity of this endemic species, its adaptive potential and evidence for local adaptation, along with its ability to respond to current threats by human pressure in the Balearic Islands. Finally, we expect that future genome analyses will have a critical impact on conservation management and policy decisions on this endangered species.

Acknowledgements

We would like to thank the following people and institutions: The Catalan Initiative for the Earth Biogenome Project (CBP) for coordination and promotion efforts of this genome project; Yumi Sims, Jo Wood and Alan Tracey (Wellcome Sanger Institute) for their help and advice on assembly curation; Marc Palmada-Flores (IBE-UPF) for his feedback on assembly and advice about available lizard genomes; Chenxi Zhou (Wellcome Sanger Institute) for his help on producing contact maps in different formats directly from the YaHS output; Emiliano Trucchi (Marche Polytechnic University, Italy) and Giorgio Bertorelle (University of Ferrara, Italy) for access to the draft *Podarcis raffonei* chromosome W sequence (unpublished at the time of this study). Finally, we are also grateful to Dovetail Genomics (Cantata Bio) for their support with the quality control and pre-processing of Omni-C data before scaffolding.

Conflict of interest

None declared.

Funding

This study was supported by the Institut d'Estudis Catalans under the Catalan Initiative for the Earth Biogenome Project (PRO2020-S02 to L.B.), the Swedish Research Council (VR 2017-03846 and VR-2021-04656 to T.U. and VR-2020-03650 to N.F.) and Starting Grant from the European Research Council (no. 948126 to N.F.). We also acknowledge the support of the Spanish Ministry of Science and Innovation to the EMBL partnership, the Centro de Excelencia Severo Ochoa, the CERCA Programme/Generalitat de Catalunya, the Spanish Ministry of Science and Innovation through the Instituto de Salud Carlos III and the Generalitat de Catalunya through Departament de Salut and Departament d'Empresa i Coneixement. Co-financing funds were obtained from the European Regional Development Fund by the Spanish Ministry of Science and Innovation corresponding to the Programa Operativo FEDER Plurirregional de España (POPE) 2014-2020 and by the Secretaria d'Universitats i Recerca, Departament d'Empresa i Coneixement of the Generalitat de Catalunya corresponding to the Programa Operatiu FEDER de Catalunya 2014-2020.

Author contributions

L.B., G.T., N.F. and T.U. conceived and advised the project. L.B. and K.E.O.A. performed sample collection and preparation. M.G., I.S.E. and N.F. performed data production. J.G.G., F.C. and T.S.A. performed data processing and genome assembly. J.G.G. and T.S.A. performed data processing and genome annotation. All authors wrote and approved the final manuscript.

Data availability

All the sequencing data generated in this study are available at the European Nucleotide Archive (ENA) (<https://www.ebi.ac.uk/ena/browser/home>) under the accession number PRJEB50294. The assembly and annotation can also be found at ENA under the accession number GCA_947686815.1.

In addition, a genome browser, a BLAST Sequence Server and annotation data can be found at <https://denovo.cnag.cat/podarcis>. A Snakemake pipeline for genome assembly is available at https://github.com/cnag-aat/assembly_pipeline.

Supplementary Data

Supplementary data are available at DNARES online.

Table S1. Sampled specimen information

Table S2. RNA samples and quality

Table S3. ONT read statistics

Table S4. Assembly statistics, including intermediate steps

Table S5. Omni-C mapping statistics

Table S6. Scaffotyping based on Whole Genome Alignments (Yash vs PodMur1.0)

Table S7. Scaffotyping based on Whole Genome Alignments (Yash vs rLacAgi1)

Table S8. Contaminated scaffolds

Table S9. Ratio of sex chromosome coverage with respect to the autosomes mean coverage (based on ONT and Illumina reads)

Table S10. Major assembly statistics of publicly available Lacertidae genomes

Figure S1. Hexagon-binned blob plot of base coverage of filtered ONT reads against GC proportion for scaffolds in assembly rPodLil1.1. Scaffolds are coloured by phylum (according to Blast hits against the nt database; last accessed September 2022) and binned at a resolution of 30 divisions on each axis. Coloured hexagons within each bin are sized in proportion to the sum of individual scaffold lengths on a square-root scale, ranging from 1,018 to 693,200,870. Histograms show the distribution of scaffold length sum along each axis.

Figure S2. Annotation flowchart based on the non-decontaminated assembly (rPodLil1.1).

Figure S3. K-mer comparison between the Illumina reads and the rPodLil1.2 assembly. Stacked histogram of k-mer distributions obtained by comparing the assembly with Merquy v1.1 using $k = 21$ on the 10X Illumina reads. Artificial duplications corresponding to duplicate k-mers are shown in blue above the main peak ($\sim 40\times$). They only account for 0.68% of the k-mers.

Figure S4. Whole genome alignment of *P. lilfordi* to *P. muralis*. Chromosomal sequences, named according to corresponding chromosomal sequences in *P. muralis*, are ordered from largest to smallest in *P. lilfordi* and oriented with respect to *P. muralis*, which inverts the order of chromosomes 16 and 17. Alignments longer than 100 kb were selected and visualized with the *pafr* R library.

Figure S5. Sex chromosome assignment. (A) Hi-C contact map showing scaffolds corresponding to the sexual chromosomes. Illumina coverage is plotted in pink. (B) Alignment of the scaffold assigned to the Z chromosome in *L. agilis* against the corresponding scaffold in *P. lilfordi*. (C) Alignment of the scaffold assigned to the W chromosome in *L. agilis* against the corresponding scaffold in *P. lilfordi*.

References

- Pérez-Cembranos, A., Pérez-Mellado, V., Alemany, I., et al. 2020, Morphological and genetic diversity of the Balearic lizard, *Podarcis lilfordi* (Günther, 1874): is it relevant to its conservation?, *Divers. Distrib.*, **26**, 1122–41.
- Castilla, A.M. and Bauwens, D. 2000, Reproductive characteristics of the Island lacertid lizard *Podarcis lilfordi*, *J. Herpetol.*, **34**, 390–6.
- Rotger, A., Igual, J. M., and Tavecchia, G. 2020, Contrasting size-dependent life history strategies of an insular lizard. *Curr. Zool.*, **66**, 625–33.
- Pérez-Mellado, V., Hernández-Estévez, J.A., García-Díez, T., et al. 2008, Population density in *Podarcis lilfordi* (Squamata, Lacertidae), a lizard species endemic to small islets in the Balearic Islands (Spain), *Amphib-Reptilia*, **29**, 49–60.
- Terrasa, B., Pérez-Mellado, V., Brown, R.P., Picornell, A., Castro, J.A., and Ramon, M.M. 2009, Foundations for conservation of intraspecific genetic diversity revealed by analysis of phylogeographical structure in the endangered endemic lizard *Podarcis lilfordi*, *Divers. Distrib.*, **15**, 207–21.
- Rotger, A., Tenan, S., Igual, J.-M., Bonner, S., and Tavecchia, G. 2023, Life span, growth, senescence and island syndrome: accounting for imperfect detection and continuous growth, *J. Anim. Ecol.*, **92**, 183–94.
- Alemany, I., Pérez-Cembranos, A., Pérez-Mellado, V., et al. 2022, DNA metabarcoding the diet of *Podarcis* lizards endemic to the Balearic Islands Fuller, R., (ed.), *Curr. Zool.*, zoac073. doi:10.1093/cz/zoac073
- Rotger, A., Igual, J.M., Genovart, M., et al. 2021, Contrasting adult body-size in sister populations of the balearic lizard, *Podarcis lilfordi* (Günther 1874) suggests anthropogenic selective pressures, *Herpetol. Monogr.*, **35**, 53–64.
- Bassitta, M., Brown, R.P., Pérez-Cembranos, A., et al. 2021, Genomic signatures of drift and selection driven by predation and human pressure in an insular lizard, *Sci. Rep.*, **11**.
- Alemany, I., Pérez-Cembranos, A., Pérez-Mellado, V., et al. 2022, Faecal microbiota divergence in allopatric populations of *Podarcis lilfordi* and *P. pityusensis*, two lizard species endemic to the Balearic Islands, *Microb. Ecol.* doi:10.1007/s00248-022-02019-3
- Baldo, L., Riera, J.L., Mitsi, K., and Pretus, J.L. 2018, Processes shaping gut microbiota diversity in allopatric populations of the endemic lizard *Podarcis lilfordi* from Menorcan islets (Balearic Islands), *FEMS Microbiol. Ecol.*, **94**, 1–14.
- Baldo, L., Tavecchia, G., Rotger Vallespir, A., Igual, J., and Riera, J. 2022, Insular holobionts: the role of phylogeography and seasonal fluctuations in shaping the gut microbiotas of the Balearic Wall lizard *Podarcis lilfordi*, *Anim. Ecol.*
- Andrade, P., Pinho, C., Pérez i de Lanuza, G., et al. 2019, Regulatory changes in pterin and carotenoid genes underlie balanced color polymorphisms in the wall lizard, *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 5633–42.
- Yang, W., Feiner, N., Pinho, C., et al. 2021, Extensive introgression and mosaic genomes of Mediterranean endemic lizards, *Nat. Commun.*, **12**, 2762.
- Yang, W., Feiner, N., Salvi, D., et al. 2022, Population genomics of wall lizards reflects the dynamic history of the Mediterranean basin, *Mol. Biol. Evol.*, **39**, msab311.
- Salvador, A. 2009, Lagartija balear – *Podarcis lilfordi*. En: Enciclopedia Virtual de los Vertebrados Españoles. Salvador, A., Marco, A. (Eds.). Museo Nacional de Ciencias Naturales, Madrid.
- Mölder, F., Jablonski, K.P., Letcher, B., et al. 2021, Sustainable data analysis with Snakemake, *F1000Res*, **10**, 33.
- Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A., and Zdobnov, E.M. 2021, BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes, *Mol. Biol. Evol.*, **38**, 4647–54.
- Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. 2015, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs, *Bioinformatics*, **31**, 3210–2.
- Rhie, A., Walenz, B.P., Koren, S., and Phillippy, A.M. 2020, Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies, *Genome Biol.*, **21**, 245.
- Zimin, A.V., Puiu, D., Luo, M.-C., et al. 2017, Hybrid assembly of the large and highly repetitive genome of *Aegilops tauschii*, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm, *Genome Res.*, **27**, 787–92.
- Zimin, A.V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S.L., and Yorke, J.A. 2013, The MaSuRCA genome assembler, *Bioinformatics*, **29**, 2669–77.
- Li, H. 2018, Minimap2: pairwise alignment for nucleotide sequences, *Bioinformatics*, **34**, 3094–100.
- Li, H. 2013, May 26, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*.
- Hu, J., Fan, J., Sun, Z., and Liu, S. 2020, NextPolish: a fast and efficient genome polishing tool for long-read assembly, *Bioinformatics*, **36**, 2253–5.
- Guan, D., McCarthy, S.A., Wood, J., Howe, K., Wang, Y., and Durbin, R. 2020, Identifying and removing haplotypic duplication in primary genome assemblies, *Bioinformatics*, **36**, 2896–8.
- Jackman, S.D., Coombe, L., Chu, J., et al. 2018, Tigrint: correcting assembly errors using linked reads from large molecules, *BMC Bioinf.*, **19**, 393.
- Coombe, L., Zhang, J., Vandervalk, B.P., et al. 2018, ARKS: chromosome-scale scaffolding of human genome drafts with linked read kmers, *BMC Bioinf.*, **19**, 234.
- Warren, R.L., Vandervalk, B.P., Jones, S.J., and Birol, I. 2015, LINKS: scaffolding genome assemblies with kilobase-long nanopore reads, *bioRxiv*.
- Zhou, C., McCarthy, S. A., and Durbin, R. 2022, YaHS: yet another Hi-C scaffolding tool. *bioRxiv*, 2022.06.09.495093.
- Li, H., Handsaker, B., Wysoker, A., et al.; 1000 Genome Project Data Processing Subgroup. 2009, The sequence alignment/map format and SAMtools, *Bioinformatics*, **25**, 2078–9.
- Laetsch, D.R. and Blaxter, M.L. 2017, BlobTools: interrogation of genome assemblies [version 1; peer review: 2 approved with reservations], *F1000Research*, **6**, 1287.
- Marçais, G., Delcher, A.L., Phillippy, A.M., Coston, R., Salzberg, S.L., and Zimin, A. 2018, MUMmer4: a fast and versatile genome alignment system, *PLoS Comput. Biol.*, **14**, e1005944.
- Challis, R., Richards, E., Rajan, J., Cochrane, G., and Blaxter, M. 2020, BlobToolKit – interactive quality assessment of genome assemblies, *G3 (Bethesda)*, **10**, 1361–74.
- Flynn, J.M., Hubley, R., Goubert, C., et al. 2020, RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci U S A*, **117**, 9451–7.
- Dobin, A., Davis, C.A., Schlesinger, F., et al. 2013, STAR: ultrafast universal RNA-seq aligner, *Bioinformatics*, **29**, 15–21.
- Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.-C., Mendell, J.T., and Salzberg, S.L. 2015, StringTie enables improved reconstruction of a transcriptome from RNA-seq reads, *Nat. Biotechnol.*, **33**, 290–5.
- Niknafs, Y.S., Pandian, B., Iyer, H.K., Chinnaiyan, A.M., and Iyer, M.K. 2017, TACO produces robust multisample transcriptome assemblies from RNA-seq, *Nat. Methods*, **14**, 68–70.
- Haas, B.J., Salzberg, S.L., Zhu, W., et al. 2008, Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments, *Genome Biol.*, **9**, R7.

40. Mapleson, D., Venturini, L., Kaithakottil, G., and Swarbreck, D. 2018, Efficient and accurate detection of splice junctions from RNA-seq with Portcullis, *GigaScience*, 7.
41. Iwata, H. and Gotoh, O. 2012, Benchmarking spliced alignment programs including Spaln2, an extended version of Spaln that incorporates additional species-specific features, *Nucleic Acids Res.*, 40, e161.
42. Alioto, T., Blanco, E., Parra, G., and Guigó, R. 2018, Using geneid to Identify Genes, *Curr Protoc Bioinformatics*, 64, e56.
43. Stanke, M., Schöffmann, O., Morgenstern, B., and Waack, S. 2006, Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources, *BMC Bioinf.*, 7, 62.
44. Lomsadze, A., Burns, P.D., and Borodovsky, M. 2014, Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Res.*, 42, e119.
45. Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., and Robles, M. 2005, Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research, *Bioinformatics*, 21, 3674–6.
46. Buchfink, B., Reuter, K., and Drost, H.-G. 2021, Sensitive protein alignments at tree-of-life scale using DIAMOND, *Nat. Methods*, 18, 366–8.
47. Jones, P., Binns, D., Chang, H.-Y., et al. 2014, InterProScan 5: genome-scale protein function classification, *Bioinformatics*, 30, 1236–40.
48. Cui, X., Lu, Z., Wang, S., Jing-Yan Wang, J., and Gao, X. 2016, CMsearch: simultaneous exploration of protein sequence space and structure space improves not only protein homology detection but also protein structure prediction, *Bioinformatics*, 32, i332–40.
49. Nawrocki, E.P. and Eddy, S.R. 2013, Infernal 1.1: 100-fold faster RNA homology searches, *Bioinformatics*, 29, 2933–5.
50. Chan, P.P. and Lowe, T.M. 2019, tRNAscan-SE: searching for tRNA genes in genomic sequences, *Methods Mol. Biol.*, 1962, 1–14.
51. Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P.A. 2019, Assembly of long, error-prone reads using repeat graphs, *Nat. Biotechnol.*, 37, 540–6.
52. Bernt, M., Donath, A., Jühling, F., et al. 2013, MITOS: improved de novo metazoan mitochondrial genome annotation, *Mol. Phylogenet. Evol.*, 69, 313–9.
53. Rovatsos, M., Vukić, J., Mrugała, A., Suwala, G., Lymberakis, P., and Kratochvíl, L. 2019, Little evidence for switches to environmental sex determination and turnover of sex chromosomes in lacertid lizards, *Sci. Rep.*, 9, 7832.
54. Vujošević, M. and Blagojević, J. 1999, The distribution of constitutive heterochromatin and nucleolus organizers in lizards of the family Lacertidae (Sauria), *Genetika*, 31, 269–76.
55. Suwala, G., Altmanová, M., Mazzoleni, S., et al. 2020, Evolutionary variability of W-linked repetitive content in lacertid lizards, *Genes*, 11, 531.
56. Olmo, E., Odierna, G., and Capriglione, T. 1987, Evolution of sex-chromosomes in lacertid lizards, *Chromosoma*, 96, 33–8.
57. Pasquesi, G.I.M., Adams, R.H., Card, D.C., et al. 2018, Squamate reptiles challenge paradigms of genomic repeat element evolution set by birds and mammals, *Nat. Commun.*, 9, 2774.
58. Formenti, G., Rhie, A., Balacco, J., et al.; Vertebrate Genomes Project Consortium. 2021, Complete vertebrate mitogenomes reveal widespread repeats and gene duplications, *Genome Biol.*, 22, 120.

Population genetics and phylogeographic history of the insular lizard *Podarcis lilfordi* (Gunther, 1874) from the Balearic Islands based on genome-wide polymorphic data

Katherin Otalora^{1,2} | Joan Lluís Riera¹ | Giacomo Tavecchia³ | Andreu Rotger³ | José Manuel Igual³ | Jean-Remi Paul Trotta⁴ | Laura Baldo^{1,5} 

¹Department of Evolutionary Biology, Ecology and Environmental Sciences, Biology Faculty, University of Barcelona (UB), Barcelona, Spain

²Fundación FUNMAJO, EBA, RAIEC, Biodiversity Branch, Tunja, Boyacá, Colombia

³Animal Demography and Ecology Unit (GEDA-IMEDEA, CSIC-UIB), Esporles, Spain

⁴CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain

⁵Institute for Research on Biodiversity (IRBio), University of Barcelona (UB), Barcelona, Spain

Correspondence

Laura Baldo, Department of Evolutionary Biology, Ecology and Environmental Sciences, Biology Faculty, University of Barcelona (UB), Av. Diagonal 645, Barcelona 08028, Spain.
Email: baldo.laura@ub.edu

Funding information

Consejería d'Educació i Universitat - Govern de les Illes Balears (GOIB), Grant/Award Number: PRD2018/25; Ministerio de Ciencia, Tecnología e Innovación de Colombia, Colciencias, Grant/Award Number: MINCIENCIAS885/2020; Institut d'Estudis Catalans under the Catalan Biogenome Project Initiative, Grant/Award Number: PRO2020-S02; Ministerio de Ciencia, Innovación y Universidades (MICIU)/Agencia Estatal de Investigación (AEI)/10.13039/501100011033/FEDER, UE, Grant/Award Number: PID2022-141578NB-C22

Abstract

Islands provide a great system to explore the processes that maintain genetic diversity and promote local adaptation. We explored the genomic diversity of the Balearic lizard *Podarcis lilfordi*, an endemic species characterized by numerous small insular populations with large phenotypic diversity. Using the newly available genome for this species, we characterized more than 300,000 SNPs, merging genotyping-by-sequencing (GBS) data with previously published restriction site-associated DNA sequencing (RAD-Seq) data, providing a dataset of 16 island populations (191 individuals) across the range of species distribution (Menorca, Mallorca, and Cabrera). Results indicate that each islet hosts a well-differentiated population ($F_{ST}=0.247 \pm 0.09$), with no recent immigration/translocation events. Contrary to expectations, most populations harbor a considerable genetic diversity (mean nucleotide diversity, $P_i=0.144 \pm 0.021$), characterized by overall low inbreeding values ($F_{IS} < 0.1$). While the genetic diversity significantly decreased with decreasing islet surface, maintenance of substantial genetic diversity even in tiny islets suggests variable selection or other mechanisms that buffer genetic drift. Maximum-likelihood tree based on concatenated SNP data confirmed the existence of the two major independent lineages of Menorca and Mallorca/Cabrera. Multiple lines of evidence, including admixture and root testing, robustly placed the origin of the species in the Mallorca Island, rather than in Menorca. Outlier analysis mainly retrieved a strong signature of genome differentiation between the two major archipelagos, especially in the sexual chromosome Z. A set of proteins were target of multiple outliers and primarily associated with binding and catalytic activity, providing interesting candidates for future selection studies. This study provides the framework to explore crucial aspects of the genetic basis of phenotypic divergence and insular adaptation.

KEYWORDS

genotyping-by-sequencing, island, outlier, restriction site-associated DNA sequencing

TAXONOMY CLASSIFICATION

Genomics

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Ecology and Evolution* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Island populations represent biological scenarios of greatest evolutionary interest due to their geographic isolation, which accelerates the process of organismal diversification (Warren et al., 2015; Whittaker et al., 2017), and their restricted scale of study, that allow close monitoring (Drake et al., 2002; Losos & Ricklefs, 2009). They also typically host a high rate of endemism, making them a global conservation priority (Manes et al., 2021; Sivaperuman et al., 2008). The Balearic Islands, encompassing the Pityusic (Ibiza and Formentera) and Gimnesic (Mallorca, Menorca, Cabrera, and associated islets) islands, comprise a large number of islands that originated following a complex geological and climatic history related to the Mediterranean sea level variations during the Pleistocene (0.2-Mya) (Cuerda, 1989; Goy et al., 1997). Among the several endemic species, they host the two species of Balearic lizard (i.e., *Podarcis lilfordi* from the Gimnesic islands and *P. pityusensis* from the Pityusic islands). They are among the most extensively studied vertebrates of the archipelago, due to their remarkable phenotypic diversity (Pérez-Cembranos et al., 2020), high population densities, and reduction of antipredator behavior associated to the "island syndrome" (Cooper & Pérez-Mellado, 2012; Hawlena et al., 2009; Novosolov et al., 2013; Rotger et al., 2023).

Podarcis lilfordi (Günther, A., 1874) presently occurs on 43 off-the-coast islets of Mallorca and Menorca, as well as in the Cabrera archipelago. Current distribution is the result of a vicariance process during the Mediterranean Sea desiccation and subsequent reflooding (5.33Mya) (Krijgsman et al., 1999; Terrasa et al., 2009). The disappearance of *P. lilfordi* from the main islands of Menorca and Mallorca dates back more than 2000 years, following the introduction of predators by humans (Alcover, 2000). The species is now confined on islets that provide an effective refuge for to date populations. Within these islets, the species evolved in a context of limited resources and absence of major predators and competitors, greatly diversifying in morphology, pigmentation (Pérez-Cembranos et al., 2020; Rotger et al., 2021), behavior, and life history traits (Rotger et al., 2023; Salvador, 2015). Major forces driving this diversification can be ascribed to the high intrapopulation competition (density-dependent factors) (Grant & Benton, 2000; Le Galliard et al., 2010; Massot et al., 1992), as well as intense genetic drift resulting from recurrent bottlenecks, primarily affecting the smallest islets (Bassitta et al., 2021; Charlesworth et al., 2003; Rotger et al., 2021; Terrasa et al., 2009). This context of rapid phenotypic differentiation among insular populations makes *P. lilfordi* an interesting vertebrate model for ecological and evolutionary studies (Camargo et al., 2010), particularly those addressing the genomic basis of local adaptation and persistence of small populations (Bassitta et al., 2021; Yang et al., 2022).

Genetic analyses of this species have been long limited to mtDNA, few nuclear genes (Brown et al., 2008; Terrasa et al., 2009) and microsatellites for few populations (Bloor et al., 2011; Rotger et al., 2021). A recent study based on single-nucleotide polymorphisms (SNPs) data by restriction site-associated DNA sequencing

(RAD-Seq) (Bassitta et al., 2021) has expanded on previous knowledge on genetic diversity and phylogeographic pattern of this species, confirming the existence of two well-discriminated genetic lineages separating populations from the major archipelagos of Menorca and Mallorca/Cabrera (Brown et al., 2008; Terrasa et al., 2009). Recently, the *P. lilfordi* genome has been sequenced (Gomez-Garrido et al., 2023) providing the framework for a comprehensive exploration of population-level genomics of this species.

Here, we used for the first time the newly available reference genome of *P. lilfordi* to map and annotate newly generated genome-wide polymorphic markers obtained by two independent sequencing methods, genotyping-by-sequencing (GBS) and RAD-Seq. We specifically characterized SNPs of eight new populations (100 individuals) by GBS, and integrated RAD-Seq data of additional 10 populations (91 individuals from Bassitta et al. (2021)), for a comprehensive analysis of 16 populations (2 were common to both studies), spanning the main range of distribution of this species. Individual GBS and RAD-Seq datasets were independently analyzed as well as merged to obtain a representative set of common SNPs loci across all populations, which allowed for a robust comparative genomics within *P. lilfordi*.

Our main objectives were to (a) assess the level of intraspecific genetic diversity and structuring of the different populations of *P. lilfordi*; (b) reconstruct the species' main phylogeographic scenario of colonization of the Balearic Islands; and (c) identify potential signatures of genome diversification and loci underpinning the phenotypic diversity and insular adaptation of this species.

2 | MATERIALS AND METHODS

2.1 | Sampling and DNA extractions

Tails tissues of 100 specimens of *P. lilfordi* were collected between 2015 and 2016 from eight islets in the Menorca and Mallorca archipelagos (Balearic Islands) (10–22 samples per islet) (Figure 1 and Table 1). Islets encompassed a representative subsample of the current species' geographic distribution and vary in several geographic and ecological aspects, including surface area, maximum altitude, vegetation cover and type, presence of human settlements, and other vertebrates, as well as lizard demographic traits (Mayol et al., 2020; Pérez-Mellado et al., 2008; Rotger et al., 2021). Specimens were captured using pitfall traps placed along paths and vegetation edges, sexed according to visual examination and morphology (Rotger et al., 2016), weighed, and body size measured as snout to vent length (SVL) (see Table S1 for specimen information). Tails were preserved in 100% ethanol and kept at -80°C until processed. Permits for sampling were provided by Conselleria d'Agricultura i Medi Ambient i Territori, Govern de les Illes Balears (CEP 31/2015 to LB and CEP 6/14 to GT). Genomic DNA was extracted with DNAeasy Tissue and Blood kit (Qiagen) with RNase treatment. DNA concentration was measured with Qubit 2 fluorometer (Invitrogen, Waltham). About 1 μg

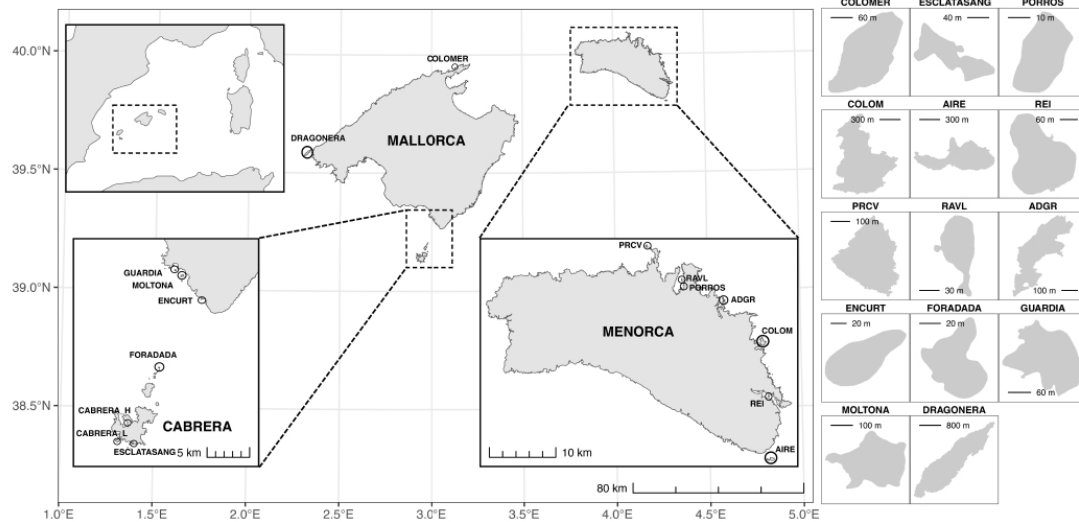


FIGURE 1 Geographic map of the 16 localities (15 islands) under study, sampled in the archipelagos of Mallorca, Cabrera, and Menorca.

of genomic DNA per sample was sent to *Centro Nacional de Analisis Genomico* (CNAG, Spain) for sequencing. DNA integrity and concentration was further measured on a fragment analyzer (Agilent Bioanalyzer) and high-quality samples with at least 10 ng/μL were selected for sequencing.

2.2 | Sequencing

A restriction site-associated DNA library was generated by GBS method (Elshire et al., 2011). To select for the best enzyme with reproducible genomic fragments across samples (a critical criterion for GBS experiments), we first generated a pilot library (four samples) with Illumina MiSeq (150 bp), using two common restriction enzymes, PstI and ApeKI. PstI provided the best reproducibility and number of variants and was therefore chosen to generate the final library.

Genomic DNA was digested with the PstI, fragments were tagged with individual barcodes, PCR-amplified, multiplexed, and sequenced on dual lanes on an Illumina HiSeq 2500 platform (2 × 100 bp). Raw read sequences were demultiplexed and quality checked using FastQC v0.10.1 (Wingett & Andrews, 2018; available at <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).

Additionally, raw RAD-Seq genomic data of *P. lilfordi* generated on a previous study (Bassitta et al., 2021) were downloaded from the online repository (PRJNA645796, 91 samples from 10 populations). The GBS dataset (this study) and the RAD-Seq dataset (Bassitta et al., 2021) shared the restriction enzyme PstI, which allowed data integration for comparative purposes (see below). Two localities of Colom and Aire were sampled and sequenced in both studies (different individuals) (Table 1 and Table S1).

2.3 | SNPs calling and filtering

Raw sequence data from GBS and RAD-Seq were independently processed as well as merged (COMBINED dataset). Thereafter, the three generated datasets (GBS, RAD-Seq and COMBINED) were processed through the same pipeline.

All barcoded raw reads were mapped to the *P. lilfordi* genome (rPodLil1.1, available at <https://denovo.cnag.cat/podarcis>) (Gomez-Garrido et al., 2023) using the Burrows-Wheeler Aligner (BWA) software (v2.1) to generate individual Binary Alignment Map (BAM) files. The genome corresponds to a single female individual sampled in the Aire population (the population was included in our dataset). Variant calling per sample was performed with Genome Analysis Toolkit HaplotypeCaller from GATK v. 3.7 (Van der Auwerera & O'Connor, 2020) to produce gVCF files (minimum-mapping-quality 20). gVCF files were finally genotyped for variant calling (i.e., SNPs).

Prior to data filtering, SNP and sample statistics were explored using VCFTools v. 0.1.16 (Danecek et al., 2011; De la Cruz & Raska, 2014). SNPs were then filtered using the following settings: --max-alleles 2, --remove-indels, --min-meanDP 10, --max-meanDP 50, --minQ 30, --maf 0.05, --max-missing 0.75 (individual datasets) or 0.90 (combined).

To retrieve common loci between GBS and RAD-Seq datasets, the max-missing parameter was set to 0.9, thus retaining loci shared by at least 90% of the specimens and effectively removing all loci unique to either GBS or RAD-Seq.

2.4 | Population genetic diversity

We used the *populations* module in STACKS v.2.61 (Catchen et al., 2013) to calculate diversity statistics per population, including the observed

TABLE 1 Sampling design and environmental variables.

Major archipelagos	Archipelago	Type	Official name ^a (locality)	No. of individuals				Geographic variables ^a					Lizard density ^c (ind/ha)
				Short name	GBS	RAD-Seq ^b	Total	X-coord	Y-coord	Surface area (ha)	Maximum altitude (msnm)	SDI	
Menorca	Menorca	Islet	Illa Gran d'Addaia	ADGR	10		10	603284.8	4430301.1	7.5	21.7	2.251	1062
Menorca	Menorca	Islet	Illa de l'Aire	AIRE	11	12	23	610443.5	4406455.4	30.5	16.1	1.946	4098
Mallorca/Cabrera	Cabrera	Inland	Cabrera (Harbour)	CABRERA_H		12	12	494459.1	4333606.0	1142.0	171.4	3.524	45059.0
Mallorca/Cabrera	Cabrera	Inland	Cabrera (Lighthouse)	CABRERA_L		8	8	493279.9	4331172.0	1142.0	171.4	3.524	45059.0
Menorca	Menorca	Islet	Illa d'en Colom	COLOM	10	10	20	609129.9	4424151.2	59.3	44.0	2.117	1616
Mallorca/Cabrera	Mallorca	Islet	el Colomer	COLOMER		4	4	511211.5	4421664.5	2.8	105.4	1.187	3339
Mallorca/Cabrera	Mallorca	Islet	sa Dragonera	DRAGONERA		4	4	441589.9	4381854.5	272.7	352.8	2.049	729
Mallorca/Cabrera	Mallorca	Islet	Illot d'en Curt	ENCURT	23		23	503125.0	4347776.2	0.4	3.1	1.194	1799
Mallorca/Cabrera	Cabrera	Islet	Esclatasang	ESCLTS		10	10	495084.6	4330730.1	0.6	39.2	1.73	3118
Mallorca/Cabrera	Cabrera	Islet	Illot de na Foradada	FORADADA		10	10	498200.0	4339516.6	0.3	5.7	1.294	2595
Mallorca/Cabrera	Mallorca	Islet	na Guardis	GUARDIA	15		15	500100.3	4351217.2	2.5	8.6	1.392	1784
Mallorca/Cabrera	Mallorca	Islet	na Moltona	MOLTONA	11		11	500926.8	4350573.0	5.7	7.9	1.395	2093
Menorca	Menorca	Islet	Porros de Fornells	PORROS		11	11	597159.9	4432654.7	0.1	2.2	1.128	1676
Menorca	Menorca	Islet	Illa des Porros	PRCV	10		10	591657.0	4438581.0	8.6	18.4	1.621	1100
Menorca	Menorca	Islet	Illa des Revells	RAVL	10		10	597032.1	4433305.5	0.4	8.0	1.518	NA
Menorca	Menorca	Islet	Illa del Rei	REI		10	10	610060.7	4415982.1	4.3	15.2	1.221	403

^aVector geographic information was extracted from Base Topogràfica de les Illes Balears (https://ideib.caib.es/geoservels/rest/services/public/GOIB_BTIB_IB/MapServer/27), publicly available from Institut Cartogràfic i Geogràfic de les Illes Balears (ICGIB). Elevations were obtained from the 2-m resolution MDT02 Digital Elevation Model (DEM) publicly available from the Centro Nacional de Información Geográfica (<https://centrodedescargas.cnig.es>).

^bFrom Bassitta et al. (2021).

^cBased on Pérez-Mellado et al., 2008 and Bassitta et al., 2021.

heterozygosity (H_o), expected heterozygosity (H_e), nucleotide diversity (P_i), inbreeding coefficient (F_{is}), and number of private alleles (PAs). Allelic richness (Ar) was calculated with the "hierfstat" package in R (Goudet et al., 2015) using four random individuals per population, that is, the minimum number of samples per population (minimum number of alleles, min.all=8). Significant associations between genetic diversity indexes (H_e , P_i , Ar , and F_{is}) and geographic variables (islet surface area, maximum altitude, and the shoreline development index, SDI, i.e., a measure of island perimeter complexity, all log-transformed) were assessed with simple and multiple linear regression in R.

2.5 | Population genetic differentiation and clustering

For differentiation and structural analyses, loci under Linkage Disequilibrium (LD) were removed using plink v2.00a3.3 (Purcell et al., 2007) (--indep-pairwise: 50 ["kb"] 5 0.20) (Table 2). Pairwise F_{ST} values (Weir & Cockerham, 1984) were estimated with the function *stamppFst* from "StAMPP" R package (Pembleton et al., 2013), and significant differences for each population pair were evaluated with confidence limits at 95% after 1000 bootstrap iterations. Heatmaps based on F_{ST} were created using *pheatmap* in R (Kolde, 2019).

Isolation by distance (IBD) was evaluated by correlating the F_{ST} genetic distances with linear geographic distances among populations, using a Mantel test (Mantel, 1967) with function *mantel.randtest* in the "ade4" package in R (Dray & Dufour, 2007) with 1000 permutations. The test was performed for Menorca and Mallorca/Cabrera archipelago separately to control for distance effect between major archipelagos.

To reconstruct evolutionary relationships among island populations, we first used a discriminant analysis of principal components (DAPC), integrated in the "Adegenet" R package (Jombart, 2008). In this multivariate statistical approach, sample genetic variance is partitioned into groups to maximize discrimination, without prior knowledge on genetic structure. The number of meaningful PCs retained was chosen by a cross-validation with the function *xvalDapc* (Jombart & Collins, 2015) with default parameters (1000 replicates).

Population structure, including contemporary and historic admixture among populations and their relationships, were further investigated using ADMIXTURE v1.3.0 (Alexander et al., 2009),

setting the numbers of co-ancestry clusters (K) from 2 to 16 (maximum number of populations). Cross-validation was performed with the -cv option (10-fold) to retain the best value of K clusters (lowest cv error). Membership probabilities were calculated according to the retained discriminant functions and specimens showing above 90% of assignment probability to an island different from their source were considered as recently translocated, or potential mislabeled and removed (none detected).

2.6 | Phylogeographic pattern

TreeMix v. 1.13 (Pickrell & Pritchard, 2012) was used to infer population splitting and mixture patterns. The method employs allele frequencies to build a graph-based model of the population network (as opposed to a bifurcating tree) by first building a maximum likelihood (ML) tree and then searching for migration events that increase the composite likelihood (Flesch et al., 2020; Pickrell & Pritchard, 2012). The program utilizes a Gaussian approximation to model genetic drift (drift parameter) along each population (Flesch et al., 2020; Pickrell & Pritchard, 2012). Using the combined dataset filtered for LD, a ML tree was built with a window size (k) of 500 SNPs, evaluating from 0 to 16 migrations edges (m), 10 iterations per edge, using the "-noss" option to prevent overcorrection of sample size. The optimal number of significant migration edges was then inferred from the second-order rate of change in likelihood (Δm) across incremental values of m with the *OptM* package in R (Fitak, 2021).

In addition to network analyses and as a complementary result, a ML tree was built using the concatenated SNP dataset (2867 bp). Prior to concatenation, the combined VCF file was filtered using BCFtools (Li, 2011) to remove individuals with more than 10% of missing data ($N=8$) and candidate SNP outliers ($N=18$, see Section 2.7). The filtered VCF file was then converted to PHYLIP format using the *vcf2phyip.py* script (Ortiz, 2019). Consensus sequences for each population were estimated with the function *consensusString* in the R package "Biostrings." The two Cabrera localities were considered as a single population ($F_{ST}=0.03$) and data merged. A ML tree was then built on the consensus alignment with IQ-TREE 2 v2.2.0.8 (Minh et al., 2020) using variable sites only and applying an ascertainment bias correction for SNP data (model GTR+ASC) (Lewis, 2001), with 10,000 pseudo-replicates.

TABLE 2 Summary of raw and filtered SNP count per dataset.

Dataset	Study	N ind	N pop	SNP count		
				Raw	Filtered ^a	LD pruned ^b
GBS	This study	100	8	18,88,392	143,272	34,065
RAD-Seq	Bassitta et al. (2021)	91	10	48,51,070	166,943	58,207
COMBINED	Both studies	191	16	63,94,354	5523	2876

^aFilters: --max-alleles 2, --remove-indels, --min-meanDP 10, --max-meanDP 50, --minQ 30, --maf 0.05, --max-missing 0.75 (individual datasets) or 0.90 (combined).

^bAfter linkage disequilibrium (LD): --indep-pair-wise 50 ["kb"] 5 0.20.

To identify the most ancestral population in our dataset we used IQ-TREE 2 with nonreversible substitution models (model 12.12) (Naser-Khdour et al., 2022) with 1000 ultrafast bootstrap replicates, using both the consensus alignment as well as one random specimen per population (all sites or only variables). The program performs a bootstrap analysis to obtain several ML rooted bootstrap trees; it then computes *rootstrap* support values for each branch in the tree, as the proportion of rooted bootstrap trees that have the root on that branch. A root testing was then performed with option `--root-test` to compare the log-likelihoods of the trees being rooted on every branch of the ML tree. The resulting trees were visualized and edited in Figtree v1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

2.7 | Identification and functional annotation of SNPs outliers

To detect footprint of genomic diversification and potential selection, outlier loci were identified using two independent approaches: PCAdapt software v 4.3.5 (Luu et al., 2017) and BayeScan v 2.1 (Foll & Gaggiotti, 2008). The PCAdapt package implemented in R software detects outlier loci based on PCA by assuming that markers excessively related to population structure are candidates for potential adaptation. The number of PCs to retain was chosen after checking score plots for population structuring, setting $maf=0.01$ and distance to "mahalanobis." The distribution of loadings (SNP contribution to each PC) was uniform, indicating no relevant LD effect. p -Values were corrected for false discovery rate (FDR) using a cutoff $q<0.01$ for outlier retention. BayeScan is designed to detect potential genetic loci under selection by analyzing variations in allele frequencies among specific groups with a Multinomial-Dirichlet model. The prior odd (PO) for neutrality indicates the ratio of selected:neutral sites (e.g., 1:1000) and provides a measure of uncertainty on the likelihood of the neutral model compared to the selection model (Lotterhos & Whitlock, 2014). The sensitivity of the analysis to the PO was evaluated using alternative values (1:100, 1:10,000). The final MCMC chain was run for 20 short pilot runs with 5000 integrations, 50,000 burn-in, thinning interval of 10, and PO set to 100. Loci were filtered for q -value <0.01 .

PCAdat and BayeScan results were finally intersected, retaining outliers with $F_{ST}>0.8$, thus providing a conservative set of candidate loci. The above analysis was performed separately for each dataset (GBS, RAD-Seq and COMBINED data).

The candidate SNP outliers were annotated by cross-referencing the SNP position against the GFF file of the *rPodLi1.1* genome assembly (Gomez-Garrido et al., 2023) for Gene ID association. For outliers falling in protein-coding region, a gene ontology (GO) annotation was performed, followed by a functional enrichment analysis with gGOST in g:Profiler (<https://biit.cs.ut.ee/gprofiler/gost>) on individual datasets (GBS and RAD-Seq). The *Podarcis muralis* genome was used as a reference to determine the functional categories

(biological processes [BP], molecular functions [MF], and cellular components [CC]) that were significantly enriched (FDR <0.05). For the subset of outliers falling within coding regions (CDS), we annotated the codon position and assessed whether the alternative allele (ALT) translated into a synonymous or nonsynonymous substitution with respect to the reference position in the genome (REF).

2.8 | Outlier association with environmental variables

A potential association between outliers and geographic/environmental variables (major archipelagos, islet surface, maximum altitude, and SDI) was assessed through a permutational multivariate analysis of variance (PERMANOVA) with the function *adonis2* in the "vegan" R package (Oksanen et al., 2022). A Euclidean distance matrix among individuals was built on allele frequencies calculated with *gl.percent.freq* in the "dartR" package (Gruber et al., 2018). Previously, missing genotype values were imputed using the function "gl.impute" in the dartR package. Model selection was performed by first fitting a model with all two-way interactions and all main terms, and then refitting the model after excluding nonsignificant interactions. Significance of model terms was assessed using marginal tests (1000 permutations). SNPs that significantly discriminated between major archipelagos were identified with logistic regression models using $p<0.01$ after adjusting p -values to minimize FDR. A genotype matrix coded as 0/1/2 was obtained using function *extract.gt* in the "vcfR" R package (Knaus & Grünwald, 2017) and used to visualize geographic clustering of individual SNP outliers by a principal component analysis (PCA).

3 | RESULTS

We generated a total of 15 billion paired-end GBS raw reads for 100 specimens from eight distinct localities/islands (99% passing the filters, Table S2). Newly generated data and previous RAD-Seq data from Bassitta et al. (2021) were individually processed through the same pipeline and combined for an integrative analysis of all 16 populations. After genome mapping, variant calling and filtering, we obtained a total of 143,272 biallelic SNPs for the GBS dataset (100 individuals and 8 localities), 166,943 for the RAD-Seq dataset (91 individuals, 10 localities), and 5523 for the COMBINED dataset (191 individuals, 16 localities, including only common loci between GBS and RAD-Seq datasets) (Table 2). As expected, multi-allelic variants represented a minor fraction of the total polymorphic sites (2.1% for GBS and 2.5% for RAD-Seq) and were excluded from the study.

Both raw and filtered SNPs were widely distributed across the 18 autosomal and two sexual chromosomes (Z and W), with SNP frequencies being proportional to chromosome size (Figure 2; chromosome W representation strictly depends on the number of males in the dataset, see Table S1). Both sequencing methods covered the

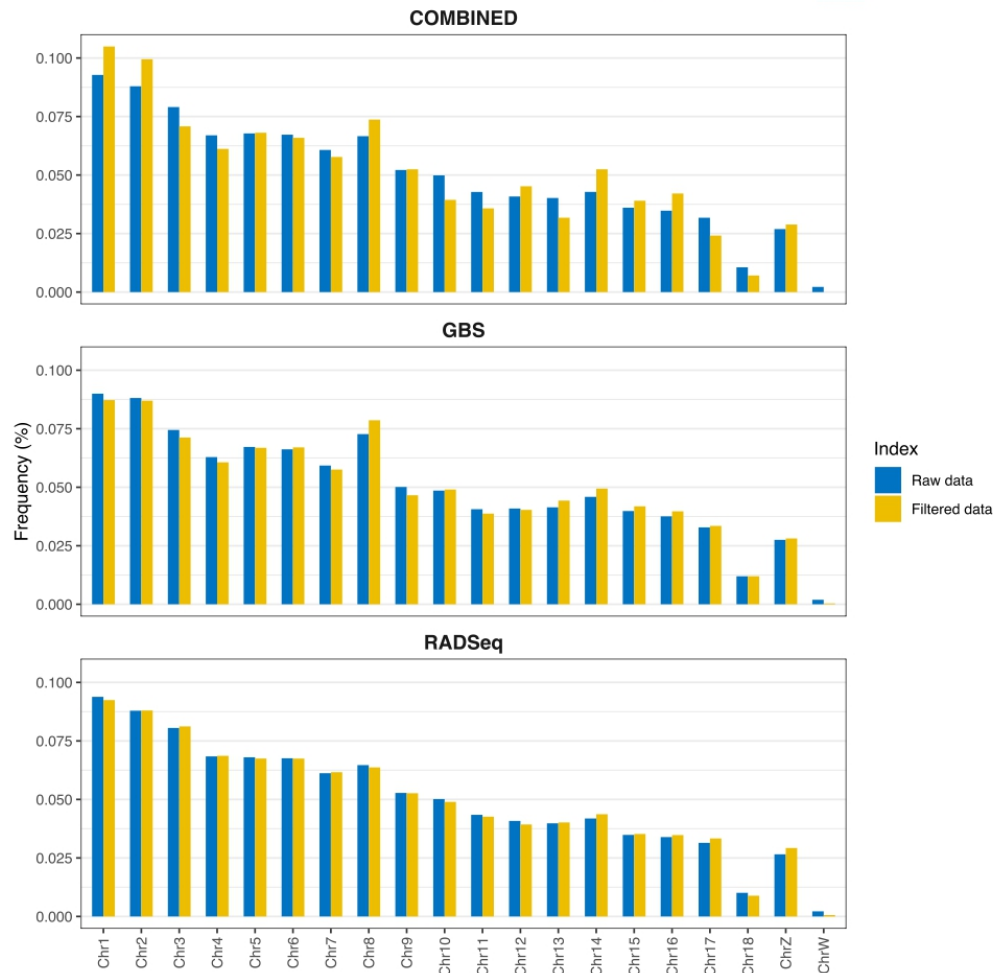


FIGURE 2 Chromosome distribution of raw and filtered SNP data for each of the three datasets (COMBINED, GBS, and RAD-Seq).

entire genome diversity, with chromosome representation being highly congruent among the three datasets (Figure 2).

3.1 | Population/islet genetic diversity

Genetic diversity based on the COMBINED dataset is summarized in Figure 3a. While diversity statistics for the individual datasets (GBS and RAD-Seq) was slightly higher in the absolute estimates per locality (Table S3), the relative pattern found among localities was highly preserved, indicating that the reduced COMBINED dataset is overall representative of the relative genetic diversity within this species.

Population nucleotide diversity (P_i) ranged between 0.104 and 0.174, with the three archipelagos showing similar ranges: 0.123–0.156 (Cabrera), 0.109–0.174 (Mallorca), and 0.104–0.158 (Menorca) (Figure 3a). The observed heterozygosity (H_o) showed comparable

values as P_i , varying between 0.095 and 0.140. Most diverse populations/islets per archipelago were the two Cabrera localities (Cabrera), Guardia and Moltona (Mallorca), and Aire and Colom (Menorca); the least diverse populations were the small islets of Porros (Menorca), Colomer (Mallorca), and Esclatasang (Cabrera) (Figure 3a). F_{IS} values were positive for all populations and range from 0.009 (Porros) to 0.107 (Moltona), indicating low inbreeding effect despite the small size of some islands. Individual datasets (GBS and RAD-Seq) based on a more extensive number of SNPs confirmed low F_{IS} values and major trend among populations (Table S3). The Allelic richness (A_r) was also comparable across localities, ranging between 1.245 (Porros) and 1.438 (Dragonera). The number of private alleles (PA) estimated on individual datasets (Table S3) vary greatly, with highest values found in the four smallest islets of Porros, en Curt, Esclatasang, and Foradada ($n > 1000$, island area < 1.50 ha), with the small islet of en Curt presenting the largest number ($n = 5316$) (Table S3).

(a)

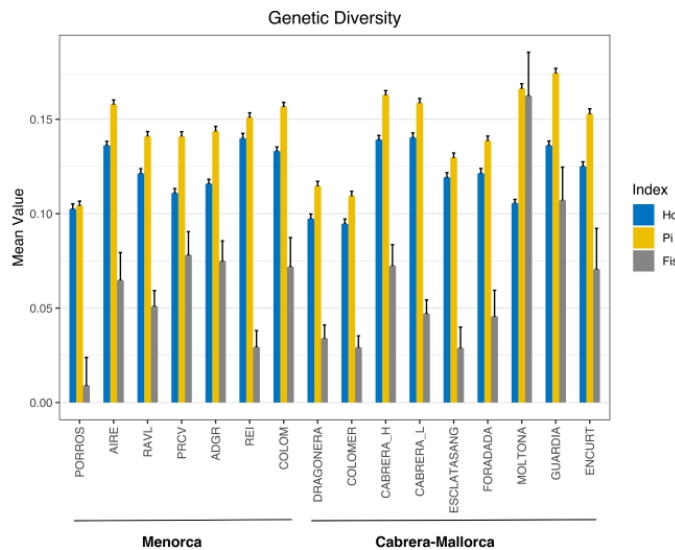
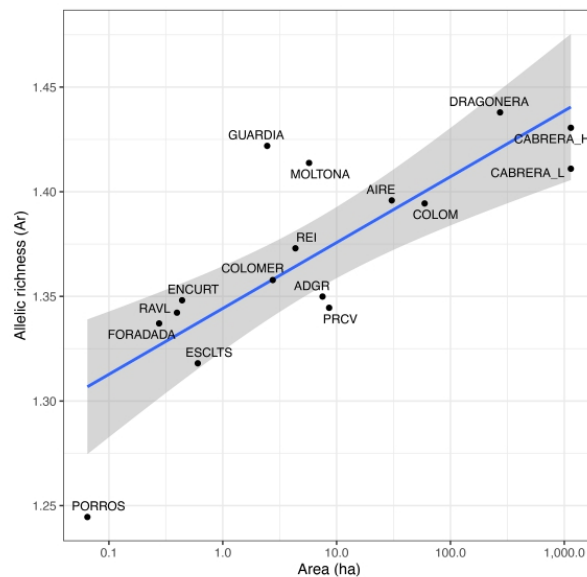


FIGURE 3 (a) Genetic diversity estimates per population (P_i : nucleotide diversity, H_o : observed heterozygosity, and F_{is} : inbreeding coefficient) with corresponding error bars. (b) Scatter plot of population allelic richness (Ar) and surface area (on log scale). The solid line shows the ordinary linear regression ($Ar = 1.34 + 0.014 \log(\text{area})$, $R^2 = .64$, $p = .0002$), while the ribbon shows the 95% confidence interval around mean expected values. All estimates are based on 5523 SNPs (COMBINED dataset).

(b)



Of the genetic diversity indexes, Ar , P_i , and F_{is} showed positive coefficients with the islet surface (Figure 3b, for Ar only) and negative with maximum altitude in multiple regression models (Ar : $p < .001$ only for area, P_i : $p = .001$ for area and $p < .005$ for max altitude, F_{is} : $p < .05$ for area and max altitude), indicating that the smallest islets typically host a more reduced genetic diversity, although no evident inbreeding depression.

3.2 | Population genetic differentiation and structure

Pairwise F_{ST} values (Weir & Cockerham, 1984) estimated for the combined dataset (2876 SNPs after LD pruning, Table 2) indicated that all populations were significantly differentiated ($p < .01$ for all pair-wises) (Figure 4). Global levels of differentiation among islet

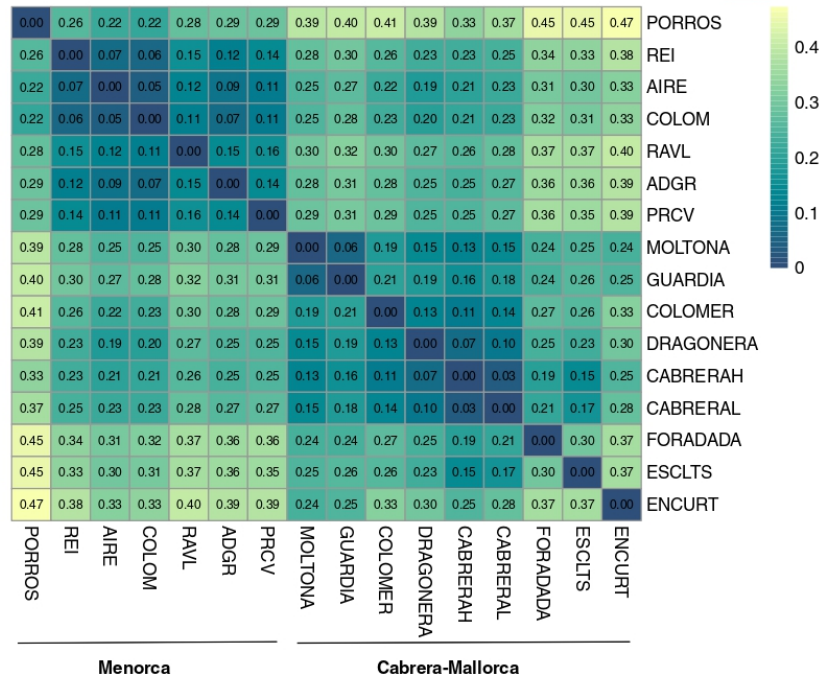


FIGURE 4 Heatmap of pairwise genetic differentiation (F_{ST}) between the 16 populations of *Podarcis lilfordi* (COMBINED dataset, 2876 SNPs). All F_{ST} estimates were significant ($p < .01$ after permutations).

populations were generally high, with a mean F_{ST} = 0.247 (ranging between 0.030 and 0.473). These values were highly comparable to those estimated on individual datasets, indicating that this reduced SNP dataset is also representative of the genetic differentiation among the studied populations.

The most discriminated population was the small islet of en Curt (Mallorca), showing an average F_{ST} of 0.338 with respect to all other islands (minimum value with Moltona, F_{ST} = 0.24, Figure 4). The least differentiated populations were Aire and Colom (0.047, Menorca), Na Guardis and Na Moltona (0.064, Mallorca), and the two Cabrera localities (0.030). In this latter case, F_{ST} value < 0.05 supports the existence of a unique panmictic population within the major Cabrera Island.

Within major archipelagos, no significant correlation was found between genetic (F_{ST}) and geographic distances ($p > .1$), excluding a scenario of isolation-by-distance.

Population genetic clustering with DAPC (Figure 5a) was consistent with F_{ST} -based distances (Figure 4) and the geographic distribution of these populations (Figure 1). Membership probabilities according to the retained discriminant function (64 PCs) correctly assigned all individuals to their source population/islet, excluding recent translocations or mislabeling. According to the first discriminant function (DF) (35% of variance), three major genetic clusters were identified (Figure 5a): (I) the group of Menorca islands, highly homogeneous and with poor internal discrimination, (II) the group of Cabrera and Mallorca islands, characterized by a stronger genetic differentiation among populations, and (III) the islet of en Curt, forming an independent lineage. DAPC on the subset of Menorcan

populations only (30 PCs retained) slightly increased the population resolution, highlighting a clear divergence of the islet of Porros (Figure 5a, top left panel).

ADMIXTURE analyses were used to further explore population structure, as well as shared ancestry (Figure 5b). Genetic clustering clearly evidenced a major discrimination between Menorca and Cabrera/Mallorca populations ($K = 2$), followed by progressive substructuring within Cabrera/Mallorca (from $K = 3$), and ultimately within Menorca (from $K = 5$). The best supported number of present genetic clusters is 9 ($CV = 0.386$), corresponding to four genetic clusters in Menorca and five in Mallorca/Cabrera, in line with the major structure recovered by DAPC analysis (Figure 5a). The proportion of shared ancestry among populations varies from zero (independent lineages of the islets of en Curt, Porros, Foradada, and Esclatasang) up to 40% (the highly admixed population of Colom). Cabrera island and distant islands from northwest of Mallorca (Dragonera and Colomer) formed a unique genetic cluster, with a small proportion of shared ancestries also with the small islets of Foradada and Esclatasang (North and South of Cabrera, respectively) (see map in Figure 1). According to the optimal current grouping ($K = 9$), no shared ancestry was detected between the two major archipelagos of Menorca and Mallorca/Cabrera, indicating a strong genetic differentiation. However, lower substructuring, from $K = 2$ to 9 retrieved a clear signature of shared ancestry between the Northern Mallorca populations of Dragonera and Colomer with the Menorcan populations of Colom, Rei, and Aire.

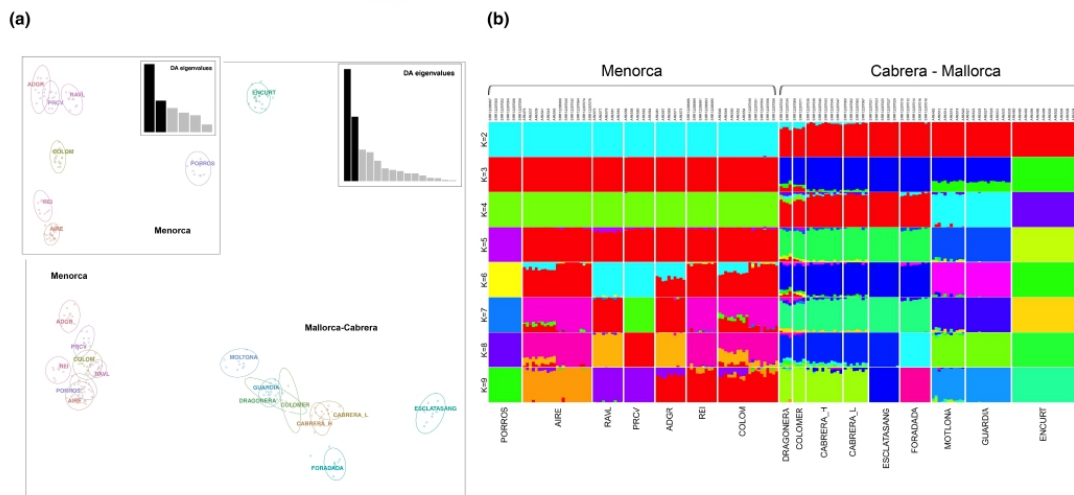


FIGURE 5 Population genetic structure based on 2876 SNPs (COMBINED dataset, after LD). (a) DAPC retaining the first 64 PCs (65% of the explained variance). (b) Unsupervised clustering by ADMIXTURE depicting $K=2-9$ (best value) clusters.

3.3 | Phylogeographic scenario

Prior to phylogenetic reconstruction, we assessed past migration events among populations using allele frequencies on Treemix and OptM analysis to support the best number of edges. OptM analysis indicated a single migration edge explaining 98.9% of the variance. The event with the greatest weight occurred from Colomer into Cabrera-Escalasang (mean migration weight: 0.454032, mean p -value: 3.22478×10^{-11}), followed by a migration event from Dragonera into Escalasang-Cabrera (mean migration weight: 0.380029, mean p -value: 6.71×10^{-5}) (Figure S1).

A midpoint-rooted ML tree was then built using consensus sequences for each population on the concatenated SNP dataset (Figure 6). The phylogenetic tree supports the major grouping obtained by the ADMIXTURE, F_{ST} , and DAPC analyses, largely in line with the population geographic distribution (Figure 1). Menorcan and Cabrera/Mallorca populations formed clearly distinct phylogenetic clades (bootstrap, $bs=100\%$). Within the Mallorca/Cabrera clade, further substructuring indicated three well-differentiated clades: (i) the southern Mallorca islets of en Curt, Moltona, and Guardia ($bs=99\%$), (ii) Cabrera islets of Foradada and Escalasang ($bs=76\%$), with Cabrera as a basal lineage, and (iii) the two northern Mallorcan islands of Colomer and Dragonera ($bs=80\%$) (Figures 1 and 6). Within the Menorca clade, substantial support is given only to the grouping of Porros de Cavalleria (PRCV) and Revells (RAVL) ($bs=73\%$), while the small islet of Porros de Fornells (PORROS) represents an independent lineage, highly differentiated from the rest of Menorcan populations ($bs=89\%$). The remaining Menorcan populations were poorly resolved, consistently with their lower F_{ST} values (Figure 4) and higher admixture levels (Figure 5b).

Tree rooting using nonreversible models on IQ-tree consistently identified the population of Colomer (Mallorca) as the most likely ancestral lineage. The island showed the highest rootstrap value (86.13%, after 10K bootstrapping on consensus sequences, all sites; values $<2\%$ for all other branches) (Figure S2). This result was robust to the use of consensus sequences with variant sites only and to subsampling to one random specimen per population, all consistently recovering the highest rootstrap value for Colomer (71.35 and 69.54, respectively). Root testing provided the maximum p -value of the approximately unbiased (AU) test for the branch leading to Colomer ($p=.85$). All branches within the Menorca clade showed a $p < .05$, significantly excluding Menorca as a potential source of the ancestral population. Rootstrap values and root testing considering only the archipelago of Menorca did not provide any clear root support within Menorca.

3.4 | Outlier identification

We explored potential selective pressures driving population genomic differentiation by identification of outlier SNPs. Intersection of outputs from two independent methods (PCAdapt and BayeScan) provided a total of 18 candidate outliers for the COMBINED dataset, 222 for the GBS and 190 for the RAD-Seq dataset (n total = 446, $q < 0.01$ and $F_{ST} > 0.8$, see Table S4 for a complete list).

Pattern of distribution of outliers per chromosome was highly concordant among the three datasets, with most outliers falling within the sexual chromosome Z ($>16\%$ for all datasets), followed by a large representation on chromosome 2 ($>9\%$) and 15 (12% and 8%, respectively) (Figure 7a). Outliers in chromosomes 3 and 4 were detected only by RAD-Seq.

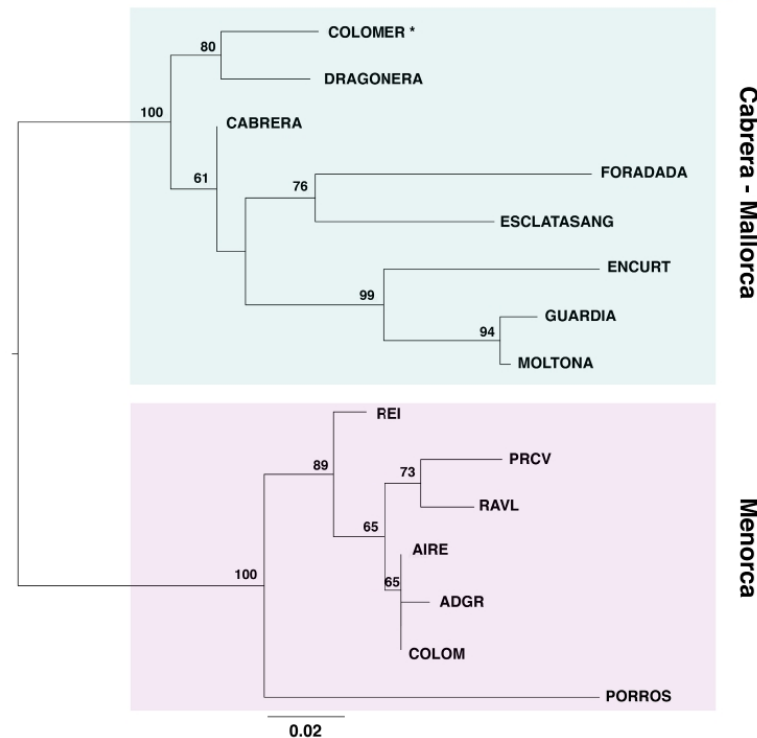


FIGURE 6 Midpoint-rooted ML tree based on consensus sequences per population after SNP concatenation (2876 SNPs, only variable sites, model GTR+ASC, 10,000 bootstrap replicates). Node numbers indicate bootstrap support. The asterisk indicates the most probable ancestral population according to IQ-TREE 2 root analysis.

About 60% of the outliers fell within protein-coding genes (44% for the COMBINED) corresponding to a total of 182 proteins (Figure 7a, pie charts). Only a minor proportion fell within CDS ($n=14$ outliers, corresponding to 11 annotated proteins). Five of these SNPs caused a nonsynonymous substitution (ALT) with respect to the reference genome (REF) and were detected in the glypican-3, B-cell lymphoma 6 protein, the cytochrome P450 2C19-like, the pore complex Nup214 genes, and an uncharacterized protein (see details in Table S4).

Functional enrichment analysis of annotated proteins with outliers indicated a comparable profile for GBS and RAD-Seq, showing a significant overrepresentation of binding proteins and catalytic activity (MF category), and metabolic process, growth, and cellular homeostasis (BP category) (BH, $p<.001$, Figure S3).

Of the overall 182 proteins with outliers, a small subset ($n=13$) was consistently retrieved by more than one dataset and most ($n=9$) were target of multiple outliers per dataset (up to seven) (Figure 7b). Three of these proteins were ion channels: the calcium channels inositol 1,4,5-trisphosphate receptor type 2 and the voltage-dependent L-type calcium channel subunit alpha-1C-like, and the potassium voltage-gated channel subfamily KQT member 1-like. Two proteins presented an outlier position identified by more than one dataset: the cullin-1 protein (SNP Position POS:937545) detected by all three datasets, and

the voltage-dependent L-type calcium channel subunit alpha-1C-like (POS:55226525) detected by both GBS and COMBINED (Table S4).

3.5 | Major archipelago-segregating SNPs

Finally, we explored environmental and geographic variables driving outlier distribution. PERMANOVA analyses indicated that major archipelago (Menorca and Mallorca/Cabrera, Table 1) was the only explanatory variable for all three datasets ($p<.001$, 1000 permutations), with all individual outliers, except one, significantly contributing to geographic between these two major phylogenetic lineages (Figure 6) (binomial test $p<.01$). PCA based on outlier genotype matrix clearly evidenced this major clustering (85.5% of variance explained for the COMBINED dataset, Figure 7c, see also Figure S4 for individual outlier genotype clustering). Only one SNP departed from this trend (SNP Chr2_24907248), recovering a clustering of South Mallorca with a subset of the Menorcan islands (Figure 7c and Figure S4). The SNP falls within the sidekick2 protein (Table S4).

A high percentage of outliers had allelic variants that perfectly segregated between major archipelagos, that is, had null frequency

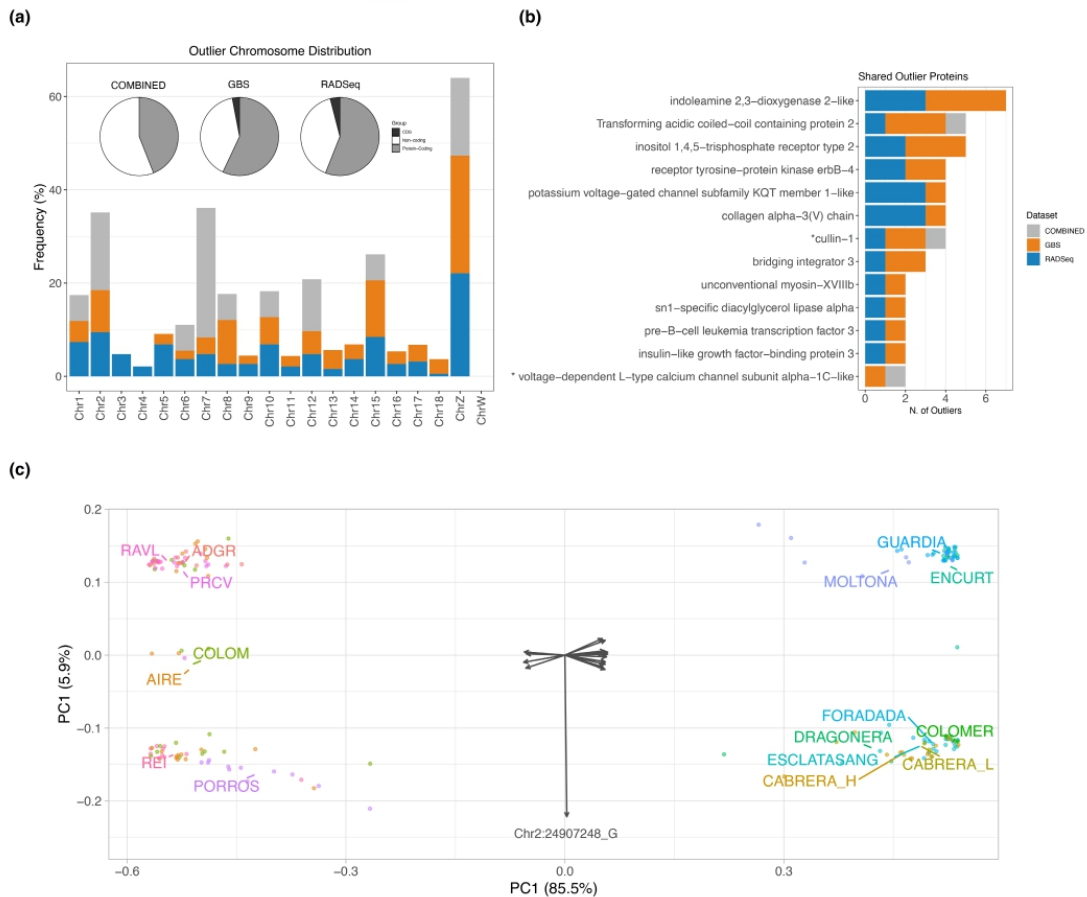


FIGURE 7 Candidate SNP outliers identified for GBS ($n=222$), RAD-Seq ($n=190$), and COMBINED ($n=18$). (a) Outliers chromosome distribution (barplot) and proportion per gene category (coding, CDS and noncoding regions) (pie chart). (b) Proteins with outliers detected by multiple datasets and corresponding number of outliers. The asterisk indicates proteins with shared SNP outliers among datasets. (c) Distance biplot of a principal component analysis (PCA) based on the outlier genotype distance matrix for the COMBINED dataset (18 SNPs). Circles are specimens plot to optimally reproduce their Euclidean distances; arrows are unscaled SNPs scores showing their contribution to each principal component.

in one archipelago while being present in the other (n total=154; 11% for COMBINED, 42% for GBS, and 31% for RAD-Seq) (Table S5). These fixed allelic differences can be considered as specific markers of the two major phylogenetic clades (Figure 6).

4 | DISCUSSION

4.1 | Genetic diversity of *P. lilfordi*: Evolution in islands and population conservation

A fundamental goal in population ecology and evolution is to understand the processes that maintain genetic diversity, and those that drive intraspecific/interpopulation divergence across geographic

space and time (Avice, 2000). In small insular populations, free of predators and with no immigrants, genetic diversity is primarily driven by genetic drift (due to environmental and demographic stochasticity, both particularly important in tiny islets) and density-dependent selection due to competition for the limited local resources (Hoffmann et al., 2021; Hunt et al., 2022; Travis et al., 2023). Founder effect and high inbreeding levels are also expected to further reduce the population genetic variance and accelerate the process of divergence from the original source population (Keller & Waller, 2002). Each *P. lilfordi* island clearly hosted a distinct population (all F_{ST} distances were significant), with no contemporary gene flow (no recent translocations detected). Few small islets showed a remarkably high differentiation from all other islets (i.e., Porros and en Curt, $F_{ST} > 0.2$ for all pairs, see Figure 4), indicating either

a strong founder effect and/or intense local drift and selection due to their reduced island size (Sendell-Price et al., 2021). Despite being effectively closed, *P. lilfordi* populations harbored relatively high genetic diversity ($H_o > 0.09$ and $P_i > 0.1$ for all populations), even higher than other continental and insular species of *Podarcis* (Sabolić & Štambuk, 2021). Diversity was overall largely comparable among all 16 populations studied and across their range of distribution (Figures 1 and 3a); nonetheless, association analyses supported the expected pattern of decreasing genetic diversity with decreasing island size (Furlan et al., 2012). Smaller islands also presented the highest number of private alleles, indicative of an ongoing process of differentiation. Mallorca/Cabrera archipelago typically hosted a higher genetic diversity than Menorca (Figure 3a), with the islands of Cabrera, Guardia, and Moltona (South of Mallorca) currently harboring the largest genetic diversity present in this species.

As previously reported by Bassitta et al. (2021), we also observed low levels of inbreeding for most populations ($F_{IS} < 0.08$), even for tiny islets such as Esclatasang, Foradada, Porros, Revells, and en Curt (all with area < 0.7 ha, $F_{IS} < 0.5$, Table 1 and Figure 3a). While small populations are expected to show a reduced heterozygosity (Keller & Waller, 2002), F_{IS} values in *P. lilfordi* indicated a slight trend of positive increase with increasing islet surface, with the highest estimate found for the island of Moltona ($F_{IS} = 0.16$). A small inbreeding coefficient, negatively associated with island size, was also found in a previous study based on microsatellites from the three neighboring populations of Moltona, Na Guardia, and en Curt (Rotger et al., 2021). A potential explanation is the existence of a population spatial substructure within larger islands, which would cause a nonrandom mating of individuals effectively reducing the observed heterozygosity. Nonetheless, we invite caution in interpretation of these results as SNP-based estimation of the inbreeding depression in wild populations is notoriously challenging (Schmidt et al., 2021) and would require further analyses based on pedigree and identification of runs of homozygosity (ROH) for a reliable quantification (Kardos et al., 2016).

Both the relatively high genetic diversity and low F_{IS} values observed in *P. lilfordi* would suggest the existence of potential mechanisms for buffering inbreeding depression in these lizard populations, increasing their chances of persistence. In line with this, recent studies support the ability of insular lizards to counteract genetic depletion even in presence of a strong founder effect (see Sherpa et al., 2023).

4.2 | A phylogeographic scenario of *P. lilfordi* colonization of the Balearic Islands

Podarcis lilfordi genetic diversity was largely structured according to the main geographic distribution, with all analyses supporting the existence of two major genetic lineages separating the large archipelagos of Menorca from Mallorca/Cabrera (Figures 5 and 6). The tree topology was largely consistent with the ones previously reported based on mtDNA and SNPs data (Bassitta et al., 2021; Brown

et al., 2008; Pérez-Cembranos et al., 2020; Terrasa et al., 2009). While Mallorca/Cabrera lineage is further sub-structured in four well-supported clades ($bs > 80\%$, Figure 6), phylogenetic relationships among Menorcan populations remain largely unresolved, as also indicated by their low degree of differentiation (Figure 4, F_{ST} values < 0.1). Menorcan islands were once largely panmictic (Brown et al., 2008; Terrasa et al., 2009) and a signature of past admixture can still be observed among most populations of this archipelago (Figure 5b), with the exclusion of the small islet of Porros, now representing an independent lineage (Figure 6).

Admixture analysis ($K=2$, Figure 5b), root inference by IQ-tree (Figure S2) and root testing all strongly suggest that the origin of the species occurred in the main island of Mallorca, with most ancestral populations in our dataset being identified in the North of Mallorca. The highest rootstrap support (86.13%) specifically points to the small island of Colomer as the population retaining nowadays most ancestral polymorphisms, in line with previous proposals (Bassitta et al., 2021; Terrasa et al., 2009). This root placement is consistent with the geographic history of both Dragonera and Colomer islands which were once connected to the Serra of Tramuntana, a mountain chain running up North of Mallorca, before their detachment from the main Mallorca Island about 2.3 Mya (Brown et al., 2008; Terrasa et al., 2009). The Serra of Tramuntana is considered as the last refuge of *P. lilfordi* (Bailón, 2004) before its extinction in the main Mallorca Island, as testified by fossil records dating around 2000 years old found in the Muleta cave, in the central part of the Tramuntana (Alcover, 2000; Kotsakis, 1981). The Colomer islet is characterized by high lizard density (Pérez-Mellado et al., 2008) and limited accessibility (due to high altitude and coastal shape). Although introduction of lizards from other islands cannot be ruled out, the islet has reduced chances of external gene flow, compatible with the retention of ancestral polymorphisms (Bassitta et al., 2021; Terrasa et al., 2009).

The following process of colonization within the archipelago of Mallorca/Cabrera did not follow an isolation by distance model (Mantel test, $p > .1$). According to admixture, phylogenetic analyses and Treemix reconstructions of past migration events, North of Mallorca was then the source of colonization of main Cabrera Island, with which they form a unique genetic clade (see Admixture results, Figure 5b). This closed genetic relationship was also reported by Bassitta et al. (2021), although it partly differs from mtDNA results (Terrasa et al., 2009).

Following this scenario, species expansion within the archipelago of Cabrera proceeded with the colonization of small surrounding islets (Esclatasang and Foradada), as supported by the observed coancestry with Cabrera Island (see Figure 5b). This suggests a founder effect in these small islands, followed by independent evolution leading to the nowadays well-differentiated genetic clade of Esclatasang and Foradada (Figures 5 and 6). Similarly, admixture and ML tree analyses indicated that islets on North of Cabrera could have been a source of colonization for Southern Mallorca islets, specifically the islet of Moltona, with which they still share a small proportion of coancestry (Figures 5b and 6), while the geographically closed sister

population of Na Guardis has currently lost all signature of coancestry with Northern Cabrera (Figure 5b). The observed clustering of Cabrera with South Mallorca islets confirms previous haplotype grouping based on mtDNA data, which also suggested a directionality of gene flow from Cabrera to South Mallorca (Network III; Terrasa et al., 2009). Moltona would have then seeded the small islet of en Curt, with which they form a unique clade (bs = 99%, Figure 6), as evidenced by their shared coancestry (Figure 5b). The tiny islet of en Curt (0.44 ha) would have followed a process of accelerated differentiation presumably driven by a founder effect, extensive genetic drift, and intense density-dependent selection with more than 1500 ind/ha (Ruiz De Infante Anton et al., 2013). As an alternative scenario, the three islets on the South of Mallorca directly derived from the ancestral mainland Mallorca population, now extinct, a scenario that we presently cannot exclude.

According to bathymetric, geological, and genetic data subsequent colonization of Menorca occurred around ~2.8 Mya (Brown et al., 2008; Terrasa et al., 2009). We were not able to assess the root within this archipelago due to high levels of admixture. Menorcan populations were highly panmictic during a large period of time and the colonization and subsequent isolation of most Menorcan islands was gradual (Pretus et al., 2004), following a vicariance type of colonization, with an asymmetric flow from south to north of the distribution of the species (Terrasa et al., 2009).

4.3 | Signature of genomic diversification of *P. lilfordi* populations

Local adaptation in effectively closed populations is a trade-off between genetic drift and selective pressures (Savolainen et al., 2013). The two processes are intrinsically associated with the island surface and its exposure to open sea, both affecting the impact of environmental stochasticity (stronger in smaller islets) and resource availability (reduced in smaller islands). Particularly, the vegetation cover, highly correlated to the island size, provides a strong cascade effect on lizard resource availability, including pollen and fruits production, insect visitor frequency and diversity, and even seabirds presence, all potential sources of dietary items (Pérez-Mellado et al., 2008; Ruiz De Infante Anton et al., 2013; Salvador, 2015; Santamaría et al., 2020). Along with stochasticity, this gives rise to a heterogeneous landscape across islands, which is expected to drive independent processes of lizard phenotypic diversification.

We explored the signature of genomic diversification that might underpin this lizard phenotypic diversity through outlier analysis. Outlier distribution showed a clearly skewed chromosome representation, with most outliers falling within the sexual chromosome Z, consistently for all datasets analyzed (Figure 7a). Recent studies in birds and lizards showed that closely related species often present high differentiation on the Z chromosome (Kulikova et al., 2022; Rovatsos et al., 2019), a pattern that is typically explained as faster Z evolution and lower recombination rates (Irwin, 2018; Lima, 2014; Mank et al., 2010; Wright et al., 2015). The existence of potential

"islands of differentiation" within the sexual chromosomes has also been hypothesized (Lavretsky et al., 2019). At present, detection of such "islands" in *P. lilfordi* Z chromosome would require a higher sequencing coverage than the one currently provided by GBS and RAD-Seq (i.e., genome resequencing).

Most outliers fell within protein-coding genes, although the large majority were intronic (Figure 7a). Substantial evidence supports the notion that introns have a crucial and evolutionarily conserved function in controlling gene expression in eukaryotes (Kumari et al., 2022; Rose, 2019), providing a particularly rapid mechanism for increasing variation in proteome variance by production of a diverse array of alternative splicing variants (AS) (Reixachs-Solé & Eyra, 2022; Wang et al., 2015). Moreover, most genes with outliers were associated with protein binding and catalytic activities and were involved in metabolic and growth processes (Figure S3), critical molecular functions for gene expression regulation and phenotypic diversification (Van Nostrand et al., 2020). In small, isolated populations, extensive modulation of gene expression conferring phenotypic plasticity could represent a major mechanism to counteract the loss of genetic diversity (Buglione et al., 2019; Fulgione et al., 2023; Sherpa et al., 2023), a hypothesis that needs to be validated by whole transcriptome and epigenome data (Chapelle & Silvestre, 2022; Fulgione et al., 2023).

We found a subset of protein-coding genes targets of multiple outliers according to both independent sequencing methods, GBS and RAD-Seq (Figure 7b), which might represent interesting candidates for further exploration of population genomic diversification in *P. lilfordi*. Of these proteins, few represented potassium and calcium channels, which could hint to an important regulatory role of osmotic pressure in lizards (Dantzer & Bradshaw, 2009). We also highlight the cullin-1 protein, with a shared outlier SNP detected by all datasets. The protein is known to be involved in ubiquitination and subsequent proteasomal degradation of target proteins (Duan et al., 2011; Gao et al., 2011; Scott et al., 2016). Recent evidence indicates a critical role of this protein in modulation of the transcription factor c-MYC, a major regulator of gene expression and cell proliferation (Sweeney et al., 2020). While there are no specific studies on cullin-1 protein in lizards, c-MYC was linked to the cellular regenerative response after tail amputation (Alibardi, 2017). The gradual loss of tail autotomy ability in insular lizards is a hallmark of their reduced antipredator response following insular adaptation (part of the island syndrome) (Cooper et al., 2004; Pafilis et al., 2009; Pérez-Mellado et al., 1997).

Exploration of the major variables driving outlier genotypes primarily recovered a signature of past geographic separation between Menorca and Cabrera/Mallorca, with a large proportion of loci presenting fixed allelic differences between archipelagos (Figure 7c). Given Mallorca as the species motherland, SNP allelic variants unique to the Mallorca/Cabrera archipelago could be considered as ancestral polymorphisms of the species, putatively retained since separation from the mainland. On the other hand, Menorca-specific allelic variants might be derived from a past founder effect that occurred during initial colonization of the Menorca archipelago (Brown et al., 2008;

Terrasa et al., 2009) or be the result of unclear selective forces, including potential selective sweeps (Brown et al., 2023; Campagna et al., 2022; Stephan, 2019). In all cases, these archipelago-specific outliers support no recent secondary contact between Archipelagos, in line with previous mtDNA-based phylogeographic reconstructions (Terrasa et al., 2009). We finally note that given the major confounding effect of archipelago, the significance/impact of other environmental variables in driving genome diversification cannot be reliably assessed with the current sampling design (Bassitta et al., 2021). Additional sampling, along with phenotypic data are required to clarify the putative adaptive role of these genomic changes.

5 | CONCLUSIONS

The Balearic lizard *P. lilfordi* has recently entered the genomic era, offering a new interesting insular system to explore trends in vertebrate genome evolution. Our study presents a first use of the *P. lilfordi* genome to reliably mapped extensive SNP data. Comparative analysis using independent methods and datasets provided robustness to our findings. We found support for the origin of this species colonization in the Mallorca Island, confirming previous proposals and providing a framework for future studies of the evolutionary trajectories in genome diversification for this insular species. The substantial genetic diversity observed in these effectively closed populations suggest that they have potential mechanisms to partially counteract genetic drift and inbreeding depression that are worth further investigation by genome resequencing and inclusion of additional populations. Based on the extensive divergence observed among most populations, we propose to consider each islet as a conservation unit for effective protection of the genetic diversity of this endemism. Future integration of population demographic data (ongoing) and collection of detailed phenotypic data, still scarce for this species, will be critical to evaluate the genomic plasticity and ability to persist of these insular populations.

AUTHOR CONTRIBUTIONS

Katherin Otalora: Conceptualization (supporting); data curation (equal); formal analysis (equal); funding acquisition (supporting); methodology (equal); writing – original draft (lead). **Joan Lluís Riera:** Data curation (supporting); formal analysis (supporting); methodology (supporting); writing – review and editing (supporting). **Giacomo Tavecchia:** Conceptualization (equal); funding acquisition (equal); methodology (equal); resources (equal); writing – review and editing (supporting). **Andreu Rotger:** Conceptualization (supporting); resources (equal); writing – review and editing (supporting). **José Manuel Igual:** Methodology (supporting); resources (supporting); writing – review and editing (supporting). **Jean-Remi Paul Trotta:** Formal analysis (supporting); writing – review and editing (supporting). **Laura Baldo:** Conceptualization (lead); data curation (lead); formal analysis (lead); funding acquisition (equal); investigation (lead); methodology (lead); project administration (lead); resources (equal);

validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (lead).

ACKNOWLEDGEMENTS

We thank Joan Lluís Pretus (UB), Sam Pons and Jordi García-Darocas for aid in sampling.

FUNDING INFORMATION

This study was supported by the CAIB – Government of the Balearic Islands (project PRD2018/25 to G.T.), the Institut d'Estudis Catalans under the Catalan Biogenome Project initiative (PRO2020-S02 to L.B.), the MCIN/AEI/10.13039/501100011033/FEDER, UE (PID2022-141578NB-C22 to L.B.), and the Doctoral Scholarship, Colombian Ministry of Science, Technology and Innovation (MINCIENCIAS885/2020 to K.O.).

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

The raw data with the individual sequences are available at the Sequence Read Archive (SRA) (BioProject ID: PRJNA1070579).

ORCID

Laura Baldo  <https://orcid.org/0000-0003-4528-6674>

REFERENCES

- Alcover, J. A. (2000). Vertebrate evolution and extinction on western and central Mediterranean islands. *Tropics*, 10, 103–123. <https://doi.org/10.3759/tropics.10.103>
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19, 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Alibardi, L. (2017). Immunolocalization of c-myc-positive cells in lizard tail after amputation suggests cell activation and proliferation for tail regeneration. *Acta Zoologica*, 98, 114–124. <https://doi.org/10.1111/azo.12153>
- Avice, J. C. (2000). *Phylogeography: The history and formation of species*. Harvard University Press. <https://doi.org/10.2307/j.ctv1nzfgj7>
- Bailón, S. (2004). Fossil records of Lacertidae in mediterranean islands: The state of the art. En: Pérez-mellado, V., Riera, N., Perera, A. (Eds.), *The biology of lacertid lizards. Evolutionary and ecological perspectives* (pp. 37–62). Institut Menorquí d'Estudis. Recerca, 8. Maó.
- Bassitta, M., Brown, R. P., Pérez-Cembranos, A., Pérez-Mellado, V., Castro, J. A., Picornell, A., & Ramon, C. (2021). Genomic signatures of drift and selection driven by predation and human pressure in an insular lizard. *Scientific Reports*, 11, 6136. <https://doi.org/10.1038/s41598-021-85591-x>
- Bloor, P., Rodríguez, V., Terrasa, B., Brown, R. P., Pérez-Mellado, V., Castro, J. A., Picornell, A., & Ramon, M. M. (2011). Polymorphic microsatellite loci for the Balearic Island lizard *Podarcis lilfordi* (Squamata: Lacertidae). *Conservation Genetics Resources*, 3, 323–325. <https://doi.org/10.1007/s12686-010-9354-7>
- Brown, R. P., Sun, H., Jin, Y., & Meloro, C. (2023). Habitat-associated genomic variation in a wall lizard from an oceanic island. *Genome Biology and Evolution*, 15, evad193. <https://doi.org/10.1093/gbe/evad193>

- Brown, R. P., Terrasa, B., Pérez-Mellado, V., Castro, J. A., Hoskisson, P. A., Picornell, A., & Ramon, M. M. (2008). Bayesian estimation of post-Messinian divergence times in Balearic Island lizards. *Molecular Phylogenetics and Evolution*, 48, 350–358. <https://doi.org/10.1016/j.ympev.2008.04.013>
- Buglione, M., Petrelli, S., Maselli, V., Trapanese, M., Salvemini, M., Aceto, S., Di Cosmo, A., & Fulgione, D. (2019). Fixation of genetic variation and optimization of gene expression: The speed of evolution in isolated lizard populations undergoing reverse island syndrome. *PLoS One*, 14(11), e0224607. <https://doi.org/10.1371/journal.pone.0224607>
- Camargo, A., Sinervo, B., & Sites, J. W. (2010). Lizards as model organisms for linking phylogeographic and speciation studies: Lizard phylogeography and speciation. *Molecular Ecology*, 19, 3250–3270. <https://doi.org/10.1111/j.1365-294X.2010.04722.x>
- Campagna, L., Mo, Z., Siepel, A., & Uy, J. A. C. (2022). Selective sweeps on different pigmentation genes mediate convergent evolution of Island melanism in two incipient bird species. *PLoS Genetics*, 18, e1010474. <https://doi.org/10.1371/journal.pgen.1010474>
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22, 3124–3140. <https://doi.org/10.1111/mec.12354>
- Chapelle, V., & Silvestre, F. (2022). Population epigenetics: The extent of DNA methylation variation in wild animal populations. *Epigenomes*, 6, 31. <https://doi.org/10.3390/epigenomes6040031>
- Charlesworth, B., Charlesworth, D., & Barton, N. H. (2003). The effects of genetic and geographic structure on neutral variation. *Annual Review of Ecology, Evolution, and Systematics*, 34, 99–125. <https://doi.org/10.1146/annurev.ecolsys.34.011802.123259>
- Cooper, W. E., & Pérez-Mellado, V. (2012). Historical influence of predation pressure on escape by *Podarcis* lizards in the Balearic Islands: Predation pressure and escape. *Biological Journal of the Linnean Society*, 107, 254–268. <https://doi.org/10.1111/j.1095-8312.2012.01933.x>
- Cooper, W. E., Pérez-Mellado, V., & Vitt, L. J. (2004). Ease and effectiveness of costly autotomy vary with predation intensity among lizard populations. *Journal of Zoology*, 262, 243–255. <https://doi.org/10.1017/S095283690300462X>
- Cuerda, J. B. (1989). *Los tiempos cuaternarios en Baleares*. Conselleria de Cultura, Educació i Esports, Direcció General de Cultura, Govern Balear.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27, 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Dantzler, W. H., & Bradshaw, D. (2009). Osmotic and ionic regulation in reptiles. In D. H. Evans (Ed.), *Osmotic and ionic regulation: Cells and animals* (pp. 443–503). CRC Press.
- De la Cruz, O., & Raska, P. (2014). Population structure at different minor allele frequency levels. *BMC Proceedings*, 8, S55. <https://doi.org/10.1186/1753-6561-8-S1-S55>
- Drake, D. R., Mulder, C. P. H., Towns, D. R., & Daugherty, C. H. (2002). The biology of insularity: An introduction. *Journal of Biogeography*, 29, 563–569. <https://doi.org/10.1046/j.1365-2699.2002.00706.x>
- Dray, S., & Dufour, A.-B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, 22, 1–20. <https://doi.org/10.18637/jss.v022.i04>
- Duan, S., Skaar, J. R., Kuchay, S., Toschi, A., Kanarek, N., Ben-Neriah, Y., & Pagano, M. (2011). mTOR generates an auto-amplification loop by triggering the β TrCP- and CK1 α -dependent degradation of DEPTOR. *Molecular Cell*, 44, 317–324. <https://doi.org/10.1016/j.molcel.2011.09.005>
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6, e19379. <https://doi.org/10.1371/journal.pone.0019379>
- Fitak, R. R. (2021). OptM: Estimating the optimal number of migration edges on population trees using Treemix. *Biology Methods & Protocols*, 6, bpab017. <https://doi.org/10.1093/biomethods/bpab017>
- Flesch, E. P., Graves, T. A., Thomson, J. M., Proffitt, K. M., White, P. J., Stephenson, T. R., & Garrott, R. A. (2020). Evaluating wildlife translocations using genomics: A bighorn sheep case study. *Ecology and Evolution*, 10, 13687–13704. <https://doi.org/10.1002/ece3.6942>
- Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180, 977–993. <https://doi.org/10.1534/genetics.108.092221>
- Fulgione, D., Maselli, V., Riveccio, E., Aceto, S., Salvemini, M., & Buglione, M. (2023). Evolutionary plasticity in insular lizard, adapting over reproduction, metabolism, and color variation. *Biology*, 12, 1478. <https://doi.org/10.3390/biology12121478>
- Furlan, E., Stoklosa, J., Griffiths, J., Gust, N., Ellis, R., Huggins, R. M., & Weeks, A. R. (2012). Small population size and extremely low levels of genetic diversity in island populations of the platypus, *Ornithorhynchus anatinus*: Low genetic diversity in island platypuses. *Ecology and Evolution*, 2, 844–857. <https://doi.org/10.1002/ece3.195>
- Gao, D., Inuzuka, H., Tan, M.-K. M., Fukushima, H., Locasale, J. W., Liu, P., Wan, L., Zhai, B., Chin, Y. R., Shaik, S., Lyssiotis, C. A., Gygi, S. P., Tokar, A., Cantley, L. C., Asara, J. M., Harper, J. W., & Wei, W. (2011). mTOR drives its own activation via SCF β TrCP-dependent degradation of the mTOR inhibitor DEPTOR. *Molecular Cell*, 44, 290–303. <https://doi.org/10.1016/j.molcel.2011.08.030>
- Gomez-Garrido, J., Cruz, F., Alioto, T. S., Feiner, N., Uller, T., Gut, M., Sanchez Escudero, I., Tavecchia, G., Rotger, A., Otalora Acevedo, K. E., & Baldo, L. (2023). Chromosome-level genome assembly of Lillford's wall lizard, *Podarcis lilfordi* (Günther, 1874) from the Balearic Islands (Spain). *DNA Research*, 30, dsad008. <https://doi.org/10.1093/dnares/dsad008>
- Goudet, J., Jombart, T., Kamvar, Z., Archer, E., & Hardy, O. (2015). *Hierfstat: Estimation and tests of hierarchical F-statistics*. R package version 0.04-22. <http://cran.r-project.org/>
- Goy, J. L. G., Zazo, C., Barceló, J. C., Goy, J. L. G., Zazo, C., & Barceló, J. C. (1997). Evolución de las áreas margino-litorales de la costa de Mallorca (I. Baleares) durante el Último y Presente Ingerglacial: Nivel del mar holoceno y clima. *Boletín Geológico y Minero*, 108, 127–136.
- Grant, A., & Benton, T. G. (2000). Elasticity analysis for density-dependent populations in stochastic environments. *Ecology*, 81, 680–693. [https://doi.org/10.1890/0012-9658\(2000\)081\[0680:EAFFDP\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[0680:EAFFDP]2.0.CO;2)
- Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). DART: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources*, 18, 691–699. <https://doi.org/10.1111/1755-0998.12745>
- Günther, A. (1874). Description of a new European species of *Zootoca*. *Annals and Magazine of Natural History*, 14, 158–159.
- Hawlana, D., Pérez-Mellado, V., & Cooper, J. (2009). Morphological traits affect escape behaviour of the Balearic lizards (*Podarcis lilfordi*). *Amphibia-Reptilia*, 30, 587–592. <https://doi.org/10.1163/156853809789647167>
- Hoffmann, A. A., Miller, A. D., & Weeks, A. R. (2021). Genetic mixing for population management: From genetic rescue to provenancing. *Evolutionary Applications*, 14, 634–652. <https://doi.org/10.1111/eva.13154>
- Hunt, D. A. G. A., DiBattista, J. D., & Hendry, A. P. (2022). Effects of insularity on genetic diversity within and among natural populations.

- Ecology and Evolution*, 12, e8887. <https://doi.org/10.1002/ece3.8887>
- Irwin, D. E. (2018). Sex chromosomes and speciation in birds and other ZW systems. *Molecular Ecology*, 27, 3831–3851. <https://doi.org/10.1111/mec.14537>
- Jombart, T. (2008). *Adegenet*: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jombart, T., & Collins, C. (2015). A tutorial for discriminant analysis of principal components (DAPC) using *adeget* 2.0.0. Lond. Imp. Coll. Lond. MRC Cent. Outbreak Anal. Model.
- Kardos, M., Taylor, H. R., Ellegren, H., Luikart, G., & Allendorf, F. W. (2016). Genomics advances the study of inbreeding depression in the wild. *Evolutionary Applications*, 9, 1205–1218. <https://doi.org/10.1111/eva.12414>
- Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, 17, 230–241. [https://doi.org/10.1016/S0169-5347\(02\)02489-8](https://doi.org/10.1016/S0169-5347(02)02489-8)
- Knaus, B. J., & Grünwald, N. J. (2017). *vcfR*: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, 17, 44–53. <https://doi.org/10.1111/1755-0998.12549>
- Kolde, R. (2019). *Pheatmap*: Pretty heatmaps. R package version 1.0.12.
- Kotsakis, T. (1981). Le Lucertole (Lacertidae, Squamata) del Pliocene, Pleistocene e Olocene delle Baleari. *Bolletti de la Societat d'Historia Natural de les Balears*, 25, 135–150.
- Krijgsman, W., Hilgen, F. J., Raffi, I., Sierro, F. J., & Wilson, D. S. (1999). Chronology, causes and progression of the Messinian salinity crisis. *Nature*, 400, 652–655. <https://doi.org/10.1038/23231>
- Kulikova, I. V., Shedko, S. V., Zhuravlev, Y. N., Lavretsky, P., & Peters, J. L. (2022). Z-chromosome outliers as diagnostic markers to discriminate mallard and Chinese spot-billed duck (Anatidae). *Zoologica Scripta*, 51, 401–414. <https://doi.org/10.1111/zsc.12540>
- Kumari, A., Sedehizadeh, S., Brook, J. D., Kozłowski, P., & Wojciechowska, M. (2022). Differential fates of introns in gene expression due to global alternative splicing. *Human Genetics*, 141, 31–47. <https://doi.org/10.1007/s00439-021-02409-6>
- Lavretsky, P., DaCosta, J. M., Sorenson, M. D., McCracken, K. G., & Peters, J. L. (2019). ddRAD-seq data reveal significant genome-wide population structure and divergent genomic regions that distinguish the mallard and close relatives in North America. *Molecular Ecology*, 28, 2594–2609. <https://doi.org/10.1111/mec.15091>
- Le Galliard, J. F., Marquis, O., & Massot, M. (2010). Cohort variation, climate effects and population dynamics in a short-lived lizard. *The Journal of Animal Ecology*, 79, 1296–1307. <https://doi.org/10.1111/j.1365-2656.2010.01732.x>
- Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, 50, 913–925. <https://doi.org/10.1080/106351501753462876>
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27, 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>
- Lima, T. G. (2014). Higher levels of sex chromosome heteromorphism are associated with markedly stronger reproductive isolation. *Nature Communications*, 5, 4743. <https://doi.org/10.1038/ncomms5743>
- Losos, J. B., & Ricklefs, R. E. (2009). Adaptation and diversification on islands. *Nature*, 457, 830–836. <https://doi.org/10.1038/nature07893>
- Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology*, 23, 2178–2192. <https://doi.org/10.1111/mec.12725>
- Luu, K., Bazin, E., & Blum, M. G. B. (2017). *Pcadapt*: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17, 67–77. <https://doi.org/10.1111/1755-0998.12592>
- Manes, S., Costello, M. J., Beckett, H., Debnath, A., Devenish-Nelson, E., Grey, K.-A., Jenkins, R., Khan, T. M., Kiessling, W., Krause, C., Maharaj, S. S., Midgley, G. F., Price, J., Talukdar, G., & Vale, M. M. (2021). Endemism increases species' climate change risk in areas of global biodiversity importance. *Biological Conservation*, 257, 109070. <https://doi.org/10.1016/j.biocon.2021.109070>
- Mank, J. E., Vicoso, B., Berlin, S., & Charlesworth, B. (2010). Effective population size and the faster-X effect: Empirical results and their interpretation. *Evolution*, 64, 663–674. <https://doi.org/10.1111/j.1558-5646.2009.00853.x>
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27, 209–220.
- Massot, M., Clobert, J., Pilorge, T., Lecomte, J., & Barbault, R. (1992). Density dependence in the common lizard: Demographic consequences of a density manipulation. *Ecology*, 73, 1742–1756. <https://doi.org/10.2307/1940026>
- Mayol, J., Bibiloni, G., Pons, G., Ballesteros, E., & Díaz, D. (2020). Descripció General de les Petites illes i els illots de les Balears. *Centro Oceanográfico de Baleares*, 1, 8–350. http://ibdigital.uib.es/greenstone/sites/localsite/collec/monografiesHistoriaNatural/index/assoc/MonografiesSHNB/_/2020vol0/29.dir/MonografieSHNB_2020vol029.pdf
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37, 1530–1534. <https://doi.org/10.1093/molbev/msaa015>
- Naser-Khdour, S., Quang Minh, B., & Lanfear, R. (2022). Assessing confidence in root placement on phylogenies: An empirical study using nonreversible models for mammals. *Systematic Biology*, 71, 959–972. <https://doi.org/10.1093/sysbio/syab067>
- Novosolov, M., Raia, P., & Meiri, S. (2013). The island syndrome in lizards. *Global Ecology and Biogeography*, 22, 184–191. <https://doi.org/10.1111/j.1466-8238.2012.00791.x>
- Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Solymos, P., Stevens, M., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., ... Weedon, J. (2022). *vegan*: Community ecology package. R package version 2.6-4.
- Ortiz, E. M. (2019). *vcf2phylip v2.0: Convert a VCF matrix into several matrix formats for phylogenetic analysis*. <https://zenodo.org/record/2540861>
- Pafilis, P., Foutopoulos, J., Poulakakis, N., Lymberakis, P., & Valakos, E. D. (2009). Tail shedding in island lizards [Lacertidae, Reptilia]: Decline of antipredator defenses in relaxed predation environments. *Evolution*, 63, 1262–1278. <https://doi.org/10.1111/j.1558-5646.2009.00635.x>
- Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). St AMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, 13, 946–952. <https://doi.org/10.1111/1755-0998.12129>
- Pérez-Cembranos, A., Pérez-Mellado, V., Alemany, I., Bassitta, M., Terrasa, B., Picornell, A., Castro, J. A., Brown, R. P., & Ramon, C. (2020). Morphological and genetic diversity of the Balearic lizard, *Podarcis lilfordi* (Günther, 1874): Is it relevant to its conservation? *Diversity and Distributions*, 26, 1122–1141. <https://doi.org/10.1111/ddi.13107>
- Pérez-Mellado, V., Corti, C., & Lo Cascio, P. (1997). Tail autotomy and extinction in Mediterranean lizards. A preliminary study of continental and insular populations. *Journal of Zoology*, 243, 533–541.
- Pérez-Mellado, V., Martín-Vallejo, J., Brown, R., Picornell, A., Castro, J., Ramón, M. M., Terrasa, B., García-Díez, T., & Hernández-Estévez, J. Á. (2008). Population density in *Podarcis lilfordi* (Squamata, Lacertidae), a lizard species endemic to small islets in the Balearic Islands (Spain). *Amphibia-Reptilia*, 29, 49–60. <https://doi.org/10.1163/156853808783431587>

- Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, 8, e1002967. <https://doi.org/10.1371/journal.pgen.1002967>
- Pretus, J., Marques, R., & Pérez-Mellado, V. (2004). Holocene sea level rise and range fragmentation of *Podarcis lilfordi* on Minorcan islets: A vicariance scenario reviewed through a mtDNA tree. In V. Pérez-Mellado, N. Riera, & A. Perera (Eds.), *The biology of lacertid lizards: Evolutionary and ecological perspectives* (Vol. 8, pp. 279–291). Institut Menorquí d'Estudis.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81, 559–575. <https://doi.org/10.1086/519795>
- Reixachs-Solé, M., & Eyra, E. (2022). Uncovering the impacts of alternative splicing on the proteome with current omics techniques. *WIREs RNA*, 13, e1707. <https://doi.org/10.1002/wrna.1707>
- Rose, A. B. (2019). Introns as gene regulators: A brick on the accelerator. *Frontiers in Genetics*, 9, 672. <https://doi.org/10.3389/fgene.2018.00672>
- Rotger, A., Igual, J. M., Genovart, M., Rodríguez, V., Ramon, C., Pérez-Mellado, V., Bibiloni, G., Rita, J., & Tavecchia, G. (2021). Contrasting adult body-size in sister populations of the Balearic lizard, *Podarcis lilfordi* (Günther 1874) suggests anthropogenic selective pressures. *Herpetological Monographs*, 35, 53–64.
- Rotger, A., Igual, J. M., Smith, J. J., & Tavecchia, G. (2016). Relative role of population density and climatic factors in shaping the body growth rate of *Lilford's* wall lizard (*Podarcis lilfordi*). *Canadian Journal of Zoology*, 94, 207–215. <https://doi.org/10.1139/cjz-2015-0188>
- Rotger, A., Tenan, S., Igual, J., Bonner, S., & Tavecchia, G. (2023). Life span, growth, senescence and island syndrome: Accounting for imperfect detection and continuous growth. *The Journal of Animal Ecology*, 92, 183–194. <https://doi.org/10.1111/1365-2656.13842>
- Rovatsos, M., Rehák, I., Velenský, P., & Kratochvíl, L. (2019). Shared ancient sex chromosomes in varanids, bearded lizards, and alligator lizards. *Molecular Biology and Evolution*, 36, 1113–1120. <https://doi.org/10.1093/molbev/msz024>
- Ruiz De Infante Anton, J., Rotger, A., Igual, J. M., & Tavecchia, G. (2013). Estimating lizard population density: An empirical comparison between line-transect and capture-recapture methods. *Wildlife Research*, 40, 552. <https://doi.org/10.1071/WR13127>
- Sabolić, I., & Štambuk, A. (2021). *Genomic patterns of phenotypic adaptation in Italian wall lizard Podarcis siculus (Rafinesque-schmaltz, 1810)* [Doctoral dissertation]. University of Zagreb, Faculty of Science, Department of Biology.
- Salvador, A. (2015). Lagartija balear – *Podarcis lilfordi*. In A. Salvador & A. Marco (Eds.), *Enciclopedia virtual de los vertebrados Españoles*. Museo Nacional de Ciencias Naturales. <http://www.vertebradosibericos.org/>
- Santamaría, S., Enoksen, C. A., Olesen, J. M., Tavecchia, G., Rotger, A., Igual, J. M., & Traveset, A. (2020). Diet composition of the lizard *Podarcis lilfordi* (Lacertidae) on 2 small islands: An individual-resource network approach. *Current Zoology*, 66, 39–49. <https://doi.org/10.1093/cz/zoz028>
- Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, 14, 807–820. <https://doi.org/10.1038/nrg3522>
- Schmidt, T. L., Jasper, M., Weeks, A. R., & Hoffmann, A. A. (2021). Unbiased population heterozygosity estimates from genome-wide sequence data. *Methods in Ecology and Evolution*, 12, 1888–1898. <https://doi.org/10.1111/2041-210X.13659>
- Scott, D. C., Rhee, D. Y., Duda, D. M., Kelsall, I. R., Olszewski, J. L., Paulo, J. A., De Jong, A., Ova, H., Alpi, A. F., Harper, J. W., & Schulman, B. A. (2016). Two distinct types of E3 ligases work in unison to regulate substrate ubiquitylation. *Cell*, 166, 1198–1214.e24. <https://doi.org/10.1016/j.cell.2016.07.027>
- Sendell-Price, A. T., Ruegg, K. C., Robertson, B. C., & Clegg, S. M. (2021). An Island-hopping bird reveals how founder events shape genome-wide divergence. *Molecular Ecology*, 30, 2495–2510. <https://doi.org/10.1111/mec.15898>
- Sherpa, S., Paris, J. R., Silva-Rocha, I., Di Canio, V., Carretero, M. A., Ficetola, G. F., & Salvi, D. (2023). Genetic depletion does not prevent rapid evolution in Island-introduced lizards. *Ecology and Evolution*, 13, e10721. <https://doi.org/10.1002/ece3.10721>
- Sivaperuman, C., Jaisankar, I., Velmurugan, A., & Sharma, T. V. R. S. (2008). Tropical Islands: Ecosystem and endemism. In *Biodiversity and climate change adaptation in tropical islands* (pp. 31–52). Academic Press, Elsevier. <https://doi.org/10.1016/B978-0-12-813064-3.00002-8>
- Stephan, W. (2019). Selective sweeps. *Genetics*, 211, 5–13. <https://doi.org/10.1534/genetics.118.301319>
- Sweeney, M. A., Iakova, P., Maneix, L., Shih, F.-Y., Cho, H. E., Sahin, E., & Catic, A. (2020). The ubiquitin ligase Cullin-1 associates with chromatin and regulates transcription of specific c-MYC target genes. *Scientific Reports*, 10, 13942. <https://doi.org/10.1038/s41598-020-70610-0>
- Terrasa, B., Pérez-Mellado, V., Brown, R. P., Picornell, A., Castro, J. A., & Ramon, M. M. (2009). Foundations for conservation of intraspecific genetic diversity revealed by analysis of phylogeographical structure in the endangered endemic lizard *Podarcis lilfordi*. *Diversity and Distributions*, 15, 207–221. <https://doi.org/10.1111/j.1472-4642.2008.00520.x>
- Travis, J., Bassar, R. D., Coulson, T., Reznick, D., & Walsh, M. (2023). Density-dependent selection. *Annual Review of Ecology, Evolution, and Systematics*, 54, 85–105. <https://doi.org/10.1146/annurev-ecolsys-110321-055345>
- Van der Auwera, G. A., & O'Connor, B. D. (2020). *Genomics in the cloud: Using Docker, GATK, and WDL in Terra* (1st ed.). O'Reilly Media.
- Van Nostrand, E. L., Freese, P., Pratt, G. A., Wang, X., Wei, X., Xiao, R., Blue, S. M., Chen, J.-Y., Cody, N. A. L., Dominguez, D., Olson, S., Sundaraman, B., Zhan, L., Bazile, C., Bouvrette, L. P. B., Bergalet, J., Duff, M. O., Garcia, K. E., Gelboin-Burkhart, C., ... Yeo, G. W. (2020). A large-scale binding and functional map of human RNA-binding proteins. *Nature*, 583, 711–719. <https://doi.org/10.1038/s41586-020-2077-3>
- Wang, Y., Liu, J., Huang, B., Xu, Y.-M., Li, J., Huang, L.-F., Lin, J., Zhang, J., Min, Q.-H., Yang, W.-M., & Wang, X.-Z. (2015). Mechanism of alternative splicing and its regulation. *Biomedical Reports*, 3, 152–158. <https://doi.org/10.3892/br.2014.407>
- Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilée, R., Condamine, F. L., Gravel, D., Morlon, H., Mouquet, N., Rosindell, J., Casquet, J., Conti, E., Cornuault, J., Fernández-Palacios, J. M., Hengl, T., Norder, S. J., Rijdsdijk, K. F., Sanmartín, I., Strasberg, D., Triantis, K. A., ... Thébaud, C. (2015). Islands as model systems in ecology and evolution: Prospects fifty years after MacArthur-Wilson. *Ecology Letters*, 18, 200–217. <https://doi.org/10.1111/ele.12398>
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38, 1358–1370.
- Whittaker, R. J., Fernández-Palacios, J. M., Matthews, T. J., Borregaard, M. K., & Triantis, K. A. (2017). Island biogeography: Taking the long view of nature's laboratories. *Science*, 357, eaam8326. <https://doi.org/10.1126/science.aam8326>
- Wingett, S. W., & Andrews, S. (2018). FastQ screen: A tool for multi-genome mapping and quality control. *F1000Research*, 7, 1338. <https://doi.org/10.12688/f1000research.15931.2>
- Wright, A. E., Harrison, P. W., Zimmer, F., Montgomery, S. H., Pointer, M. A., & Mank, J. E. (2015). Variation in promiscuity and sexual selection drives avian rate of faster-Z evolution. *Molecular Ecology*, 24, 1218–1235. <https://doi.org/10.1111/mec.13113>

Yang, W., Feiner, N., Pinho, C., While, G. M., Kaliontzopoulou, A., Harris, D. J., Salvi, D., & Uller, T. (2022). Extensive introgression and mosaic genomes of Mediterranean endemic lizards. *Nature Communications*, 12, 2762. <https://doi.org/10.1038/s41467-021-22949-9>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Otalora, K., Riera, J. L., Tavecchia, G., Rotger, A., Igual, J. M., Trotta, J.-R., & Baldo, L. (2024). Population genetics and phylogeographic history of the insular lizard *Podarcis lilfordi* (Gunther, 1874) from the Balearic Islands based on genome-wide polymorphic data. *Ecology and Evolution*, 14, e11407. <https://doi.org/10.1002/ece3.11407>

RESEARCH

Open Access



Antimicrobial peptides (AMPs) of lizards: a first comprehensive characterization of beta-defensins, ovo-defensins and cathelicidins from *Podarcis lilfordi* and closely related Lacertidae species

Katherin Otalora^{1,2}, Jessica Gómez-Garrido^{3,4} and Laura Baldo^{1,5*}

Abstract

Background Reptiles exhibit remarkable resistance to infections, making them a critical model for studying the molecular basis of innate immunity. Key components of their innate immune response are the antimicrobial peptides (AMPs), small, rapidly evolving molecules with a broad spectrum of antimicrobial activity. Despite their importance, AMPs in reptiles remain largely understudied, with no current studies available on one of the most successful reptilian lineages, lizards from the Lacertidae family. In this study, we performed an extensive mining of published Lacertidae genomes and present the first comprehensive characterization of their major AMP families —beta-defensins (BDs), ovo-defensins (OVODs), and cathelicidins (CATHs) — with the aim of understanding their genomic diversity and evolution.

Results We first identified a nearly complete catalogue of antimicrobial proteins from *Podarcis lilfordi*, an endemic lizard of the Balearic Islands (Spain): this included 63 BDs, eight OVODs, three of which are proline-rich proteins (OVOD-PrAMPs), and four CATHs. These AMPs are arranged in chromosome clusters of mostly contiguous peptides, flanked by highly conserved marker proteins. Using this fine-scale annotation, we mined corresponding orthologues and closed paralogues from *P. muralis*, *P. raffonei* and *Zootoca vivipara*, identifying a total of 58 AMPs. All AMPs consistently locate on chromosome 3 (BDs and OVODs) and chromosome 12 (CATHs), supporting a monophyletic origin for the reptilian antimicrobial defense. All Lacertidae AMPs show a multiple exon structure (two to four exons) and a characteristic cysteine motif (six cysteines in BDs, eight in OVODs, and four in CATHs), consistently with previous findings for vertebrates. Comparative sequence analyses indicate pervasive duplication events in tandem of both BDs and OVODs, with clear evidence of recent gene expansion. In contrast, OVOD-PrAMPs and CATHs mostly present a one-to-one ortholog across species. Despite this high level of intra-genomic diversity, we also identified multiple instances in which different species from distantly related Squamata families (Lacertidae, Varanidae and Dactyloidae) share identical or nearly identical peptides, suggesting convergent evolution.

*Correspondence:

Laura Baldo
baldo.laura@ub.edu

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Conclusions Overall, our findings support a remarkable diversity of AMPs in lizards, driven by an ongoing process of gene expansion, primarily via duplication in tandem, and rapid diversification. The observed instances of peptide identity or similarity, along with the putative maintenance of specific polymorphisms across species, provide a basis for investigating the potential functional equivalence of these peptides among lizards and underscore the need for population-level studies of allele intraspecific diversity.

Keywords Gene duplication, Proline, *Zootoca vivipara*, *Podarcis muralis*, *Podarcis raffonei*

Background

Antimicrobial peptides (AMPs) are a large class of small peptides (15–150 amino acids) that are found abundantly in nature and play a crucial role in the innate immune systems of all organisms, from bacteria to vertebrates [1–4]. They exhibit an ample spectrum of antimicrobial effects, primarily targeting bacteria, viruses, fungi and parasites, as well as potent immunomodulatory activity [2, 3]. In addition to serving as a primary defense against pathogens, AMPs act as key modulators of host–microbe interactions [5].

AMPs are typically derived from post-translational modifications involving the proteolytic cleavage of large precursor proteins [6]. Most active peptides present a net positive charge and an amphipathic nature, enabling direct membrane binding and disruption in target microbes [3, 4, 6]. Anionic peptides have been also reported; however, their mechanisms of actions remain mostly unclear [7].

Coevolution of AMPs and pathogen resistance has led to a highly functionally and structurally diverse repertoire of AMPs [2, 4] resulting from extensive gene duplication, pseudogenization, deletion as well as strong positive selection [1, 8]. Nonetheless, evidence of polymorphism maintenance by balancing selection and convergent evolution has also been documented [1, 3] indicating an overall complex pattern of molecular evolution. Hence, despite their ample spectrum of microbial targets, recent studies have shown that AMPs response can be remarkably specific, with single polymorphisms significantly influencing resistance to infections [1, 9].

Among the plethora of AMP-producing proteins found in vertebrates, the main classes are represented by defensins and cathelicidins (with the notable exception of caecilian amphibians that appear to lack defensins) [2, 10]. Both are characterized by the presence of cysteine residues (ranging from four to eight) arranged in conserved motifs and forming disulfide bonds that contribute to their structural stability. The number of cysteine residues and spacing of cysteine-pair alignment currently define the known subclasses [11, 12]. Beta-defensins are the most abundant and diverse subclass found in all vertebrates, known for their immunomodulatory and toxic activities, as well as their involvement in developmental processes, fertility, wound healing and cancer, among others [12]. They produce small active peptides

(20–50 residues) with a characteristic six-cysteine motif [12], and are typically arranged in large, discrete clusters within a chromosome [13, 14]. In the same syntenic block of the beta-defensins are the ovo-defensins, a small subclass of AMP-producing proteins structurally similar to beta-defensins (containing a six to eight-cysteine motif). Ovo-defensins have been described so far only in avian and reptile species, and are named for their expression in the oviduct and putative role in egg protection [15, 16].

Cathelicidins are a small class of largely studied, vertebrate-specific proteins with multifunctional activities marked by potent antimicrobial and immunomodulatory properties [17, 18]. They contain a typical four-cysteine motif and are synthesized as large precursor protein comprising a conserved cathelin-like domain (CLD) and a functional, highly variable antimicrobial domain [17, 19].

While AMPs have been extensively studied in humans, other mammals and amphibians (particularly frogs), a comprehensive genomic characterization of AMP diversity is still lacking in most other vertebrate groups [20]. Reptiles are an increasingly attractive model for antimicrobial defense studies, as they represent one of the earliest branches to evolve an innate immune response among extant vertebrates, and are known for their remarkable resistance to infections [21, 22]. Yet, AMP studies in reptiles are still at their beginnings [23, 24].

Over the past two decades, research has identified genes associated with major classes of antimicrobial peptides in a limited number of representative species across the four major reptilian orders (i.e. Squamata, Testudines, Rhynchocephalia and Crocodylia) [23]. Within Squamata (snakes and lizards), AMP diversity have been documented in species such as the Komodo dragon (*Varanus komodoensis*) [25], the green anole lizard (*Anolis carolinensis*) [26], and few snake species, including the king cobra (*Ophiophagus hannah*) [27] and the Burmese python (*Python bivittatus*) [28]. No formal AMP characterization is yet available for any members of the most successful lineage of the Squamata order: the family Lacertidae.

Lacertids are a highly species-diverse group of true lizards, primarily encompassing wall lizards that are widely distributed throughout the Mediterranean region. Few Lacertidae genomes have been recently published, including the common wall lizard, *Podarcis muralis* [29],

the Aeolian wall lizard, *P. raffonei* [30] and the viviparous lizard, *Zootoca vivipara* [31], thus providing the genomic resources for a first comprehensive overview of the AMP diversity within this reptilian family. In addition, our research group has recently produced a high-quality chromosome-level genome assembly and annotation for Lilford's wall lizard *Podarcis lilfordi* [32], and endemic species of the Balearic Islands where it inhabits multiple islets surrounding the islands of Menorca, Mallorca and the Cabrera Island. Due to the insularity of this species and the long-term monitoring of several of these populations [33], *P. lilfordi* is a growing natural model system for integrative studies of host and associated microbial symbionts [34–36] as well as pathogen resistance (unpublished) in a context of a rapidly changing environment.

Here we present the first and comprehensive identification of major classes of AMPs (defensins and cathelicidins) in the *P. lilfordi* genome by means of extensive genomic search. To further understand AMP diversity and evolution, we additionally mined AMPs from all currently published Lacertidae genomes (*P. muralis*, *P. raffonei* and *Z. vivipara*). Comparative analyses of these newly identified AMPs, along with previously published data from the Komodo dragon (family Varanidae) [25] and the green anole lizard (family Dactyloidae) [26], provided a first overview at the antimicrobial defense repertoire for this reptilian family, while also shedding light on major patterns of AMP evolution within Squamata.

Methods

Identification of candidate defensin and cathelicidin genes in *P. lilfordi*

We employed a multiple genomic approach to identify candidate defensin and cathelicidin genes within the *P. lilfordi* genome, including genome-wide TBLASTN screening of reference AMPs, filtering by characteristic cysteine motifs, genomic cluster delineation via conserved flanking markers, Open Reading Frame (ORF) identification, transcriptomic data validation, and physical properties and structural homology prediction. Each step was manually curated to ensure accurate and robust annotation.

Specifically, we first used the protein definitions file (PODLIA.protein_definitions.txt) available at <https://denovo.cnag.cat/podarcis> to retrieve annotated AMPs from the recently published genome assembly [32]. This dataset included four putative beta-defensins, one cathelicidin, and no ovo-defensins.

Second, we queried all available AMPs from the closest reptile species, the *V. komodoensis* (Komodo dragon) [25], and the *A. carolinensis* (the green anole) [26], in a TBLASTN search [37] against the *P. lilfordi* genome. The Komodo dragon includes 65 described beta-defensins peptides, six ovo-defensins (only mature peptides),

and three cathelicidins (complete genes) [25]; the green anole lizard includes 32 beta-defensins (complete mRNA), three cathelicidins (complete genes) and no ovo-defensins [26]. For each protein family, we applied relaxed TBLASTN search parameters (e-value=1, gap open/extension=8/2) to maximize output retrieval. The BLAST outputs were then filtered according to the presence of a conserved cysteine (C) motifs for each AMP family: six cysteines for beta-defensins, with a relaxed cysteine spacing from van Hoek et al. (2019); a terminal "CC" motif and between six to eight total cysteine residues for ovo-defensins [25, 26]; at least four cysteine residues for cathelicidins [19].

Third, we identified the chromosome and boundaries of defensin and cathelicidin clusters, by searching for conserved flanking marker genes previously documented in literature [25, 26, 38, 39]. For beta and ovo-defensins we targeted the exportin-1 protein (XPO1), the cathepsin B protein (CTSB), the XK, Kell blood group complex subunit-related family member 6 (XKR6) and myotubularin phosphatase domain-containing protein 9 (MTMR9) [25, 40]. For cathelicidins we targeted the fas-activated serine/threonine kinase (FASTK) and kelch-like family member 18 (KLHL18) [25, 38].

Throughout this systematic approach, we verified and corrected the orientation of all sequences, prioritizing non-overlapping sequences with the highest bit score. All analyses were performed with custom scripts in R software (2023).

Gene annotation

To obtain the complete ORFs for the candidate AMPs that passed the above filters, we queried all available full mRNA and proteins from *A. carolinensis* (32 beta-defensins mRNA and four cathelicidin complete proteins) and *V. komodoensis* (only three cathelicidins proteins) using the program exonerate v2.4.0 [41] in protein2genome mode against chr-3 (defensins) and chr-12 (for cathelicidins) of the *P. lilfordi* genome. For ovo-defensins, as no complete proteins were available, partial ovo-defensin proteins newly identified from *P. lilfordi* were searched against the NCBI nucleotide (nt) database: best RNA hits were translated to proteins and searched back against the *P. lilfordi* genome.

For each identified AMP gene, we used *P. lilfordi* Illumina and PacBio RNA-seq data from multiple individuals and tissues (Bioproject PRJNA897120) to accurately determine exon–intron structures, using this transcriptional evidence to guide the annotation process. This approach enabled us to validate most of the genes previously identified using partial sequences as well as to annotate novel genes.

Finally, we further validated all candidate peptides/proteins by their physical properties using EMBOSS

tool pepstats https://www.ebi.ac.uk/jdispatcher/seqstats/emboss_pepstats, and structural homology searches by protein prediction of secondary and tertiary structure in Phyre2 [42].

We named newly identified proteins and peptides in *P. lilfordi* using a standard nomenclature applied in other species consisting of the genus and species initials as a prefix, followed by a protein code and a number indicating their genomic order [25, 26]. Newly identified candidate proteins from Lacertidae were similarly named and numbered according to the *P. lilfordi* closest sequence, when available.

AMP chromosome location was visualized with the R function GC_chart in *geneviewer* R library. For the UPGMA (unweighted pair group method with arithmetic mean) dendrogram, a protein distance matrix was estimated with the R function 'dist.alignment' (matrix = "identity", gap = T) in *seqinr* library, and a dendrogram generated using *hclust* with height corresponding to the square root of identity distances.

Comparative AMP analysis across reptile species

To explore patterns of diversity and evolution of AMPs in Lacertidae, we queried all *P. lilfordi* AMP proteins in a BlastP and TBLASTN search against all published Lacertidae genomes to retrieve corresponding orthologs and paralogs. Available Lacertidae genomes included the Aeolian wall lizard, *P. raffonei* [30], the common wall lizard, *P. muralis* [29], and the viviparous lizard, *Z. vivipara* [31]. Additionally, we used available genome annotation and transcriptome data to guide our search. For comparative purposes, we also integrated available sequences from the Komodo dragon and the green lizard. Amino acid sequences were aligned using MUSCLE vs. 5.0 [43], visualized and manually curated in Jalview vs. 2.11.3.3 [44]. Due to the high rate of amino acid substitutions and extensive gene duplications, homology inference within the same protein family is limited. Therefore, alignment curation was performed according to the following guidelines: a) we first aligned C-residues within the conserved motif, b) we maximized homology among highly similar sequences, c) we optimized similarity within the conserved regions among all sequences, and c) for highly variable regions (i.e. functional peptide), we minimized indels among divergent sequences.

To identify clusters of similar peptides, as well as potential orthologs and paralogs within and across species, a Neighbor Joining (NJ) tree [45] was built with MEGA 4.0 [46] with 1000 bootstrap values. The resulting trees were visualized using the *ggtree* R package [47]. The use of distance-based methods (NJ and UPGMA), rather than model-based phylogenetic inference approaches such as Maximum Likelihood (ML), was guided by the complex nature of the AMP dataset. This complexity

includes extensive gene paralogy, low sequence homology beyond the conserved cysteine residues, and incomplete gene recovery across the studied species. Distance-based methods represent a conservative and widely used approach in AMP research (see, for example, [14]), as they minimize the risk of false inferences related to substitution model selection and do not assume true aminoacidic homology. While this approach limits formal phylogenetic interpretation, it allows for the identification of clusters of similar peptides and broad patterns of sequence similarity without overinterpreting evolutionary relationships.

Results

AMP genomic clusters in *P. lilfordi*

Genome-wide screening of *P. lilfordi* enabled the identification of two major AMP gene clusters, located on chromosomes 3 and 12 (Fig. 1).

Chr-3 cluster (112164644–117711852 bp, about 5.5 Mb) includes 71 putative defensins: 63 beta-defensins (BDs) and eight ovo-defensins (OVODs). The defensin block is flanked by the exportin-1 protein (XPO1) and the retrovirus-related Pol polyprotein from transposon TNT 1–94 (POL), with the cathepsin B protein (CTSB) falling within the block.

Ovo-defensins are found within the beta-defensin block, flanked by the myotubularin phosphatase domain-containing protein 9 (MTMR9) and the XK, Kell blood group complex subunit-related family, member 6 protein (XKR6).

Chr-12 cluster (54832594–58419520 bp, about 3.5 Mb) includes four cathelicidins (CATHs) arranged in a single block, flanked by the kelch-like family member 18 protein (KLHL18) and the fas-activated serine/threonine kinase protein (FASTK) (Fig. 1).

Characterization of *P. lilfordi* beta-defensins

All 63 BDs are arranged into two distinct clusters within chr-3, separated by ovo-defensins: one encompassing PIBD1-PIBD55 and the other PIBD56-63, separated by more than 2,000,000 bases (Fig. 1 and Supplemental Table S1). All BDs identified present a six-cysteine motif (29 and 35 amino acid (AA)) with the following conserved spacing: C-X(6–7)-C-X(3–5)-C-X(8–14)-C-X(5–7)-CC (Fig. 2). In addition, most sequences show two highly conserved glycine (G) residues at positions 7 and 28, and an isoleucine (I) at position 27 of the alignment (Fig. 2). Both PIBD12 and PIBD22 carry a double cysteine motif in tandem at the C-terminal, summing a total of 65 AMPs.

Most peptides present a cationic charge, with isoelectric points ranging between 6.8 to 10.7, with only 13 out of the 65 peptides showing a negative or no charge (Supplemental Table S1). According to protein homology

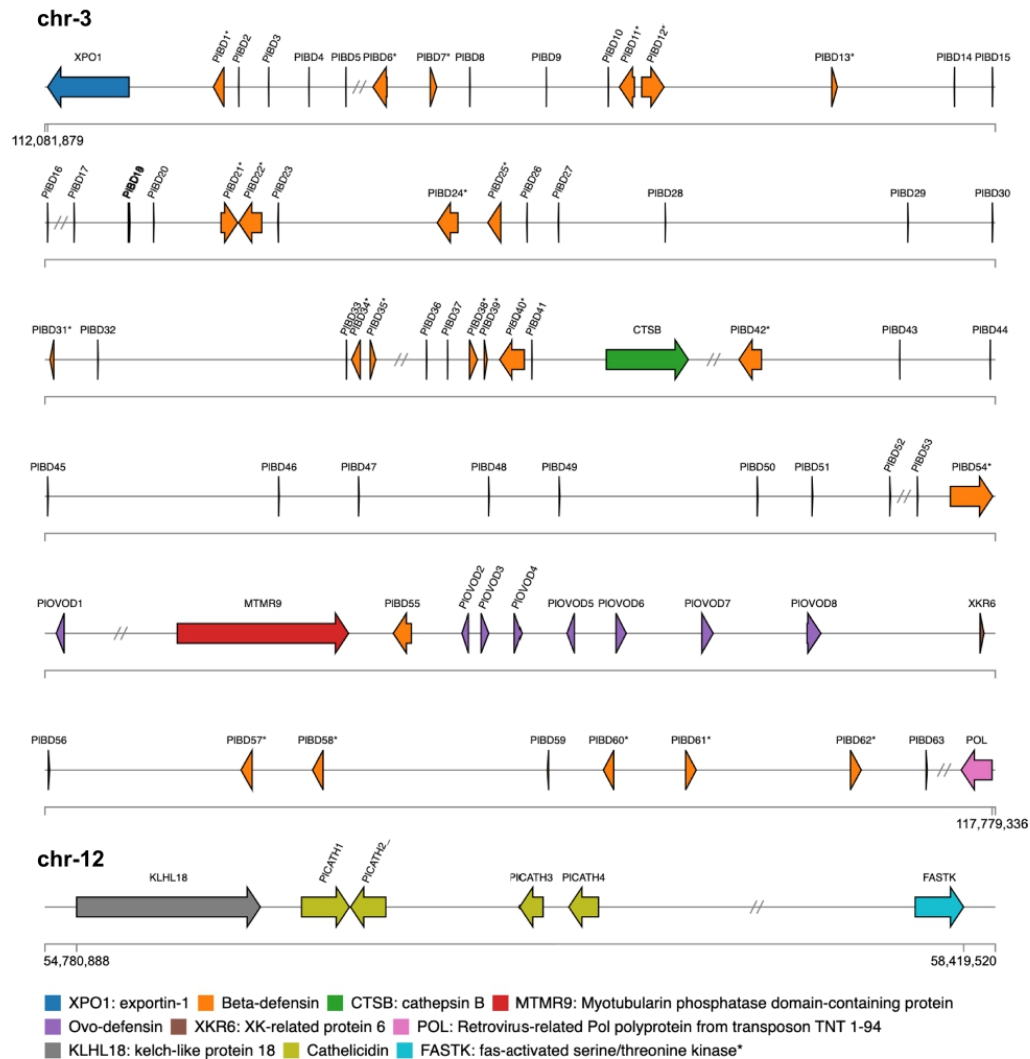


Fig. 1 *P. lilfordi* genome localization of beta-defensins (chr-3), ovo-defensins (chr-3) and cathelicidins (chr-12). Beta-defensins include both partial and complete ORFs (the latter marked with an asterisk). Marker genes for major clusters are also shown. Long intergenic regions are shortened and marked by a double //

searches, all peptides (at least 30 AA) have their best hits to an antimicrobial protein/beta-defensin, with confidence > 80% and sequence coverage > 70%, except for PIBD39 (Supplemental Table S2). Prediction of secondary structure indicates that most cysteine domains contain two to three beta-sheets.

The UPGMA (unweighted pair group method with arithmetic mean) dendrogram identified six clusters of similar sequences (< 25% dissimilarity based on an arbitrary cutoff), typically grouping contiguous or spatially

proximal peptides within the chromosome (e.g., PIBD51-52, PIBD47-48, PIBD19-20, see also Fig. 1). Two of these clusters encompass multiple BDs with identical motifs (PIBD57-58, 60-62 and PIBD59, 63), suggesting recent duplications in tandem.

Of the 65 beta-defensin motifs identified, we were able to detect the complete open reading frame (ORF) for 24 of them (Fig. 3). According to published *P. lilfordi* transcriptomic data from multiple individuals and tissues (Bioproject: PRJNA897120, but also see genome browser

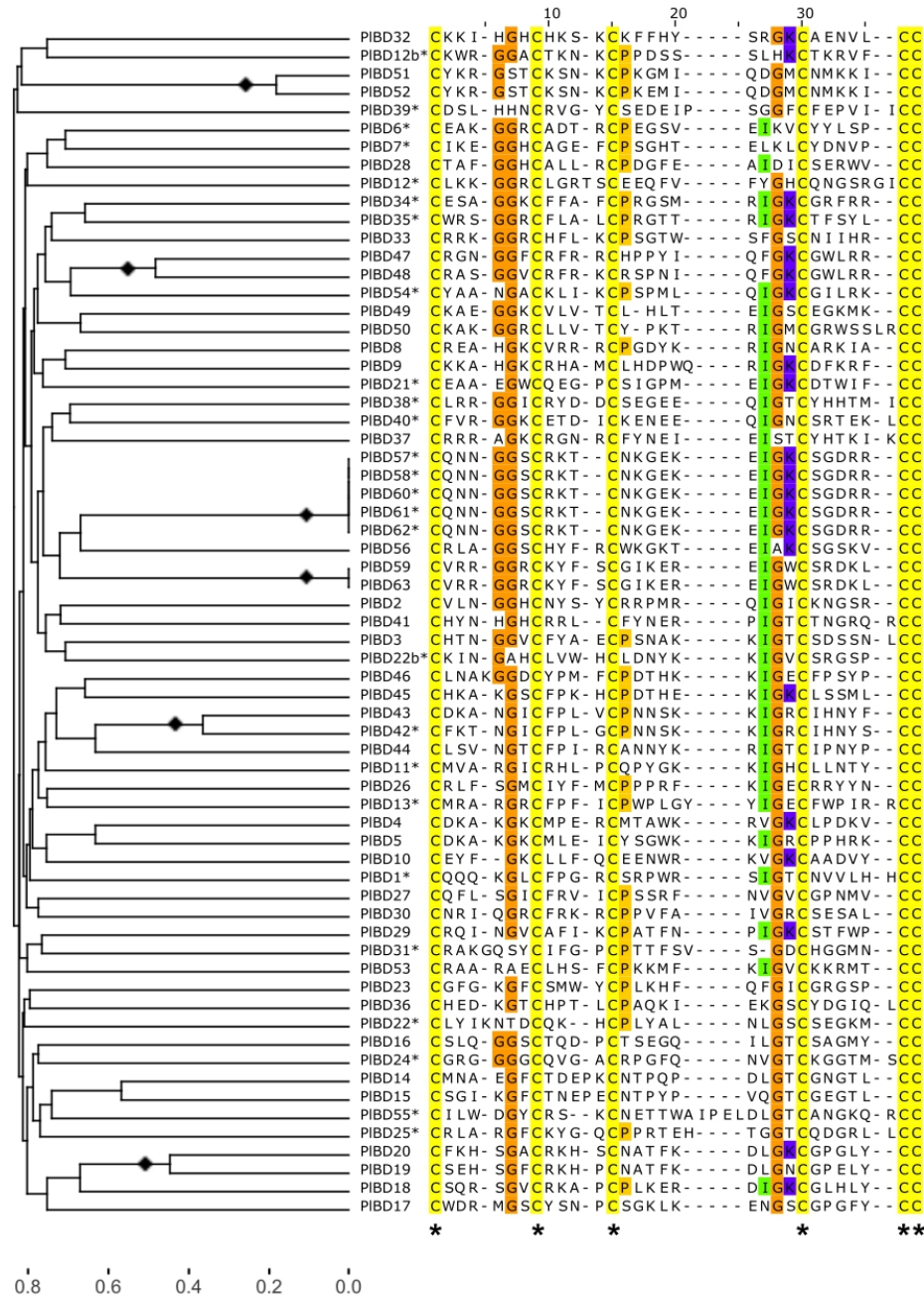


Fig. 2 Alignment and UPGMA dendrogram of *P. lilfordi* putative beta-defensin peptides (N=65) showing the six-cysteine domain (C-X(6-7)-C-X(3-5)-C-X(8-14)-C-X(5-7)-CC). Alignment colors show residues above 30% identity threshold. Shared cysteine residues are marked with an asterisk. Dendrogram height corresponds to the square root of sequence identity distances. Branch squares highlight closely related peptides (cutoff at 0.5 height, corresponding to <25% of dissimilarity). An asterisk in sequence IDs correspond to peptides with complete ORFs. PIBD12 and PIBD22 carried a double motif, and the second motif was labelled with "b"

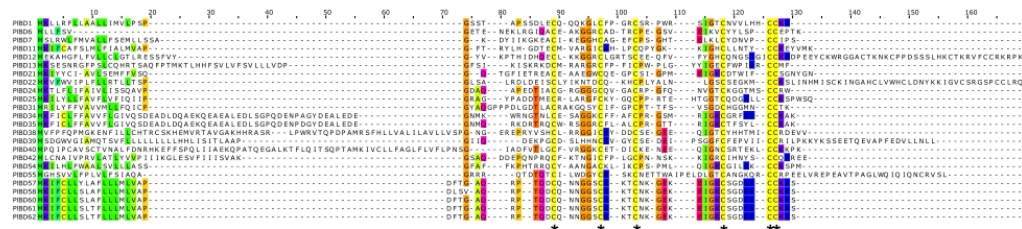


Fig. 3 Alignment of *P. lilfordi* complete beta-defensin proteins ($N=24$). Proteins encompass two to four exons (see Supplementary Table 1 for intron breaks). The six-cysteine residues are marked with an asterisk. Alignment colors show residues above 30% identity threshold. PIBD12 and 22 carry a double cysteine motif (the second motif was not aligned)

https://denovo.cnag.cat/genomes/podarcis/browse/rPo dLil1.2_annot/), fourteen of these BDs were transcribed, guiding our gene annotation (Supplemental Table S1).

Complete protein size ranges from 53 to 118 AA arranged from two to four exons, with the majority presenting two–three exons, with large exon size variation (exon-1: 6–55 AA, exon-2: 21–69 AA, exon-3: 0–49 AA, exon-4: 0–45 AA) (see Supplemental Table S1 for intron position). They contain a N-terminal with a putative signal peptide (first exon and part of the second exon), followed by the cysteine motif at the C-terminal encoding for the AMP. Structural homology searches based on complete proteins confirm strong similarity to previously identified beta-defensins, with confidence >90%, although with generally low sequence identity (35–65%) (Supplemental Table S2). Again, only PIBD39 shows poor confidence (<80%).

Notably, the BD cluster encompassing identical peptides (PIBD58–62, Fig. 2) also displays near AA identity at full protein length (Fig. 3), supporting recent whole-gene duplication.

Comparative analysis of beta-defensins across reptiles

To investigate patterns of beta-defensin diversity and evolution within the Lacertidae, we performed extensive BLAST searches against the genomes of closely related species to *P. lilfordi* (hereafter named Pl), including *P. raffonei* (hereafter Pr), *P. muralis* (Pm) and *Z. vivipara* (Zv). This analysis yielded several complete beta-defensin proteins: 11 from Pr (eight newly identified), seven from Pm (six new) and nine from Zv (six new). Notably, many of these newly characterized beta-defensins were previously annotated as uncharacterized non-coding RNAs or mRNAs. Detailed accession numbers for these sequences are provided in Supplemental Table S3.

BD peptides from all reptile species share the following six-cysteine motif: C-X(6–7)-C-X(3–7)-C-X(8–14)-C-X(4–7)-CC (see Supplemental Figure S1 for alignment).

The Neighbor-Joining (NJ) tree (Fig. 4A) identified fourteen multispecies clades (bootstrap support >70) and only one clade of *P. lilfordi*-specific sequences (clade 34).

Clades 54, 56 and 57 reflect species relative distances consistent with their taxonomic grouping: all *Podarcis* spp. cluster together, with Zv positioned at a basal node, supporting gene orthology. Most other clades do not provide species or genus resolution: these include clade-38 and 40 (grouping *Podarcis* spp. and Zv), clade-1 (Pm-Pl), clade-2 (Pm-Zv) and clade-7 and 11 (Pl-Pr) (Fig. 4A). Interestingly, clade-7 shows full identity of BD7 between Pr and Pl, suggesting putative convergent evolution (but see Discussion).

Due to a potential incomplete sampling of the full repertoire of beta-defensin full proteins from *Podarcis* spp. and Zv, we further explored pattern of BD evolution in reptiles based on cysteine motifs only, using a larger dataset that encompassed also the Komodo dragon and the green anole ($N=190$ sequences: Pl=65, Pm=7, Pr=12, Zv=9, *A. carolinensis* (Ac)=32, and *V. komodoensis* (Vk)=65) (Supplemental Table S3). The NJ tree (Fig. 4B) highlights three main aspects: within species recent duplications, species-specific lineages, and across species peptide convergence. Aside from instances of duplication in Pl, no other recent intraspecific duplication was observed for the any other species, except for the previously reported duplication in Vk (i.e. VkBD80a and b). Moreover, most Ac and Vk lineages appear to be species-specific. Nonetheless, both patterns could also result from pervasive gene loss and/or incomplete sampling of beta-defensin diversity in all six species under study, unclear at present. Notably, we identified eight clusters showing motif identity across distinct species (Fig. 4B). Such instances of motif identity are found for all species pairwise, although more frequently within the *Podarcis* genus (Pl-Pr: five clades; Pm-Pl: two clades; Ac-Vk: two clades, Pl-Pr-Zv-Ac-Vk: one clade). Remarkably, the clade encompassing AcBD29, PIBD38, VkBD4, ZvBD11 and PrBD11 (corresponding to clade-38 in Fig. 4A) shows motif identity among all species but Pm, thus encompassing all three reptilian families under study (Lacertidae, Varanidae and Dactyloidae).

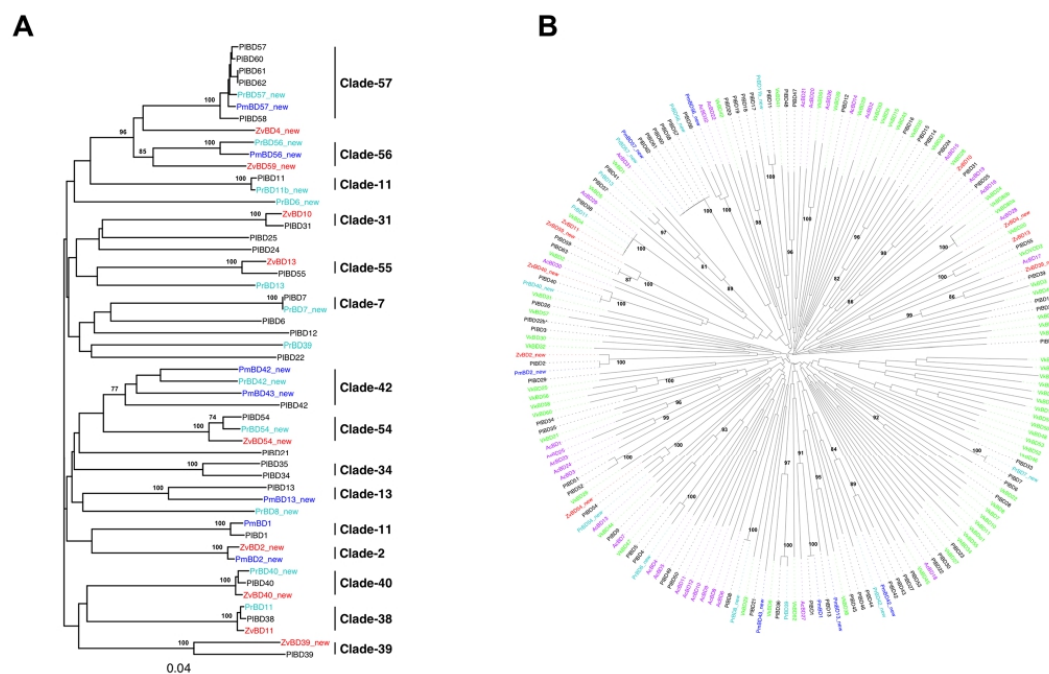


Fig. 4 Beta-defensin comparative analysis across multiple related reptile species. **A** NJ tree based on complete proteins ($N=47$). Major clusters (bootstrap support > 70) are highlighted and named according to the PIBD reference within the clade. **B** NJ tree based on the shared cysteine domain from all five species (C-X(6-7)-C-X(3-7)-C-X(8-14)-C-X(4-7)-CC) ($N=190$ sequences), highlighting major clusters (bootstrap support > 80). Proteins were named according to their public annotation; if uncharacterized, they were numbered according to their closest *P. lilfordi* sequence and marked with 'new' (see Supplementary Table 3 for species AccNos). Pl: *P. lilfordi*; Pm: *P. muralis*; Pr: *P. rafonei*; Zv: *Z. vivipara*; Vv: *V. komodoensis*; Ac: *A. carolinensis*

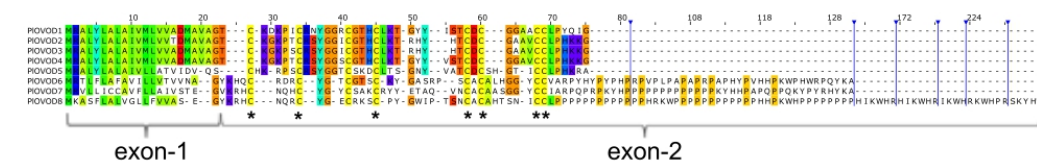


Fig. 5 Alignment of *P. lilfordi* complete ovo-defensin proteins ($N=8$). All proteins share a two-exons structure with a common Intron position and the eight-cysteine motif C-X(3-5)-C-X(3-6)-C-X(8-9)-C-X1-C-X(4-6)-CC, with a positive net charge. Blue lines mark the start of long stretches of prolines (> 10) at OVOD7 and OVOD8. Shared cysteine residues are highlighted with an asterisk

Characterization of *P. lilfordi* ovo-defensins

We detected a total of eight complete ovo-defensin proteins in chr-3 of *P. lilfordi*, tightly clustered within or near the region flanked by MTMR9 and XKR6 proteins (Fig. 1 and Supplemental Table S4). Protein size ranges from 64 to 239 AA arranged in two exons: the first exon putatively encodes for the signal peptide (about 21 AA), whereas the second exon includes the putative AMP (Fig. 5).

Intron size is relatively small, ranging from 1.3 kb and 2.3 kb, except for OVOD1, which has substantially larger intron (3.5 kb). All OVODs show an eight-cysteine motif with a highly conserved spacing: C-X(3-5)-C-X(3-6)-C-X3-C-X(8-9)-C-X1-C-X(4-6)-CC. PIOVOD7 contains

two additional CC in the signal peptide. All peptides show a conserved Tyrosine (Y) and a Glycine (G) at position 37 and 38, respectively (Fig. 5). PIOVOD6-8 are proline-rich proteins (PrAMPs), with PIOVOD7-8 displaying long stretches of up to 31 consecutive prolines (see Fig. 5 for position of proline repeats and Supplemental Table S4 for complete sequences). In PIOVOD7-8, proline repeats are interspersed with repetitive lysine-tryptophan-histidine (KWH) and lysine-tyrosine-histidine (KYH) motifs. Transcriptome data indicate that two of the OVODs (PIOVOD1 and PIOVOD4) are expressed (see Supplemental Table S4).

All OVODs exhibit a highly cationic charge (3.5 to 20) at exon 2 (i.e. the putative AMP), with isoelectric points ranging from 8.1 and 10.5 (Supplemental Table S4). Protein homology search revealed low-confidence assignment (<71%) and coverage (<58%) for most peptides (Supplemental Table S2). The best hit was given for PIOVOD5 to a hydrolase/hydrolase inhibitor (carboxypeptidase inhibitor) (96.8% of confidence and 57% of coverage), and for PIOVOD1 which matched an antimicrobial peptide (82% and 52%) (Supplemental Table S2). The remaining proteins did show any significant matches (>75% of confidence), likely due to the limited classification currently available for this group of defensins and their intrinsically disordered structures.

PIOVOD2-4 share large similarity (90–97% of AA identity), with PIOVOD2 and PIOVOD3 differing by a single AA (Fig. 5) likely due to a recent duplication event. OVODs that are located proximally on the chromosome tend to share higher protein similarity (Fig. 1), consistent with a duplication in tandem, as previously observed for BDs. PIOVOD1-5 show little homology with PIOVOD6-8 (i.e., OVOD-PrAMPs), with only 7–24% of AA identity between them. Apart from a conserved eight-cysteine residue motif, a two-exon structure, and a closed proximity on chr-3 (Fig. 1), these two groups of OVODs may represent distinct protein families or subfamilies.

Comparative analysis of ovo-defensins across reptiles

Through extensive BLAST searches and genome mining, we successfully retrieved several complete PIOVOD orthologs/paralogs from closely related species, none of which has been previously identified: five from *Zv*, four from *Pm*, and eight from *Pr* (Supplemental Table S3). Additionally, we included three previously described partial OVODs from *Vk* that met the minimum requirement of eight cysteine residues (i.e., *VkOVOD3*, 4 and 6). Due to the high fragmentation of current genome draft for *Varanus*, we were unable to retrieve the full protein sequences for this species.

The eight-cysteine motif was highly conserved across all reptile species, with a consistent spacing between the third and fourth cysteine, and between the fifth and the sixth cysteine: C-X(3–6)-C-X(3–6)-C-X3-C-X(8–12)-C-X1-C-X(4–6)-CC (Supplemental Figure S2). Notably, OVOD6-8 are proline-rich proteins in all Lacertidae, with OVOD7-8 showing long stretches of prolines interspaced by the conserved motif repeats KW(H) and KYH in all species (for *Pm*, only in OVOD8) (Supplemental Table S3 and Figure S2).

Based on the NJ tree constructed from complete proteins, we identified five main clades (A–D, bootstrap support >80) (Fig. 6).

Within Clade A, we observe several instances of recent intraspecific duplication within all *Podarcis* spp.,

including *Pl* (OVOD2-3, with one AA difference, and OVOD1-4), *Pr* (OVOD2-3, OVOD1-4, each pair sharing identical cysteine motifs) and *Pm* (OVOD2-3 with full protein identity). Clade A and B group together with high support, suggesting a common origin from a past lineage duplication. The remaining three clades (C–E) include the proline-rich proteins OVOD6-8, all of which have a one-to-one ortholog in all Lacertidae, with no recent duplications. OVOD8 shows cysteine motif identity (though not full gene identity) across *Pl*, *Pr*, *Pm* and *Zv*, and OVOD6 shows identity between *Pl* and *Pr* (Supplemental Figure S2).

Characterization of *P. lilfordi* cathelicidins

We identified four complete cathelicidin proteins on chr-12 of *P. lilfordi* (PICATH1-4), flanked by KLHL18 and FASTK proteins (Fig. 1 and Supplemental Table S5). PICATH2 was labelled as PICATH2-OH following the annotation of a closed related protein in *Zv*, named as *ZvCATH2-OH* to maintain the previous annotation from the king cobra (*Ophiophagus hannah*), where this protein was first described [27]. PICATH1 and PICATH2-OH are separated by only 266 bp, although located on opposite strands (Supplemental Table S5).

Protein size of the four CATHs range from 161 to 171 AA and are all organized into four exons (Fig. 7). All four CATHs were expressed according to transcriptome data. They exhibit a conserved four-cysteines domain with a highly conserved spacing: C-X10-C-X10-C-X(16–19)-C.

Homology searches returned best matches to the porcine CATHs, with 99.9% of confidence and coverage above 50% for all four proteins (Supplemental Table S2). Folding predictions indicate that all four PICATHs represent inactive precursors, structured into three domains: 1) a variable N-terminal domain with a signal peptide of slightly varying length (first 23–27 AA), 2) a conserved “cathelin-like domain” (CLD) containing the cysteine motif (100–101 AA long), with one alpha helix and four beta-sheets, and 3) a highly variable C-terminal domain containing the putative functional AMP (36–42 AA), with proteolytic cleavage predicted at position 130 (Fig. 7).

Except for PICATH1, all other AMP domains present a highly cationic charge (11.5–15) due to multiple Lys (K) and Arg (R) residues, with isoelectric points ranging between 12 to 13 (Supplemental Table S5). The AMP domain from PICATH1 largely differs in AA content from all other CATHs, lacking K and G residues and showing a net negative charge (–2).

Overall protein similarity among the four CATHs is low (<40%), indicating no recent duplications.

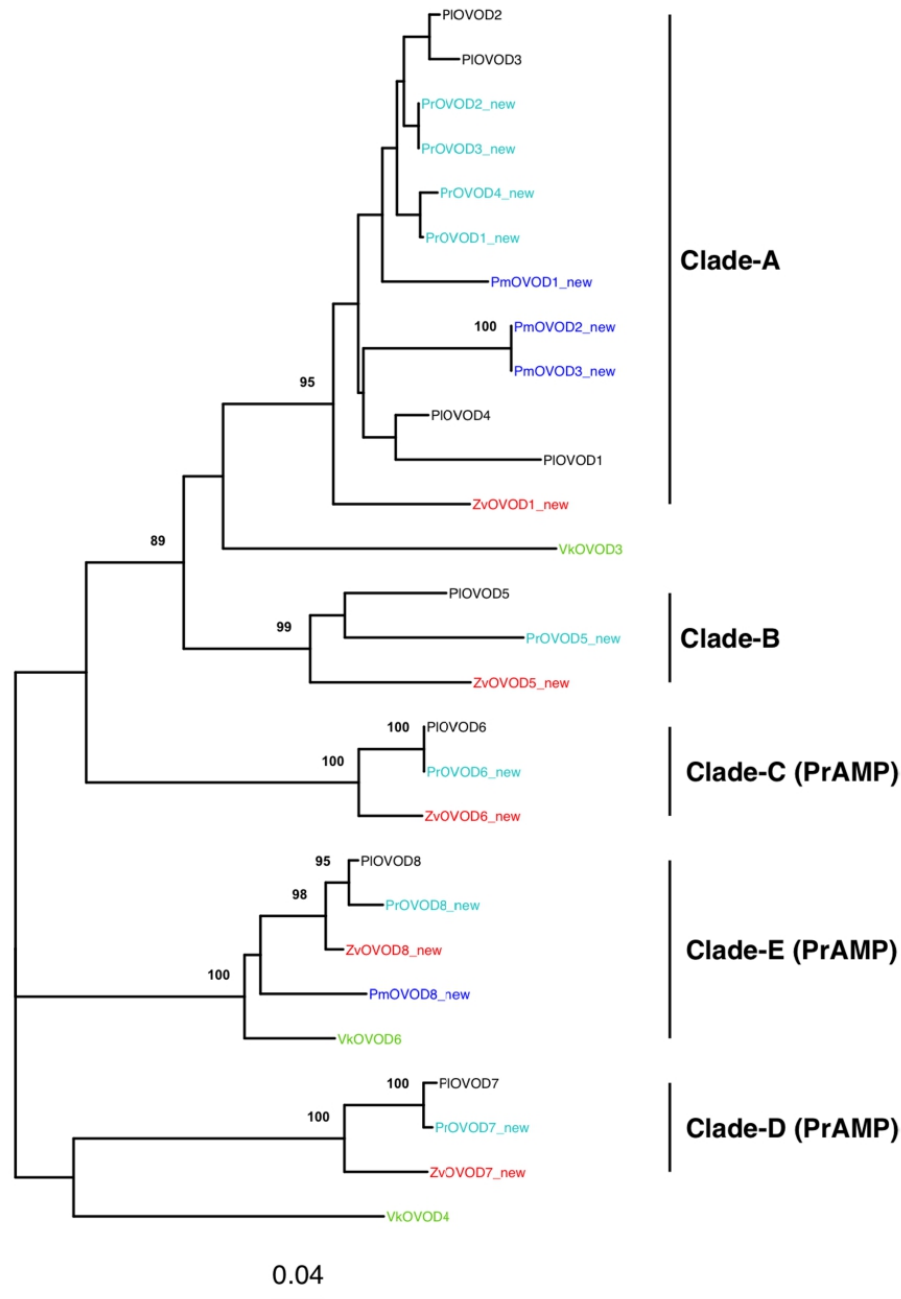


Fig. 6 NJ tree based on complete ovo-defensin proteins (partial only for Vkl) from multiple related reptile species ($N=17$). All species shared the following cysteine motif: C-X(3-6)-C-X(3-6)-C-X3-C-X(8-12)-C-X1-C-X(4-6)-CC. Major clades (bootstrap support > 80) are labelled as A-E. New proteins were marked with 'new' and numbered according to their closest PI sequence for Zv or to their sequential order in the chromosome for Pr and Pm (see Supplementary Table 3). Pl: *P. lilfordi*; Pm: *P. muralis*; Pr: *P. raffonei*; Zv: *Z. vivipara*; Vkl: *V. komodoensis*

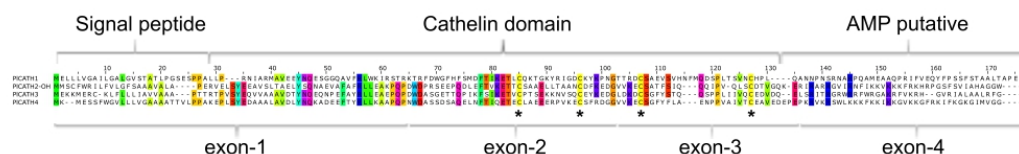


Fig. 7 Alignment of *P. lilfordi* complete cathelicidin proteins (inactive precursor) ($N=4$). All proteins share a four-exons structure and the four-cysteine motif C-X10-C-X10-C-X(16–19)-C. Main domain structure (signal peptide, cathelin domain and putative AMP) was predicted by homology searches. Shared cysteine residues are highlighted with an asterisk

Comparative analysis of cathelicidins across reptiles

For comparative purposes, we retrieved all available CATHs from the Lacertidae under study and additionally identified six novel ones: five from Pm (one newly identified), four from Pr (three new), and four from Zv (two new). We also included the three previously described CATHs from Vk and Ac. Most of the novel CATHs were previously annotated as uncharacterized mRNA (see Supplemental Table S3 for AccNos).

During sequence alignment, we found that a single CATH-coding mRNA sequence from Pm, Pr and Zv contained two consecutive cathelicidin proteins: the first protein showed high similarity to PICATH4, followed by a second protein with high similarity to PICATH3. Coexpression of CATH4-CATH3 was not observed in Pl, where the two proteins are separated by 5901 bp (Fig. 1 and Supplemental Table S5). For alignment purposes, we manually split the mRNA to recover the individual proteins and aligned them along with the other sequences.

The multispecies alignment indicates that all reptiles under study share a highly conserved four-cysteine motif: C-X(9–10)-C-X10-C-X(17–21)-C (Supplemental Figure S3). The Valine-130 (following the last cysteine residue), predicted to be the site of proteolytic cleavage for the active peptide in Pl, is also conserved across all reptiles in CATH2-4, but not in CATH1 (Supplemental Figure S3).

The NJ tree based on full proteins (Fig. 8) clearly depicts four highly supported clusters of one-to-one orthologous proteins (Clade 1–4), largely reflecting the main phylogenetic position of the two Lacertidae genera: all *Podarcis* species together (with no further resolution at species level), while Zv locates at a basal node.

Moreover, the species Ac and Vk are basal to both clade 1 and 3, with Vk presenting two closely related paralogs in CATH3. Both species lack orthologues to clade 2-OH or clade 4. However, given the incomplete genome assembly of *V. komodoensis*, this absence may also be attributed to a failure in sequence identification for this species. A recent CATH duplication is observed only in Pm, with two proteins falling within clade 2-OH.

Discussion

The innate immune system of lizards remains a largely understudied area of research [21, 39]. In this study we provided the first comprehensive genomic characterization of major AMP classes in *P. lilfordi* and closely related species within the family Lacertidae, with a comparative analysis extended to members of other Squamata families, the Komodo dragon and the green anole.

For all Squamata species under study, AMPs consistently clustered in two chromosomes: chr-3 for beta- and ovo-defensins, and chr-12 for cathelicidins, making these regions important repositories for the reptile antimicrobial defense. This genomic organization has been previously reported in crocodiles and birds, but not in mammals [23], supporting an avian and reptilian monophyletic origin. Within these regions, *P. lilfordi* flanking proteins of AMP clusters (i.e., XPO1 and CTSB for beta-defensins, XKR6 and MTMR9 for ovo-defensins, and FASTK and KHL18 for cathelicidins) were consistent with those previously described in other reptiles, but also mammals [23, 26, 39], underscoring their importance as a critical reference markers for AMP cluster identification in vertebrates.

Beta-defensins diversity and evolution in reptiles

Beta-defensins represent the most expanded and diverse AMP family in vertebrates [14]. Currently, no studies are available for Lacertidae, with Komodo dragon and the green anole being the closest species with a comprehensive characterization of beta-defensins [25, 26].

In *P. lilfordi*, the diversity of beta-defensins (65 peptides, 24 with complete ORFs) largely aligned with that previously described in the green anole ($N=32$) [26] and the Komodo dragon ($N=65$) [25]. All proteins presented substantial diversity in amino acid sequence, length, molecular weight, and isoelectric points. However, they all shared the canonical six-cysteine motif and a predominantly cationic charge, as reported for all vertebrates [12, 14, 47]. Moreover, complete beta-defensins exhibited broad structural variation, with two to four exons (two being the most recurrent), consistent with observations in the green anole [26] and the Komodo dragon [25].

In *P. lilfordi*, we observed that sequence similarity was typically higher among chromosomally adjacent peptides

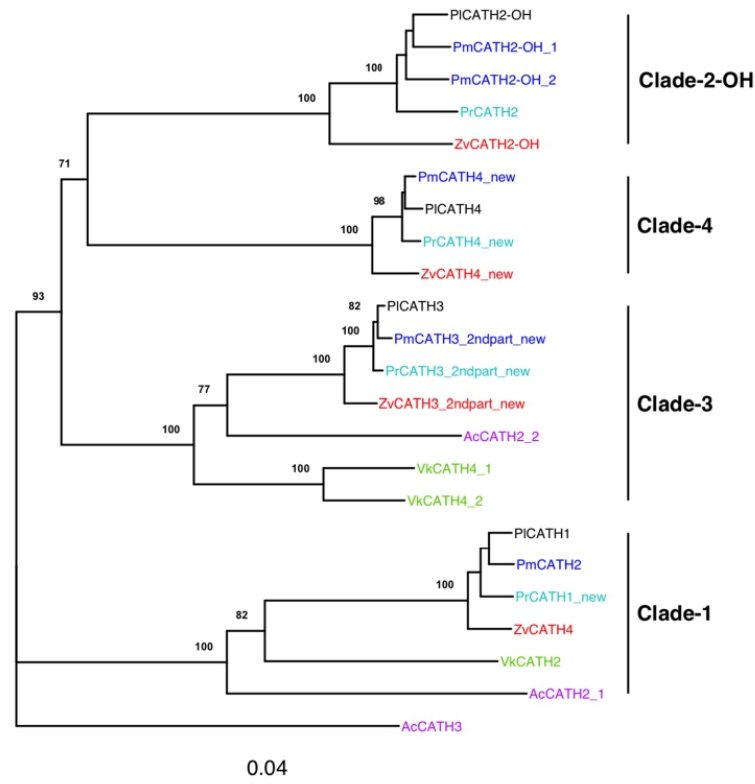


Fig. 8 NJ tree based on complete cathelicidin proteins from multiple related reptile species ($N=23$). All reptiles shared the four-cysteine motif C-X(9–10)-C-X10-C-X(17–21)-C. Major clades (bootstrap support >70) are labelled based on the *P. lilfordi* reference protein

or full proteins, suggesting a process of paralogous expansion via duplication in tandem. Instances of identity or near identity among pairs of contiguous peptides further indicate that this process is still ongoing. These observations align with previous research highlighting a major role of gene duplication in the increasing diversity of this gene family [1, 12, 14]. Notably, we also identified instances of beta-defensins (PIBD12 and 22) carrying two consecutive cysteine motifs at the C-terminal, a structural organization that was previously observed in other reptiles, but not in mammals [48], and that suggests the concomitant expression of multiple AMPs. Given the low sequence similarity between the two consecutive domains, these double motifs are unlikely to result from intra-protein duplication in tandem; rather they most likely arose from sequence translocation or other chromosomal rearrangements that remain to be clarified.

In line with this complex pattern of intra-genomic diversification, comparative beta-defensin analyses across reptiles recovered a generally low phylogenetic signal due to pervasive gene duplication, as well as putative gene

loss and rapid diversification at the active peptide [49]. While the extensive AMP genome screening of *P. lilfordi* provided a nearly complete catalogue of beta-defensins for this species, the present AMP characterization in the other Lacertidae genomes may be incomplete, limiting our ability to infer the relative role of gene loss and specie-specific gene duplication in the observed diversity. Nonetheless, we identified several clades of putative true orthologs across all Lacertidae (Fig. 4A), supporting a substantial level of evolutionary conservation for some beta-defensins, as previously observed in mammals [1, 47]. These clades showed variable branch lengths, suggesting heterogeneous rates of protein changes along species-specific lineages, likely driven by protein niche functional specialization in host–pathogen interactions [3, 11]. However, identification of true orthologues is needed to properly address rates of evolution and selection.

Interestingly, we observed several instances of cysteine motif identity or high similarity between members of the genus *Podarcis*, particularly between the two closest

related species *P. lilfordi* and *P. raffonei* [50]. Instances of identical peptides were also observed among members of distantly related genera and families, for example, a beta-defensin peptide from clade-38 was shared across all three studied families: Lacertidae, Dactyloidae, and Varanidae (Fig. 4B). In the green anole, this anionic peptide (i.e. AcBD29) is expressed in all tissues, although its function remains unknown [26]. Similarly, the beta-defensin peptide AcBD27, which presents high similarity to clade-1 peptides from *P. lilfordi* (PlBD1) and *P. muralis* (PmBD1), is known to be highly expressed in wounded lizards, particularly following tail loss, where it plays a role in preventing infection in the absence of inflammation, a key factor in successful regeneration [51, 52]. These findings underscore the need for species-specific studies to clarify whether identical or highly similar beta-defensin peptides or specific polymorphisms arise in parallel and serve functionally equivalent roles across different species.

Convergent evolution of AMPs has been previously reported in both vertebrates and invertebrates, with multiple studies documenting convergence of critical residues under similar antimicrobial pressures across diverse taxa [1, 53]. In anurans, for instance, equivalent signal motifs are thought to have arisen independently in three distinct evolutionary lineages [54]. While the instances of identical peptides observed across our studied reptiles may suggest convergent evolution, establishing a framework of true one-to-one orthologs is essential to rule out alternative explanations such as ancestral polymorphism. Population-level studies are also needed to assess allele diversity and investigate the potential role of balancing selection in polymorphism maintenance [1]. Finally, phenotypic analyses targeting specific residues or motifs will be critical to clarify selective pressures and the molecular mechanisms underlying AMP evolution.

Ovo-defensins and PrAMPs diversity and evolution in Lacertidae

Unlike beta-defensins, ovo-defensins represents a smaller subfamily of still largely understudied defensins, currently described only in birds and a few reptiles (but none in Lacertidae), where they appear to be specifically expressed in the oviduct and contribute to egg immune defense [15, 25, 40].

The eight OVODs identified in *P. lilfordi* were all located within the same chr-3 cluster region, flanked by conserved marker genes, MTMR9 and XKR6, although PIOVOD1 was located immediately outside the gene cluster. This chromosomal arrangement was previously described in other vertebrates [40], supporting a monophyletic origin for this subfamily. We confirmed the two-exon structure and the higher divergence at exon-2, as previously reported for avian and reptilian ovo-defensins

[15]. In contrast to beta-defensins, OVODs present two extra cysteines (eight in total), which contrasts with previous reports of a six-cysteine motif for OVODs in reptiles [15]. Notably, this eight cys-motif and two-exon structure was observed in all reptiles under study (both Lacertidae species and the Komodo dragon), strongly suggesting that these features represent specific structural characteristics of this protein subfamily. Differences in the number and arrangement of cysteines between beta- and ovo-defensins may underscore fundamental functional differences that are currently unclear [4]. Despite structural commonalities and a similar chromosomal arrangement, the eight PIOVODs encompassed a remarkable amino acid diversity, suggesting the potential for further categorization into distinct defensin subfamilies. While ovo-defensins from clade A appear to have undergone an extensive process of gene expansion through duplication in tandem in all *Podarcis* species (Fig. 6), the remaining clades (B-E) show a single ortholog for each species, indicating relative evolutionary stability. We note that presence of OVODs in the viviparous species *Z. vivipara* suggests that protein expression may not be restricted to eggs.

Interestingly, these clades encompass proline-rich AMPs (PrAMPs) [51], containing either multiple non-adjacent prolines (P) (OVOD6) or proline repeats at the C-terminal, typically interspersed with motifs rich in lysine (K), tryptophan (W), histidine (H), arginine (R) and tyrosine (Y) (OVOD7 and 8) (Fig. 5). Their structure and lack of any significant protein folding suggest they may represent “extended AMPs”, a class of AMPs known for containing high proportions of proline (P) or tryptophan (W) residues and lacking regular secondary structure elements [52]. Notably, both P and W-rich peptides are known for their potent antimicrobial activity due to their ability to permeate the outer membrane and additionally target intracellular pathways [51, 55, 56]. For instance, recent studies have shown that the polyproline host defense peptide batenecin 7 (Bac7) can bind to the wall polysaccharide biofilm of hypermucoviscous bacteria, such as *Klebsiella pneumoniae*, actively disrupting the bacteria cell wall [57, 58]. Additionally, KW repeats (observed here in OVOD8) have been shown to enhance AMP bactericidal activity [59], while KY residues (seen in OVOD7) cooperate with H to facilitate DNA cleavage in DNase proteins [60].

Interestingly, we found several instances of cys-motif identity in OVOD-PrAMPs across species, with OVOD8 carrying the same motif in all Lacertidae species. As observed for beta-defensins, this suggests a putative role of convergent evolution in shaping PrAMPs diversity, although alternative scenarios remain to be explored.

Cathelicidins diversity and evolution in Lacertidae

Cathelicidins are a small and ancient family of potent AMPs, known for their functional plasticity and broad spectrum of activity against a wide range of microorganisms [17, 61, 62]. To date, they have been described in several reptile species within the order Squamata [62], including the Komodo dragon [25], but no formal description exists for any Lacertidae species, apart from gene annotation of few representative cathelicidin proteins in public genomes.

In *P. lilfordi*, we identified four cathelicidin proteins (CATH1-4), all of which are tightly clustered on chr-12, flanked by the conserved genes FASTK and KLHL18, previously reported in birds and reptiles [25, 38, 62]. These proteins all exhibit the classical four-exon precursor structure common to vertebrates, characterized by a variable signal peptide, a conserved four-cysteine motif, and a highly divergent C-terminal that carries the active peptide [63–65]. Of the four cathelicidins, three present mature peptides with a highly cationic net charge (> 10), in line with the average positive charge found for reptiles, including the Komodo dragon [61]. Interestingly, CATH1 is the only protein presenting an anionic AMP across all reptiles studied, with a net charge –1 to –2 and a significantly distinct AA composition, suggesting it may have a rather different mechanism of action compared to the other cathelicidins. To date, a single anionic cathelicidin has been formally described and functionally characterized in salamanders, where it is produced by the skin and exhibits potent anti-inflammatory and wound-healing activities, although not direct antimicrobial effects [66]. Functional studies are required to assess whether this cathelicidin plays a similar role in lizard wound healing following injuries or tail autotomy, a conserved antipredatory strategy observed in all lizard species [67].

All Lacertidae species studied presented four cathelicidins, except for *P. muralis* (with five in total), all located on chr-12, with a one-to-one ortholog in all species, supporting a monophyletic origin and phylogenetic conservation within this taxonomic group. Three of these orthologous clades (2–4) clustered into a major, well-supported clade, including also representatives of the Komodo dragon and the green anole, and pointing to a common origin by a past duplication within Squamata. A fifth CATH was identified in *P. muralis*, resulting from a recent duplication in Clade-2-OH. This event indicates that the process of gene duplication in cathelicidins is still ongoing, although it is less pervasive than in defensins [62]. Interestingly, we found that CATH3 and CATH4 were coexpressed as a single large RNA in all Lacertidae, except in *P. lilfordi*, suggesting a further post-translational process to concomitantly generate the two distinct active peptides. We found no previous reports in literature describing such an RNA arrangement for

cathelicidins. However, we observed similar instances of a multidomain structure in beta-defensins (see above), suggesting it might provide some functional gain, unclear at present [48]. In line with this, recent AMP coexpression models in experimental systems have shown enhanced antimicrobial activity and immunity [68, 69], pointing to potential cooperative interactions among naturally coexpressed AMPs, which warrants further investigation.

Conclusions

Our findings substantially expand current knowledge on the diversity of AMPs in reptiles and their mechanisms of evolution, characterized by gene duplications and rapid evolution.

For the Lilford's wall lizard, the identification of key players in innate immunity sets the basis for exploring the role of AMPs in their eco-immunology and the potential responses of these insular populations to rising infection threats associated with increasing anthropogenic impacts on these islands. Healthy *P. lilfordi* populations play significant roles in their ecosystems and a fine understanding of their genetic potential in immune response can offer new insights into resilience mechanisms, guiding species-specific conservation efforts for these isolated populations.

The comprehensive characterization of major AMP classes in *P. lilfordi* will also offer a valuable resource for fine-scale annotation of these peptides in other reptile species, facilitating comparative studies across Squamata and guiding a comprehensive understanding of lineage-specific patterns of gene duplication/loss. Particularly, the identified AMP cysteine motif identity and whole-protein similarity across species lay the groundwork for exploring conserved mechanisms of antimicrobial action in reptiles and assessing their value as candidates for novel therapeutic applications.

Nonetheless, we also emphasize that relying on a single reference genome per studied species currently limits our ability to fully capture intraspecific AMP diversity. This limitation constrains our conclusions regarding the extent of gene duplications and losses, the impact of convergent evolution and the role of balancing-selection in maintaining polymorphisms. Population-level studies will be critical to address these aspects. In this context, future integrative analyses combining population genomics and expression data from multiple island populations of Lilford's wall lizard offer an exciting opportunity to uncover key evolutionary trajectories and functional roles of antimicrobial peptides in lizards.

Abbreviations

AMP	Antimicrobial peptide
Chr	Chromosome
AA	Amino acid
Pl	Podarcis lilfordi

Pm	Podarcis muralis
Pr	Podarcis raffonei
Zv	Zootoca vivipara
Vk	Varanus komodoensis
ORF	Open reading frame
BD	Beta-defensin
OVD	Ovo-defensin
PrAMP	Proline-rich AMP
UPGMA	Unweighted pair group method with arithmetic mean
NJ	Neighbor Joining

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11797-2>.

Supplementary Material 1: Table 1. Beta-defensin peptides and complete proteins from *P. lilfordi*. Table 2. Summary of beta-defensins, ovo-defensins and cathelicidin protein homology results by Phyre2. Table 3. List of AccNos of complete defensin and cathelicidin proteins from published reptile genomes. Newly identified proteins are indicated with 'new'. Table 4. Ovo-defensin complete proteins from *P. lilfordi*. Table 5. Cathelicidin proteins from *P. lilfordi*.

Supplementary Material 2: Figure S1. Multispecies alignment of beta-defensins proteins (partial) from the Lacertidae species under study (Pm, Pr, Pl and Zv). Figure S2. Multispecies alignment of ovo-defensins proteins (complete) from the Lacertidae species under study and the Komodo dragon (Pm, Pr, Pl, Zv and Vk). Figure S3. Multispecies alignment of cathelicidins proteins (partial) from Lacertidae species, the Komodo dragon and the green anole lizard (Pm, Pr, Pl, Zv, Vk and Ac).

Supplementary Material 3: AMP fasta sequences from *P. lilfordi*.

Acknowledgements

Institutional support to CNAG was given from the Spanish Ministry of Science, Innovation and Universities, Fondo de Investigaciones Sanitarias cofunded with ERDF funds (PI19/01772), the 2014–2020 Smart Growth Operating Program, and the Generalitat de Catalunya through the Departament de Recerca i Universitats and Departament de Salut.

Authors' contributions

L.B. conceived the study. L.B., K.O. and J.G. analyzed the results. L.B. and K.O. wrote the paper. All authors read and approved the final manuscript.

Funding

This study was funded by MICIU/AEI/ 10.13039/501100011033 and by "ERDF A way of making Europe", grant PID2022-141578NB-C22 to L.B. Doctoral support to KO was from the Doctoral Scholarship, Colombian Ministry of Science, Technology and Innovation (MINCIENCIAS885/2020 to KO).

Data availability

All sequence data analysed in this study are included in the Supplementary Tables, along with the accession numbers (AccNos) for published sequences. All complete nucleotide and protein sequences newly described for *P. lilfordi* have been uploaded at European Nucleotide Archive (ENA) under AccNo PRJEB89980. All complete and partial *P. lilfordi* AMPs are also available in the additional file.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Evolutionary Biology, Ecology and Environmental Sciences, Biology Faculty, University of Barcelona (UB), Av. Diagonal 643, Barcelona 08028, Spain

²Fundación FUNMAJO, RAIEC, Biodiversity Branch, EBA, Cucaita, Boyacá, Colombia

³Centro Nacional de Análisis Genómico (CNAG), Baldiri Reixac 4, Barcelona 08028, Spain

⁴Universitat de Barcelona (UB), Barcelona, Spain

⁵Institute for Research On Biodiversity (IRBio), University of Barcelona (UB), Barcelona, Spain

Received: 23 March 2025 / Accepted: 6 June 2025

Published online: 21 July 2025

References

- Lazzaro BP, Zasloff M, Rolff J. Antimicrobial peptides: application informed by evolution. *Science*. 2020;368:eau5480.
- Huan Y, Kong Q, Mou H, Yi H. Antimicrobial peptides: classification, design, application and research progress in multiple fields. *Front Microbiol*. 2020;11:582779.
- Chakraborty S, Chatterjee R, Chakravorty D. Evolving and assembling to pierce through: evolutionary and structural aspects of antimicrobial peptides. *Comput Struct Biotechnol J*. 2022;20:2247–58.
- Erdem Büyükkiraz M, Kesmen Z. Antimicrobial peptides (AMPs): a promising class of antimicrobial compounds. *J Appl Microbiol*. 2022;132:1573–96.
- Cardoso MH, Meneguetti BT, Oliveira-Júnior NG, Macedo MLR, Franco OL. Antimicrobial peptide production in response to gut microbiota imbalance. *Peptides*. 2022;157:170865.
- Zhang LJ, Gallo RL. Antimicrobial peptides. *Curr Biol*. 2016;26:R14–9.
- Harris F, Dennison SR, Phoenix DA. Anionic antimicrobial peptides from eukaryotic organisms. *Curr Protein Pept Sci*. 2009;10:585–606.
- Semple CAM, Maxwell A, Gautier P, Kilanowski FM, Eastwood H, Barran PE, et al. The complexity of selection at the major primate beta-defensin locus. *BMC Evol Biol*. 2005;5:32.
- Hanson MA. When the microbiome shapes the host: immune evolution implications for infectious disease. *Philos Trans R Soc B Biol Sci*. 2024;379:20230061.
- Benítez-Prián M, Lorente-Martínez H, Agorreta A, Gower DJ, Wilkinson M, Roelants K, et al. Diversity and molecular evolution of antimicrobial peptides in Caecilian Amphibians. *Toxins*. 2024;16:150.
- Semple CA, Gautier P, Taylor K, Dorin JR. The changing of the guard: Molecular diversity and rapid evolution of beta-defensins. *Mol Divers*. 2006;10:575–84.
- Semple F, Dorin JR. β -Defensins: multifunctional modulators of infection, inflammation and more? *J Innate Immun*. 2012;4:337–48.
- Zhu & Gao. Evolutionary origin of β -defensins. *Dev Comp Immunol*. 2013;39:79–84.
- Tu J, Li D, Li Q, Zhang L, Zhu Q, Gaur U, et al. Molecular evolutionary analysis of β -defensin peptides in vertebrates. *Evol Bioinformatics Online*. 2015;11:105.
- Whenham N, Lu TC, Maidin MBM, Wilson PW, Bain MM, Stevenson ML, et al. Ovoidensins, an oviduct-specific antimicrobial gene family, have evolved in birds and reptiles to protect the egg by both sequence and intra-six-cysteine sequence motif spacing. *Biol Reprod*. 2015;92:154.
- Gong D, Wilson PW, Bain MM, McDade K, Kalina J, Hervé-Grépinet V, et al. Gallin, an antimicrobial peptide member of a new avian defensin family, the ovoidensins, has been subject to recent gene duplication. *BMC Immunol*. 2010;11:1–15.
- Kościczkuk EM, Lisowski P, Jarczak J, Strzałkowska N, Jóźwik A, Horbańczuk J, et al. Cathelicidins: family of antimicrobial peptides. A review. *Mol Biol Rep*. 2012;39:10957–70.
- Bals R, Wilson JM. Cathelicidins - a family of multifunctional antimicrobial peptides. *Cell Mol Life Sci*. 2003;60:711–20.
- Tomasinsig L, Zanetti M. The cathelicidins—structure, function and evolution. *Curr Protein Pept Sci*. 2005;6:23–34.
- Wang G, Li X, Wang Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res*. 2016;44:D1087–93.
- Field EK, Hartzheim A, Terry J, Dawson G, Haydn N, Neuman-Lee LA. Reptilian innate immunology and ecoinmunology: what do we know and where are we going? *Integ Comp Biol*. 2022;62:1557–71.

22. Buchmann K. Evolution of innate immunity: clues from invertebrates via fish to mammals. *Front Immunol.* 2014;5: 108394.
23. Santana FL, Estrada K, Ortiz E, Corzo G. Reptilian β -defensins: expanding the repertoire of known crocodylian peptides. *Peptides.* 2021;136:170473.
24. Siddiqui R, Madver SK, Khan NA. Gut microbiome-immune system interaction in reptiles. *J Appl Microbiol.* 2022;132:2558–71.
25. van Hoek ML, Prickett MD, Settlage RE, Kang L, Michalak P, Miet KA, et al. The Komodo dragon (*Varanus komodoensis*) genome and identification of innate immunity genes and clusters. *BMC Genomics.* 2019;20:1–18.
26. Dalla Valle L, Benato F, Maistro S, Quinzani S, Alibardi L. Bioinformatic and molecular characterization of beta-defensin-like peptides isolated from the green lizard *Anolis carolinensis*. *Dev Comp Immunol.* 2012;36:222–9.
27. Zhao H, Gan TX, Liu XD, Jin Y, Lee WH, Shen JH, et al. Identification and characterization of novel reptile cathelicidins from elapid snakes. *Peptides.* 2008;29:1685–91.
28. Kim D, Soundararajan N, Lee J, Cho HS, Choi M, Cha SY, et al. Genomewide analysis of the antimicrobial peptides in python *bivittatus* and characterization of cathelicidins with potent antimicrobial activity and low cytotoxicity. *Antimicrob Agents Chemother.* 2017;61:e00530–17.
29. Andrade P, Pinho C, Pérez I, de Lanuza G, Afonso S, Brejcha J, Rubin CJ, et al. Regulatory changes in pterin and carotenoid genes underlie balanced color polymorphisms in the wall lizard. *Proc Natl Acad Sci USA.* 2019;116:5633–42.
30. Gabrielli M, Benazzo A, Biello R, Ancona L, Fuselli S, Iannucci A, et al. A high-quality reference genome for the critically endangered Aeolian wall lizard, *Podarcis raffonei*. *J Heredity.* 2023;114:279–85.
31. Yurchenko AA, Recknagel H, Elmer KR. Chromosome-level assembly of the common lizard (*Zootoca vivipara*) genome. *Genome Biol Evol.* 2020;12:1953–60.
32. Gomez-Garrido J, Cruz F, Alioto TS, Feiner N, Uller T, Gut M, et al. Chromosome-level genome assembly of Lilford's wall lizard, *Podarcis lilfordi* (Günther, 1874) from the Balearic Islands (Spain). *DNA Res.* 2023;30:dsad008.
33. Rotger A, Igual JM, Smith JJ, Tavecchia G. Relative role of population density and climatic factors in shaping the body growth rate of Lilford's Wall Lizard (*Podarcis lilfordi*). *Can J Zool.* 2016;94:207–15.
34. Otalora K, Riera JL, Tavecchia G, Rotger A, Igual JM, Trotta JRP, et al. Population genetics and phylogeographic history of the insular lizard (Günther, 1874) from the Balearic Islands based on genome-wide polymorphic data. *Ecol Evol.* 2024;14:e11407.
35. Baldo L, Riera JL, Mitsi K, Pretus JL. Processes shaping gut microbiota diversity in allopatric populations of the endemic lizard *Podarcis lilfordi* from Menorcan islets (Balearic Islands). *FEMS Microbiol Ecol.* 2018;94:fx186.
36. Baldo L, Tavecchia G, Rotger A, Igual JM, Riera JL. Insular holobionts: persistence and seasonal plasticity of the Balearic wall lizard (*Podarcis lilfordi*) gut microbiota. *PeerJ.* 2023;11:e14511.
37. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215:403–10.
38. Cheng Y, Prickett MD, Gutowska W, Kuo R, Belov K, Burt DW. Evolution of the avian β -defensin and cathelicidin genes. *BMC Evol Biol.* 2015;15:188.
39. van Hoek ML. Antimicrobial peptides in reptiles. *Pharmaceuticals (Basel).* 2014;7:723–53.
40. Zhang L, Chen D, Yu L, Wei Y, Li J, Zhou C. Genome-wide analysis of the ovodefensin gene family: monophyletic origin, independent gene duplication and presence of different selection patterns. *Infect Genet Evol.* 2019;68:265–72.
41. Slater GSC, Birney E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics.* 2005;6:1–11.
42. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015;10:845–58.
43. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32:1792–7.
44. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinform (Oxford, England).* 2009;25:1189–91.
45. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 1987;4(4):406–25.
46. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol.* 2007;24:1596–9.
47. Nava GM, Escorcia M, Castañeda MP. Molecular diversity of the antimicrobial domain of beta-defensin 3 and homologous peptides. *Comp Funct Genomics.* 2009;2009: 83636.
48. Guyot N, Landon C, Monget P. The two domains of the Avian double- β -defensin AvBD11 have different ancestors, common with potential monodomain crocodile and turtle defensins. *Biol.* 2022;11:690.
49. Tennessen JA. Molecular evolution of animal antimicrobial peptides: widespread moderate positive selection. *J Evol Biol.* 2005;18:1387–94.
50. Yang W, Feiner N, Salvi D, Laakkonen H, Jablonski D, Pinho C, et al. Population genomics of wall lizards reflects the dynamic history of the Mediterranean Basin. *Mol Biol Evol.* 2022;39:msab311.
51. Lai P-K, Tresnak DT, Hackel BJ. Identification and elucidation of proline-rich antimicrobial peptides with enhanced potency and delivery. *Biotechnol Bioeng.* 2019;116:2439–50.
52. Nguyen LT, Haney EF, Vogel HJ. The expanding scope of antimicrobial peptide structures and their modes of action. *Trends Biotechnol.* 2011;29:464–72.
53. Unckless RL, Lazzaro BP. The potential for adaptive maintenance of diversity in insect antimicrobial peptides. *Philos Trans R Soc Lond B Biol Sci.* 2016;371:20150291.
54. König E, Bininda-Emonds ORP. Evidence for convergent evolution in the antimicrobial peptide system in anuran amphibians. *Peptides.* 2011;32:20–5.
55. Mishra AK, Choi J, Moon E, Baek K-H. Tryptophan-rich and proline-rich antimicrobial peptides. *Mol.* 2018;23:815.
56. Hernandez-Gordillo V, Geisler I, Chmielewski J. Dimeric unnatural polyproline-rich peptides with enhanced antibacterial activity. *Bioorg Med Chem Lett.* 2013;24:556.
57. De Los Santos L, Beckman RL, DeBarro C, Keener JE, Torres MDT, de la Fuente-Nunez C, et al. Polyproline peptide targets *Klebsiella pneumoniae* polysaccharides to collapse biofilms. *Cell Rep Phys Sci.* 2024;5:101869.
58. Fleeman RM, Davies BW. Polyproline peptide aggregation with *Klebsiella pneumoniae* extracellular polysaccharides exposes biofilm associated bacteria. *Microbiol Spectr.* 2022;10:e0202721.
59. Gopal R, Seo CH, Song PI, Park Y. Effect of repetitive lysine-tryptophan motifs on the bactericidal activity of antimicrobial peptides. *Amino Acids.* 2013;44:645–60.
60. Korn C, Scholz SR, Gimadutdinov O, Pingoud A, Meiss G. Involvement of conserved histidine, lysine and tyrosine residues in the mechanism of DNA cleavage by the caspase-3 activated DNase CAD. *Nucleic acids research.* 2002;30:1325–32.
61. Alford MA, Baquir B, Santana FL, Haney EF, Hancock REW. Cathelicidin host defense peptides and inflammatory signaling: striking a balance. *Front Microbiol.* 2020;11:1902.
62. Tossi A, Gerdol M, Caporale A, Pacor S, Mardirossian M, Scocchi M, et al. Cathelicidins—a rich seam of antimicrobial peptides waiting for exploitation. *Front Drug Discov (Lausanne).* 2024;4.
63. Izadpanah A, Gallo RL. Antimicrobial peptides. *J Am Acad Dermatol.* 2005;52(3 Pt 1):381–90. quiz 391–2.
64. Gennaro R, Zanetti M. Structural features and biological activities of the cathelicidin-derived antimicrobial peptides. *Biopolymers.* 2000;55:31–49.
65. Zhong L, Liu J, Teng S, Xie Z. Identification of a novel cathelicidin from the genome with antibacterial activity by multiple mechanisms. *Toxins (Basel).* 2020;12:771.
66. Luo X, Ouyang J, Wang Y, Zhang M, Fu L, Xiao N, et al. A novel anionic cathelicidin lacking direct antimicrobial activity but with potent anti-inflammatory and wound healing activities from the salamander *Typhlotriton kweichowensis*. *Biochimie.* 2021;191:37–50.
67. Alibardi L. Ultrastructural immunolocalization of cathelicidin-like peptides in granulocytes of normal and regenerating lizard tissues. *Acta Histochem.* 2014;116:363–71.
68. Chen J, Peng J, Ma C, Zhang L, Wu X, Wei H, et al. Co-expression of Pig IL-2 and fusion bovine cathelicidin gene by recombinant plasmids in yeast and their promotion of mouse antibacterial defense. *Biology (Basel).* 2022;11:1491.
69. Kuddus MR, Yamano M, Rumi F, Kikukawa T, Demura M, Aizawa T. Enhanced expression of cysteine-rich antimicrobial peptide snakin-1 in *Escherichia coli* using an aggregation-prone protein coexpression system. *Biotechnol Prog.* 2017;33:1520–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com