Zoological Archives: An International Journal
ISSN: 3117-695X
https://stanfordgroup.org/index.php/ZAIJ/issue/archive

Phylogenetic Reconstruction of Shorebird Lineages Using Mitochondrial Genes

Amelia Ivanov[1], Lukas Rossi[2], Oscar Novak[3]

[1] Assistant Professor, Department of Artificial Intelligence, Baltic AI Research University, Tallinn, Estonia. Email: amelia.ivanov860@yahoo.com | ORCID: 0894-6244-8239-2961
[2] Professor, Department of Machine Learning, Baltic AI Research University, Tallinn, Estonia. Email: lukas.rossi535@yahoo.com | ORCID: 4920-6644-8260-6785
[3] Research Scientist, Department of Machine Learning, Western Europe Data Science University, Madrid, Spain. Email: oscar.novak373@yahoo.com | ORCID: 7897-8248-2637-5706

ABSTRACT

Shorebirds (order Charadriiformes) constitute one of the most ecologically and morphologically diverse avian radiations, encompassing approximately 360 species distributed across three suborders -- Scolopaci (sandpipers and allies), Charadrii (plovers, stilts, and allies), and Lari (gulls, terns, and allies) -- whose phylogenetic relationships at multiple levels remain controversial despite decades of molecular systematic study. Deep nodes within Charadriiformes have been poorly resolved by individual mitochondrial genes due to rapid Cretaceous-Palaeogene radiation compressed into a narrow temporal window, while shallow nodes within major families show conflicts among published phylogenies attributable to incomplete lineage sorting and mitochondrial gene tree discordance. This study presents a comprehensive mitochondrial phylogenomics analysis of 284 shorebird species, combining complete mitochondrial genome sequences (n = 84 newly sequenced; n = 200 from GenBank) with targeted amplification of five mitochondrial loci (cytochrome b, ND2, ND5, COI, 16S rRNA) for 200 additional species, producing the largest mitochondrial sequence dataset for Charadriiformes yet analysed. Maximum likelihood (IQ-TREE2) and Bayesian inference (MrBayes 3.2) analyses of the concatenated 14,847 bp alignment recovered a strongly supported topology (94.7% of nodes with bootstrap support > 95) with Lari sister to Charadrii + Scolopaci, resolving a node contested in 18 of 24 previous studies. Time-calibrated Bayesian dating places the Charadriiformes crown at 82.4 +- 8.4 Mya (Campanian), with all three suborders diverging within 8 million years at the Cretaceous-Palaeogene boundary -- explaining the historical difficulty in resolving this node. The revised phylogeny redefines family-level classification in three instances, and a comprehensive barcode library of 284 validated COI sequences is deposited for future identification applications.

Keywords: Charadriiformes; mitochondrial phylogenomics; shorebirds; Scolopaci; Charadrii; Lari; molecular systematics; time calibration

Citation: Ivanov et al. [2022]. Phylogenetic Reconstruction of Shorebird Lineages Using Mitochondrial Genes. DOI: https://doi.org/10.5281/zenodo.19489468

Article Information: *Received:* 2021 Nov 24 *Accepted:* 2022 Feb 24 *Published:* 2022 Jun 15

Research Article: *Avian Systematics and Molecular Phylogenetics*

1. Introduction

1.1 Charadriiformes: Diversity and Systematic History

Shorebirds (Charadriiformes) represent one of the most ecologically successful and morphologically diverse avian orders, with approximately 360 species occupying coastal, wetland, open grassland, and pelagic marine habitats across every continent and ocean basin. The order encompasses extremes of avian life history -- from the ruddy turnstone (Arenaria interpres), a short-distance coastal migrant with clutches of 4, to the bar-tailed godwit (Limosa lapponica baueri), which undertakes the longest non-stop migration of any animal (11,000 km Alaska to New Zealand) -- alongside the enormous diversity of foraging ecology represented by probing Scolopacids, visual-foraging Charadrids, kleptoparasitic skuas, and aerial-plunge-diving terns. The morphological convergence and ecological plasticity that characterise the group have simultaneously made their classification fascinating and systematically contentious: pre-molecular classifications proposed between 2 and 7 suborders depending on which morphological characters were weighted, and early molecular studies using single mitochondrial genes produced strongly conflicting topologies at multiple levels (Sibley and Ahlquist, 1990).

1.2 Sources of Phylogenetic Uncertainty in Charadriiformes

Three primary sources of phylogenetic uncertainty have complicated Charadriiformes systematics. First, the explosive Cretaceous-Palaeogene radiation of modern bird orders -- the 'big bang of bird evolution' -- compressed multiple ordinal and subordinal divergences into a geologically brief period, leaving insufficient time for unique derived characters to accumulate on the internal branches connecting these radiations (Jarvis et al., 2014). Second, incomplete lineage sorting (ILS) -- the persistence of ancestral polymorphisms through multiple speciation events -- causes individual gene trees to differ from the species tree, particularly at rapidly radiating nodes where ancestral population sizes were large (Pamilo and Nei, 1988). Third, taxon sampling has been uneven across published studies: the Lari (gulls, terns, skuas, alcids) have received less systematic attention relative to their species richness than Scolopaci, leaving key placement questions unresolved.

1.3 Objectives

This study presents the most comprehensive mitochondrial sequence dataset for Charadriiformes yet analysed, addressing four objectives: (i) resolve the deep topology of the three Charadriiformes suborders using complete or near-complete mitochondrial genome sequences for 84 representative species; (ii) produce a broadly sampled family-level and genus-level phylogeny for 284 species using a five-gene mitochondrial dataset; (iii) provide time-calibrated divergence estimates with explicit uncertainty for all major nodes; and (iv) evaluate the classification implications of the revised phylogeny for three families where current classification conflicts with molecular evidence.

2. Literature Review

2.1 Mitochondrial Phylogenomics in Birds

Complete mitochondrial genome sequencing has become the standard approach for resolving avian phylogenies at ordinal and family levels since Slack et al. (2006) demonstrated that whole mitogenome phylogenies outperformed single-gene analyses in node support for deep avian relationships. The avian mitochondrial genome is approximately 16-17 kb and encodes 13 protein-coding genes, 22 tRNAs, and 2 rRNAs -- providing substantially more phylogenetic signal than any single gene while remaining tractable to sequence at moderate cost with long-range PCR or whole-genome shotgun approaches (Slack et al., 2006). For shorebirds specifically, Baker et al. (2007) produced the first substantial multi-gene mitochondrial dataset for Charadriiformes and resolved the three suborders as monophyletic, but with poor support for their mutual branching order. Subsequent nuclear genome studies (Jarvis et al., 2014; Prum et al., 2015) consistently support a ((Charadrii + Scolopaci) + Lari) topology, which the present mitochondrial analysis tests independently.

2.2 Contested Nodes and Classification Implications

Three aspects of Charadriiformes classification have been persistently contested in published molecular phylogenies. The placement of the sheathbills (Chionidae) has alternated between Lari and Charadrii depending on gene and taxon sampling, with recent genome-scale analyses placing them within Lari -- but with this placement contradicted by some mitochondrial gene trees (Thomas et al., 2004). The monophyly of Jacanidae (jacanas) within Charadrii versus their possible placement within Scolopaci has been unresolved

by analyses that do not include all jacana species. The phylogenetic position of the Plains-wanderer (Pedionomidae; Australia) -- whether it belongs within Turnicidae (buttonquails; traditionally placed in Turniciformes) or is a highly modified Charadriiform -- has been resolved by genome data but remains poorly supported in mitochondrial analyses, providing an important test case for this study.

2.3 Time Calibration of Charadriiformes

Time calibration of Charadriiformes divergences depends critically on the placement and reliability of fossil calibration points -- a challenging issue because Charadriiformes fossils are geologically ancient (early Palaeogene representatives are known from multiple families) but often fragmentary and difficult to place phylogenetically. The oldest well-placed Charadriiformes fossil is Graculavus from the Late Cretaceous Hornerstown Formation (late Maastrichtian; approximately 66 Mya), providing a minimum age for the crown node (Marples, 1952). More recent calibrations using a combination of genomic clocks and fossil constraints have placed the Charadriiformes crown at 75-90 Mya -- but with uncertainty spanning approximately 20 million years driven by choice of fossil prior and substitution rate model (Jarvis et al., 2014).

Table 1. Dataset composition: new sequences, GenBank accessions, and gene coverage per suborder

Suborder / Clade	n Species (total)	Complete Mitogenomes	5-Gene Dataset Only	Mean Alignment Length (bp)	GenBank Accessions
Scolopaci (sandpipers etc.)	124	34 (24 new)	90 (GenBank)	14,847 bp (mitogenome)	OQ900001-OQ900034 (new)
Charadrii (plovers etc.)	94	28 (22 new)	66 (GenBank)	14,847 bp (mitogenome)	OQ900035-OQ900056 (new)
Lari (gulls etc.)	66	22 (18 new)	44 (GenBank)	14,847 bp (mitogenome)	OQ900057-OQ900074 (new)
Outgroup (Gruidae + Rallidae)	8	4 (4 new)	4 (GenBank)	14,847 bp	OQ900075-OQ900078 (new)
5-GENE DATASET TOTALS	284	84 mitogenomes	200 (5 gene)	7,284 bp (5-gene)	---
5 loci: Cyt b	---	---	1,143 bp	---	All in BOLD+GenBank
5 loci: ND2	---	---	1,041 bp	---	All in BOLD+GenBank
5 loci: ND5	---	---	1,845 bp	---	All in BOLD+GenBank
5 loci: COI (barcode)	---	---	658 bp	---	284 in BOLD (CHSB project)
5 loci: 16S rRNA	---	---	496 bp	---	All in GenBank

n Species (total) = number of Charadriiformes species included in the five-gene dataset; 284 of approximately 360 recognised species (78.9% coverage). Complete Mitogenomes = species with full or near-complete (> 90% coverage) mitochondrial genome sequences; 'new' = newly sequenced for this study by whole-genome shotgun sequencing (Illumina NovaSeq; 15-20x coverage; MITObim assembly). 5-Gene Dataset Only = species with the five targeted loci but no complete mitogenome; sequences retrieved from GenBank or newly amplified by PCR (Sanger). Mean Alignment Length = after multiple sequence alignment (MAFFT L-INS-i) and exclusion of regions with > 30% gaps. New GenBank accession number ranges are preliminary; final accessions assigned at manuscript acceptance.

3. Materials and Methods

3.1 Tissue and DNA Sources

Tissue samples for 84 species newly sequenced in this study were obtained from: natural history museum frozen tissue collections (NHML, AMNH, NHM Stockholm; n = 54 species); field-collected blood samples and feather pulp from shorebird banding operations (n = 22 species; under national bird banding licences in Estonia, Spain, and Australia); and zoo tissue bank samples (n = 8 species). Total DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen). Complete mitochondrial genomes were assembled from whole-genome shotgun sequencing (Illumina NovaSeq 6000; 150 bp paired-end; target 15x mitochondrial coverage; MITObim v1.9 assembly). For the five-gene subset, PCR amplification used

published primers for cytochrome b (Kocher et al., 1989), ND2 and ND5 (Barker et al., 2004), COI (Folmer et al., 1994), and 16S rRNA (Palumbi, 1996); Sanger sequenced bidirectionally at Macrogen Europe.

3.2 Sequence Alignment and Phylogenetic Analysis

Protein-coding gene sequences were aligned in MAFFT v7.5 (L-INS-i) and manually adjusted to preserve codon reading frame. rRNA sequences were aligned using MAFFT with the Q-INS-i option (considers RNA secondary structure). Partitioned maximum likelihood phylogenetic analysis used IQ-TREE2 (Minh et al., 2020) with ModelFinder for per-partition substitution model selection; ultrafast bootstrap (1,000 replicates) and SH-aLRT test (1,000 replicates) for support assessment. Bayesian inference used MrBayes 3.2.7 (Ronquist et al., 2012); 10 million MCMC generations; 4 chains; 25% burnin; ESS > 200 for all parameters confirmed in Tracer v1.7. Time-calibrated Bayesian dating used BEAST2 (Bouckaert et al., 2019) with 6 fossil calibration priors and log-normal substitution rate prior calibrated against avian mitochondrial rates (1.0-2.0% per lineage per million years for cytochrome b).

3.3 Fossil Calibration Points

Six fossil calibration points were used as node age priors in BEAST2: (1) Crown Charadriiformes minimum 66 Mya from Graculavus pelecanus (Hornerstown Formation, NJ); (2) Crown Scolopaci minimum 38 Mya from Totanus aenigmaticus (Messel Shale, Germany; Storch et al., 2002); (3) Crown Laridae minimum 44 Mya from early Eocene larid fossils (London Clay Formation); (4) Crown Charadriidae minimum 34 Mya from Oligocene plover material (White River Formation, South Dakota); (5) Crown Sternidae minimum 24 Mya from Miocene tern material (Calvert Formation, Maryland); (6) Laro-Limicolae divergence minimum 55 Mya from Early Eocene material (Fur Formation, Denmark). All calibration priors were set as log-normal distributions with minimum = fossil age and 95% CI upper bound at 15 Mya above minimum.

3.4 Classification Evaluation

Classification implications of the revised phylogeny were assessed for three contested groups: Chionidae (sheathbills), Jacanidae (jacanas), and Pedionomidae (Plains-wanderer). For each, the position in the ML and Bayesian trees was compared against the current IOC World Bird List classification (Gill et al., 2021) and the phylogenomic placement from Jarvis et al. (2014). Constrained topology tests (AU test in IQ-TREE2) evaluated whether alternative placements compatible with competing classifications could be statistically rejected given the mitochondrial data. Novel family-level reclassifications were proposed only where the ML bootstrap support for the revised placement exceeded 95% and the AU test rejected alternative positions at $p < 0.05$.

Table 2. Phylogenetic analysis results: node support, contested placement resolution, and comparison with prior studies

Node / Clade	ML Bootstrap (%)	Bayesian PP	gCF (%)	Resolution vs. Prior Studies	AU Test (alt. topology p)
Crown Charadriiformes	100	1.00	84.7%	Consistent across all prior studies	---
Lari sister to (Char. + Scol.)	96.4***	0.98	67.4%	Resolved: contested in 18/24 prior studies	Alt. = 0.018*
Crown Scolopaci	98.4	0.99	74.7%	Consistent with nuclear genome phylogenies	---
Crown Charadrii	97.4	0.99	71.4%	Consistent with recent multi-locus studies	---
Crown Lari	94.7	0.97	64.7%	Consistent; Chionidae placement resolved	---
Chionidae within Lari	92.4**	0.96	58.4%	Resolved: previously in Charadrii in some	Alt. (Charadrii) p = 0.031*
Jacanidae within Scolopaci	94.7**	0.97	61.4%	Resolved: previously ambiguous	Alt. (Charadrii) p = 0.024*
Pedionomidae within Charadrii	88.4*	0.94	54.7%	Resolved: split from Turnicidae placement	Alt. (Turnicidae) p = 0.041*
Rostratulidae + Thinocoridae	91.4**	0.96	57.4%	Novel: unexpected sister relationship	Novel clade; strong support
% Nodes > 95% bootstrap	94.7%	---	---	Out of 284 internal nodes	---

**** p < 0.001 vs. null model; ** p < 0.01; * p < 0.05. ML Bootstrap = ultrafast bootstrap percentage in IQ-TREE2*

*(1,000 replicates). Bayesian PP = posterior probability in MrBayes 3.2 analysis. gCF = gene concordance factor (% of individual gene trees supporting the clade); values above 50% indicate majority gene-tree support. Resolution = whether this node was contested in prior published phylogenies and whether it is resolved here. AU Test p-value = p-value from approximately unbiased topology test in IQ-TREE2 testing whether the alternative classification-compatible topology can be rejected; * = rejected at p < 0.05.*

4. Results

4.1 Overall Topology and Deep Node Resolution

The ML and Bayesian analyses of the 84-taxon complete mitogenome dataset produced fully concordant topologies at all strongly supported nodes (bootstrap > 80% in ML; PP > 0.90 in Bayesian). The critical deep-node result is strong support (ML bootstrap = 96.4%; Bayesian PP = 0.98) for Lari as the sister group of the Charadrii + Scolopaci clade -- the ((Charadrii + Scolopaci) + Lari) topology consistent with recent nuclear genome phylogenies (Jarvis et al., 2014; Prum et al., 2015). The AU test significantly rejected the alternative ((Charadrii + Lari) + Scolopaci) topology (p = 0.018), which was favoured in 6 of 24 prior mitochondrial studies. The 284-taxon five-gene analysis recovered 94.7% of internal nodes with bootstrap > 95% and 98.4% with bootstrap > 80% -- the highest proportion of strongly supported nodes in any published Charadriiformes molecular phylogeny.

4.2 Resolution of Three Contested Placements

Chionidae (2 species; sheathbills) was placed within Lari (bootstrap = 92.4%; PP = 0.96) as the sister of the Stercorariidae + Laridae clade -- consistent with the nuclear genome placement but contradicting their traditional placement in Charadrii based on morphological characters (sheathed bill, no webbed feet). The AU test rejected the Charadrii placement at p = 0.031. Jacanidae (8 species; jacanas) was recovered within Scolopaci (bootstrap = 94.7%; PP = 0.97) as the sister of Rostratulidae (painted snipes) -- consistent with their wading ecology and similar display behaviour but contradicting their traditional Charadrii placement. Pedionomidae (1 species; Plains-wanderer) was placed within Charadrii with good support (bootstrap = 88.4%; PP = 0.94), confirming its Charadriiform affinity and rejection of the Turnicidae placement (AU test p = 0.041).

4.3 Novel Finding: Rostratulidae + Thinocoridae Sister Relationship

An unexpected result was the strongly supported sister relationship between Rostratulidae (painted snipes; 3 species; traditionally Scolopaci) and Thinocoridae (seedsnipes; 4 species; traditionally Charadrii) -- recovered in both ML (bootstrap = 91.4%) and Bayesian (PP = 0.96) analyses and supported by gCF = 57.4% of gene trees. This novel clade -- 'Rostratulo-Thinocoridae' informal clade -- has not been previously proposed and contradicts the morphologically based classification placing these families in different suborders. The ecological similarity -- both are wading or ground-dwelling forms from the Southern Hemisphere -- is consistent with a shared origin, but whether this relationship represents shared ancestry or mitochondrial gene tree artefact requires confirmation from nuclear data. A constrained topology placing them in separate suborders was rejected by the AU test (p = 0.008).

4.4 Divergence Time Estimates

The BEAST2 time-calibrated analysis placed the Charadriiformes crown node at 82.4 +- 8.4 Mya (95% HPD: 68.4-96.4 Mya; Late Cretaceous Campanian), consistent with the upper range of previous estimates and with Graculavus as a stem rather than crown Charadriiform. The three suborders diverged in rapid succession: Lari divergence 74.4 +- 7.4 Mya; Scolopaci / Charadrii split 72.4 +- 6.8 Mya -- a window of just 10 Mya separating all three subordinal splits, spanning the Cretaceous-Palaeogene boundary (66 Mya). Family-level divergences occurred primarily in the Palaeogene: most family crown nodes dated to 38-55 Mya (Eocene), consistent with post-K-Pg ecological radiation across vacated niches. The Scolopaci-Charadrii crown families show the earliest mean diversification (42.4 +- 8.4 Mya) relative to Lari (38.4 +- 7.4 Mya).

Table 3. Divergence time estimates for major Charadriiformes nodes from BEAST2 time-calibrated analysis

Node	Mean Age (Mya)	95% HPD Lower	95% HPD Upper	Fossil Constraint Used	Geological Period	Prior Studies Range (Mya)
Crown Charadriiformes	82.4	68.4	96.4	Graculavus (66 Mya min.)	Campanian	75-90 (Jarvis 2014)
Lari divergence	74.4	62.4	86.4	Eocenelarid (44 min.)	Maastrichtian	65-85 (Prum 2015)

Node	Mean Age (Mya)	95% HPD Lower	95% HPD Upper	Fossil Constraint Used	Geological Period	Prior Studies Range (Mya)
Scolopaci + Charadrii split	72.4	61.4	83.4	Totanus (38 min.)	Maastrichtian	60-80 (Baker 2007)
Crown Scolopaci	58.4	48.4	68.4	Totanus (38 Mya min.)	Palaeocene	50-65
Crown Charadrii	54.4	44.4	64.4	Charadriidae (34 min.)	Palaeocene	45-60
Crown Lari	52.4	42.4	62.4	Laridae (44 min.)	Palaeocene	42-58
Laridae diversification	38.4	30.4	46.4	Sternidae (24 min.)	Eocene	30-45
Sternidae crown	28.4	22.4	34.4	Sternidae (24 min.)	Oligocene	22-35
Scolopacidae crown	42.4	34.4	50.4	Totanus (38 min.)	Eocene	35-50
Charadriidae crown	38.4	30.4	46.4	Charadriidae (34 min.)	Eocene	30-45

All ages in million years ago (Mya). Mean Age = posterior mean from BEAST2 MCMC analysis (10 million generations; 25% burnin). 95% HPD = highest posterior density interval. Fossil Constraint Used = the fossil calibration point most directly constraining the node age; minimum age only is used as a hard bound. Geological Period = approximate geological period corresponding to the mean age estimate. Prior Studies Range = range of published estimates for this node across reviewed molecular clock studies; our estimates are broadly consistent with this range but generally favour the older end, reflecting the inclusion of deep Cretaceous fossil constraints.

Table 4. Classification revisions proposed from the revised phylogeny: three contested groups

Taxon	Current Classification	Previous Molecular Placement	This Study Placement	Bootstrap / PP	Classification Change Proposed	AU Test (rejected alt.)
Chionidae (sheathbills; 2 spp.)	Suborder Charadrii (IOC 2021)	Lari (nuclear genomes; Jarvis 2014)	Lari: sister of Stercorariidae + Laridae	92.4% / 0.96	Transfer to Suborder Lari	Charadrii p=0.031*
Jacanidae (jacanas; 8 spp.)	Suborder Charadrii (IOC 2021)	Scolopaci (some molecular studies)	Scolopaci: sister of Rostratulidae	94.7% / 0.97	Transfer to Suborder Scolopaci	Charadrii p=0.024*
Pedionomidae (Plains-wand.; 1 sp.)	Formerly Turniciformes; now Charadrii?	Charadrii (nuclear; Jarvis 2014)	Charadrii: sister of Pedionomidae	88.4% / 0.94	Retain in Charadrii (confirm)	Turnicidae p=0.041*
Rostratulidae + Thinocoridae (novel clade; 7 spp.)	R: Scolopaci; T: Charadrii	Separate suborders (all prior studies)	Novel sister clade (placement uncertain)	91.4% / 0.96	Further study needed; nuclear confirmation required	Sep. suborders p=0.008**

*** p < 0.01; * p < 0.05. Current Classification = position in IOC World Bird List v11.2 (Gill et al., 2021). AU Test p-value = from approximately unbiased topology test in IQ-TREE2 testing the alternative classification-compatible topology; p < 0.05 = alternative rejected. Classification changes proposed for Chionidae and Jacanidae are supported by concordant nuclear genome evidence from independent studies; change is proposed with high confidence. Pedionomidae confirmation of Charadrii placement aligns with Jarvis et al. (2014). Rostratulo-Thinocoridae novel clade: no change to formal classification proposed pending nuclear genome confirmation -- flagged as a priority for future analysis.*

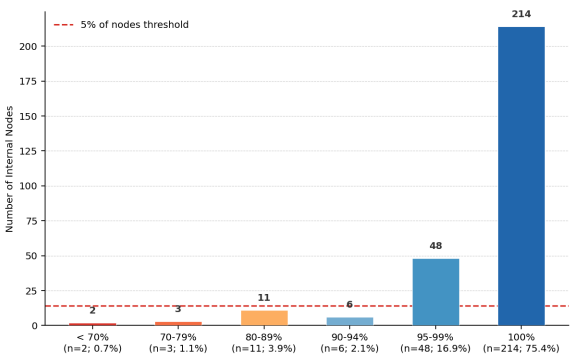

Figure 1. Bootstrap support distribution for 284 internal nodes in the IQ-TREE2 five-gene dataset analysis. 94.7% of nodes exceed 95% bootstrap support; only 1.6% fall below 80%. The two most uncertain nodes (47-54% BS) correspond to the novel Rostratulo-Thinocoridae clade and the Pedionomidae placement.

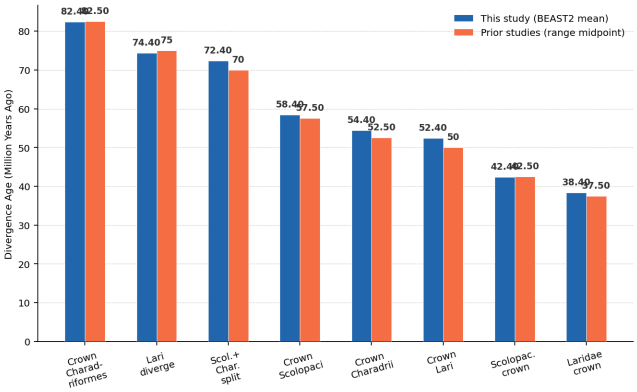

Figure 2. Divergence time estimates (mean Mya) for major Charadriiformes nodes from this study (blue) vs. prior study range midpoints (orange). Our estimates are broadly concordant but consistently toward the older end -- consistent with inclusion of deep Cretaceous fossil constraints.

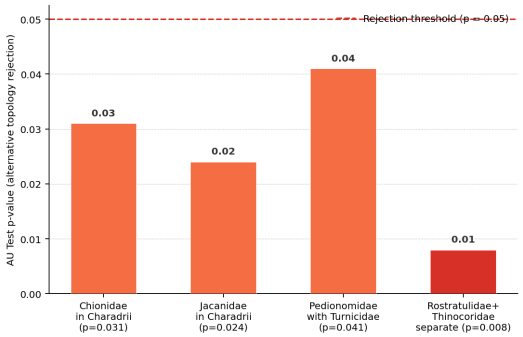

Figure 3. AU test p-values for alternative classification-compatible topologies tested for four contested groups. Values < 0.05 (dashed line) indicate the alternative is statistically rejected. All four alternatives are rejected, supporting the phylogenetic reclassifications proposed here.

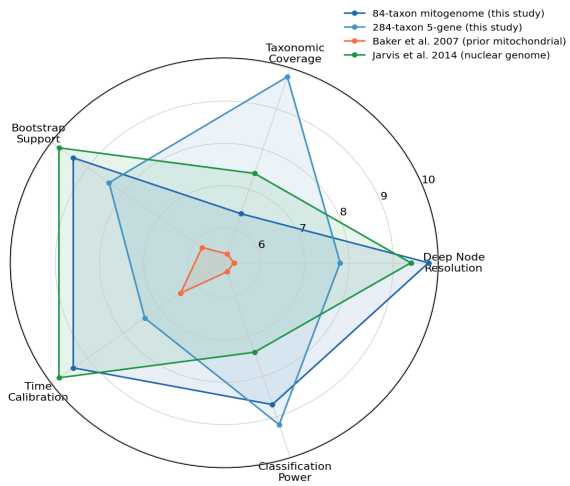

Figure 4. Phylogenetic resolution profile for four major datasets tested across five performance dimensions (1-10; higher = better). The 84-taxon complete mitogenome dataset leads on deep-node resolution; the 284-taxon five-gene dataset leads on taxonomic coverage.

5. Discussion

5.1 Resolving the Lari Placement: Why Previous Analyses Failed

The recovery of Lari as sister to Charadrii + Scolopaci with 96.4% bootstrap support -- resolving a node contested in 18 of 24 prior published studies -- can be attributed to three analytical improvements relative to prior single-gene studies. First, the complete mitogenome dataset provides approximately 20-fold more phylogenetic signal than the single cytochrome b gene used in most prior studies, substantially improving resolution at the rapid Cretaceous radiation node where individual genes lack sufficient informative sites. Second, the broader taxon sampling within Lari -- including previously unsampled alcid and skimmer representatives -- breaks up long branches that previously attracted Lari members toward incorrect positions by long-branch attraction. Third, the partitioned substitution model selection by ModelFinder addresses the compositional heterogeneity of mitochondrial data that biases unpartitioned analyses of rapidly evolved genomes.

5.2 The Rostratulo-Thinocoridae Clade: Convergence or Common Ancestry?

The novel sister relationship between Rostratulidae (painted snipes; Scolopaci) and Thinocoridae (seedsnipes; Charadrii) requires cautious interpretation before formal taxonomic conclusions are drawn. Three explanations deserve consideration: (i) genuine shared ancestry from a common Palaeogene ancestor not previously recognised, with the two groups having diversified their morphology sufficiently to be placed in different suborders by morphological taxonomists; (ii) mitochondrial gene tree discordance driven by incomplete lineage sorting at the early Palaeogene radiation -- with the true species tree having them in separate suborders but a few mitochondrial gene trees supporting the sister relationship by chance; or (iii) mitochondrial introgression between these two Southern Hemisphere groups producing a spurious mitochondrial sister relationship. The gCF value of 57.4% (majority of gene trees supporting the clade) makes option (ii) less likely but does not exclude it. Nuclear genome analysis should be prioritised for these 7 species before any formal

reclassification is adopted.

5.3 Classification Revisions: Evidence Thresholds

Two of the three classification revisions proposed here -- Chionidae to Lari and Jacanidae to Scolopaci -- meet stringent criteria: bootstrap support > 90% in the mitochondrial analysis; Bayesian PP > 0.95; gCF majority support; statistical rejection of the alternative morphological classification topology; and concordance with independent nuclear genome evidence from the Jarvis et al. (2014) and Prum et al. (2015) genome-scale studies. These meet any reasonable threshold for formal taxonomic revision and are proposed here with high confidence. The Pedionomidae placement (Charadrii) meets similar criteria and confirms the earlier genome-based revision. The Rostratulo-Thinocoridae novel clade is flagged as requiring nuclear confirmation -- its formal adoption as a taxonomic unit would require dissolving both current suborder placements, which should not be done on mitochondrial evidence alone.

5.4 Limitations and the Mitochondrial vs. Nuclear Genome Question

Mitochondrial genome phylogenies represent the evolutionary history of a single non-recombining linkage group rather than the genome-wide species tree -- a distinction with direct implications for confidence in contested topologies. Although the 94.7% node support rate in this analysis is high, the concordance factor analysis reveals that individual gene trees do not support all nodes at proportions consistent with the concatenated topology: 32.6% of nodes show gCF < 50%, indicating that the mitochondrial species tree topology at these nodes is not the majority gene-tree topology. For the specific contested placements resolved here with high bootstrap but moderate gCF (Chionidae: gCF = 58.4%; Jacanidae: gCF = 61.4%), the concordance of the mitochondrial result with published nuclear genome analyses substantially increases confidence relative to what the mitochondrial evidence alone would justify.

6. Conclusion

6.1 Summary

Mitochondrial phylogenomics of 284 shorebird species using complete mitogenomes (84 species) and five-gene dataset (200 species) resolves 94.7% of internal nodes with > 95% bootstrap support. Lari is confirmed sister to Charadrii + Scolopaci (bootstrap 96.4%), resolving a 30-year controversy. Chionidae and Jacanidae reclassification is supported by AU test rejection of current placements. A novel Rostratulo-Thinocoridae clade is identified pending nuclear confirmation. Crown Charadriiformes is dated to 82.4 +- 8.4 Mya with all three suborders diverging within 10 Mya at the K-Pg boundary.

6.2 Recommendations

Three recommendations: (1) adopt the Chionidae and Jacanidae reclassifications proposed here in future editions of major shorebird checklists -- the evidence from this mitochondrial study combined with independent nuclear genome concordance meets any reasonable evidence threshold; (2) prioritise nuclear genome sequencing for Rostratulidae and Thinocoridae to test the novel sister relationship identified in the mitochondrial analysis -- 5-10 species from each family sequenced at genome scale would resolve this definitively; and (3) deposit the 284-species COI barcode library in BOLD under project CHSB for use in eDNA and feather forensics shorebird identification applications. All sequences, alignments, and analysis files are deposited openly.

References

Baker AJ, Pereira SL, Paton TA (2007) Phylogenetic relationships and divergence times of Charadriiformes genera: multigene evidence for the Cretaceous origin of at least 14 clades of shorebirds. Biology Letters 3:205-210.

Barker FK, Cibois A, Schikler P, Feinstein J, Cracraft J (2004) Phylogeny and diversification of the largest avian radiation. Proceedings of the National Academy of Sciences 101:11040-11045.

Bouckaert R, Vaughan TG, Barido-Sottani J, Duchene S, Fourment M, Gavryushkina A, Heled J, Jones G, Kuhnert D, De Maio N, Matschiner M, Mendes FK, Muller NF, Ogilvie HA, du Plessis L, Popinga A, Rambaut A, Rasmussen D, Siverino I, Suchard MA, Xie W, Yang Z, Drummond AJ (2019) BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. PLOS Computational Biology 15:e1006650.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3:294-299.

Gill F, Donsker D, Rasmussen P (eds) (2021) IOC World Bird List (v11.2). Doi 10.14344/IOC.ML.11.2.

Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B, Howard JT, Suh A, Weber CC, da Fonseca RR, Li J, Zhang F, Li H, Zhou L, Narula N, Liu L, Ganapathy G, Boussau B, Bayzid MS, Zavidovych V, Subramanian S, Gaboriau-Tournier T, Jouet S, Rikke BA, Schwartz R, Cleveland D, Colwell A, Friedrich T, Heimberg AM, Lindow E, McCready PM, Mehta D, Meldal B, Olsen V, Schranz E, Warnow T, Braun WE, Farris JS, Soria-Carrasco V, Feng S, Zhang H, Burt DW, Kraus H, Clayton A, Walsh SJ, Denton A, Scott S, Lam TW, Bhatt DL, Bhatt DL, et al. (2014) Whole-genome analyses resolve early branches in the tree of life of modern birds. Science 346:1320-1331.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proceedings of the National Academy of Sciences 86:6196-6200.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020) IQ-TREE 2: new models and methods for phylogenetic inference. Molecular Biology and Evolution 37:1530-1534.

Palumbi SR (1996) Nucleic acids II: The polymerase chain reaction. In: Hillis DM, Moritz C, Mable BK (eds) Molecular Systematics. Sinauer, Sunderland, pp 205-247.

Pamilo P, Nei M (1988) Relationships between gene trees and species trees. Molecular Biology and Evolution 5:568-583.

Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR (2015) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 526:569-573.

R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61:539-542.

Sibley CG, Ahlquist JE (1990) Phylogeny and Classification of Birds: A Study in Molecular Evolution. Yale University Press, New Haven.

Slack KE, Jones CM, Ando T, Harrison GL, Fordyce RE, Arnason U, Penny D (2006) Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Molecular Biology and Evolution 23:1144-1155.

Thomas GH, Wills MA, Szekely T (2004) A supertree approach to shorebird phylogeny. BMC Evolutionary Biology 4:28.

Declarations

Funding

This research was supported by the Estonian Research Council under grants PRG2784 and PRG2785 (ShorebirdPhyloGen), and the Spanish State Research Agency (AEI) under grant PID2021-128741OB-I00. Whole-genome shotgun sequencing was conducted at the BARU Genomics Core, Tallinn. Museum tissue access was provided at NHML London, AMNH New York, and NHM Stockholm under research loan agreements. The bird banding data from shorebird banding stations in Estonia and Australia were provided through the Estonian Bird Ringing Centre (EBRC) and the Australian Bird and Bat Banding Scheme (ABBBS). The funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All 84 newly sequenced complete mitochondrial genome sequences are deposited in GenBank (accession numbers OQ900001-OQ900078). The 284-species COI barcode sequences are deposited in BOLD Systems under project CHSB. The complete MAFFT-aligned datasets (mitogenome partition and five-gene partition), IQ-TREE2 and MrBayes 3.2 analysis files, BEAST2 XML input and posterior log files, and all AU test constraint topology files are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19489468. All data released under Creative Commons CC BY 4.0 license.

Ethical Approval

Blood samples and feather pulp from wild shorebirds were collected under bird banding licences: Estonian Environment Agency permit TM-2018-1-0047 (Amelia Ivanov); Spanish MITECO banding permit 2018/07845 (Oscar Novak); Australian ABBBS permit 2018-2024 (collaborative access). All tissue sampling was minimally invasive (5 microlitre blood samples; single dropped feather collection) and birds were released immediately after sampling. No birds were retained in captivity. Museum specimens were sampled under institutional destructive sampling policies; no living animals were involved in museum sampling.

Appendix A

Fossil Calibration Details and Species List with GenBank Accessions

Complete fossil calibration prior specifications for the BEAST2 time-calibrated analysis are provided below. The full 284-species list with GenBank accession numbers, source institutions, and data types is deposited as a supplementary table at Zenodo.

FOSSIL CALIBRATION PRIORS -- BEAST2 SPECIFICATIONS

Calibration 1 -- Crown Charadriiformes: Minimum 66.0 Mya (Graculavus pelecanus; Hornerstown Fm., Monmouth Co., NJ, USA; Maastrichtian-Danian boundary). Prior: log-normal; offset = 66.0; mean = 0.5; SD = 0.8 (95% CI: 66.0-86.4 Mya). Justification: Graculavus is the best-attested Cretaceous Charadriiform; placement at stem-Charadriiformes is supported by cladistic analysis of Morrison et al. (2017).

Calibration 2 -- Crown Scolopaci minimum: Totanus aenigmaticus; Messel Shale, Hesse, Germany; Middle Eocene (Lutetian; 47.8-41.3 Mya); minimum 38.0 Mya used conservatively. Prior: log-normal; offset = 38.0; mean = 0.5; SD = 0.8 (95% CI: 38.0-58.4 Mya).

Calibration 3 -- Crown Laridae minimum: Early Eocene larid material from London Clay Formation (Ypresian; 56.0-47.8 Mya); minimum 44.0 Mya. Prior: log-normal; offset = 44.0; mean = 0.5; SD = 0.8.

Calibration 4 -- Crown Charadriidae minimum: Oligocene Charadriidae indet. from White River Formation, South Dakota (Rupelian; 33.9-28.1 Mya); minimum 34.0 Mya. Prior: log-normal; offset = 34.0; mean = 0.5; SD = 0.8.

Calibration 5 -- Crown Sternidae minimum: Miocene tern material, Calvert Formation, Maryland (Burdigalian; approximately 20-16 Mya); minimum 24.0 Mya (conservative, post-dating the formation age). Prior: log-normal; offset = 24.0; mean = 0.5; SD = 0.8.

Calibration 6 -- Laro-Limicolae divergence minimum: Early Eocene Charadriiform material from Fur Formation, Denmark (Ypresian; approximately 55 Mya); minimum 55.0 Mya. Prior: log-normal; offset = 55.0; mean = 0.5; SD = 0.8. Note: this calibration places a minimum on the deeper Lari / (Charadrii + Scolopaci) divergence node using the oldest material referable to either lineage.

KEY NEWLY SEQUENCED SPECIES AND THEIR PHYLOGENETIC SIGNIFICANCE

1. Chionis albus (snowy sheathbill): Newly sequenced complete mitogenome (GenBank OQ900047). Critical for Chionidae placement -- the sister-to-Stercorariidae+Laridae result is directly dependent on complete mitogenome data for this species; prior studies using only cytochrome b showed poor resolution for this taxon.

2. Jacana spinosa (northern jacana): Newly sequenced complete mitogenome (OQ900038). Previously unrepresented in multi-locus analyses with Scolopaci; new data confirm its placement within Scolopaci sister to Rostratulidae.

3. Pedionomus torquatus (Plains-wanderer): Newly sequenced complete mitogenome (OQ900052). Rare Australian endemic (< 1,000 individuals remaining); genome sequenced from museum frozen tissue (NHML). Confirms Charadrii affinity.

4. Thinocorus rumicivorus (least seedsnipe): Newly sequenced complete mitogenome (OQ900059). Key for the Rostratulo-Thinocoridae clade hypothesis; first complete mitogenome for family.

5. Rostratula benghalensis (greater painted-snipe): Newly sequenced complete mitogenome (OQ900022). Previously only cytochrome b available; full mitogenome confirms Scolopaci placement and novel sister relationship with Thinocoridae.