ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

PHYLOGEOGRAPHY OF THE SAVI'S PIPISTRELLE (VESPERTILIONIDAE, CHIROPTERA) COMPLEX BASED ON WHOLE MITOCHONDRIAL GENOME ANALYSIS

M.Sc. THESIS

Yeliz ERGÖL

Department of Climate and Marine Sciences

Earth System Science Programme

JULY 2024

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

PHYLOGEOGRAPHY OF THE SAVI'S PIPISTRELLE (VESPERTILIONIDAE, CHIROPTERA) COMPLEX BASED ON WHOLE MITOCHONDRIAL GENOME ANALYSIS

M.Sc. THESIS

Yeliz ERGÖL
(601211010)

Department of Climate and Marine Sciences

Earth System Science Programme

Thesis Advisor: Asisst. Prof. Dr. Emrah ÇORAMAN

JULY 2024

ISTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

SAVİ'NİN CÜCE YARASASI KOMPLEKSİNİN (VESPERTILIONIDAE, CHIROPTERA) FİLOCOĞRAFYASININ TÜM MİTOKONDRİYAL GENOM İLE ANALİZİ

YÜKSEK LİSANS TEZİ

Yeliz ERGÖL
(601211010)

İklim ve Deniz Bilimleri Anabilim Dalı

Yer Sistem Bilimleri Programı

Tez Danışmanı: Dr. Öğr. Üy. Emrah ÇORAMAN

TEMMUZ 2024

Yeliz Ergöl, a M.Sc student of İTU Graduate School student ID 601211010, successfully defended the thesis/dissertation entitled "PHYLOGEOGRAPHY OF THE SAVI'S PIPISTRELLE (VESPERTILIONIDAE, CHIROPTERA) COMPLEX BASED ON WHOLE MITOCHONDRIAL GENOME ANALYSIS", which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Assist. Prof. Dr. Emrah ÇORAMAN** İstanbul Technical University	.....
Jury Members :	**Assoc. Prof. Dr. Demet BİLTEKİN** Istanbul Technical University	.....
	Prof. Dr. Raşit BİLGİN Bogazici University	.....

Date of Submission : 24 May 2024
Date of Defense : 01 July 2024

To the pale blue dot,

FOREWORD

This thesis was written with the immense support of many great people. Foremostly, I would like to present my utmost thanks to my advisor Dr. Emrah Çoraman for his great guidance and continuous support. Besides the academic knowledge he has given me, he influenced my intellectuality and my interest in life and nature.

I also wish thanks to my defense committee, Assos. Prof. Dr. Demet Biltekin and Prof. Dr. Raşit Bilgin for giving their time to evaluate and contribute to this study.

I would like to thank PD Dr. Frieder Mayer for allowing access to the samples in Museum für Naturkunde, sharing his precious insights and his kind support during this study.

I am very grateful to be a part of Çoraman Lab and I would like to thank each member of our lab, Alhas Can Arslan, Bengisu Şensoy, Gamze Özbay, Songül Sönmez and Tuğçe Nur İlbaş. It was a wonderful journey but I had kept sane thanks to your company in some trying times. Bengisu and Tuğçe, we have come a long way and I enjoyed the ride with you. We made it "idiosyncratically"!

I owe a special thanks to Dr. Sercan Irmak, who taught me many things about how to be an ecologist and gave me all the knowledge about bats. He was a great mentor with his humongous support and positive energy.

All of my dear friends gave tremendous support and patience during my master's studies. I want to thank all of them. My close friends, Reyhan Şimşek, Ayşe Özkan and Yağmur Varkal were always with me during the most difficult times. I am in great debt of gratitude for them to be in my life. I owe many thanks to my dear flatmate Kübra Turgut who had listened to a lot of my complaints about life during the terrific three years we shared together.

I heartily thank Mehmet Kerem Avcı for sharing my excitement and curiosity for life, and supporting my passions.

Last but not least, I truly thank my sisters, Duygu Usta and Gülay Ergöl Pınar, who enlightened my way and supported me greatly in every moment of my life. I was capable of approaching my dreams thanks to embracing the love of my dear parents Naile Ergöl and Cantürk Ergöl. Thank you for believing in me and backing my decisions! I couldn't achieve any of this without my family.

This study is supported by the Istanbul Technical University Scientific Research Projects Coordination Unit (Project Number: TYL-2021-43345) and the National Center for High-Performance Computing (Grant Number: 4009622021). Scientific and Technological Research Council of Türkiye has provided scholarship with ARDEB projects during this study.

TABLE OF CONTENTS

Page

ABBREVIATIONS

Ala : Alanine
Arg : Arginine
Asn : Asparagine
Asp : Aspartic Acid
ATP : Adenosine Triphosphate
bp : Base Pair
CO : Cytochrome c Oxidase
Cys : Cysteine
CYTB : Cytochrome b
D-loop : Displacement Loop
DNA : Deoxyribonucleic Acid
g : Gram
Gln : Glutamine
Glu : Glutamic Acid
Gly : Glycine
His : Histidine
Ile : Isoleucine
IUCN : International Union for Conservation of Nature
kb : Kilobase
kHz : Kilohertz
Leu : Leucine
Lys : Lysine
m : Meter
Met : Methionine
ML : Maximum-Likelihood
mm : Millimeter
mtDNA : Mitochondrial DNA
NADH : Nicotinamide Adenine Dinucleotide (Reduced form)
ND : NADH Dehydrogenase
PCGs : Protein-Coding Genes
Phe : Phenylalanine
Pro : Proline
RAD-seq : Restriction site-associated DNA sequencing

RAG2 **:** Recombination Activating Gene 2

RNA **:** Ribonucleic Acid

Ser **:** Serine

Thr **:** Threonine

Trp **:** Tryptophan

Tyr **:** Tyrosine

Val **:** Valine

SYMBOLS

A	**:** Adenine
C	**:** Cytosine
G	**:** Guanine
T	**:** Thymine
X	**:** Depth of Coverage
℃	**:** Degree Celcius
%	**:** Percent

LIST OF TABLES

Page

LIST OF FIGURES

Page

PHYLOGEOGRAPHY OF THE SAVI'S PIPISTRELLE (VESPERTILIONIDAE, CHIROPTERA) COMPLEX BASED ON WHOLE MITOCHONDRIAL GENOME ANALYSIS

SUMMARY

Understanding the phylogeography of species provides insight into the historical processes that lead the formation of their current geographic distributions and also how they evolve and adapt to variable environments over time. Intraspecific and interspecific genetic variations stand as a great indicator for elucidation of the evolutionary history of organisms in diverged phylogeographical patterns. Bats represent an immensely diverged group among mammals. They inhabit a wide variety of ecosystems, with the tropical belt hosting the highest number of bat species, similar to many other life forms.

While the oldest known fossil of Chiroptera is approximately 52 million years old, they are hypothesized to be evolved around the Cretaceous-Tertiary boundary. During their evolution, they settled to the nocturnal niche. Echolocation and flight capabilities have affected the phylogenetic classification in the era of morphological and physiological taxonomy. However, the taxonomic classification of Chiroptera is a more complex phenomenon.

The developments of modern genetic techniques transformed the historical perception of taxonomy. The introduction of high-throughput sequencing, which enables the sequencing of an organism's whole genetic material, has made genomic studies increasingly popular in biodiversity research. This powerful technique has revealed that organisms can exhibit significant genetic variation, even when their phenotypic characteristics do not reflect this diversity. Cryptic species, which arise from the discordance between morphological similarity and genetic divergence, exemplify this phenomenon. Resolving cryptic diversity is cruicial for identying evolutionary significant units and provides a new dimension for investigating the ecological dynamics of species. Furthermore, it is a significant concept in biodiversity assessment and monitoring, essential for inforimg conservation actions.

Past studies have showned that the Palearctic Region hosts a rich fauna with signifacnt cryptic diversity, including among bats. Savi's Pipistrelle, *Hypsugo savii*, is a small-sized, vespertilionid bat species with a broad distribution range across various ecoregions in Europe, Asia, and North Africa. The accumulation of studies showing intraspecific variation within the species has drawn attention to the investigation of possible divergence within the taxon. Several studies idenfied deeply diverged mitochondrial lineages of *H. savii*. Two of such lineages have been proposed as distinct species statuses, *H. darwinii* and *H. stubbei*, based on their significant divergence from *H.savii*. *Hypsugo savii* hosts possibly further cryptic diversity. It is composed of deeply diverged clades, with sympatric occurences in Northwestern Africa, the Iberian Peninsula, Italy, and some Mediterranean islands (Sardinia, Sicily, and Malta). These studies, however, are based on limited sampling from a very broad

geographical range. Furthermore, they utilize realtively short mitochondrial markers and marker selection is not usually consistent among studies.

This study aims to invesitage the whole mitogemones of the previously identified *H. savii* lineages and the recently suggested related species. There are three main objectives: 1) establishing a reproducible workflow to de novo assemble mitogenome from whole genome sequencing data; 2) de novo assembling complete mitogenome of *H. savii* and the related species; and 3) reconstructing their phylogenetic relationships based on whole mitogenome sequences.

Thirty samples from various regions including Central Asia, Sinai Peninsula, North Africa, continental Europea, and Mediterranean islands were analysed. High-throughput shotgun sequencing data was used for the analysis. The analysis workflow covers data filtering, de novo assembling of complete mitogenomes, and annotating the obtained mitogenomes. Complete mitochondrial genomes were succesfully de novo assembled for thirty samples, representing all of the previously identified lineages, as well as the closely related species, *H. stubbei* and *H. alaschanicus*. Mitochondrial genes were annotated on the assembled genomes. The read pool of each sample was mapped to the assembled mitogenomes to assess their coverage and also to edit possible misallignments and gaps. The samples which had uncircular genomes were manual edited and circularized. The tRNA profiles were analyzed with tRNAscan-SE. Phylogenetic relations of the analysed samples were investigated reconstructing phylogenetic trees. The sequences were aligned and the pairwise distance of the sequences were calculated with MEGA11. The phylogenies were obtained with Maximum-Likelihood model with IQ-TREE and RAxML tools. The trees were visualized with FigTree abd iTOL WebServer. The alignments were analysed with PopArt for a haplotype network analysis. The highly variable non-coding D-loop regions were removed from the sequences for all these analysis .

All thirty-seven mitochondrial genes were annotated: two of them were ribosomal RNAs; twenty-two were transfer RNAs; and thirteen were protein-coding genes. D-loop regions of mitogenomes were also annotated. The secondary stuctures of the tRNA profiles were calculated and illustrated. The tRNA for serine amino acid was lacking the D- arm in the secondary stucture, which had no significant impact on its functionality.

Phylogenetic analysis revealed that there are three main *H. savii* lineages in the Western Palearctic region. The related species, *H. ariel*, *H.alaschanicus*, and *H. stubbei*, also formed three distinct clusters. The latter was identified in Kyrgyzstan, which is outside of its known range. The distribution of *Hypsugo savii* lineages were in three main regions: Eastern Mediterranean, Western Mediterranean, and North African-Southwestern Mediterranean. In Sardinia, the latter of these three lineages were found together. The pairwise genetic distance between the clades in Europe were between 8 and 10%, close to the difference observed between *H. alaschanicus* and the other species. Similar levels of divergences were also found between newly proposed species, *H. stubbei*, and the European clades. These observed high levels of mitogenomic differences suggest that the *H. savii* complex probably harbours further cryptic diversity.

SAVİ'NİN CÜCE YARASASI KOMPLEKSİNİN (VESPERTILIONIDAE, CHIROPTERA) FİLOCOĞRAFYASININ TÜM MİTOKONDRİYAL GENOM İLE ANALİZİ

ÖZET

Türlerin filocoğrafyasını anlamak, onların mevcut coğrafi dağılımlarının oluşumuna yol açan tarihsel süreçlere, zaman içindeki evrimsel değişimlere ve değişen koşullarda ortamlara uyum sağladıklarına dair fikir verir. Tür içi ve türler arası genetik varyasyonlar, farklı filocoğrafik örüntülere sahip organizmaların, evrimsel tarihinin aydınlatılmasında büyük bir gösterge olarak kullanılabilmektedir. Yarasaların ait olduğu memeli takımı Chiroptera, dünya çapında yaklaşık 1.400 türe sahiptir ve Memeliler sınıfı içerisinde son derece farklı bir gruptur. Yaşam alanları Antartika hariç tüm dünyayı kapsamaktadır. Haibtat alanlarının ekosistem çeşitliliği de dikkate değerdir. Kuru ve sıcak çöl ekosistemlerinden sulak alanlara, dağ bozkırlarından aşırı nemli tropiklere kadar yayılmış hâlde bulunurlar. Tropikal kuşak, çoğu yaşam formunda olduğu gibi, yarasalar için de en fazla sayıda türe sahip bölgedir.

Bilinen en eski yarasa fosili yaklaşık 52 milyon yıl yaşında olsa da, takımın Kretase-Tersiyer sınırı civarında evrimleştiği varsayılmaktadır. Evrimleri sırasında gece aktivitesine yönelmişlerdir. Gece aktivitesi ve uçuşun beraber evrimleşmesi, aktif uçuşa adapte olmuş tek memeli olarak sahip oldukları eşsiz özelliklerden birinin, ekolokasyonun ortaya çıkmasına neden olmuştur. Yarasalar, karanlıktaki hassas ve keskin avlanma becerilerini, karanlıkta ekolokasyonla uçmalarına borçludur. Bu iki özellik, morfolojik ve fizyolojik taksonomi çağında filogenetik sınıflandırmayı da etkilemiştir. Ancak Chiroptera takımındaki taksonomik sınıflandırma ve filogenetik örüntüler sanıldığından daha karmaşık bir olgudur.

Modern genetik teknik ve araçlarının gelişmesi, taksonominin tarihsel algısının değişmesine de neden olmuştur. Tür kavramlarından başlayarak takson tanımının sınırları, popülasyonlar arasındaki belirgin olmayan farklılıklara doğru kaymıştır. Genomik, bir organizmanın tüm genetik materyalinin dizilenmesine olanak sağlayan yüksek verimli dizilemenin başlaması ve geliştirilmesiyle birçok alanda egemen bir teknoloji olmuştur. Bu güçlü konsept, fenotipik özellikler bunu temsil etmese bile organizmaların genetik düzeyde büyük ölçüde farklılık gösterebileceğinin anlaşılmasına katkıda bulunmuştur. Kriptik türler, türlerin morfolojik ve genetik yapısı arasındaki bu uyumsuzluğun bir sonucudur. Genetik düzeyde önemli bir şekilde farklılaşan en az iki benzer türün tek bir nominal türde sınıflandırılması durumunda, bu tür kriptik türlerin bir örneğini teşkil etmektedir. Kriptik çeşitliliğin araştırılıp anlaşılması, evrimsel önemli birimlerin belirlenmesi için mühim bir fırsat sağlar. Aynı zamanda türlerin ekolojik dinamiklerinin araştırılmasına da yeni bir boyut getirmektedir. Ayrıca, koruma eylemi için değerlendirilmesi gereken temel unsurlar olarak ekolojik anlamda biyolojik çeşitliliğin değerlendirilmesi ve izlenmesi açısından önemli bir kavramdır.

Geçmiş çalışmalar Palearktik'in yarasalar da dahil olmak üzere kriptik çeşitliliğe sahip zengin bir faunaya ev sahipliği yaptığını gösterdi. Savi'nin cüce yarasası, *Hypsugo savii*; Avrupa, Asya ve Kuzey Afrika'daki çeşitli ekolojik bölgelerde geniş bir dağılıma sahip küçük boyutlu, vespertilionid bir yarasa türüdür. Türün genetik varyasyonunu gösteren çalışmaların birikmesi, takson içindeki olası farklılıkların araştırılmasına dikkat çekmiştir. Literatürde *H. savii*'nin mitokondriyal soylarının oldukça farklı olduğu ileri sürülmektedir. *H. darwinii* ve *H. stubbei, H. savii* ile mitokondriyal farklılıklarından ötürü yeni türler olarak önerilmişlerdir. Türün filocoğrafik dağılımı çeşitli sınıflara bölünmüştür. Kuzeybatı Afrika, İber Yarımadası, İtalya ve Akdeniz adaları (Sardunya, Sicilya, Malta vb.) dahil olmak üzere bazı bölgelerde derinden farklılaşmış sınıfların simpatrik dağılımları olduğu düşünülmektedir. Bu çalışmalardaki sorun, çok geniş bir coğrafyadan sınırlı sayıda örnekleme yapılması, kısa ve farklı mitokondriyal dizilerin genetik belirteç olarak kullanılması ve yeniden oluşturulan filogeniler arasında uyumsuzluk tespit edilmesidir.

Bu çalışmada, *H. savii* kriptik tür kompleksinin kendi içinde Batı Palearktik'te birbirinden çok farklı en az üç taksonu içerdiği düşünülmektedir. Bunlar Doğu Akdeniz dağılım gösteren takson, Batı Akdeniz'de dağılım gösteren takson (Literatürde *H. ochromixtus* olarak da isimlendirilmekte) ve Kuzey Afrika ile Güneybatı Akdeniz'de yayılan takson (Literatürde *H. darwinii* olarak ayrı bir tür olması önerilmekte) gruplarıdır. Farklı bölgelerden toplanan çalışma sayısının artması, İber Yarımadası, İtalya ve Akdeniz adalarında (Sardunya, Sicilya, Malta vb.) bu dallardan bazılarının simpatrik dağılım örüntülerini göstermektedir. Bu dalların yeni türler olarak tanınması için, simpatrik bölgelerdeki genetik çeşitliliği yansıtan iyi bir örnekleme ile çoklu ve tutarlı genetik analizlerle daha fazla kanıta ihtiyaç vardır. Bu çalışma, tüm mitokondriyal genomların kullanılması ve *H. savii* kriptik tür kompleksi içindeki mitokondriyal filocoğrafik örüntüleri anlamayı amaçlamaktadır. Çalışma bu hedefe uygun olarak üç ana hedef üzerinde planlanmıştır: 1) Tüm genom dizileme verilerinden mitogenom oluşturmak için tekrarlanabilinir bir iş akışı oluşturulması, 2) *H. savii* ve yakın türlerin tam mitogenomlarının elde edilmesi, 3) *H. savii* kriptik tür kompleksinin tam mitogenom dizileri ile filogenetik ilişkilerinin anlaşılması.

Bu çalışma öncesinde Orta Asya, Sina Yarımadası, Avrupa kıtası ve adaları, Fas gibi çeşitli bölgelerden otuz örnek çalışılmış ve yüksek verimli dizileme ile analiz edilmiştir. Bu örnekleri incelemek için birkaç adımdan oluşan bir iş akışı oluşturulmuştur. İş akışı, verilerin fastp ile niteliklere göre filtrelenmesini, MitoZ ile "de novo" algoritmasını kullanarak tam mitogenomların bir araya getirilmesini ve elde edilen mitogenomlara gen bölgesi anotasyonlarının yapılmasını kapsar. Buna göre, toplanan otuz örnek için tam mitokondriyal genomlar de novo algoritması ile analiz edilmiştir. Ayrıca örneklere Doğu Asya'dan yakın akraba olan *H. stubbei* ve *H. alaschanicus* türleri de dahil edilmiştir. Mitokondriyal genlerin, oluşturulan genom dizilerindeki yerleri tayin edilmiştir. Her numunenin okuma havuzu, kapsama dağılımını görmek ve bir yanlış hizalama veya boşluk olup olmadığını kontrol etmek için bwa ile bir araya getirilmiş mitogenomlarla haritalandırılmıştır. Haritalama Geneious Prime ile görselleştirilmiş ve dairesel olmayan genomların manuel olarak düzenlenmesi tamamlanmıştır. Mitogenomlar, tRNA profilini ortaya çıkarmak için tRNAscan-SE ile analiz edilmiştir.

Filogeni ağaçları oluşturmak için, oldukça değişken, kodlamayan mitokondriyal dizi bölgeleri, dizilerin geri kalanından çıkarılmıştır. Dizilerin geri kalanı hizalanmış ve dizilerin genetik uzaklık mesafesi MEGA11 ile hesaplanmıştır. Filogeniler, IQ-TREE ve RAxML araçlarıyla "Maksimum Olabilirlik" modeliyle elde edilmiştir. Ağaçlar FigTree ve iTOL WebServer ile görselleştiril. Hizalamalar bir haplotip ağ analizi için PopArt ile analiz edilmiştir. Ortaya çıkan filogenetik gruplar hem ağaçlarda hem de haplotip ağında gösterilmiştir. R Studio'da ggplot2 ile dalları gösteren bir dağıtım haritası çizilmiştir.

Tüm numuneler için tam mitokondriyal genomlar elde edilmiştir. Mitogenomların gen bölgelerinin belirlenmesinde, memeli mitogenomlarına özgü otuz yedi mitokondriyal genin tamamı tespit edilmiştir; bunlardan ikisi ribozomal RNA, yirmi ikisi transfer RNA ve on üçü protein kodlayan genlerdir. Mitogenomların kodlamayan kontrol bölgeleri de tanımlanmıştır. tRNA profilinin analizinde tRNA'ların ikincil yapıları tanımlanmış ve görselleştirilmiştir. Serin amino asit taşımasında görevli tRNA'nın ikincil yapısında D-kolunun bulunmadığı tespit edilmiştir. Bu eksikliğin işlevsellik üzerinde önemli bir etkisi olmadığı literatürde destek görmüştür. Oluşturulan mitokondriyal genomlar arasındaki farklar, taksonlar arasındaki çeşitliliği gösterecek şekilde ölçülmüştür. Genetik soylar arasındaki çeşitlilik seviyeleri mesafe matrisinde sunulmuştur. Tanımlanan mitokondriyal soyların evrimsel tarihini araştırmak için filogenetik ağaç yöntemleri kullanılmıştır. Filogenetik analizler, *H. savii* türünün dört ana gruptan oluştuğunu ortaya koymuştur. Yeni önerilen *H. stubbei* türünün bulunduğu Kırgızistan grubundan örnekleri içeren küme bu dört kümeden biridir. *Hypsugo savii* türünün diğer üç grubu Batı Palearktik'te tespit edilmiştir. Örneklerin çoğu Doğu Akdeniz ve Batı Akdeniz olmak üzere iki grupta yer almıştır. Sardinya'dan iki örnek, doğu ve batı dallarından tamamen ayrı bir dalda konumlanmıştır. Bu iki örneğin Kanarya Adaları, Fas, İber Yarımadası ve Akdeniz adalarında bulunan, daha önceki çalışmalarda tespit edilmiş Kuzey Afrika/Güneybatı Akdeniz soyundan bireyler olduğu tahmin edilmektedir. İkili genetik mesafe matrisi, Avrupa'daki dallar arasındaki farklılaşma düzeyinin (%8-10) *H. alaschanicus*'un, *H. savii* gruplarından farklılığına çok yakın olduğunu göstermiştir. Ayrıca, yeni önerilen *H. stubbei* türü ile Avrupa soyları arasında da benzer düzeyde çeşitlilik bulunmuştur. Bu fark, dallar arasında çok belirgin bir farklılık olarak değerlendirilmektedir. *H. savii*'nin sahip olduğu tür içi kriptik çeşitlilik bu çalışma ile desteklenmiştir.

Çalışmanın kapsamı önemli derecede farklılaşmış soyların potansiyel temas bölgelerinden sınırlı örnek incelenebilmesinden ötürü belirli sınırlar içerisinde kalmıştır. Ayrıca filogenetik ağaçlardaki kladları doğrulamak için tam mitokondriyal genomlarla daha farklı filogenetik modeller kullanılabilir.

Bu çalışma, *H. savii* kompleksinin filocoğrafyası ve evrimsel tarihine kapsamlı bir bakış sunmakta ve aynı zamanda yarasa evrimini anlamlandırmamıza da katkıda bulunmaktadır. Bu kapsamda genetik çeşitliliğe dayalı filocoğrafik örüntüleri yeniden oluşturmak için bir metodoloji oluşturulmuştur. Bu çalışma, *H. savii* türü üzerine yapılan çalışmalarıda, filogenetik analizlerde yüksek çözünürlüğü sağlamak için ilk kez tüm mitogenom dizilerinden yararlanmıştır. Bu çalışmadan elde edilen bulgular, türlerin evrimsel dinamiklerinin anlaşılmasında kapsamlı genomik analizlerin öneminin altını çizmektedir.

1. INTRODUCTION

1.1 Bats and Their Evolutionary History

In 1968, Martin Eisentraut said "No other mammal group can surpass the diversity and curiosity of the face design of bats." (Dietz & Kiefer, 2016). Not only the face of them, but with almost 1,400 species recognized worldwide, the bats are one of the most diverse groups in class Mammalia (Simmons et al, 2024). In addition to their rich taxonomic diversity, they inhabit all continents except Antarctica. Their habitats span a wide range from harsh desert conditions to tropical forests, and from wetlands to alpine steppes. As with many forms of life, the majority of bat species cluster around the equator zone (Figure 1.1), and their distribution is also influenced by geographical features (Mickleburgh et al, 2002). While a large number of bats are insectivorous, there are also carnivorous, frugivorous, or sanguivorous bats (Dietz & Kiefer, 2016).

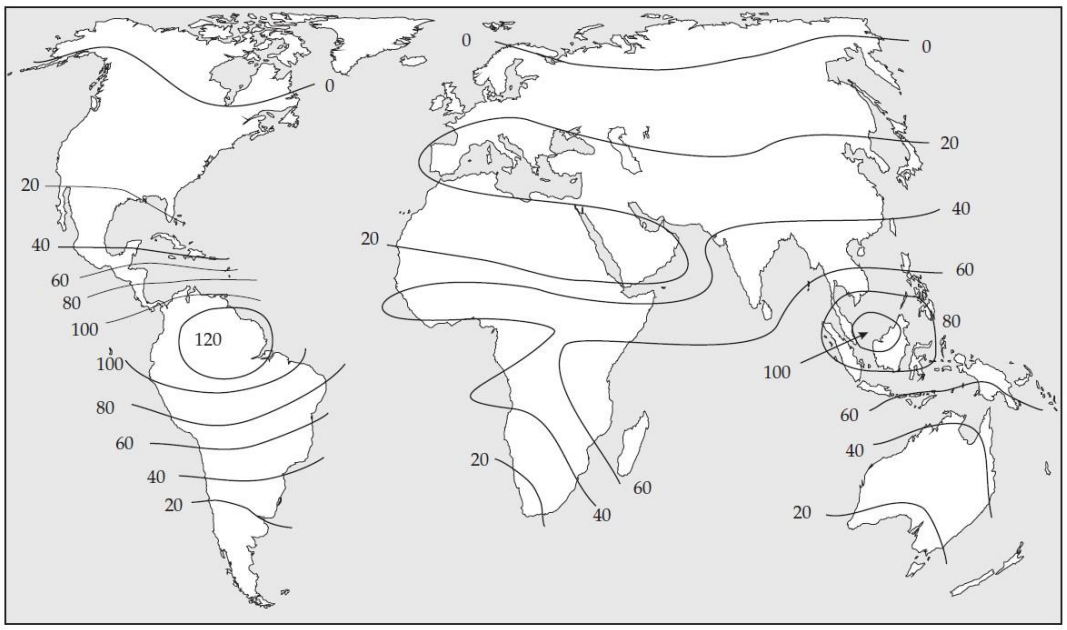

Figure 1.1: The distribution of number of bat species in the world adapted from Findley (1993).

Chiroptera, the order of bats, stands apart from other mammalian groups. Studies reveal that Chiroptera lack close phylogenetic ties with other groups in mammalian

orders. The closest – but still the distant – relatives to bats are pangolins. They are thought to share a common ancestor in the distant reaches of geological history (Murphy et al, 2001). The earliest fossil records of bats were found in Green River, Wyoming dating back to the Eocene epoch, approximately 52 million years ago (Dietz & Kiefer, 2016). Molecular phylogenetics analysis traces the last common ancestor of bats even further back, to around 64 million years ago. The emergence of bats appears to have coincided with an upswing in the diversity of flowering plants and insects toward the end of the Cretaceous period (Lillegraven et al, 1987). Telling et al. (2005) propose that the major diversification of microbat taxa during the Tertiary period correlates with a remarkable increase in mean temperature, approximately 7 ℃ higher than previous levels. This warming likely led to a critical upsurge of the plant species and the insects. As these flying mammals adapted to their ecological niche, characterized by nocturnal activity, they experienced immense diversification, facilitated by the reduction of strong predator pressure (Altringham, 2011).

For quite a long time, debates surrounding the evolutionary history of bats were revolved around their unique feature: echolocation. With a highly specialized larynx muscle that can contract 200 times per second, the echolocating bats can make high-frequency ultrasound calls up to 212 kHz (Dietz & Kiefer, 2016). From a broad perspective, echolocation enables nocturnal flight and efficient hunting. Consequently, the central question in understanding bat divergence has been whether echolocation or flight evolved first. Notably, these two critical traits appear to be intertwined, as precise echolocation likely sharpened the effectiveness of nocturnal flight and feeding (Altringham, 2011).

1.1.1 Cryptic diversity in bat species

Modern genomics challenges the strict boundaries of species definitions, including Mayr's biological species concept, which traditionally relies on reproductive isolation (De Queiroz, 2005). From this perspective, the common perception of cryptic species as "sibling species" is also ambiguous due to the lack of a universally accepted concept. Bickford et al. (2006) define cryptic species as instances where two or more species are classified as a single nominal species due to the difficulty of morphological differentiation.

Cryptic diversity has gained significance in light of recent biodiversity advances. Espindola et al. (2016) highlight its importance based on three key factors. First, cryptic diversity plays a crucial role in biodiversity assessments. Parameters such as species abundance and endemism are central to local conservation efforts, and unraveling cryptic diversity would provide more accurate estimates. Second, investigating the cryptic diversity aids in identifying evolutionary significant units. A critical task considering that many species face a "silent" extinction without proper taxonomic recognition (Löbl et al, 2023). Lastly, cryptic diversity is essential to comprehend the effects of anthropogenic activities on species and ecosystem health. Recognizing these factors underscores the need for increased research and support in cryptic diversity, particularly given the challenges posed by climate change and the conservation of richness of life (Espindola et al, 2016).

Bats hold great potential for cryptic diversity assessments. While pioneering molecular studies had already identified several cryptic species, the increased resolution of genomic analyses reveals that much remains to be discovered (Mayer & Helversen, 2001). Çoraman et al. (2013) showed an important example of rich intraspecific genetic diversity and potential cryptic diversity among several bat species in the Western Palearctic region, particularly in Anatolia. Among these, *Hypsugo savii* stands out as a candidate for respresting a cryptic species complex.

1.1.2 Study organism: *Hypsugo savii* sensu lato

Before the advent of molecular genetic studies, *Hypsugo, Vespertilio*, *Eptesicus*, and *Pipistrellus* were grouped together as a genus. These classifications were primarily based on phenotypic similarities, including dental characteristics (Kipson et al, 2023). Later on, the genus *Hypsugo* was also classified as a subgenus of *Pipistrellus* (Hill & Harrison, 1987). Menu et al. (1987), on the other hand, claimed that the *savii* group and its closely related relatives warrant recognition as a distinct genus due to their unique characteristics, such as phallus morphology in males and karyology. Horáček et al. (2000) further supported this view. Finally, pioneering mitochondrial DNA marker analyses have clarified the taxonomic position of several vespetilionid bat genera, including *Hypsugo* (Hoofer, 2003; Mayer et al, 2007).

Currently, there are eighteen recognized *Hypsugo* species (Table 1.1). These species are distributed all around the Palearctic region, extending to sub-Saharan and Southern

latitudes of the Arabian Peninsula in the south (Kipson et al, 2023). Recent studies suggested that the sub-Saharan taxa might also represent a distinct genus, *Parahypsugo* (Hutterer et al, 2019).

Table 1.1 : Recognized species of the genus *Hypsugo* (Simmons & Cirranello, 2024).

Species Name	Common Name	Described
Hypsugo affinis	Chocolate Pipistrelle	(Dobson, 1871)
Hypsugo alaschanicus	Alashanian Pipistrelle	(Bobrinskii, 1926).
Hypsugo arabicus	Arabian Pipistrelle	(Harrison, 1979)
Hypsugo ariel	Fairy Pipistrelle	(Thomas, 1904)
Hypsugo cadornae	Cadorna's Pipistrelle	(Thomas, 1916)
Hypsugo dolichodon	Long-toothed Pipistrelle	(Görföl *et al*, 2014)
Hypsugo imbricatus	Brown Pipistrelle	(Horsfield, 1824)
Hypsugo kitcheneri	Red-brown Pipistrelle	(Thomas, 1915)
Hypsugo lanzai	Lanza's Pipistrelle	(Thomas, 1915)
Hypsugo lophurus	Burmese Pipistrelle	(Peters, 1866)
Hypsugo macrotis	Big-eared Pipistrelle	(Thomas, 1914)
Hypsugo mordax	Pungent Pipistrelle	(Peters, 1866)
Hypsugo musciculus	Mouse-like Pipistrelle	(Meyer, 1899)
Hypsugo petersi	Peter’s Pipistrelle	(Peters, 1871)
Hypsugo pulveratus	Chinese Pipistrelle	(Peters, 1870)
Hypsugo savii	Savi’s Pipistrelle	(Bonaparte, 1837)
Hypsugo stubbei	Stubbe’s Pipistrelle	(Stubbe, 2002)
Hypsugo vordermanni	Vordermann’s Pipistrelle	(Jentink, 1890)

Savi’s pipistrelle is a small-sized vespertilionid bat that closely resembles *Pipistrellus* species (Figure 1.2). The dark skin color of its wings, ears, and nose is one of the distinct features of this species. Its fur exhibits golden hair tips. Additionally, it has short, round ears with a blunt and broader tragus (Dietz & Kiefer, 2016). The penial morphology, which is a distinctive character among vespertilionid bats in the Palearctic region, has also its own unique shape. The baculum (penis bone) resembles the pipistrelloid bats in length, but it is more flat and has a roof-like section and disc-like endings, setting it apart from other species (Dietz & Kiefer, 2016; Kipson et al, 2023). The forearm length of *H. savii* varies between 31.4 mm and 37.9 mm; the fifth

finger, between 38 mm and 47 mm; the third finger, between 52 mm and 63 mm, and its weight, between 5 and 9 g (Dietz & Kiefer, 2016). *Hypsugo savii* can also be acoustically identified to a certain degree. It has quasi-contant frequency calls peaking between 30 and 35kHz, which are up to 16ms long (Dietz & Kiefer, 2016).

Figure 1.2 : A mature individual of *Hypsugo savii*.

The first record of Savi's pipistrelle, *Hypsugo savii,* belongs to Bonaparte (1837) from Pisa, Italy. It was recorded as *Vespertilio savii* since the genus *Hypsugo* was not recognized at the time (Ellerman & Morrison-Scott, 1966). Also, *Hypsugo* cf *darwinii* (Tomes, 1859) was described from the materials collected from Canary Islands by Charles Darwin and later it was classified with *Hypsugo savii* (Dondini et al. 2016).*Hypsugo savii* is a thermophilous species, primarily inhabiting rocky terrains and mountain regions, reaching up to 3,300 m above sea level (Kipson et al, 2023). It has a wide distribution range across the Western Palearctic region, spanning from Central Asia to the Western Mediterranean (Figure 1.3). Especially in the Middle East, it is one of the most common species (Uhrin et al, 2016; Kipson et al, 2023).

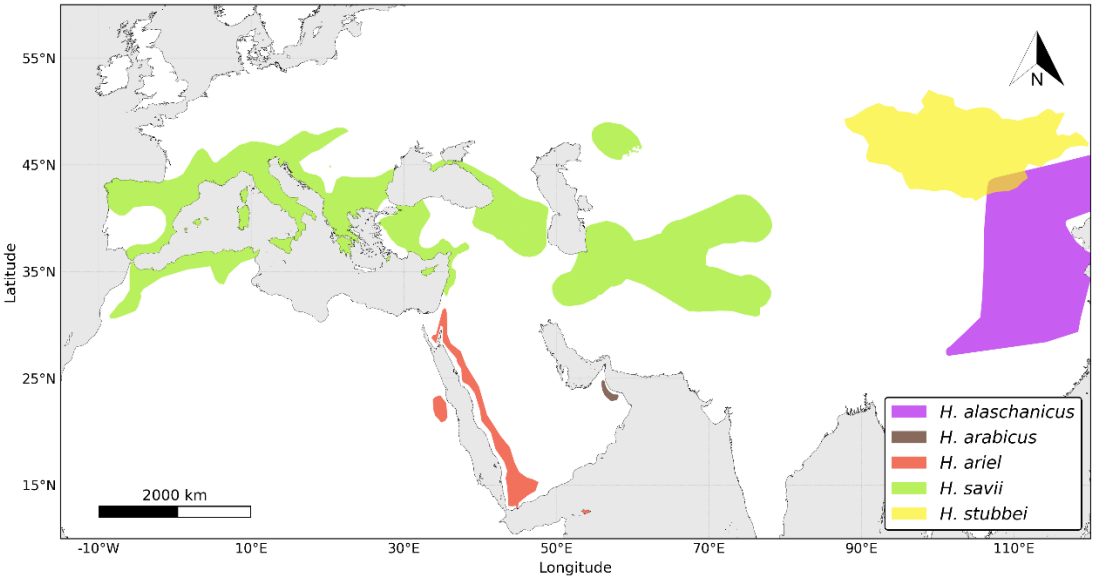

Figure 1.3 : The distribution map of *H. savii* based on IUCN. The distribution data for *H. stubbei* was adapted from Dolch et al. (2021).

1.1.3 Closely related species

Hypsugo savii has four closely related species in the region, some of which have been recently described (Figure 1.3):

1) *H. ariel:* Distributed along the west coast of the Arabian Peninsula, extending toward the eastern Mediterranean coast. It is also found in Sudan and Yemen (Dietz, 2005; Wilson & Reeder, 2005).

2) *H. arabicus*: Occurs in a very limited range in Oman and southern Iran (Benda et al. 2002, 2012).

3) *H. alaschanicus*: Found in East Asia, ranging from Mongolia to Japan (Kipson et al, 2023).

4) *H. stubbei*: A recently described species, distributed in Mongolia (Dolch et al, 2021).

The conservation status for *Hypsugo savii* in the International Union for Conservation of Nature (IUCN) is categorized as “Least Concern” and the population trend is listed as stable in the last assessment (Juste & Paunović, 2016). This species is strictly protected, listed under the Bern Convention on the Conservation of European Wildlife and Natural Habitats. The status for *Hypsugo alaschanicus* is also evaluated in “Least Concern” category. However, there is not enough knowledge of population sizes to

designate a trend (Fukui et al, 2019). *Hypsugo ariel* is listed as “Data Deficient” in the IUCN Red List and the current population trend is considered as unknown (Benda & Aulagnier, 2020*). Hypsugo arabicus*, formerly known as *Pipistrellus arabicus* as well, is also considered as “Data Deficient” in IUCN Red List based on the last assessment (Srinivasulu & Srinivasulu, 2019). Current population trend is unknown, however, it has potential to be common in the distribution range according to Benda et al. (2012). There is no available assessment report for *Hypsugo stubbei* included in IUCN Red List.

1.2 Genomic Analyses and Its Implications in Phylogenetics

1.2.1 Mitochonrial DNA

Mitochondria represent a fundamental element in eukaryotic cells. It is hypothesized that mitochondria originated through an ancient symbiotic event between ancestral eukaryotes and alphaproteobacteria, approximately two billion years ago. This evolutionary hypothesis initially emerged from the distinctive genomic characteristics of mitochondria. Mitochondrial genomes (often referred to as mitogenome of mtDNA) are circular, lack introns, and have relatively small sizes. They encode proteins for subunits within oxidative phosphorylation enzyme complexes, as well as components of the ribosome (rRNA) and transfer RNA (tRNA). Additionally, there is a small section of non-coding control region, commonly known as D-loop, which likely contains regulatory elements.

The animal mitochondrial genome is typically between 15 to 17 kilobases long making it smaller than the mitochondrial genomes of plants and fungi. There are thirty-seven genes and a control region on a typical vertebrate mitogenome. Among these genes, thirteen of them are protein-coding genes (PCGs), coding for NADH dehydrogenase complex (ND), cytochrome c oxidase complex (CO), cytochrome b (CYTB), and ATP synthase complex. Two genes are transcribed as rRNA for 12S and 16S subunits of ribosomes. Remaining twenty-two genes encode transfer RNAs of twenty amino acids which are Alanine (A), Arginine (R), Asparagine (N), Aspartic acid (D), Cysteine (C), Glutamic acid (E), Glutamine (Q), Glycine (G), Histidine (H), Isoleucine (I), Leucine (L), Lysine (K), Methionine (M), Phenylalanine (F), Proline (P), Serine (S), Threonine (T), Tryptophan (W), Tyrosine (Y), and Valine (V).

Mitochondrial DNA markers have been frequently utilized in phylogenetic studies and population genetics analysis, as they have several advantages. Firstly, they are easier to sequence as they have several copies. Secondly, mtDNA has a high mutation rate (reference). Higher mutation rate enables the accumulation of genetic differences in relatively shorter time periods. Lastly, mtDNA is maternally inherited. Lack of recombination allows tracking the maternal lineages. All of these features make mtDNA markers a powerful approach for DNA barcoding and taxonomic identification, population genetic and demographic analyses, and phylogenetic studies.

1.2.2 Increased resolution with high-throughput sequencing

In recent years, the reduced cost and increased practicality of high-throughput sequencing (also referred as next-generation sequencing) have revolutionized genomics (Sætre & Ravinet, 2019). All genetic material extracted from an individual sample is sequenced via high-throughput technologies in a method called whole genome approach. Whole genome sequencing is the most substantial method to reveal the genetic landscape and variations of an individual (Ng & Kirkness, 2010). Before the enhanced feasibility of this approach, genotyping of a certain loci on genomes and partial fragments of genes were widely used as "genetic markers." Target sequencing methods such as Restriction-site Associated DNA sequencing (RAD-seq) also provide high resolution, especially for wild populations and non-model organisms (Davey & Blaxter, 2010). However, whole-genome sequencing is taking over as the leading method for all genomic studies. A recent study showed the robustness of low-coverage whole genome sequencing (lcWGS) in investigating population structure across both model and non-model organisms. Notably, lcWGS proved effective across varying coverage levels and sampling sizes (Lou et al, 2021).

1.3 Literature Overview

Hypsugo savii has been considered a polytypic species for a very long time, with diverse subspecies notifications, ranging between five to ten (Ellerman & Morrison-Scott, 1966; Corbet, 1978). Horáček (2000) asserted that four of these subspecies, *savii*, *ochromixtus*, *caucasicus* and *austenianus*, are distinct and potentially represent cryptic species. Despite this differentiation, *H. savii* has maintained its overarching status across all claimed groups.

Pestano et al. (2003) revealed a deep genetic divergence between individuals collected from mainland Spain and Canary Islands. The differentiation was between 6.3% to 7.2% in CYTB and at least 3.8% in 16S rRNA gene. A similar finding was reported by Mayer et al. (2007) with a comparison between an individual from Morocco and a European individual. Partial ND1 gene of the Moroccan sample was at least 9.6% different from the European *H. savii*. In this study, a separate species, *Hypsugo darwinii* (which was historically described) was proposed to include a Moroccan sample and samples from Canary Islands in Pestano et al. (2003). More samples from Sicily, Sardinia (Veith et al, 2011), island of Tuscany, Italy (Dondini et al, 2016) and Malta (Mifsud and Vella, 2019) were examined and included in haplotypes of proposed *H.* cf *darwinii*. Ibáñez et al. (2006) stated that there are three lineages based on several genetic markers (CYTB, ND1 and RAG2) in Europe: 1) Southern Iberia, 2) Iberia to Switzerland, and 3) Eastern Mediterranean. Sympatric occurrence of 1 and 2 was shown. Similarly, García-Mudarra et al. (2009) presented three lineages as North African, West European, and East European. They also included CYTB and ND1 markers. Molecular diversity of *H. savii* was once again divided into three main groups with more samples from different areas like Turkey in Çoraman et al. (2013). The differentiations of Eastern Mediterranean samples from Western Mediterranean was around 7-8% and from Southwestern - North African, *H.* cf *darwinii*, was 9-8% in ND1 gene. These findings are consistent with other supporting studies (Bogdanowicz et al, 2015; Dolch et al, 2021), collectively enhancing our comprehension of the intricate genetic landscape of *H. savii*.

1.4 Aim of the Study

Previous studies have shown that *Hypsugo savii* harbours several deeply diverged lineages, some of which might represent distinct species (Kipson et al, 2023). Although most of the studies recovered these identified lineages, their relationships are not fully resolved. This lack of resolution was mostly due to the heterogenous marker selections. In this study, we aim to understand the phylogeographic patterns within the *H. savii* cryptic species complex with utilization of complete mitochondrial genomes. There are three main objectives of this thesis: 1) establishing a reproducible pipeline for assembling mitogenomes from whole genome sequencing data; 2) assembling the complete mitogenomes of *H. savii* and the related species; and 3)

reconstructing a phylogenetic tree for the *H. savii* complex based on whole mitogenomes.

2. MATERIALS AND METHODS

2.1 Data collection and Whole Genome Sequencing

For the genomic analysis, whole genome sequencing data from a total of 30 samples were utilized (Table 2.1). The specific locations are shown on the world map in Figure 2.1.Dr. Frieder Mayer at the Berlin Natural History Museum generated this dataset, which originates from the museum's sample collection. The following protocols were employed for DNA extraction and whole genome sequencing.

Table 2.1 : Specimen list and localities

Isolation Number	***Genus***	***Species***	***Location***	***Region***
125	*Hypsugo*	*stubbei*	Mongolia	Tsenkhar gol
128	*Hypsugo*	*stubbei*	Mongolia	Hoyer Tsenkher gol
213	*Hypsugo*	*alaschanicus*	Mongolia	Orog Nuur
332	*Hypsugo*	*alaschanicus*	Mongolia	Chatanbulag
4198	*Hypsugo*	*savii*	Morocco	Gorges du Todra
4674	*Hypsugo*	*savii*	Morocco	Gorges du Dades
5297	*Hypsugo*	*ariel*	Egypt	Wadi Nasb
5686	*Hypsugo*	*savii*	Greece	Agia Anastasia
5687	*Hypsugo*	*savii*	Greece	Kapelle am Neda-Fluss
5856	*Hypsugo*	*savii*	Greece	Crete
8842	*Hypsugo*	*savii*	Israel	Banyas
8843	*Hypsugo*	*savii*	Israel	Haslaga pools
9451	*Hypsugo*	*savii*	Kyrgyzstan	Aravan, Mandaniat
9668	*Hypsugo*	*savii*	Croatia	Trilj-lokacija
9681	*Hypsugo*	*savii*	Croatia	Cavtat, Miljasi, Lokva
9769	*Hypsugo*	*savii*	Armenia	North of Meghri
9840	*Hypsugo*	*savii*	Armenia	Noravank-Canyon
10102	*Hypsugo*	*savii*	Kyrgyztan	Sasyk-Unkur cave
10206	*Hypsugo*	*savii*	Slovenia	Roznik, Ljubljana
10207	*Hypsugo*	*savii*	Slovenia	Pond at Poletic
10426	*Hypsugo*	*savii*	Sardinia	Casa Oppes, Via De Martini
10431	*Hypsugo*	*savii*	Sardinia	Teletottes
10442	*Hypsugo*	*savii*	Sardinia	Rio di Curadoreddu
10445	*Hypsugo*	*savii*	Sardinia	Rio di Curadoreddu
10462	*Hypsugo*	*savii*	Sardinia	Vascone Cumbida Prantas
10522	*Hypsugo*	*savii*	Italy	Roma
10523	*Hypsugo*	*savii*	Italy	Roma
10524	*Hypsugo*	*savii*	Italy	Naples
10525	*Hypsugo*	*savii*	Italy	Naples
10526	*Hypsugo*	*savii*	Italy	Naples

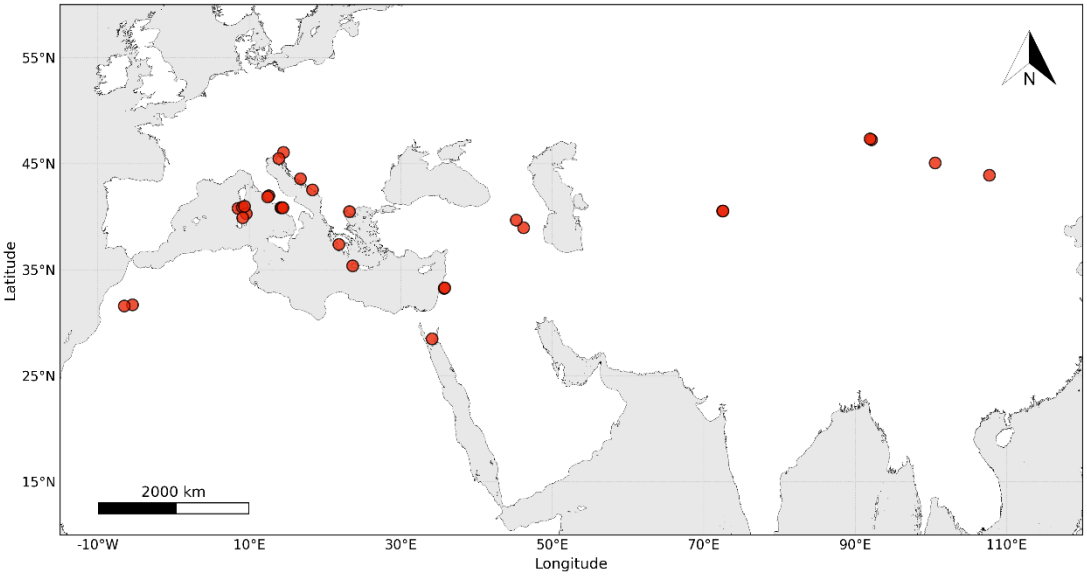

Figure 2.1: The distribution of localities of the thirty samples investigated in this study.

For DNA isolation, commercially available QIAgen™ Dneasy Blood & Tissue Kit was utilized, following the manufacturer's protocol. The isolated DNA concentrations were quantified using a Qubit fluorometer using and the Invitrogen™ double-stranded DNA (dsDNA) high-sensitivity kit. For whole genome sequencing, two commercial library kits were utilized, QIAgen™ QIAseq FX DNA Library Kit and illumina™ TruSeq Library Prep Kit, adhering to the respective manufacturer's protocols. During the library preparation, the DNA fragment size were controlled with the Agilent™ TapeStation instrument. The resulting DNA libraries had an average insert size of 280 bp.

DNA libraries containing adapter-ligated DNA fragments were sequenced using the illumina™ NovaSeq 6000 instrument. The sequencing was performed in paired-end mode with 150 bp reads. The resulting sequencing data were reported in Illumina 1.9 FastQ file format.

2.2 Sequence Analysis

A pipeline was constructed to assemble the complete mitogenomes of the analyzed samples (Figure 2.2)

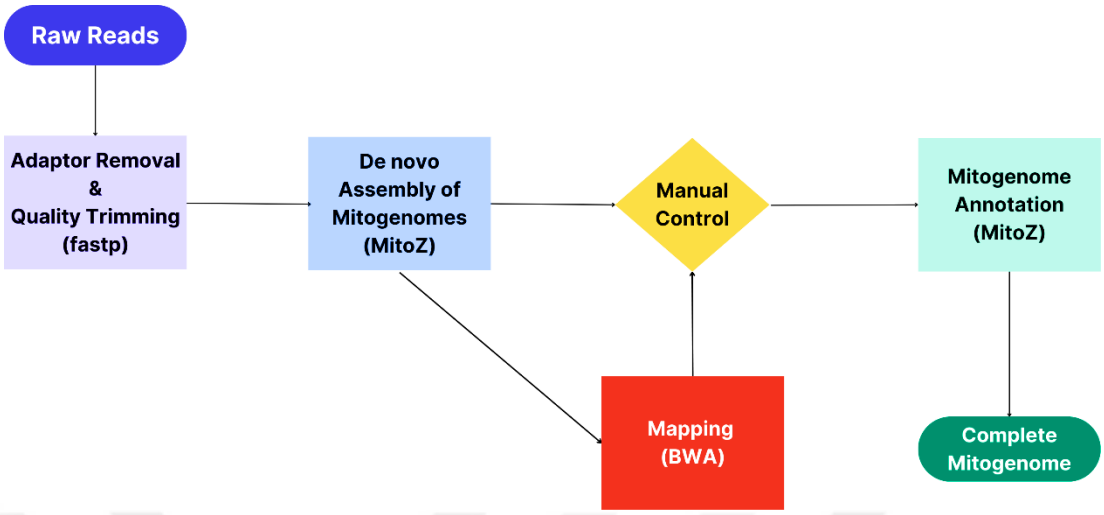

Figure 2.2: The pipeline constructed for the assemblying of complete mitogenomes.

2.2.1 Quality assessment

The quality of the raw FastQ files were assessed by utilizing fastp software (Chen et al, 2018). This algorithm detects the adapter sequences that were attached to the sequencing reads. These adapter sequences and low quality reads were removed from the data pool. Bases that have a Phred quality score above 33 were kept. If a Phred score is 33 for a base, this indicates that the probability of the error for this base is 0.05% and this base is a 99.95% correctly called base during the sequencing.

The reads dropped below than 35 bp-long after the quality and adapter trimming and PCR duplicates were removed. To handle the assembly of the repetitive regions, pate-pairs were not merged. The quality assessment of the whole process were recorded.

2.2.2 Complete mitochondrial assembly

The mitochondrial genomes were de novo assembled with the MitoZ software (Meng et al, 2019). MitoZ is specifically designed for animal mitochondrial genomes with and has faster and more accurate algorithm in comparison to other methods available (Mahar et al, 2023).

MitoZ offers separate modes for different phyla in animals. For each phylum, a distinct database is utilized to initialize the assembly. Additionally, various assembler options,

including data size, time-bound, and k-mer sizes can be set. The de novo analyses in this study were run in 'chordata' mode using the 'megahit' assembler with 59, 79 and 99 k-mer sizes.

2.2.3 Mapping of reads

The initial assembly products may include gaps, particularly if there are repetitive regions in the sequences. By mapping the cleaned data set to the assembled mitogenomes, these gaps can be closed. Additionally, mapping serves as a means of validation for the de novo assemblies.

For the mapping analyses, the Burrow-Wheeler Aligner (bwa) algorithm (Li, 2013) was utilized. The FASTA files of the de novo assembled mitogenomes were used as reference. The associated reads of the specimens were mapped onto these reference sequences with default parameters. Both forward and reverse sequences were utilized. The mapping statistics, including the number of reads mapped to the reference, the amount and percent of covered bases, mean coverage, and depth, were recorded.

The mappings were individually inspected and the repeat regions in the D-loop segment were manually edited. Geneious Prime (Geneious Prime, 2024) was used for the manual inspection and editing. As a result, a complete mitochondrial genome sequence was generated for each individual in FASTA format.

2.2.4 Mitogenome annotation and tRNA profiling

In the final step, the mitogenome of each studied species was annotated using two different algorithms: The MITOS software (Bernt et al, 2013) via the Galaxy Platform (https://usegalaxy.eu/root?tool_id=toolshed.g2.bx.psu.edu%2Frepos%2Fiuc%2Fmitos2%2Fmitos2%2F2.1.3%20galaxy0) and MitoZ. Through these analyses, the positions, lengths, and orders of the mitochondrial genes were determined. The results from both algorithms were compared to ensure the reliability of the annotations. A gene list was created, providing exact location information on the sequences. Additionally, the MitoZ algorithm's annotation command allowed visualization of gene locations and the circular structure of the mitogenomes using a subpackage called OrganellarGenomeDRAW (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html). Furthermore, the same annotation visualizations displayed the mapping coverage of the mitogenomes.

Finally, the transfer RNA (tRNA) genes in the mitogenome of *H. savii* were identified using tRNAscan-SE Search Server (http://lowelab.ucsc.edu/tRNAscan-SE/). The positions of the relevant genes were then confirmed with the previous annotations. Additionally, the secondary structures of the tRNAs were predicted.

For the annotations, the following samples were selected: Sample 332, *H. alaschanicus* from Mongolia; Sample 5297, *H. ariel* from Egypt; Sample 4198, *H. savii* from Morocco; and Sample 128, *H. stubbei* from Mongolia.

2.3 Phylogenetic Analyses

The phylogenetic relationships of the analysed samples were investigated by reconstructing a phylogenetic tree and a haplotype network (Figure 2.3). Given the high variability and repetitive sequences in D-loop regions, they were extracted from the mitogenomes, leaving the remaining sequences for further analysis.

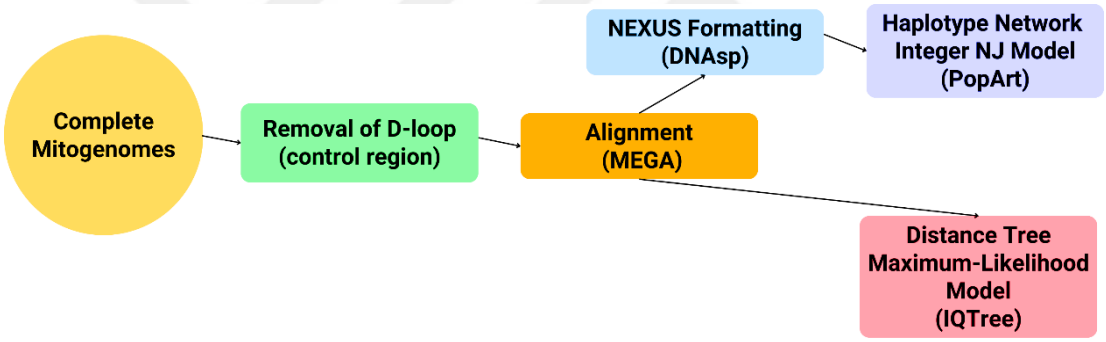

Figure 2.3: The illustration of the workflow of phylogenetic analyses.

The whole mitogenomes (excluding the D-loop regions) were aligned using the MUSCLE (MUltiple Sequence Comparison by Log-Expectation) alignment model (Edgar, 2004), as implemented in MEGA (Molecular Evolutionary Genetics Analysis) 11.0 software (Tamura et al, 2021). The publicly available complete mitochondrial sequence of *Pipistrellus kuhlii* (GenBank accession number: KU058655) was used as an outgroup for the phylogenetic reconstructions. Subsequently, the final alignments were visually inspected in Geneious Prime to identify any poorly aligned regions.

Finally, the alignments were exported as a MEGA file (.meg) and converted to a NEXUS file (.nex) using the DnaSP 6.0 software (Rozas et al, 2017). The pairwise genetic distance between mitochondrial sequences of individuals was calculated in

MEGA 11.0, using the Tamura-Nei model (Tamura & Nei, 1993) to estimate the nucleotide differentiation ratio of each sample

The Maximum-Likelihood estimation model was selected for phylogenetic reconstructions. Two different methods were employed to construct the phylogenetic distance tree. In the first method, IQ-TREE v.2.3.5 (http://www.iqtree.org/) was used with default settings. In the second approach, RAxML v.8 (Stamakis, 2014) was utilized via CIPRES Science Gateway v.3.3 (https://www.phylo.org/portal2/). The analysis ran with 100 bootstrap iterations, using a random seed value of 12345 for rapid bootstrapping. The GTRCAT model was selected for the bootstrapping phase. The best substitution model was chosen using the ModelFinder algorithm (Kalyaanamoorthy et al,. 2017) with 1000 bootstrapping iterations. Reconstructed phylogenetic trees were visualized with FigTree v.1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) and iTOL:Interactive Tree of Life Web Server (http://itol.embl.de).

In the final step, a haplotype network with complete mitochondrial sequences (D-loop regions excluded) was built with PopART v.1.7 (Leigh et al, 2015) utilizing the integer Neighbor-Joining model (Saitou & Nei, 1987).

3. RESULTS

3.1 Evaluation of Data Filtering

A total number of filtered paired-end reads (forward and reverse) obtained for each sample varied between 28 and 233 million reads, with an average of 123 million (Table 3.1). Mean length of reads was 150 bp. The average quality of reads was around 31 to 36, which are considered to be high quality. The bases that have quality scores lower than 33 were generally located at the end of the reads, between the 140^{th} and 150^{th} bases. The average duplication rate in the readpools was 3,58%, ranging between 1,96% and 9,06%.

In the filtering step, the number of reads reduced by an average of 4,38%, ranging between 0,85% and 18,48%. The average number of filtered reads which are utilized in the mitochondrial assembly pipeline was around 123 million reads, ranging between 29 and 232 million reads.

3.2 Mitochondrial Genomes and tRNA Profiles

The mitochondrial sequences assembled with MitoZ differentiated in their length. Although some of these sequences were not automatically circularized by the algorithm, all of the mitochondrial genes were present in all of the assemblies. For *H. savii* samples, the algorithm estimated the closest complete mitochondrial sequence as the mitogenome of *Pipistrellus kuhlii* in its database.

The obtained mitogenomes were inspected in *Geneious Prime* through alignment. The highest variation among the sequences were located in the control region (D-loop region). This region was also the only part of the mitogenome where alignment gaps were identified. The D-loops region is characterized by two main repeating motifs; the first one consists of 81 bp repeat sequences, occurring at least three times in the beginning of the D-loop.

The second motif features a microsatellite sequence pattern ‘TACGCA,’ reperating multiple time. The exact repeat number of this latter microsatellite region could not be

estimated due to its length exceeding the read length (150 bp). Notably, this repeat region exhibited variation both among individuals and among species.

To assess the quality of assemblies and circularize those that were not already circular, the reads were mapped onto the de novo assembled mitogenomes. The highest number of reads mapped to the assembled mitogenomes was around 1,5 million, while the lowest was around 15 thousand.

Overall, the mean depths were consistent across the mitogenome, except for the D-loop region. The average mean deapth and the median mean depth were calculated as 1176X and 600X, respectively. The maximum mean depth reached approximately 13000X, whereas the minimum was around 150X. All of the mitogenomes were succesfully circularized after manual editing.

The de novo assembled mitogenomes exhibited typical circular mammalian mitochondrial genomes (Figure 3.1). The average mitogenome length was approximately 16,600 bp, with the exception of *H. stubbei*, which was approximately 600 bp shorter (Table 3.1).

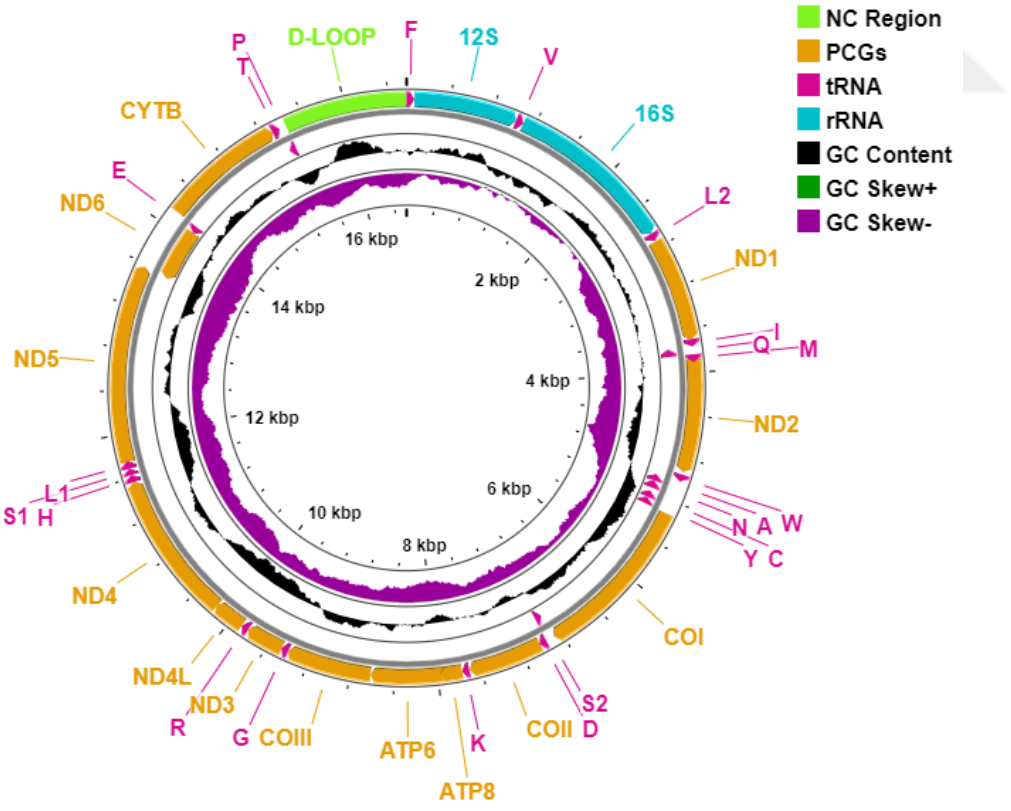

Figure 3.1: The annotated complete mitogenome of *H. savii*. Transfer RNA genes are indicated by the single letter for their corresponding amino acid.

Among conspecifics, slight variations were observed due to differences in the length of the D-loop region, overall lengths remained comparable. The nucleotide composition of the mitogenomes were asymmetrical in accordance with the AT-GC skew.

The average GC content was approximately 35% (Table 3.1). For all the analysed species, all 37 mammalian mitochondrial genes were identified (Table 3.2 and Figure 3.1).

Among these genes, two encoded ribosomal RNA (12S and 16S subunits), while twenty-two encoded transfer RNAs. Each tRNA gene on the mitochondrial sequence is shown in Figure 3.1 with the symbol of the corresponding amino acid. The list of amino acids are given in Table 3.2. The remaining thirteen genes encoded proteins for mitochondrial enzyme complexes. Those complexes are the elements of the mitochondrial respiratory chain as members of the center for cell metabolism. Additionally, a long non-coding control region (D-loop) was identified within the sequence. This sequence was around 1.1 kb and varied from individual to individual.

The profiles of the tRNA genes were investigated (Figure 3.2), assessing their length and positions. In the *H. savii* mitogenome The total length of the tRNA genes were 1,598 bp . The length of the tRNA genes ranged between 58 bp and 74 bp. The predicted secondary structures of the tRNA genes exhibited characteristic cloverleaf structures, with the exception of $tRNA^{Ser}$, where the D arm of a complete cloverleaf was missing.

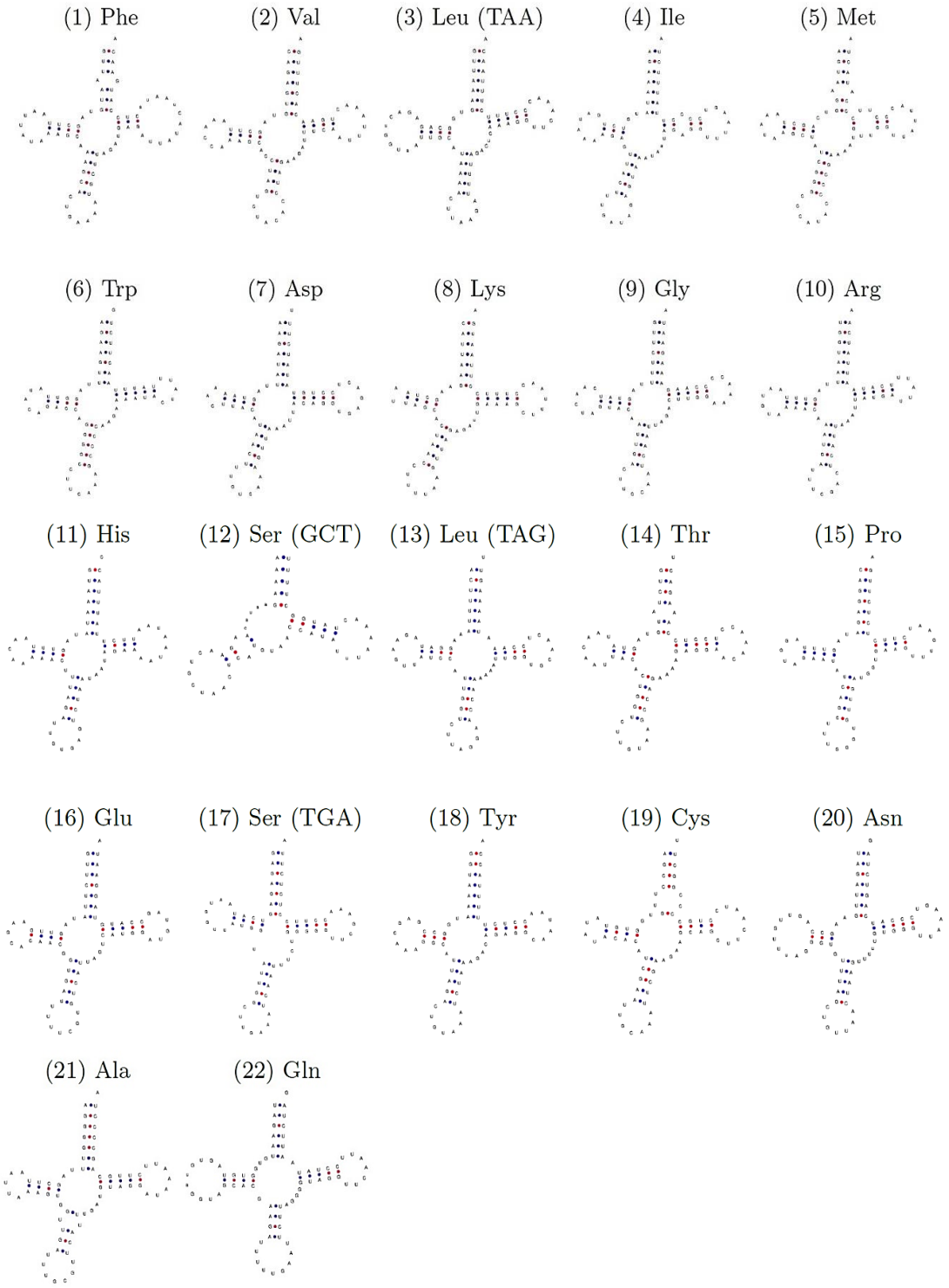

Figure 3.2: The secondary structures of tRNAs, encoded with *H. savii* mitogenome.

Table 3.1 : Filtration, assembly and mapping statistics of the read pools for each sample.

			Number of Reads			Elimination in Filtering (%)	Length of Mitogenomes (base-pair)		GC Content (%)	Coverage (%)	Mean Depth (X)
Isolation Number	Genus	Species	Raw Pool	Filtered Pool	Mapped on Mitogenome		Assembled	D-loop extracted	D-loop extracted		
125	*Hypsugo*	*stubbei*	72,958,720	69,688,062	69,111	4.48	16,822	15,435	35.2	100	600
128	*Hypsugo*	*stubbei*	181,234,670	172,183,802	81,665	4.99	15,987	15,435	35.2	100	750
213	*Hypsugo*	*alaschanicus*	241,444,810	233,495,648	1,515,095	3.29	16,618	15,432	35.3	100	13000
332	*Hypsugo*	*alaschanicus*	140,599,254	136,826,534	458,924	2.68	16,579	15,433	35.3	100	5000
4198	*Hypsugo*	*savii*	131,362,690	127,326,516	73,753	3.07	16,571	15,432	35.7	100	600
4674	*Hypsugo*	*savii*	130,114,166	127,298,956	83,695	2.16	16,727	15,432	35.7	100	700
5297	*Hypsugo*	*ariel*	100,290,796	81,761,528	62,497	18.48	16,655	15,419	34.9	100	300
5686	*Hypsugo*	*savii*	131,387,124	124,744,774	84,464	5.06	15,942	15,437	35.8	100	750
5687	*Hypsugo*	*savii*	112,728,448	108,126,210	66,103	4.08	15,906	15,435	35.7	100	550
5856	*Hypsugo*	*savii*	56,297,310	51,665,172	34,469	8.23	15,870	15,435	35.8	100	300
8842	*Hypsugo*	*savii*	98,062,814	95,545,512	64,890	2.57	16,684	15,433	35.8	100	600
8843	*Hypsugo*	*savii*	81,843,276	79,113,604	41,426	3.34	16,599	15,433	35.8	100	400
9451	*Hypsugo*	*savii*	98,057,316	93,561,474	19,048	4.58	16,514	15,434	35.2	100	150
9668	*Hypsugo*	*savii*	89,256,250	84,143,118	20,611	5.73	15,899	15,435	35.8	100	150
9681	*Hypsugo*	*savii*	94,620,174	89,011,720	14,921	5.93	15,907	15,436	35.8	100	100
9769	*Hypsugo*	*savii*	137,917,380	132,256,884	144,507	4.10	16,528	15,432	35.7	100	1200
9840	*Hypsugo*	*savii*	72,499,182	68,908,562	144,292	4.95	16,686	15,431	35.7	100	1200
10102	*Hypsugo*	*savii*	117,732,824	116,748,492	76,235	0.84	16,560	15,435	35.2	100	650
10206	*Hypsugo*	*savii*	87,897,852	84,070,950	56,132	4.35	16,570	15,431	35.6	100	500
10207	*Hypsugo*	*savii*	30,078,636	28,748,126	140,794	4.42	16,624	15,431	35.6	100	1200
10426	*Hypsugo*	*savii*	157,671,152	150,494,832	54,686	4.55	15,984	15,431	35.6	100	500
10431	*Hypsugo*	*savii*	189,762,620	183,314,346	28,142	3.40	16,610	15,435	35.6	100	200
10442	*Hypsugo*	*savii*	152,757,912	144,625,460	79,204	3.02	16,688	15,432	35.7	100	650
10445	*Hypsugo*	*savii*	158,360,520	152,629,830	14,514	3.62	15,972	15,431	35.7	100	150
10462	*Hypsugo*	*savii*	161,231,430	154,983,704	74,230	3.88	16,617	15,431	35.6	100	650
10522	*Hypsugo*	*savii*	132,917,780	127,420,348	141,354	4.14	16,653	15,431	35.6	100	1250
10523	*Hypsugo*	*savii*	131,983,814	127,784,062	69,901	3.18	16,651	15,431	35.6	100	600
10524	*Hypsugo*	*savii*	233,921,564	231,722,800	115,731	0.94	16,651	15,431	35.6	100	1000
10525	*Hypsugo*	*savii*	212,582,992	204,090,722	116,167	3.99	16,714	15,431	35.6	100	1000
10526	*Hypsugo*	*savii*	126,577,414	122,314,960	69,624	3,37	16,639	15,431	35.6	100	600

Table 3.2 : The list of genes annotated on the complete mitogenome of *H.savii*.

Gene	*Type*	*Abbreviation*	*Start*	*End*	*Strand*
tRNA-Phenylalanine	transfer RNA	$tRNA^{Phe}$	0	72	+
Ribosomal RNA for small subunit : SSU (12S)	ribosomal RNA	*srRNA*	72	1027	+
tRNA-Valine	transfer RNA	$tRNA^{Val}$	1027	1096	+
Ribosomal RNA for large subunit : LSU (16S)	ribosomal RNA	*lrRNA*	1095	2657	+
tRNA-Leucine (Leu-CUN)	transfer RNA	$tRNA^{Leu}2$	2659	2734	+
NADH dehydrogenase, subunit 1	protein coding	*ND1*	2739	3690	+
tRNA-Isoleucine	transfer RNA	$tRNA^{Ile}$	3695	3764	+
tRNA-Glutamine	transfer RNA	$tRNA^{Gln}$	3761	3835	-
tRNA-Methionine	transfer RNA	$tRNA^{Met}$	3835	3904	+
NADH dehydrogenase, subunit 2	protein coding	*ND2*	3904	4942	+
tRNA-Tryptophan	transfer RNA	$tRNA^{Trp}$	4946	5013	+
tRNA-Alanine	transfer RNA	$tRNA^{Ala}$	5017	5086	-
tRNA-Asparagine	transfer RNA	$tRNA^{Asn}$	5088	5161	-
tRNA-Cysteine	transfer RNA	$tRNA^{Cys}$	5194	5261	-
tRNA-Tyrosine	transfer RNA	$tRNA^{Try}$	5261	5329	-
Cytochrome c oxidase, subunit 1	protein coding	*COI*	5330	6863	+
tRNA-Serine (Ser-AGY)	transfer RNA	$tRNA^{Ser}2$	6878	6947	-
tRNA-Aspartic acid	transfer RNA	$tRNA^{Asp}$	6954	7021	+
Cytochrome c oxidase, subunit 2	protein coding	*COII*	7021	7702	+
tRNA-Lysine	transfer RNA	$tRNA^{Lys}$	7708	7775	+
ATP synthase, Fo subunit 8	protein coding	*ATP8*	7776	7971	+
ATP synthase, Fo subunit 6	protein coding	*ATP6*	7937	8612	+
Cytochrome c oxidase, subunit 3	protein coding	*COIII*	8617	9400	+
tRNA-Glycine	transfer RNA	$tRNA^{Gly}$	9401	9470	+
NADH dehydrogenase, subunit 3	protein coding	*ND3*	9470	9815	+
tRNA-Arginine	transfer RNA	$tRNA^{Arg}$	9817	9885	+
NADH dehydrogenase, subunit 4L	protein coding	*ND4L*	9886	10180	+
NADH dehydrogenase, subunit 4	protein coding	*ND4*	10176	11544	+
tRNA-Histidine	transfer RNA	$tRNA^{His}$	11554	11622	+
tRNA-Serine (Ser-UCN)	transfer RNA	$tRNA^{Ser}1$	11622	11681	+
tRNA-Leucine (Leu-UUR)	transfer RNA	$tRNA^{Leu}1$	11682	11752	+
NADH dehydrogenase, subunit 5	protein coding	*ND5*	11758	13555	+
NADH dehydrogenase, subunit 6	protein coding	*ND6*	13562	14081	-
tRNA-Glutamic acid	transfer RNA	$tRNA^{Glu}$	14084	14152	-
Cytochrome b	protein coding	*CYTB*	14157	15291	+
tRNA-Threonine	transfer RNA	$tRNA^{Thr}$	15297	15366	+
tRNA-Proline	transfer RNA	$tRNA^{Pro}$	15365	15432	-

3.3 Phylogenetic Analyses

Based on the results, the IQ-TREE analysis was done with the TIM2 substitution model with Gamma rate variation. Both of the phylogenetic reconstruction methods produced identical topologies, revealing six main clades (Figure 3.3). The individual representing *H. ariel* was located at the basal node of the group. The second major branch separated two individuals representing *H. alaschanicus* from the rest of the group. In the subsequent branch, two *H. stubbei* individuals from Mongolia and two individuals recorded as *H. savii* from Kyrgyzstan clustered together. Finally, the *H. savii* individuals formed three main clades: i) one clade in Sardinia; ii) another clade spanning Croatia, Greece, Armenia, Israel; and iii) a third clade in western Europe (including Sardinia) and Morroco. The first of these clustered together and then merged with the latter.

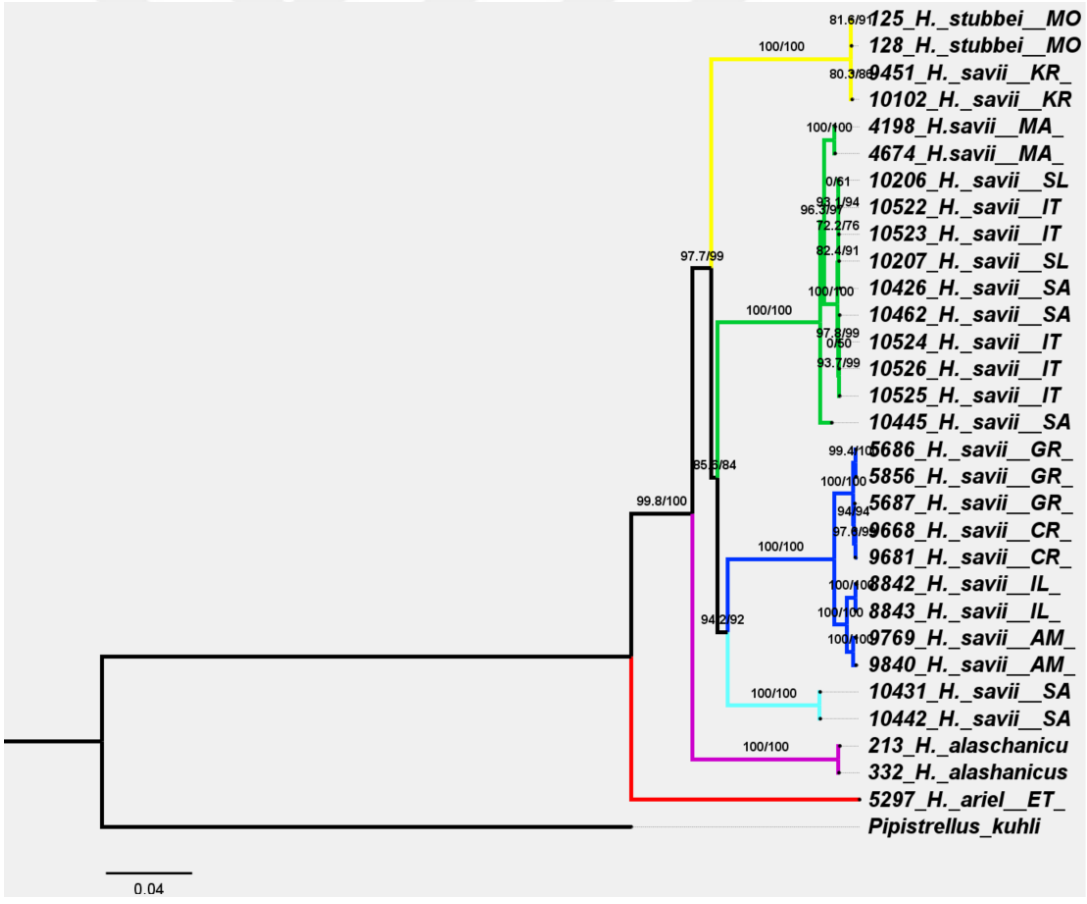

Figure 3.3: The distance tree of the samples analyzed with the ML model.

	OG	5297	332	213	125	128	9451	10102	10442	10431	8842	8843	9769	9840	5856	5686	5687	9668	9681	10445	4674	4198	10462	10525	10524	10526	10426	10206	10207	10522	10523
OG																															
5297	0.203																														
332	0.200	0.123																													
213	0.201	0.123	0.001																												
125	0.201	0.127	0.100	0.101																											
128	0.201	0.127	0.101	0.101	0.000																										
9451	0.202	0.127	0.101	0.101	0.001	0.001																									
10102	0.201	0.127	0.101	0.101	0.001	0.001	0.001																								
10442	0.196	0.119	0.093	0.094	0.085	0.085	0.086	0.086																							
10431	0.196	0.119	0.093	0.093	0.085	0.085	0.085	0.085	0.001																						
8842	0.203	0.130	0.100	0.101	0.094	0.094	0.094	0.094	0.080	0.080																					
8843	0.203	0.130	0.100	0.101	0.094	0.094	0.094	0.094	0.079	0.079	0.001																				
9769	0.203	0.130	0.101	0.101	0.095	0.095	0.095	0.095	0.080	0.080	0.008	0.008																			
9840	0.203	0.130	0.101	0.102	0.094	0.094	0.095	0.094	0.080	0.080	0.008	0.008	0.003																		
5856	0.203	0.131	0.102	0.102	0.095	0.095	0.096	0.095	0.080	0.080	0.020	0.019	0.020	0.020																	
5658	0.202	0.131	0.102	0.102	0.095	0.095	0.096	0.095	0.080	0.080	0.020	0.019	0.020	0.020	0.001																
5687	0.202	0.131	0.102	0.102	0.094	0.094	0.095	0.094	0.079	0.080	0.019	0.019	0.019	0.019	0.002	0.002															
9668	0.202	0.131	0.102	0.102	0.094	0.094	0.095	0.095	0.079	0.080	0.019	0.019	0.019	0.019	0.002	0.002	0.001														
9681	0.202	0.131	0.102	0.102	0.094	0.094	0.095	0.094	0.079	0.080	0.019	0.019	0.019	0.019	0.002	0.002	0.001	0.000													
10445	0.197	0.122	0.094	0.094	0.090	0.090	0.090	0.090	0.078	0.078	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.086	0.086												
4674	0.199	0.122	0.095	0.095	0.092	0.092	0.092	0.092	0.079	0.079	0.088	0.088	0.089	0.088	0.088	0.088	0.088	0.088	0.088	0.012											
4198	0.198	0.122	0.095	0.095	0.091	0.091	0.092	0.091	0.078	0.078	0.088	0.088	0.088	0.088	0.088	0.088	0.088	0.088	0.088	0.012	0.001										
10462	0.198	0.122	0.096	0.096	0.091	0.091	0.091	0.091	0.080	0.080	0.089	0.089	0.089	0.089	0.088	0.088	0.088	0.088	0.088	0.014	0.012	0.012									
10525	0.198	0.122	0.096	0.096	0.091	0.091	0.091	0.091	0.081	0.081	0.089	0.089	0.089	0.089	0.088	0.088	0.088	0.088	0.088	0.014	0.012	0.012	0.001								
10524	0.199	0.122	0.096	0.096	0.091	0.091	0.091	0.091	0.080	0.080	0.089	0.089	0.089	0.089	0.088	0.088	0.088	0.088	0.088	0.013	0.012	0.012	0.001	0.000							
10526	0.199	0.122	0.096	0.096	0.091	0.091	0.091	0.091	0.080	0.080	0.089	0.089	0.089	0.089	0.088	0.088	0.088	0.088	0.088	0.013	0.012	0.012	0.001	0.000	0.000						
10426	0.198	0.122	0.096	0.096	0.091	0.091	0.092	0.091	0.080	0.080	0.089	0.089	0.089	0.089	0.089	0.088	0.088	0.088	0.088	0.014	0.012	0.012	0.002	0.002	0.002	0.002					
10206	0.198	0.122	0.095	0.095	0.091	0.091	0.091	0.091	0.080	0.080	0.089	0.089	0.089	0.089	0.088	0.088	0.088	0.088	0.088	0.013	0.012	0.012	0.001	0.001	0.001	0.001	0.001				
10207	0.198	0.122	0.096	0.096	0.091	0.091	0.091	0.091	0.080	0.080	0.089	0.089	0.089	0.089	0.088	0.088	0.088	0.088	0.088	0.014	0.012	0.012	0.002	0.002	0.001	0.001	0.002	0.001			
10522	0.197	0.122	0.095	0.095	0.091	0.091	0.091	0.091	0.080	0.080	0.089	0.089	0.089	0.089	0.088	0.088	0.088	0.088	0.088	0.013	0.011	0.011	0.001	0.001	0.001	0.001	0.001	0.000	0.001		
10523	0.198	0.122	0.095	0.095	0.091	0.091	0.091	0.091	0.080	0.080	0.089	0.089	0.089	0.089	0.088	0.088	0.088	0.088	0.088	0.013	0.012	0.012	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.000	

Figure 3.4: The distance matrix of the samples based on ML model.

The pairwise genetic distance analysis revealed distinct differentiation levels among various *Hypsugo* samples (Figure 3.4). *Pipistrellus kuhlii* mitogenome used as an outgroup, exhibited the highest variation, with at least a 19.6% difference from all other mitogenomes. *H. ariel* (Sample 5297) showed differentiation ranging from 11.9% (with samples 10442 and 10431) to 13.1% (with samples in the blue group: 5856, 5686, 5687, 9668, and 9681). *H. alaschanicus* (Samples 213 and 332) differed from other *Hypsugo* samples by 9.3% to 10.2%. The yellow cluster including *H. stubbei* (Samples 125, 128, 9451, 10102) were closely related based on morphological assessment. The turquoise cluster (Samples 10442 and 10431) exhibited slight differentiation from each other, however, showed an 8% (±0.01) difference from other *Hypsugo* samples. Within the blue cluster, Israeli samples (Samples 8842, 8843) and Armenian samples (Samples 9769, 9840) were closely related, with only a 0.08% difference between these subgroups. However, these subgroups within the blue cluster differentiated from the rest of the group samples by around 2%, and by at least 8.6% from other *Hypsugo savii* samples. The Moroccan samples (Sample 4674, 4198) in the green cluster, were in close proximity to each other but with about 1.2% differentiation from the other green cluster samples. One of the Sardinian samples (Sample 10445) within the green cluster exhibited a 1.3% difference from the rest of the green cluster, while the remaining of this cluster samples (Samples 10462, 10525, 10526, 10426, 10206, 10207, 10522, and 10523) had no major differences, showing less than a 0.3% variation among them.

Accordingly, with the IQ-TREE phylogenetic tree and the pairwise distance matrix, the tree generated with RAxML is edited with colors based on clades (Figure 3.5). The yellow cluster includes *H. stubbei* samples and *H. savii* samples from Kryzygstan in Central Asia (CA), the blue cluster shows Eastern Mediterranean (EM) samples of *H. savii*, and the green cluster represents Western Mediterranean (WM) samples of H.savii. Following the literature (Pestano et al, 2003; Ibáñez et al, 2006; Mayer et al, 2007; García-Mudarra et al, 2009; Veith et al, 2011; Çoraman et al, 2013), three solid clades occupy European continent. Therefore, the remaining Samples 10442 and 10431 from Sardinia are claimed as the third major clade from the North African/ Southwestern Mediterranean (SWM) (as Mayer et al, (2007) proposed, H. cf darwinii) in the turquoise group.

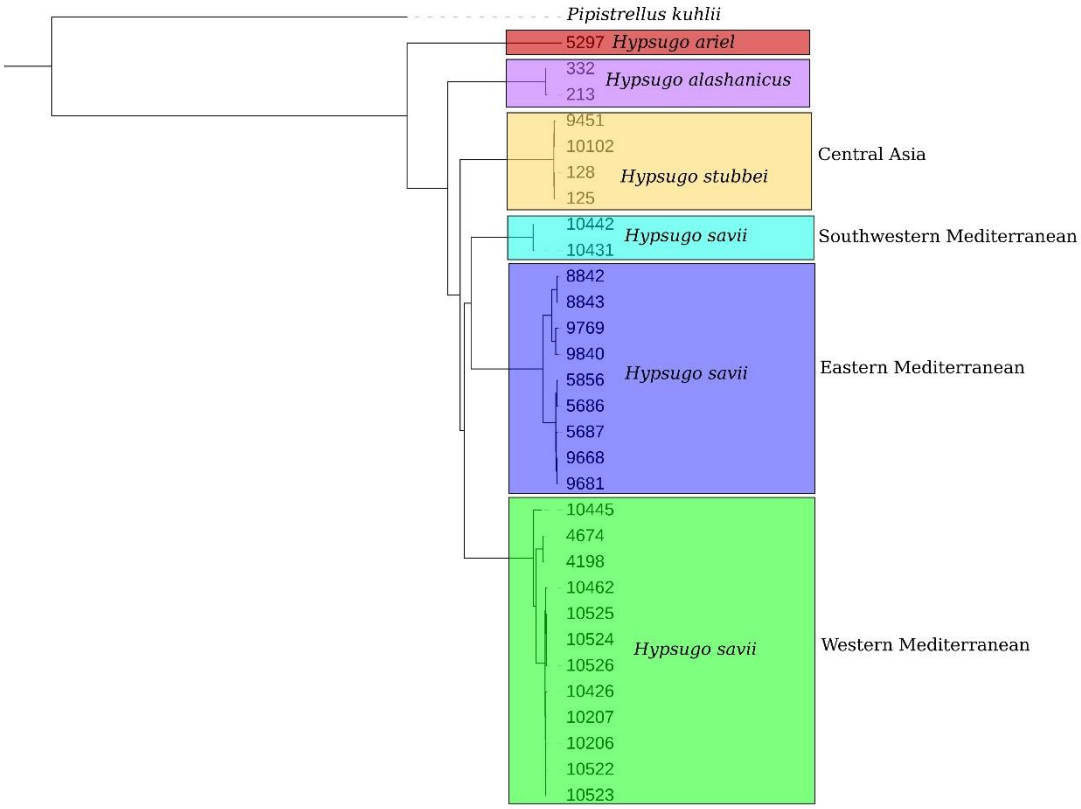

Figure 3.5: Estimated geographical clades are labeled on the phylogenetic tree generated with the ML model.

A haplotype network analysis was performed with PopArt using the integer NJ model. The general structure of the network and the grouping of the individuals were concordant with the phylogenetic trees (Figure 3.6).

Based on the objective of the study, an optimal pipeline was generated to assemble complete mitogenomes from whole genome sequencing, the complete mitogenomes of of all thirty samples were obtained. The phylogenetic analyses were performed with the obtained complete mitogenomes to establish an understanding of the phylogeography of *Hypsugo savii* cryptic species complex. Based on the phylogenetic analyses, the species complex was divided into related subgroups according to their geographic patterns.

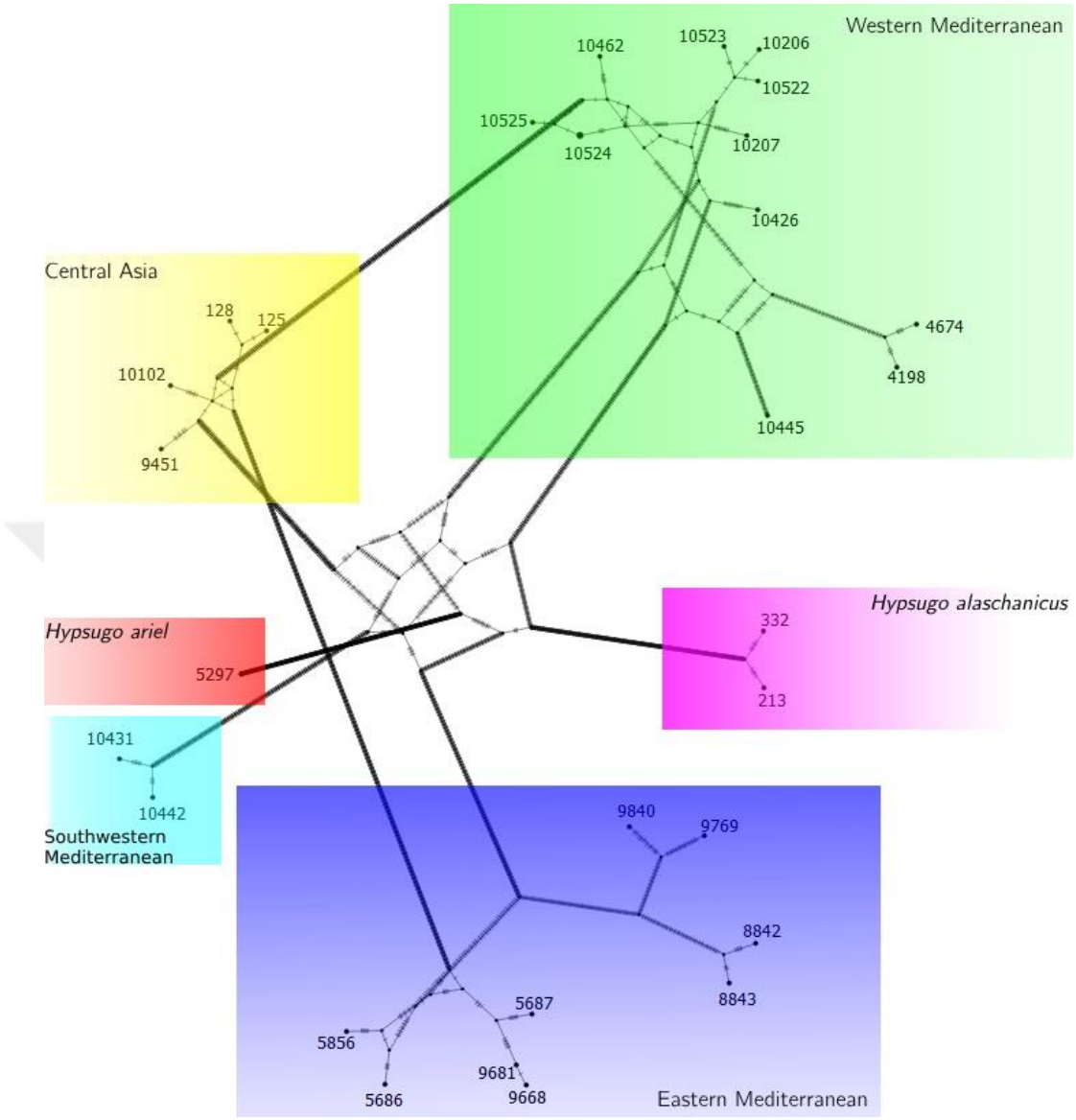

Figure 3.6: The haplotype network of the complete mitogenomes of the samples.

4. DISCUSSION

This study provides an integrated and homogenous approach for the investigation of phylogeographic patterns at the genomic level with complete mitogenomes. The utilization of mixed and limited genetic markers is the root cause of the lack of validation between accumulating work on the species of interest (Kipson et al, 2023). An alternate way of phylogenetic reconstruction is presented. Whole genome perspective contributes to the establishment of the consensus for previously proposed intraspecific diversity of *Hypsugo savii* cryptic species complex (Çoraman et al, 2013).

Three main objectives were determined for this study. The first objective was accomplished with the pipeline for the assembly of mitogenomes. A regeneratable methodological installment was aimed to be used in future studies for many species. Therefore, different algorithms were trialed to achieve the fastest and most accurate results with easy use.

Secondly, the complete mitogenomes were constructed as it was targeted. Except for the complete mitogenome of *Hypsugo alaschanicus* (GenBank ID: MK135784), there is no complete mitogenome available for the species included in this study. The mitogenomes produced in this study will be added to GenBank to contribute to further studies.

As the last objective, the phylogeographic patterns were investigated through the construction of phylograms with complete mitogenomes as a highly resolved, homogenous dataset. Mutational distance, gene rearrangements and mitochondrial morphology in a complete mitogenome are valuable in the sense of the phylogenetic position of closely related taxa (Cameron, 2014).

4.1 Data Assessment

The sample collection, library preparation and sequencing of the isolated DNA from the samples were performed prior to this study. The raw reads are polished with fastp tool for elimination of adaptors, low-quality reads and PCR duplications. fastp is a

multi-functioned and quite fast tool for preprocessing high-throughput sequencing data (Chen et al, 2018). It is shown to have high performance in comparison with other tools (Hao et al,2022) and the all-in-one functions ease the multi-stepped preprocessing.

The reports from fastp analyses have shown the previous and past situations of the read pools during data filtering. The number of reads in the readpool of each sample was highly different. The largest number was 8 times higher than the smallest read pool. This is thought to be a reason for varied sampling from various regions, different times of storage, and dissimilar pre-sequencing and sequencing protocols.

Even though the numbers of total reads diverged, the rates of eliminated reads during the filtering were consistently proportional to their total reads. There was only Sample 5297 as an exception with 18,48% eliminated reads. The report and mapping statistics were not incongruent in comparison with other samples. The relatively higher number of eliminated reads might be a result of the procedures prior to sequencing or sequencing itself. The mapping statistics of this sample also showed no abnormal pattern to reconsider the inclusion of the sample in the study.

4.2 Assessment of the Assembled Mitogenomes

Among many, MitoZ v.3.6. tool, which run with k-mer based approach, was considered as powerful tool for de novo assembly of novel species (Mahar et al, 2023). After testing algorithms such as MITObim, GetOrganelle, NOVOplasty, MitoZ was chosen as the best fitting for the data pool we have. It provided the mitogenomes with minimum mismatch ratio and gaps.

Whole genome sequencing approach is followed with mitochondrial DNA. The inclusion of all protein-coding genes (PCGs) became a common practice to increase the resolution of mitochondrial lineages. However, the inclusion of rRNAs and tRNAs is less common without a solid justification of exclusion. In some studies, these regions also shown to be effective tools in phylogenetic comparisons (Domes et al, 2008; Yu et al, 2019). Even the addition of these genes to analyses can enhance the confidence in nodal clustering and branching (Cameron, 2014).

The length of mitogenomes has shown variation intraspecifically. D-loop regions had gaps in the sequences. The further annotations and alignments indicated variations in

the D-loop was the main cause of variation in the length of mitogenomes and the incomplete circularization of some of the samples. Two perspectives were evaluated for evaluation of the cause. The first is that significant variability of nucleotide composition and length variations was detected in intra- and interspecific comparisons of mtDNA. Studies showed length and sequence variation in the D-loop of mammalians including bats (Wilkinson & Chapman, 1991). On the other hand, the repeat regions on the D-loop are prone to be inflated due to PCR in sequencing processes (Wu et al, 2015).

Each of the perspectives were evaluated with the analysis of D-loop regions which lay between $tRNA^{Pro}$ and $tRNA^{Phe}$. Two distinct tandem repeat regions were detected. The first region which was was 81 base-pair long and close to $tRNA^{Pro}$. This tandem region was studied on Vespertilionidae previously and it was shown that they are between 78 to 85 base-pair length and composed of two to nine repeats. As the control region of mtDNA, the D-loop is thought to act as a protein-binding site and ensure signal redundancy (Wilkinson, 1997). At least two repeats were detected in this region of our samples. However, the average coverage for this region in all samples were notably high from the rest of the sequence. This may indicate an underestimated number of repeats in the sequence or poor filtration of PCR duplicates in prior to assembly. By following the pre and post reads, the actual number of repeats were estimated for each sample and curated manually.

Second repeat region was microsatellite-like short tandem repeats. The major differentiation of the lengths in the samples was resulted from this region. De novo assembly underestimates the number of repeats. This is interpreted to be challenging situation if the insert size of the paired-end reads was smaller than the actual length of in this region. Therefore, an average number of repeats for this region was manually curated with some question marks.

D-loop sequences were hypervariant and manually curated. Microsatellites in mtDNA D-loop may fail to represent polymorphic patterns and accumulation of recurrent mutations in this heteroplasmic sequence may blur the distinct branches in the trees (Cameron, 2014; Dobrowolski et al, 1998; Zhan et al, 2012). Therefore, the D-loop regions were removed from the coding sites for phylogenetic analyses to maximize phylogeny resolution.

All PCGs, rRNAs, and tRNAs were presented in the sequences. The tRNAs were evaluated with their transcriptions to reveal codons for further studies. Codon usage bias analysis is a trending way of understanding selection pressure on mitochondrial genomes during evolutionary history (Wei et al, 2014). Since the typical form of secondary structure of tRNAs is cloverleaf, tRNASer1, the only tRNA without this form, was investigated. In metazoan mitochondrial genomes, the deletion of D-arm in tRNA is considered a common feature in previous studies and there is no significant functionality with C-arm being used as an amino acid acceptor (Yoon & Park, 2015).

4.3 Phylogenetic Relations

Previous studies investigating the phylogenetics of *Hypsugo savii* utilized relatively short sequences from different mitochondrial and nuclear markers. In this study, the phylogeny of the investigated lineages was analysed with complete mitochondrial genomes. Two commonly used Maximum Likelihood-based interference tools were used, aaa and bbb (Stamatakis & Kozlov, 2020).

The structures of the phylogenetic trees and the haplotype network were concordant. Related species, *H. alaschanicus* and *H. ariel*, were located at the basal nodes of the phylogeny, indicating their earlier split from the rest of the group. *Hypsugo savii* samples were grouped into four major clusters. These clusters exhibit spatial structuring (Figure 3.5).

The levels of genetic distances among the European clades of *H. savii* (8-10%) were within the range of their differentiation from *H. alaschanicus*. The newly proposed species, *H. stubbei* also exhibits similar levels of differentiation from these lineages (Dolch et al, 2021). These high genetic divergences among the *H. savii* indicate that this group might harbor further cryptic diversity.

Two samples originated from Kyrgyzstan grouped with the *H. stubbei* samples from Mongolia. In Dolch et al. (2021), this species was only identified in Mongolia. Based on their analysis with partial ND1 sequences, the samples of *H. savii* that they analyzed from Kyrgyzstan were diffentiated from *H. stubbei* samples. These findings suggest that these distinct lineages occur in Kyrgyzstan, potentially with a contact zone.

The Eastern and Western Mediterranean lineages exhibit further substructuring. The subgroups within the former exhibit around 2% genetic differentiation and are geographically separated; one is distributed in Europe and the other in Asia. These lineages might have a contact zone in Anatolia. Çoraman et al. (2013) have shown the distribution of this clade in Anatolia. There were notable variations in ND1 and CYTB genes within the Eastern Mediterranean clade compared to other clades. Likewise, in the Western Mediterranean lineage, the Moroccan subgroup distinctly diverges from the European subgroup with 1,2% genetic differentiation. García-Mudarra et al. (2009) claimed that there is no genetic patterns shaped directly by the Straits of Gibraltar. Hence, more samples from Morocco and Iberia can clarify this subdivision. Also, one sample from Sardinia was different from the rest of the lineage with average 1,3% variation.

The co-occurrence of deeply diverged lineages (SWM and WM) on Sardinia and Italy assured the studies sympatric existence of these clades on the Mediterranean island (Veith et al, 2011; Bogdanowicz et al, 2015; Dondini et al, 2016; Mifsud & Vella et al, 2019). Samples 10442 and 10431 from Sardinia were species-level differentiated from other Sardinian samples showing a totally different lineage. Our estimation is the turquoise group represents the North African/ Southwestern Mediterranean (*H.* cf *darwinii*) clade which was researched in the studies showing the cryptic diversity in *H. savii* (Pestano et al, 2003; Ibáñez et al, 2006; Mayer et al, 2007; García-Mudarra et al, 2009; Veith et al, 2011; Çoraman et al, 2013; Bogdanowicz et al, 2015; Borloti et al, 2020). Sample 10445 from Sardinia which was grouped with Western Mediterranean clade showed relatively small but notable differentiation within the clade it was included. Further sampling may also help to understand what is behind this variation.

The Moroccon samples were clustered in the Western Mediterranean clade showing another sympatry region for deeply diverged clades since the individual proposed as *H.* cf *darwinii* by Mayer et al. (2007) was from Morocco.

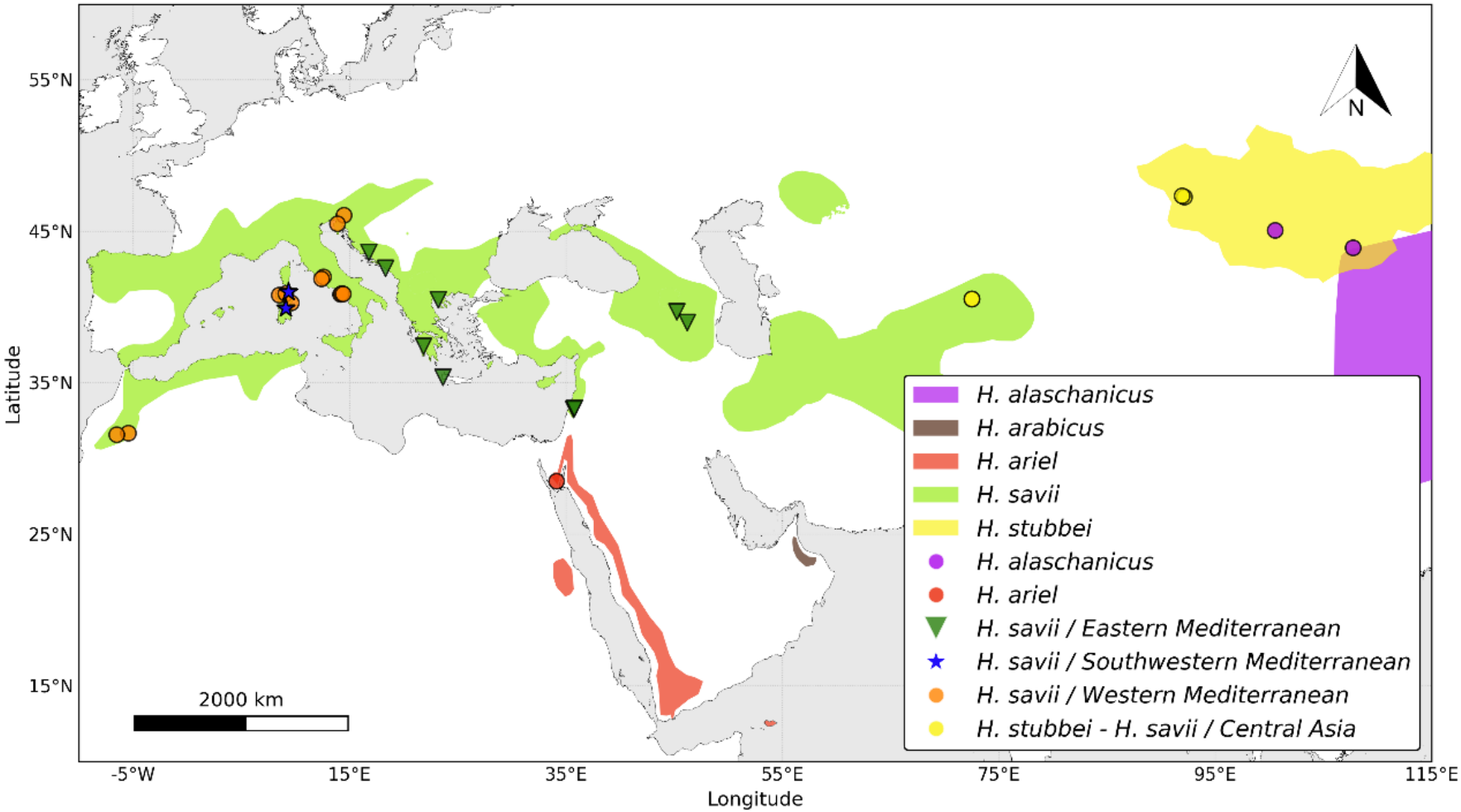

Figure 4.1: The distribution of the estimated geographical clades is shown in the Palearctic region including the deeply diverged lineages sympatrically occurring in Sardinia.

5. CONCLUSION

This study aimed to enhance our understanding of the evolutionary history of the *Hypsugo savii* species complex by de novo assembling whole mitogenomes and analyzing their phylogenetic relations. Constructing a robust phylogenetic tree based on mitogenomic data provided insight into the relationships among different lineages within the complex, as well as their closely related relatives. While the data collection spanned a broad geographical range, it remains limited in sample size. Furthermore, the study focused exclusively on mitochondrial genomes. Further investigations that incorporate nuclear genomes and explore putative contact zones among the identified lineages would yield additional insights about the cryptic diversity within this group. Elucidating biodiversity patterns and ensuring accurate taxonomic assignments are critical for effective conservation planning. In this context, this study provides foundational information about this species complex for subsequent research endeavors.

REFERENCES

Altringham, J. D. (2011). *Bats: from evolution to conservation*. Oxford University Press.

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*(5), 1530-1534.

Benda, P., Andreas, M., & Reiter, A. (2002). Record of Hypsugo arabicus from Baluchistan, Iran, with remarks on its ecology and systematic status. *Bat Research News*, 43, 75-76.

Benda, P. & Aulagnier, S. (2020). *Hypsugo ariel*. *The IUCN Red List of Threatened Species* 2020: e.T171619155A22071929.

Benda, P., Dietz, C., Andreas, M., Hotový, J., Lučan, R. K., Maltby, A., ... & Vallo, P. (2008). Bats (Mammalia: Chiroptera) of the Eastern Mediterranean and Middle East. Part 6. Bats of Sinai (Egypt) with some taxonomic, ecological and echolocation data on that fauna. *Acta Societatis Zoologicae Bohemicae*, *72*(1-2), 1-103.

Benda, P., Faizolâhi, K., Andreas, M., Obuch, J., Reiter, A., Ševčík, M., ... & Ashrafi, S. (2012). Bats (Mammalia: Chiroptera) of the Eastern Mediterranean and Middle East. Part 10. Bat fauna of Iran. *Acta Societatis Zoologicae Bohemicae*, 76(1/4), 163-582.

Bern Convention on the Conservation of European Wildlife and Natural Habitats. (1979). Appendix 2. Retrieved from [http://conventions.coe.int/Treaty/FR/Treaties/Html/104-2.htm]

Bernt, M., Bleidorn, C., Braband, A., Dambach, J., Donath, A., Fritzsch, G., ... & Struck, T. H. (2013). A comprehensive analysis of bilaterian mitochondrial genomes and phylogeny. Molecular Phylogenetics and Evolution, 69(2), 352-364.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K., Meier, R., Winker, K., ... & Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in ecology & evolution*, *22*(3), 148-155.

Bogdanowicz, W., Hulva, P., Černá Bolfíková, B., Buś, M. M., Rychlicka, E., Sztencel-Jabłonka, A., ... & Russo, D. (2015). Cryptic diversity of Italian bats and the role of the Apennine refugium in the phylogeography of the western Palaearctic. *Zoological Journal of the Linnean Society*, *174*(3), 635-648.

Boore, J. L. (1999). Animal mitochondrial genomes. Nucleic acids research, 27(8), 1767-1780.

Borloti, I., Dinis, H., & Vasconcelos, R. (2020). Bats out of Africa: Disentangling the systematic position and biogeography of bats in Cabo Verde. *Genes*, *11*(8), 877.

Cameron, S. L. (2014). Insect mitochondrial genomics: implications for evolution and phylogeny. Annual review of entomology, 59, 95-117.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884-i890.

Coraman, E., Furman, A., Karataş, A., & Bilgin, R. (2013). Phylogeographic analysis of Anatolian bats highlights the importance of the region for preserving the Chiropteran mitochondrial genetic diversity in the Western Palaearctic. *Conservation Genetics*, *14*, 1205-1216.

Corbet, G. B. (1978). The mammals of the Palaearctic region: a taxonomic review. *British Museum (Natural History)*, *341*.

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Molecular biology and evolution*, *35*(2), 518-522.

Davey, J. W., & Blaxter, M. L. (2010). RADSeq: next-generation population genetics. *Briefings in functional genomics*, *9*(5-6), 416-423.

De Queiroz, K. (2005). Ernst Mayr and the modern concept of species. Proceedings of the National Academy of Sciences, 102(suppl_1), 6600-6607.

Dietz, C. (2005). Illustrated identification key to the bats of Egypt. *Electronic publication, version*, *1*.

Dietz, C., & Kiefer, A. (2016). *Bats of Britain and Europe*. Bloomsbury publishing.

Dobrowolski, M. P., Tommerup, I. C., & O'Brien, P. A. (1998). Microsatellites in the mitochondrial genome of Phytophthora cinnamomi failed to provide highly polymorphic markers for population genetics. *FEMS microbiology letters*, *163*(2), 243-248.

Dolch, D., Stubbe, M., Batsaikhan, N., Stubbe, A., & Steinhauser, D. (2021). Hypsugo stubbei sp. nov., is a novel cryptic bat species of the genus Hypsugo (Vespertilionidae, Chiroptera, Mammalia) from Mongolia. Results of the Mongolian-German Biological Expeditions Since 1962, No. 347. *bioRxiv*, 2021-03.

Domes, K., Maraun, M., Scheu, S., & Cameron, S. L. (2008). The complete mitochondrial genome of the sexual oribatid mite Steganacarus magnus: genome rearrangements and loss of tRNAs. *BMC genomics*, *9*, 1-13.

Dondini, G., Vergari, S., Fichera, G., & Kiefer, A. (2016). First record of Hypsugo cf darwinii (Tomes, 1859) in Tuscany, Italy. *Barbastella*, *9*, 1-3.

Drummond A.J., Rambaut A. (2003) BEAST version 1.3 [computer program]. Available: http://evolve.zoo.ox.ac.uk/beast.

Duchêne, S., Archer, F. I., Vilstrup, J., Caballero, S., & Morin, P. A. (2011). Mitogenome phylogenetics: the impact of using single regions and partitioning schemes on topology, substitution rate and divergence time estimation. PloS one, 6(11), e27138.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5), 1792-1797.

Espíndola, A., Ruffley, M., Smith, M. L., Carstens, B. C., Tank, D. C., & Sullivan, J. (2016). Identifying cryptic diversity with predictive phylogeography. Proceedings of the Royal Society B: Biological Sciences, 283(1841), 20161529.

Findley, J. S. (1993). *Bats: a community perspective*. Cambridge University Press.

Fukui, D., Sano, A. & Kruskop, S.V. (2019). *Pipistrellus alaschanicus*. *The IUCN Red List of Threatened Species* 2019: e.T136560A21995360.

Garcia-Mudarra, J. L., Ibanez, C., & Juste, J. (2009). The Straits of Gibraltar: barrier or bridge to Ibero-Moroccan bat diversity?. *Biological Journal of the Linnean Society*, *96*(2), 434-450.

Geneious Prime (2023) https://www.geneious.com .

Grant, J. R., Enns, E., Marinier, E., Mandal, A., Herman, E. K., Chen, C. Y., ... & Stothard, P. (2023). Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic acids research*, *51*(W1), W484-W492.

Greiner S, Lehwark P. and Bock R. (2019) OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Research* 47: 59-64.

Hanak, V., & Horacek, J. G. (1985) The taxonomic implication of different morphological systems in bats.

Hao, Z., Liang, X., & Li, G. (2022). Quality control and preprocessing of sequencing reads. *Bio-101*, e4454.

Harrison, R. G. (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends in Ecology & Evolution*, *4*(1), 6-11.

Hasija, Y. (2023). All About Bioinformatics: From Beginner to Expert. Elsevier.

Hoofer, S. R. (2003). *Molecular phylogenetics of the chiropteran family Vespertilionidae*. Oklahoma State University.

Horáček, I., Hanák, V., & Gaisler, J. (2000). Bats of the Palearctic region: a taxonomic and biogeographic review. In *Proceedings of the VIIIth European Bat Research Symposium* (Vol. 1, pp. 11-157). Kraków: CIC ISEZ PAN.

Hutterer, R., Decher, J., Monadjem, A., & Astrin, J. (2019). A new genus and species of vesper bat from West Africa, with notes on Hypsugo, Neoromicia, and Pipistrellus (Chiroptera: Vespertilionidae). *Acta Chiropterologica*, *21*(1), 1-22.

Ibáñez, C., García-Mudarra, J. L., Ruedi, M., Stadelmann, B., & Juste, J. (2006). The Iberian contribution to cryptic diversity in European bats. *Acta Chiropterologica*, *8*(2), 277-297.

Illumina. (2015). An introduction to next-generation sequencing technology. Retrieved from [www.illumina.com/technology/next-generation-sequencing.html]

Ellerman J. R. & Morrison-Scott T. C S. (1966). Checklist of Palearctic and Indian mammals 1758–1946. Second edition. British Museum of Natural History, London, United Kingdom.

Trifinopoulos, J., Nguyen, L. T., von Haeseler, A., & Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic acids research*, *44*(W1), W232-W235.

Juste, J. & Paunović, M. (2016). *Hypsugo savii*. *The IUCN Red List of Threatened Species* 2016: e.T44856A22072380.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods*, *14(6)*, 587-589.

Kipson M., Gazaryan S., & Horáček, I. (2023) Savi's Pipistrelle *Hypsugo savii* (Bonaparte, 1837). In Russo, D. (Ed.), Handbook of the Mammals of Europe: Chiroptera. (1st ed., pp.95-112). Springer Nature.

Leigh, J. W., Bryant, D., & Nakagawa, S. (2015). POPART: full-feature software for haplotype network construction. *Methods in Ecology & Evolution*, *6*(9).

Letunic, I., & Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, gkae268.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Lillegraven, J. A., Thompson, S. D., McNab, B. K., & Patton, J. L. (1987). The origin of eutherian mammals. *Biological Journal of the Linnean Society*, *32*(3), 281-336.

Löbl, I., Klausnitzer, B., Hartmann, M., & Krell, F. T. (2023). The silent extinction of species and taxonomists—An appeal to science policymakers and legislators. Diversity, 15(10), 1053.

Lou, R. N., Jacobs, A., Wilder, A. P., & Therkildsen, N. O. (2021). A beginner's guide to low-coverage whole genome sequencing for population genomics. Molecular Ecology, 30(23), 5966-5993.

Lowe, T. M., & Chan, P. P. (2016). tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic acids research*, *44*(W1), W54-W57.

Mahar, N. S., Satyam, R., Sundar, D., & Gupta, I. (2023). A systematic comparison of human mitochondrial genome assembly tools. *BMC bioinformatics*, *24*(1), 341.

Mayer, F., & Helversen, O. V. (2001). Cryptic diversity in European bats. Proceedings of the Royal Society of London. Series B: Biological Sciences, 268(1478), 1825-1832.

Mayer, F., Dietz, C., & Kiefer, A. (2007). Molecular species identification boosts bat diversity. *Frontiers in zoology*, *4*, 1-5.

Meng, G., Li, Y., Yang, C., & Liu, S. (2019). MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic acids research*, *47*(11), e63-e63.

Menu, H. (1987). Morphotypes dentaires actuels et fossiles des chiroptères vespertilionines. 2ème partie: implications systématique et phylogéniques. *Palaeovertebrata*.

Mickleburgh, S. P., Hutson, A. M., & Racey, P. A. (2002). A review of the global conservation status of bats. *Oryx*, *36*(1), 18-34.

Mifsud, C. M., & Vella, A. (2019). Mitochondrial genetic diversity of bat species from the Maltese Islands and applications for their conservation. *Natural and Engineering Sciences*, *4*(3), 276-292.

Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010) "Creating the CIPRES Science Gateway for inference of large phylogenetic trees" in *Proceedings of the Gateway Computing Environments Workshop (GCE)*, pp 1 - 8.

Murphy, W. J., Eizirik, E., Johnson, W. E., Zhang, Y. P., Ryder, O. A., & O'Brien, S. J. (2001). Molecular phylogenetics and the origins of placental mammals. Nature, 409(6820), 614-618.

Ng, P. C., & Kirkness, E. F. (2010). Whole genome sequencing. Genetic variation: Methods and protocols, 215-226.

Pestano, J., Brown, R. P., Suárez, N. M., & Fajardo, S. (2003). Phylogeography of pipistrelle-like bats within the Canary Islands, based on mtDNA sequences. *Molecular Phylogenetics and Evolution*, *26*(1), 56-63.

Ruedi, M., & Arlettaz, R. (1991). Biochemical systematics of the Savi's bat (Hypsugo savii)(Chiroptera: Vespertilionidae). *Journal of Zoological Systematics and Evolutionary Research*, *29*(2), 115-122.

Sætre, G. P., & Ravinet, M. (2019). *Evolutionary genetics: Concepts, analysis, and practice*. Oxford University Press.

Saitou, N. and Nei, M. (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees, *Molecular Biology and Evolution*, Vol. 4, pp.406–425.

Simmons, N.B. and A.L. Cirranello. (2024). Bat Species of the World: A taxonomic and geographic database. Version 1.5. Accessed on 05/28/2024.

Srinivasulu, C. & Srinivasulu, B. (2019). *Hypsugo arabicus*. The IUCN Red List of Threatened Species 2019: e.T17328A22131108.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312-1313.

Stamatakis, A., Kozlov, A. M., & Kozlov, A. M. (2020). Efficient maximum likelihood tree building methods. *Phylogenetics in the Genomic Era*, 1-18.

Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution*, *10*(3), 512-526.

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, *38*(7), 3022-3027.

Teeling, E. C., Springer, M. S., Madsen, O., Bates, P., O'brien, S. J., & Murphy, W. J. (2005). A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science*, *307*(5709), 580-584.

The Galaxy Community. (2024) The Galaxy platform for accessible, reproducible, and collaborative data analyses. Nucleic Acids Research. https://doi.org/10.1093/nar/gkae410

Uhrin, M., Hüttmeir, U., Kipson, M., Estók, P., Sachanowicz, K., Bücs, S., ... & Benda, P. (2016). Status of S avi's pipistrelle H ypsugo savii (Chiroptera) and range expansion in C entral and south-eastern Europe: a review. *Mammal Review*, 46(1), 1-16.

Veith, M., Mucedda, M., Kiefer, A., & Pidinchedda, E. (2011). On the presence of pipistrelle bats (Pipistrellus and Hypsugo; Chiroptera: Vespertilionidae) in Sardinia. *Acta Chiropterologica*, *13*(1), 89-99.

Wei, L., He, J., Jia, X., Qi, Q., Liang, Z., Zheng, H., ... & Sun, J. (2014). Analysis of codon usage bias of mitochondrial genome in Bombyx mori and its relation to evolution. *BMC Evolutionary Biology*, *14*, 1-12.

Wilkinson, G. S., & Chapman, A. M. (1991). Length and sequence variation in evening bat D-loop mtDNA. *Genetics*, *128*(3), 607-617.

Wilkinson, G. S., Mayer, F., Kerth, G., & Petri, B. (1997). Evolution of repeated sequence arrays in the D-loop region of bat mitochondrial DNA. *Genetics*, *146*(3), 1035-1048.Wilson, D. E., & Reeder, D. M. (Eds.). (2005). *Mammal species of the world: a taxonomic and geographic reference* (Vol. 1). JHU press.

Wu, Z., & Sloan, D. B. (2019). Recombination and intraspecific polymorphism for the presence and absence of entire chromosomes in mitochondrial genomes. Heredity, 122(5), 647-659.

Wu, Z., Cuthbert, J. M., Taylor, D. R., & Sloan, D. B. (2015). The massive mitochondrial genome of the angiosperm Silene noctiflora is evolving by gain or loss of entire chromosomes. *Proceedings of the National Academy of Sciences*, *112*(33), 10185-10191.

Yoon, K. B., & Park, Y. C. (2015). Secondary structure and feature of mitochondrial tRNA genes of the Ussurian tube-nosed bat Murina ussuriensis (Chiroptera: Vespertilionidae). *Genomics Data*, *5*, 213-217.

Yu, G., Lam, T. T. Y., Zhu, H., & Guan, Y. (2018). Two methods for mapping and visualizing associated data on phylogeny using ggtree. *Molecular biology and evolution*, *35*(12), 3041-3043.

Yu, X., Tan, W., Zhang, H., Jiang, W., Gao, H., Wang, W., ... & Tian, X. (2019). Characterization of the complete mitochondrial genome of Harpalus

sinicus and its impliations for phylogenetic analyses. *Genes*, *10*(9), 724.

Zhang, T., Lu, H., Chen, C., Jiang, H., & Wu, S. (2012). Genetic Diversity of mtDNA D-loop and Maternal Origin of Three Chinese Native Horse Breeds. *Asian-Australasian journal of animal sciences*, *25*(7), 921–926.

CURRICULUM VITAE

Name Surname **:** Yeliz Ergöl

EDUCATION **:**

- **B.Sc.** **:** 2021, Istanbul Technical University, Faculty of Science and Letters, Department of Molecular Biology and Genetics.

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Ergöl Y.**, Mayer F. and Çoraman E. 2024: Phylogeography of the Savi's Pipistrelle (Vespertilionidae, Chiroptera) Complex Based on Whole Mitochondrial Genome Analysis. International Graduate Research Symposium, May 8-10 Istanbul, Turkey.