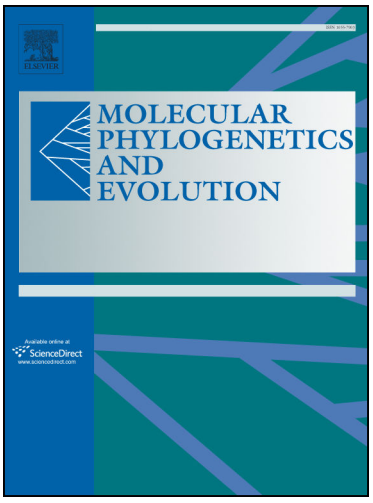
Journal Pre-proofs



Museomics help resolving the phylogeny of snowfinches (Aves, Passeridae, *Montifringilla* and allies)

Safiqul Islam, Claire Peart, Christian Kehlmaier, Yue-Hua Sun, Fumin Lei, Andreas Dahl, Sylvia Klemroth, Dimitra Alexopoulou, Maria del Mar Delgado, Paola Laiolo, Juan Carlos Illera, Sebastian Dirren, Sabine Hille, Davaa Lkhagvasuren, Till Töpfer, Martin Kaiser, Axel Gebauer, Jochen Martens, Claudia Paetzold, Martin Päckert

PII: S1055-7903(24)00127-1
DOI: https://doi.org/10.1016/j.ympev.2024.108135
Reference: YMPEV 108135

To appear in: *Molecular Phylogenetics and Evolution*

Received Date: 1 June 2023
Revised Date: 25 March 2024
Accepted Date: 16 June 2024

Please cite this article as: Islam, S., Peart, C., Kehlmaier, C., Sun, Y-H., Lei, F., Dahl, A., Klemroth, S., Alexopoulou, D., del Mar Delgado, M., Laiolo, P., Carlos Illera, J., Dirren, S., Hille, S., Lkhagvasuren, D., Töpfer, T., Kaiser, M., Gebauer, A., Martens, J., Paetzold, C., Päckert, M., Museomics help resolving the phylogeny of snowfinches (Aves, Passeridae, *Montifringilla* and allies), *Molecular Phylogenetics and Evolution* (2024), doi: https://doi.org/10.1016/j.ympev.2024.108135

Museomics help resolving the phylogeny of snowfinches (Aves, Passeridae, *Montifringilla* and allies)

Safiqul Islam[1,2,3], Claire Peart[4], Christian Kehlmaier[1], Yue-Hua Sun[5], Fumin Lei[6], Andreas Dahl[7], Sylvia Klemroth[7], Dimitra Alexopoulou[7], Maria del Mar Delgado[8], Paola Laiolo[8], Juan Carlos Illera[8], Sebastian Dirren[9], Sabine Hille[10], Davaa Lkhagvasuren[11], Till Töpfer[12], Martin Kaiser[13], Axel Gebauer[14], Jochen Martens[15], Claudia Paetzold[1], Martin Päckert[1]

1) Senckenberg Natural History Collections, Museum of Zoology, Königsbrücker Landstraße 159, 01109 Dresden, Germany

2) Max Planck-Genome-Centre Cologne, Max Planck Institute for Plant Breeding Research, Carl-von-Linne-Weg 10, 50829 Köln, Germany

3) Division of Systematic Zoology, Faculty of Biology, LMU Munich, Biocenter, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany

4) Division of Evolutionary Biology, Faculty of Biology, LMU Munich, Biocenter, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany

5) Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing, 100101, China

6) Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, 100101, China

7) Dresden-Concept Genome Center, c/o Center for Molecular and Cellular Bioengineering (CMCB), Technische Universität Dresden, Fetscherstraße 105, 1307 Dresden, Germany

8) Biodiversity Research Institute (IMIB, Universidad de Oviedo, CSIC, Principality of Asturias) – Campus de Mieres, Edificio de Investigación – 5ª planta, C. Gonzalo Gutiérrez Quirós s/n, 33600 Mieres, Spain

9) Vogelwarte Sempach, Seerose 1, 6204 Sempach, Switzerland

10) University of Natural Resources and Life Sciences, Vienna, Gregor Mendel-Strasse 33, 1180 Vienna, Austria

11) Department of Biology, School of Arts and Sciences, National University of Mongolia, P.O.Box 46A-546, Ulaanbaatar-210646, Mongolia

12) Leibniz Institute for the Analysis of Biodiversity Change, Zoologisches Forschungsmuseum Alexander Koenig, Adenauerallee, Bonn, Germany

13) Brascheweg 5, 10318 Berlin, Germany

14) Schäferei 1, 02943 Boxberg, Germany

15) Institute of Organismic and Molecular Evolution (iomE), Johannes Gutenberg University, 55099 Mainz, Germany

Corresponding author: Martin Päckert; e-mail: martin.paeckert@senckenberg.de

Results of the Himalaya expeditions of J. Martens 1969-2004, No. 297.

Results of the Mongolian-German Biological Expeditions since 1962, No 363

Abstract

Historical specimens from museum collections provide a valuable source of material also from remote areas or regions of conflict that are not easily accessible to scientists today. With this study, we are providing a taxon-complete phylogeny of snowfinches using historical DNA from whole skins of an endemic species from Afghanistan, the Afghan snowfinch, *Pyrgilauda theresae*. To resolve the strong conflict between previous phylogenetic hypotheses, we generated novel mitogenome sequences for selected taxa and genome-wide SNP data using from ddRAD sequencing for all extant snowfinch species endemic to the Qinghai-Tibet Plateau (QTP) and for an extended intraspecific sampling of the sole Central and Western Palearctic snowfinch species (*Montifringilla nivalis*).

Our phylogenetic reconstructions unanimously refuted the previously suggested paraphyly of genus *Pyrgilauda*. Misplacement of one species-level taxon (*Onychostruthus tazcanowskii*) in previous snowfinch phylogenies was undoubtedly inferred from chimeric mitogenomes that included heterospecific sequence information. Furthermore, comparison of novel and previously generated sequence data showed that the presumed sister-group relationship between *M. nivalis* and the QTP endemic *M. henrici* was suggested based on flawed taxonomy. Our phylogenetic reconstructions based on genome-wide SNP data and on mitogenomes were largely congruent and supported reciprocal monophyly of genera *Montifringilla* and *Pyrgilauda* with monotypic *Onychostruthus* being sister to the latter. The Afghan endemic *P. theresae* likely originated from a rather ancient Pliocene out-of-Tibet dispersal probably from a common ancestor with *P. ruficollis*. Our extended trans-Palearctic sampling for the white-winged snowfinch, *M. nivalis*, confirmed strong lineage divergence between an Asian and a European clade dated to 1.5 – 2.7 million years ago (mya). Genome-wide SNP data suggested subtle divergence among European samples from the Alps and from the Cantabrian mountains.

Keywords: historical DNA, chimeric mitogenomes, ddRAD sequencing, reference genomes, Qinghai-Tibet Plateau

1. **Introduction**

Natural history collections are biological archives of extant as well of past biodiversity (Meineke et al., 2019; Schindel and Cook, 2018; Hahn et al., 2020; Hilton et al., 2021) and provide basic material for the study of evolutionary changes over time of adaptive phenotypical traits (del Mar Delgado et al., 2019; Mason and Unitt, 2018; Miranda et al., 2021), range expansions in response to climate change (Mende and Hundsdörfer, 2013) or decline of genetic diversity in response to historical population bottlenecks (Glenn et al., 1999; Godoy et al., 2004; Kuhn et al., 2013; Gauthier et al., 2020). In the field of molecular systematics, genetic analysis of type specimens has become important to trace back their geographic origin (Cong et al., 2021) or for a correct assignment of taxon names to clades of a phylogeny (Tritsch et al., 2017; Kehlmaier et al., 2019; Scherz et al., 2020). Moreover, "historical DNA" (Raxworthy and Tilston Smith, 2021) extracted from museum specimens often provides the only chance to include rare species in phylogenies, e.g. in case of extinct species (Shapiro et al., 2002; Hekkala et al., 2011; Heupink et al., 2014; Podsiadlowski et al., 2017; Kehlmaier et al., 2021; Kearns et al., 2022) or of ultra-rare species represented by a single specimen in collections worldwide (Uva et al., 2018; Jønsson et al., 2018; Schweizer et al., 2020). Collections have therefore become very important for the optimization of phylogenetic data sets based on dense taxon sampling (Fiedler de Abreu Jr et al., 2020; Sampaio et al., 2023). Moreover, along with recent advances in genomic methods, museum collections provide a great potential for a wide variety of studies in the diverse field of so-called "museomics" (Zedane et al., 2016; Gauthier et al., 2020; Card et al., 2021; Raxworthy and Tilston Smith 2021; Ernst et al. 2022; Lalueza-Fox, 2022). Nevertheless, many phylogenetic studies relied on an incomplete taxon sampling due to a lack of material from inaccessible or remote areas or regions of armed conflict. This is quite problematic, because among the driving forces of biodiversity loss, political conflict is a neglected aspect, though it has manifold negative impacts on natural resources and ecosystems and therefore constitutes a major impediment for biodiversity research (McNeely, 2003; Hanson et al. 2009; Smallwood et al., 2011). As a consequence, the endemic faunas and floras of some conflict regions have remained severely understudied to date. Here, we focus on an example from Central Asia.

Afghanistan is situated at the crossroads of two global biodiversity hotspots: the Himalayas in the East and the Central Asian Mountains in the North (Marchese, 2015; Mittermeier et al., 2004; Hanson et al., 2009; map in Fig. 1). At a narrow spatial scale, Afghanistan harbors several unique ecoregions (some of them shared with neighboring Pakistan), such as the East Afghan Montane Coniferous Forests, the Sulaiman Range Alpine Meadows or the Baluchistan Xeric Woodlands (Wikramanayake et al., 2002). An ornithological survey from 2008 recorded a total of 122 bird species for Bamiyan county (Busuttil & Ayé 2009). Despite its comparatively low regional species richness, Afghanistan harbors refuge areas for major populations of threatened animals, such as the snow leopard, *Uncia uncia*, or the Marco Polo sheep, *Ovis ammon polii* (Saidajan, 2012). However, the local and regional fauna of Afghanistan remains poorly studied. For example, in a recent assessment of the global snow leopard population, occurrence data from Afghanistan was missing (see map in Riordan et al., 2016). Recent genetic studies on material from Afghanistan are scarce, and most of them focused either on domestic animals (Karimi et al., 2016) or on cultivated plant species (Terasawa et al., 2009; Seghal et al., 2015; Gori et al., 2019; Tehseen et al., 2021). Only a few phylogenetic and DNA barcoding studies included recent samples from wild Afghan populations collected in the 21st century (reptiles: Guo et al., 2011; Khan et al., 2021; Kazemi et al., 2021; butterflies: Efetov et al., 2014; Shapoval et al., 2017). In contrast, a recent phylogenetic study on the Afghan pika (*Ochontona rufescens*) relied on material exclusively collected in Iranian populations (Khalilipur et al., 2017). That knowledge gap is typical for the Afghan fauna because over the past decades traveling and field-work have become increasingly risky in this region of continuous conflict and is currently virtually impossible (Böhme and Jablonski, 2022; Maheshwari, 2022).

Among the seven endemic vertebrate species of Afghanistan (Kanderian et al., 2011), the Afghan snowfinch, *Pyrgilauda theresae* is the only bird species (Dementiev, 1963) even though there are

occasional winter records from Turkmenistan and doubtful occurrences in neighboring Tajikistan (Busuttil et al., 2010; Tolstoj and Geipel, 1990). This species was first described by Meinertzhagen (1937) from a specimen collected on the 19th of April 1937 at Shibar Pass (Kowtal-e Shibar) at about 2990 m a.s.l. (for reliability of type specimen data, see Rasmussen & Prŷs-Jones, 2003). Though it is currently considered a species of least concern by the IUCN, the global population size of *P. theresae* has not been estimated yet due to data deficiency (BirdLife International, 2021), and today the species is known from nine breeding season localities only, all of them restricted to Afghanistan (map in Vietinghoff-Scheel, 1980).

Traditionally, snowfinches were united under the single genus *Montifringilla* (e.g. Vaurie, 1956; Clements et al. 2022), but they are nowadays divided into three genera (*Montifringilla*, *Pyrgilauda,* and *Oynchostruthus*; e.g. Gill et al., 2022; Clements et al. 2023; Table 1). Six species are endemic to the Qinghai-Tibet Plateau (QTP) and its flanking mountain regions (Fig. 1A; Gebauer et al., 2006). Only one species, *M. nivalis*, occurs across a wide trans-Palearctic range from the northern QTP margin in the Chinese Altai, central Mongolia (subspecies *M. n. groumgrzimaili*) and west of the QTP region, inhabiting the Caucasus (subspecies *M. n. alpicola*) and European mountain systems (the nominate subspecies *M. n. nivalis*). Despite a large number of genetic studies, the phylogenetic relationships among snowfinch species are still under debate. At the family level, snowfinches represent a monophyletic clade in the Passeridae. Thus they are actually sparrows and not related to the finches, Fringillidae (Päckert et al., 2020b, 2021). So far, their affiliation to Passeridae was not acknowledged by a change of English species names, except a single attempt by Summers-Smith (2009) who listed the small *Pyrgilauda* species (including the larger *P. taczanowskii*; see Table 1) as "ground-sparrows" (but still listed all *Montifringilla* species as "snowfinches"). To date, intraspecific genetic diversification was described only across the trans-Palearctic range of the most widespread species, *M. nivalis*. However these findings were based on quite limited sampling (Päckert et al., 2021). Most recent molecular phylogenies of snowfinches were either based on mitochondrial markers alone (Lei et al., 2014; Cobos et al., 2021) or were subject to a strong bias between the signal of highly variable mitochondrial markers versus one or two less variable nuclear introns (Qu et al., 2006; Päckert et al., 2020b, 2021). Tree topologies inferred from these different data sets were conflicting for example concerning the position of the white-rumped snowfinch (*Onychostruthus taczanowskii*) and the relationships between species of the larger genus *Montifringilla* (Fig. 2A, B). Recent studies on large, genome-wide marker sets of snowfinches included only one species per genus (Qu et al., 2021; She et al., 2021). The Afghan snowfinch was missing from all these phylogenies. In such cases, historical material from bird collections becomes important. A historical collection of Afghan vertebrates is hosted at Zoological Research Museum Alexander Koenig (ZFMK), Bonn, Germany (Jablonski et al., 2019; Böhme and Jablonski, 2022). The ornithological collection of ZFMK hosts 276 bird specimens collected during two expeditions to Afghanistan from 01.09. to 13.09.1965 along an itinerary from Faizabad to Qala-e-Pandja (map in Niethammer, 1973) and seven years later to Darwaz (11.07.1972 - 31.07.1972) and the Pamir range (03.08. 1972 -05.09.1972).

With this paper, we provide a first taxon-complete phylogeny of snowfinches based on two data sets of one mitochondrial marker and genome-wide single nucleotide polymorphisms (SNPs) inferred from a next-generation sequencing (NGS) approach. We aim at

i) resolving taxonomic discrepancies among published phylogenies (Fig. 2A, B)

ii) providing a first phylogenetic hypothesis on the position of the Afghan snowfinch, *Pyrgilauda theresae*

iii) exploring the intraspecific diversification of the white-winged snowfinch, *Montifringilla nivalis*.

iv) 2. **Methods**

2.1. Sampling and DNA extraction

We extracted DNA from 40 samples of snowfinches from all eight species of the three genera (*Montifringilla*, *Onychostruthus,* and *Pyrgilauda*) and one further sample of the rock sparrow, *Petronia petronia* (Table S1). All samples were either frozen blood or tissue samples preserved in ethanol or preserving buffer except two toe pad samples taken from two historical specimens of the Afghan snowfinch, *P. theresae*. These were collected at Dasht-i-Nawar on 08.06.1965 at 3000 m a.s.l. (MAR1082; field number 495) and at Imai Pass on 29.04.1965 at 2800 m a.s.l. (MAR1083; field number 171). For DNA extraction from frozen blood or tissue samples we used an InnuPREP DNA Mini Kit and an innuPREP Blood DNA Mini Kit (both Analytik Jena AG, Germany) following the manufacturer's protocol except for overnight incubation with proteinase K for cell lysis.

To avoid cross-contamination historical toe pad samples were entirely processed in the clean room of the historical DNA facility of the Museum of Zoology, Senckenberg Dresden. DNA was extracted from toe-pad tissue using the DNeasy Blood & Tissue Kit (Qiagen) and two final elution steps with 2 x 50 µl elution buffer. Before and after each step of the procedure, clean benches were cleaned with DNA away (Molecular Bio Products, Inc.), and benches and the entire room were set under UV light for four hours.

2.2. Mitochondrial DNA – single marker cytochrome-b

A limited series of cytochrome-*b* sequences (cyt*b*) from a subset of our sampling was available from previous studies (Päckert et al., 2020, 2021). To complete the data set for this mitochondrial marker we amplified a 1079-bp-long cyt*b* fragment for 19 samples using the primer combination of O-L14851/ O-H16065 primers (Weir and Schluter, 2007). The PCR protocol was as follows: i) initial denaturation at 94°C for 10 min, ii) 35 cycles with denaturation at 92°C for 60 s, annealing for 60 s at 53°C and extension at 72°C for 120 s, iii), final extension at 72°C for 10 min. We purified PCR products using the ExoSAP-IT enzymatic cleanup (USB Europe GmbH; Staufen, Germany; 1:20 dilution, modified protocol: 30 min at 37°C, 15 min at 80°C). Purified PCR products were prepared for sequencing with BigDye™ 3.1 Dye Terminator Cycle Sequencing Kits (Thermo Fisher Scientific, Waltham, MA, USA) and cycle sequencing products were purified by using Sephadex (GE Healthcare, Munich, Germany), and sequenced in both reading directions on an ABI 3730 capillary sequencer (Thermo Fisher Scientific, Waltham, MA, USA). We inspected and edited all sequences with Chromas v.2.6.5 (Technelysium Pty Ltd, Brisbane, Australia), and cross-checked chromatograms of forward and reverse sequences of each sample. We used MEGA v. 10.1.8 (Kumar et al., 2018) for editing and alignment of consensus sequences. Newly generated cyt*b* sequences were deposited at GenBank under accession numbers OQ947839 - OQ947857 (compare Table S1).

2.3. Mitogenomes – library preparation, sequence assembly, and sequence annotation

Because *P. theresae* was the single snowfinch species missing out from mitogenome-based phylogenies so far (Cobos et al. 2021), we generated whole mitochondrial genomes for this species using specific protocols for museum material (Kehlmaier et al., 2019, 2021; Stelbrink et al., 2019). From the two historical toe pads (*P. theresae*) 28 µl of each lysate (11 ng DNA for sample MAR1082, and 305 ng DNA for sample MAR1083) was converted into single-indexed, double-stranded Illumina sequencing libraries (dsLibs) following Meyer and Kircher (2010) with modifications by Fortes and Paijmans (2015).

For library preparation DNA concentration was measured with a Qubit 3.0 Fluorometer (Invitrogen, Life Technologies GmbH, Darmstadt, Germany) using the highly sensitive quantitation assay. DNA molecule fragment length was checked on a Tape Station using the D1000 screen tape. Subsequently, DNA extracted from toe pad samples was sheared down to ca. 150bp using the Covaris ultrasonicator.

To increase the amount of endogenous mitochondrial DNA in the libraries, two rounds of in-solution hybridization capture (Horn, 2012; Maricic et al., 2010) were performed in a dedicated capture-only workspace in the main laboratory, using DNA baits generated from long-range PCR products of *Pyrgilauda ruficollis* (sample MAR423). Long-range PCR (LR-PCR) and primer sequences are detailed in the Supplementary Information (Table S2). The PCR protocol was as follows: Denaturation at 93°C for 3:00 min, 35 cycles of 93°C for 0:20, 62°C for 0:30 min, 68°C for 10:00 min, and final elongation at 68°C for 20:00 min. Sequencing was performed in-house on an Illumina MiSeq platform, generating 75 bp paired-end reads. To account for intraspecific variation in the mitogenome of *M. nivalis*, we applied the same long-range PCR setup and library preparation method as described above to frozen blood samples from Mongolian, Caucasian, and European samples of this species (plus one sample each of *O. taczanowskii* and *P. ruficollis* for comparison; see Table S1). LR-PCR products were sheared down to ca. 150bp and mixed at an equimolar ratio before being converted into dsLibs and sequenced.

Mitogenomes were adapter-trimmed with Skewer 0.2.2 (Jiang et al., 2014). BBmap-suite 37.24 (Bushnell et al., 2017) was used for read merging (minimum length 35 bp), quality filtering (minimum Q-score 20), and duplicate removal. The remaining quality-filtered reads (QFR) were screened for contamination using FastQScreen 0.11.4 (Wingett and Andrews, 2018) and a set of predefined mitogenomes (Table S1). Due to no obvious contamination problems, the entire read pool was used for mitogenome assembly with MITObim (Hahn et al., 2013), a two-step baiting and iterative mapping approach, with an allowed mismatch value of 2 and a starting seed KJ148629 (*P. blanfordi*). The resulting scaffolds were visualized and checked for assembly artifacts in Tablet (Milne et al., 2013). Artifacts were manually removed from the assembled contigs and all positions with coverage below 3-fold were masked as ambiguous (N) using the maskfasta subcommand of BEDTools 2.29.2 (Quinlan and Hall, 2010). The sequence length distribution of mapped reads was calculated with a customized awk command and in Microsoft Excel. Newly generated mitogenome sequences were deposited at the European Nucleotide Archive (ENA) under accession numbers OX637969-OX637975.

2.4. Validation and quality check of mitogenome sequences

We compared our seven newly generated mitogenomes to fifteen published mitogenomes of snowfinches available at GenBank (Table S3; e.g. from Ma et al. 2014, 2015). For each snowfinch species, published mitogenome sequences were identical for intraspecific pairwise comparisons, so we used a single published mitogenome sequence per snowfinch species. For hierarchical outgroup rooting we added whole mitogenome sequences of *Passer domesticus* (MN356394) and *Hypocryptadius cinnamomeus* (MN356286) to our dataset. We constructed two alignments for whole mitogenome sequences and for a set of 13 coding mitochondrial genes (n= 17) using MEGA and checked alignments for inconsistencies, such as regions of unusually high numbers of substitutions among conspecific sequences. For comparison of intra- and interspecific divergence levels, we calculated pairwise uncorrected *p*-distances with MEGA for each of the thirteen coding markers. To control for deviations among tree topologies inferred from different markers, we reconstructed mitochondrial phylogenies for all thirteen coding markers in separate runs with BEAST v.2.6.2 (Bouckaert et al., 2014; parameters and settings, see below). Following recommendations for validation of newly generated mitogenomes (and quality check of published mitogenomes) by Botero-Castro et al. (2016), we used larger published sequence data sets of mitochondrial markers (most from phylogeographic studies on snowfinch species by Qu & Lei, 2009 and Qu et al., 2006, 2005, 2009) for

comparison with homologous fragments of our mitogenome sequences: cytochrome-oxidase subunit 1 [COI]: n= 55, 483 bp; cytochrome-*b* [cyt*b*]: n= 109, 880 bp; d-loop, control region [CR]: n= 22, 288 bp (full documentation see supplementary Table S4). From each of the three data sets, we constructed minimum-spanning networks with PopArt 1.7 (Leigh and Bryant, 2015).

2.5. ddRAD library preparation and sequencing

For inference of a genome-wide SNP data set were used for double-digest restriction site associated DNA sequencing (ddRAD seq), which is a frequently applied approach for the study of diverse phylogenomic research questions in birds, such as diversification among island populations (Martin et al. 2021; Cooper and Uy (2017) or resolution of phylogenetic relationships in tits and larks (Stervander et al., 2015, 2016), seabirds (Obiol et al., 2021) and flycatchers (Gwee et al., 2019). DdRAD seq was performed at the Deep Sequencing Facility in the Center for Molecular and Cellular Bioengineering (CMCB) Dresden. We used Qubit (Thermo Fisher Scientific, Waltham, MA, USA), dsDNA High-Sensitivity (HS) and Broad-Range (BR) assays for DNA concentration measurement following the manufacturer's protocol. According to our Qubit measurements, we selected 38 samples with sufficient DNA concentrations for ddRAD seq (compare Table S1). For sample preparation, 50 ng gDNA were double-digested with SbfI and MspI (NEB) for 120 minutes at 37°C followed by heat inactivation at 65°C for 20 min. SbfI specific library barcodes carrying Truseq-i5 Illumina adapters with cohesive ends were ligated to the cohesive ends of the SbfI restriction sites of the digested DNA fragments. The same was done for the MspI site with a MspI-specific truncated universal TruSeq-i7 adapter. Ligation was performed for 120 min at 22°C followed by heat inactivation at 65°C for 20 min. Samples with different P5 Barcodes were pooled and purified using XP beads (Beckman Coulter, Krefeld, Germany) at a ratio of 1:1 to remove non-ligated adaptors. Sample-specific library indices were added during PCR (98°C 30 sec, 15 cycles [98°C 10 sec, 65°C 75 sec], 65°C 5 min) with 1x concentrated NEBNext Q5 Hot Start Hifi PCR Master Mix and standard TruSeq i5/i7 indexing PCR primers. After an XP bead purification (1:1) libraries were quantified using the Fragment Analyzer (Agilent, Waldbronn, Germany). Libraries were equimolarly pooled before sequencing them in single-end mode on an Illumina NextSeq 500 system to a read length of 75 bp and a depth of at least 1 million reads per sample. Individual fast-Q files of ddRAD raw sequencing data were deposited at NCBI SRA (for accession numbers, see Table S1). [*Sets of SNP alignments and further data files will be deposited at Dryad upon acceptance.*]

2.6. Data assembly and SNP calling

We used the program process_radtags as implemented in Stacks software (Catchen et al., 2013; 2021) with default settings applied for demultiplexing of Illumina raw reads and for removal of barcodes, adapters and low-quality reads. We checked average sequence quality per read and GC content using FastQC v0.11.9 (Andrews, 2015). For reference-based assembly of our trimmed reads (Eaton and Ree, 2013) we relied on a multi-reference approach (Bohling, 2020; Valiente-Mullor et al., 2021). First, we mapped our cleaned read data to two reference genomes of snowfinches from Qu et al. (2021): *Pyrgilauda ruficollis* (NCBI acc. no: GCF_017590135.1) and *Onychostruthus tazcanowskii* (NCBI acc. no: GCA_017590055.1) downloaded from NCBI. Genome data for these two snowfinch species was annotated to 15,000 protein-coding genes (Qu et al., 2021) in a first "computational phase" of a two-step process (Ekblom and Wolf, 2014). Sincer a chromosome-level assembly was not available for the closely related species, andreference genomes from rather distant relatives can generally be used for read mapping and SNP calling (Galla et al., 2018), we also aligned our read data to the house sparrow (*Passer domesticus*) genome which is assembled to chromosome-level (Elgvin et al., 2017; accession no: GCA_001700915.1). We used ipyrad v.0.9.42 (Eaton and Overcast, 2020) for data assembly and read mapping to the three different reference genomes. We applied a clustering threshold of 85% and

a minimum sequencing depth for clustering ≥6X. We applied default parameter settings of the reference-based ipyrad pipeline with a maximum of 8 indels, 0.5 heterozygous sites, 20% SNPs per locus, and a minimum of four samples per locus.

After these first filtering steps, the ipyrad pipeline (Eaton and Overcast, 2020) produced three independent VCF output files from mapping against three different reference genomes. For variant calling we used vcftools 1.1.5 (Danecek et al., 2011) and bcftools v1.8 (Danecek et al., 2021) with a quality value ≥30 applied to separate autosomal from Z-chromosomal data sets [which was possible only for the data set inferred from alignment to the house sparrow reference genome that was annotated to the chromosome level]. From each of the VCF files we generated final SNP data sets allowing for 0%, 10%, 20% and 30% missing data that were used as input data for phylogenetic analysis.

In a final step, we used vcftools for biallelic SNP filtering and thinning of vcf files to retain only one SNP per contig and thus to reduce the effects of physical linkage among markers (O'Leary et al., 2018). We applied the vcftools "thinning" option and tested a range of thinning factors (thin) of which thin= 200 yielded the highest proportion of unlinked SNPs in our thinned data set (i.e. 100%). Likewise, we generated SNP data sets from each of the three thinned VCF files allowing for 0%, 10%, 20% and 30% missing data.

2.7. *Phylogenetic inference and divergence time-estimates*

For Bayesian inference of the species phylogeny, we used BEAST v.2.6.2 (Bouckaert et al., 2014). A-priori we estimated best-fit substitution models with MrModeltest v.2.3 (Nylander, 2004) for all data sets processed in separate BEAST runs (for model settings see Table 2). For test runs with each of the thirteen mitochondrial coding markers, and for genome-wide SNP alignments with varying fractions of missing data, the parameters of BEAST were consistent: an MCMC chain length of 10,000,000 generations (with all model settings estimated and parameters logged and trees sampled every 1,000 generations) under the uncorrelated lognormal clock model, the "auto-optimize" option activated and a Yule prior.

Final optimized runs were performed separately for i) the set of thirteen mitochondrial coding markers (thirteen partitions, with fixed best-fit model settings for each partition; Table 2), ii) the cyt*b* data set for our full sampling and further sequences from GenBank (n= 51; supplementary Tables 1, 4), iii) the three SNP data sets (generated using different reference genomes) with the percentage of missing data that yielded the highest node support in the test runs. We performed three independent runs with BEAST for 30,000,000 generations for the mitogenome data set (with parameters logged and trees sampled every 3,000 generations) and for 50,000,000 generations for the SNP data sets (with parameters logged and trees sampled every 5,000 generations) under the uncorrelated lognormal clock model for all loci, the "auto-optimize" option activated and a Yule prior applied to the tree. We combined log files and tree files from independent BEAST runs using LOGCOMBINER v.1.8.1 (with a burnin of 30% applied to all tree files) and checked the combined log file in TRACER v. 1.4 (Rambaut & Drummond, 2007) to ensure adequate ESS values for all parameters (all ESS > 200). We constructed consensus trees using TreeAnnotator (burnin of 30% applied to the combined tree files). We also conducted a tree set analysis as implemented in DensiTree that allows comparing sets of all trees with sets of consensus trees for every possible topology (Bouckaert, 2010).

All obtained phylograms were visualized in FigTree vers. 1.4.2 (Rambaut, 2009). As a control, we repeated phylogenetic reconstructions for all five data sets using a maximum-likelihood approach as implemented in RaxML (Stamatakis, 2006, 2014). The mitogenome data set was partitioned by gene (13 coding markers with the GTR + I + Γ model consistently applied to all partitions), whereas the SNP

data sets were left unpartitioned (GTR + I + Γ model applied). For all reconstructions with RaxML, node support was inferred from 1000 thorough bootstrap replicates.

To distinguish between a lack of support due to non-informativeness of the data set and conflicting information between information from different loci we relied on the Quartet Sampling method by Pease et al. (2018), that has previously been applied to that end in phylogenetic studies (Paetzold et al., 2019; Bybee et al., 2021; Kong et al. 2022). Quartet Sampling [QS] quantifies and qualifies discord in a given input tree by calculating quartet topologies for internal branches by sampling subtending terminal taxa to create replicates. The results are given as four scores: Quartet concordance (QC: -1 <= QC <= 1) indicating the amount of discordance in the dataset for a given branch; Quartet Differential (QD: 0 <= QD <= 1) indicates the skew in inferred discordant topologies for that branch (if present); Quartet Informativeness (QI: 0 <= QD <= 1) indicates informativeness of the data for a given branch; and QF (Quartet Fidelity: QF; 0 <= QF <= 1) indicates the number of concordant replicates for a given taxon (Pease et al., 2016). Considered together the four values can indicate whether the given dataset has variation of sufficient quality and quantity to inform a given branch (QI), whether this information contradicts the one represented by the tree (QC), and if so, whether this discord is random or skewed (QD). The QD value represents any discordant quartet topology evaluated and the distribution between the two possible discordant relationships across replicates. If the ratio is even (QD=1), it indicates incomplete lineage sorting as a likely source for the discord, if it is skewed (QD << 1) this indicates horizontal gene transfer as a source of discord. Finally, the proportion of discordant replicates a specific sample is producing (QF), indicates a possible rogue behavior or low informativeness for the given sample (Pease et al., 2016). Inclusion of “rogue taxa” in a phylogeny might obscure the true evolutionary relationships among clades, often due to long-branch attraction (Westover et al., 2013). To avoid such an effect and to test robustness of resolved relationships, QS analysis can be repeated under exclusion of taxa showing low QF scores.

For QS analysis, we selected the BEAST tree inferred from read mapping to the annotated *P. domesticus* reference genome with 20% missing (autosomal) data and the corresponding full-length alignment (including invariable sites; alignment length: 517,763 sites). We performed QS with 100 replicates and the default parameters.

We applied two independent approaches to estimate divergence times among extant snowfinch species. First, we calibrated the mitogenome phylogeny by assigning empirical substitution rates by Lerner et al. (2011) to 13 coding markers (with a normal prior distribution, means according to marker-specific estimates by Lerner et al. [2011; see Table 2] and SD set to 10% of the mean estimate). The same was done for the cyt*b* data set applying the empirical substitution rates by the latter study and by Weir & Schluter (2007) in separate runs. For a secondary calibration of the genome-wide SNP data set, we assigned a fixed node age to the most recent common ancestor (tmcra) of all snowfinches derived from a fossil calibration of the Passeroidea tree by Päckert et al. (2020b, 2021). Accordingly, we set the mean tmrca prior for the node uniting all snowfinches to 8.23 Ma and adjusted a normal prior distribution (with sigma= 1.0) according to the 95% highest posterior density interval (HPDI = [6.45 – 10.24] Ma; Päckert et al. 2020b).

For the reconstruction of unrooted neighbor net networks and TCS networks, we used SplitsTree 4.17.1 (Huson & Bryant, 2006). We examined divergence among species-level taxa and population structure in *M. nivalis* using principal component analysis (PCA) as implemented in the ipyrad.pca tool from the ipyrad pipeline. We performed two separate PCAs for *Montifringilla* one the one hand and for *Pyrgilauda* and *Onychostruthus* on the other hand. For calculation of pairwise F_{ST} values among snowfinch species and among populations of *M. nivalis* we used Stacks 2.60 software (Rochette and Catchen, 2017). Both, PCA and FSt calculations relied on the autosomal SNP data set inferred from read alignment to the annotated *Passer domesticus* genome.

Results

3.1. Mitochondrial DNA

We found mismatches in the position of mitogenome sequences in tree topologies for nine out of thirteen mitochondrial coding markers suggesting the inclusion of heterospecific DNA fragments in the published mitogenomes of *O. tazcanowskii* and *P. davidiana*. In tree topologies inferred from the four remaining coding markers (COI, COII, COIII, ND5) newly generated sequences of the latter two species were resolved within clades of conspecific mitogenome sequences from GenBank at low pairwise distances (< 1%; identical sequences for COII and COIII; Table S3, Figs 2C-F, S1). The topology resulting from COII showed only poor resolution of the deeper splits with the three outgroups nested in *Pyrgilauda*; (Fig. 2D). In the minimum spanning network inferred from a larger data set of the DNA barcoding marker all COI fragments from mitogenomes were nested in the correct cluster (Fig. 2G). In contrast, we found a striking mismatch in the position of the *P. davidiana* mitogenome sequence in six out of thirteen mtDNA markers (Fig. S2): In tree topologies inferred from ATP8, ATP6, ND3, ND4L, ND4, and cyt*b* the fragments representing the *P. davidiana* mitogenome clustered with the two mitogenomes of *O. tazcanowskii* at unexpectedly short branch lengths. For these six markers, the two *O. taczanowskii* mitogenomes and the *P. davidiana* mitogenome were extremely similar (with pairwise comparisons ranging at *p*-distances below 0.01, identical sequences for ATP8, ND3 and ND4L; Table S3, Fig. S1), whereas *p*-distances between *P. davidiana* and *P. blanfordi* were about ten times higher than expected (Fig. S1). That mismatch was even found within a single marker: In the haploytpe network inferred from the first fragment of cyt*b* (403 bp), all sequences inferred from mitogenomes were nested in the correct cluster, whereas in the haplotype network based on the second half of the same marker (477 bp), the mitogenome fragment of *P. davidiana* was identical to the central haplotype of the *O. tazcanowskii* cluster (Fig. S3). The remaining three markers and the non-coding control region showed strong mismatches in the position of two species (Fig. S4): For ND1 and ND2, *O. tazcanowskii* clustered with the sister-species pair *P. davidiana* and *P. blanfordi* (identical ND2 sequences for *O. tazcanowskii* and *P. davidiana;* Table S3, Figs S1, S5). Finally, tree topologies for ND6 and the flanking non-coding control region showed mismatches for mitogenomes of both *O. tazcanowskii* and *P. davidiana* (also confirmed by the minimum-spanning network based on a larger data set for the control region [n= 24]; Fig. S4). Unexpectedly, for the latter three coding markers (ND1, ND2, ND6), *p*-distances between the two *O. taczanowskii* mitogenomes were about twenty times higher (mean p-dist.= 10%) than for pairwise comparisons with the remaining ten markers (mean p-dist= 0.5%; Table S1, Fig. S1).

According to this comparison of topologies and pairwise p-distance values across 13 mitochondrial markers we needed to take into account that GenBank mitogenomes of both *P. davidiana* and *O. taczanowskii* included heterospecific DNA sequences. We therefore repeated phylogenetic reconstructions based on the mitochondrial data set after replacing all unreliable fragments of coding markers in the consensus mitogenome sequence of the two latter species by anonymous data („N“; compare Table S1, Fig. S1). The only major difference among consensus trees inferred from the two original mitogenome data sets referred to the position of *Onychostruthus* that was nested in *Pyrgilauda* in the uncorrected consensus tree on the one hand (Fig. 3A), whereas *Montifringilla* and *Pyrgilauda* were reciprocally monophyletic and *Onychostruthus* was sister to the latter in the corrected consensus tree (Fig. 3B). The latter topology was fully congruent with that of the cyt*b* tree inferred from the full set of samples (Fig. S5). Within *Pyrgilauda*, the Afghan snowfinch, *P. theresae*, was resolved as sister to *P. ruficollis* with full support (Fig. 3). In *Montifringilla*, mitogenomes for *M. nivalis* and *M. henrici* downloaded from GenBank formed a monophyletic clade with extremely short within-clade branch lengths (Fig. 3) and *p*-distances ranged between 0.05 and 0.2% (Table S1). In contrast, our specimens of *M. nivalis* from Asia and Europe were clearly separated from their sister clade *M. adamsi,* and from a third monophyletic clade including mitogenome sequences from GenBank for *M. henrici* and *M. nivalis* (Fig. 3).

3.2. Genome-wide SNPs

3.2.1. Effects of reference genome choice and missing data

For most samples (n= 35) more than 1 million reads were recovered from the Illumina run. A lower number of reads was recovered only for three frozen tissue samples and one historical toe pad sample of *P. theresae* (compare Table S1). Mapping rates ranged between 70 – 80%, and were slightly higher for the two ingroup reference genomes with *Pyrgilauda* species and *O. tazcanowskii* scoring best with the respective congeneric reference genome (Table S1; Fig. S6). As the only exception, all *Montifringilla henrici* had lower mapping rates for the two ingroup reference genomes as compared to that of the outgroup (Fig. S6). As could be expected, the historical *P. theresae* sample showed the lowest mapping rate (Fig. S6). After trimming and first filtering steps the sequence matrices inferred from each of the three reference genomes had sizes of 1,664,040 – 2,242,210 sites (with 52.12 – 53.25% missing sites) and contained 66,729 – 107,218 SNPs (38.41 – 39.28% missing sites in the SNP matrices; Table S5).

Phylogenies inferred from test runs (MCMC chain length 10 million generations) with unthinned SNP data sets from different reference-mapping yielded congruent tree topologies (Fig. S7). For all three approaches, the highest node support from posterior probabilities was received from data sets with 20% missing data allowed (Fig. S7). For higher proportions of missing data (30%) posterior probabilities decreased from strong to poor support at eight nodes for the *O. tazcanowskii* reference genome, at three nodes for the *P. domesticus* reference genome, and at a single node for the *P. ruficollis* reference genome (Fig. S7). Thinning of data yielded a strong increase of posterior probabilities at the node uniting all snowfinches (full support in all reconstructions), and at several nodes of the *Montifringilla* clade for the phylogeny inferred from read mapping to the *O. tazcanowskii* reference genome (Fig. S7). 16 out of 18 tree topologies [thinned and unthinned SNP data] supported i) a sister-group relationship of *Onychostruthus* and *Pyrgilauda* (Fig. S7), and ii) a sister-group relationship between *M. adamsi* and *M. nivalis* (Fig. S7). Three of the four deviating topologies that suggested a sister-group relationship between *M. adamsi* and *M. henrici* (thinned data set, 10% and 20% missing data; posterior probabilities 0.97 and 0.79; not shown), were inferred from mapping to the outgroup reference genome (*P. domesticus*; Fig. S7).

For final BEAST runs we selected SNP data sets inferred from mapping to all three reference genomes i) allowing 20% missing data for unthinned SNP data (matrix sizes ranged from 30,550 – 40,695 SNPs), ii) allowing 30% missing data for the thinned data set (matrix sizes 6,263 – 8,949 biallelic unlinked SNPs; Table S5).

3.2.2. Inference of phylogeny and divergence times

Tree topologies were largely congruent i) between BEAST trees (Fig. 4) and RaxML trees (Fig. S8) based on genome-wide SNPs inferred from read mapping to each of the three reference genomes, ii) between all trees based on genome-wide SNPs (Figs 4, S8), the corrected mitogenome tree and the cyt*b* tree (Figs 3, S5). The three snowfinch genera were reciprocally monophyletic in all phylogenetic reconstructions, with *Onychostruthus* and *Pyrgilauda* being sister taxa and separated from *Montifringilla* by a deep split dated to 7.8 million years ago (mya) in the genome-wide SNP trees and to 5.0 mya in the mitogenome tree (Figs 3, 4). Tree set analysis of the autosomal data set with DensiTree (Fig. 5A) showed little uncertainty of topology and good concordance of consensus trees with the BEAST trees (Fig. 4). In contrast, the same analysis based on the Z-chromosomal SNP data set showed a much greater uncertainty of topology (decreasing intensity towards the root of both, sets of all trees and of consensus trees; Fig. 5B), which is reflected by the greater reticulation of the Z-chromosomal neighbor net network compared to that inferred from autosomal SNPs (Fig. 5C, D).

The sole incongruence between phylogenies inferred from genome-wide SNPs and from mitochondrial markers concerned the position of the Afghan endemic *P. theresae*. It was resolved as descending from the most ancient offshoot of the *Pyrgilauda* clade in the SNP-based phylogenies (Figs 4, 5A), instead of being sister to *P. ruficollis* in mitogenome- and cyt*b*-based phylogenies (Figs 3, S5). In PCA (inferred from the SNP data set with 20% missing data allowed), *O. tazcanowski* was strongly separated from all *Pyrgilauda* species along the x-axis (PC1 explaining 49.3% of the total variation; Fig. 6B). Members of *Pyrgilauda* were mainly separated along PC2 (explaining 20.3% of the total variation) with clusters of *P. blanfordi* and *P. davidiana* showing some overlap (Fig. 6B). The latter two sister taxa (Figs 4, 5), appeared as separate clusters only in the scatterplots of PC1 *vs* PC3 (0% missing data) and PC1 *vs* PC4 (20% missing data; Fig. S9). Inclusion of the historical sample *P. theresae* strongly reduced the number of SNPs (from 8,863 filtered SNPs and 3,233 unlinked SNPs to only 341 filtered and 175 unlinked SNPs with and without *P. theresae*), however, it had no effect on the divisive clustering pattern among *Pyrgilauda* and *Onychostruthus* species (Fig. S9).

For within-clade relationships of *Montifringilla*, all phylogenetic reconstructions suggested a sister-group relationship of *M. nivalis* and *M. adamsi* with *M. henrici* being sister to the latter two (Figs 4, 5). In PCA, *Montifringilla nivalis* was strongly separated from the other two congenerics along the x-axis (PC1 explaining 37.5% of the total variation; Fig. 6A) and showed remarkable intraspecific divergence between nominate *M. n. nivalis* from Europe and the two Asian subspecies *M. n. alpicola* and *M. n. groumgrzimaili* (Fig. 6A; PC2 explaining another 20.2% of the total variation). The split between the Asian and European lineages of *M. nivalis* was dated to 2.3 – 2.7 mya (mitogenomes: 1.53 [1.40 – 1.69] mya). F_{ST} values between the two major lineages (0.39 – 0.48; Table 3) ranged at a similar magnitude like F_{ST} values among *Montifringilla* species (*M. nivalis* vs. *M. adamsi*: 0.34 – 0.41; *M. nivalis* vs. *M. henrici*: 0.39 – 0.47; *M. adamsi* vs. *M. henrici*: 0.34; Table 3). The most recent divergences among terminal sister taxa (Figs 4, 5) were dated to 1.7 – 1.9 mya for splits among Caucasian *M. n. alpicola* and Mongolian *M. n. groumgrzimaili* (mitogenomes: 0.10 [0.07-0.13] mya) and at 2.3 – 2.4 mya between *P. davidiana* and *P. blanfordi* (mitogenomes: 0.40 [0.32-0.48] mya; smallest F_{ST} value for comparison among species: 0.18; Table 3).

Within European *M. n. nivalis*, samples from the Alps and from the Cantabrian Mountains were reciprocally monophyletic only in Bayesian trees based on SNP data sets inferred from mapping on either of the ingroup reference genomes (Fig. 5). In BEAST trees based on SNPs inferred from mapping on the house sparrow reference genome (and in all three RaxML trees), the Cantabrian snowfinch populations appeared as a monophyletic group that was nested in all other *M. n. nivalis* as a fully supported terminal clade. Accordingly, genetic divergence between the two European populations was limited (F_{ST} value: 0.087; Table 3). In PCA, the two European populations appeared as separate clusters only in the scatterplot of PC1 *vs* PC4 (Fig. S10).

3.2.3. Quartet sampling

Quartet sampling (QS) scores for the total sampling indicated full support for all terminal clades (species-level or subspecific taxa) and most of the deeper nodes of the snowfinch phylogeny (Fig. 7). Lower QC scores indicated some discord with a skewed distribution of discordant signals in the sister-group relationship of *M. adamsi* and *M. nivalis* and strong discordance again with a skewed ratio in the sister-group relationship of *Onychostruthus* and *Pyrgilauda* (Fig. 7). Quartet informativeness was lowest at nodes i) uniting all *Pyrgilauda* species and ii) uniting the *Pyrgilauda* crown group except *P. theresae* (with 65% and 48% of all quartets having passed the likelihood cut-off; Fig. 7). To test for an effect of single branches on nodal support we repeated the QS analysis under exclusion of two individuals that had the lowest quartet fidelity (QF) scores (*P. theresae*: QF = 0.56; one *M. n. nivalis* from the Swiss Alps, L40989 [FKN1]: QF = 0.45; Fig. 7). The second analysis yielded full support for all nodes except for the sister-group relationship between *Onychostruthus* and *Pyrgilauda* (*P. theresae*

excluded), which showed a poor QC and QD = 0 despite a moderate QI score (signifying weak support with significant skew towards one alternative relationship; Fig. 7). The Cantabrian *M. nivalis* clade received strong support from QS scores, too (Fig. 7).

4. Discussion

4.1. Quality assessment of mitogenome sequences

Next-generation sequencing (NGS) techniques have recently revolutionized molecular systematics and have been applied to generate high-quality mitogenomes of non-model organisms in the plant and animal kingdoms (Briscoe et al., 2016; Kehlmaier et al., 2019, 2021; Maddock et al., 2016; Yuan et al., 2016). Apart from the finding that a notable percentage of mitogenomes deposited at GenBank was backed by deficient meta-data (Strohm et al., 2016), NGS methods have revealed a variety of errors in published mitogenome sequences such as incorrect gene annotation, missing reads or duplications, inclusion of numts or concatenation of heterospecific DNA fragments (Päckert, 2021; Prada and Boore, 2019; Sangster and Luksenburg, 2020, 2021; Skujina et al., 2017). In a re-evaluation of 1,876 avian mitogenomes, Sangster and Luksenburg (2021) demonstrated that every second phylogenetic study that incorporated mitogenome data from GenBank included at least one problematic sequence. Therefore, a thorough quality assessment of sequence data is recommended prior to phylogenetic analyses. For validation of newly generated mitogenome sequences, Botero-Castro et al. (2016) provided a set of quality control guidelines, such as using larger sets of mtDNA markers (e.g. the DNA-barcoding marker COI) for comparison.

Following those recommendations, we could identify two typical sources of error in published mitogenome sequences of snowfinches. First, mitogenomes of *O. tazcanowskii* and *P. davidiana* available at GenBank contained heterospecific DNA fragments of another snowfinch species (for nine out of thirteen coding markers). First, paraphyly of *Pyrgilauda* in our uncorrected mitogenome tree is very likely due to the effect of these two chimeric mitogenomes (compare Sangster and Luksenburg, 2020). For birds, Sangster and Luksenburg (2021) identified 23 published chimeras across the entire class of Aves (1.5% of their mitogenome data set) and the same team reported an even greater proportion of chimeras in published mitogenomes of fishes (5.7%, Sangster and Luksenburg, 2020). Inclusion of chimeric mitochondrial DNA sequences in genetic data sets has previously blurred phylogenetic relationships of extinct taxa, such as the Mascarene parrot, *Mascarinus mascarin* (Podsiadlowski et al. 2017). Second, the single published mitogenome assigned to the taxon "*Montifringilla nivalis*" turned out to be nearly identical to that of *M. henrici* and both were deeply diverged from our newly generated mitogenomes for Asian *M. n. groumgrzimaili* and European *M. n. nivalis*. This mismatch is certainly due to flawed taxonomy, because the Tibetan snowfinch was previously included in *M. nivalis* as subspecies *M. n. henrici* (Table 1; compare Cheng, 1987, map on p. 943; Glutz v. Blotzheim and Bauer, 1997). Incorrect assignment of GenBank sequences to species-level taxa due to dated nomenclature or to simple misidentifications was reported as another conspicuous pitfall in phylogenetic reconstructions (Hofstetter et al., 2019; Päckert, 2021; Salvi et al., 2021; Tritsch et al., 2017). Therefore, the use of published mitogenome sequences for meta-analyses (such as trait evolution or niche evolution) requires particular care. For example, the phylogenetic backbone used for the reconstruction of niche evolution in snowfinches by Cobos et al. (2021; Fig. 1B) suffered from both sources of error, causing misplacement of *O. taczanowskii* (chimeric mitogenome) and of *M. henrici* (flawed taxonomy; *M. nivalis* is actually missing from their phylogeny; compare Fig. 3).

4.2. Effects of reference genomes and filtering on SNP calling and phylogeny

Generally, independent SNP analyses yielded highly consistent results regardless of the percentage of missing sites or of reference genome choice. For snowfinches, mapping rates (and thus sequence and SNP matrix sized) were slightly higher for the ingroup reference genomes (*P. ruficollis*, *O. taczanowskii*) than for the outgroup reference genome (*P. domesticus*), a general trend that was found in other bird groups (Charadriiformes; Galla et al. 2019) or in ray-finned fishes (Bohling 2020). However, single outlier taxa might not follow that general pattern, like *M. henrici* in our taxon sampling, that consistently showed lower rates for ingroup-mapping across all samples. Previous studies confirmed possible effects of reference genome choice on divergence patterns reflected by clustering methods or on tree topologies especially in cases of high incomplete lineage sorting (Valiente-Mullor et al., 2012; Rick et al. 2023). However, such effects might have played a minor role for our snowfinch data set, because the outgroup *Passer domesticus* still is a rather close relative to the snowfinches, considering that previous studies had relied on the annotated zebra finch (*Taeniopygia guttata*; Estrildidae) genome for phylogenetic reconstructions of different passerine families, such as larks (Alaudidae; Stervander et al., 2016) or tits (Paridae; Stervander et al., 2015). Nevertheless, an ingroup reference genome might be advantageous for uncovering shallow divergences reflecting intraspecific diversification, as reciprocal monophyly of the two European mountain clades of *M. n. nivalis* could be resolved only by SNP data sets inferred from ingroup-mapping. For population genomic studies, even a local or regional reference individual can be used to decrease reference bias, as was exemplified for European sticklebacks (*Gasterosteus aculeatus*; Thorburn et al., 2023).

Apart from reference genome choice, filtering thresholds, like proportions of missing data and minor allele count were shown to have major effects on phylogenetic analyses, such as possible removal of true rare or private alleles that are relevant for detection of intraspecific phylogeographic patterns or past demographic events (O' Leary et al., 2018; Rick et al., 2023). The latter effect might be of minor importance for time-calibrated trees covering intergeneric separation across evolutionary time spans of more than 15 myr. However, while we could confirm a postulated effect of missing data proportions on tree topologies or node support (see Eaton et al., 2017), this was not the case for clustering patterns inferred from PCA. Previous studies had inferred fully resolved phylogenies for evolutionary radiations with even greater amounts of missing data (60% missing data per SNP: Crotti et al., 2019; up to 90% missing data: Tripp et al., 2017). For our snowfinch data set, data thinning (extracting only biallelic unlinked SNPs) improved node support generally for the most ancient node (uniting the three snowfinch genera), and specifically also for within-clade relationships of *Montifringilla* in the outgroup-reference data set.

4.3. Quartet discordance and mito-nuclear discordance

Generally, QF values for tip clades of the snowfinch phylogeny were similar to those in previous phylogenies that included higher numbers of taxa (e.g. Kong et al. 2020). Neither of the QS analyses supported the phylogenetic hypothesis previously inferred from mitogenome sequences (Cobos et al., 2021) that had placed *Onychostruthus* firmly nested in a paraphyletic *Pyrgilauda* (Fig. 2B). QS scores suggested an effect of poorly informed branches (Pease et al., 2018) only for the *Pyrgilauda* clade, QC and QI scores for this clade improved when the analysis was repeated after removal of two terminal taxa with the lowest QF scores (*P. theresae* and one Swiss sample [FK1] of *M. n. nivalis*). Removal of *P. theresae* from the taxon sampling did not affect the phylogenetic relationships among the remaining snowfinch species, which suggests that *P. theresae* is not a "rogue taxon" per se (Westover et al. 2013), but instead the poor scores for the *Pyrgilauda* clade were caused by the high proportion of missing data in the SNP sequence inferred from the historical *P. theresae* sample. Exclusion of "poor-quality individuals" can improve the robustness of phylogenetic trees (O' Leary et al. 2018), and in fact, in the second QS analysis (without poor-quality individuals) only a single node showed a significant discord, i.e. the sister-group relationship of *Onychostruthus* and *Pyrgilauda*. This might also explain why the position of *P. theresae* is the only major conflict between nuclear and mitochondrial trees in our study

(due to the high percentage of missing data inferred from that historical toe-pad sample). The inclusion of poor-quality individuals might also have affected QC and QI scores in the *Montifringilla* clade, because the sister-group relationship of *M. adamsi* and *M. nivalis* received full support from QS analysis only when the two poor-quality samples were excluded.

The inferred QD value of 0 suggests that discordant quartets exclusively resolved one of the two possible alternative quartet topologies for each of the three most ancient nodes of the *Onychostruthus/Pyrgilauda* clade. This indicates that the discord might be a result of past reticulate evolution, i.e from past horizontal gene transfer between the two sister clades concerned (Pease et al., 2018; Paetzold et al., 2019). Genome-wide sequence data sets provide a useful matrix for tracing past horizontal gene flow even among deeply divergent clades of a phylogeny, as demonstrated for Himalayan and Tibetan pikas (*Ochotona*; Dahal et al., 2023) or for more shallow divergences among species-level taxa during island radiations (Manthey et al., 2020; Martin et al. 2021). Nevertheless, incomplete lineage sorting (ILS) must be considered as a reason for unresolved phylogenetic relationships at shallow levels of divergence (Palacios et al., 2019; Peters et al., 2007; Wang et al., 2018; Liu et al., 2022) or associated with recent range expansions (Mende and Hundsdoerfer, 2013) in snow finches as well. In snowfinches, poor support for reciprocal monophyly of Alpine and Cantabrian populations of *M. n. nivalis* might be (partly) due to ILS, for example (only the latter group received constant support from different analyses, even from QS analysis). Notably, for such shallow divergences reference genome choice was shown to have strong effects in phylogenetic data sets including high levels of ILS (Rick et al. 2023), and accordingly we found strongest support for two reciprocally monophyletic European snowfinch clades from the two data sets inferred from ingroup-mapping.

In summary, strong node support values and consistency of tree topologies across phylogenetic reconstructions adds further support to previous findings that genome-wide SNP data provide more robust and more reliable phylogenetic hypotheses as compared to a small number of conventional loci combined in multi-locus analyses, for example of mtDNA markers and nuclear introns (Glon et al., 2021). Future studies shall rely on expanded taxon samplings (e.g. subspecific taxa of widespread *M. nivalis*) and data allowing the measurement of gene/locus tree discord to investigate further patterns of diversification and possible reticulate evolution in snowfinches.

4.5. Phylogeny and evolutionary history of snowfinches

Our genome-wide data sets (SNPs and corrected mitogenomes) clearly rejected the postulated paraphyly of the snowfinch genus *Pyrgilauda* (Fig. 2A; Cobos et al., 2021) and confirmed that in accordance with current taxonomy (Gill et al., 2022), *Montifringilla* and *Pyrgilauda* are reciprocally monophyletic (compare Päckert et al., 2021). Furthermore, our phylogenetic reconstructions clearly supported a sister-group relationship of the two genera *Onychostruthus* and *Pyrgilauda*, that were previously united under genus *Pyrgilauda* (Mlíkovský, 1998; Summers-Smith, 2009). Snowfinch vocalizations were shown to facilitate distinction among genera, e.g. according to different flight call types (Gebauer et al., 2006) and to phylogenetic signal of syllable types in songs (Lei et al., 2005). However, a comprehensive taxon-complete bioacoustics analysis for snowfinches is missing so far.

Furthermore, all tree topologies inferred from mitogenomes and genome-wide SNPs clearly confirmed *M. nivalis* as a monophyletic taxon and sister to *M. adamsi*. These findings clearly reject the previously postulated i) paraphyly of *M. nivalis* based on limited sequence information (Päckert et al., 2020b, 2021) and ii) sister-group relationship of *M. nivalis* and *M. henrici* (Cobos et al., 2021), both of which were often treated as conspecifics (Table 1; see also Mayr, 1927; Moreau and Greenway Jr, 1962; Portenko and Vietinghoff-Scheel 1974). Previous authors either argued in favor of a distinctiveness of *M. henrici* against the conspecific *M. nivalis nivalis* and *M. nivalis adamsi* (e.g. Hartert (1921-1922);

Cramp and Perrins, 1994), or for a separation of *M. adamsi* against conspecific *M. n. nivalis* and *M. n. henrici* (e.g. Vaurie, 1959). However, due to the long evolutionary time span of their separation (3.5 – 4.5 mya) and reciprocal monophyly they should represent distinct species-level taxa *M. nivalis*, *M. adamsi* and *M. henrici* (in accordance with Gill et al. 2022). Martens and Eck (1995) agreed on a closer relationship of *M. nivalis* and *M. adamsi* based on morphological traits as confirmed by our phylogeny. Vaurie (1959; p. 587) had synonymized *Montifringillia nivalis groumgrzimaili* with *Montifringilla nivalis alpicola*, however, genetic distinctiveness of the latter two warrants recognition of the latter two as separate (at least subspecific) taxa. The to-date unstudied populations of *M. n. kwenluensis* from the Kunlun Shan, Xinjiang Province (Cheng, 1987; Cramp and Perrins, 1994) could shed some light on any as yet undetected cryptic diversification or on the question of whether the phenotypical cline is associated with past or recent gene flow.

Our SNP-based divergence time estimate for the onset of the snowfinch radiation during the mid-Miocene (7.8 mya; see also Päckert et al., 2020b) was in good accordance with the time estimate inferred from whole-genome data (Qu et al., 2021: 8.3 [5.5 – 11. 4] mya). Taking into account, that mitochondrial DNA is inherited only from females and therefore can only elucidate the evolutionary history of female lineages, we consider the more ancient split ages inferred from the genome-wide SNP data set as the more reliable estimates for the timing of the snowfinch radiation. Generally, divergence time estimates can greatly differ between phylogenies inferred from mitochondrial and from nuclear markers even when the same calibration is applied to the data sets (McCormack et al., 2011; Ritchie et al., 2022). Päckert et al. (2021) identified a single Pleistocene out-of-Tibet dispersal event in snowfinches, because molecular sequence data for the Afghan snowfinch, *P. theresae*, was missing at that time. Our taxon-complete phylogeny showed that *P. theresae* likely originated from a more ancient Pliocene out-of-Tibet split between 5.8 mya (SNP data set) and 3.3 mya probably from a common ancestor with *P. ruficollis* (mitogenome data set). The Afghan endemic *P. theresae* would thus represent a comparatively ancient species-level lineage as compared to many other avian endemics from the Western Himalayas and adjacent mountain ranges that represent taxa of Pleistocene origin (deRaad et al., 2022; Wolfgramm et al., 2021).

The second, more recent out-of-Tibet colonization event in snowfinches was associated with westward dispersal of founder populations of extant *M. n. nivalis* into Europe during the early Pleistocene (2.7 – 1.5 mya; see Päckert et al., 2020). As two major physical barriers, the Pyrenees (and their westward extensions to the Cantabrian Mountains) and the Alps have been identified as centers of recent diversification of several cold-adapted species in interglacial refuges (Hausdorf and Walther, 2021; Stewart et al., 2010; Tribsch and Schönswetter, 2003) and as hotspots of secondary contact zones including hybridization and gene flow across a latitudinal gradient (Ebdon et al., 2021; Hewitt, 2000; Schmitt, 2007). Accordingly, Pleistocene genetic divergence between montane populations from these European mountain systems was detected in European *M. n. nivalis* (this study), conifers (*Abies alba*, Scotti-Saintagne et al., 2021) and butterflies (genus *Erebia*; Schmitt et al., 2016; Vila et al., 2011). Repeated range contractions and expansions due to glacial cycles might have enhanced gene flow within and across mountain systems, particularly in conjunction with range expansion towards lower elevations during glacial periods (Stewart et al., 2010).

6. Conclusions

For current and future biodiversity research in a rapidly changing world, museum collections have become increasingly relevant as archives of occurrence records, species distributions and of phenotypic and genetic diversity (review in Meineke et al., 2019). Research on understudied regional faunal and floral assemblages benefits greatly from inclusion of collection materials to fill in current data and knowledge gaps (Figueira and Lages, 2019). Böhme and Jablonski (2022) emphasized the basic need for multilateral co-operations among museums and universities from Afghanistan and partner

countries for future sustainable research on the country's biodiversity and natural heritage. To date, knowledge on Afghan biodiversity has been almost entirely inferred from collection-based research on material collected prior to the era of molecular genetics (Jablonski et al., 2019, 2021). All endemic vertebrate species of Afghanistan were described from material collected during expeditions in the 1960s and earlier, such as freshwater fishes (Coad and Bogutskaya, 2012) or racerunner lizards (genus *Eremias*; Anderson and Leviton, 1967; Böhme et al., 1991). Genetic research mainly relied on historical collection material for inclusion of Afghan taxa/populations in molecular phylogenies (lizards: Orlova et al., 2022; birds: Päckert et al., 2020a).

Newer material from wild Afghan populations for genetic analyses could only be collected during a short period of time when international armed forces were present in the country, for example from military camps (Krüger et al., 2011) or from Air Force bases (GenBank sequences included in Khan et al., 2021). Field studies during this short time window enabled the discovery of Afghan populations of rare or narrowly distributed species that had been overlooked or misidentified in ornithological collections like the large-billed reed warbler *Acrocephalus orinus* (Ayé et al. 2011; Timmins et al. 2009, 2010). Based on that extended knowledge of Central Asian breeding populations of that species, re-examination and DNA barcoding of Central Asian museum specimens of reed warblers revealed that several historical *A. orinus* specimens had been mistaken for a more widespread congeneric *A. dumentorum* (Koblik et al., 2011). Such novel information and material from modern field research helped refining breeding ranges, but also offered new perspectives for population genetic or genomic research. For example, genetic analyses of rare herbarium specimens from the 20th century (1955 – 1978) documented a decline of genetic diversity over time in an Afghan landrace of wheat (Terasawa et al., 2009). On the one hand, extension of such time-limited series from collections with fresh material from the 21st century will facilitate the detection of up-to-date demographic changes for example in comparison with climate or land-use change.

On the other hand, along with optimization of wet-lab protocols and bioinformatics pipelines for NGS-based data sets, previous (partly incomplete) phylogenetic hypotheses can be tested using high-quality sequence data using freshly collected material. Likewise, typical sources of error in published sequence data can be detected, like in mitogenomes of *O. tazcanowskii* and *P. davidiana*. Dense taxon-sampling and increased robustness of molecular trees will also provide improved phylogenetic backbones for future meta-analyses of traits' evolution, niche evolution and many other research questions. As a perspective for snowfinch research, future population genetic studies should rely on a broader, trans-European sampling to identify further possible relict lineages on the Italian Peninsula, the Balkan Peninsula or Corsica. Similarly, filling in the sampling gaps in Asia would be necessary to identify a possible isolation-by-distance pattern in the most recent dispersal event of Asian *M. nivalis* to the Caucasus in a stepping-stone process *via* Central Asia (*M. n. tianshanica*), the Hindukush, the Elburz (ssp. *alpicola*) and the Zagros mountains (*M. n. raddei;* see del Hoyo and Collar, 2016). The as yet unstudied populations from Asia Minor (*M. n. leucura* from southern Anatolia) might play a key role in the understanding of geographical limits or potential overlap between Asian and European genetic lineages of the white-winged snowfinch, *M. nivalis*.

Acknowledgements

This study was substantially funded by Deutsche Forschungsgemeinschaft, DFG [grant number PA1818/3-2]. J.M. was granted financial aid for field research in Asia by the Feldbausch-Stiftung and the Wagner-Stiftung at the Fachbereich Biologie of Mainz University [no grant numbers]. M.M.D. was financially supported by Ayudas de Incorporación Científico Titular (#202230I042; CSIC). Mitogenome sequencing was performed at the Molecular Laboratory Senckenberg Dresden (SGN-SNSD-Mol-Lab).

We thank S. Birks from the Burke Museum of Natural History and Culture (Seattle, USA) and A. Lischke (Switzerland) for providing further samples. We would also like to thank C. Bettega for her help in collecting blood samples in Spain. MMD collected blood samples under the permission #2016/032415 of the Principality of Asturias, PL under the permit CO/08/058/2012 of Picos de Europa National Park. Tissue samples from Switzerland were collected from birds that were incidentally found dead within the research project 530200 on the population dynamics of snowfinches of the Swiss Ornithological Institute. Voluntary notification of access and utilization of genetic resources from Switzerland was submitted to the Swiss Federal Office for the Environment (FOEN) and is registered under the register ID N210001. We cordially thank F. Bosshard for advice and support during the formal registration process. We thank M. Stubbe for information on the Mongolian-German Biological Expeditions. We are particularly grateful to I. Overcast (Virginia Museum of Natural History, Martinsville, VA, USA) for his kind support with some specifications of ipyrad protocols.

Declaration of interest: None

Declaration of generative AI in scientific writing: No AI-assisted technologies were used in the writing process of this manuscript, except the automatic check for spelling and grammar as implemented in WORD.

CRediT authorship contribution statement:

Safiqul Islam: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, writing – Original draft, review and editing, Visualization. **Claire Peart**: Methodology, Formal analysis, Writing – review and editing. **Christian Kehlmaier**: Validation, Formal analysis, Investigation, Data curation, Writing – Review and editing. **Andreas Dahl, Sylvia Klemroth:** Methodology, Validation, Investigation, Writing – Review and editing. **Yue-Hua Sun, Fumin Lei, Maria del Mar Delgado, Paola Laiolo, Juan Carlos Illera, Sebastian Dirren, Sabine Hille, Davaa Lkhagvasuren, Till Töpfer, Martin Kaiser, Axel Gebauer**: Resources, Writing – review and editing. **Dimitra Alexopoulou**: Formal analysis, Writing – review and editing. **Claudia Paetzold**: Software, Validation, Formal analysis, Writing – Original draft, review and editing. **Jochen Martens**: Conceptualization, Resources, Funding acquisition, Writing – review and editing. **Martin Päckert**: Conceptualization, Investigation, Data curation, writing – Original draft, review and editing, Visualization, Supervision, Project administration, Funding acquisition

References

Anderson, S. C., & Leviton, A. E. (1967). A new species of *Eremias* (Reptilia: Lacertidae) from Afghanistan. *Occasional Papers of the California Academy of Sciences, 64*, 1–4.

Andrews, S. (2015). *FastQC: A quality control tool for high throughput sequence data* (p. 0.11.4). https://www. bioinformatics.babraham.ac.uk/projects/fastqc/.

Ayé, R., Hertwig, S. T., & Schweizer, M. (2011). Discovery of a breeding area of the enigmatic Large-billed Reed Warbler *Acrocephalus orinus*. *Journal of Avian Biology, 41*(5), 452–459. https://doi.org/10.1111/j.1600-048X.2010.05031.x

BirdLife International. (2021). *Species factsheet: Pyrgilauda theresae. Downloaded from* https://www.birdlife.org on 05/11/2021.

Bohling, J. (2020). Evaluating the effect of reference genome divergence on the analysis of empirical RADseq datasets. *Ecology and Evolution*, *10*(14), 7585–7601. https://doi.org/10.1002/ece3.6483

Böhme, W., & Jablonski, D. (2022). Making forgotten information available: An early study on the Afghanistan Mountain Salamander *Paradactylodon* (*Afghanodon*) *mustersi* (Smith, 1940) (Caudata: Hynobiidae). *Bonn Zoological Bulletin, 71*(1), 1–7. https://doi.org/10.20363/BZB-2022.71.1.001

Böhme, W., & Ščerbak, N. N. (1991). Ein neuer Wüstenrenner aus dem Hochland Afghanistans, *Eremias* (*Eremias*) *afghanistanica* sp.n. (Reptilia: Sauria: Lacertidae). *Bonner zoologische Beiträge, 42*(2), 137–141.

Botero-Castro, F., Delsuc, F., & Douzery, E. J. P. (2016). Thrice better than once: quality control guidelines to validate new mitogenomes. *Mitochondrial DNA. Part A, DNA Mapping, Sequencing, and Analysis*, *27*(1), 449–454. https://doi.org/10.3109/19401736.2014.900666

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*(4), 1–6.https://doi.org/10.1371/journal.pcbi.1003537

Bouckaert, R. R. (2010). DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*, *26*(10), 1372–1373. https://doi.org/10.1093/bioinformatics/btq110

Bourgeois, Y. X. C., Lhuillier, E., Cézard, T., Bertrand, J. A. M., Delahaie, B., Cornuault, J., Duval, T., Bouchez, O., Milá, B., & Thébaud, C. (2013). Mass production of SNP markers in a nonmodel passerine bird through RAD sequencing and contig mapping to the zebra finch genome. *Molecular Ecology Resources*, *13*(5), 899–907. https://doi.org/10.1111/1755-0998.12137

Briscoe, A. G., Hopkins, K. P., & Waeschenbach, A. (2016). High-Throughput Sequencing of Complete Mitochondrial Genomes. *Methods in Molecular Biology (Clifton, N.J.)*, *1452*, 45–64. https://doi.org/10.1007/978-1-4939-3774-5_3

Bushnell, B., Rood, J., & Singer, E. (2017). BBMerge – Accurate paired shotgun read merging via overlap. *PLOS ONE*, *12*(10), e0185056. https://doi.org/10.1371/journal.pone.0185056

Busuttil, S., & Ayé, R. (2009). Ornithological surveys in Bamiyan province, Islamic republic of Afghanistan. *Sandgrouse, 31*, 146–59.

Busuttil, S., Ayé, R., & Shank, C. (2010). Notes on nest and breeding of Afghan snowfinch *Pyrgilauda theresae*, Bamiyan province, Afghanistan. *Sandgrouse, 32*, 13–14.

Bybee, S. M., Kalkman, V. J., Erickson, R. J., Frandsen, P. B., Breinholt, J. W., Suvorov, A., Dijkstra, K.-D. B., Cordero-Rivera, A., Skevington, J. H., Abbott, J. C., Sanchez Herrera, M., Lemmon, A. R., Lemmon, E. M., & Ware, J. (2021). Phylogeny and classification of Odonata using targeted genomics. *Molecular Phylogenetics and Evolution, 160*, 107115. https://doi.org/10.1016/j.ympev.2021.107115

Card, D. C., Shapiro, B., Giribet, G., Moritz, C., & Edwards, S. V. (2021). Museum Genomics. *Annual Review Genetics, 55*, 633–59. https://doi.org/10.1146/annurev-genet-071719-020506

Catchen, J., Hohenlohe, P. A., Bassham, S., & Amores, A. (2013). Stacks : an analysis tool set for population genomics. *Molecular Ecology*, 3124–3140. https://doi.org/10.1111/mec.12354

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *G3 Genes|Genomes|Genetics*, *1*(3), 171–182. https://doi.org/10.1534/g3.111.000240

Cheng, T. (1987). A synopsis to the avifauna of China. Paul Parey Scientific Publ., Hamburg, Berlin.

Clements, J. F., Schulenberg, T. S., Iliff, M. J., Fredericks, T. A., Gerbracht, J. A., Lepage, D., Billerman, S. M., Sullivan, B. L. & Wood, C. L. (2022). The eBird/Clements checklist of Birds of the World: v2022. Downloaded from https://www.birds.cornell.edu/clementschecklist/download/ (last accessed on 15.02.2024)

Clements, J. F., Rasmussen, P. C., Schulenberg, T. S., Iliff, M. J., Fredericks, T. A., Gerbracht, J. A., Lepage, D., Spencer, A., Billerman, S. M., Sullivan, B. L., & Wood, C. L. (2023). The eBird/Clements checklist of Birds of the World: v2023. Downloaded from https://www.birds.cornell.edu/clementschecklist/download/ (last accessed on 15.02.2024)

Coad, B. W., & Bogutskaya, N. G. (2012). A new species of riffle minnow, *Alburnoides holciki*, from the Hari River basin in Afghanistan and Iran (Actinopterygii: Cyprinidae). *Zootaxa, 3453*, 43–55. https://doi.org/10.11646/zootaxa.3453.1.3

Cobos, M. E., Cheng, Y., Song, G., Lei, F., & Peterson, T. A. (2021). New distributional opportunities with niche innovation in Eurasian snowfinches. *Journal of Avian Biology, 52*, e02868. https://doi.org/10.1101/2021.04.06.438738

Cong, Q., Shen, J., Zhang, J., Li, W., Kinch, L. N., Calhoun, J. V., Warren, A. D., & Grishin N. V. (2021). Genomics reveals the origins of historical specimens. *Molecular Biology and Evolution, 38*, 2166–2176. https://doi.org/10.1093/molbev/msab013

Cooper, E. A., & Uy, J. A. C. (2017). Genomic evidence for convergent evolution of a key trait underlying divergence in island birds. *Molecular Ecology, 26*, 3760-3774. https://doi.org/10.1111/mec.14116

Cramp, S., Perrins, C. M. (1994). Handbook of the birds of Europe the Middle East and North Africa, Volume VIII, Crows to Finches. Oxford University Press, Oxford,New York.

Crotti, M., Barratt, C. D., Loader, S. P., Gower, D. J., & Streicher, J. W. (2019). Causes and analytical impacts of missing data in RADseq phylogenetics: Insights from an African frog (*Afrixalus*). *Zoologica Scripta*, *48*(2), 157–167. https://doi.org/10.1111/zsc.12335

Dahal, N., Romine, M. G., Khatiwara, S., Ramakrishnan, U., & Lamichhaney, S. (2023). Gene flow drives genomic diversity in Asian Pikas distributed along the core and range-edge habitats in the Himalayas. *Ecology and Evolution, 13*, e10129. https://doi.org/10.1002/ece3.10129

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & Group, 1000 Genomes Project Analysis. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), giab008. https://doi.org/10.1093/gigascience/giab008

de Raad, J., Päckert, M., Irestedt, M., Janke, A., Kryukov, A. P., Martens, J., Red'kin, Y. A., Sun, Y. H., Töpfer, T., Schleuning, M., Neuschulz, E. L. & Nilsson, M. A. (2022). Speciation and population divergence in a mutualistic seed dispersing bird. *Communications Biology, 5*(1), 429. https://doi.org/10.1038/s42003-022-03364-2

del Hoyo, J., & Collar, N. J. (2016). Illustrated checklist of the birds of the world. Volume 2 Passerines.

Lynx Edicions, Barcelona.

del Mar Delgado, M., Bettega, C., Martens, J., & Päckert, M. (2019). Ecotypic changes of alpine birds to climate change. *Scientific Reports*, *9*(1), 16082. https://doi.org/10.1038/s41598-019-52483-0

Dementiev, G. (1963). Theresa's Ground Finch *Montifringilla theresae*. *Ibis*, *105*(1), 107. https://doi.org/10.1111/j.1474-919X.1963.tb02479.x

Dickinson, E. C. (2003). The Howard and Moore Complete Checklist of the Birds of the World. Third edition. Princeton University Press, Princeton and Oxford.

Dickinson, E. C., & Christidis, L. (Eds) (2014). The Howard and Moore Complete Checklist of the Birds of the World. Fourth edition, Vol. 2: Passerines. Aves Press, Eastbourne, UK.

Eaton, D. A. R. (2014). PyRAD: assembly of de novo RADseq loci for phylogenetic analyses . *Bioinformatics*, *30*(13), 1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Eaton, D. A. R., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btz966

Eaton, D. A. R., & Ree, R. H. (2013). Inferring Phylogeny and Introgression using RADseq Data: An Example from Flowering Plants (*Pedicularis*: Orobanchaceae). *Systematic Biology*, *62*(5), 689–706. https://doi.org/10.1093/sysbio/syt032

Eaton, D. A. R., Spriggs, E. L., Park, B., & Donoghue, M. J. (2017). Misconceptions on Missing Data in RAD-seq Phylogenetics with a Deep-scale Example from Flowering Plants. *Systematic Biology*, *66*(3), 399–412. https://doi.org/10.1093/sysbio/syw092

Ebdon, S., Laetsch, D. R., Dapporto, L., Hayward, A., Ritchie, M. G., Dinca, V., Vila. R., & Lohse, K. (2021). The Pleistocene species pump past its prime: Evidence from European butterfly sister species. *Molecular Ecology, 30*(14), 3575–3589. https://doi.org/10.1111/mec.15981

Efetov, K. A., Hofmann, A., & Tarmann, G. M. (2014). Application of two molecular approaches (use of sex attractants and DNA barcoding) allowed to rediscover *Zygaenoprocris eberti* (Alberti, 1968) (Lepidoptera, Zygaenidae, Procridinae), hitherto known only from the female holotype. *Nota Lepidopterologica 37*(2), 151-160. https://doi.org/10.3897/nl.37.7871

Elgvin, T. O., Trier, C. N., Tørresen, O. K., Hagen, I. J., Lien, S., Nederbragt, A. J., Ravinet, M., Jensen, H., & Sætre, G. P. (2017). The genomic mosaicism of hybrid speciation. *Science Advances, 3*, e1602996. https://doi.org/10.1126/sciadv.1602996

Ekblom, R., & Wolf, J. B. W. (2014). A field guide to whole-genome sequencing, assembly and annotation. *Evolutionary Applications*, *7*(9), 1026–1042. https://doi.org/https://doi.org/10.1111/eva.12178

Ernst, M., Jonsson, K. A., Ericson, P. G. P., Blom, M. P. K., & Irestedt, M. (2022). Utilizing museomics to trace the complex history and species boundaries in an avian-study system of conservation concern. *Heredity, 128*, 159–168. https://doi.org/10.1038/s41437-022-00499-0

Fiedler de Abreu Jr, E., Pavan, S. E., Tschuiya, M. T. N., Wilson, D. E., Percequillo, A. R., & Maldonado, J. E. (2020). Museomics of tree squirrels: a dense taxon sampling of mitogenomes reveals hidden diversity, phenotypic convergence, and the need of a taxonomic overhaul. *BMC Evolutionary Biology, 20*, 77. https://doi.org/10.1186/s12862-020-01639-y

Figueira, R., & Lages, F. (2019). Museum and Herbarium Collections for Biodiversity Research in Angola. In: Huntley, B., J., Russo, V., Lages, F., Ferrand, N. (Eds), Biodiversity of Angola. Springer, Cham, Switzerland, 513-542.

Fortes, G. G., & Paijmans, J. L. A. (2015). Analysis of Whole Mitogenomes from Ancient Samples BT - Whole Genome Amplification: Methods and Protocols. In T. Kroneis (Ed.), *Whole Genome Amplification. Methods in Molecular Biology, vol 1347*. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-2990-0_13

Galla, S. J., Forsdick, N. J., Brown, L., Hoeppner, M. P., Knapp, M., Maloney, R. F., Moraga, R., Santure, A. W., & Steeves, T. E. (2018). Reference Genomes from Distantly Related Species Can Be Used for Discovery of Single Nucleotide Polymorphisms to Inform Conservation Management. *Genes*, *10*(1). https://doi.org/10.3390/genes10010009

Gauthier, J., Pajkovic, M., Neuenschwader, S., Kaila, L., Schmid, S., Orlando, L., & Alvarez, N. (2020). Museomics identifies genetic erosion in two butterfly species across the 20th century in Finland. *Molecular Ecology Resources, 20*, 1191–1205. https://doi.org/10.1111/1755-0998.13167

Gebauer, A., Eck, S., Kaiser, M., Lei, FM, & Martens, J. (2006). The Qinghai-Tibet Plateau : center of evolution for snow sparrows (*Montifringilla s. str.*) and mountain-steppe sparrows (*Onychostruthus , Pyrgilauda*). *Acta Zoologica Sinica, 52*, 305–309.

Gill, F., Donsker, D., & Rasmussen, P. (Eds)(2023). IOC World Bird List (v13.2). Doi 10.14344/IOC.ML.13.2. http://www.worldbirdnames.org/ (last accessed on 27.02.2024)

Glenn, T. C., Stephan, W., & Braun, M. J. (1999). Effects of a Population Bottleneck on Whooping Crane Mitochondrial DNA Variation. *Conservation Biology*, *13*(5), 1097–1107. http://www.jstor.org/stable/2641743

Glon, H., Quattrini, A., Rodríguez, E., Titus, B. M., & Daly, M. (2021). Comparison of sequence-capture and ddRAD approaches in resolving species and populations in hexacorallian anthozoans. *Molecular Phylogenetics and Evolution*, *163*, 107233. https://doi.org/10.1016/j.ympev.2021.107233

Glutz von Blotzheim, U., & Bauer, K. M. (1997). Handbuch der Vögel Mitteleuropas. Band 14/1. Passeriformes (5. Teil). Aula-Verlag, Wiesbaden, Germany.

Godoy, J. A., Negro, J. J., Hiraldo, F., & Donázar, J. A. (2004). Phylogeography, genetic structure and diversity in the endangered bearded vulture (*Gypaetus barbatus*, L) as revealed by mitochondrial DNA. *Molecular Ecology*, *13*(2), 371–390. https://doi.org/10.1046/j.1365-294x.2003.02075.x

Gori, M., Pecchioli, S., Giordani, E., Saeedi, M. A., Wafa, F. H., & Biricolti, S. (2019). Barcoding assessment of the *Citrus* species cultivated in eastern Afghanistan. *Advances in Horticultural Science, 33*(3), 403-408. https://doi.org/10.13128/ahs-25648

Guo, X., Dai, X., Chen, D., Papenfuss, T.J., Ananjeva, N.B., Melnikov, D.A., & Wang, Y. (2011). Phylogeny and divergence times of some racerunner lizards (Lacertidae: *Eremias*) inferred from mitochondrial 16S rRNA gene segments. *Molecular Phylogenetics and Evolution, 61*, 400–412. https://doi.org/10.1016/j.ympev.2011.06.022

Gwee, C. Y., Eaton, J. A., Garg, K. M., Alström, P., Van Balen, S., Hutchinson, R. O., Prawiradilaga, D. M., Le, M. H., & Rheindt, F. E. (2019). Cryptic diversity in *Cyornis* (Aves: Muscicapidae) jungle-flycatchers flagged by simple bioacoustic approaches. *Zoological Journal of the Linnean Society*, *186*(3), 725–741. https://doi.org/10.1093/zoolinnean/zlz003

Hahn, C., Bachmann, L., & Chevreux, B. (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads - a baiting and iterative mapping approach. *Nucleic Acids Research*, *41*(13), e129. https://doi.org/10.1093/nar/gkt371

Hahn, E. E., Grealy, A., Alexander, M., & Holleley, C. E. (2020). Museum Epigenomics: Charting the Future by Unlocking the Past. Trends in Ecology and Evolution, 35(4), 295–300. https://doi.org/10.1016/j.tree.2019.12.005

Hanson, T., Brooks, T. M., Da Fonseca, G. A. B., Hoffmann, M., Lamoreux, J. F., Machlis, G., Mittermeier, C. G., Mittermeier, R. A., & Pilgrim, J. D. (2009). Warfare in biodiversity hotspots. *Conservation Biology*, *23*(3), 578–587. https://doi.org/10.1111/j.1523-1739.2009.01166.x

Hartert, E. (1910). Die Vögel der Paläarktischen Fauna. Friedländer & Sohn, Berlin.

Hartert, E. (1921-1922). Die Vögel der Paläarktischen Fauna. Volume III. Friedländer & Sohn, Berlin.

Hartert, E., & Steinbacher, F. (1932). Die Vögel der Paläarktischen Fauna, Ergänzungsband. Friedländer & Sohn, Berlin.

Hausdorf, B., & Walther, F. (2021). A forgotten subspecies of the land snail species *Arianta arbustorum* from a Pleistocene refuge in the Western Alps. *Journal of Molluscan Studies*, *87*, 1–7. https://doi.org/10.1093/mollus/eyab017

Hekkala, E. R., Saumure, R. A., Jaeger, J. R., Herrmann, H.-W., Sredl, M. J., Bradford, D. F., Drabeck, D., & Blum, M. J. (2011). Resurrecting an extinct species: archival DNA, taxonomy, and conservation of the Vegas Valley leopard frog. *Conservation Genetics, 12*, 1379–1385. https://doi.org/10.1007/s10592-011-0229-6

Heupink, T. H., van Grouw, H., & Lambert, D. M. (2014). The mysterious Spotted Green Pigeon and its relation to the Dodo and its kindred. *BMC Evolutionary Biology*, *14*(1), 136. https://doi.org/10.1186/1471-2148-14-136

Hewitt, G. (2000). The genetic legacy of the Ice Ages. *Nature, 405*, 907–913. https://doi.org/10.1038/35016000

Hilton, E. J., Watkins-Colwell, G. J., & Huber, S. K. (2021). The Expanding Role of Natural History Collections. *Ichthyology and Herpetology*, *109*(2), 379–391. https://doi.org/10.1643/t2020018

Hofstetter, V., Buyck, B., Eyssartier, G., Schnee, S., & Gindro, K. (2019). The unbearable lightness of sequenced-based identification. *Fungal Diversity*, *96*(1), 243–284. https://doi.org/10.1007/s13225-019-00428-3

Horn, S. (2012). Target enrichment via DNA hybridization capture. *Methods in Molecular Biology*, *840*, 177–188. https://doi.org/10.1007/978-1-61779-516-9_21

Huson, D. H., & Bryant, D. (2006). Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution*, *23*(2), 254–267. https://doi.org/10.1093/molbev/msj030

Jablonski, D., Urošević, A., Andjelković, M., & Džukić, G. (2019). An unknown collection of lizards from Afghanistan. *ZooKeys*, *843*, 129–147. https://doi.org/10.3897/ZOOKEYS.843.29420

Jablonski, D., Basit, A., Farooqi, J., Masroor, R., & Böhme, W. (2021). Biodiversity research in a changing Afghanistan. *Science, 6549*, 1402. https://doi.org/10.1126/science.abj8118

Jiang, H., Lei, R., Ding, S.-W., & Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics*, *15*(1), 182. https://doi.org/10.1186/1471-2105-15-182

Jønsson, K. A., Blom, M. P. K., Päckert, M., Ericson, P. G. P., & Irestedt, M. (2018). Relicts of the lost arc: High-throughput sequencing of the *Eutrichomyias rowleyi* (Aves: Passeriformes) holotype uncovers an ancient biogeographic link between the Philippines and Fiji. *Molecular Phylogenetics and Evolution*, *120*, 28–32. https://doi.org/10.1016/j.ympev.2017.11.021

Kanderian, N., Lawson, D., & Zahler, P. (2011). Current status of wildlife and conservation in Afghanistan. *International Journal of Environmental Studies, 68*(93), 281–298. https://doi.org/10.1080/00207233.2011.573960

Karimi, M. O, Shariati, M. M,. Zerehdaran, S., Moradi, M. H., & Javadmanesh, A. (2016). Study of genetic diversity of sheep breeds In Afghanistan. *Bioscience Biotechnology Research Asia, 13*(1), 573–581. http://dx.doi.org/10.13005/bbra/2072

Kazemi, E., Nazarizadeh, M., Fatemizadeh, F., Khani, A., & Kaboli, M. (2021), The phylogeny, phylogeography, and diversification history of the westernmost Asian cobra (Serpentes: Elapidae: *Naja oxiana*) in the Trans-Caspian region. *Ecology and Evolution, 11*, 2024–2039. https://doi.org/10.1002/ece3.7144

Kearns, A.M., Campana, M. G., Slikas, B., Berry, L., Saitoh, T., Cibois, A., & Fleischer, R. C. (2022). Conservation genomics and systematics of a near-extinct island radiation. *Molecular Ecology, 31*(7), 1995–2012. https://doi.org/10.1111/mec.16382

Kehlmaier, C., Albury, N. A., Steadman, D. W., Graciá, E., Franz, R., & Fritz, U. (2021). Ancient mitogenomics elucidates diversity of extinct West Indian tortoises. *Scientific Reports*, *11*(1), 1–9. https://doi.org/10.1038/s41598-021-82299-w

Kehlmaier, C., Zhang, X., Georges, A., Campbell, P., Thomson, S., & Fritz, U. (2019). Mitogenomics of historical type specimens of Australasian turtles: clarification of taxonomic confusion and old mitochondrial introgression. *Scientific Reports*, *9*, 5841. https://doi.org/10.1038/s41598-019-42310-x

Khalilipur, O., Rezaei, H. R., Shabani, A. A., Kaboli, M., Ashrafi, S., & Ashrafzadeh, M. R. (2017). Assessment of the Phylogenetic Status of Afghan Pika (*Ochotona rufescens*). *Journal of Wildlife and Biodiversity, 1*(1), 12–18. https://doi.org/10.22120/jwb.2017.26905

Khan, M. A., Jablonski, D., Nadeem, M. S., Masroor, R., Kehlmaier, C., Spitzweg, C. ,& Fritz, U. (2021). Molecular phylogeny of *Eremias* spp. from Pakistan contributes to a better understanding of the diversity of racerunners. *Journal of Zoological Systematics and Evolutionary Research, 59*, 466–483. https://doi.org/10.1111/jzs.12426

Koblik, E. A, Red'kin, Y. A., Meer, M. S., Derelle, R., Golenkina, S. A., Kondrashov, F. A., & Arkhipov, V. Y. (2011). *Acrocephalus orinus*: A Case of Mistaken Identity. *PLoS One, 6*(4), e17716. https://doi.org/10.1371/journal.pone.0017716

Kong, H., Condamina, F. L., Yang, L., Harris, A., Feng, C., Wen, F., & Kang, M. (2022). Phylogenomic and Macroevolutionary Evidence for an Explosive Radiation of a Plant Genus in the Miocene. Systematic Biology, 71(3), 589-609. https://doi.org/10.1093/sysbio/syab068

Krüger, A., Strüven, L., Post, R. J., & Faulde, M. (2011). The sandflies (Diptera: Psychodidae, Phlebotominae) in military camps in northern Afghanistan (2007–2009), as identified by morphology and DNA 'barcoding'. *Annals of Tropical Medicine & Parasitology, 105*(2), 163-176.

https://doi.org/10.1179/136485911X12899838683241

Kuhn, K., Schwenk, K., Both, C., Canal, D., Johansson, U. S., van der Mije, S., Töpfer, T., & Päckert, M. (2013). Differentiation in neutral genes and a candidate gene in the pied flycatcher: using biological archives to track global climate change. *Ecology and Evolution*, *3*(14), 4799–4814. https://doi.org/10.1002/ece3.855

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Lalueza-Fox, C. (2022). Museomics. *Current Biology, 32*, 1214–1215. https://doi.org/10.1016/j.cub.2022.09.019

Lei, F. M., Wang, A. Z., Wang, G., & Yin, Z. H. (2005). Vocalizations of red-necked snow finch, *Pyrgilauda ruficollis* on the Tibetan Plateau, China - A syllable taxonomic signal? *Folia Zoologica*, *54*(1–2), 135–146.

Lei, F., Qu, Y., & Song, G. (2014). Species diversification and phylogeographical patterns of birds in response to the uplift of the Qinghai-Tibet Plateau and Quaternary glaciations. *Current Zoology 60*(2), 149–161. http://dx.doi.org/10.1093/czoolo/60.2.149

Leigh, J. W., & Bryant, D. (2015). *POPART : full-feature software for haplotype network construction*. 1110–1116. https://doi.org/10.1111/2041-210X.12410

Lerner, H. R. L., Meyer, M., James, H. F., Hofreiter, M., & Fleischer, R. C. (2011). Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. *Current Biolology, 21*, 1838–1844. https://doi.org/10.1016/j.cub.2011.09.039

Liu, X., Zhang, S., Cai, Z., & Li, K. (2022). Genomic insights into zokors' phylogeny and speciation in China. *Proceedings of the National Academy of Sciences, 119*(19), e2121819119. https://doi.org/10.1073/pnas.2121819119

Ma, Y., Huang, Y., & Lei, F. (2014). Sequencing and phylogenetic analysis of the *Pyrgilauda ruficollis* (Aves,Passeridae) complete mitochondrial genome. *Zoological Research, 35*(2), 81–91. https://doi.org/10.11813/j.issn.0254-5853.2014.2.081

Ma, Y., Wang, X., Hunag, Y., & Lei F. (2016). Complete mitochondrial genome of the *Onychostruthus taczanowskii* (Aves, Passeridae) and its phylogenetic analysis. *Mitochondrial DNA, 27*(4), 2984–2985. https://doi.org/10.3109/19401736.2015.1060473

Maddock, S., Briscoe, A., Wilkinson, M., Waeschenbach, A., San Mauro, D., Day, J. J., Littlewood, T. J., Foster, P. G., Nussbaum, R. A., & Gower, D. J. (2016). Next-generation mitogenomics: A comparison of approaches applied to caecilian amphibian phylogeny. In *PLOS One*. PLOS One. https://doi.org/10.1371/journal.pone.0156757

Maheshwari, A. (2022). Biodiversity conservation in Afghanistan under the returned Taliban. *Nature Ecology & Evolution*, *6*(4), 342—343. https://doi.org/10.1038/s41559-021-01655-1

Manthey, J. D., Oliveros, C. H., Andersen, M. H., Filardi, C. E., & Moyle, R. G. (2020). Gene flow and rapid differentiation characterize a rapid insular radiation in the southwest Pacific (Aves: *Zosterops*). Evolution, 74(8), 1788-1803. https://doi.org/10.1111/evo.14043

Marchese, C. (2015). Biodiversity hotspots: A shortcut for a more complicated concept. *Global*

Ecology and Conservation, 3, 297–309. https://doi.org/10.1016/j.gecco.2014.12.008

Maricic, T., Whitten, M., & Pääbo, S. (2010). Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLOS ONE*, *5*(11), 1–5. https://doi.org/10.1371/journal.pone.0014004

Martens, J., & Eck, S. (1995). Towards an ornithology of the Himalayas - Systematics, ecology and vocalizations of Nepal birds. *Bonner Zoologische Monographien, 38*, 1–445.

Martin, C. A., Armstrong, C., Illera, J. C., Emerson, B. C., Richardson, D. S., & Spurgin, L. G. (2021). Genomic variation, population history and within-archipelago adaptation between island bird populations. *Royal Society Open Science, 8*, 201146. https://doi.org/10.1098/rsos.201146

Mason, N. A., & Unitt, P. (2018). Rapid phenotypic change in a native bird population following conversion of the Colorado Desert to agriculture. *Journal of Avian Biology*, *49*(2), jav-01507. https://doi.org/https://doi.org/10.1111/jav.01507

Mayr, E. (1927) Die Schneefinken (Gattungen *Montifringilla* und *Leucosticte*). *Journal für Ornithologie, 75*, 596 - 619.

Mayr, E., & Greenway, Jr. (Eds)(1962). Checklist of the Birds of the World. Volume XV. Museum of Comparative Zoology, Cambridge, Massachusetts.

McCormack, J. E., Heled, J., Delaney, K. S., Peterson, A. T., & Knowles, L. L. (2011). Calibrating Divergence Times On Species Trees Versus Gene Trees: Implications For Speciation History Of Aphelocoma Jays. *Evolution*, *65*(1), 184–202. https://doi.org/10.1111/j.1558-5646.2010.01097.x

McNeely, J. A. (2003). Conserving forest biodiversity in times of violent conflict. *Oryx*, *37*(2), 142–152. https://doi.org/10.1017/S0030605303000334

Meineke, E. K., Davies, T. J., Daru, B. H., & Davis, C. C. (2019). Biological collections for understanding biodiversity in the Anthropocene. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1763), 20170386. https://doi.org/10.1098/rstb.2017.0386

Meinertzhagen, R. (1937). *Montifringilla theresae* sp. nov. *Bulletin of the British Ornithological Club, 58*(407), 10.

Mende, M. B., & Hundsdoerfer, A. K. (2013). Mitochondrial lineage sorting in action - Historical biogeography of the *Hyles euphorbiae* complex (Sphingidae, Lepidoptera) in Italy. *BMC Evolutionary Biology*, *13*(1). https://doi.org/10.1186/1471-2148-13-83

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *2010*(6), pdb.prot5448. https://doi.org/10.1101/pdb.prot5448

Milne, I., Stephen, G., Bayer, M., Cock, P. J. A., Pritchard, L., Cardle, L., Shaw, P. D., & Marshall, D. (2013). Using Tablet for visual exploration of second-generation sequencing data. *Briefings in Bioinformatics*, *14*(2), 193–202. https://doi.org/10.1093/bib/bbs012

Miranda, I., Giska, I., Farelo, L., Pimenta, J., Zimova, M., Bryk, J., Dalén, L., Mills, L. S., Zub, K., & Melo-Ferreira, J. (2021). Museomics Dissects the Genetic Basis for Adaptive Seasonal Coloration in the Least Weasel. *Molecular Biology and Evolution, 38*(10), 4388–4402. https://doi.org/10.1093/molbev/msab177

Mittermeier, R. A., Seligmann, P. A., & Ford, H. (Eds)(2004). Hotspots revisited. CEMEX, Mexico City.

Mlíkovský, J. (1998). Generic name of southern snowfinches. *Forktail, 14*, 85.

Moreau, R. E., & Greenway, Jr. J. C. (1962). Family Ploceidae, Weaverbirds. In: Mayr E, Greenway Jr JC (eds) Check-List of the Birds of the World. Volume XV. Museum of Comparative Zoology, Cambridge Massachusetts, USA.

Niethammer, G. (1973). Zur Vogelwelt des afghanischen Pamir und des Darwaz. *Bonner Zoologische Beiträge*, *24*, 270–284.

Nylander, J. A. A. (2004). MrModeltest 2.3. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology, 27*, 3193–3206. https://doi.org/10.1111/mec.14792

Obiol, F., James, J., James, H. F., Chesser, R. T., Bretagnolle, V., González-Solís, J., Rozas, J., Riutort, M., & Welch, A. J. (2021). Integrating Sequence Capture and Restriction Site-Associated DNA Sequencing to Resolve Recent Radiations of Pelagic Seabirds. *Systematic Biology*, *70*(5), 976–996. https://doi.org/10.1093/sysbio/syaa101

Orlova, V. F., Solovyeva, E. N., Dunayev, E. A., & Ananjeva, N. B. (2022). Integrative taxonomy within *Eremias multiocellata* complex (Sauria, Lacertidae) from the western part of range: Evidence from historical DNA. *Genes, 13*, 941. https://doi.org/10.3390/genes13060941

Päckert, M., Bader-Blukott, M., Künzelmann, B., Sun, Y.H., Hsu, Y.C., Kehlmaier, C., Albrecht, F., Illera, J.C., Martens, J., (2020a). A revised phylogeny of nuthatches (Aves, Passeriformes, *Sitta*) reveals insight in intra- and interspecific diversification patterns in the Palearctic. *Vertebrate Zoology, 70*(2), 241–262. https://doi.org/10.26049/VZ70-2-2020-10

Päckert, M., Favre, A., Schnitzler, J., Martens, J., Sun, Y. H., Tietze, D. T., Hailer, F., Michalak, I., & Strutzenberger, P. (2020b). Out of Tibet and out of Himalayas: centers of origin and diversification of passerine birds in the Qinghai-Tibet Plateau region. *Ecology and Evolution, 10*(17), 9283–9300. https://doi.org/10.1002/ece3.6615

Päckert, M., Hering, J., Belkacem, A. A., Sun, Y.-H., Hille, S., Lkhagvasuren, D., Islam, S., & Martens, J. (2021). A revised multilocus phylogeny of old world sparrows (Aves: Passeridae). *Vertebrate Zoology*, *71*. https://doi.org/10.3897/VZ.71.E65952

Paetzold, C., Wood, K. R., Eaton, D. A. R., Wagner, W. L., & Appelhans, M. S. (2019). Phylogeny of Hawaiian Melicope (Rutaceae): RAD-seq Resolves Species Relationships and Reveals Ancient Introgression. *Frontiers of Plant Science, 10*, 1074. https://doi.org/10.3389/fpls.2019.01074

Palacios, C., García-R, S., Parra, J. L., Cuervo, A. M., Stiles, F. G., McCormack, J. E., & Cadena, C. D. (2019). Shallow genetic divergence and distinct phenotypic differences between two Andean hummingbirds: Speciation with gene flow? *Auk*, *136*(4), 1–21. https://doi.org/10.1093/auk/ukz046

Pease, J. B., Brown, J. W., Walker, J. F., Hinchliff, C. E., & Smith, S. A. (2018). Quartet Sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany 105*(3): 385–403. https://doi.org/10.1002/ajb2.1016

Peters, J. L., Zhuravlev, Y., Fefelov, I., Logie, A., & Omland, K. E. (2007). Nuclear Loci and Coalescent Methods Support Ancient Hybridization as Cause of Mitochondrial Paraphyly between Gadwall and Falcated Duck (*Anas* spp.). *Evolution*, *61*(8), 1992–2006. https://doi.org/10.1111/j.1558-

5646.2007.00149.x

Podsiadlowski, L., Gamauf, A., & Töpfer, T. (2017). Revising the phylogenetic position of the extinct Mascarene Parrot *Mascarinus mascarin* (Linnaeus 1771) (Aves: Psittaciformes: Psittacidae). *Molecular Phylogenetics and Evolution, 107*, 499–502. https://doi.org/10.1016/j.ympev.2016.12.022

Portenko, L. A., & von Vietinghoff-Scheel. E. (1974). *Montifringilla nivalis* (L.). In: Stresemann E, Portenko LA, Dathe H, Mauersberger G (Eds) Atlas der Verbreitung Palaearktischer Vögel. 4. Lieferung. Akademie Verlag, Berlin.

Prada, C. F., & Boore, J. L. (2019). Gene annotation errors are common in the mammalian mitochondrial genomes database. *BMC Genomics*, *20*, 73. https://doi.org/10.1186/s12864-019-5447-1

Qu, Y., & Lei, F. (2009). Comparative phylogeography of two endemic birds of the Tibetan plateau, the white-rumped snow finch (*Onychostruthus taczanowskii*) and the Hume's ground tit (*Pseudopodoces humilis*). *Molecular Phylogenetics and Evolution, 51*(2), 312–326. https://doi.org/10.1016/j.ympev.2009.01.013

Qu,Y., Ericson, P. G. P., Lei ,F., & Li, S. H. (2005). Postglacial colonization of the Tibetan plateau inferred from the matrilineal genetic structure of the endemic red-necked snow finch, *Pyrgilauda ruficollis*. *Molecular Ecology, 14*(6), 1767–1781. https://doi.org/10.1111/j.1365-294X.2005.02528.x

Qu, Y., Ericson, P. G. P., Lei, F., Gebauer, A., Kaiser, M., & Helbig, A. J. (2006). Molecular phylogenetic relationship of snow finch complex (genera *Montifringilla*, *Pyrgilauda*, and *Onychostruthus*) from the Tibetan plateau. *Molecular Phylogenetics and Evolution*, *40*(1), 218–226. https://doi.org/10.1016/j.ympev.2006.02.020

Qu, Y., Lei, F., Zhang, R., & Lu, X. (2010). Comparative phylogeography of five avian species: implications for Pleistocene evolutionary history in the Qinghai-Tibetan plateau. *Molecular Ecology, 19*(2), 338–351. https://doi.org/10.1111/j.1365-294X.2009.04445.x

Qu, Y., Chen, C., Chen, X., Hao, Y., She, H., Wang, M., Ericson, P. G. P., Lin, H., Cai, T., Song, G., Jia, C., Chen, C., Zhang, H., Li, J., Liang, L., Wu, T., Zhao, J., Gao, Q., Zhang, G., … Lei, F. (2021). The evolution of ancestral and species-specific adaptations in snowfinches at the Qinghai–Tibet Plateau. *Proceedings of the National Academy of Sciences*, *118*(13). https://doi.org/10.1073/pnas.2012398118

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Rambaut, A., & Drummond, A. (2007). *TRACER v1.4. http://beast.bio.ed.ac.uk/Tracer.*

Rambaut A. (2009). *FigTree version 1.2.2 - Computer program distributed by the author, website:* http://tree.bio.ed.ac.uk (last *accessed on 29.11.2010).*

Rasmussen, P. C., & Prŷs-Jones, R. P. (2003). History *vs* mystery: the reliability of museum specimen data. Bulletin of the British Ornithological Club, 123A, 66–94.

Raxworthy, C.J., & Tilston Smith, B. (2021). Mining museums for historical DNA: advances and challenges in museomics. *Trends in Ecology and Evolution, 36*(11), 1049–1060. https://doi.org/10.1016/j.tree.2021.07.009

Rick, J. A., Brock, C. D., Lewanski, A. L., Golcher-Benavides, J., & Wagner, C. E. (2023). Reference Genome Choice and Filtering Thresholds Jointly Influence Phylogenomic Analyses. *Systematic Biology*, syad065. https://doi.org/10.1093/sysbio/syad065

Riordan, P., Cushman, S. A., Mallon, D., Shi, K., & Hughes, J. (2016). Predicting global population connectivity and targeting conservation action for snow leopard across its range. *Ecography, 39*, 419–426. https://doi.org/10.1111/ecog.01691

Ritchie, A. M., Hua, X., & Bromham, L. (2022). Investigating the reliability of molecular estimates of evolutionary time when substitution rates and speciation rates vary. *BMC Ecology and Evolution*, *22*(1), 1–19. https://doi.org/10.1186/s12862-022-02015-8

Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols*, *12*(12). https://doi.org/10.1038/nprot.2017.123

Saidajan, A. (2012). Effects of war on biodiversity and sustainable agricultural development in Afghanistan. *Journal of Developments in Sustainable Agriculture*, *7*(1), 9–13. https://doi.org/10.11178/jdsa.7.9

Salvi, D., Berrilli, E., Garzia, M., & Mariottini, P. (2021). Yet Another Mitochondrial Genome of the Pacific Cupped Oyster: The Published Mitogenome of *Alectryonella plicatula* (Ostreinae) Is Based on a Misidentified *Magallana gigas* (Crassostreinae). *Frontiers in Marine Science*, *8*(September), 8–13. https://doi.org/https://doi.org/10.3389/fmars.2021.741455

Sampaio, F. L., Day, J. J., Mendis Wickramasinghe, L. J., Cyriac, V. P., Papadopoulou, A., Brace, S., Rajendran, A., Simon-Nutbrown, C., Flouris, T., Kapli, P., Ranga Vidanapathirana, D., Kotharambath, R., Kodandaramaiah, U., & Gower, D. J. (2023). A near-complete species-level phylogeny of uropeltid snakes harnessing historical museum collections as a DNA source. *Molecular Phylogenetics and Evolution, 178*, 107651. https://doi.org/10.1016/j.ympev.2022.107651

Sangster, G., & Luksenburg, J. (2020). Chimeric mitochondrial genomes: a hazard for phylogenetics and environmental DNA identification of fishes-chimeric-mitochondrial-genomes-a-hazard-for-phylogenetics-and-environmental-dna-identification-of-fishes. *Authorea*. https://doi.org/10.22541/au.160205226.65255244/v1

Sangster, G., & Luksenburg, J. A. (2021). Scientific data laundering: Chimeric mitogenomes of a sparrowhawk and a nightjar covered-up by forged phylogenies. *Biochemical Systematics and Ecology*, *96*, 104263. https://doi.org/https://doi.org/10.1016/j.bse.2021.104263

Scherz, M. D., Rasolonjatovo, S. M., Köhler, J., Rancilhac, L., Rakotoarison, A., Raselimanana, A. P., Ohler, A., Preick, M., Hofreiter, M., Glaw, F., & Vences, M. (2020). Barcode fishing' for archival DNA from historical type material overcomes taxonomic hurdles, enabling the description of a new frog species. *Scientific Reports*, *10*(1), 19109. https://doi.org/10.1038/s41598-020-75431-9

Schindel, D. E., & Cook, J. A. (2018). The next generation of natural history collections. *PLOS Biology*, *16*(7), 1–8. https://doi.org/10.1371/journal.pbio.2006125

Schmitt, T. (2007). Molecular biogeography of Europe: pleistocene cycles and postglacial trends. *Frontiers in Zoology, 4*, 11. https://doi.org/10.1186/1742-9994-4-11

Schmitt, T., Louy, D., Zimmermann, E., & Habel, J. C. (2016). Species radiation in the Alps: multiple range shifts caused diversification in ringlet butterflies in the European high mountains. *Organisms Diversity and Evolution*, *16*(4), 791–808. https://doi.org/10.1007/s13127-016-0282-6

Schweizer, M., Etzbacher, C., Shirihai, H., Töpfer, T., & Kirwan, G. (2020) A molecular analysis of the mysterious Vaurie's Nightjar *Caprimulgus centralasicus* yields fresh insight into its taxonomic status. Journal of Ornithology, 161, 635-650. https://doi.org/10.1007/s10336-020-01767-8

Scotti-Saintagne, C., Boivin, T., Suez, M., Musch, B., Scotti, I., & Fady, B. (2021). Signature of mid-Pleistocene lineages in the European silver fir (*Abies alba* Mill.) at its geographic distribution margin. *Ecology and Evolution*, *11*(16), 10984–10999. https://doi.org/https://doi.org/10.1002/ece3.7886

Seghal, D., Vikram, P., Sansaloni, C. P., Ortiz, C., Pierre, C. S., Payne, T., Ellis, M., Amri, A., Petroli, C. D., Wenzl, P., & Singh, S. (2015). Exploring and mobilizing the gene bank biodiversity for wheat improvement. *PLoS ONE, 10*(7), e0132112. https://doi.org/10.1371/journal.pone.0132112

Shapiro, B., Sibthorpe, D., Rambaut, A., Austin, J., Wragg, G. M., Bininda-Emonds, O. R. P., Lee, P. L. M., & Cooper, A. (2002). Flight of the dodo. *Science (New York, N.Y.)*, *295*(5560), 1683. https://doi.org/10.1126/science.295.5560.1683

Shapoval, N. A., Yakolev, R. V., & Kuftina, G. N. (2017). The taxonomic status of *Cossus cossus* afghanistanus (Lepidoptera, Cossidae) from Afghanistan: insights from molecular and morphological data. Ukrainian Journal of Ecology, 7(3), 134-138.

She, H., Jiang, Z., Song, G., Ericson, P. G. P., Luo, X., Shao, S., Lei, F., & Qu, Y. (2021). Quantifying adaptive divergence of the snowfinches in a common landscape. *Diversity and Distributions*, *n/a*(n/a), 1–14. https://doi.org/https://doi.org/10.1111/ddi.13383

Skujina, I., McMahon, R., & Hegarty, M. (2017). Are Nuclear/Mitochondrial Sequence Duplications Correctly Characterised in Published Sequence Databases? *Insights in Genetics and Genomics*, *1*, 6.1.

Smallwood, P., Shank, C., Dehgan, A., & Zahler, P. (2011). Wildlife conservation … in Afghanistan? *BioScience, 61*, 506–511. https://doi.org/10.1525/bio.2011.61.7.4

Smith, S. E., Gregory, R. D., Anderson, B. J., & Thomas, C. D. (2013). The past, present and potential future distributions of cold-adapted bird species. *Diversity and Distributions*, *19*(3), 352–362. https://doi.org/https://doi.org/10.1111/ddi.12025

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, *22*(21), 2688–2690. https://doi.org/10.1093/bioinformatics/btl446

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stelbrink, B., Kehlmaier, C., Wilke, T., & Albrecht, C. (2019). The near-complete mitogenome of the critically endangered *Pseudocleopatra dartevellei* (Caenogastropoda: Paludomidae) from the Congo River assembled from historical museum material. *Mitochondrial DNA. Part B, Resources*, *4*(2), 3229–3231. https://doi.org/10.1080/23802359.2019.1669081

Stervander, M., Alström, P., Olsson, U., Ottosson, U., Hansson, B., & Bensch, S. (2016). Multiple instances of paraphyletic species and cryptic taxa revealed by mitochondrial and nuclear RAD data for *Calandrella* larks (Aves: Alaudidae). *Molecular Phylogenetics and Evolution*, *102*, 233–245. https://doi.org/10.1016/j.ympev.2016.05.032

Stervander, M., Illera, J. C., Kvist, L., Barbosa, P., Keehnen, N. P., Pruisscher, P., Bensch, S., & Hansson, B. (2015). Disentangling the complex evolutionary history of the Western Palearctic blue tits (*Cyanistes* spp.) – phylogenomic analyses suggest radiation by multiple colonization events and

subsequent isolation. *Molecular Ecology*, *24*(10), 2477–2494. https://doi.org/https://doi.org/10.1111/mec.13145

Stewart, J. R., Lister, A. M., Barnes, I., & Dalén, L. (2010). Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1682), 661–671. https://doi.org/10.1098/rspb.2009.1272

Strohm, J. H. T., Gwiazdowski, R. A., & Hanner, R. (2016). Mitogenome metadata: current trends and proposed standards. *Mitochondrial DNA. Part A, DNA Mapping, Sequencing, and Analysis*, *27*(5), 3263–3269. https://doi.org/10.3109/19401736.2015.1015003

Summers-Smith, J. D. (2009). Family Passeridae (Old-World Sparrows). In: del Hoyo, J., Elliott, A., & Christie, D. (Eds), Handbook of the Birds of the World, Vol. 14 – Bush-shrikes to Old World Sparrows. Lynx Edicions, Barcelona.

Summers-Smith, J. D., & van Balen, S. (2016). Family Passeridae (Old-World Sparrows). In: del Hoyo, J., Elliott, A., & Christie, D. (Eds), Illustrated Checklist of the Birds of the World, Vol. 2 – Passerines. Lynx Edicions, Barcelona.

Tehseen, M. M., Istipliler, D., Kehel, Z., Sansaloni, C. P., da Silva Lopes, M., Kurtulus, E., Muazzam, S., & Nazari, K. (2021). Genetic diversity and population structure analysis of *Triticum aestivum* L. landrace panel from Afghanistan. *Genes, 12*(3), 340. https://doi.org/10.3390/genes12030340

Terasawa, Y., Kawahara, T., Sasakuma, T., & Sasanuma, T. (2018). Evaluation of the genetic diversity of an Afghan wheat collection based on morphological variation, HMW glutenin subunit polymorphisms, and AFLP. *Breeding Science, 59*, 361–371. https://doi.org/10.1270/jsbbs.59.361

Thorburn, D.-M. J., Sagonas, K., Binzer-Panchal, M., Chain, M. J. J., Feulner, P. G.D., Bornberg-Bauer, E., Reusch, T. B. H., Samonte-Padilla, I. E., Milinski, M., Lenz, T.L., & Eizaguirre, C. (2023). Origin matters: Using a local reference genome improves measures in population genomics. *Molecular Ecology, 23*(7), 1706-1723. https://doi.org/10.1111/1755-0998.13838

Timmins, R. J., Mostafawi, N., Rajabi, A. M., Noori, H., Ostrowski, S., Olsson, U., Svensson, L., & Poole, C. M. (2009). The discovery of Long-billed Reed Warblers *Acrocephalus orinus* in north-eastern Afghanistan. *BirdingASIA, 12*, 42–45.

Timmins, R. J., Ostrowski, S., Mostafawi, N., Noorin, H., Rajabi, A. M., Svensson, L., Olsson, U., & Poole, C. (2010). New information on the Large-billed Reed Warbler *Acrocephalus orinus*, including its song and breeding habitat in north-eastern Afghanistan. *Forktail, 26*, 9–23.

Tolstoj, V. A., & Geipel, E. F. (1990). First record of Theresa's snow finch (*Pyrgilauda theresae*) in Tadjikistan. *Ornitologiya, 24*, 164.

Tribsch, A., & Schönswetter, P. (2003). Patterns of Endemism and Comparative Phylogeography Confirm Palaeoenvironmental Evidence for Pleistocene Refugia in the Eastern Alps. *Taxon*, *52*(3), 477–497. https://doi.org/10.2307/3647447

Tripp, E. A., Tsai, Y. E., Zhuang, Y., & Dexter, K. G. (2017). RADseq dataset with 90% missing data fully resolves recent radiation of *Petalidium* (Acanthaceae) in the ultra-arid deserts of Namibia. *Ecology and Evolution*, *7*(19), 7920–7936. https://doi.org/10.1002/ece3.3274

Tritsch, C., Martens, J., Sun, Y.-H., Heim, W., Strutzenberger, P., & Päckert, M. (2017). Improved sampling at the subspecies level solves a taxonomic dilemma - A case study of two enigmatic Chinese tit species (Aves, Passeriformes, Paridae, *Poecile*). *Molecular Phylogenetics and Evolution*,

107, 538–550. https://doi.org/10.1016/j.ympev.2016.12.014

Uva, V., Päckert, M., Cibois, A., Fumagalli, L., & Roulin, A. (2018). Comprehensive molecular phylogeny of barn owls and relatives (Family: Tytonidae), and their six major Pleistocene radiations. *Molecular Phylogenetics and Evolution*, *125*, 127—137. https://doi.org/10.1016/j.ympev.2018.03.013

Valiente-Mullor, C., Beamud, B., Ansari, I., Francés-Cuesta, C., García-González, N., Mejía, L., Ruiz-Hueso, P., & González-Candelas, F. (2021). One is not enough: On the effects of reference genome for the mapping and subsequent analyses of short-reads. *PLoS Computational Biology*, *17*(1). https://doi.org/10.1371/journal.pcbi.1008678

Vaurie, C. (1956). Systematic notes on Palearctic birds. No 24. Ploceidae: the genera *Passer*, *Petronia* and *Montifringilla*. *American Museum Novitates, 1814*, 1–27. http://hdl.handle.net/2246/5394

Vaurie, C. (1959) The Birds of the Palearctic Fauna – Passeriformes. H. F. & G. Witherby Limited, London.

Vietinghoff-Scheel, E. (1980). *Pyrgilauda theresae* Meinertzhagen. In: Dathe, H., & Neufeldt, I.A. (Eds)*, Atlas der Verbreitung Palaearktischer Vögel, 8. Lieferung.* Akademie Verlag, Berlin.

Vila, M., Marí-Mena, N., Guerrero, A., & Schmitt, T. (2011). Some butterflies do not care much about topography: A single genetic lineage of *Erebia euryale* (Nymphalidae) along the northern Iberian mountains. *Journal of Zoological Systematics and Evolutionary Research*, *49*(2), 119–132. https://doi.org/10.1111/j.1439-0469.2010.00587.x

Wang, K., Lenstra, J. A., Liu, L., Hu, Q., Ma, T., Qiu, Q., & Liu, J. (2018). Incomplete lineage sorting rather than hybridization explains the inconsistent phylogeny of the wisent. *Communications Biology*, *1*, 169. https://doi.org/10.1038/s42003-018-0176-6

Weir, J. T., & Schluter, D. (2007). The Latitudinal Gradient in Recent Speciation and Extinction Rates of Birds and Mammals. *Science*, *315*(5818), 1574 LP – 1576. https://doi.org/10.1126/science.1135590

Westover, K. M., Rusinko, J. P., Hoin, J., & Neal, M. (2013). Rogue taxa phenomenon: A biological companion to simulation analysis. *Molecular Phylogenetics and Evolution, 69*(1), 1–3. https://doi.org/10.1016/j.ympev.2013.05.010

Wikramanayake, E. D., Dinerstein, E., & Loucks, C. J. (2002). Terrestrial ecoregions of the Indo-Pacific: A conservation assessment. Island Press, Washington, DC.

Wingett, S. W., & Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research*, *7*, 1338. https://doi.org/10.12688/f1000research.15931.2

Wolfgramm, H., Martens, J., Töpfer, T., Vamberger, M., Pathak, A., Stuckas, H., & Päckert, M. (2021). Asymmetric allelic introgression across a hybrid zone of the coal tit (*Periparus ater*) in the central Himalayas. *Ecology and Evolution, 11*, 17332–17351. https://doi.org/10.1002/ece3.8369

Wolters, H. E. (1979). Die Vogelarten der Erde. 4. Lieferung. Paul Parey, Hamburg und Berlin.

Yuan, S., Xia, Y., Zheng, Y., & Zeng, X. (2016). Next-generation sequencing of mixed genomic DNA allows efficient assembly of rearranged mitochondrial genomes in *Amolops chunganensis* and *Quasipaa boulengeri*. *PeerJ*, *2016*(12). https://doi.org/10.7717/peerj.2786

Zedane, L., Hong-Wa, C., Murienne, J., Jeziorski, C., Baldwin, B. G., & Besnard, G. (2016). Museomics illuminate the history of an extinct, paleoendemic plant lineage (*Hesperelaea*, Oleaceae) known from

an 1875 collection from Guadalupe Island, Mexico. *Biological Journal of the Linnean Society, 117*, 44–57. https://doi.org/10.1111/bij.12509

Figure captions

Fig 1: Distribution of snowfinches (genera *Montifringilla*, *Pyrgilauda*, *Onychostruthus*) A) entire range across Eurasian mountain systems; B) zoom on the Qinghai-Tibet Plateau (QTP) showing distributions and range overlap for seven sympatric or parapatric species. Ranges of *M. nivalis* modified according to Gebauer et al. (2006); collection sites of samples used for this study indicated by yellow dots.

Fig. 2: Competing phylogenetic hypotheses for snowfinches based on A) a multi-locus data set (Päckert et al., 2020, 2021), B) mitochondrial DNA (Cobos et al., 2021); C-F) Phylogenies of snowfinches inferred from four mitochondrial coding markers (extracted from whole mitogenome sequences; GB= from GenBank) that did not show evidence of heterospecific fragments in the consensus mitogenome sequences (COI, COII, COIII, ND5); the two outgroup taxa (*Passer domesticus* and *Hypocryptadius cinnamomeus*) were pruned from the trees except for COI; G) TCS network for an extended data set of the DNA barcoding marker (COI; n= 53; 483 bp), circles denote haplotypes, dashes indicate substitutions between connected haplotypes, black circles indicate haplotype of COI sequences inferred from whole mitogenome data (sequence/sample ID indicated at corresponding haplotypes).

Fig. 3: Phylogenies of snowfinches based on 13 mitochondrial coding regions; the two outgroup taxa (*Passer domesticus* and *Hypocryptadius cinnamomeus*) were pruned from the trees; Bayesian inference of phylogeny with BEAST v.2.6.2, MCMC chain length: 50,000,000 generations, burnin: 30% of sampled trees; A) uncorrected alignments containing full mitogenome sequences for all species; *Pyrgilauda* is paraphyletic; B) alignments containing missing data for heterospecific markers of the putative chimeric mitogenomes of *P. davidiana* (KJ148632; missing data for ND3, ND4, ND4L, ND6, ATP6, ATP8 and cyt*b* 2nd half) and *O. taczanowskii* (KJ148631; missing data for ND6, ND1 and ND2); *Pyrgilauda* is monophyletic.

Fig. 4: Phylogeny of snowfinches inferred from three sets of SNPs from cleaned ddRAD seq reads mapped on three different reference genomes: *Passer domesticus*, *Pyrgilauda ruficollis* and *Onychostruthus tazcanowskii*; Bayesian inference of phylogeny relied on three combined runs with BEAST for each data set (MCMC chain length: 50 million generations, trees sampled every 5000 generation, burnin : 30% of sampled trees); the figure shows the time-calibrated topology for the data set inferred from mapping to the *O. taczanowskii* genome (unthinned data); node support values from Bayesian posterior probabilities: above nodes: unthinned data, 20% missing (matrix sizes: 30,550 – 40,695 – 39,952 SNPs), below nodes: thinned data (biallelic unlinked SNPs), 30% missing (matrix sizes: 6,263 – 8,949 – 8,893 SNPs); full node support for all three data sets is indicated by an asterisk.

Fig. 5: Phylogenetic relationships of snowfinches inferred from cleaned ddRAD seq reads mapped on the *Passer domesticus* reference genome for two data sets of autosomal SNPs (A, C; unthinned data; n= 30,550 sites; with 20% missing data allowed) and Z-chromosomal SNPs (B, D; unthinned data; n= 899 SNPs; with 30% missing data allowed; due to a disproportionately high amount of missing data *P. theresae* was excluded from this analysis); A, B) tree set analysis from three independent runs with BEAST using DensiTree (all trees: green; consensus trees: blue); C, D) unrooted neighbor networks; lineages of *Montifringilla nivalis*: QTP= Qinghai-Tibet Plateau, CAU= Caucasus, EU= Europe.

Fig. 6: Genetic divergence among snowfinch taxa as reflected by principal component analysis based on thinned data sets inferred from alignment on the *Passer domesticus* reference genome allowing for 20% missing data; scatterplots of PC1 vs PC2, color codes for taxa in PCA (right) correspond to colors of taxon ranges on the maps (left); A) *Montifringilla*; for *M. nivalis* only the ranges of the three subspecies included in the analysis received a color code, whereas the remaining range (other ssp.) was left transparent; B) *Pyrgilauda* and *Onychostruthus*.

Fig. 7: Phylogeny of snowfinches inferred from cleaned ddRAD seq reads mapped on the *Passer domesticus* reference genome (30,550 SNPs); quartet fidelity (QF) scores shown behind terminal taxa; quartet sampling (QS) results with scores for quartet concordance (QC), quartet differential (QD; NA= not applicable) and quartet informativeness (QI) indicated at nodes (scores for total data set above nodes; scores for reduced data set excluding two individuals with lowest QF score (marked in light grey) below nodes; within-clade scores for terminal taxa not shown).

Supplementary figures

Fig. S1: Mismatches among 13 coding markers across individual mitochondrial genomes for pairwise comparisons of seven mitogenome sequences of snowfinches; bar chart shows uncorrected *p*-distances for 13 coding genes (cytochrome-*b* divided in a first and a second fragment) for five pairwise comparisons. As a reference, divergence levels (*p*-distances) between *P. davidiana* and *P. blanfordi* on the one hand and *P. davidiana* and *O. taczanowskii* on the other hand were inferred from COI (n= 53; 483 bp) and cyt*b* data sets (n= 109; 880 bp) and are indicated as thin and thick lines, respectively. Fully correct mitogenomes should range at low p-dist values for pairwise comparisons among mitogenomes of the same species (like for *P. ruficollis*: rufi 423/ GB, all below 0.01) and high p-dist values for pairwise comparison between species (e.g. our *P. ruficollis* vs. our O. *taczanowskii* mitogenome: rufi423 vs. tacz433, between 0.06 and 0.12). Strong deviations at unexpectedly high p-dist values, for example for intraspecific comparison of ND6, ND1 and ND2 from *O. taczanowskii* mitogenomes, suggest that a heterospecific mitochondrial lineage was incorporated into the mitogenome sequence, i.e. the published mitogenome is a chimeric sequence. For phylogenetic reconstructions based on corrected mitogenomes, the heterospecific markers were replaced by missing data ("X" for coding markers; "cds" above bar chart).

Fig. S2: Phylogenies of snowfinches inferred from six mitochondrial coding markers (extracted from whole mitogenome sequences; GB= from GenBank) that showed evidence of heterospecific fragments in the consensus mitogenome sequence of *P. davidiana* marked in red (ATP6, ATP8, ND3, ND4, ND4L, cyt*b*).

Fig. S3: TCS network for an extended data set of cytochrome-*b* (n= 109; 880 bp), circles denote haplotypes, dashes indicate substitutions between connected haplotypes, black circles indicate haplotype of COI sequences inferred from whole mitogenome data (sequence/sample ID indicated at corresponding haplotypes; A) first fragment, B) second fragment; this fragment of the *P. davidiana* mitogenome is identical with the central haplotype of the *O. tazcanowskii* cluster.

Fig. S4: Phylogenies of snowfinches inferred from four mitochondrial coding markers (extracted from whole mitogenome sequences; GB= from GenBank) that showed evidence of heterospecific fragments in the consensus mitogenome sequences of *P. davidiana* and/or *O. taczanowskii* marked in red (A, B, C, D: ND1, ND2, ND6, CR); E: tcs network for an extended data set of the non-coding control region (CR; n= 22, 288 bp), circles denote haplotypes, dashes indicate substitutions between connected haplotypes, black circles indicate haplotype of COI sequences inferred from whole mitogenome data, red circles indicate the two misplaced heterospecific CR haplotypes (sequence/sample ID indicated at corresponding haplotypes).

Fig. S5: Phylogenies of snowfinches inferred from cytochrome-*b* sequences (n= 51); Bayesian inference of phylogeny with BEAST v.2.6.2, MCMC chain length: 50 million generations, burnin: 30% of sampled trees.

Fig. S6: Mapping rates (% successfully mapped reads) from read mapping to three different reference genomes (outgroup: *Passer domesticus*; ingroups: snowfinch species *Pyrgilauda ruficollis* and *Onychostruthus tazcanowskii*).

Fig. S7: Phylogeny of snowfinches inferred from three sets of SNPs from cleaned ddRAD seq reads mapped on three different reference genomes: outgroup *Passer domesticus*; ingroups *Pyrgilauda ruficollis* and *Onychostruthus tazcanowskii*; MCMC chain length; 10 million generations, trees sampled every 1000 generation burnin: 30% of sampled trees; one run with BEAST data sets with different proportions of missing data allowed (10%, 20% and 30%; tree topologies refer to 20% missing data); node support from Bayesian posterior probabilities: above nodes: unthinned data sets; below nodes: thinned data sets (biallelic unlinked SNPs); full node support for all three data sets (BI= 1, ML= 100) is indicated by an asterisk.

Fig. S8: Maximum Likelihood phylogeny of snowfinches inferred from three sets of SNPs from cleaned ddRAD seq reads mapped on three different reference genomes (20% missing data allowed).

Fig. S9: PCAs for samples of small-sized species of *Pyrgilauda* and *Onychostruthus*; filtered and thinned data, 0% missing: without *P. theresae*: 1501 unlinked SNPs; 20% missing, with and without *P. theresae*: 175 and 3233 unlinked SNPs, respectively; scatterplots of PC1 versus PC2, PC3 and PC4.

Fig. S10: PCAs for samples of large-sized *Montifringilla* species; filtered and thinned data, 0% missing: 2500 unlinked SNPs, 20% missing: 3251 unlinked SNPs, scatterplots of PC1 versus PC2, PC3 and PC4, respectively

Table 1: Classification of snowfinches according to different taxonomic authorities (subspecific taxa, only those listed treated in this study).

	Hartert (1910)	**Hartert & Steinbacher (1932)**	**Vaurie (1956), Mayr & Greenway Jr (1962), Wolters (1979)**	**Summers-Smith (2009)**	**Dickinson (2003), Dickinson & Christidis (2014), Summers-Smith & van Balen (2016), Gill et al. (2022)**
genera	1	1	1	2	3
species	5	6	7	8	8
White-winged Snowfinch	*Montifringilla nivalis* *Montifringilla n. nivalis* *Montifringilla n. alpicola* *Montifringilla n. groumgrzimaili* *Montifringilla n. henrici* *Montifringilla n. adamsi*	*Montifringilla nivalis* *Montifringilla n. nivalis* *Montifringilla n. alpicola* *Montifringilla n. groumgrzimaili* *Montifringilla n. henrici*	*Montifringilla nivalis* *Montifringilla n. nivalis* *Montifringilla n. alpicola* *Montifringilla n. groumgrzimaili* *Montifringilla n. henrici*	*Montifringilla nivalis* *Montifringilla n. nivalis* *Montifringilla n. alpicola* *Montifringilla n. groumgrzimaili*	*Montifringilla nivalis* *Montifringilla n. nivalis* *Montifringilla n. alpicola* *Montifringilla n. groumgrzimaili*
Tibetan Snowfinch				*Montifringilla henrici*	*Montifringilla henrici*
Black-winged Snowfinch		*Montifringilla adamsi*	*Montifringilla adamsi*	*Montifringilla adamsi*	*Montifringilla adamsi*

Plain-backed Snowfinch	*Montifringilla blanfordi*	*Montifringilla blanfordi*	*Montifringilla blanfordi*	*Pyrgilauda blanfordi*	*Pyrgilauda blanfordi*
Small snowfinch	*Montifringilla davidiana*	*Montifringilla davidiana*	*Montifringilla davidiana*	*Pyrgilauda davidiana*	*Pyrgilauda davidiana*
Rufous-necked Snowfinch	*Montifringilla ruficollis*	*Montifringilla ruficollis*	*Montifringilla ruficollis*	*Pyrgilauda ruficollis*	*Pyrgilauda ruficollis*
Afghan Snowfinch	-	-	*Montifringilla theresae*	*Pyrgilauda theresae*	*Pyrgilauda theresae*
White-rumped Snowfinch	*Montifringilla mandelli*	*Montifringilla taczanowskii*	*Montifringilla taczanowskii*	*Pyrgilauda taczanowskii*	*Onychostruthus taczanowskii*

Table 2: Model settings as inferred from MrModeltest applied to data sets of genome-wide SNP and of mitochondrial markers for inference of phylogeny using BEAST 2; reference genomes (alignment length and model settings for unthinned SNP-data sets with 20% missing data allowed): ref. 1= *Onychostruthus taczanowskii*; ref. 2= *Pyrgilauda blanfordi*; *Passer*= *Passer domesticus*; Rate(mito): fixed mean rates of mitochondrial markers according to Lerner et al. (2011) assigned to each of the thirteen coding mtDNA markers.

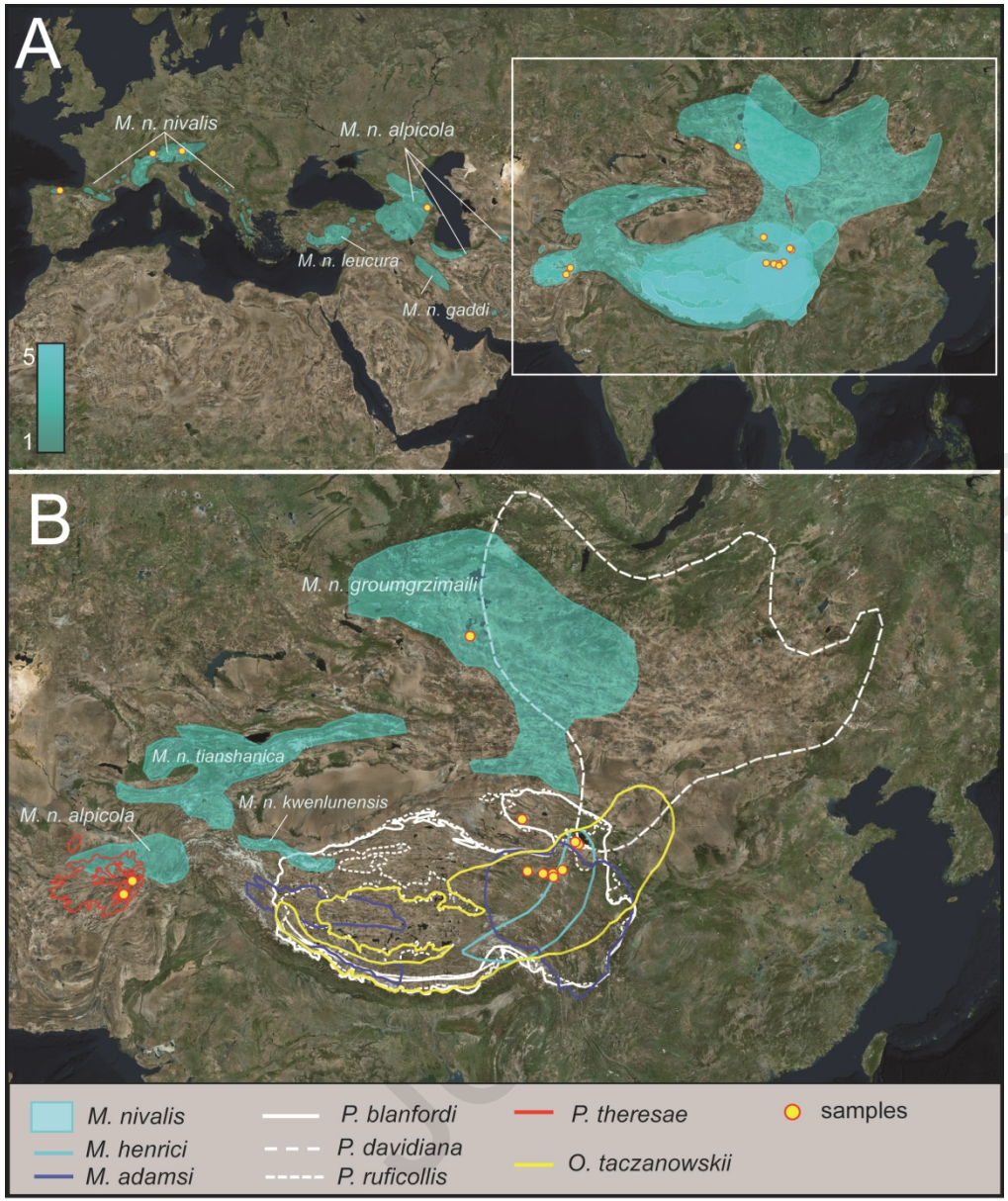
	genome-wide SNPs		**SNPs autosomes**	**SNPS Z-chrom.**	**Mitogenome (13 coding markers)**											
marker	**ref. 1**	**ref. 2**	**ref. *Passer***	**ref. *Passer***	**ND1**	**ND2**	**COI**	**COII**	**ATP8**	**ATP6**	**COIII**	**ND3**	**ND4L**	**ND4**	**ND5**	**Cytb**	**ND6**
bp	39952	40695	30550	899	978	1040	1551	685	168	684	784	351	297	1378	1817	1144	519
model	GTR	GTR	GTR	HKY	GTR + I + Γ	GTR + Γ	GTR + Γ	GTR + Γ	GTR + Γ	GTR + Γ	GTR + Γ	GTR + I + Γ	GTR + Γ	GTR + I + Γ	GTR + I + Γ	GTR + Γ	HKY + I + Γ
πA	0.1856	0.1846	0.3255	0.2353	0.3035	0.3170	0.2842	0.3255	0.3175	0.3366	0.3021	0.3202	0.2828	0.3309	0.3385	0.3044	0.4025
πC	0.3152	0.3166	0.3393	0.2692	0.3543	0.3668	0.2923	0.3393	0.3887	0.4039	0.3830	0.3687	0.3828	0.3895	0.3836	0.3787	0.3832
πT	0.3147	0.3146	0.1302	0.2845	0.1287	0.1064	0.1409	0.1302	0.0662	0.0783	0.1164	0.0979	0.1204	0.0877	0.0990	0.1237	0.0896

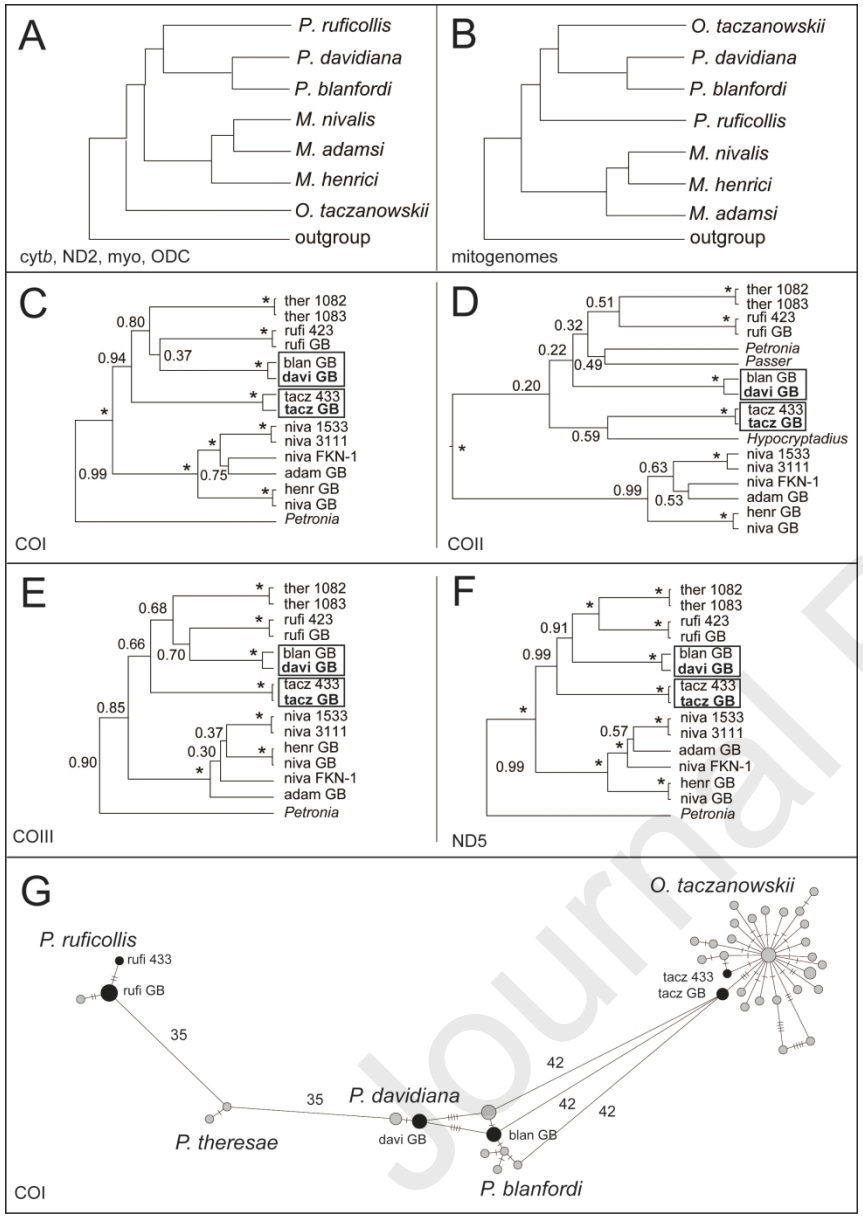
πG	0.1846	0.1841	0.2050	0.2110	0.2135	0.2097	0.2826	0.2050	0.2276	0.1812	0.1985	0.2132	0.2140	0.1918	0.1789	0.1932	0.1247
α	-	-	-	-	1.6711	1.1969	0.2795	0.1363	0.2408	0.1374	0.0949	5.1547	0.1310	15.149	2.3496	0.1491	1.0221
I	0	0	0	0	0.5829	0	0	0	0	0	0	0.5879	0	0.5775	0.5785	0	0.5303
Ti/Tv ratio	-	-	-	2.6541	-	-	-	-	-	-	-	-	-	-	-	-	13.8496
R(a)[A-C]	1.1094	0.9889	1.0733	-	2.2496	0.6892	1.6620	0.3604	163478.45	0.08066	0.6243	1289.5616	0.7773	1.1767	1.0733	1408.1095	-
R(b)[A-G]	5.1571	5.0965	5.4676	-	71.8733	30.9663	7.1987	10.0887	5373712.00	57.2492	24.7197	36274.0742	15.1338	31.9667	30.8831	11189.6514	-
R(c)[A-T]	1.2321	1.2493	1.0343	-	1.7034	0.6659	2.0029	1.2930	107937.78	1.9787	0.7366	2030.5664	1.0047	0.8836	2.0316	1531.7542	-
R(d)[C-G]	1.0337	1.0175	1.2287	-	1.2602	0.3265	1.0715	0.0000	0.0000	0.4327	0.4522	583.7142	0.0000	0.0000	0.5698	311.8066	-
R(e)[C-T]	5.1723	5.0834	5.4201	-	36.9278	12.9552	10.0005	11.4824	1093238.88	21.1377	14.7202	16167.6572	10.8711	15.2253	23.3920	15823.3740	-
Rate (mito)					0.025	0.029	0.016	0.019	0.019	0.026	0.019	0.024	0.025	0.022	0.021	0.014	0.024

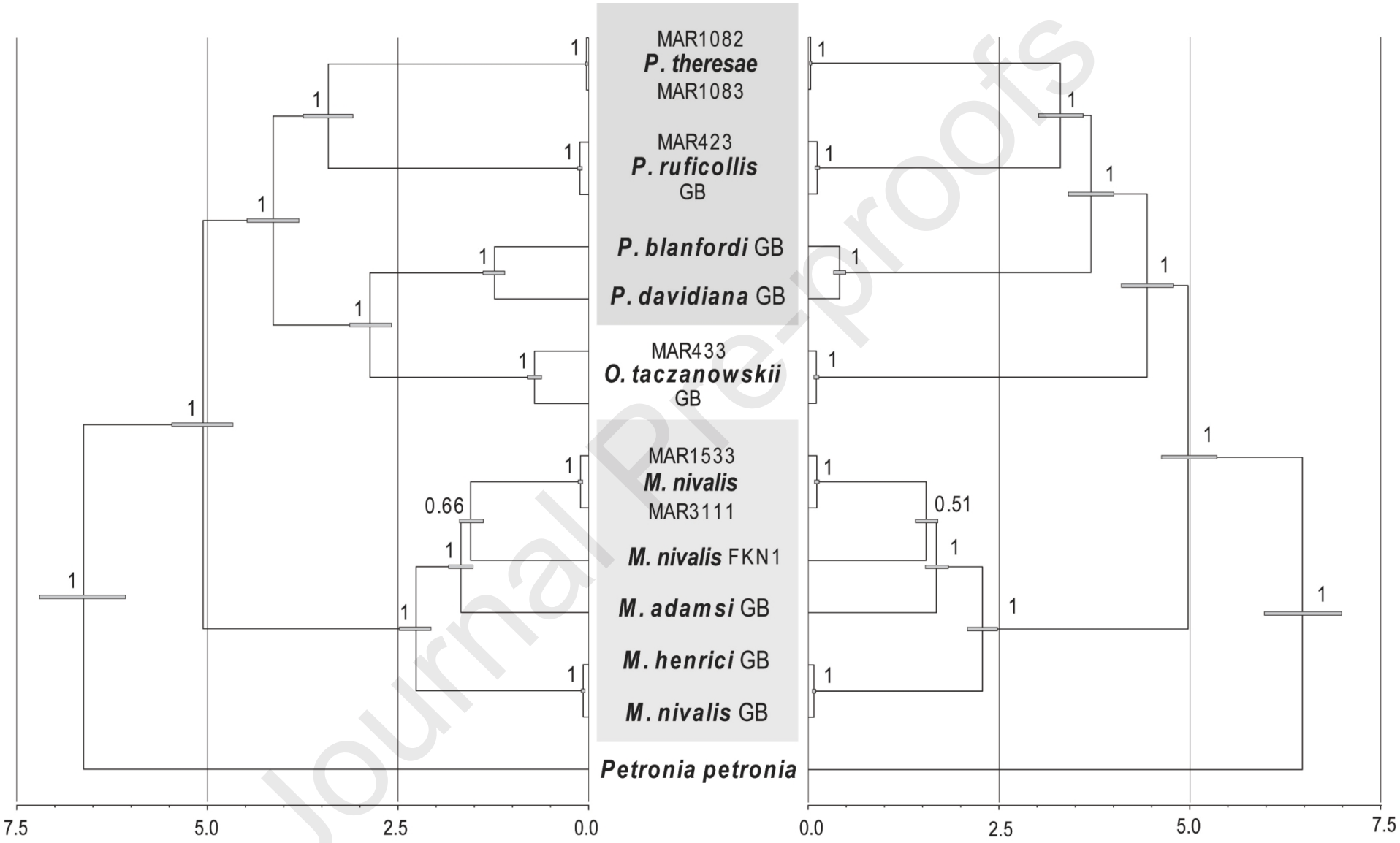
Table 3: Intra- and interspecific genetic differentiation and distance among snowfinch taxa; F_{ST} values inferred from genome-wide autosomal SNPs (upper diagonal; reference genome *Passer domesticus*; 20% missing data allowed; n= 30,550 sites); p-distance values inferred from 1144 bp of the mitochondrial cytochrome-*b* (lower diagonal).

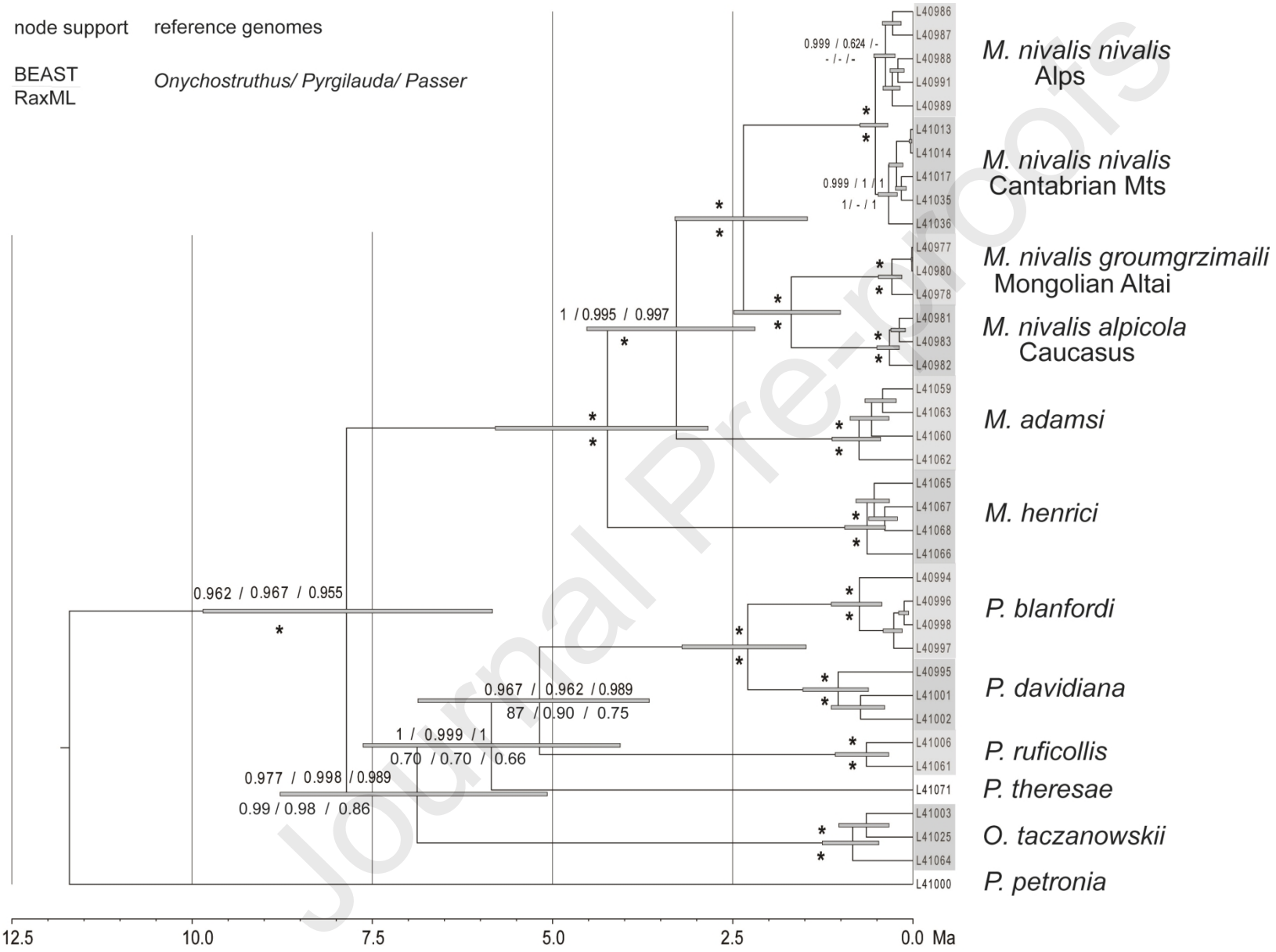
	M. n. nivalis Cant.	*M. n. nivalis* Alps	*M. n. alpicola*	*M. n. groumgrzimaili*	*M. adamsi*	*M. henrici*	*P. davidiana*	*P. blanfordi*	*P. ruficollis*	*P. theresae*	*O. tazcanowskii*	*P. petronia*
M. n. nivalis Cantabrian Mts	-	0.0873355	0.421948	0.475532	0.410394	0.468640	0.625886	0.534988	0.617837	0.775211	0.646510	0.816053
M. n. nivalis Alps	0.0021	-	0.386614	0.436230	0.386671	0.443385	0.601708	0.516821	0.590741	0.755804	0.624559	0.797324
M. n. alpicola	0.0442	0.0461	-	0.337359	0.338370	0.394863	0.568617	0.480926	0.559353	0.729820	0.596060	0.781067
M. n. groumgrzimaili	0.0466	0.0485	0.0043	-	0.366867	0.430914	0.606667	0.506102	0.596076	0.722274	0.625105	0.810012
M. adamsi	0.0400	0.0405	0.0411	0.0431	-	0.340002	0.511840	0.448769	0.500629	0.629115	0.546497	0.715263
M. henrici	0.0561	0.0563	0.0597	0.0602	0.0470	-	0.520177	0.451539	0.508726	0.605885	0.552999	0.730608
P. davidiana	0.0776	0.0795	0.0694	0.0713	0.0696	0.0810	-	0.18083	0.458616	0.659060	0.552037	0.753774
P. blanfordi	0.0784	0.0803	0.0704	0.0724	0.0716	0.0818	0.0045	-	0.376991	0.454128	0.485414	0.669589
P. ruficollis	0.0907	0.0914	0.0892	0.0908	0.0812	0.0926	0.0749	0.0761	-	0.559843	0.539417	0.746129

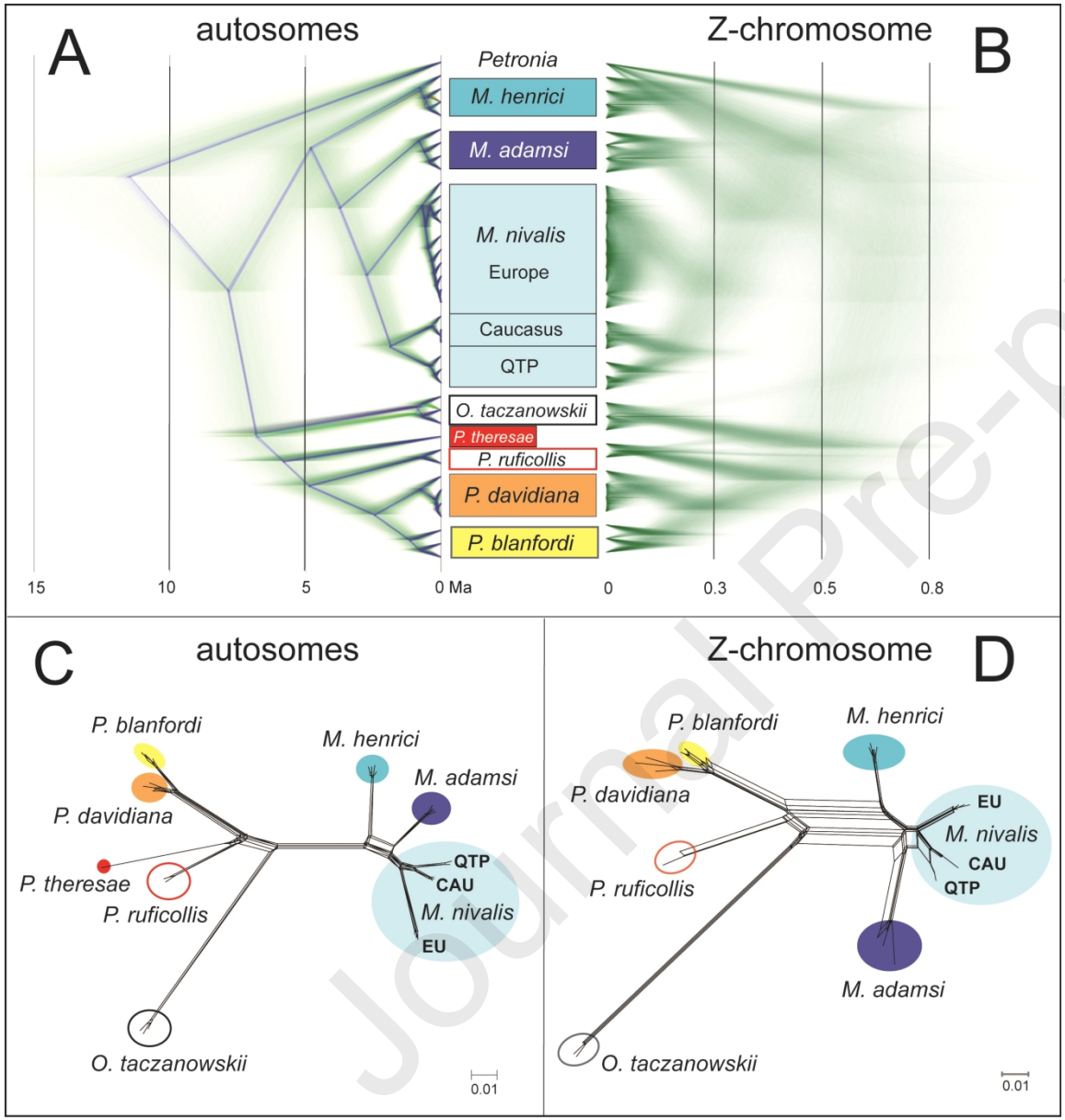
P. theresae	0.0832	0.0849	0.0818	0.0817	0.0710	0.0873	0.0576	0.0586	0.0675	-	0.612398	0.875000
O. tazcanowskii	0.0954	0.0962	0.0978	0.1009	0.0867	0.0977	0.0864	0.0887	0.0978	0.0889	-	0.737817

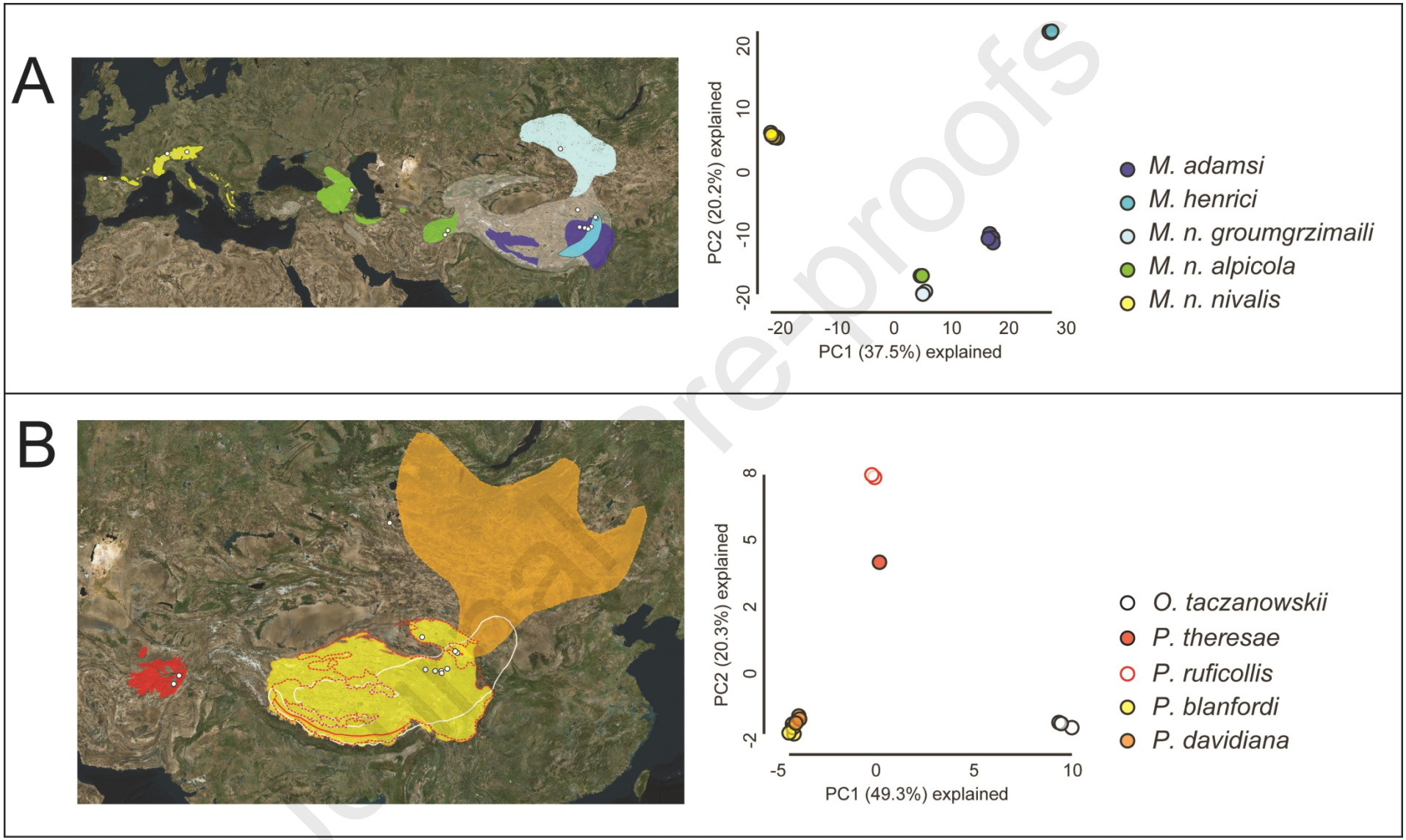

A
M. n. nivalis
M. n. alpicola
M. n. leucura
M. n. gaddi
5
1
B
M. n. groumgrzimaili
M. n. tianshanica
M. n. kwenlunensis
M. n. alpicola
M. nivalis
M. henrici
M. adamsi
P. blanfordi
P. davidiana
P. ruficollis
P. theresae
O. taczanowskii
samples

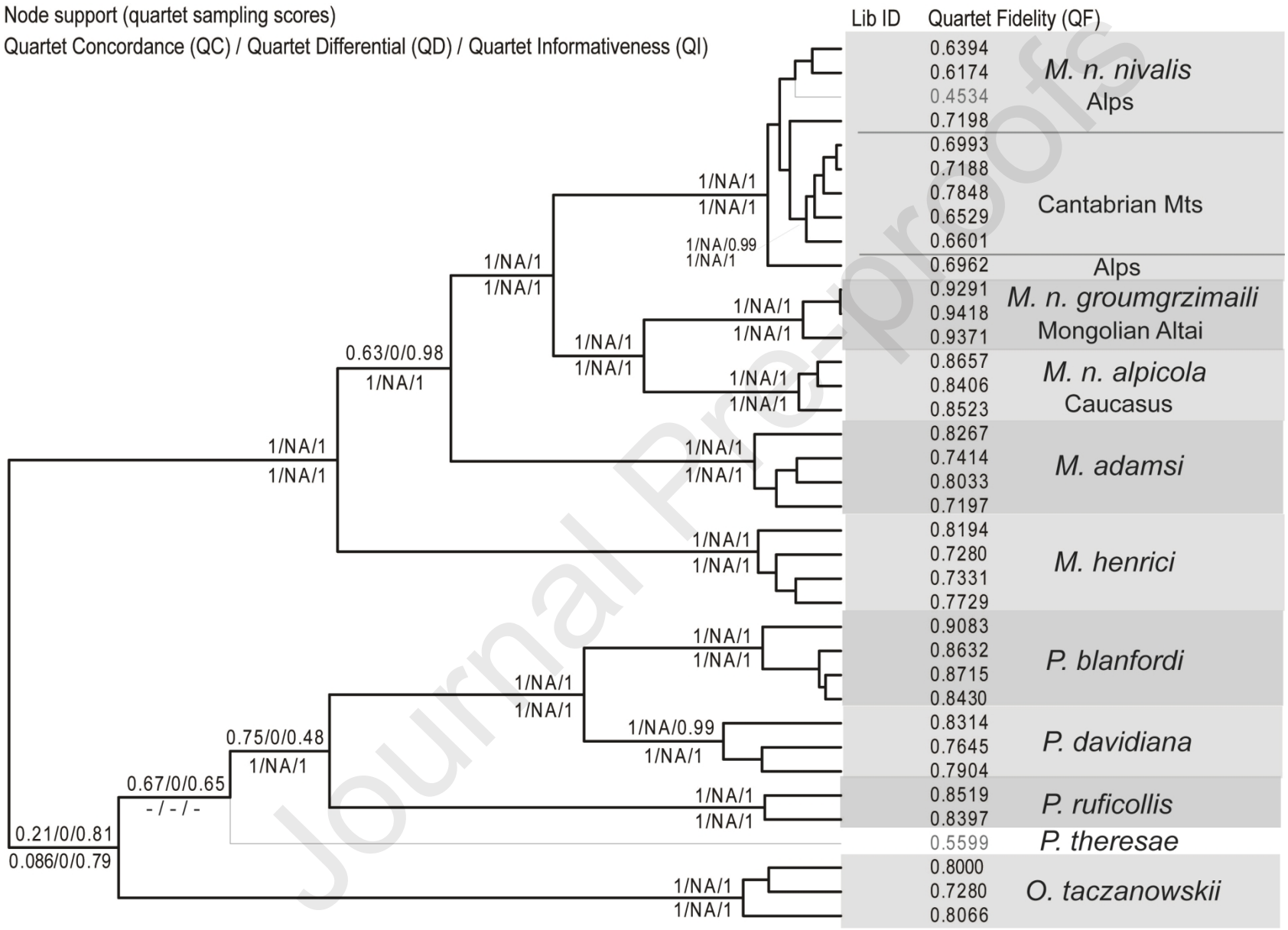


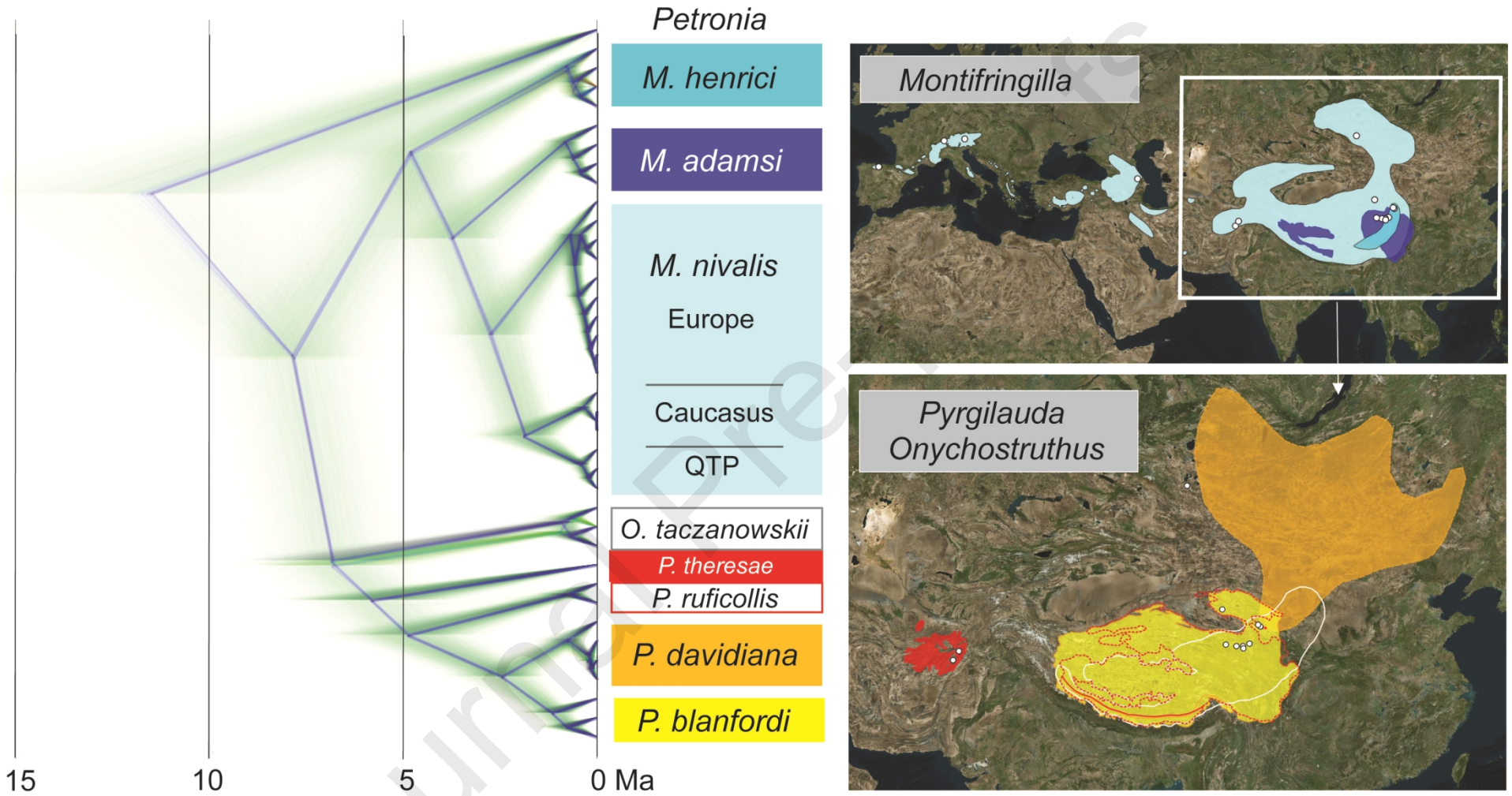












Highlights

1. Previously published mitogenomes included heterospecific sequence information.

2. Previous phylogenies suffered from further sources of error such as flawed taxonomy.

3. Genera *Montifringilla* and *Pyrgilauda* are reciprocally monophyletic.

4. Afghan *Pyrgilauda theresae* originated from a Pliocene out-of-Tibet dispersal.

5. European and Asian lineages of *M. nivalis* separated roughly 2 million years ago.

Declaration of interest: None

CRediT authorship contribution statement:

Safiqul Islam: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, writing – Original draft, review and editing, Visualization. **Claire Peart**: Methodology, Formal analysis, Writing – review and editing. **Christian Kehlmaier**: Validation, Formal analysis, Investigation, Data curation, Writing – Review and editing. **Andreas Dahl, Sylvia Klemroth:** Methodology, Validation, Investigation, Writing – Review and editing. **Yue-Hua Sun, Fumin Lei, Maria del Mar Delgado, Paola Laiolo, Juan Carlos Illera, Sebastian Dirren, Sabine Hille, Davaa Lkhagvasuren, Till Töpfer, Martin Kaiser, Axel Gebauer**: Resources, Writing – review and editing. **Dimitra Alexopoulou**: Formal analysis, Writing – review and editing. **Claudia Paetzold**: Software, Validation, Formal analysis, Writing – Original draft, review and editing. **Jochen Martens**: Conceptualization, Resources, Funding acquisition, Writing – review and editing. **Martin Päckert**: Conceptualization, Investigation, Data curation, writing – Original draft, review and editing, Visualization, Supervision, Project administration, Funding acquisition