A novel mitochondrial genome of *Arborophila* and new insight into *Arborophila* evolutionary history

Chaochao Yan[1], Biqin Mou[1], Yang Meng[1], Feiyun Tu[1,2], Zhenxin Fan[3], Megan Price[3], Bisong Yue[3], Xiuyue Zhang[1]*

1 Key Laboratory of Bio-Resources and Eco-Environment (Ministry of Education), College of Life Sciences, Sichuan University, Chengdu, P.R. China, **2** Institute of Wildlife Conservation, Jiangxi Academy of Forestry, Nanchang, P.R. China, **3** Sichuan Key Laboratory of Conservation Biology on Endangered Wildlife, College of Life Sciences, Sichuan University, Chengdu, P.R. China

* zhangxy317@126.com

Abstract

The lineage of the Bar-backed Partridge (*Arborophila brunneopectus*) was investigated to determine the phylogenetic relationships within *Arborophila* as the species is centrally distributed within an area covered by the distributions of 22 South-east Asian hill partridge species. The complete mitochondrial genome (mitogenome) of *A. brunneopectus* was determined and compared with four other hill partridge species mitogenomes. NADH subunit genes are radical in hill partridge mitogenomes and contain the most potential positive selective sites around where variable sites are abundant. Together with 44 other mitogenomes of closely related species, we reconstructed highly resolved phylogenetic trees using maximum likelihood (ML) and Bayesian inference (BI) analyses and calculated the divergence and dispersal history of *Arborophila* using combined datasets composed of their 13-protein coding sequences. *Arborophila* is reportedly be the oldest group in Phasianidae whose ancestors probably originated in Asia. *A. rufipectus* shares a closer relationship with *A. ardens* and *A. brunneopectus* compared to *A. gingica* and *A. rufogularis*, and such relationships were supported and profiled by NADH dehydrogenase subunit 5 (*ND5*). The intra-genus divergence of all five *Arborophila* species occurred in the Miocene (16.84~5.69 Mya) when there were periods of climate cooling. We propose that these cooling events in the Miocene forced hill partridges from higher to lower altitudes, which led to geographic isolation and speciation. We demonstrated that the apparently deleterious +1 frameshift mutation in NADH dehydrogenase subunit 3 (*ND3*) found in all *Arborophila* is an ancient trait that has been eliminated in some younger lineages, such as Passeriformes. It is unclear of the biological advantages of this elimination for the relevant taxa and this requires further investigation.

Introduction

Species of the *Arborophila* (commonly: hill partridge) are found in Asia from the Himalayas eastwards to Taiwan and south to Java. *Arborophila* is a very diverse genus, being the second most species rich genus within Galliformes [1–3]. Hill partridges are fairly small and their body typically ranges between 24 and 30 cm, while their beaks and tails are of equal length [2]. These small partridges are often brightly marked, but this marking provides camouflage in the leaf litter of eastern and southern Asia forests. Hill partridges are plump-bodied with thick necks and moderately long legs, stronger and longer tarsometatarsal, and rounded and rather short wings. Thus hill partridges are poor fliers with low aerial maneuverability.

The *Arborophila* genus contains over 20 species with 10 of those being found in China [2,4]. In 1998, eight of the 10 Chinese hill partridge species were listed in the *China red data book of endangered animals* as 'Rare' or 'Endangered' [5,6], among which the Sichuan Hill Partridge (*A. rufipectus*) and Hainan Partridge (*A. ardens*) are currently ranked as Endangered and Vulnerable, respectively, by IUCN Red List of Threatened Species [6]. Although some of the hill partridge species have a relatively wide distribution and inhabit a variety of habitats, their poor flying ability and ground-dwelling nature makes them vulnerable to human activities, non-indigenous predators and rapid loss of habitat and are thus rare within their respective ranges.

Research regarding the taxonomic status and interspecific relationships of *Arborophila*, is still in its infancy and is currently being debated [7–9]. Recent studies tend to place the *Arborophila* near the root position of Phasianidae [3,10–12], however, the phylogenetic relationship within *Arborophila* has been little studied.

Mitochondrial DNA (mtDNA) has frequently and widely been used as an ideal marker to resolve intractable phylogenies, because it typically accumulates mutations at a faster rate than nuclear DNA [8]. The accumulation of mutations makes it more effective to resolve relationships among recently diverged species or populations. However this rapid accumulation of mutations makes mtDNA incompatible to resolve older relationships because the more variable DNA sequences from remote species can easily become saturation [13]. Therefore we applied careful testing to avoid saturation occurring in our dataset. Another limitation of mtDNA is that it is considered prone to introgression among species and only reflects a one-sided perspective on the ancestry [14]. However, mtDNA has merits such as lack of recombination and is easily accessible. Moreover, the high rates of anatomical evolution match the rapid mitoDNA evolution in birds somehow[4,15–17].

Unlike the phylogenic distribution of *Arborophila*, the geographic distribution is relatively well understood. The widely distributed Bar-backed Partridge (*A. brunneopectus*) ranges across an area that sits centrally to the distributions of 22 extant hill partridge species, and overlaps with three of the 22 species (Fig 1). It is reported that *A. brunneopectus* shares a closer relationship with *A. ardens*, although the number of intermediate species between *A. brunneopectus* and *A. ardens* is unknown. Complete ecological data of these species are lacking [18–20].

To date, only four mitogenomes within the genus *Arborophila* have been determined, while two are incomplete and lackspecific descriptions [21]. In this study, we determined the complete mitochondrial genome of *A. brunneopectus*. We also reconstructed the phylogenetic tree of 45 related birds and estimated contributions of every single mitogene to the consensus tree. Based on the divergence time and biogeographic data of the studied species, a hypothesis model of settlement and migration of hill partridge species was proposed. We also examined and discussed the relationship of variable sites and positive selective sites as well as the +1 frameshift mutation in NADH dehydrogenase subunit 3 (*ND3*) through the mitochondrial
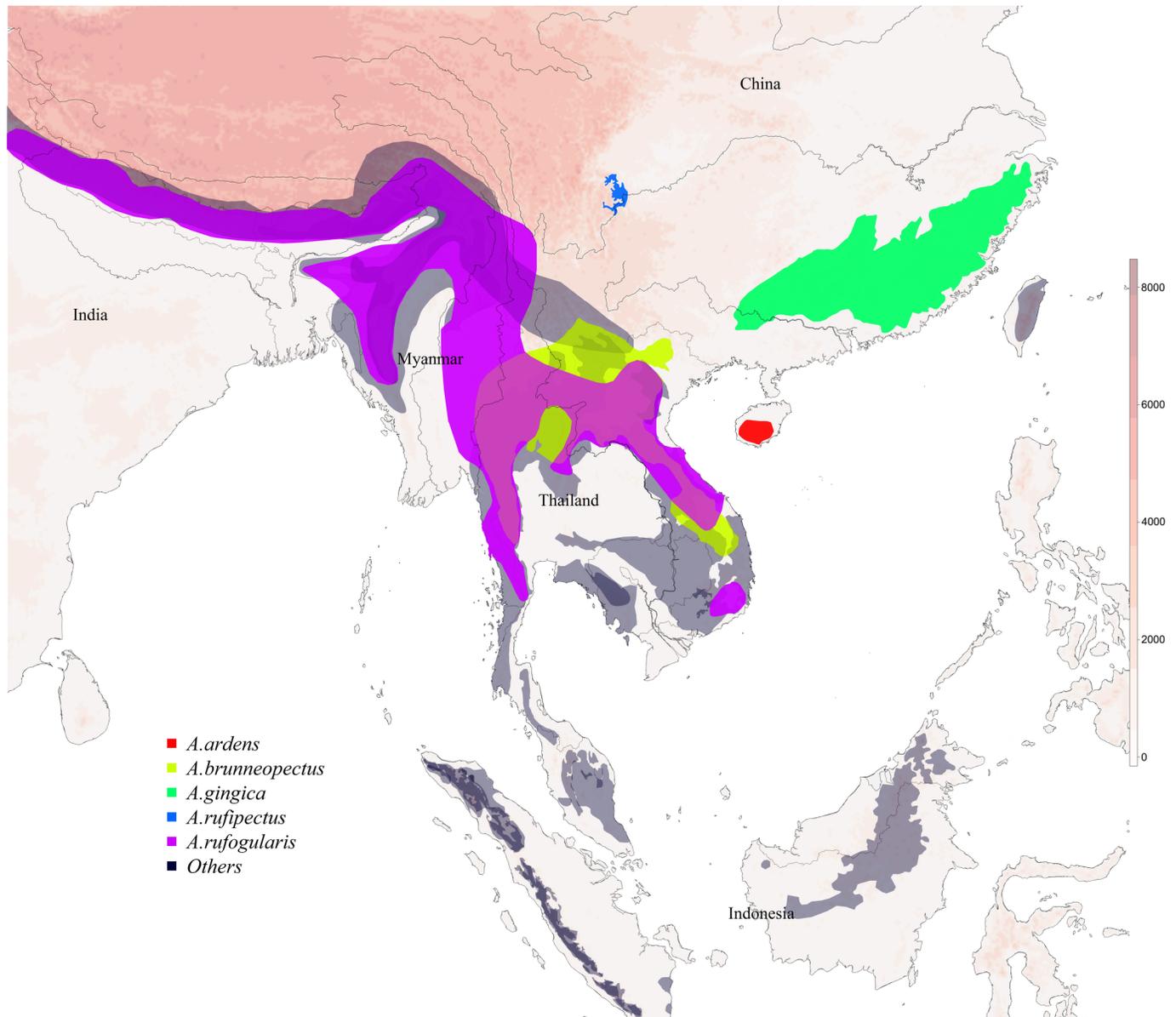
Fig 1. Geographic distribution of the five hill partridges within this study in comparison to the overall distribution of the other 17 hill partridges in South-east Asia. The five studied hill partridge distributions have been assigned different colors while the others 17 are all colored midnight blue. Elevation information (altitude: meters above sea level) is represented by different heat colors.

DNA alignments. The publication of the mitogenome of *A. brunneopectus* will greatly contribute to understanding of relationships within *Arborophila*.

Materials and methods

DNA extraction, amplification and sequencing

The protocol of this research has been reviewed and approved by the Ethics and Experimental Animal Committee of Sichuan Institute of Zoology, Chinese Academy of Science. The Bar-backed Partridge sample was offered by Sichuan Key Laboratory of Conservation Biology on

Endangered Wildlife. Total genomic DNA was extracted from breast muscle tissue using standard phenol/chloroform methods [22]. Target segments were amplified with a long and accurate polymerase chain reaction (LA-PCR) technique according to the manufacturer's instructions (TaKaRa, China). For gaps or low quality terminus between target segments, normal PCR was applied as supplementary to complete the sequencing. Twenty pairs of primers were used to amplify our mtDNA fragments, among which the first eight pairs of primers were used in LA-PCR while the remaining 12 pairs were used in normal PCR. The primers used for both PCR amplification and sequencing are listed in S1 Table. LA-PCR reaction was carried out in a 25μl reaction volume containing 100ng template DNA, 1×LA PCR buffer (TaKaRa), 0.2mM dNTP, 2.5mM MgCl$_2$, 1μM each primer, 1.0U LA Taq polymerase (TaKaRa). After a pre-denaturation at 95˚C for 4min, 35 thermal cycle were carried out with denaturation at 94˚C for 30s, annealing at optimal temperature based on each pair of primers for 45s, and extension at 72˚C for 1-2min, and after the thermal cycle, a final extension was executed at 72˚C for 10min. LA-PCR products were electrophoresed on a 1.0% Agarose Gel and purified with the DNA Agarose Gel Extraction Kit (Omega, USA). The purified fragments were directly sequenced from both strands using primer walking method on the ABI 373/377 capillary sequencer. To fill gaps between fragments, normal PCR was also carried out with corresponding PCR buffer, Taq polymerase in PCR master mix and relatively short extension time according to lengths of their target segments in thermal cycle.

The electropherogram had no double peaks throughout the whole mitogenome and the sequence fragments of genes showed no frameshift mutation stop codons when translated into amino acid except *ND3*. In addition, the total cellular DNA was extracted from muscle tissue where it is particularly rich in mtDNA. For those reasons, the data represents the true mtDNA sequence rather than nuclear mitochondrial pseudogenes.

Sequences analyses

Original sequences were assembled in Lasergene SeqMan Pro (version7.1.0 DNASTAR, Inc., Madison, WI, USA) and optimized manually. To annotate the mitochondrial genome, tRNAscan-SE 1.21 was used to determine positions of 20 tRNAs in organellar (mitochondrial/chloroplast) tRNAs search mode [23]. Remarkably, the tRNA-Trp was annotated as tRNA-SeC (Selenocysteine tRNAs) in this program and the other tRNA-Ser (GCT) was undetected but found through homological alignment and determined by checking its second structure. The locations of protein-coding genes and ribosomal RNA (rRNA) genes were determined by comparison with known homological sequences.

As a typical character of *Arborophila* mitogenomes, the widely emerged extra 174th cytosine insertion in gene *ND3* was carefully investigated in our five mitogenomes, then extended to all birds accessible in Genbank (S3 Table). Subsequently, these sequences were aligned in MEGA5 followed by removing redundant sequences. The distribution of the two types of birds categorized by *ND3* with extra cytosine (WEC) or without extra cytosine (WOC) was determined and further divided according to their taxonomic orders. The relationships of these orders were drawn mainly from references to recent large scale molecular basedstudies [24–26]. It should be noted that, not all sequence authors are aware of the existence of a +1 frameshift mutation in the avian mitogene and may simply remove the insertion, so that some species exhibiting WOC *ND3* may have misled our results [7,27].

Phylogenetic and biogeographic analyses

Our sequenced genome was applied as a query using nucleotide BLAST online to produce a robust phylogenetic tree. We preferentially chose the unique sequence curated by NCBI staff

from the nearly 150 homological results and finally got those 44 showed in Table 1. In these species, all the five main lineages (OW partridges, peafowls, junglefowls, tragopans and gallo-pheasants) of Phasianidae have representative species (Fig 2). The traditional non-pheasant species *Alectura lathami* and *Meleagris gallopavo* were also included in our alignment. However, some genus *Argusianus* were missing and *Polyplectron* were sparse because of there are no complete mitogenomes to date. *Anseranas semipalmata* and *Anas platyrhynchos* were set as outgroups as they were neither too far nor too close to the ingroup in evolution. Moreover, they were proved powerful in constructed robust ingroup relationships in our study. To confirm whether trees inferred from mitoDNA differed from nuclear DNA trees, we collected 35 of the 45 species who shared six nuclear intron sequences amplified in Wang's study [28].

After eliminating terminators, the 13 protein-coding genes sequences of each species were concatenated using custom made Perl scripts and were aligned in MAGE 5.2 [29]. The NADH dehydrogenase subunit 6 (*ND6*) gene was included. Although it is encoded by the light-strand that has a significantly different base composition compared to heavy-chain [30], it still fits the best substitution model of all and can develop Bayesian posterior probabilities (BBP) and bootstrap proportions (BSP) in both Bayesian inference (BI) and maximum likelihood (ML) phylogenetic trees. The alignment was checked by comparing with several sequences aligned and concatenated manually. All sites with alignment gaps or missing/ambiguous data were excluded making all sequences a uniform 11313bp.

In order to investigate the possibility of a bias owing to substitution saturation, we estimated the substitution saturation of alignments with Xia's [31] method and plotted transitions and transversions against pairwise sequence divergence using 13 protein-coding mito-genes with DAMBE 5.3.10 [32]. Codon positions 1 and 2 were not saturated. Codon position 3 was saturated only when adding outgroups. The third codon position was not as constrained as the other two positions according to the Wobble hypothesis when tRNA base pairing. Furthermore, the results illustrated that our alignments had little substitute saturate (ISS > ISS.C, p<0.001) (S1 Fig), and are suitable for phylogenetic analyses.

jModelTest 2.1.1 [33] was used in our combined datasets to select the preferred models of evolution for Bayesian inference (BI) under the Akaike Information Criterion [34,35]. The GTR+I+G model was chosen as the best model for the concatenated datasets. Two partitioned schemes, three codon position to be separated or 13 protein-coding genes to be divided into 3 groups: 1) ATP synthase F0 subunit 6 (*ATP6*), ATP synthase F0 subunit 8 (*ATP8*), cytochrome b (*CYTB*), NADH dehydrogenase subunit 1 (*ND1*), NADH dehydrogenase subunit 2 (*ND2*), *ND3*, NADH dehydrogenase subunit 4 (*ND4*), NADH dehydrogenase subunit 4L (*ND4L*), NADH dehydrogenase subunit 5 (*ND5*); 2) cytochrome c oxidase subunit 1 (*COX1*), cytochrome c oxidase subunit 2 (*COX2*), cytochrome c oxidase subunit 2 (*COX3*); 3) *ND6* suggested in PartitionFinder v1.1.1 [36] were also applied to Partitioned BI analysis respectively with the same alignments to strengthen the results.

Bayesian inference (BI) was performed in MrBayes v3.2 [37,38]. Two separate runs were performed with four Markov chains, each run for 10, 000, 000 generations and sampled every 1000 generations, the consensus tree was calculated after omitting the first 25% trees as burn-in. Maximum likelihood was analyzed in PhyML v3.0 [39]. Basic parameters were estimated with ML methods. A substitution model was also set as GTR from jModeltest result; tree topology search was set as BEST (best of NNI and SPR search). Four relative substitution rate categories were run for 1,000 bootstrap replicates todetermine best tree. We also performed RAxML method on the website http://phylobench.vital-it.ch/raxml-bb/ with GTR+GAMMA model [40] to reconstruct ML phylogenetic trees with default parameters, which yielded a consensus tree with PhyML and BI (Fig 2). The same phylogenetic protocol was applied to construct combined nuclear DNA tree as well as combined mito-nuclear DNA tree with the six

Table 1. Taxa and sequence Genbank accession numbers of the 45 species used for phylogenetic analysis.

Family/subfamily	Species and common name	Genbank accession NO.	Tranditional taxonomy
Anatidae	*Anas platyrhynchos*	NC_009684	Outgroup
Megapodiidae	*Alectura lathami*	NC_007227	Megapodiidae
Meleagridinae	*Meleagris gallopavo*	NC_010195	Meleagrididae
Perdicinae	*Alectoris chukar*	FJ752426	Old World quails
Perdicinae	*Arborophila gingica*	FJ752425	Old World quails
Perdicinae	*Arborophila rufipectus*	NC_012453	Old World quails
Perdicinae	*Arborophila rufogularis*	FJ752424	Old World quails
Perdicinae	*Arborophila ardens*	NC_022683.1	Old World quails
Perdicinae	*Arborophila brunneopectus*	NC_022684.1*	Old World quails
Perdicinae	*Bambusicola fytchii*	FJ752423	Old World quails
Perdicinae	*Bambusicola thoracica*	NC_011816	Old World quails
Perdicinae	*Coturnix chinensis*	NC_004575	Old World quails
Perdicinae	*Coturnix japonica*	NC_003408	Old World quails
Perdicinae	*Francolinus pintadeanus*	NC_011817	Old World quails
Perdicinae	*Perdix dauurica*	FJ752431	Old World quails
Perdicinae	*Tetraophasis obscurus*	NC_018034	Tragopans
Perdicinae	*Tetraophasis szechenyii*	FJ799728	Tragopans
Phasianinae	*Chrysolophus amherstiae*	FJ752434	Gallopheasants
Phasianinae	*Chrysolophus pictus*	NC_014576	Gallopheasants
Phasianinae	*Crossoptilon auritum*	NC_015897	Gallopheasants
Phasianinae	*Crossoptilon crossoptilon*	NC_016679	Gallopheasants
Phasianinae	*Gallus gallus*	NC_001323	Junglefowls
Phasianinae	*Gallus lafayetii*	NC_007239	Junglefowls
Phasianinae	*Gallus sonneratii*	NC_007240	Junglefowls
Phasianinae	*Gallus varius*	NC_007238	Junglefowls
Phasianinae	*Ithaginis cruentus*	NC_018033	Junglefowls
Phasianinae	*Lophophorus lhuysii*	NC_013979	Tragopans
Phasianinae	*Lophophorus sclateri*	FJ752432	Tragopans
Phasianinae	*Lophura ignita*	NC_010781	Gallopheasants
Phasianinae	*Lophura nycthemera*	NC_012895	Gallopheasants
Phasianinae	*Phasianus colchicus*	NC_015526	Gallopheasants
Phasianinae	*Phasianus versicolor*	NC_010778	Gallopheasants
Phasianinae	*Pavo muticus*	NC_012897	Peafowls
Phasianinae	*Polyplectron bicalcaratum*	NC_012900	Peafowls
Phasianinae	*Pucrasia macrolopha*	FJ752429	Tragopans
Phasianinae	*Syrmaticus ellioti*	NC_010771	Gallopheasants
Phasianinae	*Syrmaticus humiae*	NC_010774	Gallopheasants
Phasianinae	*Syrmaticus reevesii*	NC_010770	Gallopheasants
Phasianinae	*Syrmaticus soemmerringi*	NC_010767	Gallopheasants
Phasianinae	*Tragopan caboti*	NC_013619	Tragopans
Phasianinae	*Tragopan temminckii*	FJ752427	Tragopans
Tetraoninae	*Bonasa bonasia*	FJ752435	Tetraoninae
Numididae	*Acryllium vulturinum*	NC_014180	Numididae
Anseranatidae	*Anseranas semipalmata*	NC_005933	outgroup
Numididae	*Numida meleagris*	NC_006382	Numididae

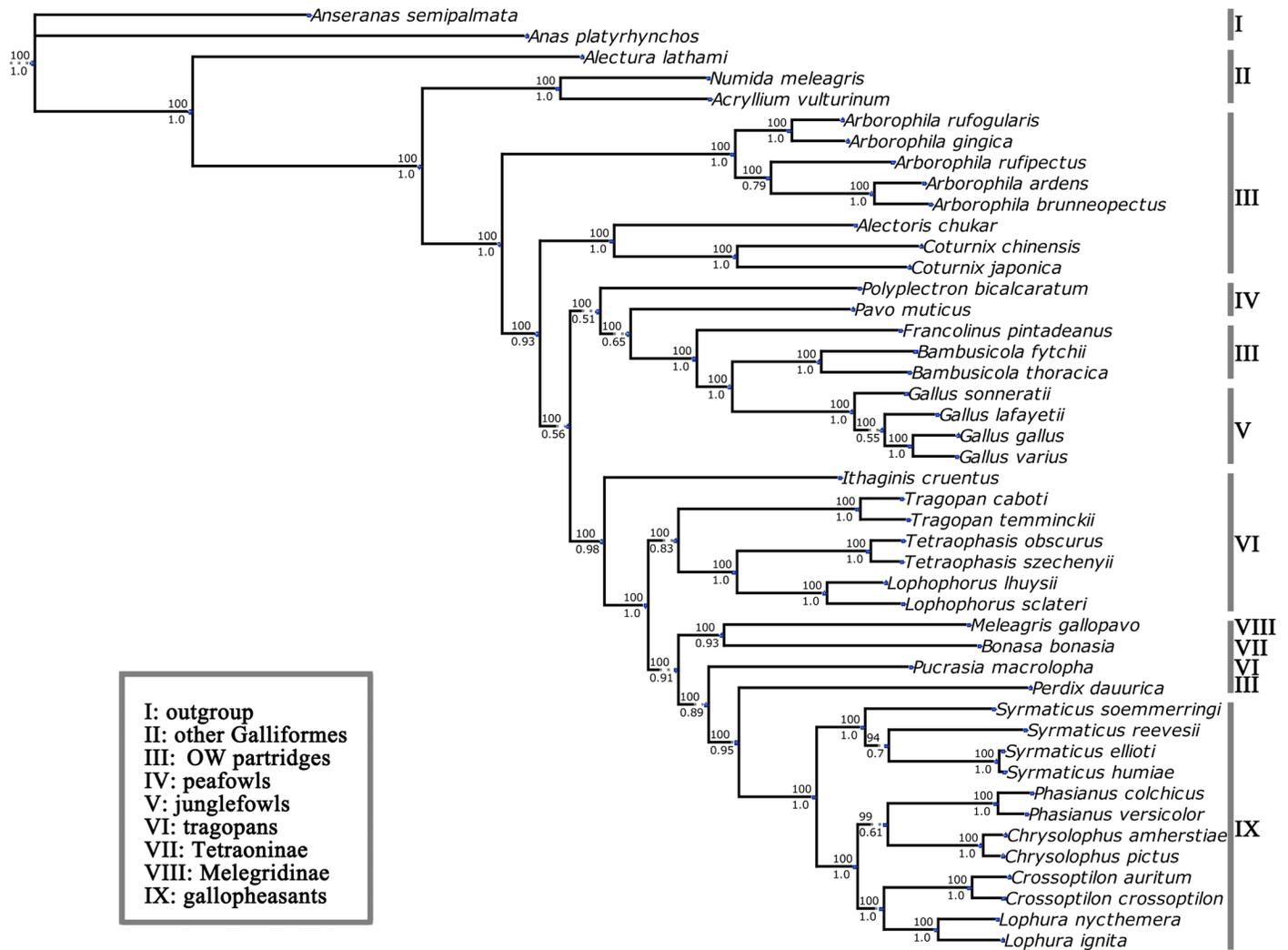*the sequence determined in this study

Fig 2. Molecular phylogenetic tree based on 13 mitochondrial protein-coding genes using Bayesian inference (BI) and maximum likelihood (ML) analysis. The numbers beside the nodes are Bayesian posterior probabilities and ML bootstrap proportions. *Anseranas semipalmata* and *Anas platyrhynchos* were set as outgroups.

nuclear genes. To test the contributions of a single mitogene to the consensus tree, the partitioned Bremer support (PBS) value of every mitogene was calculated in TreeRot V.3 [41]. Every mitogene and gene jackknifing (excluding one gene each run) was also analyzed in BI and the remaining data was used to reconstruct tree. Pairwise Robinson-Foulds distances among trees were calculated with R package phytools [42].

The consensus phylogenetic tree was also used in PAML [43] to detect positive selective loci in mitogenomes of *Arborophila* lineage using branch-sites model A test (model = 2, NSsites = 2) with the clade of hill partridge species set as foreground while the others species were in the background. We run the null hypothesis (fix_momega = 1, omega = 1) and alternative hypothesis (fix_omega = 0, omega = 2) respectively and test the likelihood value (lnL) with chi-square. Bayes Empirical Bayes (BEB) analyses were applied to test positive sites for foreground lineages [44].

Fossils are essential as accurate molecular calibrations for constructing modern bird evolutionary timescale [45]. To calculate the divergence time of *Arborophila,* five compatible fossil calibrations (C1~C5 in Fig 3) were chosen to infer Absolute time of each differentiation event: 1) Fowl and ducks diverged at 89.8±7 million years ago (Mya) [46,47]; 2) Numididae-Phasianidae split at 52.4±5.1 Mya as the fossil *Gallinuloides wyomingensis* date from 54 Mya [48,49]; 3) Ancestors of *Arborophila* diverged from the other lineages in Phasianidae at around 39 Mya [46,50]; 4) The divergence between the *Coturnix* and *Gallus* was around 35 Mya [46]; 5) *G. gallus* and *G. sonneratii* diverged around 8 Mya [7]. Divergence times of all 45 birds were estimated in BEAST v2.3.3 [51]. In BEAST, we set uncorrelated lognormal relaxed clock model rate variation among branches. Four of five (C1~C4) compatible fossil calibrations were added to infer absolute time of each differentiation event while C5 was use to verify accuracy and of the inference. 95% confidence intervals included these constraints as calibration priors with a Yule process on the tree. The log-likelihood values and posterior probabilities were checked using TRACER v1.6 [52] to confirm that the chains had become stationary so that the ESS values of our parameters were all high (over 200 or so). The final analysis was run for 10 million generations with trees sampled every 1,000 generations. TreeAnnotator was then used to discard the first 20% trees and to generate the consensus tree. Geographic data of endemic bird areas of the 22 hill partridge species and the other 23 birds were obtained from the BirdLife database [53,54]. Distribution data of the 22 hill partridges were analyzed with R package to produce species distribution maps [55]. Altitude data was also analyzed using R package to produce raster data [56]. The BEAST output consensus tree and endemic bird area geographic data was analyzed in BioGeoBEARS to infer the biogeographic history of phylogenies of the 45 species. We used Dispersal-Extinction-Cladogenesis (DEC) and DEC+J models with AIC method and found no significant difference (p = 0.313) between them.

Results

Comparative mitochondrial genomes of *Arborophila*

The newly sequenced mitochondrial genome of *A. brunneopectus* is similar to the four previously published *Arborophila* mitochondrial genomes, consisting of 13 protein-coding genes, 22 tRNA genes, two ribosomal RNA genes, and a control region (D-loop) (Genbank accession Number: NC_022684.1) (S2 Table). Most mitochondrial genes are encoded on the H strand, except for the *ND6* gene and eight tRNA genes (tRNA-Gln, tRNA-Ala, tRNA-Asn, tRNA-Cys, tRNA-Tyr, tRNA-Ser (TGA), tRNA-Pro, tRNA-Glu). The base composition of the H strand of the Bar-backed Partridge mt-genome is: 30.40% A; 13.41% G; 24.86% T; 31.3% C, which reflects the typical A-T rich pattern in vertebrate mitochondrial genomes. The length of the *A. brunneopectus* mitogenome is 16719bp, which is marginally less than that of the other four hill partridges (16726~16728bp). Insertion/deletion variations occurred mainly in D-loop (loci 215, 219, 227, 298: 1bp per site), 12s rRNA (loci 1937: 6bp del) and 16s rRNA (loci 2834: 7bp del). There are 43 potential positive selective sites (see Fig 4 and S3 Fig) among which 11 sites are significant in mitogene alignments of *Arborophila* compared to that of the other 41 selected species. Notably, rate of variable sites around those potential positive selective loci are significantly higher than the remaining parts of the whole mitogenomes (p < 0.01, Fisher-Pitman Permutation Test). Single-nucleotide polymorphism sites (SNPs) are also intensive in *ND1*, *ND2*, *ND4* and *ND5* other than D-loop and rRNAs. Whereas, the conserved loci are often found in tRNAs and some specific areas of mitogenome components indicating functional importance, such as gene domain of active center of genes. SNPs distributions in *COX1*, *COX2*, *COX3*, *ATP6*, *ATP8* and *CYTB* show more homogeneous than that in the subunits of NADH dehydrogenase.
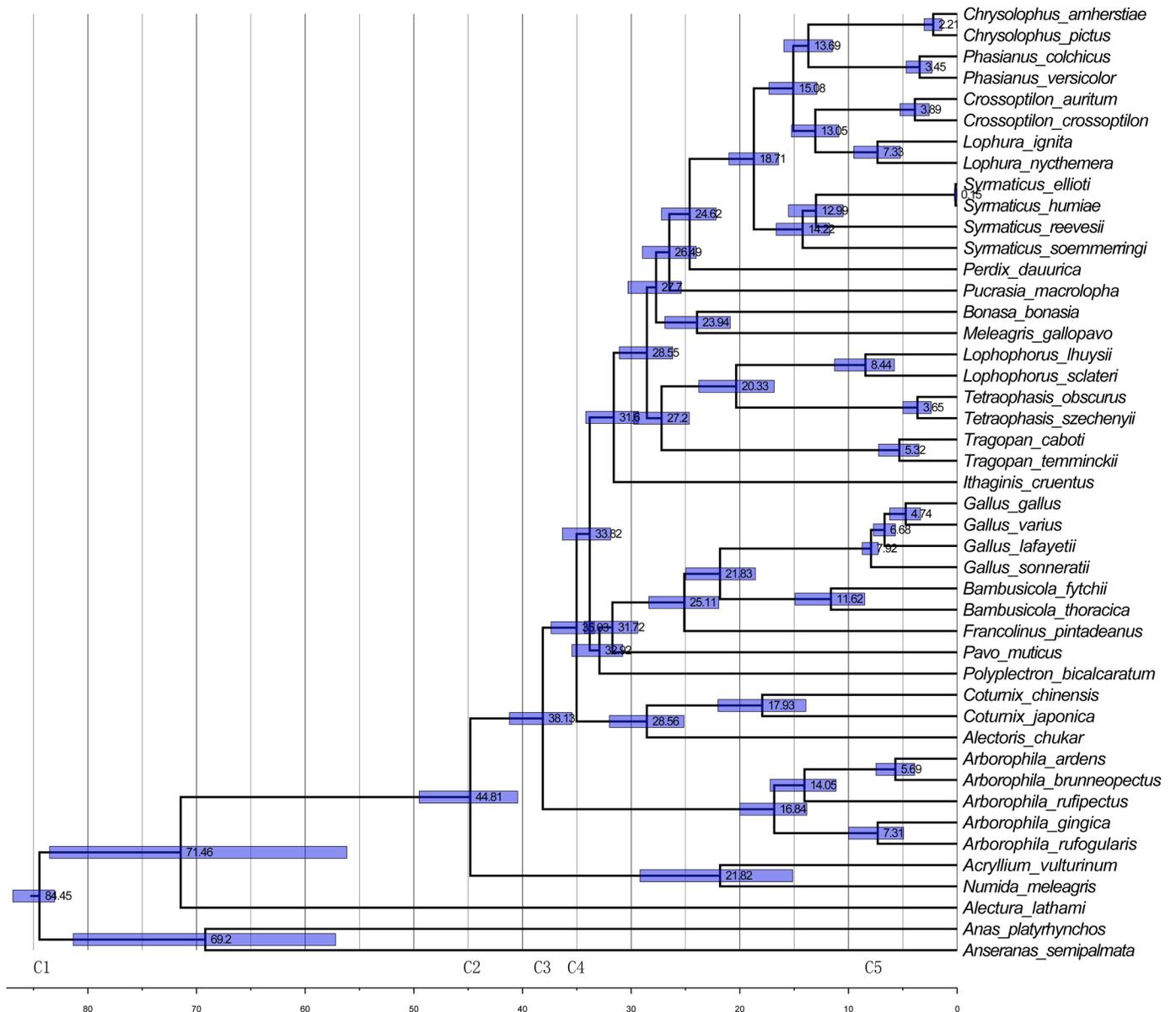
Fig 3. Species tree with estimated divergence time. The divergence times were estimated in BEAST with five fossil calibrations (C1-C5). Mean ages are included beside corresponding nodes. Horizontal gray bars on each node indicate the 95% credible interval of divergence time in millions of years, with the unit of divergence time as Mya.

The single insertion mutation in *ND3*

The insertion of an extra cytosine at position 147 of the *ND3* protein coding sequence that results in premature stopping of translation was confirmed in all the five hill partridge species [57,58]. The extra cytosine insertion information of 426 bird species from 30 orders is depicted in Fig 5. Ten of the 30 orders in our data contain WOC (without extra cytosine) *ND3*. All 166 Passeriformes species and one Apodiformes species have WOC *ND3*, but it is expected that additional information will be found for Apodiformes and thus this is preliminary. Falconiformes, Ciconiiformes and Cuculiformes have nearly equal WOC and WEC (with extra
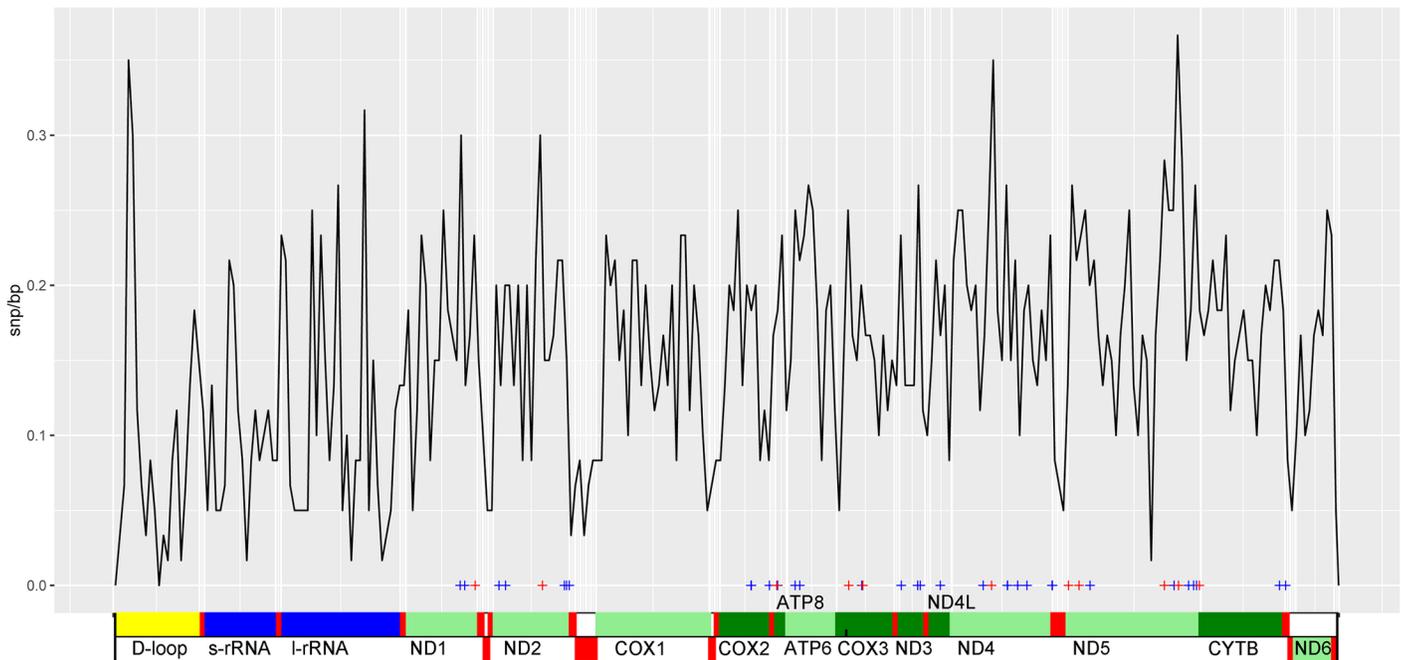
Fig 4. Plot of SNP rates in 60bp-length bins through the whole mitogenome of five hill partridges in this study. Potential positive selective sites detected in PAML are assigned a blue cross, and significant positive selective sites are assigned a red cross.

cytosine) *ND3*, while WEC *ND3* is dominant in the other five orders containing both WOC and WEC *ND3*. The near crown taxa shows a tendency to lose the extra cytosine insertion, however, the species without the insertion have not grouped but dispersed in different taxa indicating a complex evolutionary history of birds.

Phylogeny of *Arborophila* in Phasianidae

Phylogenetic analyses revealed robust tree, with similar topological structures (Fig 2). All Phasianidae species were divided into four deep lineages, as depicted in the consensus tree, among which *Arborophila* was given a basal phylogenetic position, apparently branching earlier than all other Phasianidae genera. The previously suggested taxa (clade I~IX in Fig 2), especially the Old World partridges, were separated and recombined in many situations. Among the five hill partridge species, *A. gingica* and *A. rufogularis* are a sister-group of a common ancestors shared by of the other three species, *A. ardens*, *A. brunneopectus* and *A. rufipectus*. *Arborophila. ardens* and *A. brunneopectus* show a closer relationship and diverged later than *A. rufipectus*. Such relationships of these five species are defined as type I and supported by two single gene trees (*ND5* and *ND6*). Eight genes (*ATP6*, *COX1*, *COX2*, *CYTB*, *ND2*, *ND3*, *ND4L* and *ND4*) support alternative relationships defined as type II that place *A. rufipectus* closer to *A. gingica* and *A. rufogularis* rather than *A. ardens* and *A. brunneopectus*. In all single gene trees, *ND5* gene tree has the least distance (RF = 8) relative to the consensus tree. Meanwhile, the tree inferred from combined data exclude *ND5* as it has the biggest distance (RF = 6) compared to the consensus tree (Table 2). The PBS analyses indicated that *ND5* (PBS = 686.83) contributed most to our consensus tree followed by *ND4* (PBS = 554.75) and *ND2* (PBS = 535.58), while *ATP8* (PBS = 21.91) made the least contribution,
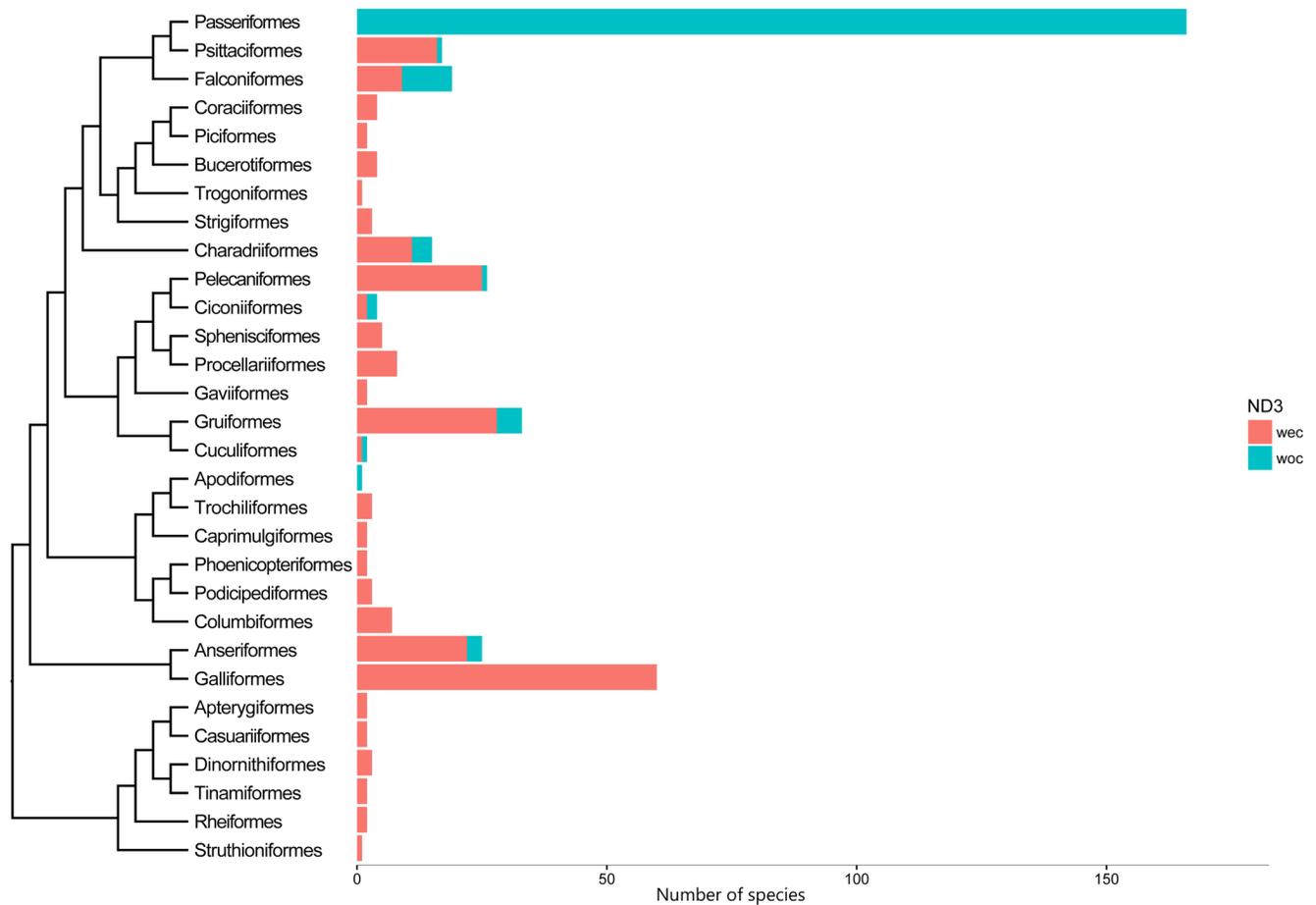
Fig 5. Barplot for counts of *ND3* with extra cytosine (WEC) or without extra cytosine (WOC) in each of the 30 bird orders studied. The data is extracted from 426 species mitogenomes. The phylogenetic relationships of the bird orders are constructed based on three recent studies [24–26].

Divergence and dispersal

The divergence time calculated from fossil calibrations (C1~C5 in Fig 3) suggests that common ancestors of Phasianidae, *Numida* and *Acryllium* diverged at around 45 Mya. *Arborophila* diverged from Phasianidae 7 Mya later, in Asia, which is consistent with Crow's and He's inferences [7,50]. After that time, ancestors of *A. rufogularis* and *A. gingica* diverged from *Arborophila* at around 16.84 Mya. *A. rufipectus* split from ancestors of *A. brunneopectus* and *A. ardens* at round 14.05 Mya, and these two species diverged at about 5.69 Mya, While *A. rufogularis* and *A. gingica* diverged at 7.31 Mya. All hill partridge species divergence events occurred during the Miocene when global temperatures underwent gradual cooling.

We collected GIS spatial data of the 22 South-east Asian hill partridge species from BirdLife databases [54]. The distribution areas of the five studied hill partridge species ranged across a wide area with little overlap and often geographically distinct and isolated populations (Fig 1). Only *A. brunneopectus* and *A. rufogularis* shared overlapping geographical distributions. *Arborophila rufipectus* and *A. ardens* represented the north and south, respectively, extents of the five species' distributions, while *A. rufogularis* and *A. gingica* represented the west and east (respectively). The geographic distributions of the other 17 south-east Asian hill partridge species were usually across mountainous terrain in Indochina and Maritime Southeastern Asia,

Table 2. Phylogenetic analyses using jackknifing of single mitogene.

mitogene	PBS	single gene mode			gene jackknifing mode		
		model	topology	RF distance	model	topology	RF distance
ND1	451.67	GTR+I+G	-	36	TVM+I+G	type I	2
ND2	535.58	GTR+I+G	type II	20	GTR+I+G	type I	0
COX1	351.75	GTR+I+G	type II	14	TVM+I+G	type I	0
COX2	160.41	TVM+I+G	type II	28	TVM+I+G	type I	0
ATP8	21.91	GTR+I+G	-	42	TVM+I+G	type I	0
ATP6	290	GTR+I+G	type II	26	TVM+I+G	type I	0
COX3	207	TVM+I+G	-	32	TVM+I+G	type I	0
ND3	51.42	TVM+I+G	type II	28	TVM+I+G	type I	0
ND4L	176	TIM+I+G	type II	36	TVM+I+G	type I	0
ND4	554.75	GTR+I+G	type II	18	GTR+I+G	type I	4
ND5	686.83	TVM+I+G	type I	8	TVM+I+G	type II	6
CYTB	466.16	GTR+I+G	type II	36	TVM+I+G	type I	2
ND6	199.5	K81uf+I+G	type I	28	GTR+I+G	type I	0

PBS: partitioned Bremer support value to the consensus tree; topology refer to clade of *Arborophila*, type I: (((*A. arden*, *A. brunneopectus*)*A. rufipectus*)(*A. gingica*, *A.rufogularis*)); type II: (((*A.gingica*, *A.rufogularis*)*A. rufipectus*)(*A. arden*, *A. brunneopectus*)); '-' means the inferred tree is not typical type I and type II; RF distance: Robinson-Foulds distance of gene tree to the consensus tree.

where human activities and disturbance can be uncommon. These 17 species only overlapped with the distributions of *A. rufogularis* and *A. brunneopectus* while are quite distant from the other three studied species(Fig 1). In addition to our phylogenetic analysis of the 45 species, the geographic distribution data allowed a reconstruction of historical/phylogenetic biogeography (S2 Fig) of these species. Historical/phylogenetic biogeography suggests that *Arborophila* split with other Phasianidae species within Asia and, similarly, that the intra-genus divergence likely occurred in Asia.

Discussion

SNPs bias positive selective sites within *Arborophila*

It was shown that all positive selection sites experienced nonsynonymous substitution within the five hill partridge species (S3 Fig), whereas, unique AAs were produced from mutations. For example, the mutation ND2: 35V>M (GTA>ATG) occurred in four of five species' mitogenomes, except *A. rufipectus*. It placed Methionine at the site where no Methionine had ever presented. Factors such as high-altitude adaptability [59] may be used to explain the selective but require systematically investigation on ecological data such as the altitude distribution of the Galliformes species. All the positive selection sites might potentially affect the function of NADH dehydrogenase. However, whether enzyme activity of NADH dehydrogenase is different between the studied species or lineages of other hill partridge species and other pheasants requires further verification.

Mitochondrial genes of the studied hill partridge species accumulate SNPs around loci under positive selection relative to other lineages observed. Significantly positive selected sites in *Arborophila* mitogenes are surrounded by a higher level of SNPs rate (proportion of SNP sites in a 60bp-length bin). For example, the SNPs abundant *ND5* contain the most potential positive selective sites (S3 Fig). According to population genetics theory, hitchhiking effort or selective sweep can lead to less genetic diversity around positive selective sites in a population during their fixation. However, whether this effect also applies to individuals from different

species living in diverse habitats is uncertain. Every selective site within hill partridge species is composed of different AAs suggests that they might have evolved in different directions. The uneven SNPs distribution in the hill partridge species mitogenomes reflects that selective pressure has not acted equally on the whole mitogenome. The interweaved conserved areas in the mitogenomes could be a specific character for specific lineages, which provides clues of functional domains of a mitogene or conserved sequence boxes in the domain II in D-loop [7]. Additionally, the relative conserved regions of genes may also provide robust and stable markers in the construction of phylogenetic trees, as is noted that the *COXI* or *CYTB* usually yield a consensus tree agreeing with combined data [8]. *ND5* presents a high level of variable sites even compared to D-loop, which indicates that *ND5* may be an alternative in resolving phylogenetic relationships within lower ranks of taxon and may be used intraspecific analysis [12].

Elimination of cytosine insertion in bird taxa

Clearly, the extra cytosine insertion or +1 frameshift in *ND3* is not a novel phenomenon in *Arborophila* as it is common across bird taxa and absent in others. For example, all species near the root of Palaeognathae have the insertion cytosine, while the younger and highly diverse Passeriformes does not ([Fig 5](#)). Nearly all Galloanserae species contain the extra cytosine, except for three Anseriformes species. The exact number/identity of species with and without the extra cytosine requires confirmation, as the uncommon insertion in a mito-gene might be overlooked or regarded as an error in sequencing.

Since many tortoise have the +1 frameshift mutation [58,60], we believe the 147th cytosine insertion is a phenotype inherited from ancestors of all birds. We analyzed the *ND3* insertion in Falconiformes mitogenomes to verify this ancestral heritage of cytosine insertion. Falconiformes is split with 10 WEC *ND3* and 10 WOC *ND3*. The WEC taxa include *Spizaetus*, *Nisaetus*, *Aquila*, *Buteo*, *Accipiter*, *Aegypius* and *Spilornis* and are placed near root while WOC taxainclude *Pandion*, *Sagittarius*, *Falco* and *Micrastur* are near the crown of phylogenetic tree (reconstructed in the same way described in our method). It is hypothesized that certain mitochondrial translation systems have the ability to tolerate frameshift insertions using programmed translational frame shifting [58], but why species in some younger taxa attempt to eliminate the insertion is unclear. One possible interpretation is that mitochondria are material and energetically costly (to sustain normal operations) and mitochondria decrease or offset the influence of frameshift insertion. While species without the deleterious mutation save energy and material and thus is advantageous for these species. It is still unclear why these WOC species dispersed among (relatively) unrelated bird taxa. The loss of an extra cytosine insertion is presumably independent because introgressive hybridization is rare in species from different orders.

Recent research into phylogenetic discrepancies

The results of our study mirror those of similar recent phylogenetic studies. The clades of the consensus tree formed confounding relationships for pheasants and partridges involving the families: Tetraonidae and Meleagrididae [10,11,28,61,62] ([Fig 2](#)). Similarly, our research found that lineages of pheasants and partridges are not monophyletic in as accordance with recent studies [8,62–64]. However recent studies have not been conclusive and the lineages of some species are ambiguous or controversial, such as the blood pheasant (*Ithaginis cruentus*) and peafowl inferred from diverse DNA markers [3,10,11,27,65–67]. Blood pheasant (*Ithaginis cruentus*) was placed at root of taxa composed of gallopheasants, Perdix, tragopans, Tetraoninae and Meleagrididae (BPP = 1; BSP = 0.98) in our consensus tree and some recent studies [10,56–58], but was placed as sister group of Tragopans or at root of family Phasianidae in

some other studies [10,59, 62]. Peafowls was commonly placed at root of junglefowls and some Old World partridges lineages as presented in our consensus tree [10,11,56,60,63,64], while some nuclear or combined phylogenetic trees rooted these species within tragopans or even gallopheasants[3,57,58].

Despite this study's results, the location of hill partridge species' root at the base of Phasianidae is still controversial. Some nuclear DNA markers, such as CR1 or nuclear ovomucoid intron G place this lineage at a near crown status [9,63]. However, few studies have studied phylogenetic relationships within *Arborophila* [3,12,28,68] akin to this study. Zheng et al. [4] constructed the relationships of 10 hill partridge species in China based on morphology and biogeographic distribution. Chen et al. [68] reconstructed phylogenetic tree of 10 *Arborophila* species with three mito-fragments (CYTB, ND2, COI) and 3 nuclear introns (OVOG, G3PDH, ALDOB). Wang et al. [28] developed trees for seven *Arborophila* species based on two mito-genes and six nuclear intron sequence. The relationships of the five hill-partridge species in our study are consistent with Zheng et al. [4] but are in conflict with Chen et al. [68] and Wang et al. [24] regarding the lineage of *A. rufipectus* in *Arborophila*. All four studies agreed regarding the closer relationship between *A. ardens* and *A. brunneopectus*, which was in accordance with their closer spatial distribution. However, the closely (phylogenetically) related *A. gingica* and *A. rufogularis* are geographically distant. Interestingly, the geographical distribution of *A. rufogularis* overlaps with *A. brunneopectus* but they are more phylogenetically remote than to other hill partridge species (Fig 1). The chief inconsistency regarding the relationships of hill partridge species is whether *A. rufipectus* shares a closer relationship with *A. gingica* and *A. rufogularis* or with *A. ardens* and *A. brunneopectus*. As *A. rufipectus* is geographically isolated and far from all other hill partridge species, its phylogeny may be key to discovering the evolutionary history of *Arborophila*.

The disparity between Galliformes phylogenetic trees and species-level divergence, and between mitochondrial and ambiguous results may be a result of protein-coding sequence convergence and high levels of incomplete lineage sorting [69] that occurred during their rapid radiation [24]. The disagreements of our consensus tree and previous studies may be due to two reasons. Firstly, evolutionary rates and modes of inheritance are different between nuclear DNA and mitoDNA. With faster evolutionary rates, mitogenes are advantageous for studying relationships of crown groups, while nuclear genes offer abundant information in resolving relationships of deeper clades of a tree. Phylogenetic relationships inferred from nuclear DNA might be bias for taxa in which introgressive hybridization is common [70], while such interference decreases in mitogenomes if it is restricted to male introgression [71]. Secondly, different mitogenes can be discrepant in resolving relationships of species, for evolutionary rate varied through the whole mitogenome. DNA fragments from different sources can have different evolutionary histories in one species.

To further understand which mitogene is responsible for different topologies in *Arborophila* taxa, we carefully reconstructed the phylogenetic trees with combined datasets as well as a single mitogene. The results suggested that *ND5* strongly affected the status of *A. rufipectus* within *Arborophila* from combined datasets. *ND5* was demonstrated under positive selection including five of the 11 significant positive selective sites and had faster evolutionary rate. As the longest component in 13-mitogenes combined data, *ND5* contributed most to the consensus tree with the highest PBS value (Table 2). Nevertheless, *ND5*-excluded combined data produced the tree of biggest RF distance relative to consensus tree, while tree inferred from only the *ND5* gene was closest to the consensus tree suggesting its dominance in influencing the topology of the consensus tree. Interestingly, the influence of *ND5* of the consensus tree was restricted to the taxa of *Arborophila*, while it changed the topology of the five hill partridge species from type I to type II. In an additional test, the *ND5* even reversed the phylogenetic tree

constructed from combined data of six nuclear DNA fragments, which had strongly supported the type I topology. When adding *ND5* to the combined data, the topology of the five hill partridge species turned from type I to type II (S4 Fig). However, when adding 13 mitogenes to the six nuclear combined data, the mitogenes-nuclear data supported the type I topology, but with rather low Bayesian posterior probabilities (0.56) at node (*A. rufipectus* (*A.gingica*, *A.rufogularis*)). Chen et al. [68] and Wang et al.[28] had not used the *ND5* gene in their phylogenetic analyses. Li et al. [72] reconstructed the phylogenetic tree without the *A. brunneopectus*, and they used the whole mitogenomes rather than 13 concatenated mitogenes, and with such treatment the much longer sequences may reduce or offset the influence of *ND5*. The whole mitogenome caused discordance and low support of deep nodes of the consensus tree so was discarded in our phylogenetic analyses.

The dominant influence of *ND5* on *Arborophila* clade topology suggests that the other 12 mitogenes or the six nuclear genes can offer limited phylogenetic information in *Arborophila* taxa. It is, however, still debated whether type I or type II is closer to the true evolutionary lineage of these five hill partridge species. The results of phylogenetic studies may also be affected by different length of gene segment lengths and lack of intermediate species. In this regard, we thought mitogenomes, especially when combined with nuclear DNA, were efficient molecular markers for the study of phylogenetic relationship among Phasianidae species [73–75]. However, cautiou is advised when choosing appropriate nuclear markers, as hybridization is common in bird taxa [73]. In many cases, the phylogenetic relationships of birds become confused when adding nuclear genes [9,11,63]. Compared to single mitochondrial gene sequence, the combined multi-gene sequence of mitogenomes can provide abundant information of the evolutionary relationships of many taxa within Phasianidae [8,76].

Climate shapes distributions of hill partridges

The divergence time of these five hill partridges were generally earlier in our results than those of Chen's [68] study. The only exception was that Chen et al. [68] proposed 48.8 Mya as the divergence point for ancestors of *Arborophila* when they split from other Phasianidae species, which was much earlier than our 38.13 Mya. Chen et al. [68] indicated that ancestors of *A. ardens* and *A. crudigularis* (originating from Indochina) arrived on their islands after the sea level retreat during the last glacial period started at 2.58 Mya [68,77], evolving into new species at approximately 2.2 Mya. However, it is still unclear when *A. ardens* diverged from *A. brunneopectus* and the natal location of their common ancestor.

The current geographic distribution of species can suggest their evolutionary history [78]. Crowe et al. [50] suggested that there was an initial dispersal by the common ancestor of *Xenoperdix* and *Arborophila* from Africa to Asia. Ning et al. [67] came to a similar conclusion using combined mitogenes and nuclear DNA data. These studies are consistent with the results of our historical/phylogenetic biogeography calculated in BioGeoBEARS (S2 Fig). We further demonstrated that ancestors of *Arborophila* split with other Phasianidae species in Asia and further divided into the diverse hill partridge species. We notice that all four divergnece events happened in the Miocene (23.03 to 5.33 Mya) where there was a moderately warm climate but with two major temperature decreases at 14 Mya and 8 Mya respectively [79–81]. These two global cooling events roughly coincide with the divergence events of *A. rufipectus* and *A. ardens/A. brunneopectus*, and *A. gingica* and *A. rufogularis* (Fig 3). We propose that climate cooling forced hill partridges species from higher altitudes to lower and migrated to the margins of, and away from, the Yungui and Tibetan Plateaus. Migration to warmer climates during such extreme cooling events may explains why the five studied hill partridge species are isolated and distant from the Tibetan and Yungui Plateaus, leaving a geographic zone devoid of

hill partridge species ([Fig 1](Fig 1)). Since extreme events have caused speciation of hill partridge species, the current habitat fragmentation and geographic isolation may prompt further speciation.

As root genus of Phasianidae, the origin of *Arborophila* is important in determining the inchoate evolutionary history of this family. Similarly, the phylogeny of *Arborophila* can increase the understanding of the geographical distribution and migration of these short-ranging birds who are poor fliers with low aerial maneuverability. However, the status of *A. rufipectus* within *Arborophila* is still unclear and it recommended that further analysis of molecular data of *Arborophila* is required.

Supporting information

S1 Fig. Analyses of substitution saturation at each codon position. Maximum likelihood saturation plots were compared between the codon positions of the complete nucleotide dataset. The Y = X line marks the theoretical limit where the number of observed substitutions equals the number of inferred substitutions. The slope of the linear regression indicates the amount of substitution saturation; the smaller the slope, the greater the number of inferred multiple substitutions. The blue 'x' and green hollow triangle refer to transition rate and transversion rate of 1st positions of codons respectively; the light blue solid triangle and red hollow triangle refer to transition rate and transversion rate of 2nd positions of codons respectively; the pink solid triangle and yellow hollow square refer to transition rate and transversion rate of 3rd positions of codons respectively; the navy blue solid square and green hollow diamond refer to transition rate and transversion rate of all DNA positions.
(TIF)

S2 Fig. Ancestral area reconstruction for Phasianidae and outgroups used in this study. Pie charts in the nodes indicates the probable ancestral areas calculated in BioGeoBEARS using DEC model. Squares at end of each branch are areas the corresponding species inhabited. As: Asia, Af: Africa, Am: America, Eu: Europe, Au: Australia.
(TIF)

S3 Fig. Collection of positive selection sites through the mitogenomes. The clade of hill partridges set as foreground and the other 40 birds background in selection pressure analysis. Significantly positive selection sites are signed with red asterisk at the top of the column. Amino acids of large differences in same column are filled with dark yellow.
(TIF)

S4 Fig. Molecular phylogenetic tree based on 13 mitochondrial protein-coding genes and six nuclear introns combined dataset using Bayesian inference (BI). The tree contains 35 of 45 species in the 13 combined mitogenes tree. Ignoring absent species in tree from the combined mito-nuclear genes, their topologies are the same except the status of *A. rufipectus*. The very low Bayesian posterior probability at that node may result from the discrepancies that mitogenes and nuclear introns yield different topologies at these branches.
(TIF)

S1 Table. Amplification and sequencing primers for complete mitochondrial DNA of *Arborophila brunneopectus*. * marks the primers used in LA-PCR, while the remaining primers are used in normal PCR.
(DOCX)

S2 Table. Mitochondrial DNA structure of *Arborophila brunneopectus*. ▲mtDNA components, genes of tRNA are signed as their correspond one-letter abbreviation, genes located in

L-strand are sign "-"behind their name *Numbers correspond to the nucleotides separating adjacent genes. Negative numbers indicate overlapping nucleotides.
(DOCX)

S3 Table. Mitochondrial *ND3* genes information of 426 birds. The states of *ND3* are signed with WOC: Without extra Cytosine and WEC: With extra Cytosine.
(XLSX)

Acknowledgments

Author Contributions

Conceptualization: Chaochao Yan, Feiyun Tu, Zhenxin Fan, Xiuyue Zhang.

Data curation: Chaochao Yan, Biqin Mou.

Formal analysis: Chaochao Yan, Biqin Mou, Yang Meng.

Funding acquisition: Bisong Yue.

Investigation: Chaochao Yan, Biqin Mou, Yang Meng.

Methodology: Feiyun Tu.

Project administration: Zhenxin Fan.

Software: Chaochao Yan.

Supervision: Bisong Yue.

Visualization: Chaochao Yan, Biqin Mou.

Writing – original draft: Chaochao Yan.

Writing – review & editing: Chaochao Yan, Megan Price, Xiuyue Zhang.

Reference

1. Clements J, Schulenberg T, Iliff M, Roberson D, Fredericks T, Sullivan B, et al. The eBird/Clements checklist of birds of the world: v2015. 2015. http://www.birds.cornell.edu/clementschecklist/download/

2. Madge S, McGowan PJ, Kirwan GM. Pheasants, partridges and grouse: a guide to the pheasants, partridges, quails, grouse, guineafowl, buttonquails and sandgrouse of the world: A&C Black; 2002. 321–2.

3. Chen D, Liu Y, Davison GW, Dong L, Chang J, Gao S, et al. Revival of the genus Tropicoperdix Blyth 1859 (Phasianidae, Aves) using multilocus sequence data. Zool J Linn Soc. 2015; 175(2): 429–438. https://doi.org/10.1111/zoj.12273

4. Zheng X, editor. Chinese Fauna, Aves, Volume IV, Galliformes. Peking, China: Scince Press, Chinese Academy of Sciences; 1978. pp. 84–99.

5. Cheng K-m, Wang C-s. China red data book of endangered animals: Aves. Beijing: Science Press; 1998.

6. IUCN. The IUCN Red List of Threatened Species. Version 2017–1. 2017. www.iucnredlist.org

7. He L, Dai B, Zeng B, Zhang X, Chen B, Yue B, et al. The complete mitochondrial genome of the Sichuan Hill Partridge (*Arborophila* rufipectus) and a phylogenetic analysis with related species. Gene 2009; 435(1–2): 23–8. https://doi.org/10.1016/j.gene.2009.01.001

8. Shen YY, Liang L, Sun YB, Yue BS, Yang XJ, Murphy R, et al. A mitogenomic perspective on the ancient, rapid radiation in the Galliformes with an emphasis on the Phasianidae. BMC Evol Biol. 2010; 10(1): 132–41. https://doi.org/10.1186/1471-2148-10-132 PMID: 20444289

9. Liu Z, He L, Yuan H, Yue B, Li J. CR1 retroposons provide a new insight into the phylogeny of Phasianidae species (Aves: Galliformes). Gene. 2012; 502(2): 125–32. http://dx.doi.org/10.1016/j.gene.2012.04.068 PMID: 22565186

10. Shen YY, Dai K, Cao X, Murphy R W., Shen XJ, Zhang YP. The Updated Phylogenies of the Phasianidae Based on Combined Data of Nuclear and Mitochondrial DNA. PLoS ONE. 2014; 9(4): e95786. https://doi.org/10.1371/journal.pone.0095786 PMID: 24748132

11. Liu F, Ma L, Yang C, Tu F, Xu Y, Ran J, et al. Taxonomic Status of Tetraophasis obscurus and Tetraophasis szechenyii (Aves: Galliformes: Phasianidae) Based on the Complete Mitochondrial Genome. Zool Sci. 2014; 31(3): 160–7. https://doi.org/10.2108/zsj.31.160 PMID: 24601778

12. Meiklejohn KA, Danielson MJ, Faircloth BC, Glenn TC, Braun EL, Kimball RT. Incongruence among different mitochondrial regions: a case study using complete mitogenomes. Mol Phylogenet Evol. 2014; 78: 314–23. https://doi.org/10.1016/j.ympev.2014.06.003 PMID: 24929245

13. Wiens JJ, Kuczynski CA, Stephens PR. Discordant mitochondrial and nuclear gene phylogenies in emydid turtles: implications for speciation and conservation. Biol J Linn Soc. 2010; 99(2): 445–61. https://doi.org/10.1111/j.1095-8312.2009.01342.x

14. Fontenot BE, Makowsky R, Chippindale PT. Nuclear–mitochondrial discordance and gene flow in a recent radiation of toads. Mol Phylogenet Evol. 2011; 59(1): 66–80. http://dx.doi.org/10.1016/j.ympev.2010.12.018 PMID: 21255664

15. Armstrong MH, Braun EL, Kimball RT. Phylogenetic utility of avian ovomucoid intron G: a comparison of nuclear and mitochondrial phylogenies in Galliformes. The Auk. 2001; 118(3): 799–804. https://doi.org/10.1642/0004-8038(2001)118[0799:PUOAOI]2.0.CO;2

16. Brusatte SL, Lloyd GT, Wang SC, Norell MA. Gradual assembly of avian body plan culminated in rapid rates of evolution across the dinosaur-bird transition. Curr Biol. 2014; 24(20): 2386–92. https://doi.org/10.1016/j.cub.2014.08.034 PMID: 25264248

17. Wyles JS, Kunkel JG, Wilson AC. Birds, behavior, and anatomical evolution. Proc Natl Acad Sci USA. 1983; 80(14): 43947.

18. Hill M, Eames JC, Trai L, Cu N. Population sizes, status and habitat associations of forest birds in Chu Yang Sin Nature Reserve, Dak Lak province, Vietnam. Bird Conserv Int. 2001; 11(1): 49–70. https://doi.org/10.1017/S095927090100106X

19. Yu CX, Yang G, Li D, Zhou F. Ecological distribution and spatial niche of pheasants in the Karst mountains of southwest Guangxi Province, China. Zoological Research. 2011; 32(5): 549–55. PMID: 22006809

20. Gao Y-r. Conservation status of endemic Galliformes on Hainan Island, China. Bird Conserv Int. 1998; 8(4): 411–6. https://doi.org/10.1017/S0959270900002161

21. Shen YY, Shi P, Sun YB, Zhang YP. Relaxation of selective constraints on avian mitochondrial DNA following the degeneration of flight ability. Genome Res. 2009; 19(10): 1760–5. https://doi.org/10.1101/gr.093138.109 PMID: 19617397

22. Sambrook J, Russell DW, Janssen K, Argentine J. Molecular cloning: a laboratory manual. 3 ed. Peking: Science Press; 2002.

23. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic acids Res. 1997; 25(5): 955–64. PMID: 9023104

24. Yuri T, Kimball RT, Harshman J, Bowie RC, Braun MJ, Chojnowski JL, et al. Parsimony and model-based analyses of indels in avian nuclear genes reveal congruent and incongruent phylogenetic signals. Biology (Basel). 2013; 2(1): 419–44. https://doi.org/10.3390/biology2010419 PMID: 24832669

25. Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science. 2014; 346(6215): 1320–31. https://doi.org/10.1126/science.1253451 PMID: 25504713

26. Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, et al. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature. 2015; 526(7574): 569–73. https://doi.org/10.1038/nature15697 PMID: 26444237

27. Ma LL, Zhang XY, Yue BS, Ran JH. Complete mitochondrial genome of the Chinese Monal pheasant Lophophorus lhuysii, with phylogenetic implication in Phasianidae. Mitochondr DNA. 2010; 21(1): 5–7. http://dx.doi.org/10.3109/19401730903552026

28. Wang N, Kimball RT, Braun EL, Liang B, Zhang Z. Assessing phylogenetic relationships among Galliformes: a multigene phylogeny with expanded taxon sampling in Phasianidae. PloS one. 2013; 8(5): e64312. https://doi.org/10.1371/journal.pone.0064312 PMID: 23741315

29. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011; 28(10): 2731–9. https://doi.org/10.1093/molbev/msr121 PMID: 21546353

30. Gibson A, Gowri-Shankar V, Higgs PG, Rattray M. A comprehensive analysis of mammalian mitochondrial genome base composition and improved phylogenetic methods. Mol Biol Evol. 2005; 22(2): 251–64. https://doi.org/10.1093/molbev/msi012 PMID: 15483324

31. Xia X, Xie Z, Salemi M, Chen L, Wang Y. An index of substitution saturation and its application. Mol Phylogenet Evol. 2003; 26(1): 1–7. https://doi.org/10.1016/S1055-7903(02)00326-3 PMID: 12470932

32. Xia X, Lemey P. Assessing substitution saturation with DAMBE. in: Lemey P, Salemi M, Vandamme AM, editors. The Phylogenetic Handbook. Cambridge: Cambridge University Press; 2009: pp. 611–26.

33. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 2012; 9(8): 772-. https://doi.org/10.1038/nmeth.2109 PMID: 22847109

34. Mayr G. The fossil record of galliform birds: comments on Crowe et al.(2006). Cladistics. 2008; 24: 74–6. https://doi.org/10.1111/j.1096-0031.2007.00170.x

35. Burnham KP, Anderson DR. Multimodel inference understanding AIC and BIC in model selection. Sociol Methods Res. 2004; 33(2): 261–304. https://doi.org/10.1177/0049124104268644

36. Lanfear R, Calcott B, Ho SY, Guindon S. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol. 2012; 29(6): 1695–1701. https://doi.org/10.1093/molbev/mss020 PMID: 22319168

37. Altekar G, Dwarkadas S, Huelsenbeck JP, Ronquist F. Parallel metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics. 2004; 20(3): 407–15. https://doi.org/10.1093/bioinformatics/btg427 PMID: 14960467

38. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003; 19(12)_: 1572–4. https://doi.org/10.1093/bioinformatics/btg180 PMID: 12912839

39. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic Biol. 2010; 59(3): 307–21. https://doi.org/10.1093/sysbio/syq010

40. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML Web servers. Systematic Biol. 2008; 57(5): 758–71. https://doi.org/10.1080/10635150802429642

41. Lambkin CL, Lee MSY, Winterton SL, Yeates DK. Partitioned Bremer support and multiple trees. Cladistics. 2005; 18(4): 436–44. https://doi.org/10.1111/j.1096-0031.2002.tb00159.x

42. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol. 2012; 3(2): 217–23. https://doi.org/10.1111/j.2041-210X.2011.00169.x

43. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. MolBiol Evol. 2007; 24(8): 1586–91. https://doi.org/10.1093/molbev/msm088

44. Yang Z, Wong WS, Nielsen R. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol. 2005; 22(4): 1107–18. https://doi.org/10.1093/molbev/msi097 PMID: 15689528

45. Dyke GJ, Tuinen MV. The evolutionary radiation of modern birds (Neornithes): reconciling molecules, morphology and the fossil record. Zool J Linn Soc. 2004; 141(2): 153–77. https://doi.org/10.1111/j.1096-3642.2004.00118.x

46. Tuinen MV, Dyke GJ. Calibration of galliform molecular clocks using multiple fossils and genetic partitions. Mol Phylogenet Evol. 2004; 30(1): 74–86. https://doi.org/10.1016/S1055-7903(03)00164-7 PMID: 15022759

47. Tuinen MV, Hedges SB. Calibration of avian molecular clocks. Mol Biol Evol. 2001; 18(2): 206–13. https://doi.org/10.1093/oxfordjournals.molbev.a003794 PMID: 11158379

48. Ksepka DT. Broken gears in the avian molecular clock: new phylogenetic analyses support stem galliform status for Gallinuloides wyomingensis and rallid affinities for Amitabha urbsinterdictensis. Cladistics. 2009; 25(2): 173–97. https://doi.org/10.1111/j.1096-0031.2009.00250.x

49. Pereira SL, Baker AJ. A molecular timescale for galliform birds accounting for uncertainty in time estimates and heterogeneity of rates of DNA substitutions across lineages and sites. Mol Phylogenet Evol. 2006; 38(2): 499–509. https://doi.org/10.1016/j.ympev.2005.07.007 PMID: 16112881

50. Crowe TM, Bowie RCK, Bloomer P, Mandiwana TG, Hedderson TAJ, Randi E, et al. Phylogenetics, biogeography and classification of, and character evolution in, gamebirds (Aves: Galliformes): effects of character exclusion, data partitioning and missing data. Cladistics. 2006; 22(6): 495–532. https://doi.org/10.1111/j.1096-0031.2006.00120.x

51. Li WL, Drummond AJ. Model averaging and Bayes factor calculation of relaxed molecular clocks in Bayesian phylogenetics. Mol Biol Evol. 2012; 29: 751–61. https://doi.org/10.1093/molbev/msr232 PMID: 21940644

52. Rambaut A, Suchard M, Xie D, Drummond A. Tracer v1. 6. http://beastbioedacuk/Tracer. 2014.

53. International B. Endemic Bird Areas (EBAs). Cambridge, UK: BirdLife International. 2016;- Species distribution data download. www.birdlife.org/datazone/species.

54. NatureServe BIa. Bird species distribution maps of the world. 2016 ed. Cambridge, UK and NatureServe, Arlington, USA: BirdLife International. 2015.

55. Richard AB, Allan RW, Ray B, Thomas PM, Alex D. maps: Draw Geographical Maps [R package maps version 3.1.1]. R Package. 2016.

56. Hijmans RJ. Geographic Data Analysis and Modeling [R package raster version 2.5.8]. R Package. 2015.

57. Mindell DP, Sorenson MD, Dimcheff DE. An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. Mol Biol Evol. 1998; 15(11): 1568–71. PMID: 12572620

58. Russell RD, Beckenbach AT. Recoding of Translation in Turtle Mitochondrial Genomes: Programmed Frameshift Mutations and Evidence of a Modified Genetic Code. J Mol Evol. 2008; 67(6): 682–95. https://doi.org/10.1007/s00239-008-9179-0 PMID: 19030769

59. Zhou T, Shen X, Irwin DM, Shen Y, Zhang Y. Mitogenomic analyses propose positive selection in mitochondrial genes for high-altitude adaptation in galliform birds. Mitochondrion. 2014; 18: 70–5. https://doi.org/10.1016/j.mito.2014.07.012 PMID: 25110061

60. Parham JF, Macey JR, Papenfuss TJ, Feldman CR, Türkozan O, et al. The phylogeny of Mediterranean tortoises and their close relatives based on complete mitochondrial genome sequences from museum specimens. Mol Phylogenet Evol. 2006; 38: 50–64. https://doi.org/10.1016/j.ympev.2005.07.015 PMID: 16150614

61. Jiang L, Wang G, Peng R, Peng Q, Zou F. Phylogenetic and molecular dating analysis of Taiwan Blue Pheasant (*Lophura swinhoii*). Gene. 2014; 539(1): 21–9. https://doi.org/10.1016/j.gene.2014.01.067 PMID: 24508942

62. Kimball RT, Mary CM, Braun EL. A macroevolutionary perspective on multiple sexual traits in the Phasianidae (Galliformes). Int J Evol Biol. 2011: 423938. http://dx.doi.org/10.4061/2011/423938 PMID: 21716735

63. Meng Y, Dai B, Ran J, Li J, Yue B. Phylogenetic position of the genus Tetraophasis (Aves, Galliformes, Phasianidae) as inferred from mitochondrial and nuclear sequences. Biochem Syst Ecol. 2008; 36(8): 626–37. https://doi.org/10.1016/j.bse.2008.01.007

64. Yan C, Zhou Y, Lu L, Tu F, Huang T, Zhang X, et al. Complete mitochondrial genome of Hainan partridge, *Arborophila ardens* (Galliformes: Phasianidae). Mitochondr DNA. 2014; 25(4): 259–60. https://doi.org/10.3109/19401736.2013.800494 PMID: 23795844

65. Huang Z, Liu N, Xiao Y, Cheng Y, Mei W, Wen L, et al. Phylogenetic relationships of four endemic genera of the Phasianidae in China based on mitochondrial DNA control-region genes. Mol Phylogenet Evol. 2009; 53(2): 378–83. https://doi.org/10.1016/j.ympev.2009.07.002 PMID: 19591951

66. Kimball RT, Braun EL. A multigene phylogeny of Galliformes supports a single origin of erectile ability in non-feathered facial traits. J Avian Biol. 2008; 39(4): 438–45. https://doi.org/10.1111/j.0908-8857.2008.04270.x

67. Wang N, Kimball RT, Braun EL, Liang B, Zhang Z. Ancestral range reconstruction of Galliformes: the effects of topology and taxon sampling. J Biogeogr. 2017; 44(1): 122–35. https://doi.org/10.1111/jbi.12782

68. Chen D, Chang J, Li SH, Liu Y, Liang W, Zhou F, et al. Was the exposed continental shelf a long-distance colonization route in the ice age? The Southeast Asia origin of Hainan and Taiwan partridges. Mol Phylogenet Evol. 2015; 83C: 167–73. https://doi.org/10.1016/j.ympev.2014.11.006 PMID: 25463750

69. Rogers J, Gibbs RA. Comparative primate genomics: emerging patterns of genome content and dynamics. Nat Genet. 2014; 15: 347–59. https://doi.org/10.1038/nrg3707 PMID: 24709753

70. Roos C, Zinner D, Kubatko LS, Schwarz C, Yang M, et al. Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. BMC Evol Biol. 2011; 11(1): 77. https://doi.org/10.1186/1471-2148-11-77 PMID: 21435245

71. Mitchell SM, Muehlbauer LK, Freedberg S. Nuclear introgression without mitochondrial introgression in two turtle species exhibiting sex-specific trophic differentiation. Ecol Evol. 2016; 6(10): 3280–8. https://doi.org/10.1002/ece3.2087 PMID: 27252833

72. Li X, Huang Y, Lei F. Comparative and phylogenomic analyses on mitochondrial genomes of *Arborophila* species. Hereditas. 2014; 36(9): 912–20. PMID: 25252309

73. Grant PR, Grant R. Hybridization of bird species. Science. 1992; 256(5054): 193–97. https://doi.org/10.1126/science.256.5054.193 PMID: 17744718

74. Shaw KL. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proc Natl Acad Sci USA. 2002; 99(25): 16122–7. https://doi.org/10.1073/pnas.242585899 PMID: 12451181

75. Powell AFLA, Barker FK, Lanyon SM. Empirical evaluation of partitioning schemes for phylogenetic analyses of mitogenomic data: An avian case study. Mol Phylogenet Evol. 2013; 66(1): 69–79. https://doi.org/10.1016/j.ympev.2012.09.006 PMID: 23000817

76. Willerslev E, Gilbert MT, Binladen J, Ho SY, Campos PF, Ratan A, et al. Analysis of complete mitochondrial genomes from extinct and extant rhinoceroses reveals lack of phylogenetic resolution. BMC Evol Biol. 2009; 9: 95–106. https://doi.org/10.1186/1471-2148-9-95 PMID: 19432984

77. Raymo ME, Mitrovica JX, O'Leary MJ, DeConto RM, Hearty PJ. Departures from eustasy in Pliocene sea-level records. Nat Geosci. 2011; 4(5): 328–32. https://doi.org/10.1038/ngeo1118

78. Dynesius M, Jansson R. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. P Natl Acad Sci USA. 2000; 97(16): 9115–20. https://doi.org/10.1073/pnas.97.16.9115

79. Zachos J, Pagani M, Sloan L, Thomas E, Billups K. Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. Science. 2001; 292: 686–93. https://doi.org/10.1126/science.1059412 PMID: 11326091

80. Zhisheng A, Kutzbach JE, Prell WL, Porter SC. Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times. Nature. 2001; 411(6833): 62–6. https://doi.org/10.1038/35075035 PMID: 11333976

81. Cerling TE, Harris JM, MacFadden BJ, Leakey MG, Quade J, Eisenmann V, et al. Global vegetation change through the Miocene/Pliocene boundary. Nature. 1997; 389(6647): 153–8. https://doi.org/10.1038/38229