High-quality genome assemblies provide clues on the evolutionary advantage of blue peafowl over green peafowl

Abhisek Chakraborty, Samuel Mondal, Shruti Mahajan, Vineet K. Sharma [*]

MetaBioSys Group, Department of Biological Sciences, Indian Institute of Science Education and Research Bhopal, Bhopal, 462066, Madhya Pradesh, India

A B S T R A C T

An intriguing example of differential adaptability is the case of two Asian peafowl species, *Pavo cristatus* (blue peafowl) and *Pavo muticus* (green peafowl), where the former has a "Least Concern" conservation status and the latter is an "Endangered" species. To understand the genetic basis of this differential adaptability of the two peafowl species, a comparative analysis of these species is much needed to gain the genomic and evolutionary insights. Thus, we constructed a high-quality genome assembly of blue peafowl with an N50 value of 84.81 Mb (pseudochromosome-level assembly), and a high-confidence coding gene set to perform the genomic and evolutionary analyses of blue and green peafowls with 49 other avian species. The analyses revealed adaptive evolution of genes related to neuronal development, immunity, and skeletal muscle development in these peafowl species. Major genes related to axon guidance such as *NEO1* and *UNC5*, semaphorin (*SEMA*), and ephrin receptor showed adaptive evolution in peafowl species. However, blue peafowl showed the presence of 42% more coding genes compared to the green peafowl along with a higher number of species-specific gene clusters, segmental duplicated genes and expanded gene families, and comparatively higher evolution in neuronal and developmental pathways. Blue peafowl also showed longer branch length compared to green peafowl in the species phylogenetic tree. These genomic insights obtained from the high-quality genome assembly of *P. cristatus* constructed in this study provide new clues on the superior adaptability of the blue peafowl over green peafowl despite having a recent species divergence time.

1. Introduction

Pavo cristatus is colloquially referred to as peacock or Indian peafowl, and is known for its unique ornamental phenotypes, sexual behaviours, and evolutionary significance. The cultural importance of peacock is not only limited to the Indian subcontinent but expands to Persian, Mesopotamian and ancient Greek cultures. The Asian peafowls belong to the order Galliformes, and family Phasianidae that also includes species like chicken, turkey, quail, etc. Phasianidae family consists of terrestrial, short-winged birds, and has more than 180 species. Indian or blue peafowl had origin in the Indian subcontinent, and shows sexual dimorphism, polygamy, and intricate male display during courtship [1]. Indian peafowl and their closest relative green peafowl, which is the only other species from *Pavo* genus, had diverged around 3 million years ago (mya) [2].

Peacocks have intrigued biologists for hundreds of years and are still a fascinating specimen of study [3,4]. From Charles Darwin's explanation for the colourful plumage indicating that the vibrant plumage was selected sexually, to Amotz Zahavi proposing his handicap theory [1], a lot has been added to the understanding of peacock's unique pattern of evolution, still it remains among the most intriguing birds. In addition, factors such as the number of ocelli (eye-spots) in tail [5], certain behavioural factors and calls of peacock had led to the evolution of the distinct traits responsible for sexual selection in peacocks [6]. Gazing pattern of peacocks towards specific display regions plays an important role in intra and intersexual selection [7], and in addition the co-evolution of opsin genes and plumage colouration genes is also associated to sexual selection in birds [8]. To perform complex cognitive activities such as sexual selection, brain size has been evolved for better motor control abilities, well developed neural networks, and highly evolved brain organization [9,10]. Comparative genomic analysis also supported sexual selection in peacock, and showed that feather and immune-related gene-pairs underwent selection pressure in this species [4], consistent with Hamilton-Zuk hypothesis [11].

The native habitat of green peafowl was spread across South East Asia, however this species is now extinct or near extinct in Malaysia, Bangladesh, and India, and is also facing reduction in population size in Thailand, Laos, China, and Indonesia [12,13]. Genomic, anthropogenic, and climatic evidences also suggest the decrease in effective population size or endangerment of green peafowl species, and the role of human disturbance in it [3,14]. Habitat loss has confined this species to restricted geographical regions, which caused a reduction in gene flow and higher rate of inbreeding [14].

Blue peafowl has been categorized as species of "Least Concern" whereas green peafowl has been declared as "Endangered" by IUCN for its gradual decrease in population size [15]. The first genome sequencing of blue peafowl (1.16 Gbp) performed by Jaiswal et al. (2018) [3] showed adaptive evolution of genes related to immunity, skeletal muscle, and feather development that aids in phenotypic evolution of blue peafowl, followed by the genome sequencing of this species by other groups [16,17]. The recent genome sequencing of green peafowl revealed a genome size of 1.05 Gbp consisting of 27 pseudochromosomes [14,18]. In this study, we constructed a genome assembly of blue peafowl with the best assembly contiguity till date by performing 10x Genomics sequencing, Oxford Nanopore sequencing, and Illumina short read sequencing, and using previously available Illumina and Nanopore sequencing data of this species [3,16]. Usage of multiple sequencing technologies and hybrid assembly approaches provided better genomic contiguity and helped in better quality of gene space representation, genome annotation, genomic characterization and in revealing novel evolutionary insights to understand the genomic basis of highly valued traits [19–21].

Further, to investigate the genomic basis of their differential adaptability we performed comparative evolutionary analyses of blue and green peafowl species using the high-quality genome assembly of blue peafowl constructed in this study, which revealed adaptive evolution of genes related to neuronal development along with immunity, and skeletal muscle development related pathways in both the peafowl species. However, the genomic evidence highlights better adaptive evolution of blue peafowl species for survival compared to green peafowl.

2. Materials and methods

2.1. Genome sequencing

The DNA was extracted from the collected blood sample using DNeasy Blood and Tissue kit (Qiagen, CA, USA). DNA was used to prepare a short reads library using NEBNext Ultra II DNA library preparation kit for Illumina (New England Biolabs, England) and sequenced on Illumina HiSeq X instrument (Illumina Inc., USA) for 150 bp paired-end reads. The DNA was amplified using Genomiphi V2 DNA amplification kit. For 10x Genomics linked read sequencing, the amplified DNA library was prepared on a Chromium instrument using Gel Bead Kit v2 and Chromium Genome Library kit (10x Genomics, CA, USA). The prepared library was sequenced on Illumina NovaSeq 6000 instrument (Illumina Inc., USA) for paired-end sequencing. Nanopore long read sequencing library was prepared using the SQK-LSK108 library preparation kit and following the protocol ligation sequencing gDNA. The prepared library was sequenced on a Nanopore sequencer to generate long read data (Oxford Nanopore Technologies, UK).

2.2. Genome assembly

2.2.1. 10x data assembly
Barcoded 10x Genomics data generated from this study was used for *de novo* genome assembly using Supernova v2.1.1 with default parameters and maxreads = all option without any prior pre-processing, and the haplotype-phased genome assembly was obtained using Supernova mkoutput "pseudohap2" style [22].

2.2.2. Illumina data assembly
Illumina short read paired-end sequencing data available from the previous study [3] (generated from the same individual), and data generated from this study were used for *de novo* genome assembly. 10x Genomics data from this study was also filtered for barcode sequences using python scripts available in proc10xG (https://github.com/ucdavis-bioinformatics/proc10xG), and used in this assembly. Prior to genome assembly, all three sets of data were quality-filtered using NGSQCToolkit v2.3 [23] with the same parameters used in a previous study [3]. Quality-filtered data were *de novo* assembled using SPAdes v3.15.3 [24] with k-mer value of 101 as used in the previous study [3].

2.2.3. Nanopore data assembly
Oxford Nanopore data generated in this study and previous study [16] were used for adapter trimming using Porechop v0.2.4 with

default parameters (Oxford Nanopore technologies), and the pre-processed reads were used for *de novo* assembly using Flye v2.9 with "–genome-size 1.13g" and other default parameters [25]. The assembled genome obtained from Nanopore data was polished three times using Pilon v1.23 (default parameters with "–changes" option) [26] with the quality-filtered lllumina short read and 10x Genomics data (barcode-filtered) that were used in the *de novo* assembly performed using SPAdes.

2.2.4. Assembly post-processing and generation of final genome assembly

Three different genome assemblies obtained using three types of sequencing data were scaffolded using the quality-filtered Illumina short read paired-end data from this study and previous studies [3,16], quality-filtered mate-pair data from previous study [16], and the pre-processed Nanopore long read data from this study and previous study [16] using Platanus-allee v2.2.2 "consensus" (with "-IP", "-OP" and "-p" options), separately [27]. The resultant assemblies were further scaffolded with quality-filtered RNA-Seq reads obtained from previous study [28] using AGOUTI "scaffold" v0.3.3 (default parameters) [29]. 10x Genomics reads were used for barcode-processing using Longranger basic v2.2.2 with default parameters (https://support.10xgenomics.com/genome-exome/software/pipelines/latest/installation), and were used for further scaffolding of the assemblies using ARCS v1.2.2 with default parameters [30].

The scaffolded assemblies obtained from 10x Genomics data and Illumina short read data were used for assembly merging using Quickmerge v0.3 with 10x Genomics data-based assembly as the hybrid-assembly, minimum alignment length of 4,000 bases, and other default parameters [31]. The resultant merged assembly was further scaffolded using the Oxford Nanopore data-based assembly using Quickmerge v0.3 with the previously merged assembly as hybrid-assembly, minimum alignment length of 4,000 bases, and other default parameters [31]. The final merged and scaffolded assembly was gap-closed using Sealer v2.1.5 (with k-mer values from 30 to 120 with 10 bp interval) [32], and LR_Gapcloser (default parameters) [33] with the quality-filtered short reads, and the pre-processed Nanopore long reads, respectively, that were used for the *de novo* assembly in the previous steps.

The gap-closed assembly was polished three times to fix any mis-assembly, erroneous base, or small indel using Pilon v1.23 (default parameters with "–changes" option) [26] with the quality-filtered 10x Genomics data (barcode-filtered), and Illumina paired-end short read data that were used for *de novo* assembly in SPAdes. The same quality-filtered short read data was used to further error-correct the Pilon-polished assembly with Seqbug (default parameters) [34], after constructing the individual nucleotide position matrix for each scaffold using bam-readcount (https://github.com/genome/bam-readcount) with the parameters - minimum base quality 25, minimum mapping quality 25, maximum depth 400. Scaffolds with length of ≥5,000 bases were extracted to construct the final genome assembly of blue peafowl.

Presence of any redundant sequences in this genome assembly was checked using Redundans v0.14a with "–noscaffolding", "–nogaplcosing" options and other default parameters [35]. GC-depth distribution analysis was also performed to check any redundancy by mapping the quality-filtered and paired-end 10x Genomics and Illumina short reads onto the genome assembly using QualiMap "bamqc" v2.2.2 with default parameters [36]. To assess the genome assembly completeness, BUSCO v5.2.2 was used with aves_odb10 single-copy orthologous gene set [37]. Also, barcode-filtered 10x Genomics reads (quality-filtered), and quality-filtered Illumina reads from this study and previous study [3] were separately mapped to the final genome assembly of blue peafowl using BWA-MEM v0.7.17 with default parameters to calculate the read mapping percentage [38].

Genomic heterozygosity of blue peafowl was estimated using GenomeScope v2 [39] after constructing the k-mer count histogram using Jellyfish v2.2.10 ("-m 21" and "-s 1200M" parameters) [40] with the quality-filtered Illumina data obtained from this study and the previous study [3]. Sequence variation in blue peafowl genome assembly was identified by mapping barcode-filtered 10x Genomics data (quality-filtered), and quality-filtered Illumina data generated from the same individual using BWA-MEM v0.7.17 [38], and SAMtools v1.9 [41] (default parameters). BCFtools v1.14 was used for variant calling, and variant filtering was performed with the parameters - variant quality ≥30, sequencing depth ≥30, mapping quality ≥50 [42].

The scaffold-level genome of blue peafowl was assembled into pseudochromosomes by mapping onto green peafowl chromosome-level assembly [18] via genome-wide synteny alignment using Chromosemble in Satsuma v2 (default parameters) [43,44]. The genome assembly workflow is schematically represented in Fig. S1.

2.3. Repeat annotation

For identification of repetitive sequences in the final genome assembly of blue peafowl, RepeatModeler v2.0.2a was used with "-LTRStruct" and other default parameters to construct a *de novo* repeat library [45]. Chicken-specific repeat sequences available in Repbase library [46] were also extracted, and added to the *de novo* repeat library obtained from RepeatModeler for better prediction of repeat sequences in the improved genome assembly of blue peafowl. This combined repeat library was used for soft-masking of the blue peafowl genome assembly using RepeatMasker v4.1.2 (http://www.repeatmasker.org) with default parameters.

2.4. Construction of coding gene set

The repeat-masked genome assembly of blue peafowl was used to predict the coding gene sequences using MAKER v3.01.04 with *ab initio* and evidence alignment approaches [47]. Prior to MAKER genome annotation pipeline, quality-filtered RNA-Seq reads from a previous study [28] were used for *de novo* transcriptome assembly of blue peafowl using Trinity v2.13.2 (default parameters) [48]. The protein sequences of blue peafowl, and its phylogenetically closer species - *Gallus gallus*, *Chrysolophus pictus*, *Phasianus colchicus*, *Meleagris gallopavo*, *Coturnix japonica*, *Numida meleagris* (belonging to the same Galliformes order) available in Ensembl genome browser 105 [49] were extracted and used as empirical evidence in MAKER pipeline (first round) along with the *de novo* transcriptome

assembly obtained from Trinity. The parameters were: "augustus_species = chicken", "est2genome = 1", "protein2genome = 1", and "AED_threshold = 1" along with other default values. The first round of MAKER annotation result was used for training SNAP v2006-07-28 gene prediction program [50], and the second round of MAKER annotation. This result was further used for training AUGUSTUS v3.2.3 gene prediction program [51] for our species and SNAP v2006-07-28 gene prediction program, and the training results were used for a third round of MAKER annotation. After the third or final round of MAKER annotation, coding genes were filtered based on AED (Annotation Edit Distance) value < 0.5, and length of ≥150 bases to construct the final high-confidence coding gene set of blue peafowl. Additionally, the high-quality genome assembly of blue peafowl was also used for *de novo* prediction of tRNAs and rRNAs using tRNAscan-SE v2.0.7 with default parameters [52], and Barrnap v0.9 with "–kingdom euk" parameter (https://github.com/tseemann/barrnap), respectively, and for homology-based identification of miRNAs using MirGeneDB v2.1 database [53] using BLASTN (95% query coverage and 95% sequence identity).

2.5. Collinearity analysis

Pseudochromosome-level assembly of blue peafowl was used to identify the intra-species, and inter-species (between blue and green peafowl) collinear blocks. Collinear blocks for these peafowl species were identified using intra-species and inter-species All-versus-All BLASTP alignments (e-value 10^{-5}) using green peafowl protein sequences obtained from a previous study [18], and protein sequences of blue peafowl obtained from this study. MCScanX was used with "-b 1" for intra-species and "-b 2" for inter-species collinearity, and other default parameters [54]. Collinear blocks between the longest chromosomes (≥5 Mb) of green peafowl and the longest pseudochromosomes (≥5 Mb) of blue peafowl were visualized using Evol2Circos [55].

2.6. Phylogenetic analysis

Protein sequences of blue peafowl obtained in this study, green peafowl obtained from a previous study [18], and 49 other avian species available on Ensembl genome browser 105 were used to determine the phylogenetic position of the peafowl species across 12 phylogenetic orders. The selected species from Ensembl database are mentioned in Table S1.

Protein sequences of the selected species were used to construct the orthogroups using OrthoFinder v2.5.4 with default parameters [56]. Among these, the fuzzy one-to-one orthogroups containing sequences from all 51 species were identified using KinFin v1.0 with default parameters [57], and filtered to contain only the longest sequence per species. These filtered orthogroups were individually aligned using MAFFT v7.490 (default parameters) [58], which were concatenated after filtering the empty sites using BeforePhylo v0.9.0 with "-conc = raxml" and "-trim" options (https://github.com/qiyunzhu/BeforePhylo). The concatenated alignment was used to construct the maximum likelihood dependent species phylogenetic tree using RAxML v8.2.12 with 'PROTGAMMAAUTO' substitution model and 100 bootstrap values [59].

2.7. Identification of genes with evolutionary signatures

For the analysis of signatures of adaptive evolution in blue and green peafowl, coding gene information of the six species from Galliformes order available in Ensembl Release 105, green peafowl from previous study [18], and blue peafowl obtained in this study were used. Only the phylogenetically closer species (from Galliformes order itself) were selected for this analysis to identify the more specific genes evolved in this species compared to its closer relatives. Protein sequences of these eight species were used to construct the orthogroups using OrthoFinder v2.5.4 (default parameters) [56], and the orthogroups containing sequences from all eight species were further filtered for the presence of the longest sequence per species. The resultant orthogroups were aligned individually using MAFFT v7.490 with default parameters [58].

These multiple sequence alignments were used to extract the *Pavo* species-specific genes showing unique amino acid positions compared to the other selected species. In this analysis, any gap and ten positions around the gap present in the alignments were ignored. Functional impact of these substitutions on the protein function was analyzed using Sorting Intolerant From Tolerant (SIFT) with UniProt as reference database [60].

Protein sequence alignments of these orthogroups were used to individually construct maximum likelihood-based gene phylogenetic tree using RAxML v8.2.12 with 100 bootstrap values and 'PROTGAMMAAUTO' substitution model [59]. Blue peafowl and green peafowl genes with higher nucleotide divergence compared to genes from other selected species were identified by calculating the branch-length distance values using "adephylo" package available in R [61].

For positive selection analysis, the coding gene sequences (nucleotide) of all eight species were used for orthogroups construction, orthogroups filtering based on the longest sequence per species, and nucleotide sequence alignment using MAFFT v7.490 with default parameters [58]. The nucleotide alignments were converted into PHYLIP format, and used for positive selection analysis using "codeml" program from PAML v4.9a [62], along with the species phylogenetic tree across the eight species (constructed in a similar manner as described earlier). The parameters for "codeml" program were - Model A: model = 2, NSsites = 2, fix_omega = 0, omega = 1; Model A1: model = 2, NSsites = 2, fix_omega = 1, omega = 1. Likelihood-ratio tests were carried out, and blue peafowl and green peafowl genes qualifying against null model with FDR-corrected p-values <0.05 in chi-square analysis were extracted as genes showing positive selection. Further, genes from both the peafowl species containing codon sites with >95% probability for the foreground lineage obtained from Bayes Empirical Bayes (BEB) analysis were identified as genes with positively selected codon sites.

Genes showing at least two of the three evolutionary signatures – higher nucleotide divergence, positive selection, and unique substitution with functional impact were termed as genes with Multiple Signs of Adaptive evolution (MSA) [63,64].

2.8. Evolution of gene families

Evolution of gene families in blue and green peafowl species in terms of gene family expansion or contraction was analyzed using CAFÉ v4.2.1 with respect to 49 other avian species used for species phylogenetic tree construction in this study [65]. All-versus-All BLASTP homology search (with "-outfmt 7" and "-seg yes" options) result using the longest isoforms of the protein sequences of these species was clustered using MCL v14.137 with "-I 3" parameter [66], and gene families filtering was performed as suggested for CAFÉ analysis. The resultant gene families along with the ultrametric species phylogenetic tree that was obtained using the divergence time between blue and green peafowl (from TimeTree database) were used for CAFÉ analysis with two-lambda (λ) model, where species from Galliformes order were assigned a separate λ-value from the other selected species.

2.9. Exon expansion analysis

Exon expansion in the coding gene sequences of the two peafowl species with respect to each other was analyzed. For identification of the orthologous genes in these two species, the protein sequences of these two species were used to construct the orthogroups using OrthoFinder v2.5.4 with default parameters [56]. Only the longest protein sequence per species was retained in each orthogroup. Number of exons expanded or contracted in the orthologous genes of one peafowl species compared to the other was calculated from the respective GFF files.

2.10. Functional annotation

The coding gene set of blue peafowl was mapped against NCBI-nr database using DIAMOND BLASTP with e-value cutoff 10^{-5} and "–sensitive" option [67], and the unmapped genes were mapped against Pfam-A (e-value 10^{-5}) [68] and SwissProt (using DIAMOND BLASTP with e-value 10^{-5} and "–sensitive" option) databases [69]. Protein function prediction for genes without any match against all three databases was performed using InterProScan v5.54–87.0 with default options [70]. Genes that showed evolutionary signatures in both the peafowl species were annotated using KAAS v2.1 web server [71] and eggNOG-mapper v2 (default parameters) [72]. Shared orthologous clusters or gene families among blue peafowl and its phylogenetically closer species (according to the species phylogenetic tree) from Galliformes order were identified using OrthoVenn2 with default options [73].

3. Results

3.1. Genome assembly

A total of 502 Gb (444.3X) of Illumina short read and mate-pair data, 91.8 Gb (81.2X raw sequencing coverage) of 10x Genomics linked read data, and 8.6 Gb (7.6X) of Oxford Nanopore data was used to construct a high-quality genome assembly of blue peafowl (Table S2). A total of 92.7 Gb of transcriptome data of this species from a previous study [28] was also used in downstream analysis to help in the genome assembly and gene prediction. The separate assemblies constructed using 10x Genomics, Illumina short reads, and Nanopore data were used to build the final scaffolded and merged genome assembly of blue peafowl (≥5 Kbp) with a genome size of 1.13 Gbp (comprised of 1,665 scaffolds), N50 value of 4.95 Mb, longest scaffold size of 17.6 Mb (Table S3). After the single scaffolded assembly was constructed using Quickmerge, base correction was performed in the Pilon-polished assembly using SeqBug [34], which corrected 0.002% base positions. Further, 89.5% BUSCOs could be found in the final genome assembly. 93.75% barcode-filtered 10x Genomics linked reads from this study, and 93.82% quality-filtered Illumina short reads from this study and a previous study [3] were mapped on the assembled genome, attesting to the good quality of the genome assembly. The blue peafowl genome was estimated to contain 0.47% heterozygosity (Fig. S2). Sequence variation analysis in the final assembled genome showed that 0.17% of the base positions had single nucleotide variations that were present across 1,471 scaffolds.

Further, Redundans [35] was used to check the presence of any redundant sequences in this genome assembly that might have occurred due to the assembly of two haplotypes in some highly heterozygous regions. 319 scaffolds were identified with redundant sequences, covering a total size of only 9.8 Mb (0.86% of the total assembly), and thus, the redundancy in blue peafowl genome was estimated to be low. This was further supported by the GC-depth distribution analysis in the genome assembly, showing a unimodal distribution with only one significant peak indicating the presence of very less redundant regions (Fig. S3).

Table 1
P. cristatus pseudochromosome-level assembly statistics comparison with previous studies.

Parameters	*P. cristatus* genome assemblies			
	Jaiswal et al. [3]	Dhar et al. [16]	Liu et al. [17]	This study
Total number of scaffolds	98,687	15,025	726	69
Scaffold N50 (bases)	25,613	232,312	11,421,185	84,809,949
Scaffold L50	12,614	_a	_a	5
The longest scaffold length (bp)	286,113	2,488,982	38,857,732	150,848,701
Total repeat content	8.62%	7.33%	15.20%	10.91%

[a] Data not available.

This genome assembly was further improved by constructing a pseudochromosome-level assembly. 438 superscaffolds and 69 pseudochromosomes were constructed from the scaffolds of blue peafowl genome assembly in this study, and this pseudochromosome-level assembly comprised of 1.13 Gbp with an N50 value of 84.81 Mb and covering 89.1% BUSCOs (Table S3). The assembled genome size was same as the previously estimated genome size of 1.13 Gbp [3]. The pseudochromosome-level genome assembly statistics were significantly improved than the previously available *P. cristatus* genome assemblies (Table 1) [3,16,17].

3.2. Genome annotation

Repeat-masking of the blue peafowl genome assembly was performed using a combined repeat library comprising of 206 chicken-specific repeat sequences and 348 *de novo* constructed repeat family sequences obtained from RepeatModeler. 9.87% of the blue peafowl pseudochromosome-level genome assembly consisted of interspersed repeats (5.51% L2/CR1/Rex type of retroelements, 0.86% DNA transposons) and 0.82% of the genome was found to contain simple repeats (Table S4). LINEs were the most prevalent repeat elements in blue peafowl genome, similar to green peafowl [18].

Prior to coding gene prediction, *de novo* transcriptome assembly using previously available data identified 904,608 assembled transcripts (N50 value = 3,873 bp), that were used as empirical evidence in MAKER pipeline. Coding genes prediction using the repeat-masked genome assembly with MAKER genome annotation pipeline identified a total of 25,681 coding gene sequences after three rounds of comprehensive MAKER annotation, and AED value cut-off and length-based filtering. Only 253 (0.99%) out of the 25,681 coding genes were predicted from the contigs that were identified as redundant sequences in the genome assembly. In contrast, the recently sequenced genome of green peafowl contained only 14,935 genes with a higher percentage of genomic repeats (15.92%) [18]. 83.3% genes of the final blue peafowl coding gene set could be annotated using the publicly available databases and InterProScan, which is higher than the previously reported study [17]. 290 tRNAs, 239 miRNAs, and 28 rRNAs were also identified in blue peafowl genome assembly.

3.3. Collinearity and orthologous gene clustering

Intra-species collinearity analysis using the genome assemblies of blue and green peafowls revealed 22.44% of blue peafowl and
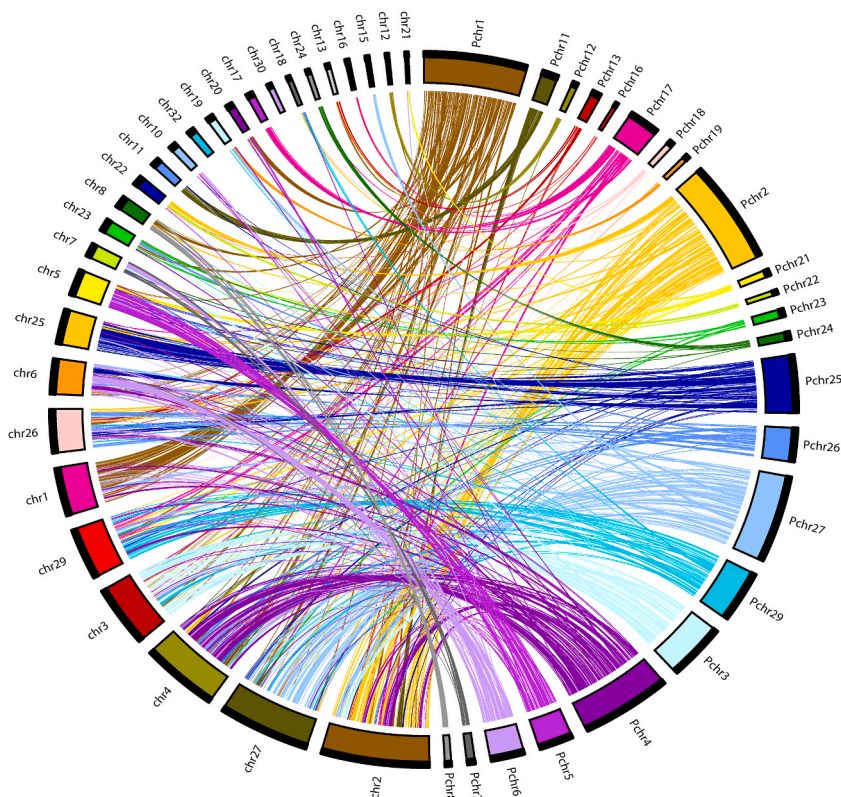

Fig. 1. Circular plot showing collinear blocks identified between the chromosomes of green peafowl and the pseudochromosomes of blue peafowl. Right side of the circle represents the pseudochromosomes of blue peafowl (Pchr = Pseudochromosome), and left side of the circle represents the chromosomes of green peafowl (chr = chromosome). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

8.09% of green peafowl coding genes to be involved in the intra-species collinear blocks of the two respective species. 583 inter-species collinear blocks between the two peafowl species were also identified, which included 48.88% and 74.52% of the coding genes of blue and green peafowls, respectively. The higher percentage of inter-species collinear genes in peafowl species is due to the recent divergence of the two peafowl species. However, the absence of a higher fraction of blue peafowl coding genes in the inter-species collinear blocks points towards a higher number of species-specific coding genes in this species. The inter-species collinear blocks between the longest chromosomes of the two *Pavo* species (≥5 Mb) are visualized in Fig. 1 and Fig. S4. No collinear blocks were observed for chr30 and chr32 of *P. muticus* since their chromosomal counterparts in *P. cristatus* (Pchr30 and Pchr32, respectively) were <5 Mb–5.7 Kb and 2.3 Mb, respectively, and thus could not be shown in Fig. 1.

Also, the number of gene clusters was found to be similar (13,210 - 14,301) in blue peafowl, green peafowl, and other species from Galliformes order, however a higher number of species-specific clusters was observed in blue peafowl species (Fig. 2A-B) that also supports the presence of a higher number of species-specific coding genes in blue peafowl. Further, a large number (11,124) of genes were identified in the species-specific gene clusters of blue peafowl, among which only 137 (1.23%) were contributed by the redundancy in the genome assembly. The genes present in blue peafowl specific gene clusters were mapped onto the green peafowl transcriptome assembly (assembled using 89 million RNA-Seq reads using Trinity v2.13.2 with default parameters [18]) using BLASTN with the parameters - query coverage >90%, sequence identity >90%, and e-value 10^{-9}, which showed the presence of only 1.74% of these genes in green peafowl. Genes included in the species-specific gene clusters of blue and green peafowl are involved in various KEGG pathways mentioned in Table S5.

3.4. Phylogenetic position

Phylogenetic position of the two peafowl species was determined with respect to 49 other avian species available in Ensembl genome browser 105 [49]. A total of 1,441 fuzzy one-to-one orthogroups were identified across all 51 selected bird species, and 899, 247 sequence alignment positions were used to construct the species phylogenetic tree. Blue and green peafowl belonging to the *Pavo* genus were positioned in the same clade and were found to be the closest to each other in the phylogenetic tree (Fig. 3). However, blue peafowl had longer branch length than green peafowl in the species phylogenetic tree. Among other species from Galliformes order, *Gallus gallus* was the closest phylogenetically placed species of these peafowl species, and *Numida meleagris* had diverged the earliest. Species from Anseriformes order were phylogenetically closest to the species from Galliformes order compared to species from other avian phylogenetic orders.
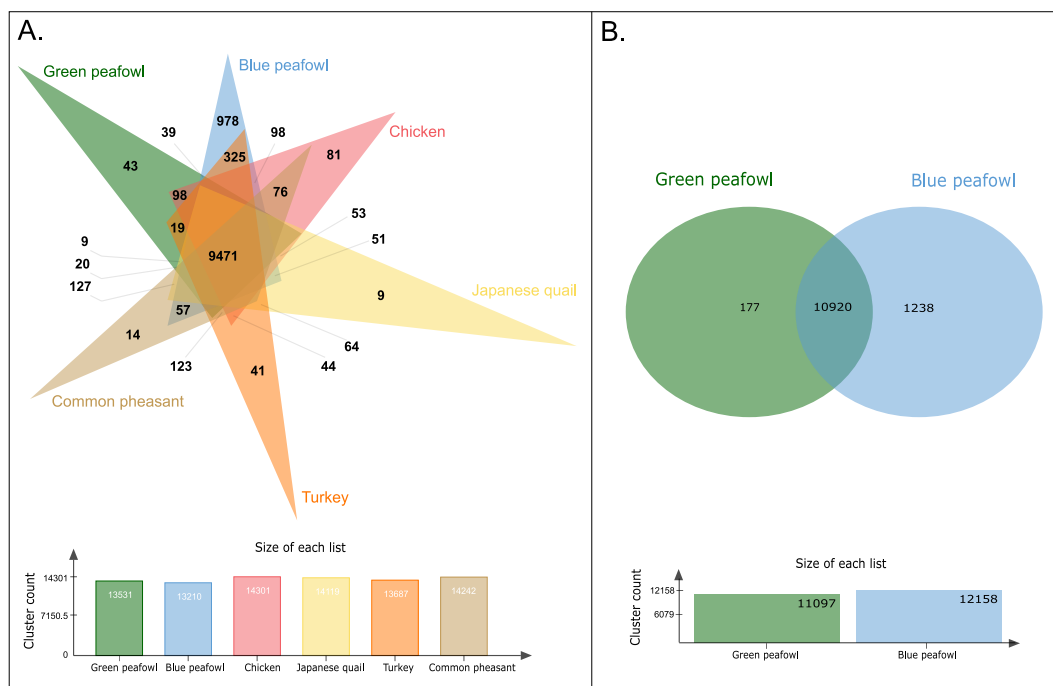
Fig. 2. Shared orthologous gene clusters obtained from OrthoVenn2 analysis. **A.** Shared orthologous clusters among blue peafowl and five other phylogenetically closer birds from Galliformes order, **B.** Shared orthologous clusters between blue and green peafowl species. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
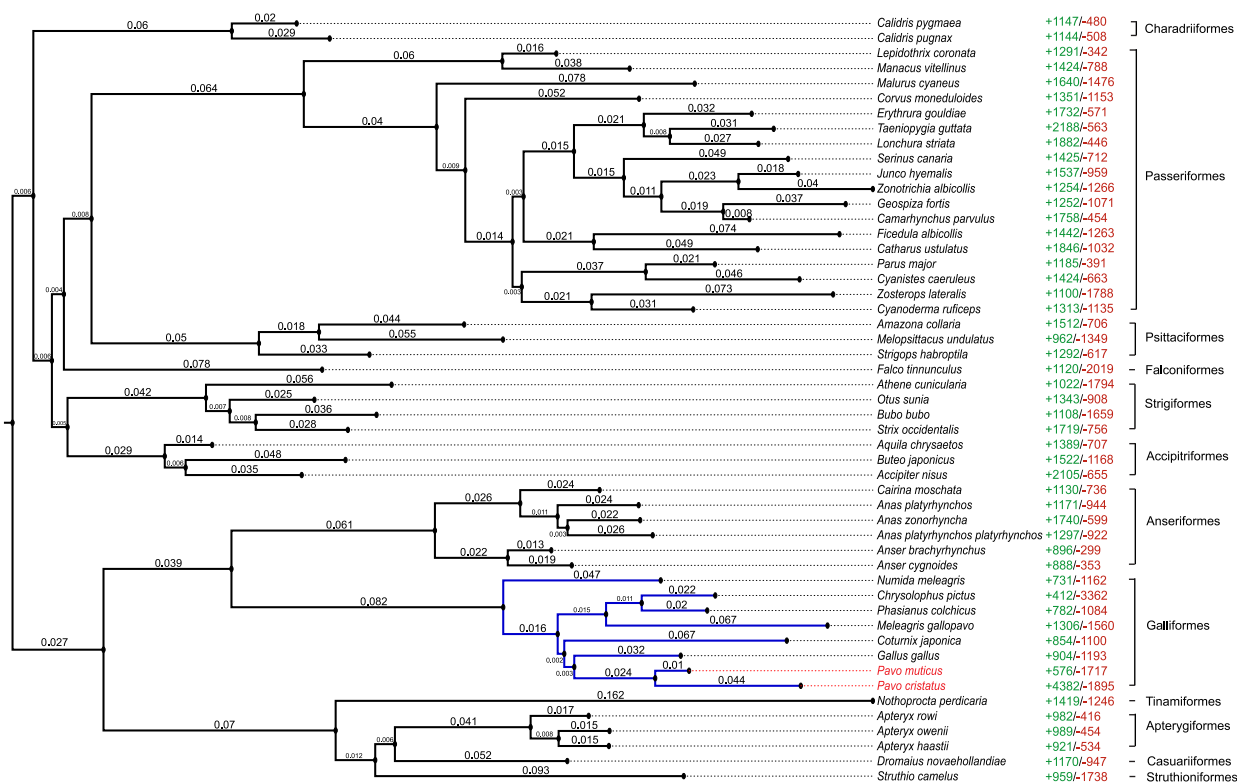
Fig. 3. Phylogenetic position of the peafowl species with respect to 49 other avian species available in Ensembl release 105. The numbers mentioned in the branches of the phylogenetic tree represent the branch length values. Numbers mentioned in green and red denote to the numbers of expanded and contracted gene families, respectively, in each species. The divergence times for each node in the phylogenetic tree were estimated from the TimeTree database v5 [74], and are shown in Fig. S5. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.5. Genome-wide exon expansion

The analysis of orthogroups (containing genes of the two peafowl species) revealed 4,119 genes in blue peafowl, and 3,960 genes in green peafowl that showed expansion in exon numbers with respect to each other. 16 genes in each of the two *Pavo* species showed ≥10 times expansion in terms of exon number along with increased coding gene length and decreased average exon length in most cases (Table S6). Among these genes, titin had the highest number of expanded exons in blue peafowl (117 exons in blue peafowl and 4 exons in green peafowl). Titin provides structural continuity of a sarcomere in vertebrate striated muscles, and thus helps in contraction of avian flight muscle [75].

In contrast, *COL11A1* showed the highest number of expanded exons in green peafowl (3 exons in blue peafowl and 65 exons in green peafowl) followed by another collagen gene *COL24A1*. Collagens are responsible for structure and strength of connective tissues to support body parts such as cartilage, muscles, and organs [76]. Association of *Pavo* genes showing exon expansion with KEGG pathways are mentioned in Table S7.

3.6. Evolution of gene families

A total of 12,107 filtered (size and clade-based filtering) gene families were identified across the 51 avian species used for species phylogenetic tree construction. In blue peafowl, 4,382 and 1,895 gene families were expanded and contracted, respectively. 126 (2.88%) out of the 4,382 expanded gene families in blue peafowl contained the redundant genes identified in the blue peafowl genome assembly. In green peafowl, 576 gene families were expanded, and 1,717 families were contracted. 174 and 75 highly expanded gene families (with >10 expanded genes) were obtained in blue and green peafowl, respectively, among which 69 were common in both the peafowl species (Table S8). Further, the genes (11,124) present in blue peafowl specific gene clusters were present in 10,641 gene families, among which 2,584 gene families did not contain any green peafowl genes. Among the 11,124 genes present in blue peafowl specific gene clusters, 6,370 genes were gained from the expansion of blue peafowl gene families, 1,183 genes were resulted from both expansion of blue peafowl and contraction of green peafowl gene families, and 471 genes were absent in green peafowl due to contraction of their corresponding gene families.

Furthermore, these 69 highly expanded common gene families were majorly involved in neuronal and signalling KEGG pathways - axon guidance, pathways of neurodegeneration, cytokine-cytokine receptor interaction, chemokine signalling pathway, Hippo signalling pathway, signalling pathways regulating pluripotency of stem cells, neuroactive ligand-receptor interaction, endocytosis, Rap1 signalling pathway, TGF-beta signalling pathway, Ras signalling pathway, cholinergic synapse, melanogenesis, MAPK signalling pathway, Wnt signalling pathway, regulation of actin cytoskeleton, Phospholipase D signalling pathway, and others. However, in terms of gene numbers, the highly expanded common gene families in blue peafowl contained a larger number of genes compared to green peafowl.

3.7. Genes with evolutionary signatures

For the identification of evolutionary signatures in the two peafowl species, a total of 8,527 orthogroups were constructed across eight species from Galliformes order. Compared to the other species, 1,077 genes showed higher nucleotide divergence, 605 genes were positively selected (p-values <0.05), and 700 genes contained unique amino acid substitutions with functional impact in blue peafowl. Among these genes, 429 genes were identified as MSA genes, and 41 genes showed all three evolutionary signatures. In green peafowl, 174 genes showed higher nucleotide divergence, 660 genes showed unique substitution with functional impact, and 142 genes showed positive selection (p-values <0.05), among which 110 genes were MSA genes, and 17 genes displayed all three evolutionary signatures. Distribution of *Pavo* genes with evolutionary signatures in different KEGG pathways and COG categories are mentioned in Tables S9–S16.

The MSA genes of blue peafowl were involved in KEGG pathways including neurodegeneration, endocytosis, axon guidance, protein processing in endoplasmic reticulum, PI3K-Akt signalling, and other pathways (Table S12). The MSA genes of green peafowl were involved in KEGG pathways - MAPK signalling pathway, Ras signalling pathway, calcium signalling pathway, PI3K-Akt signalling pathway, endocytosis, neuroactive ligand-receptor interaction, and others (Table S12). Among the genes that showed all three evolutionary signatures, genes related to neuronal development, immune response, and cytoskeletal functions were prominent in both the peafowl species.

3.7.1. Adaptive evolution in nervous system development-related genes in peafowl species

Genes related to various processes related to neuronal development such as axon guidance, and neuronal differentiation have been found among the genes showing all three evolutionary signatures and were also among the expanded gene families. Among the key *P. cristatus* genes with all three evolutionary signatures (Table S17), *NEO1* and *UNC5* are receptors known for Netrin-dependent repulsive axon guidance functions [77], *ZC3H14* and *CHRDL1* function in neuronal differentiation [78,79]. *POSTN*, an extracellular matrix protein also aids in axon regeneration and neurite outgrowth activity [80]. Genes with other neuronal functions such as facilitation of sodium-activated potassium channel activity in neurons (by *Ano3* gene) [81], and modulating neuropeptide signalling (by *NEP* gene) [82] also showed all evolutionary signatures. Among the key *P. muticus* genes showing all three evolutionary signatures (Table S18), *DUSP* functions in development of neural cells [83], *CSMD2* has a role in development of synapse and dendrites [84], *SEMA5* acts in neural development by means of axonal guidance [85], and *ADGRG6* functions in myelination and development of Schwann cells [86].

Among the highly expanded gene families that are common in both peafowl species, ephrin receptor, ADAM family, TRIM family, semaphorin, olfactory receptor, kelch-like protein, Wnt family, and nicotinic acetylcholine receptor families are noteworthy (Table S8). Ephrin receptors are one of the key regulators of axon guidance through Rho GTPase activities [87], ADAM proteins regulate ephrin signalling by means of their proteolytic cleavage, and also have roles in Notch signalling pathway that is another major pathway in neuronal development [88]. TRIM family proteins regulate netrin receptor signalling responsible for axon guidance through their ubiquitin ligase activities [89]. Semaphorins also act as major axonal guidance cues, and function in nervous system development [85]. Olfactory receptor genes exhibit adaptive evolutionary evidence in a wide range of bird species, and also showed highly expanded gene families in these peafowl species. Olfactory receptor gene also showed higher nucleotide divergence in green peafowl, unique substitution with functional impact in both the peafowl species, and had expanded exon number in blue peafowl, which suggest that olfactory cues are important for the peafowl species [90]. The actin-binding proteins in neurons - kelch-like proteins were also found among the highly expanded gene families, which supports the expansion of kelch-like protein family during progressive evolution of animal species [91]. However, in the highly expanded common gene families, gene families showed more expansion in blue peafowl compared to green peafowl in terms of gene numbers.

3.7.2. Adaptive evolution in immunity and cytoskeletal genes in Asian peafowls

Immunity-related genes were also found among the genes showing all evolutionary signatures (Tables S17–S18) and had highly expanded gene families. Among the *P. cristatus* genes with three evolutionary signatures, *RASAL2* and *MAP2K2* are two key enzymes involved in Ras signalling and MAPK signalling pathways, respectively, phospholipase C functions in regulation of Toll-like receptor (TLR) signalling [92], and *CNOT2* represses transcription of MHC II genes [93]. Among the *P. muticus* genes with three evolutionary signatures, Bruton's tyrosine kinase (*BTK*) acts in B cell proliferation and differentiation as well as neuronal differentiation [94]. Among the highly expanded gene families common between both peafowl species (Table S8), MAPK, C-type lectin domain family, chemokine receptor, forkhead box protein family, and receptor protein serine threonine kinase families are noteworthy examples of immunity-related gene families [95–97].

In both the peafowl species, genes related to cytoskeletal functions were found as genes with three evolutionary signatures (Tables S17–S18). Among the *P. cristatus*-specific genes, myosin interacting protein *MyRIP* recruits myosin VIIa to retinal melanosomes

[98], *NPL* aids in cardiac and skeletal muscle function through sialic acid catabolism [99], *POMGNT2* is responsible for functioning of extracellular matrix (ECM) proteins [100]. Among the *P. muticus*-specific genes related to cytoskeletal activities, *CDH17*, *DOCK*, *SYT13*, *CEP170*, and *ADAMTS15* were notable. Further, cytoskeletal activities related gene families - gap junction protein family, claudin, dynein heavy chain, calpain, cadherin, actin, myosin, tubulin, kinesin, septin, and *ADAMTS* were noteworthy among the highly expanded gene families of both the peafowl species (Table S8). Notably, calpain genes are known to bind with titin and stabilize the sarcomere [101], and titin also showed the highest number of exon expansion in blue peafowl.

3.7.3. Adaptive evolution in feather development and visual genes

Among the genes showing evolutionary signatures in blue peafowl, *EDNRB* and *CBL* showed higher nucleotide divergence compared to other species from Galliformes order, and *CBL* and *Nectin1* were among the positively selected genes. Further, *EDNRB* showed exon expansion in blue peafowl, and *PMEL*, *MC1R*, *Nectin1*, and *CBL* showed exon expansion in green peafowl. All these genes are involved in feather color determination and melanin deposition in avian species [17,102]. Adaptive evolution of *MC1R* is also correlated with the level of sexual dichromatism as well as sexual selection in species from Galliformes order [103]. Further, genes from highly expanded gene families common in both species, positively selected genes in both the peafowl species, and MSA genes of blue peafowl were involved in melanogenesis (KEGG pathway) (Tables S10 and S12). Additionally, genes associated with TGF-β, Wnt, and MAPK signalling pathways that are also involved in feather development were found among the adaptively evolved genes, in accordance with the previous studies [3,17].

Birds have evolved their visual system as they heavily rely on the same to adapt in various light conditions [8]. Among the opsin genes that developed in a non-neutral way in birds [8], *RRH*, *RH1*, and *VA* had greater number of exons in blue peafowl compared to green peafowl, and *OPNP* showed exon expansion in green peafowl.

4. Discussion

The peafowl species are known for their unique phenotypic characteristics and evolutionary importance. In this study, the high-quality genome assembly of blue peafowl obtained using Illumina short reads, 10x Genomics sequencing, and Oxford Nanopore long read technology helped in constructing a pseudochromosome-level assembly consisting of 69 pseudochromosomes [18] with the highest N50 value (84.81 Mb) of blue peafowl genome till date. This high-quality assembly helped in gaining important evolutionary and comparative insights on the two peafowl species and can be used as a valuable reference genome for future studies of these intriguing bird species.

The final coding gene set of blue peafowl was constructed using a combination of *ab initio* and homology-based approach with AED value and length-based filtering criteria to ensure a good quality gene set. The number of genes was 42% higher in blue peafowl compared to green peafowl. One of the reasons for the greater number of coding genes predicted in blue peafowl is likely due to the usage of larger genomic data with much more sequencing coverage compared to the previous studies on blue peafowl genome [3,16, 17]. However, to examine if the higher number of genes is due to genomic redundancy, we checked for the presence of any redundancy in the blue peafowl genome assembly and found the presence of only 0.86% (9.8 Mb) redundant sequences in the genome and 0.99% redundant coding genes.

The difference in gene numbers of the two peafowl species can also be supported by multiple observations. The number of coding genes in blue peafowl genome was also higher (23,153 genes) in a previous study [16] and noted in our study with much more sequence data and with a better assembly quality. The difference between the coding gene number in the two species does not appear to be due to the incomplete genome assembly of green peafowl since the assembled genome size (1.049 Gb) of green peafowl is close to the estimated genome size of green peafowl (1.05 Gb), and the complete BUSCO score of green peafowl genome assembly was high (97.6%) indicating its near-completeness [18]. Further, two recent independent studies have reported the presence of similar number of genes in green peafowl - 14,935 genes [18], and 15,584 genes [14], and one of the two studies has also reported the high BUSCO completeness (97.1%) of the coding gene set [18] suggesting the coding gene set of green peafowl was also nearly complete. Further, BLASTN mapping of the blue peafowl specific coding genes onto the green peafowl transcriptome with not so stringent parameters showed the presence of only a minor fraction (1.74%) of the genes in green peafowl, suggesting the genes were specific to blue peafowl. Species-specific nature of the genes in blue peafowl genome can also be explained by the presence of a lower percentage of blue peafowl coding genes (48.88%) in the inter-species collinear blocks, compared to green peafowl. Further, we found that both expansion of blue peafowl and contraction of green peafowl gene families have contributed to the higher number of coding genes in blue peafowl.

A comprehensive genome-wide phylogenetic tree constructed with all available avian species in Ensembl 105 along with green peafowl species helped in better resolution of phylogenetic position of peafowls and showed that the clade formed by both *Pavo* species was closest to *Gallus gallus* among other Galliformes order species. This observation provides further support to the previous studies [3, 16], except a study where *P. cristatus* was closer to *Meleagris gallopavo* perhaps due to the inclusion of lower (15 species) number of species in their phylogenetic tree [17].

The genes related to pathways of immunity and cytoskeleton were found to be adaptively evolved in blue peafowl in previous studies [3,17]. Comparative evolutionary analyses of blue peafowl, green peafowl, and six other species from the Galliformes order in this study showed the genes majorly involved in neuronal development along with immunity, skeletal muscle development, feather, and visual system development to be adaptively evolved in both the peafowl species. Adaptive evolution of immune response-related genes is responsible for the immunocompetence of peafowl species required for sexual selection, whereas skeletal muscle and feather development explain their large body size and decorative feathers [3,17]. However, in this study the key genes related to neuronal

development showed evolutionary signatures, gene family expansion and exon expansion that provide additional genomic insights into the adaptive evolution in the two peafowl species.

For performing specialized cognitive activities such as sound localization, foraging, and assessing rival individuals, avian species possess complicated neural networks and increased brain size, which makes birds as one of the brainiest evolved organisms among the vertebrates [10,104]. In addition, males and females evaluate their mates, and males also assess their rival males through gazing that is one of the most important cognitive functions in peacock and examples of intersexual and intrasexual selection activities [7], for which they would require an evolved nervous system. To support this, one of the key results from this study showed adaptive evolution of genes involved in functioning of nervous system such as axon guidance and neuronal differentiation in peafowl species. Neuron functioning largely relies on development of proper axonal connections, which depends on attractive or repulsive forces generated by the binding of guidance molecules with the receptors present on growth cones, a phenomenon known as axon guidance [87]. Extracellular matrix proteins such as cadherin, and extracellular developmental proteins such as Wnt are also involved in axon guidance and synapse formation in the nervous system [105,106], and were highly expanded in both peafowl species. Besides this, adaptively evolved genes related to neuronal differentiation in different regions of brain were also identified in these peafowl species. Taken together, these observations indicate adaptive evolution in nervous system in peafowl species that perhaps helps in sexual selection and other complex cognitive activities [4].

One of the interesting observations from this study was the increase in exon-intron numbers in the orthologous genes of the two peafowl species that perhaps occurred due to inclusion of new exons which can arise from splice motifs or upstream shortened introns [107], and gain of exons could be associated with increased gene expression by activating new transcription start sites (TSSs) - a phenomenon known as Exon-mediated activation of transcription starts (EMATS) [108]. The exon-intron expansion was also prominent in opsin and plumage colouration genes that can be associated with the distinct phenotypic characteristics of peafowl, along with nervous system and immunity-related genes. Previous studies have suggested an adaptive association between opsin and plumage colouration genes, which is linked to the advantages during sexual selection [8]. This corroborates well with the case of peafowl species, given the importance of visual cues in sexual selection in peafowls [7].

The two Asian peafowl species, green peafowl and blue peafowl provide an intriguing example of differential adaptability where the former is left with a limited number of individuals and is an "Endangered" species, whereas the latter is a species of "Least Concern". This study provides genomic and evolutionary clues for the difference in the survival and existence status of the two peafowl species (Fig. 4). Although blue and green peafowl had a species divergence time of only 3 mya, the striking contrast in number of coding genes present in blue peafowl (25,681 genes) compared to green peafowl (14,935 genes [18]) is indeed intriguing. This perhaps can be explained by the presence of a higher percentage of segmental duplicated genes, a greater number of expanded gene families, a lesser percentage of genes present in inter-species collinear blocks, and a higher number of species-specific gene clusters in blue peafowl (Fig. 2B) compared to green peafowl.

Further, a larger number of genes showed higher nucleotide divergence and positive selection in blue peafowl compared to green peafowl, indicating adaptive evolution to be more prominent in blue peafowl. The signatures of adaptive evolution in genes associated with nervous system development, immunity, and other functions were also higher in blue peafowl (Tables S10–S12). It was also
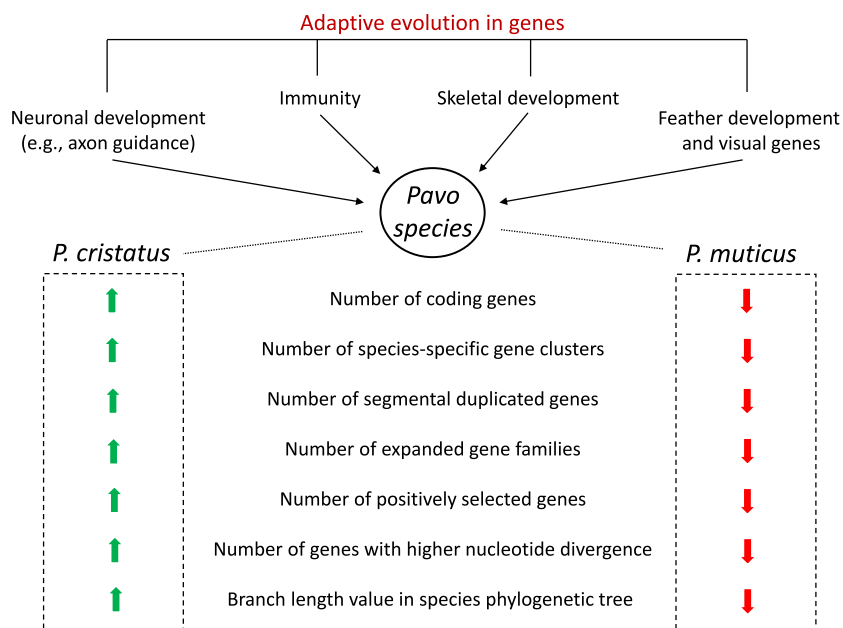
Factors contributing to the evolutionary advantage of blue peafowl over green peafowl

Fig. 4. Comparative genomic and evolutionary characteristics of the two peafowl species.

supported by the phylogenetic analysis where the branch length was observed to be comparatively longer in case of blue peafowl compared to green peafowl suggesting a higher rate of evolution in blue peafowl species (Fig. 3). These findings provide valuable clues on the better adaptability and survival of the blue peafowl compared to the green peafowl. Besides these, PSMC results suggested a more recent second population bottleneck event in green peafowl (∼20,000–10,000 years ago) [14] compared to blue peafowl (∼450, 000 years ago) [3]. The effective population size of green peafowl experienced a steep decline during the recent second population bottleneck event, whereas blue peafowl population was comparatively stable [3,14].

The social behavior and adaptability of the two Asian peafowls also appear to contribute to their contrasting populations sizes. The impact of habitat loss and exploitation by humans for food and commercial purposes seems to have affected the green peafowl population more since it is less tolerant to human activities. A reduction in population contributes to gene flow reduction, high inbreeding, low genetic diversity, and thus leads to a higher extinction possibility [14]. Thus, despite having a recent divergence time, the above-mentioned factors seem to have contributed to the distinct genomic divergence of these peafowl species, and a reduction in population size of green peafowl. This was also the case in Felidae family that evolved in eight lineages over a time period of only six million years [109], where cheetah showed lower genetic diversity and faced a population decline compared to other felid species mostly because of a recent population bottleneck event, loss of habitat, and difficulties in captive breeding, similar to the case of green peafowl [109,110].

Taken together, it is tempting to speculate that the survival and adaptability of an organism appears to be a complex interplay of adaptive evolution, social behavior and environmental influences, and the case of the two peafowls is one such intriguing example that needs more studies to determine the quantitative impacts of these factors. Further, the high-quality genome assembly of *P. cristatus* constructed in this study will act as a valuable reference for future studies of these intriguing bird species.

5. Conclusion

In this study, we constructed a high-quality assembly of blue peafowl (*P. cristatus*) genome and performed comparative analyses of blue and green peafowl species to understand the genomic basis of differential adaptability of these two species. This study revealed adaptive evolution of nervous system developmental genes along with immunity, and skeletal muscle development genes in both the peafowl species. However, blue peafowl showed better adaptive evolution compared to green peafowl in presence of higher number of expanded gene families, segmental duplicated genes, species-specific gene clusters, and genes with evolutionary signatures, which highlights the distinct genomic divergence and provides genomic clues on the contrasting population size of the two Asian peafowl species.

Ethics approval

This study was approved by the Institute Ethics Committee, IISER Bhopal.

Author contribution statement

Abhisek Chakraborty: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Samuel Mondal, Shruti Mahajan: Performed the experiments; Wrote the paper.

Vineet K. Sharma: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Data availability statement

Data associated with this study has been deposited at NCBI SRA database under the BioProject accession – PRJNA891214, Bio-Sample accession – SAMN31312556.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e18571.

References

[1] A. Zahavi, Mate selection—a selection for a handicap, J. Theor. Biol. 53 (1975) 205–214, https://doi.org/10.1016/0022-5193(75)90111-3.

[2] Y. Ouyang, Z. Yang, D. Li, J. Huo, K. Qian, Y. Miao, Genetic Divergence between Pavo muticus and Pavo Cristatus by Cyt B Gene, J. Yunnan Agric. Univ., 2009.

[3] S.K. Jaiswal, A. Gupta, R. Saxena, V.P.K. Prasoodanan, A.K. Sharma, P. Mittal, A. Roy, A.B.A. Shafer, N. Vijay, V.K. Sharma, Genome sequence of peacock reveals the peculiar case of a glittering bird, Front. Genet. (2018), https://doi.org/10.3389/fgene.2018.00392.

[4] S.K. Jaiswal, A. Gupta, A.B.A. Shafer, P.K. Vishnu Prasoodanan, N. Vijay, V.K. Sharma, Genomic insights into the molecular basis of sexual selection in birds, Front. Ecol. Evol. (2021), https://doi.org/10.3389/fevo.2021.538498.

[5] M. Petrie, H. Tim, S. Carolyn, Peahens prefer peacocks with elaborate trains, Anim. Behav. 41 (1991) 323–331, https://doi.org/10.1016/S0003-3472(05)80484-1.

[6] A. Loyau, M. Saint Jalme, G. Sorci, Intra- and intersexual selection for multiple traits in the peacock (Pavo cristatus), Ethology (2005), https://doi.org/10.1111/j.1439-0310.2005.01091.x.

[7] J.L. Yorzinski, G.L. Patricelli, S. Bykau, M.L. Platt, Selective attention in peacocks during assessment of rival males, J. Exp. Biol. (2017), https://doi.org/10.1242/jeb.150946.

[8] R. Borges, I. Khan, W.E. Johnson, M.T.P. Gilbert, G. Zhang, E.D. Jarvis, S.J. O'Brien, A. Antunes, Gene loss, adaptive evolution and the co-evolution of plumage coloration genes with opsins in birds, BMC Genom. (2015), https://doi.org/10.1186/s12864-015-1924-3.

[9] P.J. Makovicky, S. Reddy, Evolution: brainier birds, Curr. Biol. (2020), https://doi.org/10.1016/j.cub.2020.05.025.

[10] T. Nomura, E.I. Izawa, Avian brains: insights from development, behaviors and evolution, Dev. Growth Differ. (2017), https://doi.org/10.1111/dgd.12362.

[11] W.D. Hamilton, M. Zuk, Heritable true fitness and bright birds: a role for parasites? Science 80– (1982) https://doi.org/10.1126/science.7123238.

[12] P.J.K. McGowan, J.W. Duckworth, W. Xianji, B. Van Balen, Y. Xiaojun, M.K.M. Khan, S.H. Yatim, L. Thanga, I. Setiawan, R. Kaul, A review of the status of the Green Peafowl Pavo muticus and recommendations for future action, Bird. Conserv. Int. (1998), https://doi.org/10.1017/S0959270900002100.

[13] D. Kong, F. Wu, P. Shan, J. Gao, D. Yan, W. Luo, X. Yang, Status and distribution changes of the endangered Green Peafowl (Pavo muticus) in China over the past three decades (1990s-2017), Avian. Res. (2018), https://doi.org/10.1186/s40657-018-0110-0.

[14] F. Dong, H.C. Kuo, G.L. Chen, F. Wu, P.F. Shan, J. Wang, D. Chen, F.M. Lei, C.M. Hung, Y. Liu, X.J. Yang, Population genomic, climatic and anthropogenic evidence suggest the role of human forces in endangerment of green peafowl (Pavo muticus), Proc. R. Soc. A B. 288 (2021), https://doi.org/10.1098/RSPB.2021.0073.

[15] F. Wu, D.J. Kong, P.F. Shan, J. Wang, G.N. Kungu, G.Y. Lu, X.J. Yang, Ongoing green peafowl protection in China, Zool. Res. (2019), https://doi.org/10.24272/j.issn.2095-8137.2019.069.

[16] R. Dhar, A. Seethy, K. Pethusamy, S. Singh, V. Rohil, K. Purkayastha, I. Mukherjee, S. Goswami, R. Singh, A. Raj, T. Srivastava, S. Acharya, B. Rajashekhar, S. Karmakar, De novo assembly of the Indian blue peacock (Pavo cristatus) genome using Oxford Nanopore technology and Illumina sequencing, GigaScience 8 (2019), https://doi.org/10.1093/GIGASCIENCE/GIZ038.

[17] S. Liu, H. Chen, J. Ouyang, M. Huang, H. Zhang, S. Zheng, S. Xi, H. Tang, Y. Gao, Y. Xiong, D. Cheng, K. Chen, B. Liu, W. Li, J. Ren, X. Yan, H. Mao, A high-quality assembly reveals genomic characteristics, phylogenetic status, and causal genes for leucism plumage of Indian peafowl, GigaScience 11 (2022), https://doi.org/10.1093/GIGASCIENCE/GIAC018.

[18] X. Zhang, C. Lin, H. Li, S. Liu, Q. Wang, S. Yang, M. Shi, S.K. Sahu, Y. Zhu, J. Wang, J. Huang, Y. Hu, J. Yu, S. Zhang, G. Li, W. Guan, H. Lu, T. Lan, Y. Xu, Chromosome-level genome assembly of the green peafowl (Pavo muticus), Genome Biol. Evol. (2022), https://doi.org/10.1093/GBE/EVAC015.

[19] O.K. Tørresen, B. Star, S. Jentoft, W.B. Reinar, H. Grove, J.R. Miller, B.P. Walenz, J. Knight, J.M. Ekholm, P. Peluso, R.B. Edvardsen, A. Tooming-Klunderud, M. Skage, S. Lien, K.S. Jakobsen, A.J. Nederbragt, An improved genome assembly uncovers prolific tandem repeats in Atlantic cod, BMC Genom. 18 (2017) 1–23, https://doi.org/10.1186/S12864-016-3448-X/TABLES/1.

[20] W. Li, K. Li, Q. jie Zhang, T. Zhu, Y. Zhang, C. Shi, Y. long Liu, E. hua Xia, J. jun Jiang, C. Shi, L. ping Zhang, H. Huang, Y. Tong, Y. Liu, D. Zhang, Y. Zhao, W. kai Jiang, Y. jie Zhao, S. yan Mao, J. ying Jiao, P. zhen Xu, L. li Yang, G. ying Yin, L. zhi Gao, Improved hybrid de novo genome assembly and annotation of African wild rice, Oryza longistaminata, from Illumina and PacBio sequencing reads, Plant Genome 13 (2020), e20001, https://doi.org/10.1002/TPG2.20001.

[21] C. Holt, M. Campbell, D.A. Keays, N. Edelman, A. Kapusta, E. Maclary, E.T. Domyan, A. Suh, W.C. Warren, M. Yandell, M. Thomas, M.D. Shapiro, Improved genome assembly and annotation for the rock pigeon (Columba livia), G3 genes, Genomes, Genet. 8 (2018) 1391–1398, https://doi.org/10.1534/G3.117.300443/-/DC1.

[22] N.I. Weisenfeld, V. Kumar, P. Shah, D.M. Church, D.B. Jaffe, Direct determination of diploid genome sequences, Genome Res. (2017), https://doi.org/10.1101/gr.214874.116.

[23] R.K. Patel, M. Jain, N.G.S.Q.C. Toolkit, A toolkit for quality control of next generation sequencing data, PLoS One 7 (2012), e30619, https://doi.org/10.1371/JOURNAL.PONE.0030619.

[24] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prjibelski, A.V. Pyshkin, A.V. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, P.A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, J. Comput. Biol. (2012), https://doi.org/10.1089/cmb.2012.0021.

[25] M. Kolmogorov, J. Yuan, Y. Lin, P.A. Pevzner, Assembly of long, error-prone reads using repeat graphs, Nat. Biotechnol. (2019), https://doi.org/10.1038/s41587-019-0072-8.

[26] B.J. Walker, T. Abeel, T. Shea, M. Priest, A. Abouelliel, S. Sakthikumar, C.A. Cuomo, Q. Zeng, J. Wortman, S.K. Young, A.M. Earl, Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement, PLoS One (2014), https://doi.org/10.1371/journal.pone.0112963.

[27] R. Kajitani, D. Yoshimura, M. Okuno, Y. Minakuchi, H. Kagoshima, A. Fujiyama, K. Kubokawa, Y. Kohara, A. Toyoda, T. Itoh, Platanus-allee is a de novo haplotype assembler enabling a comprehensive access to divergent heterozygous regions, Nat. Commun. (2019), https://doi.org/10.1038/s41467-019-09575-2.

[28] P.W. Harrison, A.E. Wright, F. Zimmer, R. Dean, S.H. Montgomery, M.A. Pointer, J.E. Mank, Sexual selection drives evolution and rapid turnover of male gene expression, Proc. Natl. Acad. Sci. U. S. A (2015), https://doi.org/10.1073/pnas.1501339112.

[29] S.V. Zhang, L. Zhuo, M.W. Hahn, AGOUTI: improving genome assembly and annotation using transcriptome data, GigaScience (2016), https://doi.org/10.1186/s13742-016-0136-3.

[30] S. Yeo, L. Coombe, R.L. Warren, J. Chu, I. Birol, ARCS: scaffolding genome drafts with linked reads, Bioinformatics (2018), https://doi.org/10.1093/bioinformatics/btx675.

[31] M. Chakraborty, J.G. Baldwin-Brown, A.D. Long, J.J. Emerson, Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage, Nucleic Acids Res. (2016), https://doi.org/10.1093/nar/gkw654.

[32] D. Paulino, R.L. Warren, B.P. Vandervalk, A. Raymond, S.D. Jackman, I. Birol, Sealer, A scalable gap-closing application for finishing draft genomes, BMC Bioinf. (2015), https://doi.org/10.1186/s12859-015-0663-4.

[33] G.C. Xu, T.J. Xu, R. Zhu, Y. Zhang, S.Q. Li, H.W. Wang, J.T. Li, Lr-Gapcloser, A Tiling Path-Based Gap Closer that Uses Long Reads to Complete Genome Assembly, Gigascience, 2018, https://doi.org/10.1093/gigascience/giy157.

[34] P. Mittal, S.K. Jaiswal, N. Vijay, R. Saxena, V.K. Sharma, Comparative analysis of corrected tiger genome provides clues to its neuronal evolution, Sci. Rep. (2019), https://doi.org/10.1038/s41598-019-54838-z.

[35] L.P. Pryszcz, T. Gabaldón, Redundans: an assembly pipeline for highly heterozygous genomes, Nucleic Acids Res. 44 (2016) e113, https://doi.org/10.1093/nar/gkw294.

[36] K. Okonechnikov, A. Conesa, F. García-Alcalde, Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data, Bioinformatics (2016), https://doi.org/10.1093/bioinformatics/btv566.

[37] F.A. Simão, R.M. Waterhouse, P. Ioannidis, E.V. Kriventseva, E.M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs, Bioinformatics (2015), https://doi.org/10.1093/bioinformatics/btv351.

[38] H. Li, Aligning sequence reads, clone sequences and assembly contigs with, BWA-MEM 00 (2013) 1–3.

[39] G.W. Vurture, F.J. Sedlazeck, M. Nattestad, C.J. Underwood, H. Fang, J. Gurtowski, M.C. Schatz, GenomeScope: fast reference-free genome profiling from short reads, Bioinformatics (2017), https://doi.org/10.1093/bioinformatics/btx153.

[40] G. Marçais, C. Kingsford, A fast, lock-free approach for efficient parallel counting of occurrences of k-mers, Bioinformatics (2011), https://doi.org/10.1093/bioinformatics/btr011.

[41] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools, Bioinformatics (2009), https://doi.org/10.1093/bioinformatics/btp352.

[42] V. Narasimhan, P. Danecek, A. Scally, Y. Xue, C. Tyler-Smith, R. Durbin, BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data, Bioinformatics (2016), https://doi.org/10.1093/bioinformatics/btw044.

[43] M.G. Grabherr, P. Russell, M. Meyer, E. Mauceli, J. Alföldi, F. di Palma, K. Lindblad-Toh, Genome-wide synteny through highly sensitive sequence alignment: Satsuma, Bioinformatics 26 (2010) 1145–1151, https://doi.org/10.1093/BIOINFORMATICS/BTQ102.

[44] A. Chakraborty, M.S. Bisht, R. Saxena, S. Mahajan, J. Pulikkan, V.K. Sharma, Genome sequencing and de novo and reference-based genome assemblies of Bos indicus breeds, Genes Genomics 1 (2023) 1–10, https://doi.org/10.1007/S13258-023-01401-W, 2023.

[45] J.M. Flynn, R. Hubley, C. Goubert, J. Rosen, A.G. Clark, C. Feschotte, A.F. Smit, RepeatModeler2 for automated genomic discovery of transposable element families, Proc. Natl. Acad. Sci. U. S. A (2020), https://doi.org/10.1073/pnas.1921046117.

[46] W. Bao, K.K. Kojima, O. Kohany, Repbase Update, a Database of Repetitive Elements in Eukaryotic Genomes, Mob. DNA, 2015, https://doi.org/10.1186/s13100-015-0041-9.

[47] M.S. Campbell, C. Holt, B. Moore, M. Yandell, Genome annotation and curation using MAKER and MAKER-P, Curr. Protoc. Bioinfor. (2014), https://doi.org/10.1002/0471250953.bi0411s48.

[48] B.J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P.D. Blood, J. Bowden, M.B. Couger, D. Eccles, B. Li, M. Lieber, M.D. Macmanes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C.N. Dewey, R. Henschel, R.D. Leduc, N. Friedman, A. Regev, De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis, Nat. Protoc. (2013), https://doi.org/10.1038/nprot.2013.084.

[49] T. Hubbard, D. Barker, E. Birney, G. Cameron, Y. Chen, L. Clark, T. Cox, J. Cuff, V. Curwen, T. Down, R. Durbin, E. Eyras, J. Gilbert, J. Hammond, L. Huminiecki, A. Kasprzyk, H. Lehvaslaiho, P. Lijnzaad, C. Melsopp, E. Mongin, R. Pettett, M. Pocock, S. Potter, A. Rust, E. Schmidt, S. Searle, G. Slater, J. Smith, W. Spooner, A. Stabenau, J. Stalker, E. Stupka, A. Ureta-Vidal, I. Vastrik, M. Clamp, The Ensembl genome database project, Nucleic Acids Res. (2002), https://doi.org/10.1093/nar/30.1.38.

[50] I. Korf, Gene finding in novel genomes, BMC Bioinf. (2004), https://doi.org/10.1186/1471-2105-5-59.

[51] M. Stanke, O. Keller, I. Gunduz, A. Hayes, S. Waack, B. Morgenstern, AUGUSTUS: a b initio prediction of alternative transcripts, Nucleic Acids Res. (2006), https://doi.org/10.1093/nar/gkl200.

[52] P.P. Chan, T.M. Lowe, tRNAscan-Se, Searching for tRNA genes in genomic sequences, Methods Mol. Biol. (2019), https://doi.org/10.1007/978-1-4939-9173-0_1.

[53] B. Fromm, E. Høye, D. Domanska, X. Zhong, E. Aparicio-Puerta, V. Ovchinnikov, S.U. Umu, P.J. Chabot, W. Kang, M. Aslanzadeh, M. Tarbier, E. Mármol-Sánchez, G. Urgese, E. Johansen, E. Hovig, M. Hackenberg, M.R. Friedländer, K.J. Peterson, MirGeneDB 2.1: toward a complete sampling of all major animal phyla, Nucleic Acids Res. (2022), https://doi.org/10.1093/nar/gkab1101.

[54] Y. Wang, H. Tang, J.D. Debarry, X. Tan, J. Li, X. Wang, T.H. Lee, H. Jin, B. Marler, H. Guo, J.C. Kissinger, A.H. Paterson, MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity, Nucleic Acids Res. (2012), https://doi.org/10.1093/nar/gkr1293.

[55] M. Pandey, B. Kushwaha, R. Kumar, P. Srivastava, S. Saroj, M. Singh, Evol2circos: a web-based tool for genome synteny and collinearity analysis and its visualization in fishes, J. Hered. (2020), https://doi.org/10.1093/jhered/esaa025.

[56] D.M. Emms, S. Kelly, OrthoFinder: phylogenetic orthology inference for comparative genomics, Genome Biol. (2019), https://doi.org/10.1186/s13059-019-1832-y.

[57] D.R. Laetsch, M.L. Blaxter, KinFin: software for taxon-aware analysis of clustered protein sequences, G3 Genes, Genomes, Genet. https://doi.org/10.1534/g3.117.300233, 2017.

[58] K. Katoh, D.M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability, Mol. Biol. Evol. (2013), https://doi.org/10.1093/molbev/mst010.

[59] A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies, Bioinformatics (2014), https://doi.org/10.1093/bioinformatics/btu033.

[60] P.C. Ng, S. Henikoff, SIFT: predicting amino acid changes that affect protein function, Nucleic Acids Res. (2003), https://doi.org/10.1093/nar/gkg509.

[61] T. Jombart, S. Dray, Adephylo: exploratory analyses for the phylogenetic comparative method, Bioinformatics (2010), https://doi.org/10.1093/bioinformatics/btq292.

[62] Z. Yang, Paml 4: phylogenetic analysis by maximum likelihood, Mol. Biol. Evol. (2007), https://doi.org/10.1093/molbev/msm088.

[63] A. Chakraborty, S. Mahajan, S.K. Jaiswal, V.K. Sharma, Genome sequencing of turmeric provides evolutionary insights into its medicinal properties, Commun. Biol. 41 (4) (2021) 1–12, https://doi.org/10.1038/s42003-021-02720-y, 2021.

[64] A. Chakraborty, S. Mahajan, M.S. Bisht, V.K. Sharma, Genome sequencing and comparative analysis of Ficus benghalensis and Ficus religiosa species reveal evolutionary mechanisms of longevity, iScience 25 (2022), https://doi.org/10.1016/J.ISCI.2022.105100.

[65] T. De Bie, N. Cristianini, J.P. Demuth, M.W. Hahn, CAFE: a computational tool for the study of gene family evolution, Bioinformatics (2006), https://doi.org/10.1093/bioinformatics/btl097.

[66] S. Van Dongen, C. Abreu-Goodger, Using MCL to extract clusters from networks, Methods Mol. Biol. (2012), https://doi.org/10.1007/978-1-61779-361-5_15.

[67] B. Buchfink, K. Reuter, H.G. Drost, Sensitive protein alignments at tree-of-life scale using DIAMOND, 2021, Nat. Methods 184 (18) (2021) 366–368, https://doi.org/10.1038/s41592-021-01101-x.

[68] A. Bateman, The Pfam protein families database, Nucleic Acids Res. (2004), https://doi.org/10.1093/nar/gkh121.

[69] A. Bairoch, R. Apweiler, The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000, Nucleic Acids Res. (2000), https://doi.org/10.1093/nar/28.1.45.

[70] P. Jones, D. Binns, H.Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A.F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S.Y. Yong, R. Lopez, S. Hunter, InterProScan 5: genome-scale protein function classification, Bioinformatics (2014), https://doi.org/10.1093/bioinformatics/btu031.

[71] Y. Moriya, M. Itoh, S. Okuda, A.C. Yoshizawa, M. Kanehisa, Kaas, An automatic genome annotation and pathway reconstruction server, Nucleic Acids Res. (2007), https://doi.org/10.1093/nar/gkm321.

[72] J. Huerta-Cepas, K. Forslund, L.P. Coelho, D. Szklarczyk, L.J. Jensen, C. Von Mering, P. Bork, Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper, Mol. Biol. Evol. (2017), https://doi.org/10.1093/molbev/msx148.

[73] Y. Wang, D. Coleman-Derr, G. Chen, Y.Q. Gu, OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species, Nucleic Acids Res. (2015), https://doi.org/10.1093/nar/gkv487.

[74] S. Kumar, M. Suleski, J.M. Craig, A.E. Kasprowicz, M. Sanderford, M. Li, G. Stecher, S.B. Hedges, TimeTree 5: an expanded resource for species divergence times, Mol. Biol. Evol. 39 (2022) 1–6, https://doi.org/10.1093/molbev/msac174.

[75] T. Cao, J.P. Jin, Evolution of flight muscle contractility and energetic efficiency, Front. Physiol. (2020), https://doi.org/10.3389/fphys.2020.01038.

[76] S. Ricard-Blum, The collagen family, Cold Spring Harbor Perspect. Biol. (2011), https://doi.org/10.1101/cshperspect.a004978.

[77] N.P. Boyer, S.L. Gupton, Revisiting netrin-1: one who guides (Axons), Front. Cell. Neurosci. (2018), https://doi.org/10.3389/fncel.2018.00221.

[78] M. Fu, P.J. Blackshear, RNA-binding proteins in immune regulation: a focus on CCCH zinc finger proteins, Nat. Rev. Immunol. (2017), https://doi.org/10.1038/nri.2016.129.

[79] W.L. Gao, S.Q. Zhang, H. Zhang, B. Wan, Z.S. Yin, Chordin-like protein 1 promotes neuronal differentiation by inhibiting bone morphogenetic protein-4 in neural stem cells, Mol. Med. Rep. (2013), https://doi.org/10.3892/mmr.2013.1310.

[80] E. Matsunaga, S. Nambu, M. Oka, M. Tanaka, M. Taoka, A. Iriki, Periostin, a neurite outgrowth-promoting factor, is expressed at high levels in the primate cerebral cortex, Dev. Growth Differ. (2015), https://doi.org/10.1111/dgd.12194.

[81] F. Huang, X. Wang, E.M. Ostertag, T. Nuwal, B. Huang, Y.N. Jan, A.I. Basbaum, L.Y. Jan, TMEM16C facilitates Na + -activated K + currents in rat sensory neurons and regulates pain processing, Nat. Neurosci. (2013), https://doi.org/10.1038/nn.3468.

[82] N.N. Nalivaeva, N.D. Belyaev, I.A. Zhuravin, A.J. Turner, The Alzheimers amyloid-degrading peptidase, neprilysin: can we control it? Int. J. Alzheimer's Dis. 2012 (2012) https://doi.org/10.1155/2012/383796.

[83] R. Pérez-Sen, M.J. Queipo, J.C. Gil-Redondo, F. Ortega, R. Gómez-Villafuertes, M.T. Miras-Portugal, E.G. Delicado, Dual-specificity phosphatase regulation in neurons and glial cells, Int. J. Mol. Sci. (2019), https://doi.org/10.3390/ijms20081999.

[84] M.A. Gutierrez, B.E. Dwyer, S.J. Franco, Csmd2 Is a Synaptic Transmembrane Protein that Interacts with Psd-95 and Is Required for Neuronal Maturation, ENeuro, 2019, https://doi.org/10.1523/ENEURO.0434-18.2019.

[85] R.H. Adams, H. Betz, A.W. Püschel, A novel class of murine semaphorins with homology to thrombospondin is differentially expressed during early embryogenesis, Mech. Dev. (1996), https://doi.org/10.1016/0925-4773(96)00525-4.

[86] A. Mogha, A.E. Benesh, C. Patra, F.B. Engel, T. Schöneberg, I. Liebscher, K.R. Monk, Gpr126 functions in schwann cells to control differentiation and myelination via G-protein activation, J. Neurosci. (2013), https://doi.org/10.1523/JNEUROSCI.1809-13.2013.

[87] J. Huot, Ephrin signaling in axon guidance, Prog. Neuro Psychopharmacol. Biol. Psychiatr. (2004), https://doi.org/10.1016/j.pnpbp.2004.05.025.

[88] L. Atapattu, M. Lackmann, P.W. Janes, The role of proteases in regulating Eph/ephrin signaling, Cell Adhes. Migrat. (2014), https://doi.org/10.4161/19336918.2014.970026.

[89] N.P. Boyer, L.E. McCormick, S. Menon, F.L. Urbina, S.L. Gupton, A pair of E3 ubiquitin ligases compete to regulate filopodial dynamics and axon guidance, J. Cell Biol. (2020), https://doi.org/10.1083/jcb.201902088.

[90] S.S. Steiger, A.E. Fidler, J.C. Mueller, B. Kempenaers, Evidence for adaptive evolution of olfactory receptor genes in 9 bird species, J. Hered. (2010), https://doi.org/10.1093/jhered/esp105.

[91] S. Prag, J.C. Adams, Molecular phylogeny of the kelch-repeat superfamily reveals an expansion of BTB/kelch proteins in animals, BMC Bioinf. (2003), https://doi.org/10.1186/1471-2105-4-42.

[92] O.T.T. Le, T.T.N. Nguyen, S.Y. Lee, Phosphoinositide turnover in Toll-like receptor signaling and trafficking, BMB Rep. (2014), https://doi.org/10.5483/BMBRep.2014.47.7.088.

[93] A. Rodríguez-Gil, O. Ritter, V.V. Saul, J. Wilhelm, C.Y. Yang, R. Grosschedl, Y. Imai, K. Kuba, M. Kracht, M. Lienhard Schmitz, The CCR4-NOT complex contributes to repression of Major Histocompatibility Complex class II transcription, Sci. Rep. (2017), https://doi.org/10.1038/s41598-017-03708-7.

[94] E.J. Yang, J.H. Yoon, K.C. Chung, Bruton's tyrosine kinase phosphorylates cAMP-responsive element-binding protein at serine 133 during neuronal differentiation in immortalized hippocampal progenitor cells, J. Biol. Chem. 279 (2004) 1827–1837, https://doi.org/10.1074/jbc.M308722200.

[95] J.W. Griffith, C.L. Sokol, A.D. Luster, Chemokines and chemokine receptors: positioning cells for host defense and immunity, Annu. Rev. Immunol. (2014), https://doi.org/10.1146/annurev-immunol-032713-120145.

[96] M. Bermejo-Jambrina, J. Eder, L.C. Helgers, N. Hertoghs, B.M. Nijmeijer, M. Stunnenberg, T.B.H. Geijtenbeek, C-type lectin receptors in antiviral immunity and viral escape, Front. Immunol. (2018), https://doi.org/10.3389/fimmu.2018.00590.

[97] N. Nik Tavakoli, B.D. Hambly, D.R. Sullivan, S. Bao, Forkhead box protein 3: essential immune regulatory role, Int. J. Biochem. Cell Biol. (2008), https://doi.org/10.1016/j.biocel.2007.10.004.

[98] A. El-Amraoui, J.S. Schonn, P. Küssel-Andermann, S. Blanchard, C. Desnos, J.P. Henry, U. Wolfrum, F. Darchen, C. Petit, MyRIP, a novel Rab effector, enables myosin VIIa recruitment to retinal melanosomes, EMBO Rep. (2002), https://doi.org/10.1093/embo-reports/kvf090.

[99] X.Y. Wen, M. Tarailo-Graovac, K. Brand-Arzamendi, A. Willems, B. Rakic, K. Huijben, A. Da Silva, X. Pan, S. El-Rass, R. Ng, K. Selby, A.M. Philip, J. Yun, X. C. Ye, C.J. Ross, A.M. Lehman, F. Zijlstra, N. Abu Bakar, B. Drögemöller, J. Moreland, W.W. Wasserman, H. Vallance, M. van Scherpenzeel, F. Karbassi, M. Hoskings, U. Engelke, A. de Brouwer, R.A. Wevers, A.V. Pshezhetsky, C.D. van Karnebeek, D.J. Lefeber, Sialic acid catabolism by N-acetylneuraminate pyruvate lyase is essential for muscle function, JCI Insight (2018), https://doi.org/10.1172/jci.insight.122373.

[100] S.M. Halmo, D. Singh, S. Patel, S. Wang, M. Edlin, G.J. Boons, K.W. Moremen, D. Live, L. Wells, Protein O-linked mannose β-1,4-N-acetylglucosaminyltransferase 2 (POMGNT2) is a gatekeeper enzyme for functional glycosylation of α-dystroglycan, J. Biol. Chem. 292 (2017) 2101–2109, https://doi.org/10.1074/jbc.M116.764712.

[101] C. Hayashi, Y. Ono, N. Doi, F. Kitamura, M. Tagami, R. Mineki, T. Arai, H. Taguchi, M. Yanagida, S. Hirner, D. Labeit, S. Labeit, H. Sorimachi, Multiple molecular interactions implicate the connectin/titin N2A region as a modulating scaffold for p94/calpain 3 activity in skeletal muscle, J. Biol. Chem. (2008), https://doi.org/10.1074/jbc.M708262200.

[102] C. Nie, L. Qu, X. Li, Z. Jiang, K. Wang, H. Li, H. Wang, C. Qu, L. Qu, Z. Ning, Genomic regions related to white/black tail feather color in dwarf chickens identified using a genome-wide association study, Front. Genet. (2021), https://doi.org/10.3389/fgene.2021.566047.

[103] N.J. Nadeau, T. Burke, N.I. Mundy, Evolution of an avian pigmentation gene correlates with a measure of sexual selection, Proc. R. Soc. B Biol. Sci. (2007), https://doi.org/10.1098/rspb.2007.0174.

[104] E. Jarvis, O. Güntürkün, L. Bruce, A. Csillag, H. Karten, W. Kuenzel, L. Medina, G. Paxinos, D.J. Perkel, T. Shimizu, G. Striedter, J. Martin Wild, G.F. Ball, J. Dugas-Ford, S.E. Durand, G.E. Hough, S. Husband, L. Kubikova, D.W. Lee, C.V. Mello, A. Powers, C. Siang, T.V. Smulders, K. Wada, S.A. White, K. Yamamoto, J. Yu, A. Reiner, A.B. Butler, Avian brains and a new understanding of vertebrate brain evolution, Nat. Rev. Neurosci. (2005), https://doi.org/10.1038/nrn1606.

[105] B. Ranscht, Cadherins: molecular codes for axon guidance and synapse formation, Int. J. Dev. Neurosci. (2000), https://doi.org/10.1016/S0736-5748(00)00030-7.

[106] Y. Zou, Wnt signaling in axon guidance, Trends Neurosci. (2004), https://doi.org/10.1016/j.tins.2004.06.015.

[107] J.J. Merkin, P. Chen, M.S. Alexis, S.K. Hautaniemi, C.B. Burge, Origins and impacts of new mammalian exons, Cell Rep. (2015), https://doi.org/10.1016/j.celrep.2015.02.058.

[108] A. Fiszbein, K.S. Krick, B.E. Begg, C.B. Burge, Exon-mediated activation of transcription starts, Cell (2019), https://doi.org/10.1016/j.cell.2019.11.002.

[109] A. Schmidt-Küntzel, D.L. Dalton, M. Menotti-Raymond, E. Fabiano, P. Charruau, W.E. Johnson, S. Sommer, L. Marker, A. Kotzé, S.J. O'Brien, Conservation genetics of the cheetah: genetic history and implications for conservation, Cheetahs Biol. Conserv. (2018) 71, https://doi.org/10.1016/B978-0-12-804088-1.00006-X.

[110] M. Menotti-Raymond, S.J. O'brient, Dating the genetic bottleneck of the African cheetah (DNA rmgerprint/mtDNA), Proc. Natl. Acad. Sci. USA 90 (1993) 3172–3176.