

Endangered Steppe Eagle (*Aquila nipalensis*) (Aves, Accipitriformes, Accipitridae) genome and mitogenome assembly: A resource for molecular evolution and comparative genomics

Wannapol Buthasane¹, Thidathip Wongsurawat², Piroon Jenjaroenpun², Sithichoke Tangphatsornruang³, Wirulda Pootakham³, Chutima Sonthirod³, Wasitthee Kongkachana³, Alisa Wilantho³, Ratiwan Sitdhibutr⁴, Chaian Kasorndorkbua⁵, Gunnaporn Suriyaphol¹

¹ Biochemistry Unit, Department of Physiology, Faculty of Veterinary Science, Chulalongkorn University, Bangkok 10330, Thailand

² Division of Medical Bioinformatics, Research Department, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand

³ National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency, Pathum Thani 12120, Thailand

⁴ Raptor Rehabilitation Unit, Kasetsart University Veterinary Teaching Hospital Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand

⁵ Laboratory of Raptor Research and Conservation Medicine, Department of Pathology, Faculty of Veterinary Medicine, Kasetsart University, Bangkok, 10900, Thailand

Corresponding author: Gunnaporn Suriyaphol (gunnaporn.v@chula.ac.th)



Academic editor: George Sangster

Received: 20 May 2025

Accepted: 11 May 2026

Published: 3 June 2026

ZooBank: <https://zoobank.org/AAFE230E-8723-4789-BAAC-95CEF94BC24A>

Citation: Buthasane W, Wongsurawat T, Jenjaroenpun P, Tangphatsornruang S, Pootakham W, Sonthirod C, Kongkachana W, Wilantho A, Sitdhibutr R, Kasorndorkbua C, Suriyaphol G (2026) Endangered Steppe Eagle (*Aquila nipalensis*) (Aves, Accipitriformes, Accipitridae) genome and mitogenome assembly: A resource for molecular evolution and comparative genomics. ZooKeys 1281: 105–122. <https://doi.org/10.3897/zookeys.1281.158566>

Copyright: © Wannapol Buthasane et al.
This is an open access article distributed under terms of the Creative Commons Attribution License ([Attribution 4.0 International – CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

Abstract

The Steppe Eagle, *Aquila nipalensis* Hodgson, 1833, is a migratory, endangered raptor experiencing population declines due to habitat loss and human persecution. This study aims to construct a high-quality *de novo* nuclear genome and complete mitochondrial genome assembly using PromethION Oxford Nanopore long-read and MGI short-read sequencing technologies. The assembled genome size was 1.21 Gb and contained 16,192 predicted protein-coding genes. Phylogenomic reconstruction based on ultra-conserved elements (UCEs) robustly placed *A. nipalensis* within Aquilinae, with *Aquila chrysaetos* (Linnaeus, 1758) identified as its closest extant relative. Mitogenome-based analyses recovered congruent topology but revealed dataset-dependent differences in divergence time estimation. The most recent common ancestor of *A. nipalensis* and other *Aquila* species was estimated at approximately 3–13 million years ago, depending on dataset, whereas divergence between *Aquila* and *Nisaetus* occurred around 22 Mya. Comparative genomic analyses further identified positively selected genes associated with vesicle trafficking, secretion, and tissue development, suggesting potential adaptive signatures related to physiological performance. In conclusion, these genomic and evolutionary insights establish a foundational reference for future population genomic, adaptive, and conservation studies of this endangered raptor.

Key words: Comparative analysis, *de novo* genome assembly, evolution

Introduction

The Steppe Eagle (*Aquila nipalensis* Hodgson, 1833) is a large-bodied eagle belonging to the family Accipitridae. *Aquila nipalensis* is a long-distance migratory raptor widely distributed across the Palaearctic steppe. It breeds primarily in open grasslands of Eurasia, from southern Russia and Kazakhstan eastward through Mongolia to northern China, and winters across Africa, the Middle East,

and parts of South and Southeast Asia. This extensive migratory system links breeding and wintering habitats across thousands of kilometres and exposes the species to diverse ecological pressures throughout its range (Bari et al. 2020). The species has been recorded only infrequently in Thailand during the winter season. This species is classified as Endangered in the International Union for Conservation of Nature (IUCN) Red List (BirdLife International 2020). Population declines have been attributed to the conversion of steppe habitats into agricultural land, direct persecution, and electrocution from electric lines (BirdLife International 2020). The Steppe Eagle, a diurnal raptor species and primarily preys on small mammals, birds, and reptiles, but it also scavenges on carrion (Orihuela-Torres et al. 2021). Its scavenging behaviour increases its risk of exposure to chemical contaminants, similar to vultures, with cases of diclofenac toxicity reported as contributing to population declines (Sharma et al. 2014). However, the genetic aspects related to those functions have not yet been reported in the Steppe Eagle.

Genomic research provides extensive biological information, encompassing genes associated with molecular functions, biological processes, and cellular components. It also offers valuable insights into evolutionary relationships among species (Zhang et al. 2014b; Stiller et al. 2024). Within the family Accipitridae, reference genomes have been assembled for several species such as Golden Eagle (*Aquila chrysaetos* (Linnaeus, 1758)), Bald Eagle (*Haliaeetus leucocephalus* (Linnaeus, 1766)), a White-tailed Eagle (*Haliaeetus albicilla* (Linnaeus, 1758)) and Asian King Vulture (*Sarcogyps calvus* (Scopoli, 1786)) (Zhang et al. 2014a; Mead et al. 2021; Buthasane et al. 2024). *Aquila nipalensis* belongs to the booted eagles (subfamily Aquilinae), a lineage of Accipitridae characterised by feathered tarsi and complex evolutionary relationships revealed by molecular phylogenetic studies. Early multigene studies demonstrated that traditional eagle genera such as *Aquila* Brisson, 1760, *Hieraaetus* Kaup, 1844, and *Spizaetus* Vieillot, 1816 are not strictly monophyletic, indicating that morphological classifications had obscured true evolutionary relationships. Subsequent molecular phylogenies confirmed the monophyly of the booted eagle clade while revealing deep subdivisions and taxonomic restructuring within Aquilinae. More recent comprehensive analyses based on multiple genetic loci support several major clades within the group and refined the taxonomy of booted eagles, clarifying relationships among species in *Aquila* and related genera (Helbig et al. 2005; Lerner and Mindell 2005; Lerner et al. 2017).

However, a comprehensive genomic analysis of Steppe Eagle and its comparative genomic context have not yet been reported. In addition to nuclear genomes, mitochondrial genomes (mitogenomes) are widely studied in wildlife conservation and species identification. Moreover, mitogenomes are extensively used to estimate divergence times, offering crucial insights into the molecular evolution of various avian species, such as Golden Eagle, Eurasian Goshawk (*Astur gentilis* (Linnaeus, 1758)), Red Junglefowl (*Gallus gallus* (Linnaeus, 1758)), Osprey (*Pandion haliaetus* (Linnaeus, 1758)), and Secretary Bird (*Sagittarius serpentarius* (Miller, 1779)) (Arcones et al. 2021). Understanding divergence time estimation is fundamental to unravelling the historical relationships among species, reconstructing evolutionary lineages, and assessing the accumulation of genetic variation over time (Nabholz et al. 2009; Eo and DeWoody 2010). The objective of the present study was to assemble a

reference genome and mitogenome of Steppe Eagle by integrating long-read and short-read sequencing technologies. Additionally, we investigated the species' unique genetic features, its evolutionary relationships with related species, and estimated divergence times. The findings provide essential baseline data to support future population genomic studies, including breeder selection and efforts toward population restoration.

Materials and methods

Genome assembly, gene prediction, and annotation

Sample collection was conducted in accordance with ethical guidelines approved by the Institutional Animal Care and Use Committee of Kasetsart University, Thailand (approval no. ACKU64-VET-052). Genomic DNA was extracted from a whole-blood sample preserved in 70% ethanol, obtained from a male Steppe Eagle housed at the Raptor Rehabilitation Unit, Kasetsart University. High molecular weight (HMW) DNA was isolated using the Wizard HMW DNA Extraction Kit (Promega, Madison, WI, USA). DNA concentration was quantified using a NanoDrop One Microvolume UV-Vis Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and DNA quality and integrity were assessed via pulsed-field gel electrophoresis (PFGE). Sequencing was conducted using both long-read and short-read platforms to obtain comprehensive genomic data. ONT long-read sequencing was performed on a PromethION 2 Integrated (P2i) platform (Oxford Nanopore Technologies, Oxford, UK) using two PromethION flow cells (R10.4.1). HMW genomic DNA was used for library preparation with the ONT Ligation Sequencing Kit (SQK-LSK114; Oxford Nanopore Technologies) following the manufacturer's protocol. Basecalling was performed using Guppy (Oxford Nanopore Technologies), and the resulting reads were exported in FASTQ format for downstream analyses. Short-read paired-end sequencing was carried out on the MGISEQ-2000 platform (MGI Tech, Shenzhen, China). Raw ONT reads were *de novo* assembled using Flye v. 2.9.1-b1780 with the '--meta' option (Kolmogorov et al. 2019). The draft assembly was polished in two stages: (1) long-read polishing was performed using Racon v. 1.5.0 for two rounds with parameters -m 8 -x -6 -g -8 -w 500 (<https://github.com/isovic/racon>), followed by Medaka v. 1.6.1 using the r941_min_sup_g507 model (<https://github.com/nanoporetech/medaka>); and (2) short-read polishing was conducted with Pilon v. 1.24 (Walker et al. 2014), iterating until no further improvements were detected. Genome completeness was evaluated using BUSCO v. 4.0.5 with the avian single-copy ortholog gene set ($n = 8,338$) (Manni et al. 2021).

For homology-based gene prediction, protein sequences from four avian species—*G. gallus*, *Meleagris gallopavo* Linnaeus, 1758, *Taeniopygia guttata* Vieillot, 1817, and *Columba livia* Gmelin, 1789—retrieved from the UniProtKB database (<https://www.uniprot.org>) were aligned to the assembled genome using Exonerate within the MAKER annotation pipeline (Holt and Yandell 2011). The coding sequences of *A. chrysaetos chrysaetos* (Accession [GCA_900496995.4](https://www.ncbi.nlm.nih.gov/nuccore/GCA_900496995.4)) were applied as expressed sequence tag (EST) evidence in the MAKER pipeline. *Ab initio* predictions were generated using AUGUSTUS v. 3.3.3 (Augustus: Gene Prediction, RRID:SCR_008417), with parameters trained on *G. gallus*

genes. Additionally, SNAP was employed for another round of *ab initio* prediction, with training performed using gene models derived from the initial round of MAKER annotation (Korf 2004). Gene filtering criteria included a minimum coding sequence length of 300 bp and a maximum intron length of 10 kb. Gene models generated from SNAP and AUGUSTUS predictions were retained only if they exhibited an Annotation Edit Distance (AED) score ≤ 0.25 .

Phylogenetic analysis using ultraconserved elements (UCEs)

Ultraconserved elements (UCEs) were extracted using PHYLUCE v. 1.6.8. The 5K UCE probe set targeting approximately 5,060 loci was downloaded from the Faircloth Lab UCE probe repository. To assess the impact of missing data, complete data matrices were generated and filtered at 75%, 95%, and 100% occupancy thresholds. Following sequence alignment and trimming, the best-fit substitution model for each filtered dataset was determined using ModelTest-NG v. 0.1.7 under the best-fit substitution models. Nodal support was assessed using 1,000 nonparametric bootstrap replicates. Phylogenetic relationships were inferred among 26 species representing the orders Accipitriformes and Cathartiformes. Details of the species included in the analysis, along with their accession numbers, are provided in Table 1. *Gallus gallus* was designated as the outgroup.

Mitogenome assemblies, annotations, and phylogeny inference

The mitogenome was sequenced using the MGISEQ-2000 platform (MGI Tech, Shenzhen, China) and assembled with NOVOPlasty v. 3.8.2 (Dierckxsens et al. 2017). Annotation was carried out using the MITOS WebServer (Bernt et al. 2013). Protein-coding, rRNA and tRNA genes were further identified using the NCBI BLAST (Altschul et al. 1990). The circular structure of the mitogenome was visualized using OrganellarGenomeDRAW (OGDRAW) v. 1.3.122 (Greiner et al. 2019). Analyses of nucleotide and amino acid composition were conducted using MEGA X (Kumar et al. 2018).

Two mitochondrial DNA datasets were generated for phylogenetic tree analyses. In the first dataset, the complete mitogenome of *A. nipalensis* was compared with 13 previously published mitogenomes representing diverse avian orders, including Accipitriformes, Cathartiformes and Galliformes. Maximum-likelihood (ML) phylogenies were reconstructed using the concatenated nucleotide sequences of the 13 conserved mitochondrial protein-coding genes (PCGs): cytochrome b (*CYTB*); NADH dehydrogenase subunits 1 (*ND1*), *ND2*, *ND3*, *ND4*, *ND4L*, *ND5*, and *ND6*; cytochrome c oxidase subunits 1 (*COX1*), *COX2*, and *COX3*; and ATP synthase F0 subunit 6 (*ATP6*) and *ATP8*. The newly sequenced mitogenome of *A. nipalensis* were included in both phylogenetic reconstruction and divergence time estimation. Details of the species included in the analyses, along with their GenBank accession numbers, are provided in Table 1.

Multiple sequence alignments were performed using MAFFT v. 7.505, and poorly aligned regions were trimmed with trimAl v. 1.5.rev1. The best-fit nucleotide substitution models were selected using ModelFinder: GTR+F+I+G4 for the concatenated PCG dataset. ML phylogenetic trees were inferred using IQ-TREE v. 2.4.0 with 1,000 ultrafast bootstrap replicates (Nguyen et al. 2014; Hoang

Table 1. Species included in the phylogenetic analyses with their GenBank accession numbers.

Scientific name	GenBank accession number	
	Genome assembly	Mitogenome
<i>Accipiter nisus</i>	GCA_027497715.1	—
<i>Tachyspiza soloensis</i>	GCA_025727755.1	—
<i>Tachyspiza virgata</i>	GCA_025728005.1	—
<i>Aegypius monachus</i>	GCA_025309335.1	—
<i>Aquila chrysaetos</i>	GCA_900496995.4	NC_024087.1
<i>Aquila fasciata</i>	—	NC_029188.1
<i>Aquila nipalensis</i>	Present study	Present study NC_045042.1
<i>Astur gentilis</i>	GCA_929443795.1	NC_011818.1
<i>Butastur indicus</i>	GCA_026770005.1	—
<i>Butastur liventer</i>	GCA_025447895.1	—
<i>Buteo buteo</i>	GCA_025504755.1	—
<i>Buteo hemilasius</i>	GCA_025727635.1	—
<i>Buteo lagopus</i>	GCA_025727815.1	—
<i>Cathartes aura</i>	GCA_000699945.1	—
<i>Circaetus pectoralis</i>	GCA_013399555.1	NC_052805.1
<i>Gallus gallus</i>	GCA_016699485.1	MG605671
<i>Gymnogyps californianus</i>	GCA_018139145.2	BK059163.1
<i>Gyps fulvus</i>	GCA_025727715.1	—
<i>Gyps himalayensis</i>	GCA_021398385.1	—
<i>Haliaeetus albicilla</i>	GCA_000691405.1	NC_040858.1
<i>Haliastur indus</i>	GCA_027575065.1	—
<i>Milvus migrans</i>	GCA_026770275.1	—
<i>Nisaetus alboniger</i>	GCA_026109245.1	NC_007599.1
<i>Nisaetus nipalensis</i>	GCA_012487455.1	NC_007598.1
<i>Pandion haliaetus</i>	GCA_013401275.1	NC_008550.1
<i>Sagittarius serpentarius</i>	GCA_013399415.1	NC_023788.1
<i>Sarcogyps calvus</i>	GCA_040869245.1	OR896160.1
<i>Spizaetus tyrannus</i>	GCA_013399215.1	NC_052803.1

et al. 2017). The resulting trees were visualised in FigTree v. 1.4.4 (Rambaut 2018), with Galliformes species used as the outgroup.

Divergence times were estimated using MCMCTree implemented in the PAML v. 4.9j package. The gradient and Hessian matrix were calculated using CODEML, and analyses were performed with a burn-in of 1,000 iterations. A secondary fossil calibration point was constructed using divergence constraints between *A. chrysaetos* and *G. gallus*, *Sagittarius serpentarius* and *G. californianus*, *A. chrysaetos* and *P. haliaetus*, and *A. chrysaetos* and *Astur gentilis*. Minimum and maximum bounds of 86.5–101.8, 56.9–90.0, 37.7–59.0, and 8.1–50.1 Mya, respectively, were obtained from the TimeTree database (Kumar et al. 2022). Taxonomic nomenclature followed Gregory et al. (2024).

Positively selected gene (PSG) analysis

To identify genes under positive selection, coding sequences have been analysed from six accipitriform species, including *A. nipalensis*, *A. chrysaetos*, *Spizaetus tyrannus*, *Astur gentilis*, *H. albicilla*, and *Sarcogyps calvus*, with

G. gallus (Galliformes) serving as the outgroup. Positive selection analyses were conducted using the branch-site model in CODEML, implemented via PosiGene (Sahm et al. 2017). Candidate PSGs were identified based on a non-synonymous to synonymous substitution ratio ($dN/dS > 1$) and a significance threshold of False Discovery Rate (FDR) correction ($Q\text{-value} < 0.05$). Functional annotation and enrichment analyses of significant PSGs were conducted using the DAVID database to classify associated Gene Ontology (GO) biological processes, molecular functions, and cellular components (Huang et al. 2009; Sherman et al. 2022).

Results

De novo genome assembly and annotation

The Steppe Eagle reference genome was assembled using raw reads generated from ONT sequencing (36× coverage; 43.90 Gb total bases) and the MGISEQ-2000 platform (117× coverage; 142.05 Gb total bases). The final assembled genome was 1.21 Gb in length (154 × total coverage) and has been deposited in GenBank under accession number PRJNA916808. After removing low-quality and duplicated contigs, the assembly achieved a scaffold N50 of 23.49 Mb. Genome completeness assessed using BUSCO v. 4.0.5 with the Aves dataset (8,338 single-copy orthologs) revealed 97.2% complete BUSCOs, including 96.9% single-copy and 0.3% duplicated BUSCOs, along with 0.6% fragmented and 2.2% missing BUSCOs. Detailed genome assembly statistics and repeat element composition are summarized in Tables 2, 3. Genome annotation identified 16,192 protein-coding genes with an average gene length of 14,196 bp. On average, each gene contained 10.02 exons and 7.79 introns. Detailed gene structure statistics, including exon and intron lengths and GC content, are summarized in Table 4.

Table 2. Genome assembly statistics for *Aquila nipalensis*.

Statistic	Value
N50 (bp)	23,485,944
L50	18
N70 (bp)	13,799,119
L70	32
N90 (bp)	6,409,290
L90	57
Total assembly size (bp)	1,206,752,884
Number of contigs/scaffolds	1,036
Contigs/scaffolds $\geq$ 100 kbp	147
Contigs/scaffolds $\geq$ 1 Mbp	86
Contigs/scaffolds $\geq$ 10 Mbp	44
Longest contig/scaffold (bp)	52,297,487
Percentage of ambiguous bases (%N)	0
GC (%)	42.29
BUSCO completeness (%) (Aves_odb10)	97.2
BUSCO duplicated (%)	0.3
BUSCO fragmented (%)	0.6
BUSCO missing (%)	2.2

Table 3. Summary of repetitive elements in the genome of *Aquila nipalensis*.

Repeat element	Number of elements	Total length (bp)	Percentage of genome (%)
Retrotransposons			
LINE	83,855	32,712,501	2.71
SINE	2	151	0.00
LTR	14,512	4,268,267	0.35
DNA transposons	4,706	346,655	0.03
Simple repeats	279,202	10,560,709	0.88
Low complexity	51,803	2,554,678	0.21
Unclassified	124,764	42,392,404	3.51

Table 4. Summary statistics of gene annotation in the genome of *Aquila nipalensis*.

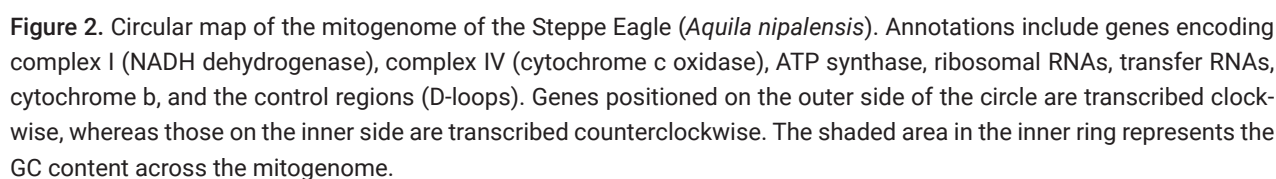
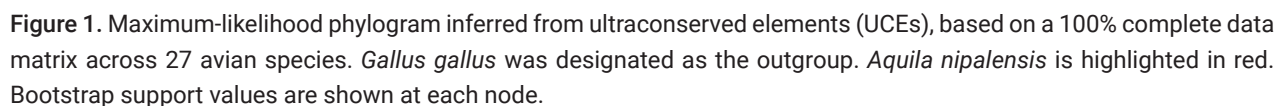
Feature	Value
Protein-coding genes	16,192
Total gene length (bp)	228,000,000
Average gene length (bp)	14,196
Total exon length (bp)	28,230,000
Average exon length (bp)	178
Total intron length (bp)	201,040,000
Average intron length (bp)	1,377
Average exons per gene	10.02
Average introns per gene	7.79
GC content in exons (%)	51.73
GC content in introns (%)	42.22

UCE-based phylogenomic analysis

After filtering the complete matrices at 75%, 95%, and 100% occupancy thresholds, 4,883, 4,698, and 4,535 UCE loci were retained, respectively. The resulting phylogenetic trees from all three filtered datasets exhibited identical topologies and consistent bootstrap support values, indicating that varying levels of missing data had no detectable impact on tree structure or nodal support. The analysis recovered *A. nipalensis* as sister to *A. chrysaetos*, identifying the latter as its closest extant relative within the dataset. At the genus level, *Aquila* formed a well-supported clade closely related to *Spizaetus*. These taxa were nested within a broader monophyletic clade of the family Accipitridae, which included representatives of the genera *Accipiter* Brisson, 1760; *Aegypius* Savigny, 1809; *Astur* Lacépède, 1801; *Butastur* Hodgson, 1843; *Buteo* Lacépède, 1799; *Circaetus* Vieillot, 1816; *Gyps* Savigny, 1809; *Haliaeetus* Savigny, 1809; *Haliaastur* Selby, 1840; *Milvus* Lacépède, 1799; *Nisaetus* Hodgson, 1836; *Sarcogyps* Lesson, 1842; *Spizaetus* Vieillot, 1816; and *Tachyspiza* Kaup, 1844 (Fig. 1).

Mitogenome assemblies and comparative mitogenome analysis

The mitogenome sequence has been deposited in GenBank under accession number: [PV670917](#). The mitogenome of the Steppe Eagle was 17,731 bp in length, and the order of mitochondrial genes is illustrated in Fig. 2. The mitogenome contained 13 PCGs, 22 transfer RNA genes (tRNAs), two ribosomal RNA



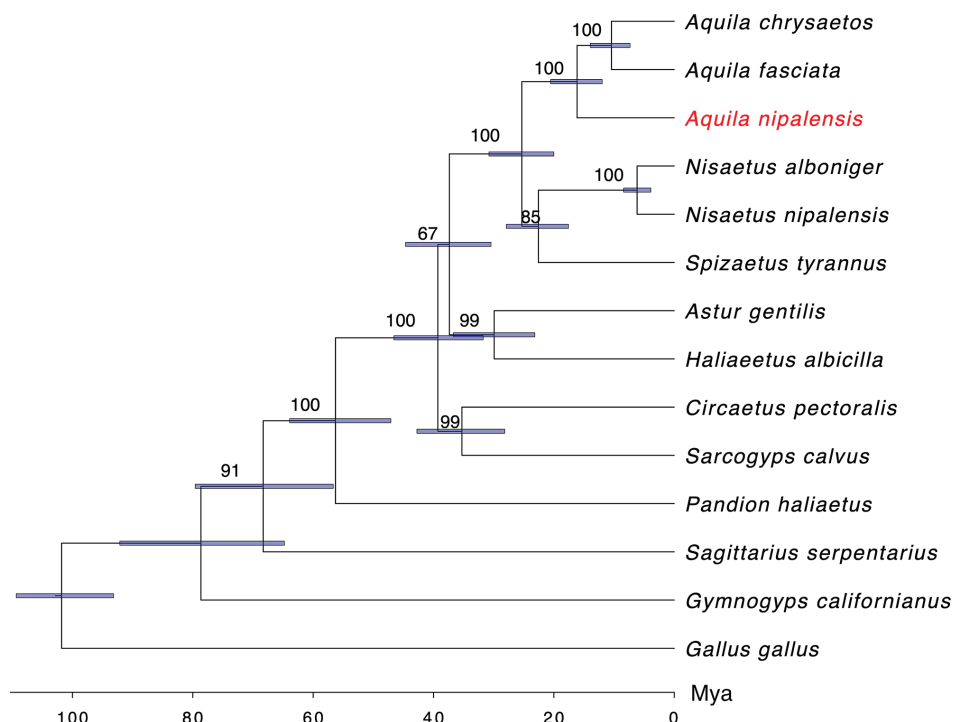


Figure 3. Maximum-likelihood phylogenetic analyses based on concatenated nucleotide sequences of the 13 mitochondrial protein-coding genes (PCGs) from 14 avian species. Time-calibrated phylogenetic tree with estimated divergence times (in million years ago, Mya) and bootstrap support values. *Gallus gallus* was used as the outgroup. *Aquila nipalensis* is highlighted in red.

genes, and two control regions (D-loops). The overall nucleotide composition was 53.5% adenine and thymine (AT) and 46.5% guanine and cytosine (GC). The protein-coding regions covered 11,391 bp, accounting for 64.2% of the total mitogenome of *A. nipalensis*. All PCGs, except *ND6*, were transcribed on the plus strand. The predominant start codon was ATG, although *COX1*, *ND5*, and *ND3* utilized GTG, GTG, and ATC, respectively. Relative synonymous codon usage (RSCU) analysis revealed that the codons CCU (Pro), CCC (Pro) and UCC (Ser) were the most frequently used (Suppl. material 1: table S1).

A ML phylogenetic analysis based on concatenated nucleotide sequences of the 13 mitochondrial PCGs from 14 avian species within Accipitriformes and Cathartiformes, including *Gallus gallus* as the outgroup, is presented in Fig. 3. The corresponding phylogram with branch lengths proportional to genetic distance is shown in Suppl. material 1: fig. S1. *Aquila nipalensis* was recovered within the subfamily Aquilinae, clustering with *Aquila chrysaetos* (Linnaeus, 1758), *Aquila fasciata* Vieillot, 1822, *Nisaetus alboniger* (Blyth, 1845), *Nisaetus nipalensis* Hodgson, 1833, and *Spizaetus tyrannus* (Wied-Neuwied, 1820). The Aquilinae formed a sister clade to Accipitrinae (represented by *Astur gentilis*) and Buteonini (represented by *H. albicilla*). Divergence time estimation indicated that the most recent common ancestor of *A. nipalensis* and its closest relatives dated to approximately 15.1 Mya (95% highest posterior density [HPD]: 11.16–19.18 Mya). Fossil calibration constraints were applied to the split between *A. chrysaetos* and *G. gallus*, *Sagittarius serpentarius* and *G. californianus*, *A. chrysaetos* and *P. haliaetus*, and *A. chrysaetos* and *Astur gentilis*. Estimated divergence times across the analysed lineages ranged from approximately 5 to 90 Mya (Fig. 3). Genetic distance analysis revealed that pairwise distances

between *A. nipalensis* and other *Aquila* species ranged from approximately 6.9% to 7.0%, while distances between *A. nipalensis* and species from other genera ranged from 9.7% to 27.1% (Suppl. material 1: table S2).

Positively selected genes

Sixteen significant PSGs were identified in *A. nipalensis*. Functional annotation indicated that these PSGs are associated with exocytosis (*EXOC7*, *SYCN*), cytoplasmic vesicle transport (*MYOF*, *SYCN*, *TGFA*), centriolar satellites (*EXOC7*, *MYOF*), and phospholipid binding (*SGIP1*, *MYOF*). The complete list of PSGs and their detailed functional annotations are provided in Tables 5, 6.

Table 5. Candidate positively selected genes identified in *Aquila nipalensis*.

Gene symbol	Gene name	dN/dS	P-value	FDR (Q-value)	Species included
SYCN	syncollin	30.23	3.93E-03	0.036	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Sarcogyps calvus</i>
TGFA	transforming growth factor alpha	20.38	1.22E-03	0.015	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Gallus gallus</i>
LOC115340598	poly(rC)-binding protein 3-like	9.76	2.41E-10	3.38E-08	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Gallus gallus</i>
LOC115352069	heat shock transcription factor, Y-linked-like	9.73	9.85E-07	4.55E-05	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Sarcogyps calvus</i>
SEBOX	SEBOX homeobox	6.80	1.10E-04	0.0020976	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Sarcogyps calvus</i>
EXOC7	exocyst complex component 7	5.69	7.52E-06	0.00024125	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Gallus gallus</i>
NEURL4	neuralized E3 ubiquitin protein ligase 4	3.87	8.96E-06	0.00024125	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Sarcogyps calvus</i>
CFAP92	cilia and flagella associated protein 92 (putative)	3.73	3.76E-07	2.02E-05	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Sarcogyps calvus</i>
PROB1	proline rich basic protein 1	3.63	2.49E-10	3.38E-08	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Sarcogyps calvus</i>

Gene symbol	Gene name	dN/dS	P-value	FDR (Q-value)	Species included
MTDH	metadherin	3.13	5.68E-03	0.0470669	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Gallus gallus</i>
LOC115351132	uncharacterized LOC115351132	2.86	2.85E-03	0.02814015	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Sarcogyps calvus</i>
SGIP1	SH3GL interacting endocytic adaptor 1	2.86	2.78E-03	0.02814015	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Gallus gallus</i>
GPAT2	glycerol-3-phosphate acyltransferase 2, mitochondria	2.64	3.82E-03	0.03623137	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Sarcogyps calvus</i>
MYOF	myoferlin	1.77	3.14E-10	3.38E-08	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Gallus gallus</i>
C1H4orf50	chromosome 1 C4orf50 homolog	1.65	5.47E-03	0.0470669	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Sarcogyps calvus</i>
DNAAF9	dynein axonemal assembly factor 9	1.37	3.90E-04	0.00617391	<i>Aquila chrysaetos</i> <i>Aquila nipalensis</i> <i>Astur gentilis</i> <i>Gallus gallus</i>

Discussion

The genome size of the Steppe Eagle, *A. nipalensis*, was estimated at 1.21 Gb, which falls within the known range for Accipitriformes (1.13–1.29 Gb) (Zhang et al. 2014a; Mead et al. 2021) and aligns with recent comparative avian genomic assessments (Catanach and Pirro 2023; Catanach et al. 2025), further supporting the conserved genome size architecture within Accipitriformes. Phylogenomic inference from all sets based on UCEs recovered *A. nipalensis* as sister to *A. chrysaetos*, supporting its placement within Aquilinae (Catanach et al. 2025). The mitogenome of *A. nipalensis* exhibits the typical avian mitochondrial gene composition and organization, including

Table 6. Gene ontology enrichment analysis of candidate positively selected genes.

Category	GO term	P-value	Fold enrichment	FDR (Q-value)	Gene involved
UP_KW_BIOLOGICAL_PROCESS	exocytosis	0.0294	56.06	0.353	<i>EXOC7</i> , <i>SYCN</i>
GOTERM_BP_DIRECT	exocytosis	0.0548	31.65	1.00e+0	<i>EXOC7</i> , <i>SYCN</i>
GOTERM_CC_DIRECT	cytoplasmic vesicle	0.0567	6.81	1.00e+0	<i>MYOF</i> , <i>SYCN</i> , <i>TGFA</i>
GOTERM_CC_DIRECT	centriolar satellite	0.0622	28.12	1.00e+0	<i>EXOC7</i> , <i>MYOF</i>
GOTERM_MF_DIRECT	phospholipid binding	0.0631	27.73	1.00e+0	<i>SGIP1</i> , <i>MYOF</i>

13 protein-coding genes, 22 tRNAs, and two rRNA genes. Comparative analysis showed that its gene order is largely conserved among Accipitriformes mitogenomes reported previously (Zhou et al. 2019; Urantówka et al. 2021). Notably, the mitogenome of *A. nipalensis* contains two control regions, which differs from some earlier reports (Zhou et al. 2019) but is consistent with several accipitriform species (Urantówka et al. 2021; Sonongbua et al. 2024). Duplication of mitochondrial control regions has been associated with increased longevity in birds (Skujina et al. 2016), suggesting potential functional or life-history implications. However, further comparative analyses are required to clarify its adaptive significance in raptors.

UCE-based phylogenomic analysis provided the most robust estimate of species relationships. UCEs, distributed across the nuclear genome and biparentally inherited, capture genome-wide evolutionary history and avoid the maternal bias inherent to mitochondrial markers (Faircloth et al. 2012; McCormack et al. 2013). Across all occupancy thresholds, tree topology and bootstrap support remained consistent, indicating strong phylogenomic signal and minimal impact of missing data. Nevertheless, concatenated phylogenomic analyses may obscure gene tree heterogeneity arising from incomplete lineage sorting or hybridization, potentially leading to inaccurate species tree inference (Kubatko and Degnan 2007).

The mitogenome of *A. nipalensis* exhibits the typical avian mitochondrial gene composition and organization, including 13 protein-coding genes, 22 tRNAs, and two rRNA genes. Comparative analysis showed that its gene order is largely conserved among Accipitriformes mitogenomes reported previously (Zhou et al. 2019; Urantówka et al. 2021). Notably, the mitogenome of *A. nipalensis* contains two control regions, which differs from some earlier reports (Zhou et al. 2019) but is consistent with several accipitriform species (Urantówka et al. 2021; Sonongbua et al. 2024). Duplication of mitochondrial control regions has been associated with increased longevity in birds (Skujina et al. 2016), suggesting potential functional or life-history implications. However, further comparative analyses are required to clarify its adaptive significance in raptors. Although complete mitogenomes provided strong phylogenetic signal due to their high copy number and extensive comparative databases (Galtier et al. 2009), mitochondrial DNA represents a single maternally inherited locus and may therefore produce phylogenetic patterns that differ from nuclear genome-based hypotheses (Toews and Brelsford 2012; Mindell et al. 2018). Such discordance may arise from mitochondrial introgression, substitution saturation, or lineage-specific rate heterogeneity. Divergence time estimates differed between the 13-PCG dataset (~15.1 Mya) and the *CYTB*+*ND2* dataset (~6.9 Mya), underscoring the impact of marker choice on molecular dating. The longer concatenated 13-PCG dataset likely provided increased phylogenetic signal and reduced stochastic variance, resulting in more stable rate estimates. In contrast, shorter loci such as *CYTB* and *ND2* are more susceptible to rate heterogeneity, selective pressures, and saturation, potentially leading to biased divergence estimates (Campillo et al. 2019; Jiang et al. 2025). Despite these discrepancies in branch lengths and node ages, overall topology remained congruent across datasets, reinforcing the stability of inferred relationships within *Aquilinae* while emphasising caution in interpreting divergence times.

Genes under positive selection further provided insight into potential adaptive processes. *EXOC7* and *SYCN* are involved in vesicle trafficking and regulated exocytosis, processes essential for membrane targeting and cellular homeostasis (Wäsle et al. 2005; He et al. 2007). *SYCN* has also been proposed to possess antimicrobial properties (Waters et al. 2021), which may be associated with dietary exposure to pathogen-rich environments. *MYOF*, which is associated with endothelial growth and tissue development (Bernatchez et al. 2009), may contribute to cardiovascular and muscular adaptations relevant to sustained flight performance. The relatively small number of PSGs detected is likely attributable to the short divergence times among closely related accipitrid species. Broader taxonomic sampling and complementary population-genomic approaches may further enhance detection of lineage-specific adaptive signals (Nielsen et al. 2005).

Conclusions

This study provides the high-quality *de novo* genome and complete mitogenome assembly of *A. nipalensis*, establishing an essential genomic reference for this endangered raptor. Integrating nuclear UCE-based phylogenomics with complete and partial mitochondrial datasets clarified its evolutionary position within Accipitridae and demonstrated how marker selection can influence divergence time estimation. In addition, the identification of PSGs associated with vesicle trafficking, exocytosis, and tissue development offers preliminary insights into potential adaptive mechanisms underlying physiological performance and environmental resilience. Together, these genomic resources provide a comprehensive framework for future investigations of adaptive evolution, population structure, and evidence-based conservation strategies for the Steppe Eagle.

Acknowledgements

We extended our sincere gratitude to the staff at Khao Kheow Open Zoo for their assistance with sample collection. Special thanks are also due to the staff of the National Omics Center for their valuable suggestions on data analysis. We are grateful to the reviewers for their constructive comments and insightful suggestions, which helped improve the quality of this manuscript.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* 215(3): 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Arcones A, Ponti R, Vieites DR (2021) Mitochondrial substitution rates estimation for divergence time analyses in modern birds based on full mitochondrial genomes. *International Journal of Avian Science* 163: 1463–1471. <https://doi.org/10.1111/ibi.12965>
- Bari F, Rehman EU, Kabir M, Ahmad S (2020) An extension to the known wintering range of the Steppe Eagle *Aquila nipalensis*, in the Poonch and Jhelum Valleys, Azad Jammu and Kashmir, Pakistan. *Ardeola* 67(2): 415–422. <https://doi.org/10.13157/arla.67.2.2020.sc1>

- Bernatchez PN, Sharma A, Kodaman P, Sessa WC (2009) Myoferlin is critical for endocytosis in endothelial cells. *American Journal of Physiology – Cell Physiology* 297: C484–C492. <https://doi.org/10.1152/ajpcell.00498.2008>
- Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsche G, Pütz J, Middendorff M, Stadler PF (2013) MITOS: improved *de novo* metazoan mitochondrial genome annotation. *Molecular Phylogenetic and Evolution* 69(2): 313–319. <https://doi.org/10.1016/j.ympev.2012.08.023>
- BirdLife International (2020) *Aquila nipalensis*. The IUCN Red List of Threatened Species 2020: e.T22696038A180479129. <https://doi.org/10.2305/IUCN.UK.2020-3.RLTS.T22696038A180479129.en>
- Buthasane W, Shotelersuk V, Chetruengchai W, Srichomthong C, Assawapitaksakul A, Tangphatsornruang S, Pootakham W, Sonthirod C, Tongsima S, Wangkumhang P, Wil-antho A, Thongphakdee A, Sanannu S, Poksawat C, Nipanunt T, Kasornrorkbua C, Koepfli K-P, Pukazhenthil BS, Suriyaphol P, Wongsurawat T, Jenjaroenpun P, Suriyaphol G (2024) Comprehensive genome assembly reveals genetic diversity and carcass consumption insights in critically endangered Asian king vultures. *Scientific Reports* 14(1): 9455. <https://doi.org/10.1038/s41598-024-59990-9>
- Campillo LC, Burns KJ, Moyle RG, Manthey JD (2019) Mitochondrial genomes of the bird genus *Piranga*: rates of sequence evolution, and discordance between mitochondrial and nuclear markers. *Mitochondrial DNA Part B: Resources* 4: 2566–2569. <https://doi.org/10.1080/23802359.2019.1637286>
- Catanach TA, Pirro S (2023) The complete genome sequences of 87 species of hawks (Accipitriformes, Aves). *Biodiversity Genomes* 2023: 1–5. <https://doi.org/10.56179/001c.67877>
- Catanach TA, Halley MR, Pirro S (2025) Enigmas no longer: using ultraconserved elements to place several unusual hawk taxa and address the non-monophyly of the genus *Accipiter* (Accipitriformes: Accipitridae). *Biological Journal of the Linnean Society* 144(2): blae028. <https://doi.org/10.1093/biolinnean/blae028>
- Dierckxsens N, Mardulyn P, Smits G (2017) NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Research* 45(4): e18. <https://doi.org/10.1093/nar/gkw955>
- Eo SH, DeWoody JA (2010) Evolutionary rates of mitochondrial genomes correspond to diversification rates and to contemporary species richness in birds and reptiles. *Proceedings of the Royal Society B: Biological Sciences* 277(1700): 3587–3592. <https://doi.org/10.1098/rspb.2010.0965>
- Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC (2012) Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology* 61(5): 717–726. <https://doi.org/10.1093/sysbio/sys004>
- Galtier N, Nabholz B, Glemin S, Hurst GDD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology* 18: 4541–4550. <https://doi.org/10.1111/j.1365-294X.2009.04380.x>
- Gregory SMS, Sangster G, Worthy TH, Scofield RP (2024) Falling through the cracks: a family-group name for a clade of hawks and eagles (Accipitridae) including *Morphnus* Dumont, 1816, *Harpia* Vieillot, 1816, *Harpyopsis* Salvadori, 1875 and *Macheiramphus* Bonaparte, 1850. *Avian Systematics* 2(II): N5–N17.
- Greiner S, Lehwark P, Bock R (2019) OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Research* 47(W1): W59–W64. <https://doi.org/10.1093/nar/gkz238>

- He B, Xi F, Zhang X, Zhang J, Guo W (2007) Exo70 interacts with phospholipids and mediates the targeting of the exocyst to the plasma membrane. *The EMBO Journal* 26: 4053–4065. <https://doi.org/10.1038/sj.emboj.7601834>
- Helbig AJ, Kocum A, Seibold I, Braun MJ (2005) A multi-gene phylogeny of aquiline eagles (Aves: Accipitriformes) reveals extensive paraphyly at the genus level. *Molecular Phylogenetics and Evolution* 35: 147–164. <https://doi.org/10.1016/j.ympev.2004.10.003>
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2017) UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution* 35(2): 518–522. <https://doi.org/10.1093/molbev/msx281>
- Holt C, Yandell M (2011) MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12: 491. <https://doi.org/10.1186/1471-2105-12-491>
- Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* 4: 44–57. <https://doi.org/10.1038/nprot.2008.211>
- Jiang C, Zhu W, Wu S, Zhao X, Mohamed N, Li B (2025) Comparative mitogenomics, phylogeny, and biogeography of selected species of *Saxicola* (Aves, Passeriformes). *ZooKeys* 1249: 69–92. <https://doi.org/10.3897/zookeys.1249.152269>
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA (2019) Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology* 37(5): 540–546. <https://doi.org/10.1038/s41587-019-0072-8>
- Korf I (2004) Gene finding in novel genomes. *BMC Bioinformatics* 5: 59. <https://doi.org/10.1186/1471-2105-5-59>
- Kubatko LS, Degnan JH (2007) Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology* 56: 17–24. <https://doi.org/10.1080/10635150601146041>
- 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>
- Kumar S, Suleski M, Craig JM, Kasprowicz AE, Sanderford M, Li M, Stecher G, Hedges SB (2022) TimeTree 5: an expanded resource for species divergence times. *Molecular Biology and Evolution* 39(8): msac174. <https://doi.org/10.1093/molbev/msac174>
- Lerner H, Mindell DP (2005) Phylogeny of eagles, Old World vultures, and other Accipitridae based on nuclear and mitochondrial DNA. *Molecular Phylogenetics and Evolution* 37: 327–346. <https://doi.org/10.1016/j.ympev.2005.04.010>
- Lerner H, Christidis L, Gamauf A, Griffiths C, Haring E, Huddleston CJ, Kabra S, Kocum A, Krosby M, Kvaløy K, Mindell D, Rasmussen P, Røv N, Wadleigh R, Wink M, Gjershaug JO (2017) Phylogeny and new taxonomy of the Booted Eagles (Accipitriformes: Aquilinae). *Zootaxa* 4216: 301–320. <https://doi.org/10.11646/zootaxa.4216.4.1>
- Manni M, Berkeley MR, Seppey M, Zdobnov EM (2021) BUSCO: assessing genomic data quality and beyond. *Current Protocols* 1(12): e323. <https://doi.org/10.1002/cpz1.323>
- McCormack JE, Harvey MG, Faircloth BC, Crawford NG, Glenn TC, Brumfield RT (2013) A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. *PLOS ONE* 8: e54848. <https://doi.org/10.1371/journal.pone.0054848>
- Mead D, Ogden R, Meredith A, Peniche G, Smith M, Corton C, Oliver K, Skelton J, Betteridge E, Doulcan J, Holmes N, Wright V, Loose M, Quail MA, McCarthy SA, Howe K, Chow W, Torrance J, Collins J, Challis R, Durbin R, Blaxter M (2021) The genome sequence of the European Golden Eagle, *Aquila chrysaetos chrysaetos* Linnaeus 1758. *Wellcome Open Research* 6: 112. <https://doi.org/10.12688/wellcomeopenres.16631.1>

- Mindell DP, Fuchs J, Johnson JA (2018) Phylogeny, taxonomy, and geographic diversity of diurnal raptors: Falconiformes, Accipitriformes, and Cathartiformes. In: Sarasola JH, Grande JM, Negro JJ (Eds) *Birds of Prey: Biology and Conservation in the XXI Century*. Springer, Cham, 3–32. https://doi.org/10.1007/978-3-319-73745-4_1
- Nabholz B, Glémin S, Galtier N (2009) The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC Evolutionary Biology* 9: 54. <https://doi.org/10.1186/1471-2148-9-54>
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2014) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32(1): 268–274. <https://doi.org/10.1093/molbev/msu300>
- Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C (2005) Genomic scans for selective sweeps using SNP data. *Genome Research* 15: 1566–1575. <https://doi.org/10.1101/gr.4252305>
- Orihuela-Torres A, Morales-Reyes Z, Pérez-García JM, Naves-Alegre L, Sánchez-Zapata JA, Sebastián-González E (2021) Unravelling the vertebrate scavenger assemblage in the Gobi Desert, Mongolia. *Journal of Arid Environments* 190: 104509. <https://doi.org/10.1016/j.jaridenv.2021.104509>
- Rambaut A (2018) FigTree v1.4.4. <http://tree.bio.ed.ac.uk/software/figtree/> [accessed on 10 January 2025]
- Sahm A, Bens M, Platzer M, Szafranski K (2017) PosiGene: automated and easy-to-use pipeline for genome-wide detection of positively selected genes. *Nucleic Acids Research* 45: e100. <https://doi.org/10.1093/nar/gkx179>
- Sharma AK, Saini M, Singh SD, Prakash V, Das A, Bharathi Dasan R, Pandey S, Bohara D, Galligan TH, Green RE, Knopp D, Cuthbert RJ (2014) Diclofenac is toxic to the Steppe Eagle *Aquila nipalensis*: widening the diversity of raptors threatened by NSAID misuse in South Asia. *Bird Conservation International* 24(3): 282–286. <https://doi.org/10.1017/S0959270913000609>
- Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, Imamichi T, Chang W (2022) DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research* 50: W216–W221. <https://doi.org/10.1093/nar/gkac194>
- Skujina I, McMahon R, Lenis VP, Gkoutos GV, Hegarty M (2016) Duplication of the mitochondrial control region is associated with increased longevity in birds. *Aging (Albany NY)* 8(8): 1781–1789. <https://doi.org/10.18632/aging.101012>
- Sonongbua J, Thong T, Panthum T, Budi T, Singchat W, Kraichak E, Chaiyes A, Muangmai N, Duengkae P, Sitdhibutr R, Kasorndorkbua C, Srikulnath K (2024) Insights into mitochondrial rearrangements and selection in accipitrid mitogenomes, with new data on *Haliastur indus* and *Accipiter badius poliopsis*. *Genes (Basel)* 15(11): 1439. <https://doi.org/10.3390/genes15111439>
- Stiller J, Feng S, Chowdhury A-A, Rivas-González I, Duchêne DA, Fang Q, Deng Y, Kozlov A, Stamatakis A, Claramunt S, Nguyen JMT, Ho SYW, Faircloth BC, Haag J, Houde P, Cracraft J, Balaban M, Mai U, Chen G, Gao R, Zhou C, Xie Y, Huang Z, Cao Z, Yan Z, Ogilvie HA, Nakhleh L, Lindow B, Morel B, Fjeldså J, Hosner PA, da Fonseca RR, Petersen B, Tobias JA, Székely T, Kennedy JD, Reeve AH, Liker A, Stervander M, Antunes A, Tietze DT, Bertelsen MF, Lei F, Rahbek C, Graves GR, Schierup MH, Warnow T, Braun EL, Gilbert MTP, Jarvis ED, Mirarab S, Zhang G (2024) Complexity of avian evolution revealed by family-level genomes. *Nature* 629(8013): 851–860. <https://doi.org/10.1038/s41586-024-07323-1>

- Toews DP, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* 21: 3907–3930. <https://doi.org/10.1111/j.1365-294X.2012.05664.x>
- Urantówka AD, Krocak A, Strzała T, Zaniewicz G, Kurkowski M, Mackiewicz P (2021) Mitogenomes of Accipitriformes and Cathartiformes were subjected to ancestral and recent duplications followed by gradual degeneration. *Genome Biology and Evolution* 13(9): evab193. <https://doi.org/10.1093/gbe/evab193>
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM (2014) Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLOS ONE* 9(11): e112963. <https://doi.org/10.1371/journal.pone.0112963>
- Wäsle B, Turvey M, Larina O, Thorn P, Skepper J, Morton AJ, Edwardson JM (2005) Syncollin is required for efficient zymogen granule exocytosis. *Biochemical Journal* 385: 721–727. <https://doi.org/10.1042/BJ20041064>
- Waters RA, Robinson J, Edwardson JM (2021) Syncollin is an antibacterial polypeptide. *Cellular Microbiology* 23: e13372. <https://doi.org/10.1111/cmi.13372>
- Zhang G, Li B, Li C, Gilbert MTP, Jarvis ED, Wang J, The Avian Genome C (2014a) Comparative genomic data of the Avian Phylogenomics Project. *GigaScience* 3(1): 26. <https://doi.org/10.1186/2047-217X-3-26>
- Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith RW, Ödeen A, Cui J, Zhou Q, Xu L, Pan H, Wang Z, Jin L, Zhang P, Hu H, Yang W, Hu J, Xiao J, Yang Z, Liu Y, Xie Q, Yu H, Lian J, Wen P, Zhang F, Li H, Zeng Y, Xiong Z, Liu S, Zhou L, Huang Z, An N, Wang J, Zheng Q, Xiong Y, Wang G, Wang B, Wang J, Fan Y, da Fonseca RR, Alfaro-Núñez A, Schubert M, Orlando L, Mourier T, Howard JT, Ganapathy G, Pfenning A, Whitney O, Rivas MV, Hara E, Smith J, Farré M, Narayan J, Slavov G, Romanov MN, Borges R, Machado JP, Khan I, Springer MS, Gatesy J, Hoffmann FG, Opazo JC, Håstad O, Sawyer RH, Kim H, Kim K-W, Kim HJ, Cho S, Li N, Huang Y, Bruford MW, Zhan X, Dixon A, Bertelsen MF, Derryberry E, Warren W, Wilson RK, Li S, Ray DA, Green RE, O'Brien SJ, Griffin D, Johnson WE, Haussler D, Ryder OA, Willerslev E, Graves GR, Alström P, Fjeldså J, Mindell DP, Edwards SV, Braun EL, Rahbek C, Burt DW, Houde P, Zhang Y, Yang H, Wang J, Consortium AG, Jarvis ED, Gilbert MTP, Wang J, Ye C, Liang S, Yan Z, Zepeda ML, Campos PF, Velazquez AMV, Samaniego JA, Avila-Arcos M, Martin MD, Barnett R, Ribeiro AM, Mello CV, Lovell PV, Almeida D, Maldonado E, Pereira J, Sunagar K, Philip S, Dominguez-Bello MG, Bunce M, Lambert D, Brumfield RT, Sheldon FH, Holmes EC, Gardner PP, Steeves TE, Stadler PF, Burge SW, Lyons E, Smith J, McCarthy F, Pitel F, Rhoads D, Froman DP (2014b) Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 346(6215): 1311–1320. <https://doi.org/10.1126/science.1251385>
- Zhou C, Tu H, Chen Y, Dou L, Meng Y, Yang N, Yue B, Wu Y (2019) The complete mitochondrial genome of *Aquila nipalensis* and its phylogenetic position. *Mitochondrial DNA Part B: Resources* 4(2): 2152–2153. <https://doi.org/10.1080/23802359.2019.1623118>

Supplementary material 1

Genetic data

Authors: Wannapol Buthasane, Thidathip Wongsurawat, Piroon Jenjaroenpun, Sithichoke Tangphatsornruang, Wirulda Pootakham, Chutima Sonthirot, Wasitthee Kongkachana, Alisa Wilantho, Ratiwan Sitdhibutr, Chaian Kasorndorkbua, Gunnaporn Suriyaphol

Data type: docx

Explanation note: **table S1**. Relative synonymous codon usage (RSCU) values of mitochondrial protein-coding genes in the steppe eagle (*Aquila nipalensis*). **figure S1**. Maximum likelihood phylogram with branch lengths proportional to genetic distance based on concatenated sequences of 13 mitochondrial PCGs from 14 avian species. **table S2**. Pairwise genetic distances between the steppe eagle (*Aquila nipalensis*) and 21 other bird species, calculated based on concatenated mitochondrial protein-coding gene sequences.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/zookeys.1281.158566.suppl1>

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

Sampling was conducted in compliance with the ethical guidelines under the Institutional Animal Care and Use Committee of Kasetsart University, Thailand (approval no. ACKU64-VET-052).

Artificial Intelligence (AI) use

The authors accept full responsibility for the content of the manuscript, including the disclosure of any use of AI.

No AI tools were used in the preparation of this manuscript.

Funding

This work was supported by the Second Century Fund (C2F), Chulalongkorn University; the 90th Anniversary of Chulalongkorn University Scholarship (grant no. GCUGR1125652078D). The funders had no role in the design of the study, data collection, analysis, interpretation or manuscript preparation.

Author contributions

Investigation: WB, TW, PJ; Formal analysis: WB, TW, PJ, WP, CS, WK, AW; Data curation: WB, TW, PJ; Methodology: WB, TW, PJ, ST, WP, CS, WK, AW; Software: TW, PJ, ST, WP, CS, WK, AW; Validation: TW, PJ, ST, WP, CS, WK, AW, GS; Supervision: TW, PJ, ST, GS; Visualization: WB; Resources: RS, CK; Conceptualization: GS; Funding acquisition: WB, GS; Project administration: GS; Writing – original draft: WB, GS; Writing – review and editing: TW, PJ, ST, CK, GS.

Author ORCIDs

W. Buthasane  <https://orcid.org/0009-0007-2819-9591>
T. Wongsurawat  <https://orcid.org/0000-0002-3659-2074>
P. Jenjaroenpun  <https://orcid.org/0000-0002-1555-401X>
S. Tangphatsornruang  <https://orcid.org/0000-0003-2673-0012>
W. Pootakham  <https://orcid.org/0000-0001-6721-6453>
C. Sonthirot  <https://orcid.org/0009-0009-3116-0078>
W. Kongkachana  <https://orcid.org/0000-0002-2356-9246>
R. Sitdhibutr  <https://orcid.org/0000-0003-0501-2839>
C. Kasorndorkbua  <https://orcid.org/0000-0002-2138-9954>
G. Suriyaphol  <https://orcid.org/0000-0002-8961-5970>

Data availability

All of the data that support the findings of this study are available in the main text or Supplementary Information.