

A high-quality reference genome for the Ural Owl (*Strix uralensis*) enables investigations of cell cultures as a genomic resource for endangered species

Ioannis Chrysostomakis¹, Annika Mozer¹, Camilla Bruno Di-Nizo¹, Dominik Fischer², Nafiseh Sargheini³, Laura von der Mark¹, Bruno Huettel³, Jonas J. Astrin¹, Till Töpfer¹, and Astrid Böhne^{1,*}

¹Leibniz Institute for the Analysis of Biodiversity Change, Museum Koenig Bonn, 53113 Bonn, Germany

²Zoo Wuppertal, 42117 Wuppertal, Germany

³Max Planck Genome-Centre Cologne, Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany

*Correspondence address. Astrid Böhne, Centre for Molecular Biodiversity Research, Leibniz Institute for the Analysis of Biodiversity Change, Museum Koenig Bonn, Adenauerallee 127, 53113 Bonn, Germany. E-mail: a.boehne@leibniz-lib.de

Abstract

Background: Reference genomes have a wide range of applications, but we are far from a complete genomic picture of the tree of life. We here contribute another piece to the puzzle by providing a high-quality reference genome for the Ural Owl (*Strix uralensis*), a species of conservation concern and efforts affected by habitat destruction and climate change.

Results: We generated a reference genome assembly for the Ural Owl based on high-fidelity (HiFi) long reads and chromosome conformation capture (Hi-C) data that figures as one of the best avian genome assemblies currently available (BUSCO completeness of 99.94%). The primary assembly had a size of 1.38 Gb with a scaffold N50 of 90.1 Mb, while the alternative assembly had a size of 1.3 Gb and a scaffold N50 of 17.0 Mb. We show an exceptionally high repeat content (21.07%) that is different from those of other bird taxa with repeat extensions. We confirm a *Strix*-characteristic chromosomal fusion and support the observation that bird microchromosomes have a higher density of genes, associated with a reduction in gene length due to shorter introns. An analysis of gene content provides evidence of changes in the keratin gene repertoire, as well as modifications of metabolism genes of owls. This opens an avenue of research if this is related to flight adaptations. The population size history of the Ural Owl decreased over long periods of time, with increases during the Eemian interglacial and a stable size during the last glacial period. Ever since, it has been declining to its currently lowest effective population size. We also investigated the cell culture of progressive passages as a tool for genetic resources. Karyotyping of passages confirmed no large variants, while a single-nucleotide polymorphism analysis revealed a low presence of short variants across cell passages.

Conclusions: The established reference genome is a valuable resource for ongoing conservation efforts but also for (avian) comparative genomics research. Further research is needed to determine whether cell culture passages can be safely used in genomic research.

Keywords: *Strix uralensis*, Strigidae, karyotyping, genome sequence, genome annotation, cell culture, SNP, variant

Background

High-quality reference genomes are rapidly becoming available for many branches of the tree of life [1, 2]. These data are now increasingly used for comparative genomic studies on large evolutionary timescales trying to link phenotypes to genotypes [3]. However, even in genomically and traditionally well-studied groups such as birds, most lineages still lack high-quality reference genome assemblies that would allow for detailed studies of genome evolution.

Typical avian karyotypes are composed of macro- and microchromosomes (but see [4, 5]). Compared to macrochromosomes, which are typically between 30 and 250 megabase pairs (Mb) in size, microchromosomes have an average size of 12 Mb, although microchromosomes as small as 3.4 Mb have been observed [6, 7]. Despite recent efforts to characterise avian genomes and understand their karyotype evolution, less than 10% of all known bird species have a characterised karyotype [8]. The diploid number of about half of these varies between 78 and 82 chromosomes [1]. Regarding the family Strigiformes (owls), karyotype

information is available for 13% of species [8]. Interestingly, microchromosomes encode half of the genes in birds, although they account for only about a quarter of the genome sequence [6, 9]. Moreover, the mutation rate of microchromosomes is significantly higher than that of macrochromosomes [10]. Therefore, avian karyotypes, genome structure, and especially the microchromosomes deserve more cytogenetic and molecular attention.

To this aim, we here provide a first high-quality reference genome for the Ural Owl (*Strix uralensis*, NCBI taxonomy ID 36305). This species is one of the largest Eurasian owls, inhabiting the Palaearctic lowlands up to the treeline, mainly in the taiga forest belt over a large uninterrupted range from Scandinavia through Siberia to Sakhalin and the Japanese islands. It also occurs in geographically isolated, mixed, and deciduous forests of south-eastern and central Europe (southern Germany, Czech Republic, Austria, Slovenia, and Poland; partly supported by reintroductions). So far, 11 subspecies have been described from its vast distribution based on differences in size and colouration [11], although not all of these have been widely accepted [12].

Received: March 31, 2025. Revised: June 6, 2025. Accepted: August 21, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of GigaScience. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

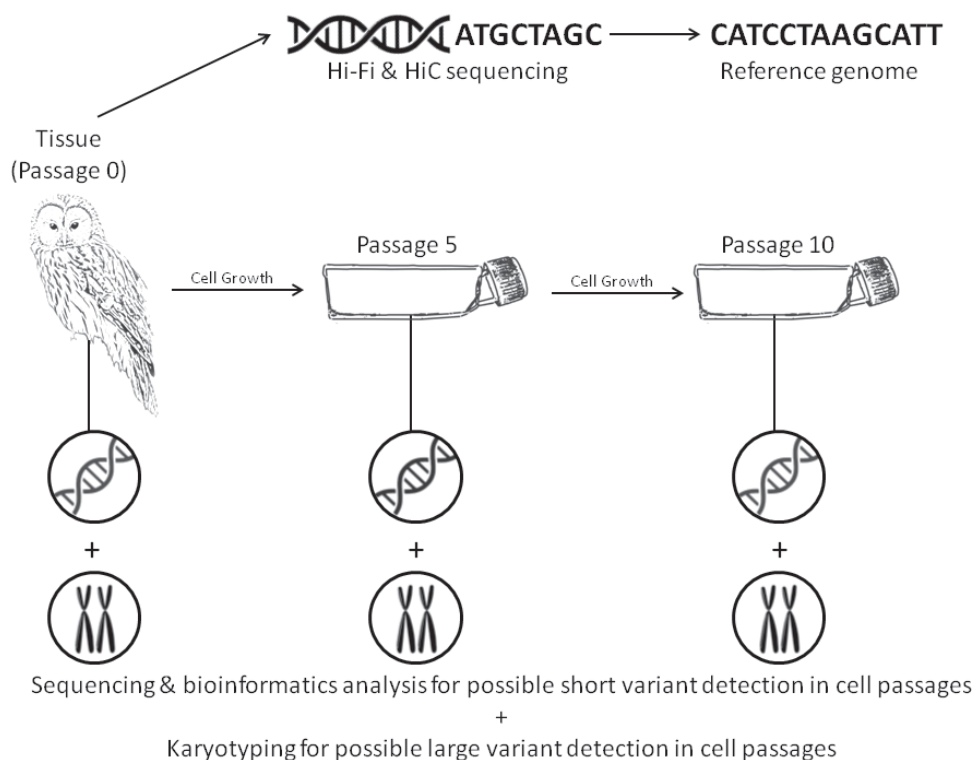


Figure 1: Reference genome and cell passage variant detection workflow. Tissue from a male Ural Owl (*Strix uralensis*) is extracted and sequenced, assembled, and annotated to provide a reference genome. Additionally, a cell culture is established from the primary tissue. From passage 0 (primary tissue), passage 5 (3 independent replicates), and passage 10 (4 independent replicates), cells are harvested for short-read sequencing and karyotyping.

Furthermore, the molecular data at hand (i.e., mitochondrial and nuclear marker genes) do not support morphology-based taxonomic distinctions [13].

Ural Owls are nocturnal hunters of small mammals and birds and usually stay in their territories throughout the year [12, 14]. As the Ural Owl is sedentary and nests in hollow stumps or tree holes [15, 16], it is affected by ecosystem degradation [17]. Nesting sites have been reduced by intensive logging activities, agricultural use, and forestry management [11]. While globally still considered under the IUCN Red List category “Least Concern,” *S. uralensis* went extinct in Austria, southern Germany, and the Czech Republic in the past century, mainly due to direct persecution [18–20]. Successful reintroductions have taken place in these countries (e.g., [18–20]). These central European reintroductions have restored gene flow between the remaining Alpine and European populations [13, 21]. The Ural Owl will likely further be affected by climate change, potentially shifting its range to more northern regions [22] and altering breeding times [23]. Correspondingly, the Ural Owl is a species that benefits from conservation measures in the European Union under the EU Birds Directive [24] and Nature Habitats Directive [25]. It is also part of the Bern Convention [26]. International trade of all Strigiformes is regulated by the Convention on International Trade in Endangered Species of Wild Fauna and Flora [27].

Cryobanking, defined as the preservation of viable cells and tissues at ultracold temperatures, typically using liquid nitrogen, is considered paramount in preserving the genetic variability of species, especially those facing population decline, such as the Ural Owl, to ensure population health and persistence [28, 29]. Although some instances have been reported where long-term cell culture generated genetic instability and heteroploidy [30, 31], it

is still unclear how frequent such a phenomenon is and at which stage of cell cultivation it occurs.

Herein, we generated a reference genome for the Ural Owl as a genomic resource to facilitate further research on this species and on Strigidae more generally. We assessed the genome assembly quality and provide a first analysis of its gene content. As a species of potential conservation concern and as a proof of principle, we assessed the application of cell culture to produce sufficient DNA in terms of quantity and quality to allow genomics for species with limited biological material. We investigated mutation as a function of passage number (i.e., the transfer of cells from vessel to vessel). To this end, we obtained a cell culture from the same individual that was genome-sequenced and cultivated the cell lines until passage 10 and subsequently sequenced replicates of passages 5 and 10 (Fig. 1).

Data Description

In order to provide valuable genomic resources to the scientific community studying avian ecology and phylogenomics and to investigate the potential of lab-grown cells for use in DNA sequencing, skin cells were harvested from a 10-year-old, recently deceased male Ural Owl individual. The skin samples were originally frozen at -80°C and later grown in an appropriate medium and used for DNA-sequencing. We performed PACBIO long-read sequencing of muscle tissue, which produces high-quality, long DNA fragments. We used cultured cells for Hi-C sequencing, which allowed us to estimate the physical proximity of DNA molecules inside the cell to create a bird genome assembly with the highest gene completeness score to date. Next, we grew the harvested skin cells for multiple generations to understand whether this process

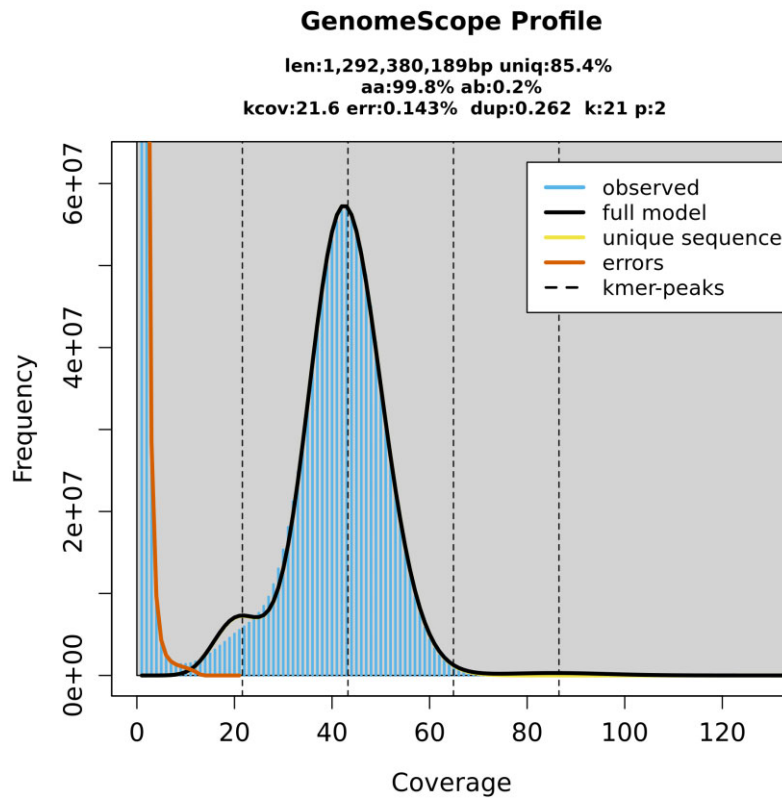


Figure 2: The k-mer genome profile of *Strix uralensis* generated from PacBio HiFi reads with GenomeScope2. The y-axis shows the k-mer counts, and the x-axis shows sequencing depth. The first peak corresponds to heterozygous k-mers and the second larger peak to homozygous k-mers with a coverage of $\sim 42\times$.

Table 1: Assembly statistics of the primary and alternate genome assembly of *Strix uralensis*

Assembly statistics	Primary	Alternate
Assembly size [bp]	1,381,008,983	1,262,176,999
GC content [%]	42.77	42.81
Contigs	512	15,615
N50	90,173,155	17,018,198
L50	6	18
L90	28	8,171
Ns per 100 kb	2.94	68.29
Merqury error (%)	3.82794e-07	1.80919e-06
Merqury QV score	64.17	57.43
Complete BUSCOs [%]	99.94	70.82

causes damage to chromosome structure and the accumulation of DNA mutations. In the future, these data can be used to study avian phylogenomics and diversity as well as further understand the unique traits of owls. All sequence data of this study can be accessed from INSDC under the BioProject ID PRJNA1212906. Processed data are available from Zenodo [32].

Analyses

Read quality control and estimation of genome size and heterozygosity

After quality control, filtering, and decontamination, the final set of HiFi reads used was composed of 5,078,732 reads with a total length of ~ 58 Gb, and the Hi-C reads used were composed of 79.7 million reads with a total length of ~ 20 Gb.

Using a k-mer size of 21, GenomeScope was able to predict a genome size of 1,292,799,460 bp, a repeat length of 188,615,362 bp, a heterozygosity of 0.2% (this would translate to two heterozygous sites per 1 kb, a commonly reported heterozygosity indicator for birds), and a read error rate of 0.14% (Supplementary Table S1). Smudgeplot and GenomeScope both verified the diploid status of the individual (Supplementary Fig. S1; Fig. 2). The genome did not reveal any large runs of homozygosity (ROH).

Reference genome

The optimal assembly was created with Hifiasm parameters “-12 -n-weight 5 -n-perturb 50000 -f-perturb 0.5 -D 10 -N 150 -s 0.4” (Supplementary Table S1).

Genome quality metrics

We could place 93.6% of assembled scaffolded genome sequence data into 41 chromosomes, which is consistent with the karyotype of the species (Table 1, Fig. 3). We also detected no contamination as all scaffolds aligned to sequences of other avian genomes (Supplementary Fig. S2). Our Hi-C contact map further supported the high contiguity of the primary assembly by showing no remaining conflicts and little to no scaffolds with strong contacts to nonrepeat regions (Fig. 4, Supplementary Figs. S3(?PMU ?) and S4).

The Merqury quality value (QV) score, which is the proportion of the assembly sequence supported by HiFi reads, was estimated for both haplomes. We obtained a score of 64.2 (equivalent to an error probability of 3.8238e-07%) for the primary and 57.4 (equivalent to an error probability of 1.80919e-06%) for the alternate assembly (Supplementary Table S2). We also found a

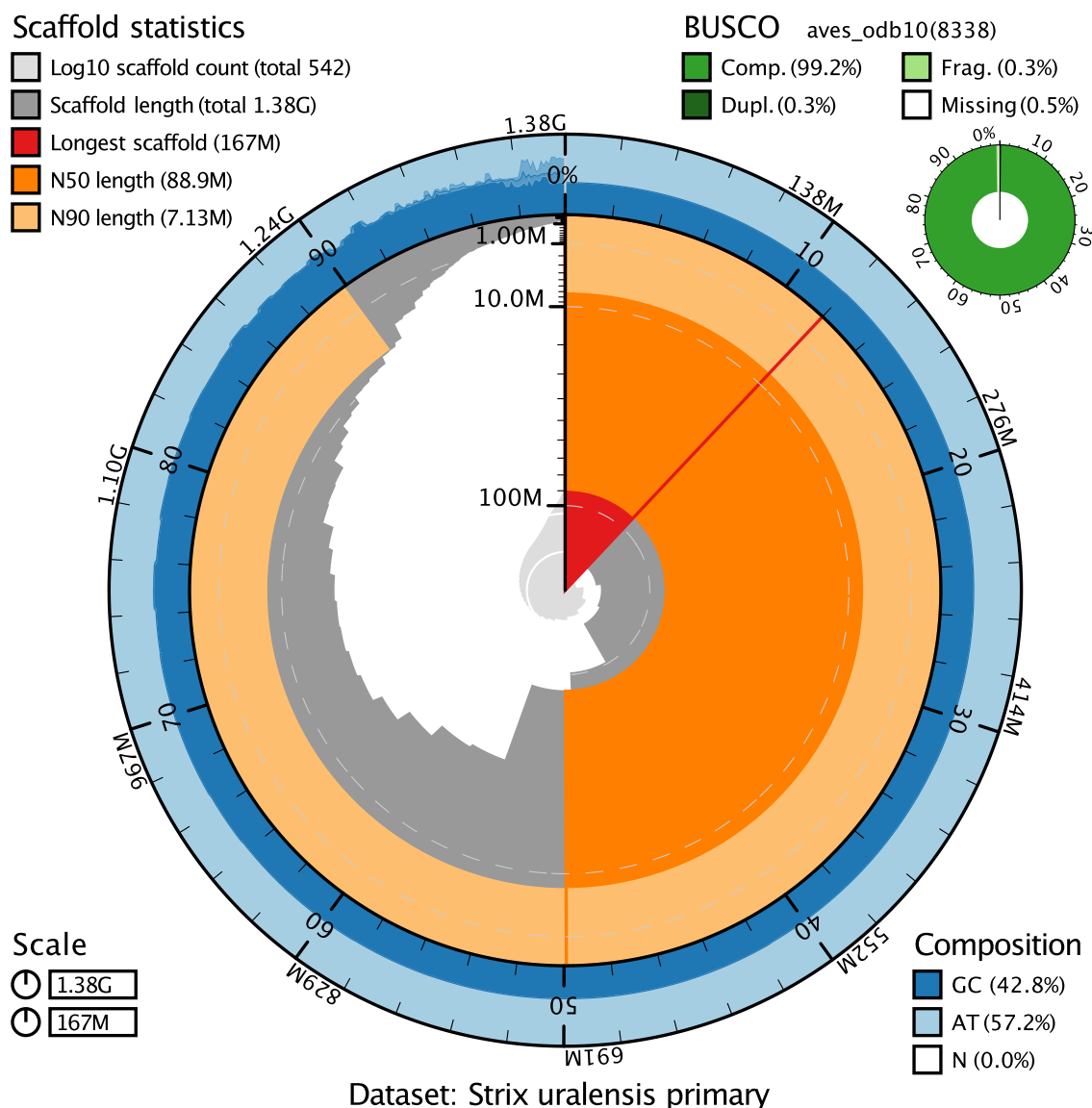


Figure 3: Snail plot summary of assembly statistics for *Strix uralensis* primary assembly. The main plot is divided into 1,000 size-ordered bins around the circumference, with each bin representing 0.1% of the 1,381,000,783 bp assembly. The distribution of sequence lengths is shown in dark grey, with the plot radius scaled to the longest sequence present in the assembly (166,530,430 bp, shown in red). Orange and pale orange arcs show the N50 and N90 sequence lengths (88,922,949 and 7,132,230 bp), respectively. The pale grey spiral shows the cumulative sequence count on a log scale, with white scale lines showing successive orders of magnitude. The blue and pale blue area around the outside of the plot shows the distribution of GC, AT, and N percentages in the same bins as the inner plot. A summary of complete, fragmented, duplicated, and missing BUSCO genes in the aves_odb10 set is shown in the top right.

completeness score of 98.36% for the primary and a combined 99.81% for the two haplomes, representing the fraction of high-quality k -mers from the reads present in the assembly. This further supports the completeness and accuracy of the assembly (Table 1, Supplementary Table S2).

Aligning the PacBio HiFi, and Illumina Hi-C reads to both haplomes revealed comparable coverage levels (primary: 41.78 ± 12.18 , 14.18 ± 91.47 -fold, respectively; alternate: 34.15 ± 23.24 , 12.10 ± 107.37 , respectively), mapping rates (primary: 99.85%, 99.92%, respectively; alternate: 80.52%, 86.6%, respectively), and mapping quality scores (primary: 36.93, 28.90, respectively; alternate: 28.11, 8.2, respectively). These results further indicate that the assembly is well phased with a minimal amount of assembly bias and errors (Supplementary Table S3).

Genome repeat and gene annotation

Repeat landscapes depict the clustering of transposable elements (TEs) in relation to their Kimura substitution rates, which measures the divergence of TEs from their respective consensus sequence. Lower Kimura substitution rates indicate recent transposition events, while higher rates suggest older events. From the landscape of the primary haplome (Fig. 5, Supplementary Figs. S5 and S6), a strong signal for a recent repeat expansion of long interspersed nuclear elements (LINEs) and an even more recent expansion of long terminal repeat (LTR) retrotransposons, as well as a well-maintained large number of older repeats, are visible. This might support a repeat expansion in the Ural Owl or the genus *Strix*. A third and older expansion is dominated by LINEs and unknown repeats, suggesting that they are a new or unique feature

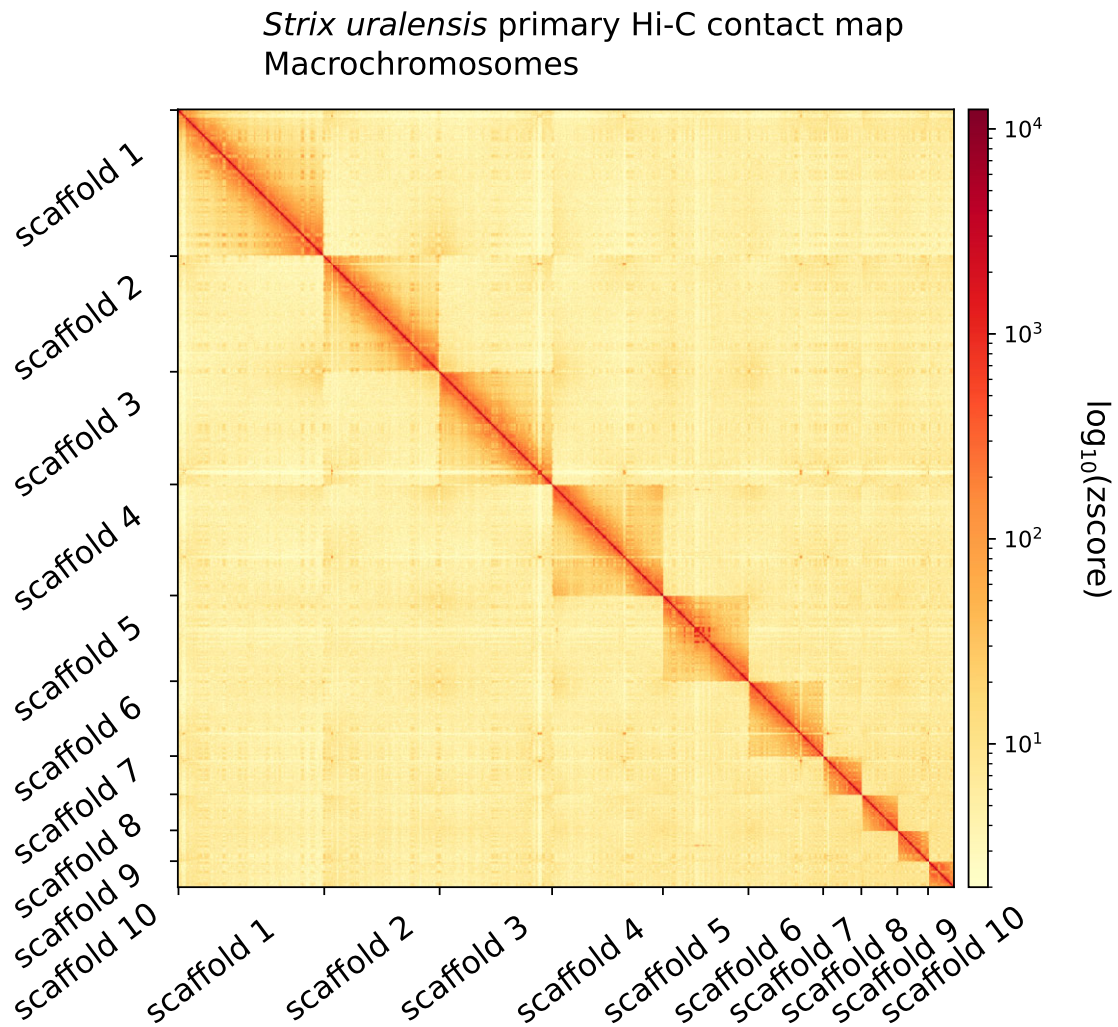


Figure 4: *Strix uralensis* primary haplome Hi-C contact map showing spatial interactions between the ten largest chromosomes. Chromosomes are ordered by size from left to right and from top to bottom. The red diagonal corresponds to intrachromosomal contacts and depicts chromosome boundaries. The frequency of contacts is shown on a logarithmic heatmap scale. Plot generated with HiCExplorer.

of *Strix*, and a reference or consensus might not yet exist in the reference databases.

Gene annotation

For the primary haplome, we were able to annotate a total of 17,977 protein-coding genes, which cover ~33.6% of the total size of the assembly (Supplementary Table S4). We detected 182,313 exons and 164,373 introns. Compared to the Swiss-Prot and UniProt databases, we were able to match 16,461 and 17,511 of our genes to annotations, respectively (Supplementary Table S5).

We next investigated gene distribution along the genome. Using a 30Mb cutoff [6, 7], we identified 10 macrochromosomes and 31 microchromosomes based on our assembly (Fig. 4 and Supplementary Fig. S7). Despite their size, microchromosomes had a higher gene density than macrochromosomes. While there are comparatively more genes on microchromosomes, these genes are shorter than those on macrochromosomes, mainly due to shorter introns (Supplementary Fig. S7).

To shed light on the genome annotation content, we compared the Ural Owl genome to other high-quality genomes of the Aves lineage, including several owl species. Gene expansion (gain) and contraction (loss) among our selected species found 316 gene fam-

ily gains in the Ural Owl and 207 losses, 168 of which were, presumably, completely lost and thus have no representative in the Ural Owl genome assembly (Fig. 6).

Additionally, we found 81 gene families unique to the Ural Owl that do not have orthologs in the other species (Supplementary Table S5). We further found that genomes of lower quality, such as those of the Ferruginous Pygmy Owl (*Glaucidium brasilianum*) and the Black-and-white Owl (*Strix nigrolineata*), had more gene losses, which are hence probably not biologically true but represent technical limitations. A Gene Ontology (GO) term analysis of the genes unique to the Ural Owl revealed many interesting gene families that, due to the low quality of the Black-and-white Owl genome, might also be interpreted as partially representing the *Strix* genus (Supplementary Fig. S8). Among these categories, we note several GO terms relevant to characteristic traits of the Ural Owl, namely, its adaptation to dim-light conditions and a sedentary and predatory hunting strategy. The “animal organ morphogenesis” parent GO term groups the child GO terms “eye development,” “sensory organ development,” “neurogenesis,” and “heart development,” all of which point to adaptations of *Strix* to their environment or lifestyle.

Next, we investigated gene gain and loss at nodes that are supported by more than one reference genome, which would make

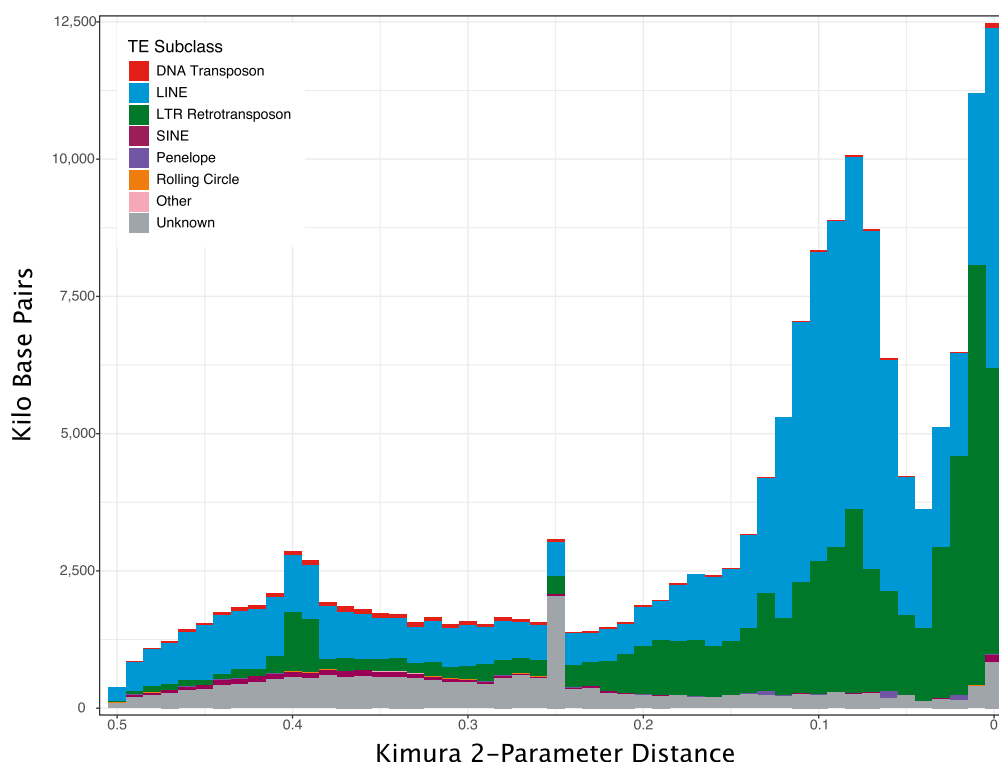


Figure 5: *Styrax uralensis* primary haplome repeat landscape. The x-axis shows the Kimura substitution of detected repeat categories and the y-axis the number of repeats detected for each TE family in kilobase pairs. Detected subclasses are colour-coded as indicated in the inset. The genome assembly was masked using EarlGrey.

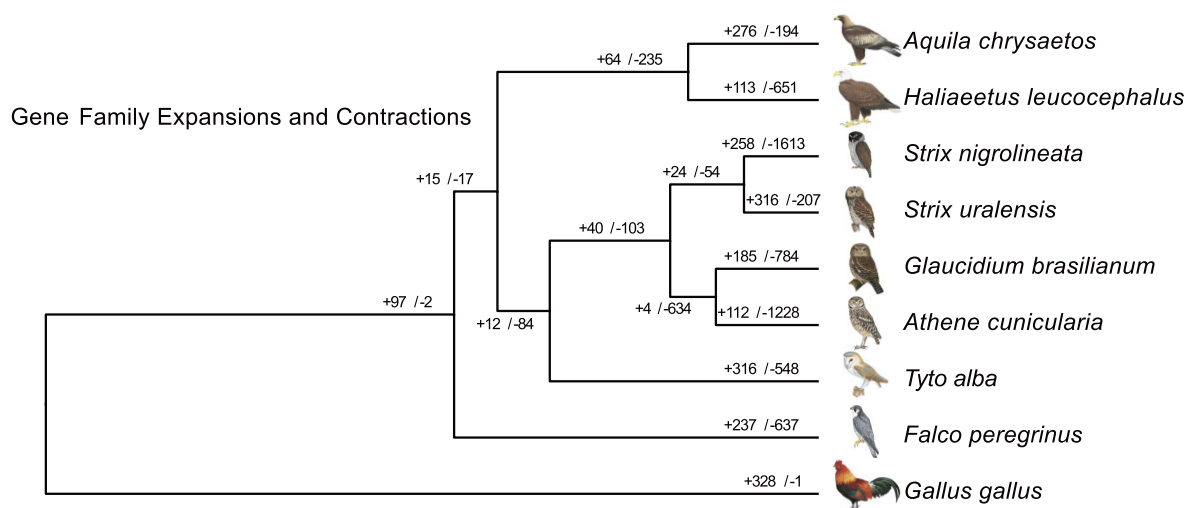


Figure 6: Ultrametric phylogenetic tree of selected Neoaves species and *Styrax uralensis*. Numbers indicate gene family expansions (+) and contractions (-). Bird drawings from Birds of the World [33].

them more robust and at the same time informative about clade-specific genomic changes.

We observed 15 gene family gains and 17 losses in the last common ancestor of Strigiformes and Accipitriformes (hawks, eagles, vultures, kites), both characterised by a predatory lifestyle. Overarching GO terms among the gained genes included “behavior,” metabolic, cellular, and developmental processes. Notably, the child GO terms contained many terms related to general and cellular metabolism (e.g., “ATP metabolic process,” “carbohydrate derivative metabolic process,” “cellular lipid catabolic process,” “cellular lipid metabolic process”). We identified three gains in

keratin genes (feather and scale keratin), two gains related to histones/histone modification and two gains related to skeletal muscle functioning (BEST3, CKB).

The gene losses comprised several mitochondrial genes, which we attribute to lower quality of mitochondrial gene annotation of the used genomes since contrastingly to the ortholog-based results, we could annotate 36 of 37 mitochondrial genes in our assembly.

The other gene losses concerned uncharacterized gene families as well as a ribonucleoprotein (IMP4), the claudin gene family encoding for tight junction proteins, and a DNA polymerase.

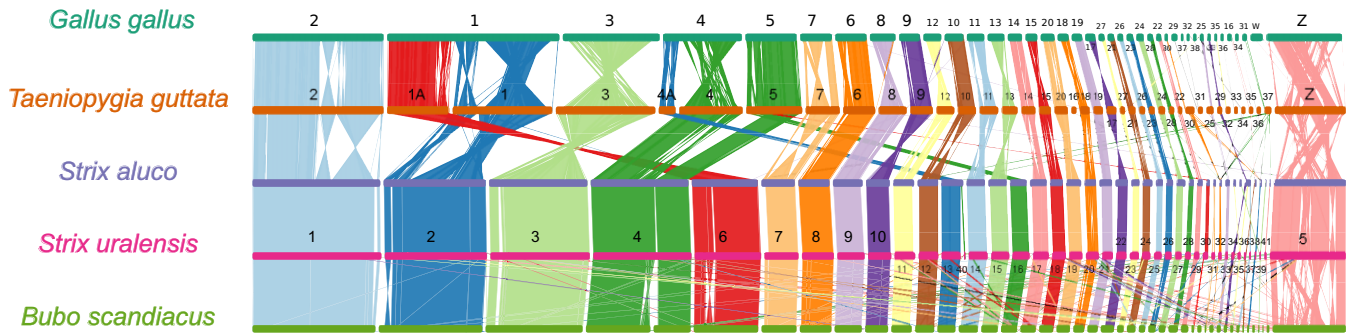


Figure 7: Chromosome-scale synteny analysis. Synteny of chromosomes of *Taeniopygia guttata*, *Gallus gallus*, *Strix aluco*, and *Bubo scandiacus* compared to the newly sequenced *Strix uralensis*. Syntenic regions among the species are indicated with a unique colour. Assignment of sex chromosomes was based on the *S. aluco* genome annotation. The chromosomes of all species have been reordered to highlight synteny relationships to *S. uralensis*. Plot made with NGenomeSynt.

We identified 12 gene gains and 84 losses reconstructed for the common ancestor of owls. We again found an expansion of the keratin gene repertoire (gain of one keratin and one scale-keratin-like gene). GO parental terms of gains pointed again to metabolic changes but also those associated with the immune system. The gains contained also an olfactory receptor. The much more numerous losses were associated even at the higher level, with many different GO categories again often related to metabolism (e.g., “regulation of amide metabolic process,” “pyridine-containing compound metabolic process”).

Chromosome-scale syntenies

Synteny with the chromosome-level assemblies of *Strix aluco* and *Bubo scandiacus* identified the Z chromosome of our male individual. It is the fifth largest chromosome in the Ural Owl assembly. The synteny between the two *Strix* genomes shows no major syntenic differences (Fig. 7). This is also mostly true in the comparison to the Snowy Owl, with the exception of the Z chromosome, which shows some internal rearrangements compared to the two *Strix* species. Whether this is caused by assembly quality and accuracy remains to be investigated. Compared to chicken and Zebra Finch, however, we identified several large-scale changes. We detected a fusion of chromosomes 5 and 6 of the Snowy Owl (corresponding to parts of chicken chromosomes 4 and 5 and Zebra Finch chromosome 4 and parts of 5; Fig. 7 and Supplementary Table S6) into chromosome 4 of the two *Strix* assemblies. This is supported by previous cytogenetic analyses [34]. The remaining part of chicken chromosome 4 corresponds to Ural Owl chromosome 13, Snowy Owl chromosome 12, and Zebra Finch 4a. The remaining part of chicken and Zebra Finch chromosome 5 corresponds to two chromosomes in the Ural Owl (chromosomes 16 and 30). As in the Zebra Finch, chicken chromosome one corresponds to two chromosomes in the Snowy Owl (chromosomes 2 and 6). For the remaining chromosomes of our Ural Owl assembly, we could mostly identify 1:1 relationships with chromosomes of Snowy Owl, chicken, and Zebra Finch with the exception of the Ural Owl microchromosomes 31, 35, 39, and 41, for which we could not unambiguously identify a corresponding chicken or Zebra Finch chromosome but homology to Snowy Owl scaffolds and Tawny Owl chromosomes.

Demographic history

The demographic history of the Ural Owl derived from our genome assembly appears to have a complex relationship with glacial and interglacial periods. The effective population was predicted to have decreased until around the Holstein interglacial period

(3.74×10^5 – 4.24×10^5 years ago [ya]), when its population size stabilized but remained low during the Saalian glacial period (4 – 1.3×10^5 ya) and began to increase as the Eemian interglacial period (1.3 – 1.15×10^5 ya) began to emerge. It continued to increase and reached a plateau during the last glacial period (Weichselian glaciation, 1.15 – 0.117×10^5 ya). Before the end of the last glacial period, at around 0.3×10^5 ya, the Ural Owl population began to decrease until it reached the current lowest effective population size (Fig. 8).

Variation analysis over progressive cell passages

Karyotype confirms chromosome numbers and reveals no large variants caused by passaging

Chromosomal analyses detected $2n = 82$ in both passages 5 and 10, corroborating the diploid chromosome number described for the Ural Owl previously (subspecies *S. uralensis uralensis* and *S. u. japonica*, [35]). No large-scale chromosomal rearrangements were observed between both passages (Fig. 9).

Short-read variants

After quality filtering, we identified 885,159 variant sites (in the following referred to as single-nucleotide polymorphisms [SNPs]) (Supplementary Fig. S9). Out of these, most (i.e., 670,463 SNPs) were fixed variant sites across all samples and hence mostly represented heterozygous sites of the individual, which are represented with just one of the two alleles in the reference genome or sites that had a wrong allele in the reference assembly.

The remaining 214,696 SNPs varied across samples, indicative of potential mutations, and were analysed in the following.

A comparison of all SNPs across all samples revealed that variant amount and type differed. The biggest differences resulted from SNPs called from the HiFi data as well as passages 5.1 and 10.1, which appear to have more SNPs than the other passages. Most of these are heterozygous first-alternate sites (Supplementary Fig. S9, Table 2), and for many of those, we suspect that they are false heterozygous calls rather than true mutations. To investigate this pattern further, we inspected genotype quality and depth focusing on sites with a genotype in a sample not found in any other sample (i.e., private sites) compared to the same metrics at all other sites of that individual (i.e., common variants and conserved heterozygous sites). This analysis revealed that all passages had similar median depth per variant site (DP $\sim 26.78 \pm 6.55$) and genotype quality (GQ ~ 99), suggesting rather consistent data quality across samples (Table 2). It further showed that median depth and quality of private SNPs consistently had

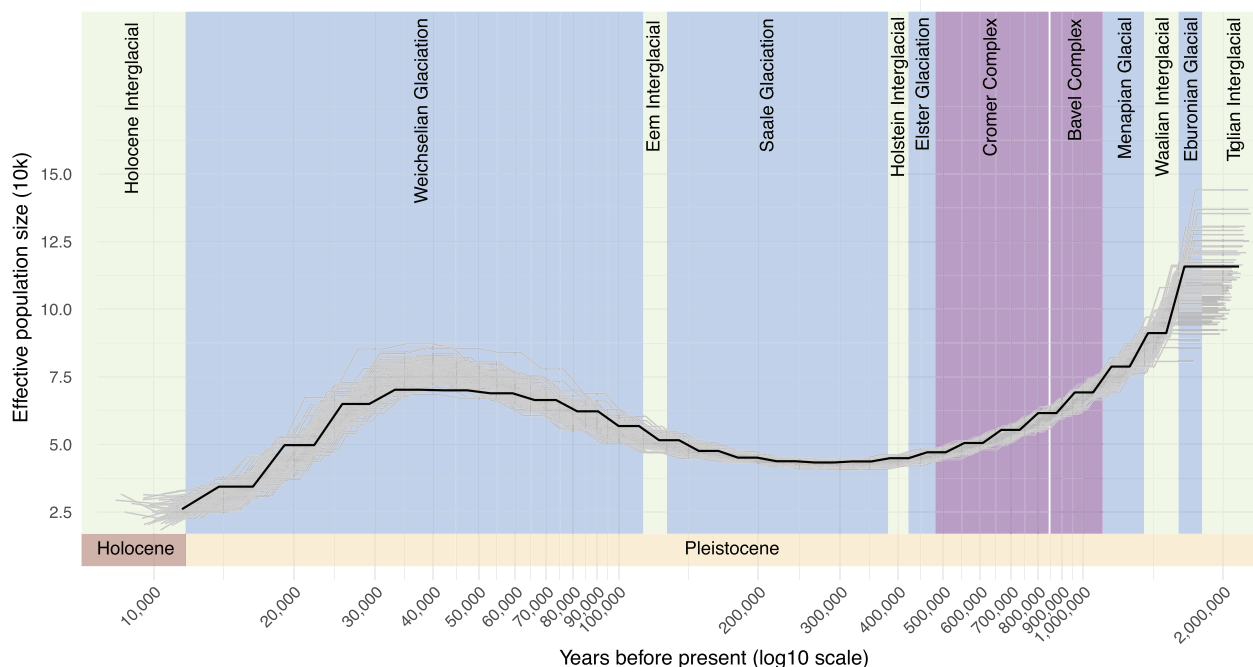


Figure 8: Inferred demographic history of *Strix uralensis*. The plot shows a pairwise sequentially Markovian coalescent (PSMC) analysis based on the primary genome assembly. The x-axis shows years before present (ya) on a logarithmic scale, and the y-axis shows the estimated effective population size. Bootstrap results are shown in light grey.

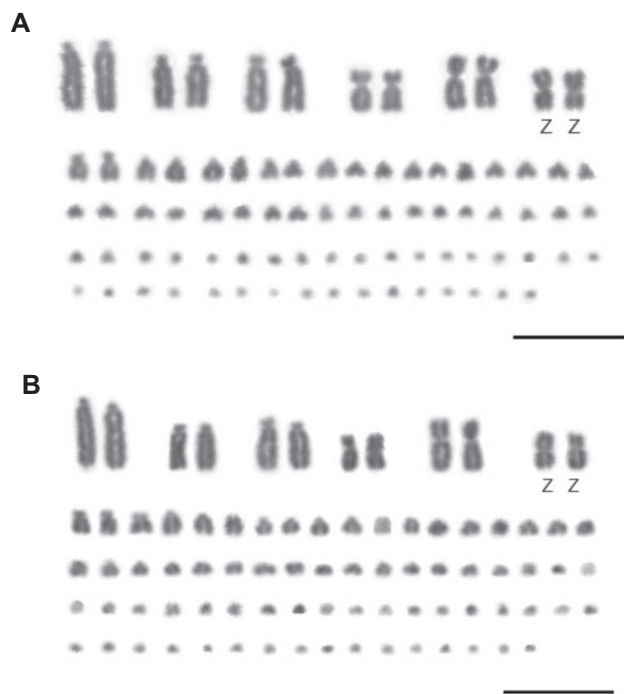


Figure 9: Karyotype analysis. Karyotype of *S. uralensis* male with $2n = 82$ after passage 5 (A) and passage 10 (B). Bar = 10 μm .

a significantly lower depth and quality than the average nonprivate site, suggesting that these alleles are to some extent wrongly called (Supplementary Fig. S10). To account for patterns potentially driven by sequencing technology, we also assessed in each replicate how many positions it differed compared to passage 0.

This revealed the same pattern of an increase of SNPs in sample passages 5.1 and 10.1.

Discussion

Reference genomes are accumulating across the tree of life, and here birds have seen special attention fueled by initiatives such as B10K [36–38]. Still, many of these genomes remain incomplete in terms of chromosomal-scale assembly type as well as gene annotation comprehensiveness. Genome assembly quality can impact phylogenomic inferences, analyses of gene prediction, gene family expansion and contraction, and, most importantly, structural evolution. In an effort to allow such analyses for the vastly understudied bird order Strigiformes, we here present a reference genome for the Ural Owl that is among the best bird genome assemblies currently, reflected by assessments of sequence and gene completeness. We could place most of the genome into chromosomal-scale scaffolds that are in line with the species karyotype, which we also confirm by cytogenetics. We further located the supposedly *Strix*-specific chromosomal fusion, which distinguishes it from the genus *Bubo* [5]. In addition, we identified several other chromosomal rearrangements between owl genomes (which are overall very syntenic) and those of chicken and zebra finch. For four microchromosomes, we could not identify a reliable homologous chromosome in chicken or Zebra Finch, but we found syntenic scaffolds in the Snowy Owl and microchromosomes in the Tawny Owl assemblies. In addition, the microchromosomes we assembled were also well supported by our HiC data. We thus hypothesize that the lack of synteny to chicken and Zebra Finch might reflect more structural changes specific to owls. Matchingly, a first analysis of the Ural Owl genome content also indicates an important increase in repetitive sequences compared to most other bird genomes. Birds, on average, have rather compact genomes compared to other

Table 2: SNP statistics over progressive cell culture passages

Sample	Median depth of variant sites	Median GQ of variant sites	SNPs other than shared heterozygous sites/private to individual	SNPs other than shared heterozygous sites as a percentage of the total genome size [%]	SNPs that are 0/1	SNPs that are not 0/1 or 1/1	SNPs other than fixed heterozy- gous sites [%]	SNPs different from passage 0	SNPs different from passage 0 [%]
HiFi reads	39	99	78,983/35,645	0.0057	77,682	1,265	0.14	–	–
Passage 0	27	99	57,929/8,212	0.0042	56,232	816	0.922	–	–
Passage 5.1	27	99	78,193/23,889	0.0057	76,709	716	0.809	58,620	6.62
Passage 5.2	29	99	55,225/3,280	0.0040	54,037	615	0.695	38,765	4.38
Passage 5.3	28	99	57,664/5,581	0.0042	56,267	781	0.882	41,301	4.67
Passage 10.1	20	99	103,086/63,058	0.0075	101,450	25	0.028	96,768	10.93
Passage 10.2	28	99	54,404/3,410	0.0039	53,373	420	0.475	38,921	4.40
Passage 10.3	28	99	55,976/3,758	0.0041	54,679	642	0.725	39,843	4.50
Passage 10.4	15	99	54,215/15,521	0.0039	52,153	750	0.847	50,033	5.65

vertebrate lineages (average ~1.1 Gb [39]), mostly owing to a low content of repetitive elements (10–15%). Until now, owls were seemingly no exception to this, with an average genome size of 1.2 Gb and a repeat content of ~8.6% [39]. Still, cytogenetic studies already suggested differently and hinted at owls being rather an exception in the avian lineage, with large-scale variations in karyotypes, which our assembly also supports. For example, the barn owls have chromosomes of a more homogeneous smaller size, while the true owls have a more classical avian karyotype with macro- and microchromosomes, suggesting chromosome fusion and fission in owls [5]. Interestingly, the recently published genome (1.6 Gb) of the Snowy Owl further extends this observation by demonstrating that it has one of the highest reported repeat contents for birds (28.34%), mainly composed of centromeric satellite DNA [40]. Our assembly's total repeat content, at 21.07% (1.5% of which is unidentified), follows this pattern. While both owl genomes' repeat expansions are largely driven by retrotransposons, the Snowy Owl had a stronger increase of LTRs compared to LINEs than the Ural Owl. Nevertheless, LTR retrotransposons are the largest repeat class also in the Ural Owl, and the youngest repeat expansion is also driven by LTRs, suggesting this pattern to be more broadly present in true owls. In the Snowy Owl, the repeats are suspected to have driven the evolution of novel centromeres. Accordingly, cytogenetic analyses already have identified large centromeric satellite blocks shared among and unique to true owls [34]. Other bird lineages with increased repeat content are woodpeckers and the Common Scimitarbill (*Rhinopomastus cyanomelas*) [39]. The cause and consequences of the repeat extensions in the genera *Strix* and *Bubo* remain unclear at this point, which is also true for the woodpecker [41].

The number of microchromosomes identified in *S. uralensis* ($n = 31$) is consistent with the average number of microchromosomes reported by Tegelström and Rytman [42] from karyotypes of over 230 bird species. However, there is no established rule for distinguishing between macrochromosomes and microchromosomes (e.g., [43]). We here confirm that in birds, microchromosomes have a higher gene density than macrochromosomes (e.g., [6, 43–45]). We could further show that the higher density of genes on microchromosomes is associated with a reduction of gene length, which in turn is due to correspondingly shorter introns. A similar pattern has been reported in chickens, where the size of the chromosome correlates with the length of the genes it harbors [46]. Thus, in true owls, microchromosomes hold up the crucial role they supposedly have played throughout vertebrate evolution [47].

Our assembly seems to be particularly well suited for an analysis of gene content due to a high completeness of gene annotation. However, due to vastly varying assembly qualities, the correctness of our gene family expansion analysis should be taken with a grain of caution. Still, this preliminary analysis suggests that in the future, we will be able to connect changes in gene content to adaptations of owls. This is supported by Ural Owl-specific gene gains in the GO-term derived function of “eye development,” “sensory organ development,” and “neurogenesis,” for example, which could be linked to adaptations required for a nocturnal, predatory lifestyle [48].

We also offer candidate genes for further investigation that characterise predatory lifestyle (i.e., genes gained in the common ancestor of Strigiformes and Accipitriformes that acquired this lifestyle). We especially observed gains of genes with a metabolic function, which could relate to the change in diet in the ancestor of these two bird orders. We also found gains of keratin genes in the common ancestor of the two predatory bird lineages and also in the ancestor of owls. Feathers are epidermal appendages. Vertebrate skin appendages consist of two fibrous proteins, α - β -keratins. Interestingly, β -keratins are exclusively found in reptiles and birds. Both keratin gene families show expansions in different lineages. The Barn Owl had the lowest number (6) of β -keratins in a study comparing 48 bird (draft) genomes. The Zebra Finch, in comparison, had 149 genes [49]. This comparison further showed that the proportion of claw β -keratins and keratinocyte β -keratins is higher in predatory birds. We support the latter finding, with the detection of three gains of feather and scale keratins in the common ancestor of Accipitriformes and Strigiformes and two further gains in the ancestor of all owls. These keratin genes deserve more attention as potential candidates that could underlie morphological adaptations of feathers in predatory birds in general but also more specifically in the mostly nocturnally hunting owls. Their silent flight is made possible by physical characteristic fringes of the feathers on the leading edge of the wings [50]. The genomic basis of this adaptation remains to be identified.

The Ural Owl is protected under the CITES convention Annex II and the Bern Convention on the Conservation of European Wildlife and Natural Habitats. While globally not (yet) under concern, the species went extinct in Germany and Austria due to habitat destruction but also direct persecution. Reintroduction programs have been started, but due to low availability of breeding couples, individuals of various origins are used for these actions [19]. An analysis of marker genes neither supported morphological subspecies nor revealed a phylogeographic population

structure for the Ural Owl, but it revealed genetic clusters that could be informative for breeding programs [13]. The here generated reference genome will facilitate future genomic studies in this direction of *S. uralensis*.

With an estimated genome-wide heterozygosity of 0.2% (2 het/kb), the here sequenced individual shows a higher level of heterozygosity than genomes of endangered bird species (IUCN Red List status accessed March 2025 [51]), such as the White-eared Night Heron (*Gorsachius magnificus*, Endangered, 0.49 het/kb) [52], Andean Condor (*Vultur gryphus*; Vulnerable, 0.75 het/kb), California Condor (*Gymnogyps californianus*; Critically Endangered, 1.34 het/kb) [53], and Crested Ibis (*Nipponia nippon*, Endangered, 0.043 het/kb) [54]. A similar heterozygosity level as the one we estimated for the Ural Owl was detected in, for example, Wild Turkey (0.24%) and Mallard (0.26%) [55]. It is somewhat lower than levels reported for other Strigiformes, such as the Little Owl (*Athene noctua*, 0.593%) [56], Tawny Owl (*S. aluco*, 0.57% to 0.70%) [57], and Barn Owl (*Tyto alba*, 0.59% to 0.71%) [58], yet twice as high as in the Burrowing Owl (*Athene cunicularia*, 0.1%) [59].

Overall, species with a threat of going extinct show reduced levels of heterozygosity compared to nonthreatened related taxa [60]. Related taxa of the same bird order, with and without risk of extinction, differ quite drastically in genome-wide heterozygosity [54]. These differences likely result in lower evolutionary potential and reduced reproductive fitness, which may contribute to species extinction [60]. It remains to be assessed at which level heterozygosity reduction causes an issue for a particular species. By generating genomic data for the Ural Owl, we contribute to the required knowledge for genetic monitoring of biodiversity.

Further, our reference genome already sheds light on the demographic history of the species, indicating both population contractions and expansions, apparently related to ecological effects of the glacial-interglacial cycle. In particular, the pattern over the past 120,000 years not only demonstrates the Ural Owl's tolerance to lower temperatures but, more importantly, reflects its flexible habitat choice of semi-open woodlands with a mixed composition of broadleaf and coniferous species [11]. From the Eemian interglacial through the Weichselian glaciation, climatic changes caused fluctuations in ice sheet extent and associated changes in vegetation composition, including a gradual and/or repeated reduction in forest cover, leading to a treeless shrubby or grassy tundra from the mid-Weichselian (e.g., [61–63]). While the open or semi-open structure of the woodland habitats favored the Ural Owl's preference for breeding and hunting grounds [12] until the mid-Weichselian, the expanding tundra substantially reduced suitable habitats, leading to a marked decrease in effective population size.

In the light of species preservation, protection, and restoration, *ex situ* efforts are gaining more attention. Cell culturing is a valuable and widely used technique, spanning applications from basic science to biotechnology research [30]. However, there is no consensus regarding the number of passages considered “safe” before cells experience metabolic changes, DNA damage, and chromosomal instability. What is deemed “high passage” for one cell culture may not lead to significant passage effects in another [31]. Thus, the effects of prolonged culture are complex and depend on the individual cell culture, tissue, and species.

The first criterion for identifying healthy and stable cells is observing cell morphology. Chromosome content serves as another critical benchmark, as normal cells maintain a stable chromosome number. Some studies using nonmodel species, such as felines [64] and fishes [65], showed no heteroploidy on karyotypes obtained by cell cultures. However, to the best of our knowledge,

this is the first study that addresses genomic and chromosome changes in wild birds and compares the effect of different cell passages on genome integrity.

Cryopreservation of cells has increasingly been considered a strategy for conservation as new technologies using genetic material from somatic cells (e.g., somatic cell nuclear transfer or induced pluripotent stem cells) are evolving [28, 66]. One of the prerequisites for nuclear donor cells and *ex situ* conservation is the stability of chromosomes [64]. Studies that investigate cell line passage and age effects are still scarce in nonmodel organisms and are crucial since altered metabolism and genomic instability no longer represent reliable models of their original source of material.

Herein, comparison between karyotypes of passage 5 and passage 10 showed no differences, suggesting that no large structural rearrangement occurred during the progressive number of passages and that it is safe to establish the diploid number of (this) bird species until at least the 10th passage. Genetic instability is well documented in cells that have undergone more than 20 passages, particularly in transformed continuous cell lines (e.g., [67]) and tumor cell lines [68, 69]. However, for primary cell cultures, a straightforward method to determine the safest passage number before cells develop mutations or genetic instability is lacking. We opted to cultivate cells up to passage 10 based on two factors: first, the uncertainty surrounding the exact passage limit at which primary cells may enter senescence (as a noncontinuous cell line has a limited *in vitro* lifetime) and, second, technical challenges observed during later passages as cells began to exhibit signs of morphological decline, including the presence of granules and debris, difficulty detaching, and a reduced growth rate, all of which would complicate further subculturing beyond passage 10. This seems to suggest safe cell culturing for this species until passage 10. To some extent, this is supported by our SNP analysis of several passage replicates, which indicated no general pattern for increased genomic changes between passages 5 and 10. However, we detected outlier samples with respect to SNP numbers among replicates of both passage numbers. At this point, we lack any point of reference expectation as to how many (potential) mutations are to be expected in a cell culture system as the one we applied. Compared to overall levels of variant sites, the number of SNPs in the individual samples that could be mutations is lower, representing between 0.0039% and 0.0075% of the genome assembly length. The effects of these variants remain to be determined, as well as the reason for between-replicate differences. We further suspect that several mutations are variant-calling artifacts, supported by lower SNP calling quality, which asks for an exploration of mutation identification and, more importantly, validation for the type of cell culture we have set up here.

Potential Implications

We were able to assemble a reference genome for the Ural Owl of gold-standard quality, which is open to the community to be used for broader comparative genomic studies and phylogenomic analysis but also provides immediate resources to researchers interested in the Ural Owl for taxonomic and conservation aspects. With the data generated, we contribute to the endeavor of sequencing all life on Earth [2]. Our analysis of genomic data derived from cell passages opens space for discussion of cell cultures as material for genomics, especially for species with limited biological material available. The workflows applied by us could be used on similar data from other species.

Methods

Species origin and sampling strategy

Skin and muscle tissue samples from a 10-year-old male individual of *S. uralensis* (ring ID ZG-14.0-10-0234) were obtained from the Raptor Center & Wildlife Park Hellenthal (Wildfreigehege und Greifvogelstation Hellenthal, Hellenthal, Germany) during necropsy in 2020. The procedure was performed by Dominik Fischer, who is a veterinarian and approved to handle animals. No further approval was needed for this study. DNA barcoding was performed (collection ID ZFMK-TIS-50475) to ensure species identity using primers for COI from Astrin and Stüben [70], and sequences matched against BOLD (Barcode of Life Data System) [71]. The barcode sequence has been uploaded to BOLD as FOGS049-22.

Reference genome

Sequencing

DNA was extracted from the skin biopsy (collection ID ZFMK-TIS-50482, stored at LIB Biobank in liquid nitrogen vapor phase) using the Monarch HMW DNA Extraction Kit (NEB). High molecular weight status was validated by quality control with capillary electrophoresis (Agilent Femto Pulse System; [RRID:SCR_018058](#)), and an SPK 3.0 PacBio HiFi library was prepared according to the recommendations by the vendor. Next, HiFi SMRT sequencing was performed on two SMRT cells on a PacBio Sequel IIe (Pacific Biosciences; [RRID:SCR_017990](#)) at the Max-Planck Genome-centre Cologne (MP-GC; Cologne, Germany). Also, a chromatin-capture library was prepared from cryopreserved cells generated for the analysis over progressive cell passages, as described below, with an Arima-Hi-C Kit according to the protocol for mammalian cell lines, followed by sequencing on an Illumina NextSeq 2000 system ([RRID:SCR_023614](#)) in paired-end read mode.

Read processing

HiFi data

Contaminant sequences were filtered from the HiFi reads using Kraken2 v2.1.3 ([RRID:SCR_026838](#)) [72, 73] with the Kraken database kraken2 PlusPFP, downloaded in March 2023, and parameters “–confidence 0.51 –use-names.” HiFi read quality was assessed using seqkit v2.8.2 ([RRID:SCR_018926](#)) [74, 75] and a *k*-mer-based approach. The *k*-mers were calculated with Meryl v1.4.1 ([RRID:SCR_026366](#)) [76] using the parameters “count *k*=21,” and the counts were converted into a histogram with the *meryl histogram* command.

To verify the ploidy of the individual, Smudgeplot v0.2.5 [77] was used. First, *k*-mers within a specific range (lower-upper), determined with the *smudgeplot.py* cutoff function, were extracted using the *meryl print less-than* command. These filtered *k*-mers were then processed with *smudgeplot.py hetkmers* to calculate the coverage of unique heterozygous *k*-mer pairs. The resulting coverage was plotted using *smudgeplot_plot.R*.

GenomeScope2 v2.0.1 [77] with default parameters was used to estimate genome size, heterozygosity, and the homozygous and heterozygous coverage peaks.

ROHan v1.0.1 [78] was used to identify large (>1 Mb) runs of homozygosity.

Hi-C data

Adapter removal and quality filtering of the raw Hi-C reads were performed using Fastp v0.23 ([RRID:SCR_016962](#)) [79] with parameters “–length_required 95, –qualified_quality_phred 20 –

adapter_fasta,” with a curated adapter list of the most common adapters used as input.

Error correction was done using *Tadpole* from BBDMap v39.01 ([RRID:SCR_016965](#)) [80], with parameters “*k*=50, reassemble=t, mode=correct, minprob=0.6, prefilter=1, prehashes=2, and prealloc=t.” To remove contamination from the short Hi-C reads, Kraken2 v2.1.3 was used similarly to the HiFi reads but in paired-read mode, with parameter “–paired.”

Initial genome assembly

The HiFi reads were used with Hifiasm v0.19.5 ([RRID:SCR_021069](#)) [81] to generate a phased genome assembly. In order to obtain an optimal, phased genome, we tested several Hifiasm parameters before choosing the ones that provided us with the assembly of the highest contiguity and completeness, with both phased haplotypes having a similar length. We tested all possible combinations of different purging level (0, 2, 3), increasing runtime, and number of iterations (“–n-weight 5 –n-perturb 50000 –f-perturb 0.5 –D 10 –N 150 –s 0.2”) and explicitly providing the homozygous peak to Hifiasm, which was estimated by GenomeScope (“–hom-cov 40”) (for more details, see [Supplementary Table S1](#)).

Genome scaffolding

The selected haplomes from Hifiasm were split at positions containing Ns using *split_fa* from the Purge_Dups package v1.2.6 ([RRID:SCR_021173](#)) [82]. The resulting sequences were mapped to themselves using Minimap2 v2.26 ([RRID:SCR_018550](#)) [83] with parameters “–x asm5 –DP” and to the HiFi reads using Minimap2 with parameters “–x map-hifi.” These mappings were used to remove assembly duplicates with Purge_Dups.

Mitochondrial genome detection

To identify and extract the mitochondrial genome, we utilized MiToHiFi v3.2.1 ([RRID:SCR_026369](#)) [84, 85], referencing the sequence NC_038218.1 from *S. uralensis* (isolate C5 mitochondrial genome, complete; NCBI Taxonomy ID: 36305). The most likely scaffold was kept and identified as the mitochondrial chromosome (MT), and all other candidate scaffolds were removed from the assembly.

Assembly manual curation

Hi-C reads were aligned to the final assemblies, and a Hi-C contact map was created using PretextView v0.0.2 ([RRID:SCR_022023](#)) [86]. A HiFi coverage track was generated from the aligned HiFi reads using bedtools *genomecov* ([RRID:SCR_006646](#)) and integrated into the PretextView map with PretextViewGraph ([RRID:SCR_026377](#)).

Manual curation was performed within PretextView v0.0.2 ([RRID:SCR_022024](#)) [87], where scaffolds were reordered and oriented based on Hi-C interaction frequencies. Following curation, the final assembly scaffolds were processed with AGP tools from the Vertebrate Genomes Project (VGP) using the rapid manual curation protocol [88] established by the Darwin Tree of Life consortium [89] to create the curated assembly. Scaffold names were further sorted and renamed by size using a combination of seqkit v2.8.2 and SAMtools v1.19.2 ([RRID:SCR_002105](#)) [90]. The final Hi-C contact map was visualized with HiCExplorer ([RRID:SCR_022111](#)) [91].

Genome quality control

The completeness of the final curated assembly was assessed using BUSCO v5.8 ([RRID:SCR_015008](#)) [92, 93] and compleasm v0.2.6 [94] with the *aves_odb10* lineage. Assembly contiguity and general assembly metrics were calculated using Quast v5.2.0 ([RRID:SCR_001228](#)) [95].

For k-mer-based analysis, k-mer counts were generated for each assembly using Meryl. These counts were analysed with Merqury v1.3 (RRID:SCR_022964) [96] to estimate assembly completeness and accuracy. The analysis yields Merqury's consensus QV, which is estimated by comparing the read and assembly k-mer counts and then transformed to a log-scaled probability of base-call errors. A higher QV indicates a more accurate assembly. We also obtained a Merqury completeness percent, which reflects the proportions of high-quality HiFi read k-mers present in the assembly.

HiFi reads were mapped to each assembly using Minimap2 with parameters “-ax map-hifi.” Alignment quality and coverage distribution were assessed using Qualimap v2.3 [97].

Potential contamination and quality were also assessed using the blobtoolkit pipeline v3.5.4 (RRID:SCR_025882) [98] and visualized using the interactive Blobtoolkit viewer in the Galaxy EU server (RRID:SCR_006281) [99].

Genome annotation

Repeat annotation

Repetitive elements in the primary assembly were identified and annotated using EarlGrey v5.1.1 [100], which was run with RepeatMasker v4.1.5 (RRID:SCR_012954) [101] and RepeatModeler v2.0.6 (RRID:SCR_015027) [102]. In addition to the RepeatModeler library, we used a previously created, custom avian TE library to mask repetitive elements [103]. The softmasked genome was used for protein-coding gene prediction.

Protein-coding gene annotation

To perform protein annotation, we created two reference protein sets. Set one contained only the merged proteomes of the following publicly available genomes, downloaded using the NCBI dataset cli v16.3.0 [104]: *S. nigrolineata* (GCA_013396715.1), *Gallus gallus* (GCF_016699485.2), *Glaucidium brasilianum* (GCA_013399595.1), *Falco peregrinus* (GCF_023634155.1), *Athene cunicularia* (GCF_003259725.1), *Aquila chrysaetos* (GCF_900496995.4) [105], *Taeniopygia guttata* (GCF_003957565.2), and *Anas platyrhynchos* (GCF_015476345.1).

Protein set two was created by merging all proteomes in set one with proteins from the following public and curated databases: (i) proteins from the BUSCO v5.4 aves_odb10 dataset; (ii) aves proteins from OrthoDB v11 [106], which were obtained using Tomas Bruna's orthodb-clades pipeline [107]; and (iii) proteomes, which were also extracted from the UniProt database for the following species: *Calypte anna* (UP000054308), *Steatomis caripensis* (UP000516988), *Cnemophilus loriae* (UP000517678), *Dasyornis broadbenti* (UP000521322), *Corythaixoides concolor* (UP000526942), *Irena cyanogastra* (UP000530962), *Bucco capensis* (UP000534107), *Cephalopterus ornatus* (UP000543364), *Molothrus ater* (UP000553862), *Ptilonorhynchus violaceus* (UP000584880), *Promerops cafer* (UP000587587), *Vidua chalybeata* (UP000634236), and *Urocolius indicus* (UP000654395).

Protein-coding genes in the *S. uralensis* genome were annotated using a combination of *ab initio*, protein similarity, and transcriptome-based protein prediction models. BRAKER3 v3.0.3 [108, 109] was run in EP mode using the protein set two described above. GALBA v1.0.11.2 [110] was run using protein set one to annotate genes. The outputs from GALBA and BRAKER3 v3.0.2 were combined using TSEBRA from the BRAKER3 package. To ensure high-quality annotations, only the longest gene orthologs for each locus were retained using the *agat_sp_keep_longest_isoform.pl* script from the AGAT package v1.4.1 (RRID:SCR_027223) [111].

Demographic history of *S. uralensis*

The demographic history of the Ural Owl was reconstructed using PSMC v0.6.5 (RRID:SCR_017229) as implemented by [112]. Variants were called per chromosome using a combination of BCFtools v1.21 [113] *mpileup* with parameters “-Q 30 -q 30” and *bcftools* call using the “-c” option. The resulting variant call format (VCF) file was converted to a consensus fastq format using the *vcfutils.pl vcf2fq* script with parameters “-d 10, -D 60, and -Q 30.” The PSMC model was run with the following parameters: -N25 -t15 -r5 -p “2+2+25*2+4+6” and 100 bootstraps, a generation time of 3 years [114, 115], and an assumed mutation rate of 4.6×10^{-9} [116, 117].

Genome synteny

In order to identify the Z sex chromosome within our genome assembly and to assess the synteny of different bird genomes, we used the *GetTwoGenomeSyn.pl* built-in script of NGenomeSyn [118] with the options “-MappingBin minimap2 -MinLenA 100000 -MinLenB 100000 -NumThreads 5 -MappingPara “-Lx asm10 -eqx -I 200 G -MD -N 1”” to estimate chromosome-scale alignments among *G. gallus* (GCA_024206055.2), *T. guttata* (GCF_048771995.1), *S. aluco* (GCA_031877795.1), *S. uralensis*, and *B. scandiacus* (GCA_965212795.1). To improve visualization and reduce noisiness in the synteny plot, we used the *get.synteny.blocks.multi* function of the *asyntr* R package [119] to aggregate single alignments into synteny blocks of incrementally increasing minimum block size of 100 to 4,000 and filtered for chromosomes smaller than 1,000 bp and for alignment lengths smaller than 200 bp. For visualization purposes, we manually removed alignments between *G. gallus* and *T. guttata* that were not supported by Luo et al. [120].

Functional gene annotation

Predicted genes were functionally annotated by performing sequence similarity searches against the Swiss-Prot database (RRID:SCR_021164) using BLASTP (RRID:SCR_001010) from BLAST v2.13.0+ (RRID:SCR_004870) with default parameters. As with our own genome's annotation, we used *agat_sp_keep_longest_isoform.pl* to keep only the longest isoform of each gene locus from the *Aquila chrysaetos*, *Gallus gallus*, *S. nigrolineata*, *Athene cunicularia*, *Glaucidium brasilianum*, *Falco peregrinus*, and proteomes of protein set 1, together with *Tyto alba* (GCA_018691265.1) and *Haliaeetus leucocephalus* (GCA_000737465.1), and used OrthoFinder v2.5.5 (RRID:SCR_017118) [121] to estimate orthologous gene families among those species. This analysis identified orthogroups and genes that have undergone expansion or contraction in the Ural Owl, as well as orthogroups unique to this species.

GO Term analysis

GO Term analysis was performed by mapping all Ural Owl genes to the Vertebrate Egglog database using egglog mapper (RRID:SCR_021165) v2.1.12 [122, 123]. Missing GO Terms were filled in with the GO Terms of the previously found Swiss-Prot gene symbols associated with each gene. Next, the genes belonging to gene families unique to the Ural Owl (found with Orthofinder) were analysed by Revigo v1.8.1 (RRID:SCR_005825) [124] with default settings and choosing the *Large* subset option. The resulting GO Terms were analysed with Revigo again with default settings and this time with the *Small* subset option. The full Biological Process Revigo table was plotted in R v4.4.2 (RRID:SCR_001905) using an edited version of the Revigo treemap plotting script.

Variation analysis over progressive cell passages

Cell culture

Primary cells were grown from a skin biopsy specimen of the same individual as used for genome sequencing (collection ID ZFMK-TIS-51054) previously stored at the LIB Biobank in liquid nitrogen, following standard protocols. Skin tissues were rapidly thawed, minced into small fragments, and transferred to cell culture flasks. Flasks were incubated at 37°C with 5% CO₂ in fibroblast growth basal medium (Lonza), supplemented with 20% fetal bovine serum (Biowest), including antibiotics (100 U/mL penicillin and 100 g/mL streptomycin; Sigma-Aldrich). Cells were visually inspected in an inverted microscope (Nikon Eclipse TS2; [RRID:SCR_025716](#)) for contamination, and cell media were changed every 2–3 days. After reaching ~80% confluence (determined visually), cells were propagated using 0.125% trypsin solution (Biowest), at a subculture ratio of 50:50. Cells were harvested for DNA extraction and chromosome analysis at passage 5 (3 different replicates) and passage 10 (4 different replicates).

Chromosome sampling for large variant analysis

In order to investigate the stability of the karyotype composition through different passages, chromosome preparations were obtained from cells for passages 5 and 10, according to [125], with modifications. Chromosomes were harvested after treatment with colchicine 0.01% for one hour, followed by hypotonic treatment with 0.075 M KCl and cell fixation in methanol/acetic acid (3:1). Slides were stained with Giemsa 5%. At least 20 metaphases for each passage were analysed to define the diploid number (2n) in a Zeiss Axio Imager Z2m system ([RRID:SCR_027233](#)).

DNA extraction of primary tissue and cell culture passages

Passage samples were extracted using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's protocol for cultured cells, while muscle tissue of the same individual (collection ID ZFMK-TIS-50476), previously stored at the LIB Biobank in 96% ethanol (passage 0), was extracted using the standard protocol of the same kit.

Sequencing of primary tissue and passages

After DNA extraction, samples were sent for purification (Vahts DNA Clean Beads; Vazyme Biotech), PCR-free library preparation (NEBNext Ultra II FS DNA PCR-free Library Prep Kit for Illumina; NEB), and subsequent paired-end sequencing on a Illumina NovaSeq 6000 Sequencing System ([RRID:SCR_016387](#)) (Illumina) using the NovaSeq 6000 S4 Reagent Kit (Illumina) to Biomarker Technologies (bmkgene).

Read mapping

The Illumina reads of all passages (passage 0; i.e., the primary tissue, three passage 5 replicate samples, and four passage 10 replicate samples) were processed with fastp v0.20.0 [79] with parameters “–length_required 95, –qualified_quality_phred 20 –adapter_fasta,” with a curated adapter list of the most common adapters used as input and decontaminated with Kraken2 v2.1.3 with the Kraken database kraken2 PlusPFP database, downloaded in March 2023, in paired-read mode, with parameter “–paired –confidence 0.51 –use-names.” These reads were then mapped to the reference genome using BWA-MEM2 v2.2.1 ([RRID:SCR_022192](#)) [126, 127] with the *mem* command and options “–M –R,” where a read group specific to each sample was used for “–R.” The resulting output was sorted using SAMtools v1.19.2, and additional process-

ing steps (SAMtools' *fixmate*, *sort*, and *markdup*) were performed to generate the final mapping files for each sample.

SNP calling

SNP calling for each cell passage sample and the HiFi reads (“reference”) was performed individually using GATK HaplotypeCaller v4.2.6.1 ([RRID:SCR_001876](#)) [128, 129] with the options “–ERC GVCF –min-base-quality-score 30 –pcr-indel-model NONE.” Joint SNP calling was performed by first combining samples using GATK GenomicsDBImport with the option “–batch-size 3.” The combined database was then used for joint SNP calling with GATK's *GenotypeGVCFs*.

Variant quality recalibration was conducted in three rounds. GATK's *BaseRecalibrator* was run with the option “–maximum-cycle-value 50000,” followed by GATK's *ApplyBQSR* for each sample before re-calling variants individually and collectively. From the final set of called genotypes, SNPs were extracted using GATK's *SelectVariants* with the option “–select-type SNP” and filtered with GATK *VariantFiltration* using the following filters: “QD < 2.0, FS > 60.0, MQ < 40.0, SOR > 3.0, MQRankSum < -12.5, ReadPosRankSum < -8.0, QUAL < 30.0.”

Variants were filtered for depth, minor allele frequency (MAF), and the fraction of missing genotypes using BCFtools filter v1.21 with the options “–e “INFO/DP<\$MIN_DEPTH || INFO/DP>\$MAX_DEPTH” ” and “–i “MAF>\$MAF && F_MISSING≤\$MISS.”

Short-read variant analysis in passages

To assess the quality of the DNA contained in cell cultures and to understand its potential for being used as an amplified genomic resource, these filtered SNPs were then analysed in R v4.4.2. Passages 5 and 10 were compared to passage 0 and the HiFi reads at sites where passages 5 and 10 differ from either of the reference passages. A Wilcoxon test from rstatix v0.7.2 was performed to test whether the depth and GQ of these SNPs of each sample were significantly different from the average DP or GQ. The variant calls were recoded so that “0|0” = 0, “0|1” = 1, “1|0” = 1, “1|1” = 2, “0|2” = 3, “2|0” = 3, “1|2” = 4, “2|1” = 4, “2|2” = 5, “0|3” = 6, “3|0” = 6 and plotted with ggplot2.

Availability of Source Code and Requirements

Project: Strix uralensis assembly, annotation, and comparative analysis

Location: Zenodo DOI: 10.5281/zenodo.15100180

Operating system(s): e.g., Platform independent

License: CC0

Additional Files

Supplementary Fig. S1. PloidyPlot of *Strix uralensis* HiFi reads.

Supplementary Fig. S2. *Strix uralensis* primary haplome BlobToolKit GC coverage and cumulative sequence plots. (A) Blob plot of base coverage in *S. uralensis* against GC proportion for sequences in *S. uralensis* primary haplome. Sequences are coloured by phylum. Circles are sized in proportion to sequence length. Histograms show the distribution of sequence length sum along each axis. (B) Cumulative sequence length for *S. uralensis* primary assembly. The grey line shows cumulative length for all sequences. Coloured lines show cumulative lengths of sequences assigned to each phylum using the BUSCO genes tax rule.

Supplementary Fig. S3. *Strix uralensis* primary haplome Hi-C contact map showing spatial interactions for all chromosomes. Chromosomes are ordered by size from left to right and from top to bottom. The red diagonal corresponds to intrachromosomal contacts and depicts chromosome boundaries. The frequency of contacts is shown on a logarithmic heatmap scale. Plot generated with HiC-Explorer.

Supplementary Fig. S4. *Strix uralensis* primary haplome Hi-C contact map showing spatial interactions between all chromosomes following the ten largest chromosomes. Chromosomes are ordered by size from left to right and from top to bottom. The red diagonal corresponds to intrachromosomal contacts and depicts chromosome boundaries. The frequency of contacts is shown on a logarithmic heatmap scale. Plot generated with HiCExplorer.

Supplementary Fig. S5. Pie chart of repeat family categories estimated by EarlGrey.

Supplementary Fig. S6. Major repeat family categories and sub-categories estimated by EarlGrey.

Supplementary Fig. S7. Gene structure of the *Strix uralensis* primary haplome. The 41 haploid chromosomes of *S. uralensis* are divided into macrochromosomes (>30 Mb; $n = 10$; yellow) and microchromosomes (<30 Mb; $n = 31$; red). (A) Chromosome size distribution of macro- and microchromosomes. (B) Gene density of macro- and microchromosomes. (C) Mean gene length of macro- and microchromosomes. (D) Mean intron length of macro- and microchromosomes. Boxplot center lines represent the median, box limits the upper and lower quartiles, and whiskers the 1.5× interquartile range. Differences were assessed using the Wilcoxon test ($***P \leq 0.001$).

Supplementary Fig. S8. Treemap plot of the most frequent GO term categories of gene families unique to the *Strix uralensis* primary haplome. Colour of sections is unique to each category, and the size scales positively with the GO term frequency.

Supplementary Fig. S9. Variant calls of each passage. SNPs are ordered by genome position as derived from the variant file, and colour indicates allele, as illustrated in the inset and referring to “0|0” = 0, “0|1” = 1, “1|0” = 1, “1|1” = 2, “0|2” = 3, “2|0” = 3, “1|2” = 4, “2|1” = 4, “2|2” = 5, “0|3” = 6, “3|0” = 6.

Supplementary Fig. S10. Quality assessment of private sites. Median depth (top) and median genotype quality (bottom) of private SNPs in each section compared to nonprivate SNPs (star).

Supplementary Table S1.xlsx. Assembly statistics of all tested assembly parameters.

Supplementary Table S2.xlsx. Primary and alternate assembly Merquy Statistics.

Supplementary Table S3.xlsx. HiFi and HiC reads mapping quality results.

Supplementary Table S4.xlsx. Protein annotation statistics.

Supplementary Table S5.xlsx. *Strix uralensis* GO Term analysis.

Supplementary Table S6.xlsx. Synteny statistics of *Strix uralensis* and three other bird species.

Abbreviations

2n: diploid chromosome number; b: bases; bp: base pair; BLAST: Basic Local Alignment Search Tool; BOLD: Barcode of Life Data System; BUSCO: Benchmarking Universal Single-Copy Orthologs; CITES: Convention on International Trade in Endangered Species of Wild Fauna and Flora; DP: depth; EU: European Union; GATK: Genome Analysis Toolkit; Gb: gigabases; GO: Gene Ontology; GQ: genotype quality; Hi-C: high-throughput chromosome conformation capture; HiFi: high-fidelity; HMW: high molecular weight; kb: kilobases; LINE: long interspersed nuclear element; LTR: long ter-

minal repeat transposable element; M: molar; MAF: minor allele frequency; Mb: megabases; MT: mitochondrial chromosome; PSMC: pairwise sequentially Markovian coalescent; QV: quality value; RG: read group; ROH: runs of homozygosity; SNP: single-nucleotide polymorphism; TE, transposable element; VCF: variant call format; VGP: Vertebrate Genomes Project; ya: years ago.

Declarations

The primary tissue used for this work was derived from a naturally deceased bird and provided by a veterinarian. We did not perform animal experimentation.

Acknowledgments

We thank Juliane Vehof and Benjamin Wipfler for enabling us to use their microscope.

Author Contributions

I.C.: conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization, writing (original draft; review & editing). A.M.: conceptualization, data curation, formal analysis, investigation, methodology, validation, visualization, writing (original draft; review & editing). C.B.D.N.: conceptualization, formal analysis, investigation, methodology, writing (original draft; review & editing). D.F.: resources, writing (review & editing). N.S.: investigation, writing (review & editing). L.v.d.M.: investigation, writing (review & editing). B.H.: investigation, writing (review & editing). J.J.A.: conceptualization, funding acquisition, supervision, writing (original draft; review & editing). T.T.: validation, supervision, writing (original draft; review & editing). A.B.: conceptualization, data curation, validation, visualization, funding acquisition, supervision, writing (original draft; review & editing).

Funding

This work was supported by the Leibniz Gemeinschaft Leibniz Association Network grant CollOmic K419/2021 to A.B. and LIB innovation fund to A.B. and J.J.A.

Data Availability

The sequencing reads, assembly, and BioSample data supporting the results of this article are available from the INSDC under BioProject number PRJNA1212906. Further datasets and code supporting the results of this article are available from Zenodo [130]. Code is available from Zenodo [131]. All supporting data and materials are available in the GigaScience database, GigaDB [132].

Competing Interests

The authors declare that they have no competing interests.

References

1. Lewin HA, Robinson GE, Kress WJ, et al. Earth BioGenome Project: sequencing life for the future of life. *Proc Natl Acad Sci USA* 2018;115:4325–33. <https://doi.org/10.1073/pnas.1720115115>.
2. Earth Biogenome Project. <https://www.earthbiogenome.org>. Accessed 4 September 2025.

3. Blaxter M, Archibald JM, Childers AK, et al. Why sequence all eukaryotes? *Proc Natl Acad Sci USA*. 2022;119:e2115636118. <https://doi.org/10.1073/pnas.2115636118>.
4. Ellegren H. Evolutionary stasis: the stable chromosomes of birds. *Trends Ecol Evol*. 2010;25:283–91. <https://doi.org/10.1016/j.tree.2009.12.004>.
5. Rebholz WER, Boer LEMD, Sasaki M, et al. The chromosomal phylogeny of owls (Strigiformes) and new karyotypes of seven species. *Cytologia (Tokyo)*. 1993;58:403–16. <https://doi.org/10.1508/cytologia.58.403>.
6. Burt DW. Origin and evolution of avian microchromosomes. *Cytogenet Genome Res*. 2002;96:97–112. <https://doi.org/10.1159/000063018>.
7. Pichugin AM, Galkina SA, Potekhin AA, et al. Estimation of the minimal size of chicken gallus gallus domesticus microchromosomes via pulsed-field electrophoresis. *Russ J Genet*. 2001;37:535–38. <https://doi.org/10.1023/A:1016622816552>.
8. Degrandi TM, Barcellos SA, Costa AL, et al. Introducing the bird chromosome database: an overview of cytogenetic studies in birds. *Cytogenet Genome Res*. 2020;160:199–205. <https://doi.org/10.1159/000507768>.
9. Smith J, Bruley CK, Paton IR, et al. Differences in gene density on chicken macrochromosomes and microchromosomes. *Anim Genet*. 2000;31:96–103. <https://doi.org/10.1046/j.1365-2052.2000.00565.x>.
10. Axelsson E, Webster MT, Smith NGC, et al. Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. *Genome Res*. 2005;15:120–25. <https://doi.org/10.1101/gr.3021305>.
11. Roselaar C. *Strix uralensis* Ural owl. In: Cramp S, ed. *Handbook of the birds of Europe, the Middle East, and North Africa: the birds of the western Palearctic*. Oxford, UK: Oxford University Press; 1985.
12. Able KP. *Handbook of the birds of the world, volume 5, barn-owls to hummingbirds*. Auk. 2000;117:532–34. <https://doi.org/10.1093/auk/117.2.532>.
13. Hausknecht R, Jacobs S, Müller J, et al. Phylogeographic analysis and genetic cluster recognition for the conservation of Ural owls (*Strix uralensis*) in Europe. *J Ornithol*. 2014;155:121–34. <https://doi.org/10.1007/s10336-013-0994-8>.
14. Cramp S, ed. *Handbook of the birds of Europe, the Middle East, and North Africa: the birds of the western Palearctic*. Oxford, UK: Oxford University Press; 1985.
15. König C, Weick F. *Owls of the world*. 2nd ed. London: A & C Black Publishers, 2010.
16. Mikkola H, Willis I. *Owls of Europe*. Calton, UK: T & A D Poyser, 1992.
17. BirdLife International. *Strix uralensis*. In *The IUCN Red List of Threatened Species 2022*;e.T22689108A199908915. <https://dx.doi.org/10.2305/IUCN.UK.2021-3.RLTS.T22689108A199908915.en>. Accessed 5 September 2025.
18. Kopij G. Population and range expansion of forest boreal owls (*Glaucidium passerinum*, *Aegolius funereus*, *Strix uralensis*, *Strix nebulosa*) in East-Central Europe. *Vogelwelt*. 2011;132: 207–14.
19. Scherzinger W. Die Wiederbegründung des Habichtskauz-Vorkommens *Strix uralensis* im Böhmerwald. *Ornithologischer Anzeiger*. 2006;45:97–156.
20. Soorae PS. *Global Re-introduction Perspectives: 2011: More case studies from around the globe*. Abu Dhabi, UAE: IUCN/SSC Re-introduction Specialist Group & Environment Agency—Abu Dhabi, 2011.
21. Scope A, Schwendenwein I, Stanclova G, et al. Exploratory plasma biochemistry reference intervals for Ural Owls (*Strix uralensis*, Pallas 1771) from the Austrian reintroduction project. *J Zoo Wildl Med*. 2016;47:486–92. <https://doi.org/10.1638/2015-0200.1>.
22. Huntley B, Green R, Collingham YC, et al. *A climatic atlas of European breeding birds*. Barcelona, Spain: Lynx; 2007.
23. Lehtikoinen A, Ranta E, Pietiäinen H, et al. The impact of climate and cyclic food abundance on the timing of breeding and brood size in four boreal owl species. *Oecologia*. 2011;165:349–55. <https://doi.org/10.1007/s00442-010-1730-1>.
24. European Commission: Directorate-General for Environment. *The Birds Directive—40 years of conserving our shared natural heritage*. 2019. https://environment.ec.europa.eu/topics/nature-and-biodiversity/birds-directive_en. Accessed 4 September 2025.
25. European Commission: Directorate-General for Environment, Sundseth K. *The Habitats Directive—celebrating 20 years of protecting biodiversity in Europe*. 2012. https://environment.ec.europa.eu/topics/nature-and-biodiversity/habitats-directive_en. Accessed 4 September 2025.
26. The Bern Convention. <https://www.coe.int/en/web/bern-convention>. Accessed 4 September 2025.
27. CITES/Appendices. <https://cites.org/eng/app/appendices.php>. Accessed 4 September 2025.
28. Mooney A, Ryder OA, Houck ML, et al. Maximizing the potential for living cell banks to contribute to global conservation priorities. *Zoo Biol*. 2023;42:697–708. <https://doi.org/10.1002/zoo.21787>.
29. Ryder OA, Onuma M. Viable cell culture banking for biodiversity characterization and conservation. *Annu Rev Anim Biosci*. 2018;6:83–98. <https://doi.org/10.1146/annurev-animal-030117-014556>.
30. Freshney RI. *Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications*. 6th ed. Hoboken, New Jersey, United States: John Wiley & Sons; 2010. <https://doi.org/10.1002/9780470649367>.
31. Hughes P, Marshall D, Reid Y, et al. The costs of using unauthenticated, over-passaged cell lines: how much more data do we need? *BioTechniques*. 2007;43:575–86. <https://doi.org/10.2144/000112598>.
32. Chrysostomakis I, Böhne A, Mozer A. Supplementary figures and tables for: a chromosome-scale reference genome for the Ural owl (*Strix uralensis*) demonstrates the applicability of cell cultures as a sources for genomics of endangered species, 2025. <https://doi.org/10.5281/zenodo.14676512>. Accessed 4 September 2025.
33. CornellLab/Birds of the world. <https://birdsoftheworld.org/>. Accessed 4 September 2025.
34. Yamada K, Nishida-Umehara C, Matsuda Y. A new family of satellite DNA sequences as a major component of centromeric heterochromatin in owls (Strigiformes). *Chromosoma*. 2004;112:277–87. <https://doi.org/10.1007/s00412-003-0267-z>.
35. Takagi N, Sasaki M. A phylogenetic study of bird karyotypes. *Chromosoma*. 1974;46:91–120. <https://doi.org/10.1007/BF00332341>.
36. Zhang G. Bird sequencing project takes off. *Nature*. 2015;522:34. <https://doi.org/10.1038/522034d>.
37. Zhang G, Li C, Li Q, et al. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*. 2014;346:1311–20. <https://doi.org/10.1126/science.1251385>.
38. B10K. <https://b10k.com/>. Accessed 4 September 2025.

39. Feng S, Stiller J, Deng Y, et al. Dense sampling of bird diversity increases power of comparative genomics. *Nature*. 2020;587:252–57. <https://doi.org/10.1038/s41586-020-2873-9>.
40. Baalsrud HT, Garmann-Aarhus B, Enevoldsen ELG, et al. Evolutionary new centromeres in the snowy owl genome putatively seeded from a transposable element. *bioRxiv*. 2024. <https://doi.org/10.1101/2024.07.05.602039>. Accessed 4 September 2025.
41. Forest T, Achaz G, Marbouty M, et al. Chromosome-level genome assembly of the European green woodpecker *Picus viridis*. *G3*. 2024;14:jkae042. <https://doi.org/10.1093/g3journal/jkae042>.
42. Tegelström H, Rytman H. Chromosomes in birds (Aves): evolutionary implications of macro- and microchromosome numbers and lengths. *Hereditas*. 1981;94:225–33. <https://doi.org/10.1111/j.1601-5223.1981.tb01757.x>.
43. Fillon V. The chicken as a model to study microchromosomes in birds: a review. *Genet Sel Evol*. 1998;30:209. <https://doi.org/10.1186/1297-9686-30-3-209>.
44. McQueen HA, Fantes J, Cross SH, et al. CpG islands of chicken are concentrated on microchromosomes. *Nat Genet*. 1996;12:321–24. <https://doi.org/10.1038/ng0396-321>.
45. Schmid M, Nanda I, Guttenbach M, et al. First report on chicken genes and chromosomes. *Cytogenet Genome Res*. 2000;90:169–218. <https://doi.org/10.1159/000056772>.
46. International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*. 2004;432:695–716. <https://doi.org/10.1038/nature03154>.
47. Waters PD, Patel HR, Ruiz-Herrera A, et al. Microchromosomes are building blocks of bird, reptile, and mammal chromosomes. *Proc Natl Acad Sci USA*. 2021;118:e2112494118. <https://doi.org/10.1073/pnas.2112494118>.
48. Borges R, Khan I, Johnson WE, et al. Gene loss, adaptive evolution and the co-evolution of plumage coloration genes with opsins in birds. *BMC Genomics*. 2015;16:751. <https://doi.org/10.1186/s12864-015-1924-3>.
49. Greenwold MJ, Bao W, Jarvis ED, et al. Dynamic evolution of the alpha (α) and beta (β) keratins has accompanied integument diversification and the adaptation of birds into novel lifestyles. *BMC Evol Biol*. 2014;14:249. <https://doi.org/10.1186/s12862-014-0249-1>.
50. Wagner H, Weger M, Klaas M, et al. Features of owl wings that promote silent flight. *Interface Focus*. 2017;7:20160078. <https://doi.org/10.1098/rsfs.2016.0078>.
51. The IUCN Red List. www.iucnredlist.org. Accessed 4 September 2025.
52. Luo H, Lin Q, Fang W, et al. Genomic insights into the endangered white-eared night heron (*Gorsachius magnificus*). *BMC Genom Data*. 2024;25:11. <https://doi.org/10.1186/s12863-024-01194-1>.
53. Robinson JA, Bowie RCK, Dudchenko O, et al. Genome-wide diversity in the California condor tracks its prehistoric abundance and decline. *Curr Biol*. 2021;31:2939–46.e5. <https://doi.org/10.1016/j.cub.2021.04.035>.
54. Li S, Li B, Cheng C, et al. Genomic signatures of near-extinction and rebirth of the crested ibis and other endangered bird species. *Genome Biol*. 2014;15:557. <https://doi.org/10.1186/s13059-014-0557-1>.
55. Li B-P, Kang N, Xu Z-X, et al. Transposable elements shape the landscape of heterozygous structural variation in a bird genome. *Zool Res*. 2025;46:75–86. <https://doi.org/10.24272/j.issn.2095-8137.2024.237>.
56. Pellegrino I, Negri A, Boano G, et al. Evidence for strong genetic structure in European populations of the little owl *Athene noctua*. *J Avian Biol*. 2015;46:462–75. <https://doi.org/10.1111/jav.00679>.
57. Brito PH. Contrasting patterns of mitochondrial and microsatellite genetic structure among Western European populations of tawny owls (*Strix aluco*). *Mol Ecol*. 2007;16:3423–37. <https://doi.org/10.1111/j.1365-294X.2007.03401.x>.
58. Antoniazza S, Burri R, Fumagalli L, et al. Local adaptation maintains clinal variation in melanin-based coloration of European barn owls (*Tyto alba*). *Evolution*. 2010;64:1944–54. <https://doi.org/10.1111/j.1558-5646.2010.00969.x>.
59. Mueller JC, Kuhl H, Boerno S, et al. Evolution of genomic variation in the burrowing owl in response to recent colonization of urban areas. *Proc R Soc B*. 2018;285:20180206. <https://doi.org/10.1098/rspb.2018.0206>.
60. Spielman D, Brook BW, Frankham R. Most species are not driven to extinction before genetic factors impact them. *Proc Natl Acad Sci USA*. 2004;101:15261–64. <https://doi.org/10.1073/pnas.0403809101>.
61. EYu N, Seifert-Eulen M, Boettger T, et al. Eemian and Early Weichselian vegetation and climate history in Central Europe: a case study from the Klinge section (Lusatia, eastern Germany). *Rev Palaeobot Palynol*. 2008;151:72–78. <https://doi.org/10.1016/j.revpalbo.2008.02.005>.
62. Velichko AA, Novenko EY, Pisareva VV, et al. Vegetation and climate changes during the Eemian interglacial in Central and Eastern Europe: comparative analysis of pollen data. *Boreas*. 2008;34:207–19. <https://doi.org/10.1111/j.1502-3885.2005.tb01016.x>.
63. Malkiewicz M. A late Saalian glaciation, Eemian interglacial and early Weichselian pollen sequence at Szklarka, SW Poland—reconstruction of vegetation and climate. *Quat Int*. 2018;467:43–53. <https://doi.org/10.1016/j.quaint.2016.09.026>.
64. Song J, Hua S, Song K, et al. Culture, characteristics and chromosome complement of Siberian tiger fibroblasts for nuclear transfer. *In Vitro CellDevBiol Animal*. 2007;43:203–9. <https://doi.org/10.1007/s11626-007-9043-3>.
65. Alvarez MC, Otis J, Amores A, et al. Short-term cell culture technique for obtaining chromosomes in marine and freshwater fish. *J Fish Biol*. 1991;39:817–24. <https://doi.org/10.1111/j.1095-8649.1991.tb04411.x>.
66. Bolton RL, Mooney A, Pettit MT, et al. Resurrecting biodiversity: advanced assisted reproductive technologies and biobanking. *Reprod Ferti*. 2022;3:R121–46. <https://doi.org/10.1530/RAF-22-0005>.
67. Odoemelam E, Raghavan N, Miller A, et al. Revised karyotyping and gene mapping of the *Biomphalaria glabrata* embryonic (Bge) cell line. *Int J Parasitol*. 2009;39:675–81. <https://doi.org/10.1016/j.ijpara.2008.11.011>.
68. He Z, Wilson A, Rich F, et al. Chromosomal instability and its effect on cell lines. *Cancer Rep*. 2023;6:e1822. <https://doi.org/10.1002/cnr2.1822>.
69. Wenger SL, Senft JR, Sargent LM, et al. Comparison of established cell lines at different passages by karyotype and comparative genomic hybridization. *Biosci Rep*. 2004;24:631–39. <https://doi.org/10.1007/s10540-005-2797-5>.
70. Astrin JJ, Stüben PE. Phylogeny in cryptic weevils: molecules, morphology and new genera of western palaearctic Cryptorhynchinae (Coleoptera:curculionidae). *Invert Systematics*. 2008;22:503. <https://doi.org/10.1071/IS07057>.
71. Ratnasingham S, Hebert PDN. BOLD : The Barcode of Life Data System (<http://www.barcodinglife.org>). *Mol Ecol Notes*. 2007.

- <https://doi.org/10.1111/j.1471-8286.2007.01678.x>. Accessed 4 September 2025.
72. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* 2019;20:257. <https://doi.org/10.1186/s13059-019-1891-0>.
 73. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 2014;15:R46. <https://doi.org/10.1186/gb-2014-15-3-r46>.
 74. Shen W, Le S, Li Y, et al. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One.* 2016;11:e0163962. <https://doi.org/10.1371/journal.pone.0163962>.
 75. Shen W, Sipos B, Zhao L. SeqKit2: a Swiss army knife for sequence and alignment processing. *iMeta.* 2024;3:e191. <https://doi.org/10.1002/imt2.191>.
 76. Miller JR, Delcher AL, Koren S, et al. Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics.* 2008;24:2818–24. <https://doi.org/10.1093/bioinformatics/btn548>.
 77. Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat Commun.* 2020;11:1432. <https://doi.org/10.1038/s41467-020-14998-3>.
 78. Renaud G, Hanghøj K, Korneliussen TS, et al. Joint estimates of heterozygosity and runs of homozygosity for modern and ancient samples. *Genetics.* 2019;212:587–614. <https://doi.org/10.1534/genetics.119.302057>.
 79. Chen S, Zhou Y, Chen Y, et al. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34:i884–90. <https://doi.org/10.1093/bioinformatics/bty560>.
 80. BBMap short read aligner, and other bioinformatic tools. <https://sourceforge.net/projects/bbmap/>. Accessed 4 September 2025.
 81. Cheng H, Concepcion GT, Feng X, et al. Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. *Nat Methods.* 2021;18:170–75. <https://doi.org/10.1038/s41592-020-01056-5>.
 82. Guan D, McCarthy SA, Wood J, et al. Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics.* 2020;36:2896–98. <https://doi.org/10.1093/bioinformatics/btaa025>.
 83. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics.* 2018;34:3094–100. <https://doi.org/10.1093/bioinformatics/bty191>.
 84. Allio R, Schomaker-Bastos A, Romiguier J, et al. Efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. *Mol Ecol Resour.* 2020;20:892–905. <https://doi.org/10.1111/1755-0998.13160>.
 85. Uliano-Silva M, Ferreira JGRN, Krashenninnikova K, et al. Mi-toHiFi: a python pipeline for mitochondrial genome assembly from PacBio high fidelity reads. *BMC Bioinf.* 2023;24:288. <https://doi.org/10.1186/s12859-023-05385-y>.
 86. sanger-tol/PretextView. GitHub. <https://github.com/sanger-tol/PretextView>. Accessed 4 September 2025.
 87. sanger-tol/PretextView. GitHub. <https://github.com/sanger-tol/PretextView>. Accessed 4 September 2025.
 88. rapid-curation. Gitlab. <https://gitlab.com/wtsi-grit/rapid-curation/-/tree/main>. Accessed 4 September 2025.
 89. Darwin Tree of Life. <https://www.darwintreeoflife.org/>. Accessed 4 September 2025.
 90. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25:2078–79. <https://doi.org/10.1093/bioinformatics/btp352>.
 91. Wolff J, Rabbani L, Gilsbach R, et al. Galaxy HiCEXplorer 3: a web server for reproducible hi-C, capture hi-C and single-cell hi-C data analysis, quality control and visualization. *Nucleic Acids Res.* 2020;48:W177–84. <https://doi.org/10.1093/nar/gkaa220>.
 92. Simão FA, Waterhouse RM, Ioannidis P, et al. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics.* 2015;31:3210–12. <https://doi.org/10.1093/bioinformatics/btv351>.
 93. Manni M, Berkeley MR, Seppely M, et al. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol.* 2021;38:4647–54. <https://doi.org/10.1093/molbev/msab199>.
 94. Huang N, Li H. Compleasm: a faster and more accurate reimplement of BUSCO. *Bioinformatics.* 2023;39:btad595. <https://doi.org/10.1093/bioinformatics/btad595>.
 95. Gurevich A, Saveliev V, Vyahhi N, et al. QUAST: quality assessment tool for genome assemblies. *Bioinformatics.* 2013;29:1072–75. <https://doi.org/10.1093/bioinformatics/btt086>.
 96. Rhie A, Walenz BP, Koren S, et al. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol.* 2020;21:245. <https://doi.org/10.1186/s13059-020-02134-9>.
 97. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics.* 2016;32:292–94. <https://doi.org/10.1093/bioinformatics/btv566>.
 98. Challis R, Richards E, Rajan J, et al. BlobToolKit—interactive quality assessment of genome assemblies. *G3.* 2020;10:1361–74. <https://doi.org/10.1534/g3.119.400908>.
 99. Afgan E, Baker D, Batut B, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 2018;46:W537–44. <https://doi.org/10.1093/nar/gky379>.
 100. Baril T, Galbraith J, Hayward A. Earl Grey: a fully automated user-friendly transposable element annotation and analysis pipeline. *Mol Biol Evol.* 2024;41:msae068. <https://doi.org/10.1093/molbev/msae068>.
 101. Tarailo-Graovac M, Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics.* 2009;Chapter 4:Unit 4.10. <https://doi.org/10.1002/0471250953.bi0410s25>.
 102. Flynn JM, Hubley R, Goubert C, et al. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci USA.* 2020;117:9451–57. <https://doi.org/10.1073/pnas.1921046117>.
 103. Kapusta A, Suh A. Evolution of bird genomes—a transposon's-eye view. *Ann NY Acad Sci.* 2017;1389:164–85. <https://doi.org/10.1111/nyas.13295>.
 104. O'Leary NA, Cox E, Holmes JB, et al. Exploring and retrieving sequence and metadata for species across the tree of life with NCBI datasets. *Sci Data.* 2024;11:732. <https://doi.org/10.1038/s41597-024-03571-y>.
 105. Mead D, Ogden R, Meredith A, et al. The genome sequence of the European golden eagle, *Aquila chrysaetos chrysaetos* Linnaeus 1758. *Wellcome Open Res.* 2021;6:112. <https://doi.org/10.12688/wellcomeopenres.16631.1>.
 106. Kriventseva EV, Tegenfeldt F, Petty TJ, et al. OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free software. *Nucleic Acids Res.* 2015;43:D250–56. <https://doi.org/10.1093/nar/gku1220>.

107. tomasbruna/orthodb-clades. GitHub. <https://github.com/tomasbruna/orthodb-clades>. Accessed 4 September 2025.
108. Hoff KJ, Lange S, Lomsadze A, et al. BRAKER1: unsupervised RNA-seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*. 2016;32:767–69. <https://doi.org/10.1093/bioinformatics/btv661>.
109. Gabriel L, Brúna T, Hoff KJ, et al. BRAKER3: fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. *Genome Res*. 2024;34:769–77. <https://doi.org/10.1101/gr.278090.123>.
110. Brúna T, Li H, Guhlín J, et al. Galba: genome annotation with miniprot and AUGUSTUS. *BMC Bioinf*. 2023;24:327. <https://doi.org/10.1186/s12859-023-05449-z>.
111. Dainat J, Hereñú D, Murray DKD, et al. NBISweden/AGAT. Zenodo. 2024. <https://doi.org/10.5281/zenodo.3552717>. Accessed 4 September 2025.
112. Li H, Durbin R. Inference of human population history from individual whole-genome sequences. *Nature*. 2011;475:493–96. <https://doi.org/10.1038/nature10231>.
113. samtools/bcftools. GitHub. <https://github.com/samtools/bcftools>. Accessed 4 September 2025.
114. Béziers P, Roulin A. Sexual maturity varies with melanic plumage traits in the barn owl. *J Avian Biol*. 2021;52:jav.02715. <https://doi.org/10.1111/jav.02715>.
115. Brommer JE, Pietiäinen H, Kolunen H. Reproduction and survival in a variable environment: ural owls (*Strix uralensis*) and the three-year vole cycle. *Auk*. 2002;119:544–50. <https://doi.org/10.1093/auk/119.2.544>.
116. Fujito NT, Hanna ZR, Levy-Sakin M, et al. Genomic variation and recent population histories of Spotted (*Strix occidentalis*) and Barred (*Strix varia*) owls. *Genome Biol Evol*. 2021;13:evab066. <https://doi.org/10.1093/gbe/evab066>.
117. Terhorst J, Kamm JA, Song YS. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat Genet*. 2017;49:303–9. <https://doi.org/10.1038/ng.3748>.
118. He W, Yang J, Jing Y, et al. NGenomeSyn: an easy-to-use and flexible tool for publication-ready visualization of syntenic relationships across multiple genomes. *Bioinformatics*. 2023;39:btad121. <https://doi.org/10.1093/bioinformatics/btad121>.
119. simonhmartin/asynt. Github. <https://github.com/simonhmartin/asynt>. Accessed 4 September 2025.
120. Luo H, Jiang X, Li B, et al. A high-quality genome assembly highlights the evolutionary history of the great bustard (*Otis tarda*, *Otidiformes*). *Commun Biol*. 2023;6:746. <https://doi.org/10.1038/s42003-023-05137-x>.
121. Emms DM, S. Kelly OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019;20:238. <https://doi.org/10.1186/s13059-019-1832-y>.
122. Huerta-Cepas J, Szklarczyk D, Heller D, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res*. 2019;47:D309–14. <https://doi.org/10.1093/nar/gky1085>.
123. Cantalapiedra CP, Hernández-Plaza A, Letunic I, et al. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021;38:5825–29. <https://doi.org/10.1093/molbev/mab293>.
124. Supek F, Bošnjak M, Škunca N, et al. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*. 2011;6:e21800. <https://doi.org/10.1371/journal.pone.0021800>.
125. Raxworthy M. Animal cell culture: a practical approach. *Biochem Educ*. 1987;15:53. [https://doi.org/10.1016/0307-4412\(87\)90173-7](https://doi.org/10.1016/0307-4412(87)90173-7).
126. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009;25:1754–60. <https://doi.org/10.1093/bioinformatics/btp324>.
127. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*. 2013. <https://doi.org/10.48550/arXiv.1303.3997>. Accessed 4 September 2025.
128. Van Der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013;43:11.10.1–11.10.33. <https://doi.org/10.1002/0471250953.b11110s43>.
129. Poplin R, Ruano-Rubio V, DePristo MA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*. 2017. <https://doi.org/10.1101/201178>. Accessed 4 September 2025.
130. Chrysostomakis I, Böhne A, Mozer A. Supplementary figures and tables for: a chromosome-scale reference genome for the Ural owl (*Strix uralensis*) demonstrates the applicability of cell cultures as a sources for genomics of endangered species. *Zenodo*. 2025. <https://doi.org/10.5281/ZENODO.14676512>. Accessed 4 September 2025.
131. Chrysostomakis I, Mozer A, Böhne A. Code and pipelines for: a chromosome-scale reference genome for the Ural owl (*Strix uralensis*) demonstrates the applicability of cell cultures as a sources for genomics of endangered species. *Zenodo*. 2025. <https://doi.org/10.5281/ZENODO.15100180>. Accessed 4 September 2025.
132. Chrysostomakis I, Mozer A, Di-Nizo CB, et al. Supporting data for “A High-Quality Reference Genome for the Ural Owl (*Strix uralensis*) Enables Investigations of Cell Cultures as a Genomic Resource for Endangered Species.” *GigaScience Database*. <https://doi.org/10.5524/102735>.