Building a Portuguese Coalition for Biodiversity Genomics

João P. Marques[a,1,2], Paulo C. Alves[a,1,2,3,4], Isabel R. Amorim[a,5], Ricardo J. Lopes[a,1,2,6], Mónica Moura[a,2,7,8], Eugene W. Myers[a,9,10], Manuela Sim-Sim[a,11,12,13], Carla Sousa-Santos[a,14], M. Judite Alves[11,12], Paulo A. V. Borges[5], Thomas Brown[9], Miguel Carneiro[1,2], Carlos Carrapato[15,16], Luís M. P. Ceríaco[1,12,17], Claúdio Ciofi[18], Luís P. da Silva[1,2], Genevieve Diedericks[19], Maria A. Diroma[18], Liliana Farelo[1,2], Giulio Formenti[20], Fátima Gil[21], Miguel Grilo[22,23], Henrique G. Leitão[19], Alessio Ianucci[18], Cristina Máguas[11,12], Ann M. Mc Cartney[24], Sofia L. Mendes[11,25], João M. Moreno[11,25,26], Marco Morselli[27], Alice Mouton[18,28], Chiara Natali[18], Fernando Pereira[5], Rúben M. C. Rego[7], Roberto Resendes[7], Guilherme Roxo[7], Hannes Svardal[19,29], Helena Trindade[11,12], Sara Vicente[11,12,30], Sylke Winkler[9,31], Marcela Alvarenga[1,2,3], Andreia J. Amaral[32,33,34], Agostinho Antunes[3,35], Paula F. Campos[35], Adelino V. M. Canário[36], Rita Castilho[36,37], L. Filipe C. Castro[3,35], Angelica Crottini[1,2,3], Mónica V. Cunha[11,12,38], Gonçalo Espregueira Themudo[11,35], Pedro J. Esteves[1,2,3], Rui Faria[1,2], Carlos Rodríguez Fernandes[11,39], Jean-Baptiste Ledoux[35], Bruno Louro[36], Sara Magalhães[11,25], Octávio S. Paulo[11,25], Gareth Pearson[36], João Pimenta[1,2], Francisco Pina-Martins[11,25], Teresa L. Santos[11,25], Ester Serrão[36], José Melo-Ferreira[b,1,2,3] and Vítor C. Sousa[b,11,25]

[a] shared first authorship

[b] shared senior authorship

Corresponding authorship: João P. Marques (joao.marques@cibio.up.pt), José Melo-Ferreira (jmeloferreira@cibio.up.pt) and Vitor C. Sousa (vmsousa@ciencias.ulisboa.pt)

Affiliations

[1] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, 4485-661 Vairão, Portugal

[2] BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão,

Portugal
[3] Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, 4099-002 Porto, Portugal
[4] EBM - Estação Biológica de Mértola,, 7750-329 Mértola, Portugal
[5] cE3c-Centre for Ecology, Evolution and Environmental Changes, Azorean Biodiversity Group, CHANGE – Global Change and Sustainability Institute, and University of the Azores, Rua Capitão João d´Ávila, Pico da Urze, 9700-042, Angra do Heroísmo, Portugal
[6] MHNC-UP, Natural History and Science Museum of the University of Porto, 4099-002 Porto, Portugal
[7] CIBIO, Research Centre in Biodiversity and Genetic Resources, InBIO Associated Laboratory, Azores Unit, University of the Azores, Campus of Ponta Delgada, Rua da Mãe de Deus 13A, 9500-321 Ponta Delgada, Azores, Portugal
[8] UNESCO Chair – Land Within Sea: Biodiversity & Sustainability in Atlantic Islands
[9] Max Planck Institute of Molecular Cell Biology and Genetics Pfotenhauerstraße 108 01307 Dresden, Germany
[10] Okinawa Institute of Science and Technology, Okinawa, Japan
[11] cE3c - Centre for Ecology, Evolution and Environmental Changes & CHANGE - Global Change and Sustainability Institute, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
[12] MUHNAC - Museu Nacional de História Natural e da Ciência, Universidade de Lisboa, Rua da Escola Politécnica, 58, 1250-102 Lisboa, Portugal
[13] Departamento de Biologia Vegetal, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
[14] MARE-Marine and Environmental Sciences Centre/ARNET-Aquatic Research Network, ISPA-Instituto Superior de Ciências Psicológicas, Sociais e da Vida, Rua Jardim do Tabaco, 34, 1149-041 Lisbon, Portugal.
[15] Instituto da Conservação da Natureza e das Florestas, IP, Portugal
[16] DCNF-Alentejo/DLAP/Parque Natural do Vale do Guadiana
[17] Universidade Federal do Rio de Janeiro, Museu Nacional, Departamento de Vertebrados, Quinta da Boa Vista, São Cristóvão, 20940-040 Rio de Janeiro, RJ, Brasil
[18] Department of Biology, University of Florence, 50019 Sesto Fiorentino (FI), Italy
[19] Department of Biology, University of Antwerp, 2610 Antwerp, Belgium
[20] The Vertebrate Genome Laboratory, The Rockefeller University, 1230 York Ave, New York City, 10021 NY, USA
[21] Aquário Vasco da Gama, R. Direita do Dafundo, 1495-718 Cruz Quebrada – Dafundo, Portugal
[22] MARE-Marine and Environmental Sciences Centre/ARNET-Aquatic Research Network, ISPA-Instituto Superior de Ciências Psicológicas, Sociais e da Vida, Rua Jardim do Tabaco, 34, 1149-041 Lisbon, Portugal.
[23] Egas Moniz Center for Interdisciplinary Research (CiiEM), Egas Moniz School of Health and Science, 2829-

511 Caparica, Almada, Portugal
[24] Genomics Institute, University of California, Santa Cruz, CA 95060 USA
[25] Departamento de Biologia Animal, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
[26] MARE – Centro de Ciências do Mar e do Ambiente (MARE), ARNET—Aquatic Research Network,
[27] Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parco Area delle Scienze 23/A, 43124 Parma, Italy
[28] InBios/Conservation Genetics Laboratory, University of Liege, Chemin de la Vallée 4, 4000 Liege, Belgium
[29] Naturalis Biodiversity Center, 2333 CR Leiden, The Netherlands
[30] ERISA – Escola Superior de Saúde Ribeiro Sanches, Instituto Politécnico da Lusofonia (IPLUSO), Rua do Telhal aos Olivais nº 8 – 8a, 1950-396 Lisboa, Portugal
[31] DRESDEN concept genome center, c/o Max Planck Institute of Molecular Cell Biology and Genetics Pfotenhauerstraße 108 01307 Dresden, Germany
[32] Departamento de Zootecnia, Escola Superior de Ciência e Tecnologia, Universidade de Évora, 7000-671 Évora, Portugal
[33] CIISA - Centre for Interdisciplinary Research in Animal Health, Faculty of Veterinary Medicine, University of Lisbon, Alto da Ajuda, 1300-477 Lisbon, Portugal
[34] AL4AnimalS - Associate Laboratory for Animal and Veterinary Sciences, Portugal
[35] CIIMAR/CIMAR, Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Porto, Portugal
[36] CCMAR/CIMAR LA Centre for Marine Sciences, Campus de Gambelas, University of Algarve, 8005-139 Faro, Portugal
[37] Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal
[38] BioISI- Biosystems and Integrative Sciences Institute Campus da Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal.
[39] Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, 1649-013 Lisboa, Portugal

Abstract

The diverse physiography of the Portuguese land and marine territory, spanning from continental Europe to the Atlantic archipelagos, has made it an important repository of biodiversity throughout the Pleistocene glacial cycles, leading to a remarkable diversity of species and ecosystems. This rich biodiversity is under threat from anthropogenic drivers, such as climate change, invasive species, land use changes, overexploitation or pathogen (re)emergence. The inventory, characterization and study of biodiversity at inter- and intra-specific levels using genomics is crucial to promote its preservation and recovery by informing biodiversity conservation policies, management measures and research. The participation of researchers from Portuguese institutions in the European Reference Genome Atlas (ERGA) initiative, and its pilot effort to generate reference genomes for European biodiversity, has reinforced the establishment of Biogenome Portugal. This nascent institutional network will connect the national community of researchers in genomics. Here, we describe the Portuguese contribution to ERGA's pilot effort, which will generate high-quality reference genomes of six species from Portugal that are endemic, iconic and/or endangered, and include plants, insects and vertebrates (fish, birds and mammals) from mainland Portugal or the Azores islands. In addition, we outline the objectives of Biogenome Portugal, which aims to (i) promote scientific collaboration, (ii) contribute to advanced training, (iii) stimulate the participation of institutions and researchers based in Portugal in international biodiversity genomics initiatives, and (iv) contribute to the transfer of knowledge to stakeholders and engaging the public to preserve biodiversity. This initiative will strengthen biodiversity genomics research in Portugal and fuel the genomic inventory of Portuguese eukaryotic species. Such efforts will be critical to the conservation of the country's rich biodiversity and will contribute to ERGA's goal of generating reference genomes for European species.

Keywords

Biodiversity, Genomics, Conservation, Portugal

Sumário

A diversa fisiografia dos territórios terrestre e marinho portugueses, que se estende desde a Europa continental até aos arquipélagos do Atlântico, promoveu a sua importância como repositório de biodiversidade ao longo dos ciclos glaciares do Pleistoceno, resultando numa notável diversidade de espécies e ecossistemas. No entanto, esta biodiversidade enfrenta ameaças decorrentes de atividades humanas, tais como as alterações climáticas, invasão de espécies, mudanças no uso do solo, exploração excessiva e (re)emergência de agentes patogénicos. A fim de promover a preservação e recuperação da biodiversidade, é crucial inventariar, caracterizar e estudar a biodiversidade a níveis inter e intraespecíficos, utilizando técnicas genómicas. Essa informação é essencial para informar políticas de conservação da biodiversidade, medidas de gestão e pesquisas futuras. A participação de investigadores de instituições portuguesas na iniciativa European Reference Genome Atlas (ERGA), e no projeto-piloto para gerar genomas de referência para a biodiversidade europeia, promoveu a criação do Biogenoma Portugal. Esta rede institucional em desenvolvimento visa conectar a comunidade nacional de investigadores em genómica. Neste contexto, descrevemos a contribuição portuguesa para o projeto-piloto do ERGA, que tem como objetivo gerar genomas de referência de alta qualidade de seis espécies de Portugal. Estas espécies são endémicas, icónicas e/ou ameaçadas e incluem plantas, insetos e vertebrados (peixes, aves e mamíferos) do território continental de Portugal e das ilhas dos Açores. Adicionalmente, delineamos os objetivos do Biogenoma Portugal que visa: (i) promover a colaboração científica, (ii) contribuir para a formação avançada, (iii) estimular a participação de instituições e investigadores baseados em Portugal em iniciativas internacionais de genómica da biodiversidade e (iv) contribuir para a transferência de conhecimento e envolvimento do público em geral na preservação da biodiversidade. Esta iniciativa fortalecerá a investigação em genómica da biodiversidade em Portugal e contribuirá para o objetivo do ERGA de gerar genomas de referência para espécies europeias. Estes esforços terão um papel fundamental na conservação da biodiversidade do país e na promoção da conscientização sobre a importância de preservar a diversidade biológica.

Introduction

The remarkable diversity of life on Earth is essential for maintaining ecosystem stability, fostering ecological interactions among species, regulating the climate, and providing fundamental resources that sustain human well-being and promote the development of society. Today, Earth's biodiversity, ecological and evolutionary heritage, and the ecosystem services they provide are under intense and extensive strain as a result of many direct and indirect anthropogenic activities. These major activities include habitat destruction and fragmentation caused by agricultural expansion and other forms of land conversion, overexploitation of natural resources, overharvesting, illegal killing and trading of wild species, climate changes, environmental pollution, and global spread of exotic species (Ceballos et al., 2017; Cowie et al., 2022; Habel et al., 2019; Karger et al., 2021; O'Hara et al., 2021). These pressures not only result in the decline and extinction of populations, species and ecosystems, but also have cascading effects on various dimensions of biodiversity, including genetic, functional or phenological diversity and the interactions within biodiversity networks. These impacts harm the health of natural populations and their survival, and reduce their potential to adapt to environmental challenges (Dauphin et al., 2021). Characterising and understanding in detail the genetic variation underlying biodiversity at all levels provides essential modern tools to delineate sound and efficient strategies to mitigate the impacts of human activities, design management plans for conservation and restoration, and support data-driven biodiversity policy development (Browne et al., 2019; Hohenlohe et al., 2021; Ralls et al., 2020; Segelbacher et al., 2022). Regional biodiversity genomics initiatives that promote research and applications under the umbrella of international coalitions, can play an important role in this global challenge.

Europe hosts a remarkable biodiversity richness despite its temperate climate and historically anthropogenic landscapes (Ette & Geburek, 2021), and has an important biodiversity hotspot along the Mediterranean basin (Habel et al., 2019; Myers et al., 2000). Most of the continental territory of Portugal, located in south-western Europe in the Iberian Peninsula, is part of the Mediterranean biodiversity hotspot (Myers et al., 2000). The country also includes the volcanic Atlantic archipelagos of Azores, Madeira and Selvagens (Figure 1). Even though Portugal is the 19th largest European country (land area ~92000 km^2), it holds the 5th largest marine territory (~1.7 million km^2) (Pauly et al., 2020). As a result of its unique geographic location, climate and geophysical features, Portugal comprises a variety of landscapes and ecosystems and harbours a large diversity of species. The Iberian Peninsula provided refugia for species throughout the glacial periods and remains one of the most important repositories of biodiversity in Europe, both at inter- and intraspecific level (Gómez & Lunt, 2007; Hewitt, 2004). In addition, the geographic isolation of the archipelagos of Azores, Madeira and Selvagens, which together with the Canary and Cabo Verde archipelagos form the biogeographical unit of Macaronesia, has led to the divergence and speciation of many lineages that resulted in many endemisms (Florencio et al., 2021). According to the Portuguese National Strategy for the Conservation of Nature and Biodiversity 2030 (ENCNB 2030; https://dre.pt/dre/detalhe/resolucao-conselho-ministros/55-2018-115226936), Portugal holds ~35,000 animal and plant species, representing a relevant proportion of these species groups present in Europe. The IUCN Red List database (https://www.iucnredlist.org; 27 April 2023) lists ~800 of the assessed species as endemic to the continental mainland or archipelagos of Portugal. There are 414 protected areas in Portugal encompassing 22.28% of the land and 2.46% of marine water, and 404 species and 102 habitats are protected under EU law (Biodiversity Information System for Europe: https://biodiversity.europa.eu/countries/portugal). The ENCNB 2030 recognizes the importance of a systematic inventory and characterization of biodiversity at all levels – ecosystems, species and genetic diversity – to properly anchor conservation strategies. Genomics provides unprecedented power to understand biodiversity at the inter- and intraspecific levels (Formenti et al., 2022; Theissinger et al., 2023). Thus, gathering scientific expertise in the area, coordinating research efforts and promoting genomic studies of native Portuguese species, can give an invaluable contribution to preserving the rich biodiversity of Portugal and its ecosystem services.

ERGA-Portugal community and the Pilot effort

The European Reference Genome Atlas (ERGA; https://www.erga-biodiversity.eu/) is an ambitious and ground-breaking initiative that aims to revolutionise our understanding of biodiversity by creating a comprehensive genomic resource for European species. ERGA is a hub for researchers, scientists, and biodiversity enthusiasts to access and explore a vast collection of genomic data. By collecting and sequencing the genomes of various organisms across Europe, ERGA aims to provide a valuable tool for studying and preserving biodiversity, understanding evolutionary processes, and addressing pressing environmental challenges. ERGA leverages the power of genomics to unlock the secrets of European biodiversity and foster collaborative efforts towards its conservation and sustainable management.

ERGA is a community of researchers, which in Portugal includes more than 50 members distributed across seven research units, five universities and two natural history museums (May 31st 2023; Figure 1). This community of ERGA members based in Portugal is called ERGA-Portugal. ERGA-Portugal covers a wide range of expertise that is relevant for biodiversity genomics research, from taxonomy and systematics across diverse biological groups (e.g., microorganisms, plants, insects, fish, amphibians, birds or mammals), to comparative and population genomics and bioinformatics. This community communicates through open meetings, mailing lists and message sharing channels to facilitate the exchange of biodiversity genomic expertise across Portugal. Many Portuguese community members are also active participants in ERGA's scientific committees.

To contribute to the establishment and mission of ERGA, ERGA-Portugal participated in the pilot test of ERGA. This pilot aimed at testing and developing the biodiversity genomic networks across Europe through the sequencing of several European species, with the contribution of expert ERGA members and European institutions and without a centralised source of funding. In Portugal, the goal of this project was to consolidate the ERGA-Portugal community, and initiate collective efforts to generate reference genomes for Portuguese species. This pilot effort in Portugal focused on endemic, endangered and iconic species from several taxonomic groups and different ecosystems in both the mainland and the Azores islands. From an initial list of 53 species proposed by members of ERGA-Portugal, 11 were prioritised based on ERGA's feasibility criteria defined by the ERGA Sampling and Sample Processing Committee (Böhne et al., 2023). These species were then ranked based on an anonymous online survey open to all members of ERGA-Portugal at the time. The six species with the highest score were included in the pilot effort. This shortlist included one fish, one mammal, one bird, one invertebrate and two plants (Figure 2).

Each species was championed by a genome team, led by one or two species ambassadors, and included ERGA-Portugal members and other ERGA members with transdisciplinary skills, who were responsible for successfully implementing each step for generating a reference genome. These steps included, e.g. sampling, taxonomic identification, vouchering, laboratory work, sequencing, assembly, annotation and downstream analysis. These six genome teams, detailed below for each species, promoted new national and international collaborations. The implementation of this project also promoted interactions with national authorities, for example, to obtain permits for capture, storage and export of samples, in coordination with the Portuguese National authority on Nature Conservation (Institute for Nature Conservation and Forests - ICNF), and the Azorean Regional Directorate for the Environment and Climate Change (DRAAC). For mainland Portugal, the capture and collection of samples of wildlife specimens were authorised by the ICNF, while for the Autonomous regions of Azores permission was given by the DRAAC. For the Azores, collections complied with the Access and Benefit Sharing provisions codified in the Nagoya Protocol (Convention on Biological Diversity, 2010). For all species, the sampling steps followed the guidelines provided by the ERGA SSP committee (Böhne et al., 2023). Sampling was optimised to ensure that High Molecular Weight (HMW) DNA and RNA could be obtained from at least one individual and ideally from multiple tissues (see details for each species below). The immediate preservation and transport of tissues in liquid nitrogen was prioritised and tissues were subsequently maintained at -80ºC conditions.

ERGA-Portugal pilot species

Iberian minnowcarp - *Anaecypris hispanica* (Steindachner, 1866)

Common name in portuguese: saramugo

The **Iberian minnowcarp** (*Anaecypris hispanica*) is a strictly freshwater fish with a short lifespan that inhabits a restricted geographical area in two river basins of the Southern Iberian

Peninsula (Cardoso, 2023; De Miguel et al., 2010), a region which is severely affected by the effects of global climate change. This endemic species is considered to be the most endangered strictly freshwater fish in Iberian rivers, and is listed as Endangered on the IUCN Red List (Crivelli, 2006) and as Critically Endangered on the Portuguese Red List (Cabral et al., 2005). These fish live in vulnerable freshwater ecosystems and are exposed to multiple threats, including increased temperatures and propensity for drought, pollution, habitat fragmentation, dams, intensive water use and invasive species proliferation (Cabral et al., 2005; Cardoso, 2023). The Iberian minnowcarp is considered an iconic species as it is the only living member of the genus *Anaecypris*, representing a phylogenetically unique old lineage that persisted in the Iberian Peninsula (Perea et al., 2010). A high-quality reference genome will open the door to reconstruct the evolutionary history of this monospecific genus, elucidating its relationship with other fish lineages. Given that the Iberian minnowcarp is fragmented into small populations, inhabiting a region highly impacted by drought and water scarcity (Sousa-Santos et al., 2014), the availability of its reference genome will be crucial to understand adaptation to intermittent environments, to inform conservation management, and predict responses to increasing water temperatures. The Iberian minnowcarp is included in ongoing *ex-situ* conservation programs, hence these genomic resources will allow quantifying the loss of genetic diversity due to consecutive generations bred from small initial stocks of wild breeders, as well as optimise and monitor future *ex-situ* efforts.

Sampling and sample processing: An adult female was collected by hand net from an open-air breeding tank at Aquário Vasco da Gama (Cruz Quebrada-Dafundo, Portugal), where an *ex-situ* conservation program is ongoing with captive-born fish derived from 37 individuals of River Chança (Guadiana river basin, Portugal). The individual was sampled, euthanized and dissected on the same day, according to permits from the ICNF (Permit P-026382/2021) and animal welfare regulations of the host institution (ORBEA-MARE 02/2021). The head of the fish was flash-frozen and used for biobanking. Fin clips were used for barcoding and preserved in ethanol. Fin clips were also used to derive cell lines, which were flash-frozen and can be used for karyotyping and further genetic studies in the future. Tissues were collected immediately and flash-frozen in liquid nitrogen. Liver, spleen, muscle, eyes and ovary tissues were collected.

Sequencing and assembly: Aliquots of these tissues were shipped to the Max Planck Institute of Molecular Cell Biology and Genetics (MPI-CBG, Germany), where the remaining steps for DNA extraction and sequencing were performed. The HMW DNA extraction and library preparation for PacBio HiFi was done using the spleen, whereas Hi-C library preparation for Illumina was done using muscle. RNA-Sequencing and Iso-Seq libraries from eye and gonad tissues were produced and sequenced. Sequencing data was used to obtain a genome assembly, which was done and completed by the genome team members from MPI-CBG (Germany) using the VGP standards.

Iberian hare - *Lepus granatensis* (Rosenhauer, 1856)

Common name in Portuguese: lebre Ibérica

The **Iberian hare** (*Lepus granatensis*), also known as Granada hare, is a lagomorph species endemic to the Iberian Peninsula, and the only hare species present in Portugal (Alves et al., 2023). It is an iconic species of major ecological and economic importance in the Iberian ecosystems, acting as an important prey to apex predators and a relevant small game species in Portugal and Spain. Genetic studies have promoted this species as an outstanding model to understand diverse evolutionary questions, such as the impact of glacial cycles in repeated changes in species ranges, and the influence of introgressive hybridization on their adaptive potential (Seixas et al., 2018). It is currently classified as Least Concern in the IUCN Red List (Soriguer & Carro, 2018), but marked decline over the last decade led to its classification as Vulnerable in the last review of the Red Book of Mammals of Portugal (Mathias et al., 2023). The rate of population decline has increased due to the emergence of a natural recombinant derived from the rabbit *Myxoma virus,* the *hare Myxoma virus* (ha-MYXV) (Águeda-Pinto et al., 2019; Dalton et al., 2019). Upgrading the genomic resources for the species by generating a high-quality reference genome will improve genomic studies, and anchor research aiming to quantify the impact of emerging pathogens on the naïve host hare populations, which can be used to inform conservation measures.

Sampling and Sample Processing: One male individual Iberian hare was sampled in Mértola, southern Portugal, during the regular permitted hunting season, duly authorised by ICNF (Permit 012/2022/CAPT). Tissues were collected shortly after the animal's death and flash frozen in liquid nitrogen. Tissues from several organs were collected during the necropsy and kept at -80ºC for biobanking: liver, kidney, spleen, lung and testis. The whole body will be prepared for deposition at the Natural History and Science Museum of the University of Porto (MHNCUP_MAM 0812).

Sequencing and Assembly: Tissue samples were shipped to the University of Antwerp for Omni-C and RNA-Sequencing library preparation, and to the University of Florence for PacBio DNA library preparation. Final Omni-C libraries were generated using kidney tissue, while RNA-Sequencing libraries were produced for liver, kidney, spleen, lung and testis tissues. These libraries along with PacBio DNA libraries were sequenced at the University of Florence. Long-read DNA sequencing was performed in HiFi mode in a PacBio Sequel IIe platform using five 8 million ZMW SMRT cells. Sequencing data was used to produce a draft genome assembly following the gold standard instructions implemented in the VGP Galaxy Pipeline (Batut et al., 2018; Lariviere et al., accessed June 2023).

Azores laurel - *Laurus azorica* (Seub.) Franco

Common name in portuguese: louro-da-terra and louro-bravo

The **Azores laurel** (*Laurus azorica*) is a micro or mesophanerophyte dioecious perennial tree endemic to the Azores Islands. It can be commonly found in submontane *Laurus* forests (Elias et al., 2016) but also occupies habitats such as coastal scrublands, among other native Azorean vegetation (Elias et al., 2016). It is listed as Least Concern in the IUCN Red List (Silva & Beech, 2016), yet it is facing habitat loss due to competition with invasive species and alteration of habitats resulting from the implementation of production forests and pastures (Borges Silva et al., 2018; Dutra Silva et al., 2019; Hortal et al., 2010; Lourenço et al., 2011).

The genus *Laurus* L. (Lauraceae) is currently restricted to isolated refugia in the southern Black Sea area, Mediterranean Basin, Northern Africa, and the Macaronesian archipelagos. Its taxonomic classification remains uncertain, with some recognizing the existence of two species, *Laurus azorica* (Seub.) Franco, endemic from the Azores and *Laurus nobilis* L. However, while some molecular studies contradict this distinction (Rodríguez-Sánchez et al., 2009), others even recognise a third species, *Laurus novo-canariensis* Rivas Mart., Lousã, Fern.Prieto, E.Dias, J.C.Costa & C.Aguiar (Rivas-Martínez et al., 2002). A high-quality reference genome will allow performing phylogenetic and population genetic studies to reconstruct the evolutionary history of this group, contributing to clarify the diversity and taxonomy of the *Laurus* genus.

Sampling and sampling processing: Fresh branches with healthy leaves were collected using gardening scissors from one mature male individual in the Azores islands, specifically, São Miguel Island, in Lombadas. The branches were loosely wrapped in moist absorbent paper, placed inside an open plastic bag and sent on the same day through express mail to the University of Lisbon, where they were kept at -20°C until DNA extraction. The corresponding voucher duplicate was dried using standard herbarium procedures and stored in the AZB herbarium at the Biology Department of the Faculty of Sciences and Technology of the University of the Azores, with voucher AZB4382, and collector code LZ001. A second sampling was conducted during March 2023 and followed the same methodology. The Nagoya certificate has CCIR number 61/2021/DRCTD.

Sequencing: The HMW DNA extraction was performed at the Department of Plant Biology of the Faculty of Sciences of the University of Lisbon (FCUL), using a modified cetyltrimethylammonium bromide (CTAB) protocol (Doyle & Doyle, 1987, later modified by Weising et al., 1994) adjusted based on the exchange of protocols within the ERGA community. DNA extraction was challenging, as extractions did not pass the quality control for PacBio Hi-Fi library preparation and sequencing. Leaf tissue was thus shipped in dry ice to the University of Florence, where the libraries for Pac-Bio will be prepared and sequenced.

Black wheatear - *Oenanthe leucura* (Gmelin, 1789)

Common name in Portuguese: chasco-preto

The **black wheatear** (*Oenanthe leucura*) is a passerine bird species confined to North Africa and southwest Europe, a region where global environmental changes (namely desertification and agriculture abandonment) are impacting the distribution of many bird species (Fernández-Nogueira & Corbelle-Rico, 2018, 2020; Fusco et al., 2021). Black wheatears are no exception, with findings supporting that some edge populations have already vanished (France) whilst others continue to decrease (Western populations in Portugal and Spain). Despite the decline and fragmentation of these populations, an isolated population located in northern Portugal (~100 pairs) still subsists, confined mainly to cultivated lands (primarily vineyards and olive orchards) of the Portuguese section of the Douro River, internationally known for the Port wine. For this reason, despite being classified globally as Least Concern by the IUCN Red List, Portuguese populations are classified as Critically Endangered in the Portuguese Red List (Cabral et al., 2005). We aim to characterise the structure and history of this population, which is closely associated with anthropogenic agricultural lands in the Douro valley, yet geographically isolated from other Iberian populations that occur mostly in natural habitats. This isolation and fragmentation could have irreversible effects on the survival of these populations and their conservation may require the implementation of specific conservation measurements. A genomic analysis of the Portuguese populations will thus provide information on these individuals' relatedness and dispersal capacity.

Sampling and sample processing: Due to the small size of the black wheatear population in Portugal, the sacrifice of a bird would be unjustified and, therefore, ICNF authorised the collection of blood samples from live birds (Permit 720/2021/REC). Three individuals (one female and two males) were caught using spring traps at the Douro Valley, near the Tua river mouth. Each bird was ringed, photographed and measured, and a blood sample (~100 microliters) was taken by venepuncture at the ulnar vein with a microcapillary. Each blood sample was immediately stored in liquid nitrogen and transferred to a -80ºC freezer on the same day.

Sequencing and assembly: Samples were sent to the ERGA partners in dry ice, to generate Omni-C (University of Antwerp) and PacBio HiFi data (University of Florence) from the female (C96097) and RNA-Sequencing (University of Antwerp) from one of the males (E29638). All DNA libraries were sequenced at the University of Florence. Long-read DNA sequencing was performed using two 8 million ZMW SMRT cells. Genome assembly is currently underway following the gold standard instructions implemented in the VGP Galaxy Pipeline (Batut et al., 2018; Lariviere et al., accessed June 2023).

Portuguese crowberry - *Corema album* (L.) D.Don

Common name in Portuguese: Camarina or Camarinha

The **Portuguese crowberry** (*Corema album*) is a dioecious perennial woody shrub endemic to the Atlantic coast of the Iberian Peninsula (ssp. *album*), and to the Azores Islands (ssp. *azoricum*). In the Iberian Peninsula it inhabits coastal areas from Galicia to Gibraltar, and is an important species in sand dune habitats which are highly valuable for conservation purposes (Council Directive 92/43/EEC of 21 May 1992 on the Conservation of Natural Habitats and of Wild Fauna and Flora, 1992). The dynamic nature of these coastal ecosystems provides a vast variety of habitats with unique floristic and animal richness. However, these dune systems face increasing disturbances as they support various economic and leisure activities, associated with the growth of the coastal population (Antunes et al., 2018). Because of habitat loss, *C. album* has been classified as Vulnerable on the Red List of Andalusia, Spain (Cabezudo et al., 2005). In the Azores Islands, it inhabits volcanic lava and ash fields (de Oliveira & Dale, 2012). The fruits of the Portuguese crowberry are edible, producing bioactive compounds that have been associated with chemoprotective activity and potential health-benefits (de Oliveira & Dale, 2012; Jacinto et al., 2021; Zunzunegui et al., 2006). Building a high-quality annotated reference genome will support ongoing projects on plant species from coastal areas, and provide the opportunity to explore evolutionary biology questions associated with adaptations to coastal environments under a scenario of global changes.

Sampling and sample processing: Two male and two female adult plants from the same population were sampled during two field trips to Azeitão, near Arrábida Natural Park. Young expanding leaves and fruits were collected after 48 hours of dark treatment (plant covered with light-opaque paper sheets with a few holes that allowed air flow), according to permits from ICNF (21PTLX00657D). Samples from leaves and fruits for DNA and RNA extraction were flash frozen at -20ºC and -80ºC, respectively. Voucher specimens were deposited at the Herbarium (LISU 270092) of the MUHNAC, Universidade de Lisboa.

Sequencing: The HMW DNA extraction was very challenging. First, it was performed at the Department of Plant Biology of FCUL using a CTAB protocol (Doyle & Doyle, 1987 later modified by Weising et al., 1994), which resulted in low DNA quality even after purification. Taking advantage of the ERGA network, we used a nuclei isolation protocol adapted from the ARIMA-HiC kit, shared by Narjes Yousefi (pers. comm.), a member of ERGA. This resulted in higher quality and quantity of extracted DNA, but it still did not pass the quality control for PacBio Hi-Fi library preparation and sequencing at the University of Florence. After several attempts, to prevent DNA degradation during shipping, leaf tissue was directly shipped in dry ice so that DNA extraction could be performed at the University of Florence, where the libraries for Pac-Bio will be prepared and sequenced.

Cave ground beetle - *Trechus terceiranus* (Machado, 1988)

Common name in portuguese: escaravelho-cavernícola-da-Ilha-Terceira

The **cave ground beetle** *Trechus terceiranus* (Coleoptera) is endemic to Terceira Island in the Azores and is restricted to subterranean habitats, including lava tubes. These tiny beetles have up to 4.3 mm in length, do not fly and have several morphological traits associated with their life in subterranean habitats (Machado, 1988). This species is listed as Vulnerable in the IUCN Red List (Borges & Amorim, 2018). Interestingly, several species of the genus *Trechus* occur in the Azores islands, most of which inhabit subterranean habitats (seven species), but two occur in surface habitats. One epigean species, *T. terrabravensis* Borges, Serrano & Amorim, 2004, co-exists on Terceira Island with *T. terceiranus*. A high-quality reference genome will be crucial to pursue comparative genomics, phylogenetic and population genetic studies to reconstruct the evolutionary history of this group, to detect genes involved in adaptation to subterranean habitats and to determine the genetic architecture of traits related with such habitats (e.g., eye development, pigmentation and biological clock).

Sampling and sample processing: Six adult individuals were sampled in a subterranean habitat (volcanic pit, lava tube) at 45-70m deep, using pitfall traps. Sampling was done according to permits issued by the Regional Government of the Azores (IRCC 23/2021/DRCTD). Adult individuals were flash frozen with liquid nitrogen. Dried vouchers were stored at the Dalberto Teixeira Pombo entomological collection at the University of the Azores, Terceira (DTP-MF1091).

Sequencing: The frozen whole organisms were shipped in dry ice to the University of Lisbon, from where they were subsequently sent to Centro Nacional de Análisis Genómico (CNAG) – Barcelona for RNA extraction and sequencing; and to the University of Florence where DNA extraction was attempted. Due to constraints to obtain high quality HMW DNA for PacBio, additional samples were sent in order to repeat the extraction to obtain enough DNA amount and quality for sequencing.

Building a national network for biodiversity genomics: Biogenome Portugal

Objectives and structure of Biogenome Portugal

While ERGA and ERGA-Portugal are a community of researchers, their collaboration in ERGA activities and in the pilot test has played a significant role in forming a national institutional network for biodiversity genomics, called Biogenome Portugal (BGP). This network is being formally established with the objective of fostering scientific collaboration, facilitating the exchange of expertise and infrastructure, organising advanced training programs, promoting Portugal's active participation in international biodiversity genomics initiatives, and contributing to knowledge transfer and outreach efforts. The network will comprise members with a diverse range of expertise, encompassing fields such as genomics, ecology, taxonomy, evolution, and more, spanning across various taxonomic groups.

By bringing together experts from different fields, the BGP network aims to facilitate the exchange of ideas, data, and resources that will lead to the development of new research projects centred on generating and using high quality reference genomes for Portuguese species, and address a plethora of biodiversity applications. Such collaboration also implies sharing institutional genomic and computational infrastructures and articulating with the national network of infrastructures related to BGP's areas of activity. Cooperation among members is crucial for promoting the exchange of expertise, a key objective of the network. By sharing knowledge and skills, the network members will strengthen their collective expertise and coordinate efforts. Activities towards the transfer of skills and expertise include organising advanced training in the field of biodiversity genomics. Through these initiatives, BGP aims to foster the training of specialised human resources able to accompany the fast development of genomic data analysis tools. To broaden its scope, the network intends to facilitate the involvement of individual researchers, even if their institution is not formally affiliated with the network.

BGP also aims to disseminate the results of its activities to stakeholders and the general public, thus promoting knowledge transfer and outreach. These activities are important not only to enrich scientific literacy about the importance of genomics for biodiversity conservation, but also to promote and encourage the use of the genomic research findings across stakeholders.

Finally, BGP also aims to coordinate and synchronise Portuguese participation in global biodiversity genomics initiatives. By reinforcing and sharing the opportunities for institutional representation in internationally funded projects, BGP can promote the contribution of Portuguese expertise to important regional, continental or worldwide initiatives.

Articulation with the Portuguese National Strategy for the Conservation of Nature and Biodiversity 2030

The Portuguese National Strategy for the Conservation of Nature and Biodiversity (ENCNB 2030; https://dre.pt/dre/detalhe/resolucao-conselho-ministros/55-2018-115226936) is the main legal Portuguese document that is used to frame all national policies until 2030. It recognizes the importance of Portugal's national biodiversity at both the European and global scale and acknowledges how the nation's biogeographic specificities have led to high levels of endemic and relict species, each with a unique evolutionary history and genetic composition, which is important to preserve. The national strategy has several main objectives, including planning and executing action promoting the conservation and recovery of species and habitats at the national level and promoting the conservation of plant and animals' genetic diversity. These objectives align with the aims of the nascent BGP. The network can provide the knowledge-based capacities necessary for the conservation and sustainable use of the national marine, terrestrial and freshwater genetic resources. BGP can reinforce a science-based approach for species conservation action plans and guide complex conservation strategies, including *in situ* and *ex situ* conservation.

Fostering national research infrastructures

In 2014, the national public agency for science, technology and innovation (Fundação para a Ciência e a Tecnologia, FCT) created a *National Roadmap for Research Infrastructures of Strategic Interest*. This roadmap aimed at promoting cooperation, excellence and internationalisation to strengthen national scientific infrastructures. A total of 56 Research Infrastructures (RIs) have been supported during the first funding cycle in key areas, such as the Environment, Health and Food, Social and Cultural Innovation, among others. BGP identified several RIs of interest, intersecting its area of intervention. Among these are GenomePT (the National Laboratory for Genome Sequencing and Analysis), Biodata.pt (ELIXIR PT - Portuguese Distributed Infrastructure for Biological Data), PORBIOTA (Portuguese E-Infrastructure for Information and Research on Biodiversity), EMBRC.PT (European Marine Biological Resource Centre – Portugal), and PRISC (Portuguese Research Infrastructure for Scientific Collections). Future action implies fostering the establishment of strategic cooperation with the active infrastructures, which can be facilitated by the current presence of several institutions of BGP in the RIs, pending the continuation of the current roadmap.

Engaging the community: establishing training programs, supporting the generation of additional reference genomes and expanding funding opportunities

The analysis of reference genomes at scale requires standardised procedures for the sampling and storing of the biological material, as well as sequencing and analysis of the genomic data. Researchers working to analyse and utilise these cutting-edge genomic resources need specific training, to maintain quality standards and streamline procedures. The development of bioinformatic pipelines by the larger ERGA community using open-source platforms such as Galaxy (The Galaxy Community, 2022) allows quick implementation of analyses across research centres. Within the BGP network, an online training programme is being implemented to cover different topics across pipelines of genomic analyses, such as genome and transcriptome assembly, annotation, comparative and population genomics. This programme aims to expand and facilitate training to enhance expertise in genomic analyses across the Portuguese research community. The first introductory course to Galaxy and VGP assembly pipeline took place in the first semester of 2023.

Further expertise sharing will come from the expansion of reference genomes projects under the BGP umbrella, which meets the technical and quality standards set by ERGA. For example, two reference genomes of Anthozoan species belonging to octocorals are being generated under the BGP initiative: the pink sea fan (*Eunicella verrucosa*) and the dead man's fingers (*Alcyonium digitatum*). Octocoral genomes remain poorly studied, with only ~10 reference genomes available from the more than 3,000 described species. These two species belong to a pool of octocorals for which the sequencing of reference genomes was supported by EASI-Genomics (H2020 824110; Project ID 10240, CoGeCo). The two species are widely distributed along the Portuguese coast in sublittoral rocky habitats, and sequencing, genome assembly and annotations are underway.

Attracting national and international funding will be crucial to support the implementation of BGP’s activities. Working as a network will lead to stronger research proposals, more successful grant applications, and increase collective participation in international consortia to foster biodiversity genomics cutting-edge research and knowledge transfer. This strategy will allow consolidating, strengthening and expanding the BGP network and its impact on fundamental and applied science and innovation.

Producing high quality chromosome-level assemblies of Portuguese species

The alignment of BGP with ERGA can foster the multiplication of opportunities for sequencing projects for Portuguese biodiversity. This aligns with the aim of ERGA of “propagating guidelines for state-of-the-art genome establishment through training and knowledge transfer”. Produced genomes can therefore take advantage of the standard sampling and molecular protocols and bioinformatics pipelines for sequencing, genome assembly and annotation established by ERGA, and be included under the ERGA umbrella. While promoting the visibility and accessibility of the genomes to the ERGA network and beyond, those genomes will contribute to the Reference Genome Atlas of European biodiversity. Such parallel initiatives under the BGP and ERGA umbrella, coordinated with ERGA projects, will expedite the production of a genomic inventory for all Portuguese eukaryotic species. Those efforts are invaluable to enable the use of genomic applications to preserve and protect the country’s rich biodiversity.

Figures

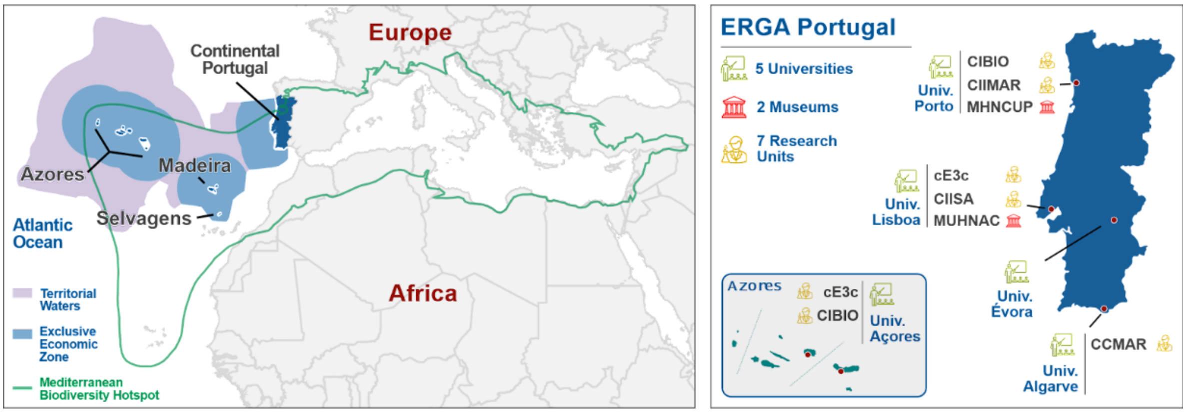

Fig.1 - a) Spatial context of the Portuguese territory, including the two Autonomous Regions (Azores and Madeira) b) Distribution of the Portugal-based members of ERGA (ERGA-Portugal), including their host institutions.

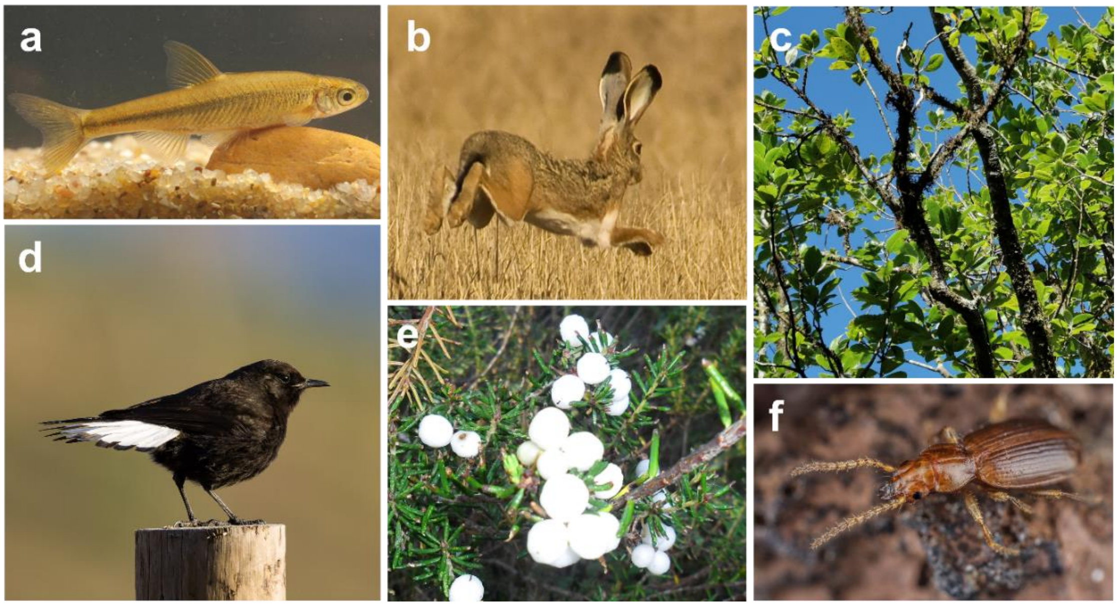

Fig.2 - Portuguese species sequenced in ERGA's pilot effort: a) Iberian minnowcarp - *Anaecypris hispanica*; b) Iberian hare - *Lepus granatensis*; c) Azores laurel - *Laurus azorica*; d) Black wheatear - *Oenanthe leucura*; e) Portuguese crowberry - *Corema album*; f) Cave ground beetle - *Trechus terceiranus*. Picture credits: a) Carlos Carrapato; b) Pedro Moreira; c) Mónica Moura; d) Ricardo Jorge Lopes; e) Manuela Sim-Sim; f) Javier Torrent (Azorean Biodiversity Group).

Acknowledgements

Co-authors: This work was supported by National Funds through FCT-Fundação para a Ciência e a Tecnologia in the scope of the project UIDP/50027/2020. Strategic funding from FCT to cE3c and BioISI Research Units (UIDB/00329/2020 and UIDB/04046/2020) and to the associate laboratory CHANGE (LA/P/0121/2020) are gratefully acknowledged. Financing by Azorean government regional funds, SRCCTD/DRCTD, project ERGA TAXA AÇORES (M2.2.A/PROJ.INT/A/001/2021 - ERGA taxa Açores). J. Melo-Ferreira acknowledges support from FCT - contract 2021.00150.CEECIND and project grant HybridChange PTDC/BIA-EVL/1307/2020, via Portuguese national funds. P.C. Alves acknowledges support by NORTE-01-0246-FEDER-000063, supported by Norte Portugal Regional Programme. T. Brown was supported by DFG (INST 269/768-1). FCT/MEC support the research contract to da L. P. Silva (CEECIND/02064/2017). R.J. Lopes was supported by a FCT - Transitory Norm contract (DL57/2016/CP1440/CT0006). P.A.V. Borges was funded by project FCT-UIDB/00329/2020-2024 (Thematic Line 1 – integrated ecological assessment of environmental change on biodiversity). S. L. Mendes is supported by an FCT PhD studentship (SFRH/BD/145153/2019). J. M. Moreno is supported by an FCT PhD studentship (SFRH/BD/143199/2019). This study received Portuguese national funds from FCT through project UIDB/04326/2020. H. Svardal was supported by the Flemish University Research Fund (BOF). A. Mouton was supported by a Postdoctoral Researcher Fellowship of the Fonds de la Recherche Scientifique – FNRS. Portuguese National Funds through FCT support the research contract to A. Crottini (2020.00823.CEECIND/CP1601/CT0003) and R. Faria (contract 2020.00275.CEECIND; grant Invcontinuum - PTDC/BIA-EVL/1614/2021). C. Fernandes thanks the support of cE3c through an assistant researcher contract (FCiência.ID contract #366) and FCT for Portuguese National Funds attributed to cE3c within the strategic project UID/BIA/00329/2020; C. Rodriguez Fernandes also thanks FPUL for a contract of invited assistant professor. Operational Programme (NORTE2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional. Project grant HYBRIDOMICS - PTDC/BIA-EVL/4345/2021, via Portuguese national funds. We gratefully acknowledge the generous support of the Max Planck Society for this research. Additional support was obtained from Horizon Europe under the Biodiversity, Circular Economy and Environment (REA.B.3); co-funded by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 22.00173; and by the UK Research and Innovation (UKRI) under the Department for Business, Energy and Industrial Strategy's Horizon Europe Guarantee Scheme. **ERGA Hubs:** Illumina and PacBio HiFi sequencing (RNA + DNA) was supported by the sequencing facility of the Department of Biology, University of Florence through the Departments of Excellence programme funded by the Italian Ministry for University and Research. We thank the Antwerp University Hospital Center of Medical Genetics and Jarl Bastianen for access to sequencing library quality control equipment. The Hi-C library preparation for *Trechus terceiranus* is being done at the Metazoa Phylogenomics Lab, Institute of Evolutionary Biology (CSIC-University Pompeu Fabra) through a collaboration with Judit Salces-Ortiz, Nuria Escudero and Rosa Fernández. **ERGA infrastructure:** We acknowledge access to the storage resources at Barcelona Supercomputing Center, which are partially funded from European Union H2020-INFRAEOSC-2018-2020 programme through the DICE project (Grant Agreement no. 101017207). We would like to thank Alisha Ahamed, Josephine Burgin, Joana Paupério, Jeena Rajan and Guy Cochrane from the European Nucleotide Archive (ENA) for their support regarding data coordination and submission of genomic data. **Commercial Partners:** We would like to acknowledge and thank all supplier partners that have kindly donated kits and reagents to the ERGA pilot Library Preparation Hubs to support species without funding to produce the generation of high-quality genomes and annotations. This support has been key to embedding a culture of diversity, equity, inclusion, and justice in the Pilot Project. Specifically, we want to thank Dovetail Genomics, Part of Cantata Bio LLC, especially Mark Daly, Thomas Swale and Lily Shuie; Arima Genomics; PacBio; Integrated DNA Technologies (IDT); MagBio Genomics Europe GmbH; Zymo Research; Agilent Technologies; Fisher Scientific Spain; Illumina Inc.

Author contributions

Species Ambassador or Executive coordination: Marques JP, Alves PC, Amorim IR, Lopes RJ, Moura M, Myers EW, Sim-Sim M, Sousa-Santos C and Melo-Ferreira J

Genome Team Member: Marques JP, Alves PC, Amorim IR, Lopes RJ, Moura M, Myers EW, Sim-Sim M, Sousa-Santos C, Alves MJ, Borges PAV, Brown T, Carneiro M, Carrapato C, Ceríaco LMP, Ciofi C, da Silva LP, Diedericks G, Diroma MA, Farelo L, Formenti G, Gil F, Grilo ML, Ianucci A, Leitão HG, Máguas C, Mc Cartney AM, Mendes SL, Moreno JM, Morselli M, Mouton A, Natali C, Pereira F, Rego RMC, Resendes R, Roxo G, Svardal H, Trindade H, Vicente S, Winkler S, Sousa VC and Melo-Ferreira J

ERGA-Portugal Member: Marques JP, Alves PC, Amorim IR, Lopes RJ, Moura M, Sim-Sim M, Sousa-Santos C, Alves MJ, Carneiro M, Ceríaco LMP, Trindade H, Vicente S, Alvarenga M, Amaral AJ, Antunes A, Campos PF, Canário AVM, Castilho R, Castro LFC, Crottini A, Cunha MV, Espregueira Themudo G, Esteves PJ, Faria R, Rodríguez Fernandes C, Ledoux JB, Louro B, Magalhães S, Paulo OS, Pearson GA, Pimenta J, Santos TL, Serrao EA, Pina-Martins F, Sousa VC and Melo-Ferreira J

ERGA-Portugal Coordination: Sousa VC and Melo-Ferreira J

Write Manuscript: Marques JP, Amorim IR, Lopes RJ, Moura M, Sim-Sim M, Sousa-Santos C, Trindade H, Espregueira Themudo G, Rodríguez Fernandes C, Ledoux JB, Paulo OS, Sousa VC and Melo-Ferreira J

Revise Manuscript: Marques JP, Alves PC, Amorim IR, Lopes RJ, Moura M, Sim-Sim M, Sousa-Santos C, Ciofi C, Mc Cartney AM, Mouton A, Trindade H, Alvarenga M, Amaral AJ, Antunes A, Canário AVM, Castilho R, Castro LFC, Crottini A, Cunha MV, Espregueira Themudo G, Esteves PJ, Faria R, Rodríguez Fernandes C, Ledoux JB, Louro B, Magalhães S, Paulo OS, Santos TL, Serrao EA, Pina-Martins F, Sousa VC and Melo-Ferreira J

Read and Approve Manuscript: Marques JP, Alves PC, Amorim IR, Lopes RJ, Moura M, Myers EW, Sim-Sim M, Sousa-Santos C, Alves MJ, Borges PAV, Brown T, Carneiro M, Carrapato C, Ceríaco LMP, Ciofi C, da Silva LP, Diedericks G, Diroma MA, Farelo L, Formenti G, Gil F, Grilo ML, Ianucci A, Leitão HG, Máguas C, Mc Cartney AM, Mendes SL, Moreno JM, Morselli M, Mouton A, Natali C, Pereira F, Rego RMC, Resendes R, Roxo G, Svardal H, Trindade H, Vicente S, Winkler S, Alvarenga M, Amaral AJ, Antunes A, Campos PF, Canário AVM, Castilho R, Castro LFC, Crottini A, Cunha MV, Espregueira Themudo G, Esteves PJ, Faria R, Rodríguez Fernandes C, Ledoux JB, Louro B, Magalhães S, Paulo OS, Pearson GA, Pimenta J, Santos TL, Serrao EA, Pina-Martins F, Sousa VC and Melo-Ferreira J

References

Águeda-Pinto, A., Lemos de Matos, A., Abrantes, M., Kraberger, S., Risalde, M. A., Gortázar, C., McFadden, G., Varsani, A., & Esteves, P. J. (2019). Genetic Characterization of a Recombinant *Myxoma virus* in the Iberian Hare (*Lepus granatensis*). *Viruses*, *11*(6), 530. https://doi.org/10.3390/v11060530

Alves, P. C., Acevedo, P., & Melo-Ferreira, J. (2023). Iberian Hare *Lepus granatensis* Rosenhauer, 1856. In K. Hackländer & F. E. Zachos (Eds.), *Handbook of the Mammals of Europe* (pp. 1–23). Springer International Publishing. https://doi.org/10.1007/978-3-319-65038-8_8-1

Antunes, C., Pereira, A. J., Fernandes, P., Ramos, M., Ascensão, L., Correia, O., & Máguas, C. (2018). Understanding plant drought resistance in a Mediterranean coastal sand dune ecosystem: Differences between native and exotic invasive species. *Journal of Plant Ecology*, *11*(1), 26–38. https://doi.org/10.1093/jpe/rtx014

Batut, B., Hiltemann, S., Bagnacani, A., Baker, D., Bhardwaj, V., Blank, C., Bretaudeau, A., Brillet-Guéguen, L., Čech, M., Chilton, J., Clements, D., Doppelt-Azeroual, O., Erxleben, A., Freeberg, M. A., Gladman, S., Hoogstrate, Y., Hotz, H.-R., Houwaart, T., Jagtap, P., … Grüning, B. (2018). Community-Driven Data Analysis Training for Biology. *Cell Systems*, *6*(6), 752-758.e1. https://doi.org/10.1016/j.cels.2018.05.012

Böhne, A., Fernández, R., Leonard, J. A., McCartney, A. M., McTaggart, S., Melo-Ferreira, J., Monteiro, R., Oomen, R. A., Pettersson, O. V., & Struck, T. H. (2023). Contextualising samples: Supporting reference genomes for European biodiversity through sample and associated metadata collection (p. 2023.06.28.546652). bioRxiv. https://doi.org/10.1101/2023.06.28.546652

Borges, P.A.V. & Amorim, I.R. 2018. *Trechus terceiranus*. The IUCN Red List of Threatened Species 2018: e.T97125072A99166594. https://dx.doi.org/10.2305/IUCN.UK.2018-1.RLTS.T97125072A99166594.en. Accessed on 22 June 2023.

Borges Silva, L., Lourenço, P., Teixeira, A., Azevedo, E. B., Alves, M., Elias, R. B., & Silva, L. (2018). Biomass valorization in the management of woody plant invaders: The case of *Pittosporum undulatum* in the Azores. *Biomass and Bioenergy*, *109*, 155–165. https://doi.org/10.1016/j.biombioe.2017.12.025

Browne, L., Wright, J. W., Fitz-Gibbon, S., Gugger, P. F., & Sork, V. L. (2019). Adaptational lag to temperature in valley oak (*Quercus lobata*) can be mitigated by genome-informed assisted gene flow. *Proceedings of the National Academy of Sciences*, *116*(50), 25179–25185. https://doi.org/10.1073/pnas.1908771116

Cabezudo, B., Talavera, S., Blanca, G., Salazar, C., Cueto, M., Valdés, B., Hernández-Bermejo, J. E., Herrera, C. M., Rodríguez-Hiraldo, C., & Navas, D. (2005). *Lista Roja de la flora vascular de Andalucía. Consejería de Medio Ambiente, Junta de Andalucía, Sevilla*.

Cabral, M. J., Almeida, J., Almeida, P. R., Dellinger, T., Ferrand de Almeida, N., Oliveira, M. E., Palmeirim, J. M., Queirós, A. I., Rogado, L., & Santos-Reis, M. (2005). *Livro Vermelho dos Vertebrados de Portugal*. Instituto da Conservação da Natureza. http://dspace.uevora.pt/rdpc/handle/10174/6006

Cardoso, A. (2023, February 7). Updating *Anaecypris hispanica* distribution and conservation status in Portugal. FiSHMED Journal. https://www.fishmedjournal.sibic.org/FiSHMED.2022.001/

Ceballos, G., Ehrlich, P. R., & Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences*, *114*(30), E6089–E6096. https://doi.org/10.1073/pnas.1704949114

Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora, CONSIL, 206 OJ L (1992). http://data.europa.eu/eli/dir/1992/43/oj/eng

Cowie, R. H., Bouchet, P., & Fontaine, B. (2022). The Sixth Mass Extinction: Fact, fiction or speculation? *Biological Reviews*, *97*(2), 640–663. https://doi.org/10.1111/brv.12816

Crivelli, A. J. (2006). IUCN Red List of Threatened Species: *Anaecypris hispanica*. *IUCN Red List of Threatened Species*. https://www.iucnredlist.org/en

Dalton, K. P., Martín, J. M., Nicieza, I., Podadera, A., de Llano, D., Casais, R., Gimenez, S., Badiola, I., Agüero, M., Duran, M., Buitrago, D., Romero, L. J., García, E., & Parra, F. (2019). *Myxoma virus* jumps species to the Iberian hare. *Transboundary and Emerging Diseases*, *66*(6), 2218–2226. https://doi.org/10.1111/tbed.13296

Dauphin, B., Rellstab, C., Schmid, M., Zoller, S., Karger, D. N., Brodbeck, S., Guillaume, F., & Gugerli, F. (2021). Genomic vulnerability to rapid climate warming in a tree species with a long generation time. *Global Change Biology*, *27*(6), 1181–1195. https://doi.org/10.1111/gcb.15469

De Miguel, R., Pino, E., Ramiro, A., Aranda, F., Peña, J. P., Doadrio, I., & Fernández-Delgado, C. (2010). On the occurrence of *Anaecypris hispanica*, an extremely endangered Iberian endemism, in the Guadalquivir River basin. *Journal of Fish Biology*, *76*(6), 1454–1465. https://doi.org/10.1111/j.1095-8649.2010.02574.x

de Oliveira, P. B., & Dale, A. (2012). *Corema album* (L.) D. Don, the white crowberry – a new crop. *Journal of Berry Research*, *2*(3), 123–133. https://doi.org/10.3233/JBR-2012-033

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19: 11–15.

Dutra Silva, L., Brito de Azevedo, E., Vieira Reis, F., Bento Elias, R., & Silva, L. (2019). Limitations of Species Distribution Models Based on Available Climate Change Data: A Case Study in the Azorean Forest. *Forests*, *10*(7), Article 7. https://doi.org/10.3390/f10070575

Elias, R. B., Gil, A., Silva, L., Fernández-Palacios, J. M., Azevedo, E. B., & Reis, F. (2016). Natural zonal vegetation of the Azores Islands: Characterization and potential distribution. *Phytocoenologia*, 107–123. https://doi.org/10.1127/phyto/2016/0132

Ette, J.-S., & Geburek, T. (2021). Why European biodiversity reporting is not reliable. *Ambio*, *50*(4), 929–941. https://doi.org/10.1007/s13280-020-01415-8

Fernández-Nogueira, D., & Corbelle-Rico, E. (2018). Land Use Changes in Iberian Peninsula 1990–2012. *Land*, *7*(3), Article 3. https://doi.org/10.3390/land7030099

Fernández-Nogueira, D., & Corbelle-Rico, E. (2020). Determinants of Land Use/Cover Change in the Iberian Peninsula (1990–2012) at Municipal Level. *Land*, *9*(1), Article 1. https://doi.org/10.3390/land9010005

Florencio, M., Patiño, J., Nogué, S., Traveset, A., Borges, P. A. V., Schaefer, H., Amorim, I. R., Arnedo, M., Ávila, S. P., Cardoso, P., de Nascimento, L., Fernández-Palacios, J. M., Gabriel, S. I., Gil, A., Gonçalves, V., Haroun, R., Illera, J. C., López-Darias, M., Martínez, A., … Santos, A. M. C. (2021). Macaronesia as a Fruitful Arena for Ecology, Evolution, and Conservation Biology. *Frontiers in Ecology and Evolution*, *9*, 718169. https://doi.org/10.3389/fevo.2021.718169

Formenti, G., Theissinger, K., Fernandes, C., Bista, I., Bombarely, A., Bleidorn, C., Ciofi, C., Crottini, A., Godoy, J. A., Höglund, J., Malukiewicz, J., Mouton, A., Oomen, R. A., Paez, S., Palsbøll, P. J., Pampoulie, C., Ruiz-López, M. J., Svardal, H., Theofanopoulou, C., … Zammit, G. (2022). The era of reference genomes in conservation genomics. *Trends in Ecology & Evolution*, *37*(3), 197–202. https://doi.org/10.1016/j.tree.2021.11.008

Fusco, J., Walker, E., Papaïx, J., Debolini, M., Bondeau, A., & Barnagaud, J.-Y. (2021). Land Use Changes Threaten Bird Taxonomic and Functional Diversity Across the Mediterranean Basin: A Spatial Analysis to Prioritize Monitoring for Conservation. *Frontiers in Ecology and Evolution*, *9*. https://www.frontiersin.org/articles/10.3389/fevo.2021.612356

Gómez, A., & Lunt, D. H. (2007). Refugia within Refugia: Patterns of Phylogeographic Concordance in the Iberian Peninsula. In S. Weiss & N. Ferrand (Eds.), *Phylogeography of Southern European Refugia: Evolutionary perspectives on the origins and conservation of European biodiversity* (pp. 155–188). Springer Netherlands. https://doi.org/10.1007/1-4020-4904-8_5

Habel, J. C., Rasche, L., Schneider, U. A., Engler, J. O., Schmid, E., Rödder, D., Meyer, S. T., Trapp, N., Sos del Diego, R., Eggermont, H., Lens, L., & Stork, N. E. (2019). Final countdown for biodiversity hotspots. *Conservation Letters*, *12*(6), e12668. https://doi.org/10.1111/conl.12668

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *359*(1442), 183–195.

Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, *30*(1), 62–82. https://doi.org/10.1111/mec.15720

Hortal, J., Borges, P. A. V., Jiménez-Valverde, A., de Azevedo, E. B., & Silva, L. (2010). Assessing the areas under risk of invasion within islands through potential distribution modelling: The case of *Pittosporum undulatum* in São Miguel, Azores. *Journal for Nature Conservation*, *18*(4), 247–257. https://doi.org/10.1016/j.jnc.2009.11.002

Jacinto, J., Giovanetti, M., Oliveira, P. B., Valdiviesso, T., Máguas, C., & Alegria, C. (2021). Quality attributes of cultivated white crowberries (*Corema album* (L.) D. Don) from a multi-origin clonal field. *Euphytica*, *217*(3), 40. https://doi.org/10.1007/s10681-021-02767-2

Karger, D. N., Kessler, M., Lehnert, M., & Jetz, W. (2021). Limited protection and ongoing loss of tropical cloud forest biodiversity and ecosystems worldwide. *Nature Ecology & Evolution*, *5*(6), Article 6. https://doi.org/10.1038/s41559-021-01450-y

Lariviere, D., Ostrovsky, A., Gallardo, C., Syme, A., Abueg, L., Pickett, B., Formenti, G., & Sozzoni, M. VGP assembly pipeline. (Galaxy Training Materials); https://training.galaxyproject.org/training-material/topics/assembly/tutorials/vgp_genome_assembly/tutorial.html, Online; accessed June 2023

Lourenço, P., Medeiros, V., Gil, A., & Silva, L. (2011). Distribution, habitat and biomass of *Pittosporum undulatum*, the most important woody plant invader in the Azores Archipelago. *Forest Ecology and Management*, *262*(2), 178–187. https://doi.org/10.1016/j.foreco.2011.03.021

Machado, A. (1988). Two new cavernicolous species of the genus *Trechus* Clairv. From the Azores (Coleoptera, Carabidae). *Bocagiana*, 119, 1-8.

Mathias, M. L., Fonseca, C., Rodrigues, L., Grilo, C., Lopes-Fernandes, M., Palmeirim, J. M., Santos-reis, M., Alves, P. C., Cabral, J. A., Ferreira, M., Mira, A., Eira, C., Negrões, N., Paupério, J., Pita, R., Rainho, A., Rosalino, L. M., Tapisso, J. T., & Vingada, J. (2023). *Livro Vermelho dos Mamíferos de Portugal Continental*. FCiências.ID / ICNF.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), Article 6772. https://doi.org/10.1038/35002501

O'Hara, C. C., Frazier, M., & Halpern, B. S. (2021). At-risk marine biodiversity faces extensive, expanding, and intensifying human impacts. *Science*, *372*(6537), 84–87. https://doi.org/10.1126/science.abe6731

Pauly, D., Zeller, D., & Palomares, M. L. D. (2020). Sea around us. Concepts, Design and Data.

Perea, S., Böhme, M., Zupančič, P., Freyhof, J., Šanda, R., Özuluğ, M., Abdoli, A., & Doadrio, I. (2010). Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. *BMC Evolutionary Biology*, *10*(1), 265. https://doi.org/10.1186/1471-2148-10-265

Ralls, K., Sunnucks, P., Lacy, R. C., & Frankham, R. (2020). Genetic rescue: A critique of the evidence supports maximizing genetic diversity rather than minimizing the introduction of putatively harmful genetic variation. *Biological Conservation*, *251*, 108784. https://doi.org/10.1016/j.biocon.2020.108784

Rivas-Martínez, S., Díaz, T. E., Fernández-González, F., Izco, J., Loidi, J., Lousa, M., & Penas, A. (2002). Vascular plant communities of Spain and Portugal: Addenda to the syntaxonomical checklist of 2001. 2. *Itinera Geobotanica (España)*. https://scholar.google.com/scholar_lookup?title=Vascular+plant+communities+of+Spain+and+Portuga l%3A+addenda+to+the+syntaxonomical+checklist+of+2001.+2&author=Rivas-Mart%C3%ADnez%2C+S.+%28Universidad+Complutense+de+Madrid.+Facultad+de+Farmacia%29&publication_year=2002

Rodríguez-Sánchez, F., Guzmán, B., Valido, A., Vargas, P., & Arroyo, J. (2009). Late Neogene history of the laurel tree (*Laurus* L., Lauraceae) based on phylogeographical analyses of Mediterranean and Macaronesian populations. *Journal of Biogeography*, *36*(7), 1270–1281. https://doi.org/10.1111/j.1365-2699.2009.02091.x

Segelbacher, G., Bosse, M., Burger, P., Galbusera, P., Godoy, J. A., Helsen, P., Hvilsom, C., Iacolina, L., Kahric, A., Manfrin, C., Nonic, M., Thizy, D., Tsvetkov, I., Veličković, N., Vilà, C., Wisely, S. M., & Buzan, E. (2022). New developments in the field of genomic technologies and their relevance to conservation management. *Conservation Genetics*, *23*(2), 217–242. https://doi.org/10.1007/s10592-021-01415-5

Seixas, F. A., Boursot, P., & Melo-Ferreira, J. (2018). The genomic impact of historical hybridization with massive mitochondrial DNA introgression. *Genome Biology*, *19*(1), 91. https://doi.org/10.1186/s13059-018-1471-8

Silva, L., & Beech, E. (2016). IUCN Red List of Threatened Species: *Laurus azorica*. *IUCN Red List of Threatened Species*. https://www.iucnredlist.org/en

Soriguer, R., & Carro, F. (2018). IUCN Red List of Threatened Species: *Lepus granatensi*s. *IUCN Red List of Threatened Species*. https://www.iucnredlist.org/en

Sousa-Santos, C., Robalo, J. I., Francisco, S. M., Carrapato, C., Cardoso, A. C., & Doadrio, I. (2014). Metapopulations in temporary streams – The role of drought–flood cycles in promoting high genetic diversity in a critically endangered freshwater fish and its consequences for the future. *Molecular Phylogenetics and Evolution*, *80*, 281–296. https://doi.org/10.1016/j.ympev.2014.08.007

The Galaxy Community. (2022). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Research*, *50*(W1), W345–W351. https://doi.org/10.1093/nar/gkac247

Theissinger, K., Fernandes, C., Formenti, G., Bista, I., Berg, P. R., Bleidorn, C., Bombarely, A., Crottini, A., Gallo, G. R., Godoy, J. A., Jentoft, S., Malukiewicz, J., Mouton, A., Oomen, R. A., Paez, S., Palsbøll, P. J., Pampoulie, C., Ruiz-López, M. J., Secomandi, S., … Zammit, G. (2023). How genomics can help biodiversity conservation. *Trends in Genetics*. https://doi.org/10.1016/j.tig.2023.01.005

Weising, K., Nybom, H., Pfenninger, M., Wolff, K., & Meyer, W. (1994). DNA Fingerprinting in Plants and Fungi. CRC Press, Boca Raton, 338 pp. Taylor & Francis. https://books.google.pt/books?id=AtkHlIks64YC

Zunzunegui, M., Díaz Barradas, M. C., Clavijo, A., Alvarez Cansino, L., Ain Lhout, F., & García Novo, F. (2006). Ecophysiology, growth timing and reproductive effort of three sexual forms of *Corema album* (Empetraceae). *Plant Ecology*, *183*(1), 35–46. https://doi.org/10.1007/s11258-005-9004-4