



Alt

PDF



Nature

Nature Publishing Group

Dense sampling of bird diversity increases power of comparative genomics

Shaohong Feng, Josefin Stiller, [...], and Guojie Zhang

[Additional article information](#)

Associated Data

- ▶ [Supplementary Materials](#)
- ▶ [Data Availability Statement](#)

Abstract

Whole-genome sequencing projects are increasingly populating the tree of life and characterizing biodiversity^{1–4}. Sparse taxon sampling has previously been proposed to confound phylogenetic inference⁵, and captures only a fraction of the genomic diversity. Here we report a substantial step towards the dense representation of avian phylogenetic and molecular diversity, by analysing 363 genomes from 92.4% of bird

families—including 267 newly sequenced genomes produced for phase II of the Bird 10,000 Genomes (B10K) Project. We use this comparative genome dataset in combination with a pipeline that leverages a reference-free whole-genome alignment to identify orthologous regions in greater numbers than has previously been possible and to recognize genomic novelties in particular bird lineages. The densely sampled alignment provides a single-base-pair map of selection, has more than doubled the fraction of bases that are confidently predicted to be under conservation and reveals extensive patterns of weak selection in predominantly non-coding DNA. Our results demonstrate that increasing the diversity of genomes used in comparative studies can reveal more shared and lineage-specific variation, and improve the investigation of genomic characteristics. We anticipate that this genomic resource will offer new perspectives on evolutionary processes in cross-species comparative analyses and assist in efforts to conserve species.

Subject terms: Evolutionary genetics, Comparative genomics

Main

Comparative genomics is rapidly growing, fuelled by the advancement of sequencing technologies. Many large-scale initiatives have been proposed with a core mission of producing genomes for hundreds of species, representing the phylogenetic diversity of particular taxa^{6–8}.

