PLOS ONE

RESEARCH ARTICLE

Molecular cytogenetic characterization of repetitive sequences comprising centromeric heterochromatin in three Anseriformes species

Yoshinobu Uno[1☯¤*], **Chizuko Nishida**[2☯], **Ayano Hata**[1], **Satoshi Ishishita**[3], **Yoichi Matsuda**[1,3*]

1 Laboratory of Avian Bioscience, Department of Animal Sciences, Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya, Aichi, Japan, **2** Department of Natural History Sciences, Faculty of Science, Hokkaido University, Sapporo, Hokkaido, Japan, **3** Avian Bioscience Research Center, Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya, Aichi, Japan

☯ These authors contributed equally to this work.
¤ Current address: Laboratory for Phyloinformatics, RIKEN Center for Biosystems Dynamics Research (BDR), Kobe, Hyogo, Japan
* yoshinobu.uno@riken.jp (YU); yoimatsu@agr.nagoya-u.ac.jp (YM)

Citation: Uno Y, Nishida C, Hata A, Ishishita S, Matsuda Y (2019) Molecular cytogenetic characterization of repetitive sequences comprising centromeric heterochromatin in three Anseriformes species. PLoS ONE 14(3): e0214028. https://doi.org/10.1371/journal.pone.0214028

Editor: Roscoe Stanyon, University of Florence, ITALY

Received: January 14, 2019

Accepted: March 5, 2019

Published: March 26, 2019

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files. All repetitive sequences identified newly in this study are available from the DDBJ database (accession numbers LC416770-LC416795).

Funding: This work was supported by a Grant-in-Aid for Scientific Research (B) (no. 22370081) from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Abstract

The highly repetitive DNA sequence of centromeric heterochromatin is an effective molecular cytogenetic marker for investigating genomic compartmentalization between macrochromosomes and microchromosomes in birds. We isolated four repetitive sequence families of centromeric heterochromatin from three Anseriformes species, viz., domestic duck (*Anas platyrhynchos*, APL), bean goose (*Anser fabalis*, AFA), and whooper swan (*Cygnus cygnus*, CCY), and characterized the sequences by molecular cytogenetic approach. The 190-bp APL-*Hae*III and 101-bp AFA-*Hin*fI-S sequences were localized in almost all chromosomes of *A. platyrhynchos* and *A. fabalis*, respectively. However, the 192-bp AFA-*Hin*fI-L and 290-bp CCY-*Apa*I sequences were distributed in almost all microchromosomes of *A. fabalis* and in approximately 10 microchromosomes of *C. cygnus*, respectively. APL-*Hae*III, AFA-*Hin*fI-L, and CCY-*Apa*I showed partial sequence homology with the chicken nuclear-membrane-associated (CNM) repeat families, which were localized primarily to the centromeric regions of microchromosomes in Galliformes, suggesting that ancestral sequences of the CNM repeat families are observed in the common ancestors of Anseriformes and Galliformes. These results collectively provide the possibility that homogenization of centromeric heterochromatin occurred between microchromosomes in Anseriformes and Galliformes; however, homogenization between macrochromosomes and microchromosomes also occurred in some centromeric repetitive sequences.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Highly repetitive DNA sequences are one of the major components of chromosomes, which are generally divided into two categories, viz., interspersed repetitive sequences and site-specific repetitive sequences, based on their genomic organization and chromosomal distribution [1]. Two subtypes of non-long terminal repeat (LTR) retrotransposons, viz., long interspersed elements (LINEs) and short interspersed elements (SINEs), are well-known as major components of interspersed-type repetitive sequences, which are distributed throughout the genome. Site-specific highly repetitive sequences constituting the heterochromatin viz., centromeric repetitive DNA sequences, non-centromeric chromosome site-specific repetitive sequences, microsatellite repeat motifs, etc., are tandem duplicated and are usually present as more than 10,000 copies in the genome, which have important roles in chromosome organization, sex chromosome differentiation, and chromatin architecture in interphase nuclei [2–4]. Centromeric heterochromatin-associated highly repetitive DNA sequences have been isolated and characterized from a high number of vertebrates. These sequences are generally susceptible to rapid nucleotide substitution; therefore, they often rapidly evolve in a concerted manner, resulting in low and intraspecific sequence variation but a higher degree of interspecific sequence variation [5–7]. This indicates that the centromeric repetitive sequence is a good taxonomic and phylogenetic marker for reconstructing the evolutionary relationships between closely related species that share the same origin of repetitive sequence families.

The typical avian karyotype consists of 6–10 pairs of macrochromosomes, including sex chromosomes, and a number of indistinguishable microchromosomes, with diploid numbers ranging from 74 to 84 [8–10]. Genome sequencing of many avian species, including chicken (*Gallus gallus*, Galliformes), has revealed that there are distinct structural differences between macrochromosomes and microchromosomes in terms of multiple factors such as of recombination rate, GC and CpG contents, gene density, and density of repetitive sequences [11–13]. In Galliformes, repetitive sequences distributed primarily on microchromosomes have been isolated from chicken [chicken nuclear-membrane-associated (CNM) repeat] [14], the Japanese quail (*Coturnix japonica*; CJA-*Bgl*II) [15], turkey (*Melleagris gallopavo*; TM repeat) [16], the Blue-breasted Quail (*Coturnix chinensis*; CCH-S) [17], bobwhite quail (*Colinus virginianus*; CVI-*Msp*I) [18], and chukar partridge (*Alectoris chukar*; ACH-*Sau3A*I) [18]. In other avian and reptilian species, microchromosome-specific centromeric repetitive sequences or microsatellite repeat motifs on microchromosomes have also been observed the lesser rhea (*Pteronemia pennata*) and greater rhea (*Rhea americana*), of Struthioniformes [19]; Japanese mountain hawk-eagle (*Nisaetus nipalensis orientalis*), of Falconiformes [20]; campo flicker (*Colaptes campestris*), green-barred woodpecker (*C. melanochloros*), and white woodpecker (*melanerpes candidus*) of Piciformes [21]; and Chinese soft-shelled turtle (*Pelodiscus sinensis*) [22]. These results suggest that the centromeric repetitive sequences are homogenized predominantly between microchromosomes in Aves and turtles, known as chromosome size-correlated genomic compartmentalization. Among these microchromosome-specific repeat families, four CNM repeat families (CNM repeat, TM repeat, CCH-S, and ACH-*Sau3A*I) contain 12–17-bp T-rich and A-rich motifs, in which the conserved A_{3-5} and T_{3-5} reiterations are separated by 6–7 bp [14,16,18]. This result indicates that the CNM repeat families originate from the same ancestral repetitive sequence in the common ancestor. However, chromosome size-correlated compartmentalization is collapsed in several avian species, in which the centromeric repetitive sequences localized to both macrochromosomes and microchromosomes have been identified [18,20,23,24]. Therefore, molecular cytogenetic characterization of the centromeric repetitive sequences in other avian orders provides important information on the evolutionary changes in genomic organization of avian macrochromosomes and microchromosomes.

The order Anseriformes consists of three extant families, viz., Anatidae, Anhimidae, and Anseranatidae, containing over 180 species; they are highly adapted for swimming on the surface of water. Anseriformes is considered to have appeared approximately 77 million years ago when the ancestral Galloanserae split into the two main lineages of Anseriformes and Galliformes [25]. Anseriformes and Galliformes are the most basal lineages of neognathous birds that appeared after Palaeognathae in the phylogeny of birds. The karyotypes of Anseriformes are characterized by high diploid chromosome numbers (2n = 78–98), and their typical karyotypes are composed of 78–80 chromosomes that consist of a small number of macrochromosomes and numerous microchromosomes, which are similar to typical avian karyotypes [8–10,26]. The tandem repetitive sequence, RBMII, which was isolated from the Red-breasted Merganser (*Mergus serrator*) [27], was found in all 22 Anseriformes species, including the domestic duck (*A. platyrhynchos*) [28]. However, in Anseriformes, molecular characterization of repetitive sequences is limited to RBMII sequences, whose chromosomal distribution is not known yet.

In this study, to improve our understanding of chromosome size-correlated genomic compartmentalization in Aves, we isolated repetitive sequences constituting centromeric heterochromatin from three Anseriformes species, viz., the domestic duck (*A. platyrhynchos*), bean goose (*A. fabalis*), and whooper swan (*C. cygnus*), and characterized the sequences molecular cytogenetically. We examined the chromosomal locations, genomic organization, and sequence conservation among the avian species using fluorescent in situ hybridization (FISH) and filter hybridization. Finally, we discussed the molecular evolution of the centromeric repetitive sequences in Anseriformes and genomic compartmentalization between macrochromosomes and microchromosomes in Aves based on the obtained data.

Materials and methods

Ethics statement

Animal care and all experimental procedures were conducted according to the guidelines for the care and use of experimental animals of Nagoya University. The animal protocols were approved by the Animal Experiment Committee, Graduate School of Bioagricultural Sciences, Nagoya University (approved no. 2009090901).

Cell culture, chromosome preparation, and chromosome banding

For cell culture, small pieces of skin tissues were collected from a 1-month-old female domestic duck (*A. platyrhynchos*, Anatidae), purchased from a breeding farm in Japan, and a female bean goose (*A. fabalis*, Anatidae) and two female whooper swans (*C. cygnus*, Anatidae) from the Asahiyama Zoo, Asahikawa and Kushiro Zoo, Kushiro, respectively, Hokkaido, Japan. Fibroblasts were cultured in Medium 199 (Thermo Fisher Scientific-GIBCO, Carlsbad, CA, USA) supplemented with 15% fetal bovine serum (Thermo Fisher Scientific-GIBCO), 100 μg/ml kanamycin, and 1% antibiotic–antimycotic solution (PSA) (Thermo Fisher Scientific-GIBCO) at 39˚C in 5% CO_2. For Giemsa-stained and C-banded karyotype analyses, fibroblasts were collected 30 min after colcemid treatment, suspended in 0.075 M KCl for 20 min, and then fixed with 3:1 methanol/acetic acid. Chromosome preparations were prepared following the standard air-drying method. Replication-banded chromosome slides were prepared for in situ hybridization as described previously [29]. The cultured cells were treated with 5-bromodeoxyuridine (BrdU) (25 μg/ml) (Sigma-Aldrich, St Louis, MO, USA) at the late replication stage for 4.5 h, including the 30-min colcemid treatment. After staining the slides with Hoechst 33258 (1 μg/ml) for 5 min, replication bands were obtained by heating at 65˚C for 3 min and

exposing to UV light at 65˚C for an additional 6 min. The slides were maintained at −80˚C until use.

To examine the chromosomal distribution of centromeric heterochromatin, C-banding was performed by the barium hydroxide/saline/Giemsa method [30] with slight modification; chromosome slides were treated with 0.2 M HCl for 5 min at room temperature and then with 5% $Ba(OH)_2$ at 50˚C for 5–7 min.

FISH

For cross-species chromosome painting, we used DNA probes of chicken (*G. gallus*, GGA) chromosomes 1–9 and Z [31] and a mixture of microchromosome-specific paints, consisting of 20 pairs of chicken microchromosomes [32] as described previously [33]. FISH probes were labeled with biotin-16-dUTP (Roche Diagnostics, Basel, Switzerland) by nick translation and hybridized to metaphase spreads. After washing, the slides were incubated with Avidin, Alexa Fluor 488 conjugate (Thermo Fisher Scientific-Molecular Probes, Carlsbad, CA, USA) and stained with 0.75 μg/ml propidium iodide (PI). For dual-color FISH, the other probe was labeled with digoxigenin (DIG)-11-dUTP (Roche Diagnostics) and stained with rhodamine-conjugated anti-DIG Fab fragments (Roche Diagnostics) [34].

Molecular cloning and nucleotide sequencing of repetitive DNA sequences

High molecular weight genomic DNA was extracted from the liver of female *A. platyrhynchos* and cultured fibroblasts of female *A. fabalis* and *C. cygnus*. Genomic DNA of *A. platyrhynchos* was digested with 10 restriction endonucleases, viz., *Alu*I, *BamH*I, *Bgl*II, *Eco*RI, *Hae*III, *Hind*II, *Msp*I, *Sae*I, *Sma*I, and *Taq*I. Genomic DNA of *A. fabalis* and *C. cygnus* was digested with 19 restriction endonucleases, viz., *Apa*I, *BamH*I, *Bgl*I, *Bgl*II, *Eco*RI, *Eco*RV, *Hae*III, *Hind*III, *Hinf*I, *Nsi*I, *Pvu*II, *Rsa*I, *Sac*I, *Sau3A*I, *Sma*I, *Sph*I, *Taq*I, *Xba*I, and *Xho*I. The digested genomic DNA was size-fractionated by 3% agarose gel electrophoresis and then stained with ethidium bromide. Prominent DNA bands of repetitive sequences were eluted from the gel using the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) and cloned into the pGEM-7f(+) vector (Promega, Madison, WI, USA). Nucleotide sequences were determined using ABI PRISM 3130 DNA Analyzer (Thermo Fisher Scientific-Applied Biosystems, Carlsbad, CA, USA) after cycle-sequencing reactions with the BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific-Applied Biosystems). Dot matrix analysis of the nucleotide sequences was performed with MAFFT version 7 (http://mafft.cbrc.jp/alignment/server/).

Southern blot hybridization

Genomic DNA was digested with restriction endonucleases, fractionated on 2% agarose gel, and transferred onto nylon membranes (Roche Diagnostics). DNA fragments of repetitive sequences were labeled with DIG-11-dUTP using PCR DIG Labeling Mix (Roche Diagnostics) and hybridized to the membrane. Hybridization was performed overnight at 45˚C in DIG Easy Hyb Solution (Roche Diagnostics). After hybridization, the membrane was washed at 45˚C in 0.1% sodium dodecyl sulfate (SDS)/2× saline sodium citrate (SSC), 0.1% SDS/1× SSC, 0.1% SDS/0.5× SSC, and 0.1% SDS/0.1× SSC for 15 min each. Chemiluminescent signals were detected with anti-DIG-AP Fab fragments and CDP-Star (Roche Diagnostics) and exposed to Biomax MS-1 Autoradiography Film (Carestream Health, Rochester, NY, USA).

Slot blot hybridization

To examine the nucleotide sequence conservation of repetitive sequences among avian species, slot blot hybridization was performed with DIG-11-dUTP-labeled DNA fragments as described in our previous studies [18,35]. We used genomic DNA from females of 17 avian species representing 10 orders, viz., ostrich (*Struthio camelus*) and emu (*Dromaius novaehollandiae*) belonging to Struthioformes; elegant crested tinamou (*Eudromia elegans*) belonging to Tinamiformes; helmeted guineafowl (*Numida meleagris*), Japanese quail (*C. japonica*), and chicken (*G. gallus*) belonging to Galliformes; bean goose (*A. fabalis*), whooper swan (*C. cygnus*), and domestic duck (*A. platyrhynchos*) belonging to Anseriformes; Siberian crane (*Grus leucogeranus*) belonging to Gruiformes; Black-faced Spoonbill (*Platalea minor*) belonging to Pelecaniformes; Blakiston's fish owl (*Bubo blakistoni*) belonging to Strigiformes; osprey (*Pandion haliaetus*) and Japanese mountain hawk-eagle (*N. nipalensis orientalis*) belonging to Accipitriformes; palm cockatoo (*Probosciger aterrimus*) and Yellow-naped Amazon (*Amazona auropalliata*) belonging to Psittaciformes; and barn swallow (*Hirundo rustica*) belonging to Passeriformes. RNase-treated whole genomic DNA (500 ng) was denatured with 0.4 N NaOH for 10 min and transferred to nylon membranes (Roche Diagnostics) using BIO-DOT SF blotting equipment (Bio-Rad, Hercules, CA, USA). DNA probes of repetitive sequences were labeled with DIG-11-dUTP using the PCR DIG Labeling Mix (Roche Diagnostics) and hybridized with the membranes at 45˚C in DIG Easy Hyb solution (Roche Diagnostics). Chemiluminescent signals were detected as described in Southern blot hybridization.

Results

Giemsa-stained and C-banded karyotypes

Giemsa-stained and C-banded karyotypes of *A. platyrhynchos* (2n = 80) were described in previous studies [36–41]. The chromosome number of *A. fabalis* was 2n = 80, consisting of two pairs of large submetacentric chromosomes, one pair each of large subtelocentric chromosomes, medium-sized metacentric chromosomes, and medium-sized subtelocentric chromosomes; four pairs of small acrocentric chromosomes; 30 pairs of indistinguishable microchromosomes; and the submetacentric Z and subtelocentric W sex chromosomes (Fig 1A). The karyotype of *C. cygnus* (2n = 80) consisted of one pair each of large submetacentric chromosomes, large metacentric chromosomes, and large acrocentric chromosomes; two pairs of medium-sized acrocentric chromosomes; four pairs of small acrocentric chromosomes; 30 pairs of indistinguishable microchromosomes; and the acrocentric Z and small acrocentric W sex chromosomes, which were similar to those of *A. platyrhynchos* (Fig 1B). The morphology of the chromosomes 4, Z, and W chromosomes was acrocentric in *A. platyrhynchos* and *C. cygnus*, whereas these three chromosomes were metacentric, submetacentric, and subtelocentric, respectively, in *A. fabalis* (Fig 1). Large C-positive heterochromatin blocks were observed in the centromeric regions of most autosomes and the Z chromosome and in whole regions of the W chromosome in these two species as well as *A. platyrhynchos* (Fig 2) [38,39,41]. In *A. platyrhynchos* and *A. fabalis*, C-positive heterochromatin on the Z chromosome was observed in only the centromeric region (Fig 2A), whereas ladder C-positive heterochromatin blocks were observed throughout the region of the Z chromosome in *C. cygnus* (Fig 2B).

Chromosome homologies with chicken chromosomes

Each of macrochromosome probes (GGA1–3, 5–9, and Z), except for GGA4, painted a single pair of chromosomes of *A. fabalis* and *C. cygnus* (Fig 3). The GGA4 probe hybridized to chromosome 4 and additionally to one pair of microchromosomes (Fig 3B and 3E). These results

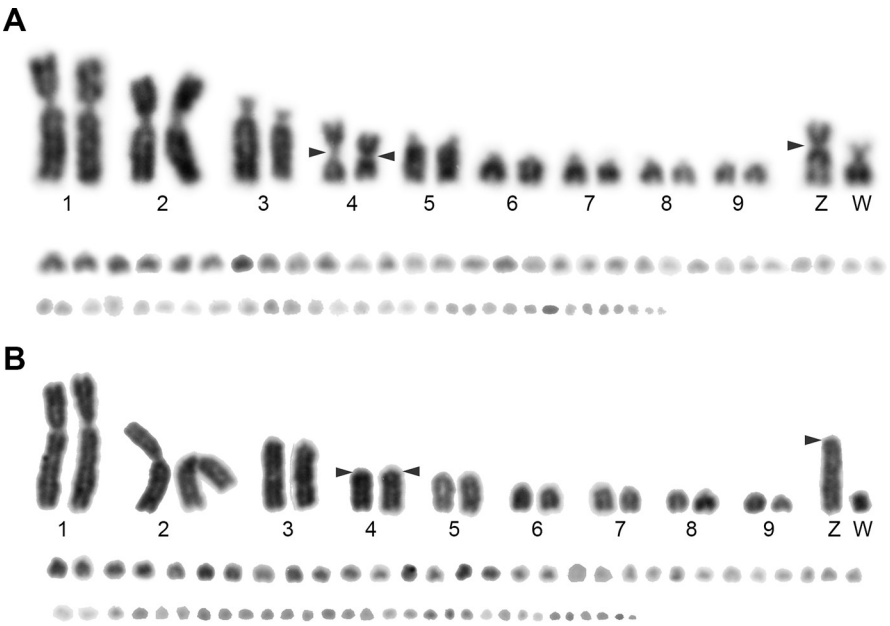


Fig 1. Giemsa-stained karyotypes of female bean goose (*A. fabalis*) (A) and whooper swan (*C. cygnus*) (B). Arrowheads indicate the positions of the centromeres in chromosome 4 and the Z and W sex chromosomes.

https://doi.org/10.1371/journal.pone.0214028.g001

of *A. fabalis* and *C. cygnus* were consistent with those of *A. platyrhynchos* in our previous study [40]. The microchromosome-specific paint pool (GGAmicro) hybridized with approximately half of microchromosomes, and no hybridization signals were detected on macrochromosomes (Fig 3F).

Repetitive sequence families and their nucleotide sequences

Prominent DNA bands of repetitive sequences were revealed by agarose gel electrophoresis of *A. platyrhynchos* (APL) genomic DNA digested with *Hae*III. A DNA band of approximately 190 bp was isolated from the gel, and then 16 clones inserted into plasmid vectors were obtained. For *A. fabalis* (AFA), 16 and 8 clones were obtained from two bands of approximately 100 and 200 bp by *Hinf*I digestion, respectively. For *C. cygnus*, 50 clones were obtained

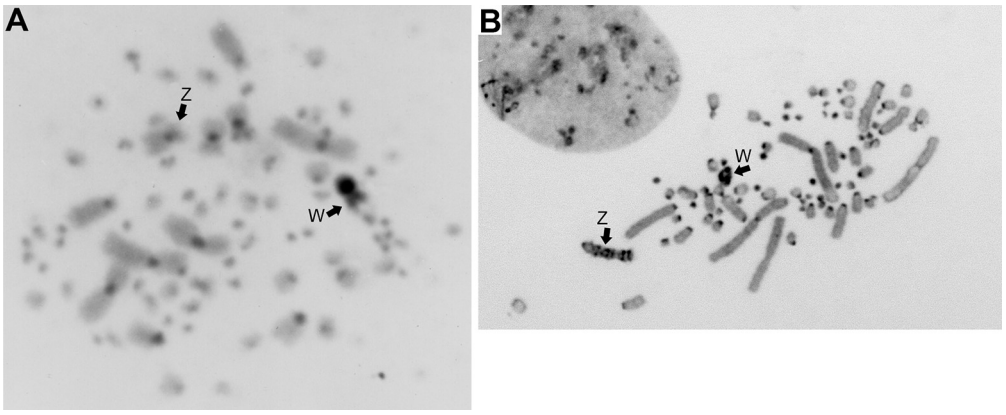


Fig 2. C-banded metaphase spreads of *A. fabalis* (A) and *C. cygnus* (B) females. Arrows indicate the Z and W sex chromosomes.

https://doi.org/10.1371/journal.pone.0214028.g002

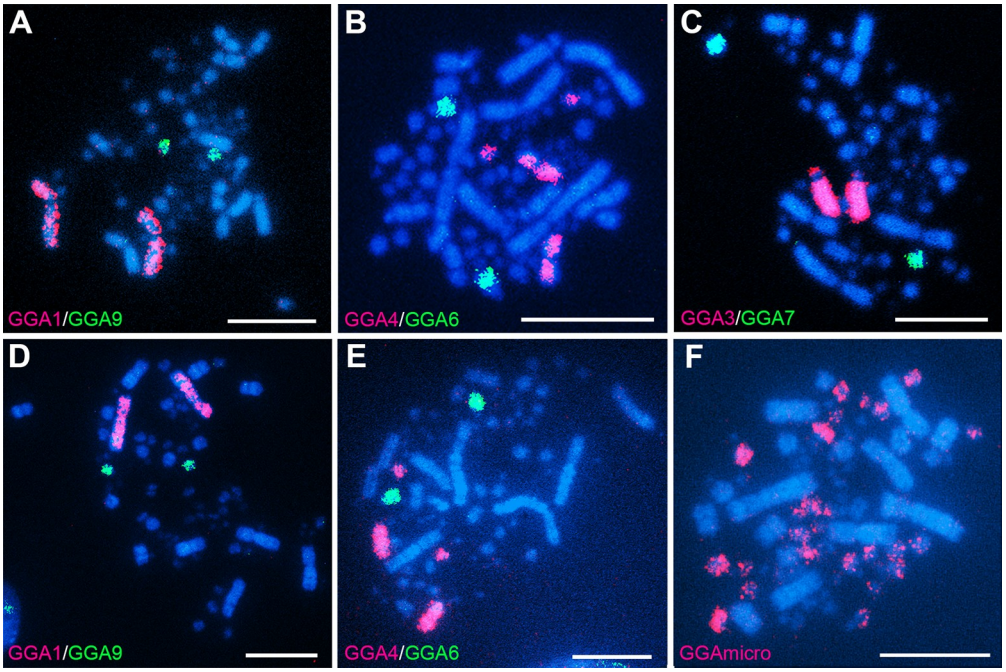


Fig 3. Chromosome painting with chicken (*G. gallus*, GGA) chromosome-specific DNA probes to Hoechst-stained chromosome spreads of *A. fabalis* (A–C) and *C. cygnus* females (D–E). DIG-labeled GGA1 (red) and biotin-labeled GGA9 (green) hybridized to chromosomes 1 and 9, respectively (A, D). DIG-labeled GGA4 (red) hybridized to chromosome 4 and a pair of microchromosomes, and biotin-labeled GGA6 (green) hybridized to chromosome 6 (B, E). DIG-labeled GGA3 (red) hybridized to chromosome 3, and biotin-labeled GGA7 (green) hybridized to chromosome 7 (C). DIG-labeled painting probe of 20 microchromosome pairs hybridized to approximately half of the microchromosomes (F). Scale bars represent 10 μm.

https://doi.org/10.1371/journal.pone.0214028.g003

from a band of approximately 300 bp by *Apa*I digestion. Nucleotide sequences were determined for all clones.

Consequently, 26 DNA fragments categorized into four families of repetitive sequences were isolated from these three species and deposited in DDBJ (http://www.ddbj.nig.ac.jp/), viz., the APL-*Hae*III family from *A. platyrhynchos*, the AFA-*Hinf*I-S and AFA-*Hinf*I-L familiy from *A. fabalis*, and the CCY-*Apa*I family from *C. cygnus* (Table 1). The sizes, GC content, and

Table 1. Repetitive sequence families isolated from *A. platyrhynchos*, *A. fabalis*, and *C. cygnus*, and their lengths, sequence identities between fragments, and GC content.

Repetitive sequence family	No. of clones	Length of concensus sequence (bp)[a]	Average sequence identity between fragments (%)[b]	Average GC content (%)[c]	Accession number
A. platyrhynchos					
APL-*Hae*III	5/16	190 (161 – 190)	84.7 (78.2 – 89.4)	51.6 (48.7 – 53.2)	LC416791 –LC416795
A. fabalis					
AFA-*Hinf*I-S	4/16	101 (101)	94.1 (91.0 – 98.1)	60.2 (58.4 – 61.4)	LC416770 –LC416773
AFA-*Hinf*I-L	2/8	192 (192)	98.4 (98.4)	51.6 (51.0 – 52.1)	LC416774, LC416775
C. cygnus					
CCY-*Apa*I	15/50	290 (286 – 294)	97.6 (94.9 – 100.0)	54.0 (53.4 – 54.6)	LC416776 –LC416790

[a]Range of fragment lengths in parentheses

[b]Range of sequence identities in parentheses

[c]Range of GC content in parentheses

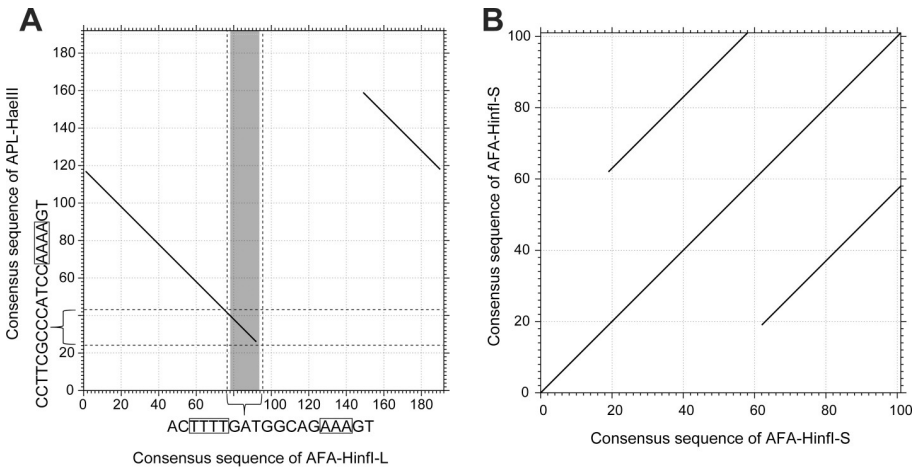
https://doi.org/10.1371/journal.pone.0214028.t001

nucleotide sequence identities between the fragments and within the same sequence family are summarized in Table 1. The length of the consensus sequence of five APL-*Hae*III fragments was 190 bp (S1 Fig). A deletion of 29 nucleotides was found at the end of two 161-bp fragments (APL-*Hae*III-08 and APL-*Hae*III-11). Nucleotide sequence identities, which were calculated by eliminating insertions and deletions but including one nucleotide gap, ranged from 78.2% to 89.4% (84.7% on an average), and the GC content was relatively high (51.6% on an average). All four AFA-*Hinf*I-S fragments were of the same length of 101 bp, and two AFA-*Hinf*I-L fragments were of 192 bp (S2 Fig). The identities of nucleotide sequences between the fragments ranged from 91.0% to 98.1% (94.1% on an average) for the AFA-*Hinf*I-S sequence family and 98.4% for the AFA-*Hinf*I-L sequence family. AFA-*Hinf*I-S showed much higher GC content (60.2%) compared to that of AFA-*Hinf*I-L (51.6%). The length of 15 CCY-*Apa*I fragments ranged from 286 to 294 bp with 97.6% sequence identify and 54.0% GC content on an average (S3 Fig).

The nucleotide sequence of AFA-*Hinf*I-L showed homology with APL-*Hae*III. Two partial sequences at positions 26–117 and 118–159 in the consensus APL-*Hae*III sequence showed 84.9% and 76.5% identities with those at positions 1–92 and 149–190 in the consensus AFA-*Hinf*I-L, respectively (Fig 4A). The AFA-*Hinf*I-S sequence contained two 42–43-bp internal repeat units (Fig 4B). Homology searches of the four repetitive sequence families were performed using the NCBI non-redundant sequence database (http://blast.ncbi.nlm.nih.gov) and Repbase (http://www.girinst.org/repbase/). The consensus sequences of APL-*Hae*III and the partial sequence of AFA-*Hinf*I-L at positions 60–192 showed similarity with the RBMII sequences isolated from several Anseriformes species, such as *A. platyrhynchos* (X61424) (88.0% and 81.4% identity, respectively) and wood duck (*Aix sponsa*, ASP) (X61410) (88.1% and 79.4% identity, respectively) (Fig 4C and S1 Table), indicating that APL-*Hae*III and AFA-*Hinf*I-L shared the same origin as that the of RBMII sequence, which is a major tandem repetitive sequence in the Anseriformes species [27,28]. Nucleotide sequence homology was observed for 20-bp partial sequences at positions 24–43 in APL-*Hae*III and at positions 76–95 in AFA-*Hinf*I-L (26.0%–47.6% and 38.0%–59.0%, respectively) and with the CNM repeat family sequences in Galliformes, chicken CNM repeat [14], turkey TM repeat [16], Blue-breasted Quail CCH-S [17], and chukar partridge ACH-*Sau3A*I (AB872160) [18] (Fig 4D). The highest identity (59.0%) was found between the partial sequence of AFA-*Hinf*I-L and the 20-bp consensus sequence of internal repeat unit of ACH-*Sau3A*I. The T-rich and A-rich motifs of the CNM repeat family were conserved in AFA-*Hinf*I-L but not found in APL-*Hae*III (Fig 4A and 4D). The partial sequence at position 179–223 of CCY-*Apa*I also showed homology with the CNM repeat family sequences of Galliformes (35.0%–55.0%; Fig 4E). The highest identity (55.0%) was observed for the TM repeat of turkey [14]; however, the T-rich and A-rich motifs were not detected in the 45-bp partial sequence of CCY-*Apa*I.

Chromosomal distribution

In *A. platyrhynchos*, the APL-*Hae*III sequence family showed intense hybridization signals in the centromeric regions of all macrochromosomes and microchromosomes, except for the chromosomes 1 and Z, i.e., no signals were observed on the chromosomes 1 and Z (Fig 5A). In *C. cygnus*, CCY-*Apa*I was localized to the centromeric regions of approximately 10 microchromosomes (Fig 5B). AFA-*Hinf*I-S and AFA-*Hinf*I-L were localized to the centromeric/pericentromeric regions of all chromosomes and almost all microchromosomes, respectively, in *A. fabalis* (Fig 5C–5E). The fluorescent signals of AFA-*Hinf*I-L and AFA-*Hinf*I-S hardly overlapped, suggesting that the two sequence families were located separately in centromeric heterochromatin on the same microchromosomes (Fig 5E).



C

```
                                    10         20         30         40         50
APL-HaeIII                          ---------- -----CCTCA GAAAACGCAT TGCAAGTCCC CTTCGCCCAT
AFA-HinfI-L position 60-192         TGAACCCGAG TTAGA..CG. A.....A.GK ACGGT.C.A. T...TG....
RBMII (APL, X61424)                 --GACTGAAA ACACGG.C.. C.......T. .C.G..C... A......A..
RBMII (ASP, X61410)                 ---------- ----C..A.. .....AT.G. .T....C... T.....T...

                                    60         70         80         90         100
APL-HaeIII                          CCAAAAGTGC CCCCAAAGGA AGTCTTTGGG TACAACTGAA CTCCCCAGGG
AFA-HinfI-L position 60-192         .A........ T.T..GC... ..C..... ....C..... .....TG...
RBMII (APL, X61424)                 .......... .......... ..G..... ....T..... ..........
RBMII (ASP, X61410)                 ...G...... .......... ..G..... .....G.... ..........

                                    110        120        130        140        150
APL-HaeIII                          GAGCCTTTAC GCAGCCCCAC ATGCAACTGG ACCCTCCGGG TTCACTCTTT
AFA-HinfI-L position 60-192         .......... ...A...... .......A.. ..T------- ----------
RBMII (APL, X61424)                 .......... .......... .......... ........T. ...G......
RBMII (ASP, X61410)                 .........T .......... .......... ........TA ...G......

                                    160        170        180        190        200
APL-HaeIII                          CACAGCAAGA CTAAATATAG CCCTCGGCGC AATGGGAGAA GACTGAAAAA GCGGG  190
AFA-HinfI-L position 60-192         ---------- ---------- ---------- ---------- ---------- -----  133
RBMII (APL, X61424)                 -.......CC .........A ....A.C.AG C....C.... ---------- -----  187
RBMII (ASP, X61410)                 ..A.....CC .........A ....G.---- ---------- ---------- -----  162
```

D

```
                                    10         20         30         40
AFA-HinfI-L position 76-95          --ACTTTTTG ATGGCAGAAA GT-------- ---------- ---  20
APL-HaeIII position 24-43           --CCC..CGC CCAT.CA... ..-------- ---------- ---  20
CNM                                 --G.G....C TCTT.GC... TCCCCCATTT AAGCGAAAAT CAC  41
TM                                  -GGTA...GT GG.AG.A... .GGGTATTTT CGTGTAAAAA CA-  41
CCH-S                               GR.T.....- --CC.SA... TAA------- ---------- ---  20
ACH-Sau3AI internal unit            --GTA....- -G..AGA... TGA------- ---------- ---  20
```

E

```
                                    10         20         30         40
CCY-ApaI position 179-223           TGTAGTTGTG GGAATCGTGT AGTTACTTAT CGTCGTTCAA AACCT  45
CNM                                 GCGTT..C.C TTCGCAAATC CCCC.T..-- --AA.CGA...T.AC  41
TM                                  G...T..... ...GAAAAAG G.G..T..-- --.....GT.. ..A.A  41
CCH-S                               -.R.T..T.T CCCSAAAATA .--------- ---------- -----  20
ACH-Sau3AI internal unit            -..T..TG.. ..GAAAATG .--------- ---------- -----  20
```

Fig 4. Comparison of the nucleotide sequences of APL-*Hae*III, AFA-*Hinf*I-S, AFA-*Hinf*I-L, and CCY-*Apa*I sequences with their homologous sequences. Dot matrix analysis between the consensus sequences of 190-bp APL-*Hae*III and 192-bp AFA-*Hinf*I-L (A). The gray region on AFA-*Hinf*I-L shows the 12–17-bp T-rich and A-rich motif that is conserved in the CNM repeat sequence family of Galliformes [14,16,18], and squares on the sequences indicate the A_{3-5} or T_{3-5} internal repeats in this motif. Dot matrix analysis of the 101-bp AFA-*Hinf*I-S consensus sequence (B). Dot matrix analysis was performed in the condition of the scoring matrix, 200PAM/K = 2 and threshold score = 22 (E = 0.00805). Alignment of the APL-*Hae*III consensus sequence and partial sequence at nucleotide position 60–192 of the AFA-*Hinf*I-L consensus sequence with the RBMII sequences of *A*. *platyrhynchos* (APL) (X61424) and *Aix sponsa* (ASP) (X61410) (C). Alignment of the partial sequences at nucleotide position 76–95 of the AFA-*Hinf*I-L consensus sequence and at position 24–43 of the APL-*Hae*III consensus sequence (D) and the partial sequence at positions 179–223 of the CCY-*Apa*I consensus sequence (E) with the four CNM sequence homologs in Galliformes, viz., CNM repeat in chicken [14], TM repeat in turkey (*M*. *gallopavo*) [16], CCH-S in Blue-breasted Quail (*C*. *chinensis*) [17], and ACH-*Sau3A*I in chukar partridge (*A*. *chukar*) [18]. Squares indicate the A_{3-5} or T_{3-5} internal repeats in the 12–17-bp T-rich and A-rich motifs conserved in the CNM repeat sequence family of Galliformes [14,16,18].

https://doi.org/10.1371/journal.pone.0214028.g004

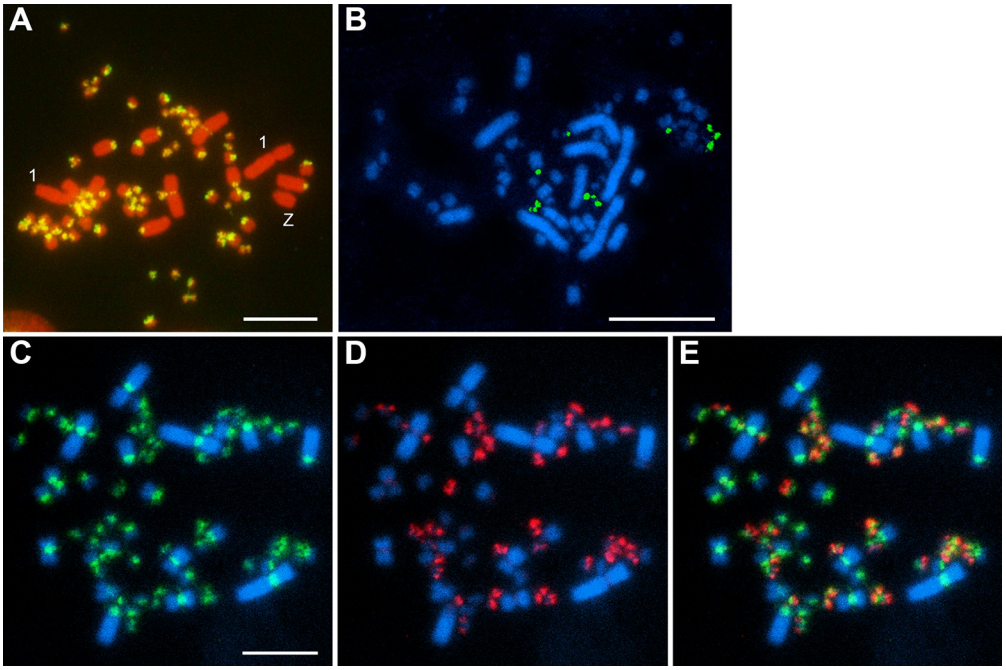


Fig 5. Chromosomal distribution of four families of repetitive sequences on metaphase spreads. Chromosomal distribution of the biotin-labeled APL-*Hae*III-04 fragment on the PI-stained metaphase chromosome spread of *A. platyrhynchos* female (A). Hybridization pattern of the biotin-labeled CCY-*Apa*I-05 fragment to the Hoechst-stained metaphase spread of *C. cygnus* female (B). Hybridization patterns of the biotin-labeled AFA-HinfI-S03 fragment (green) (C) and DIG-labeled AFA-*Hinf*I-L04 fragment (red) (D) to the Hoechst-stained metaphase spread of *A. fabalis* female and their merged image (E). Scale bars represent 10 μm.

https://doi.org/10.1371/journal.pone.0214028.g005

Organization in the genome

Southern blot hybridization was performed to examine the genomic organization of four families of repetitive sequences. APL-*Hae*III showed polymeric ladder signals of tandem repeats of the 190-bp monomer unit in *Hae*III, *Msp*I, and *Taq*I digests (Fig 6A); the monomer unit was present in the highest abundance, with decreasing copy numbers of each higher order. By contrast, the *BamH*I digest produced higher intensity of hybridization bands with increasing size of multimers. This result indicated that the *BamH*I cleavage site was not highly conserved in the tandem array of the 190-bp monomer unit. The restriction site for both *Hpa*II and *Msp*I is CCGG, and *Hpa*II does not cleave when the second cytosine is methylated, whereas *Msp*I does. In contrast to the *Msp*I digest, the intensity of ladder bands increased from low to high molecular weight in *Hpa*II digests, indicating that the APL-*Hae*III sequence was highly methylated.

Hybridization of the AFA-*Hinf*I-S sequence showed only a 101-bp monomeric band in the *Hinf*I digest, indicating that the *Hinf*I site was highly conserved with regard to the tandem array of the sequence (Fig 6B). However, no restriction sites of *BamH*I, *Hae*III, *Rsa*I, and *Msp*I were found in the AFA-*Hinf*I-S monomer unit (S2A Fig), thus, resulting in smear-like bands in the digests of these enzymes. This sequence was weakly methylated because the intensities of ladder bands at lower molecular weight were slightly higher in the *Hpa*II digest than those in the *Msp*I digest.

In hybridization with AFA-*Hinf*I-L, ladder bands at lower molecular weight were observed in the *Hinf*I, *Alu*I, *Rsa*I, and *Msp*I digests, whose restriction sites were all contained in the 192-bp monomer unit (Fig 6C and S2 Fig). Of these sites, the *Hinf*I and *Alu*I sites were

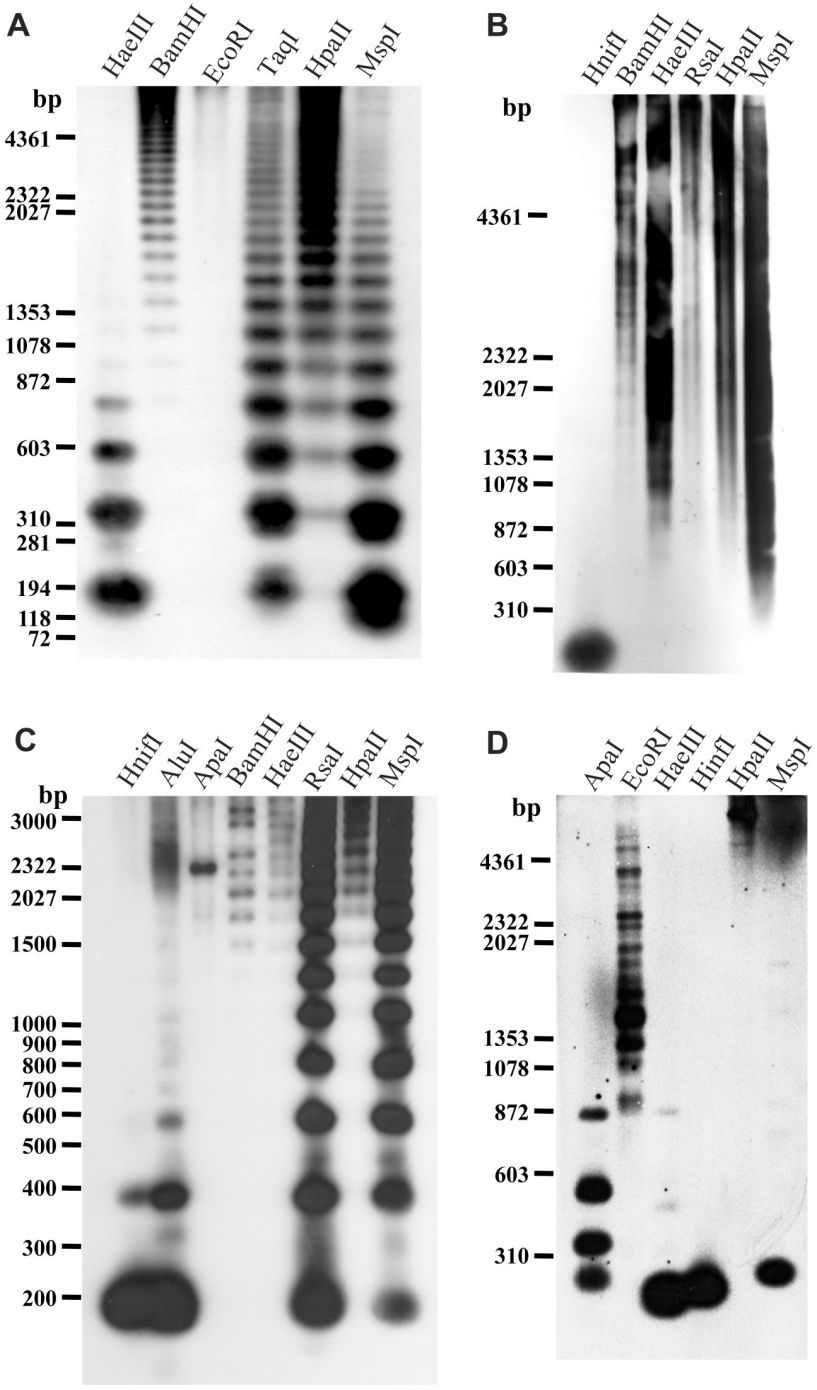


Fig 6. Southern blot hybridization patterns of four repetitive sequence families. Southern blot hybridization of *A. platyrhynchos* genomic DNA probed with the APL-*Hae*III-04 fragment (A). Southern blot hybridization of *A. fabalis* genomic DNA probed with the AFA-*Hinf*I-S03 (B) and AFA-*Hinf*I-L04 (C) fragments. Southern blot hybridization of *C. cygnus* genomic DNA probed with the CCY-*Apa*I-05 fragment (D). A mixture of λ DNA–*Hind*III and ϕX174 DNA–*Hae*III was used as a molecular size marker for (A, B, D), and a mixture of λ DNA–*Hind*III digest and 100-bp ladder digest was used for (C).

https://doi.org/10.1371/journal.pone.0214028.g006

particularly highly conserved. In contrast to the *Msp*I digest, there were no hybridization signals at a lower molecule weight in the *Hpa*II digest, indicating that this sequence was hypermethylated.

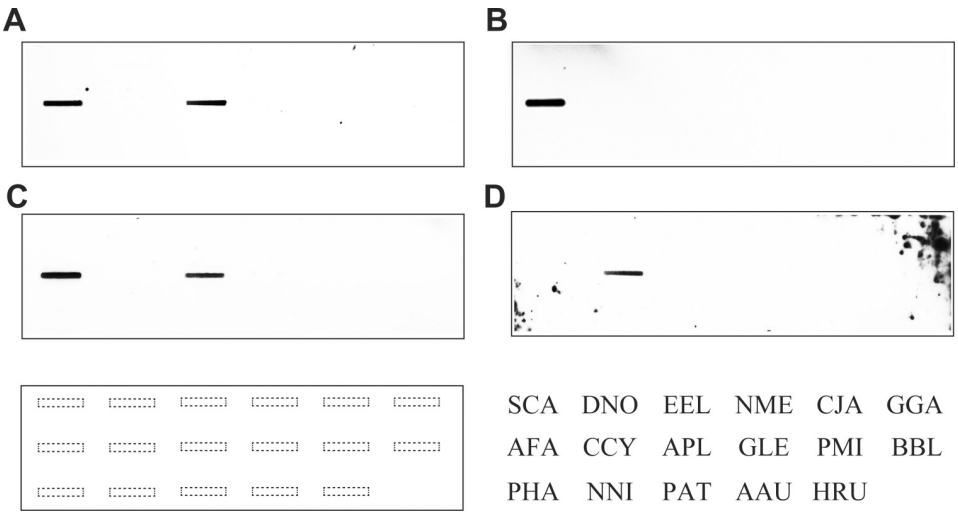


Fig 7. Slot blot hybridization of four repetitive sequences to genomic DNA of 17 species. The fragments used for this experiment are as follows: APL-*Hae*III-04 (A), AFA-*Hinf*I-S03 (B), AFA-*Hinf*I-L04 (C), and CCY-*Apa*I-05 (D). Genomic DNA used for this experiment was obtained from the following avian species of 10 orders: (1) Struthioniformes, SCA (*S. camelus*) and DNO (*D. novaehollandiae*); (2) Tinamiformes, EEL (*E. elegans*); (3) Galliformes, NME (*N. meleagris*), CJA (*C. japonica*), and GGA (*G. gallus*); (4) Anseriformes, AFA (*A. fabalis*), CCY (*C. cygnus*), and APL (*A. platyrhynchos*); (5) Gruiformes, GLE (*G. leucogeranus*); (6) Pelecaniformes, PMI (*P. minor*); (7) Strigiformes, BBL (*B. blakistoni*); (8) Accipitriformes, PHA (*P. haliaetus*) and NNI (*N. nipalensis orientalis*); (9) Psittaciformes, PAT (*P. aterrimus*) and AAU (*A. autumnalis*); and (10) Passeriformes, HRU (*H. rustica*). The locations of genomic DNA on membranes used for slot blot hybridization are represented at the bottom of the figures.

https://doi.org/10.1371/journal.pone.0214028.g007

Hybridization of the 290-bp CCY-*Apa*I fragment revealed that the *Apa*I, *Hinf*I, and *Msp*I sites were conserved in the tandem array of this repeat sequence (Fig 6D and S3 Fig). The hybridization bands positioned at approximately 580 and 870 bp corresponded to the dimeric and trimeric bands in the *Apa*I digest, respectively. The intermediate band between the 290- and 580-bp bands may have been derived from the internal restriction sites; however, this site was not found for this fragment. The <290-bp hybridization band in the *Hae*III digest was derived from multiple internal *Hae*III sites contained in the sequence (S3 Fig). Only one 290-bp band was observed in the *Msp*I digest; however, no bands were found at lower molecular weight in the *Hpa*II digest, indicating that this repetitive sequence was hypermethylated.

Nucleotide sequence conservation

Slot blot hybridization probed with four families of repetitive sequences was performed using genomic DNA from 17 avian species of 10 orders (Fig 7). Hybridization signals of APL-*Hae*III and AFA-*Hinf*I-L were detected for two species of Anseriformes, viz., *A. platyrhynchos* and *A. fabalis* (Fig 7A and 7C). The hybridization signals of AFA-*Hinf*I-S and CCY-*Apa*I were detected only in *A. fabalis* and *C. cygnus*, respectively (Fig 7B and 7D). In order to examine the chromosomal distribution of AFA-*Hinf*I-L in *A. platyrhynchos* and APL-*Hae*III in *A. fabalis*, we performed cross-species FISH mapping of AFA-*Hinf*I-L to *A. platyrhynchos* chromosomes and of APL-*Hae*III to *A. fabalis* chromosomes. However, no hybridization signals were found.

Discussion

Giemsa-stained karyotype analysis revealed that the chromosome number was 2n = 80 for both *A. fabalis* and *C. cygnus* as *A. platyrhynchos* [36–41]. The size and morphology of most macrochromosomes were similar between these two species; however, the chromosomes 4, Z,

and W of *A. platyrhynchos* and *C. cygnus* were morphologically different from those of *A. fabalis*. C-positive heterochromatin was observed in the centromeric regions of almost all autosomes and whole regions of the W chromosomes in *A. fabalis* and *C. cygnus*, which was consistent with that of *A. platyrhynchos*, as reported previously [39,41]. However, amplification of the interstitial heterochromatin blocks in the Z chromosome, as shown by the C-positive ladder signals, occurred specifically in the lineage of *C. cygnus*.

The chicken chromosome paints 1–9 and Z each hybridized to a single pair of chromosomes in *A. fabalis* and *C. cygnus*, except for GGA4 that hybridized to a single pair of macrochromosomes (chromosome 4) and a single pair of microchromosomes. The same result has been reported in four other Anseriformes species [swan goose (*Anser cygnoides*), Muscovy duck (*Cairina moschata*), *A. platyrhynchos*, and coscoroba swan (*Coscoroba coscoroba*)]; several palaeognathous birds, including ostrich and tinamous; and several Galliformes species, including pheasants, turkey, and New World quail [18,26,33,42–46], indicating that all macrochromosomal structures are highly conserved in Anseriformes, Galliformes, and Palaeognathae. However, GGA4 paint hybridized to a single pair of macrochromosomes (chromosome 4) in one Anseriformes species, greylag goose (*A. anser*) [43], and chicken and several phasianid species, viz., chukar partridge (*A. chukar*), Japanese quail (*C. japonica*), Blue-breasted Quail (*C. chinensis*), Chinese bamboo-partridge (*Bambusicola thoracica*), and common peafowl (*Pavo cristatus*) [18,4,45], indicating that the fusion of a microchromosome to the ancestral chromosome 4 occurred independently in each lineage of Galliformes and Anseriformes after they split from the common ancestor of Galloanserae.

In this study, we isolated four families of centromere-specific repetitive sequences, APL-*Hae*III, AFA-*Hinf*-S, AFA-*Hinf*-L, and CCY-*Apa*I, from three Anseriformes species. Among these four repetitive sequences, AFA-*Hinf*-S and CCY-*Apa*I were species-specific, suggesting that these repetitive sequences occurred independently in each species. APL-*Hae*III and AFA-*Hinf*-L were conserved in *A. platyrhynchos* and *A. fabalis*, both of which showed sequence similarities to the RBMII sequence that is conserved in at least 22 Anseriformes species [27,28]. However, these two sequence families showed different chromosomal distribution, and no hybridization signals were detected on chromosomes by cross-species FISH mapping. These results indicate that APL-*Hae*III and AFA-*Hinf*-L were derived from the same origin as the RBMII sequence in Anseriformes; however, they differentiated to the extent that they hybridized interspecifically by slot blot hybridization but not by FISH.

APL-*Hae*III and AFA-*Hinf*I-S were distributed in almost all chromosomes. However, AFA-*Hinf*I-L and CCY-*Apa*I were microchromosome-specific centromeric repeats that were firstly identified in Anseriformes. AFA-*Hinf*I-L was predominantly localized to all microchromosomes and CCY-*Apa*I to some microchromosomes. All the four repetitive sequences were Anseriformes-specific sequences; however, the partial sequences of APL-*Hae*III, AFA-*Hinf*I-L, and CCY-*Apa*I showed homology with the CNM family sequences, including chicken CNM, turkey TM, Blue-breasted quail CCH-S, and chukar partridge ACH-*Sau3A*I sequences, which are localized primarily to microchromosomes in Galliformes [14,16–18]. The T-rich and A-rich motifs conserved in the CNM family sequences were observed in AFA-*Hinf*I-L but not in APL-*Hae*III and CCY-*Apa*I. Consequently, the CNM family sequences of Galliformes and APL-*Hae*III, AFA-*Hinf*I-L, and CCY-*Apa*I were concluded to be partially derived from the same ancestral sequence and diverged independently in each lineage. Microchromosome-specific repetitive sequences have been isolated from Falconiformes, Galliformes, Piciformes, and Struthioniformes in Aves and from the Chinese soft-shell turtle [14–22], suggesting that chromosome size-correlated genome compartmentalization between macrochromosomes and microchromosomes is common in birds and turtle. AFA-*Hinf*I-L and CCY-*Apa*I of Anseriformes were also homogenized in a chromosome size-correlated manner. However, the

homogenization between macrochromosomes and microchromosomes also occurred in APL-*Hae*III and AFA-*Hinf*I-S, as observed in ACH-*Sau3A*I, CVI-*Hae*III, CCA-*BamH*I, and CSQ-*BamH*I in Galliformes [18].

Not much is known about how chromosome size-dependent distribution of the centromeric repetitive sequences evolved in avian genomes. One possible explanation is that such biased distribution of the centromeric repetitive sequences is caused by chromosome positioning in the nuclei. In the interphase nuclei of chicken, turkey, and Japanese quail, microchromosomes are located predominantly in the nuclear interior, and macrochromosomes are primarily located in peripheral parts of the nuclei [32,47–50]. Owing to the spatially different disposition, physical interaction of chromatin between macrochromosomes and microchromosomes may be restricted, resulting in the restriction of homogenization of the centromeric repetitive sequences between different sized chromosomes. However, the molecular basis that is responsible for the spatial structure of the centromeres of macrochromosomes and microchromosomes in the interphase nuclei are not fully understood.

The presence of microchromosomes is a common feature of Aves, Reptilia (sauropsids), except for Crocodilia [51], and also in some amphibian species [52]. Comparative genome and chromosome analyses for amphibians, reptiles, birds, and mammals suggest that ancestral tetrapods and amniotes may have retained many microchromosomes whose linkages are highly conserved in chicken microchromosomes [53–58]. Comparison of the GC content in exonic third codon positions (GC_3) of genes between macrochromosomes and microchromosomes in several reptilian species, Chinese soft-shell turtle, Japanese four-striped rat snake (*Elaphe quadrivirgata*), central bearded dragon (*Pogona vitticeps*), and green anole (*Anolis carolinensis*) demonstrated that the genes on microchromosomes tend to have higher GC_3 than those on macrochromosomes, as shown in chicken, suggesting that chromosome size-dependent GC heterogeneity was acquired in the common ancestors of sauropsids [59–62]. Further identification of the microchromosome-specific centromeric repetitive sequences from avian and reptilian species may help clarify the relationship between the genomic organization of microchromosomes and chromosome size-correlated compartmentalization between macrochromosomes and microchromosomes in tetrapods.

Supporting information

S1 Fig. Nucleotide sequences of five APL-*Hae*III fragments isolated from the *Hae*III-digested genomic DNA of *A. platyrhynchos*. Internal restriction sites of endonucleases are represented by the following underlining: *Hae*III, dots and dashes; and *Msp*I, wave. Dots indicate the same nucleotides as those of the consensus sequence shown at the top, and hyphens indicate gaps.
(TIF)

S2 Fig. Alignments of nucleotide sequences of four AFA-*Hinf*I-S fragments (A) and two AFA-*Hinf*I-L fragments (B) isolated from the *Hinf*I-digested genomic DNA of *A. fabalis*. Internal restriction sites of endonucleases are represented by the following underlining: *Alu*I, conventional; *Hinf*I, double; *Msp*I, wave; and *Rsa*I, dots. Dots indicate the same nucleotides as those of the consensus sequence shown at the top, and hyphens indicate gaps.
(TIF)

S3 Fig. Alignment of nucleotide sequences of 15 CCY-*Apa*I fragments isolated from the *Apa*I-digested genomic DNA of *C. cygnus*. Internal restriction sites of endonucleases are represented by the following underlining: *Apa*I, bold; *Hae*III, dots and dashes; *Hinf*I, double; and *Msp*I, wave. Dots indicate the same nucleotides as those of the consensus sequence shown at

the top, and hyphens indicate gaps.
(TIF)

S1 Table. Nucleotide sequences exhibiting homologies with three repetitive sequence families, APL-*Hae*III, AFA-*Hinf*I-L, and AFA-*Hinf*I-S, except for the CNM family sequence.
(XLSX)

Acknowledgments

We thank the Asahiyama Zoo, Asahikawa and Kushiro Zoo, Kushiro, Hokkaido, Japan for providing with the specimens of *A. fabalis* and *C. cygnus*, respectively, and Dr. Darren K Griffin, University of Kent, England, and Dr. Felix A. Habermann, Ludwig-Maximilians-University Munich, Germany, for providing with chicken chromosome-specific painting probes.

Author Contributions

Conceptualization: Yoshinobu Uno, Chizuko Nishida, Yoichi Matsuda.

Data curation: Yoshinobu Uno, Chizuko Nishida, Yoichi Matsuda.

Funding acquisition: Yoichi Matsuda.

Investigation: Yoshinobu Uno, Chizuko Nishida, Ayano Hata, Satoshi Ishishita.

Project administration: Yoshinobu Uno, Chizuko Nishida.

Resources: Yoshinobu Uno, Chizuko Nishida, Yoichi Matsuda.

Supervision: Yoshinobu Uno, Yoichi Matsuda.

Validation: Yoshinobu Uno, Chizuko Nishida.

Visualization: Yoshinobu Uno, Yoichi Matsuda.

Writing – original draft: Yoshinobu Uno, Chizuko Nishida, Yoichi Matsuda.

Writing – review & editing: Yoshinobu Uno, Chizuko Nishida, Ayano Hata, Satoshi Ishishita, Yoichi Matsuda.

References

1. Singer MF. Highly repeated sequences in mammalian genomes. Int Rev Cytol. 1982; 76: 67–112. PMID: 6749748
2. Kalitis P, Choo KHA. Centromere DNA of higher eukaryotes. In: Choo KHA, editor. The centromere. Oxford: Oxford University Press; 1997. pp. 97–142.
3. Burrack LS, Berman J. Flexibility of centromere and kinetochore structures. Trends Genet. 2012; 28: 204–212. https://doi.org/10.1016/j.tig.2012.02.003 PMID: 22445183
4. Matsubara K, O'Meally D, Azad B, Georges A, Sarre SD, Graves JAM, et al. Amplification of microsatellite repeat motifs is associated with the evolutionary differentiation and heterochromatinization of sex chromosomes in Sauropsida. Chromosoma. 2016; 125: 111–123. https://doi.org/10.1007/s00412-015-0531-z PMID: 26194100
5. Eichler EE. Repetitive conundrums of centromere structure and function. Hum Mol Genet. 1999; 8: 151–155. PMID: 9931322
6. Henikoff S, Ahmad K, Malik HS. The Centromere Paradox: Stable Inheritance with Rapidly Evolving DNA. Science. 2001; 293: 1098–1102. https://doi.org/10.1126/science.1062939 PMID: 11498581
7. Plohl M, Luchetti A, Meštrović N, Mantovani B. Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. Gene. 2008; 409: 72–82. https://doi.org/10.1016/j.gene.2007.11.013 PMID: 18182173
8. Takagi N, Sasaki M. A phylogenetic study of bird karyotypes. Chromosoma. 1974; 46: 91–120. PMID: 4134896

9. Belterman RHR, Boer LEMD. A karyological study of 55 species of birds, including karyotypes of 39 species new to cytology. Genetica. 1984; 65: 39–82. https://doi.org/10.1007/BF00056765
10. Christidis L. Animal cytogenetics 4: Chordata 3 B: Aves. Stuttgart: Gebrüder Borntraeger; 1990.
11. International Chicken Genome Sequencing Consortium (ICGSC). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature. 2004; 432: 695–716. https://doi.org/10.1038/nature03154 PMID: 15592404
12. Axelsson E, Webster MT, Smith NG, Burt DW, Ellegren H. Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. Genome Res. 2005; 15: 120–125. https://doi.org/10.1101/gr.3021305 PMID: 15590944
13. Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, et al. Comparative genomics reveals insights into avian genome evolution and adaptation. Science. 2014; 346: 1311–1320. https://doi.org/10.1126/science.1251385 PMID: 25504712
14. Matzke MA, Varga F, Berger H, Schernthaner J, Schweizer D, Mayr B, et al. A 41–42 bp tandemly repeated sequence isolated from nuclear envelopes of chicken erythrocytes is located predominantly on microchromosomes. Chromosoma. 1990; 99: 131–137. PMID: 2357914
15. Tanaka K, Suzuki T, Nojiri T, Yamagata T, Namikawa T, Matsuda Y. Characterization and chromosomal distribution of a novel satellite DNA sequence of Japanese quail (*Coturnix coturnix japonica*). J Hered. 2000; 91: 412–415. PMID: 10994713
16. Matzke AJM, Varga F, Gruendler P, Unfried I, Berger H, Mayr B, et al. Characterization of a new repetitive sequence that is enriched on microchromosomes of turkey. Chromosoma. 1992; 102: 9–14. https://doi.org/10.1007/BF00352284 PMID: 1291229
17. Yamada K, Shibusawa M, Tsudzuki M, Matsuda Y. Molecular cloning and characterization of novel centromeric repetitive DNA sequences in the blue-breasted quail (*Coturnix chinensis*, Galliformes). Cytogenet Genome Res. 2002a; 98: 255–261. doi:71044
18. Ishishita S, Tsuruta Y, Uno Y, Nakamura A, Nishida C, Griffin DK, et al. Chromosome size-correlated and chromosome size-uncorrelated homogenization of centromeric repetitive sequences in New World quails. Chromosome Res. 2014; 22: 15–34. https://doi.org/10.1007/s10577-014-9402-3 PMID: 24532185
19. Yamada K, Nishida-Umehara C, Matsuda Y. Characterization and chromosomal distribution of novel satellite DNA sequences of the lesser rhea (*Pterocnemia pennata*) and the greater rhea (*Rhea americana*). Chromosome Res. 2002b; 10: 513–523.
20. Nishida C, Ishijima J, Ishishita S, Yamada K, Griffin DK, Yamazaki T, et al. Karyotype Reorganization with Conserved Genomic Compartmentalization in Dot-Shaped Microchromosomes in the Japanese Mountain Hawk-Eagle (*Nisaetus nipalensis orientalis*, Accipitridae). Cytogenet Genome Res. 2013; 141: 284–294. https://doi.org/10.1159/000352067 PMID: 23838459
21. de Oliveira TD, Kretschmer R, Bertocchi NA, Degrandi TM, de Oliveira EHC, Cioffi M de B, et al. Genomic organization of repetitive DNA in woodpeckers (Aves, Piciformes): implications for karyotype and ZW sex chromosome differentiation. PLoS ONE. 2017; 12: e0169987. https://doi.org/10.1371/journal.pone.0169987 PMID: 28081238
22. Yamada K, Nishida-Umehara C, Matsuda Y. Molecular and cytogenetic characterization of site-specific repetitive DNA sequences in the Chinese soft-shelled turtle (*Pelodiscus sinensis*, Trionychidae). Chromosome Res. 2005; 13: 33–46. https://doi.org/10.1007/s10577-005-2351-0 PMID: 15791410
23. Saifitdinova AF, Derjusheva SE, Malykh AG, Zhurov VG, Andreeva TF, Gaginskaya ER. Centromeric tandem repeat from the chaffinch genome: isolation and molecular characterization. Genome. 2001; 44: 96–103. PMID: 11269362
24. Yamada K, Nishida-Umehara C, Matsuda Y. A new family of satellite DNA sequences as a major component of centromeric heterochromatin in owls (Strigiformes). Chromosoma. 2004; 112: 277–287. https://doi.org/10.1007/s00412-003-0267-z PMID: 14997323
25. Kumar S, Stecher G, Suleski M, Hedges SB. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. Mol Biol Evol. 2017; 34: 1812–1819. https://doi.org/10.1093/molbev/msx116 PMID: 28387841
26. Rodrigues BS, de Assis MFL, O'Brien PCM, Ferguson-Smith MA, de Oliveira EHC. Chromosomal studies on *Coscoroba coscoroba* (Aves: Anseriformes) reinforce the Coscoroba–Cereopsis clade. Biol. J. Linn. Soc. 2014; 111: 274–279.
27. McHugh KP, Madsen CS, de Kloet SR. A highly repeated retropseudogene-like sequence in DNA of the redbreasted merganser (*Mergus serrator*). Gene. 1990; 87: 193–197. PMID: 2332168
28. Madsen CS, McHugh KP, de Kloet SR. Characterization of a major tandemly repeated DNA sequence (RBMII) prevalent among many species of waterfowl (Anatidae). Genome. 1992; 35: 1037–1044. PMID: 1473722

29. Suzuki T, Kurosaki T, Shimada K, Kansaku N, Kuhnlein U, Zadworny D, et al. Cytogenetic mapping of 31 functional genes on chicken chromosomes by direct R-banding FISH. Cytogenet Cell Genet. 1999; 87: 32–40. https://doi.org/10.1159/000015388 PMID: 10640808

30. Sumner AT. A simple technique for demonstrating centromeric heterochromatin. Exp Cell Res. 1972; 75: 304–306 PMID: 4117921

31. Griffin DK, Haberman F, Masabanda J, O'Brien P, Bagga M, Sazanov A, et al. Micro- and macrochromosome paints generated by flow cytometry and microdissection: tools for mapping the chicken genome. Cytogenet Cell Genet. 1999; 87: 278–281. https://doi.org/10.1159/000015449 PMID: 10702695

32. Habermann FA, Cremer M, Walter J, Kreth G, Hase J von, Bauer K, et al. Arrangements of macro- and microchromosomes in chicken cells. Chromosome Res. 2001; 9: 569–584. https://doi.org/10.1023/A:1012447318535 PMID: 11721954

33. Nishida-Umehara C, Tsuda Y, Ishijima J, Ando J, Fujiwara A, Matsuda Y, et al. The molecular basis of chromosome orthologies and sex chromosomal differentiation in palaeognathous birds. Chromosome Res. 2007; 15: 721–734. https://doi.org/10.1007/s10577-007-1157-7 PMID: 17605112

34. Matsuda Y, Chapman VM. Application of fluorescence in situ hybridization in genome analysis of the mouse. Electrophoresis. 1995; 16: 261–272. PMID: 7774567

35. Uno Y, Asada Y, Nishida C, Takehana Y, Sakaizumi M, Matsuda Y. Divergence of repetitive DNA sequences in the heterochromatin of medaka fishes: Molecular cytogenetic characterization of constitutive heterochromatin in two medaka species: *Oryzias hubbsi* and *O. celebensis* (Adrianichthyidae, Beloniformes). Cytogenet Genome Res. 2013; 141: 212–226. https://doi.org/10.1159/000354668 PMID: 24028862

36. Ohno S, Stenius C, Christian LC, Beçak W, Beçak ML. Chromosomal uniformity in the avian subclass Carinatae. Chromosoma. 1964; 15: 280–288. https://doi.org/10.1007/BF00321513 PMID: 14196875

37. Takagi N, Makino S. A revised study on the chromosomes of three species of birds. Caryologia. 1966; 19: 443–455. https://doi.org/10.1080/00087114.1966.10796235

38. Denjean B, Ducos A, Darre A, Pinton A, Seguela A, Berland H, et al. Caryotypes des canards commun (*Anas platyrhynchos*), Barbarie (*Cairina moschata*) et de leur hybride. Revue Méd Vet. 1997; 148: 695–704.

39. Wójcik E, Smalec E. Description of the mallard duck (*Anas platyrhynchos*) karyotype. Folia Biol. 2007; 55: 115–120.

40. Islam FB, Uno Y, Nunome M, Nishimura O, Tarui H, Agata K, et al. Comparison of the chromosome structures between the chicken and three anserid species, the domestic duck (*Anas platyrhynchos*), Muscovy duck (*Cairina moschata*), and Chinese goose (*Anser cygnoides*), and the delineation of their karyotype evolution by comparative chromosome mapping. J Poult Sci. 2014; 51: 1–13

41. Schmid M, Steinlein C. The hypermethylated regions in Avian chromosomes. Cytogenet Genome Res. 2017; 151: 216–227. https://doi.org/10.1159/000464268 PMID: 28315870

42. Shetty S, Griffin DK, Graves JA. Comparative painting reveals strong chromosome homology over 80 million years of bird evolution. Chromosome Res. 1999; 7: 289–295. PMID: 10461874

43. Guttenbach M, Nanda I, Feichtinger W, Masabanda JS, Griffin DK, Schmid M. Comparative chromosome painting of chicken autosomal paints 1–9 in nine different bird species. Cytogenet Genome Res. 2003; 103: 173–184. doi: 76309 PMID: 15004483

44. Shibusawa M, Nishibori M, Nishida-Umehara C, Tsudzuki M, Masabanda J, Griffin DK, et al. Karyotypic evolution in the Galliformes: an examination of the process of karyotypic evolution by comparison of the molecular cytogenetic findings with the molecular phylogeny. Cytogenet Genome Res. 2004a; 106: 111–119. https://doi.org/10.1159/000078570 PMID: 15218250

45. Shibusawa M, Nishida-Umehara C, Tsudzuki M, Masabanda J, Griffin DK, Matsuda Y. A comparative karyological study of the blue-breasted quail (*Coturnix chinensis*, Phasianidae) and California quail (*Callipepla californica*, Odontophoridae). Cytogenet Genome Res. 2004b; 106: 82–90. https://doi.org/10.1159/000078569 PMID: 15218246

46. Skinner BM, Robertson LBW, Tempest HG, Langley EJ, Ioannou D, Fowler KE, et al. Comparative genomics in chicken and Pekin duck using FISH mapping and microarray analysis. BMC Genomics. 2009a; 10: 357. https://doi.org/10.1186/1471-2164-10-357 PMID: 19656363

47. Skinner BM, Völker M, Ellis M, Griffin DK. An appraisal of nuclear organisation in interphase embryonic fibroblasts of chicken, turkey and duck. Cytogenet Genome Res. 2009b; 126: 156–164. https://doi.org/10.1159/000245915

48. Berchtold D, Fesser S, Bachmann G, Kaiser A, Eilert J-C, Frohns F, et al. Nuclei of chicken neurons in tissues and three-dimensional cell cultures are organized into distinct radial zones. Chromosome Res. 2011; 19: 165–182. https://doi.org/10.1007/s10577-010-9182-3 PMID: 21249442

49. Maslova A, Zlotina A, Kosyakova N, Sidorova M, Krasikova A. Three-dimensional architecture of tandem repeats in chicken interphase nucleus. Chromosome Res. 2015; 23: 625–639. https://doi.org/10.1007/s10577-015-9485-5 PMID: 26316311

50. Zlotina A, Maslova A, Kosyakova N, Al-Rikabi ABH, Liehr T, Krasikova A. Heterochromatic regions in Japanese quail chromosomes: comprehensive molecular-cytogenetic characterization and 3D mapping in interphase nucleus. Chromosome Res. 2018; https://doi.org/10.1007/s10577-018-9597-9 PMID: 30565005

51. Olmo E, Signorino G. Chromorep: a reptile chromosomes database: an online reference. 2005; Available from http://chromorep.univpm.it/ Accessed 2018 Aug 21.

52. Sessions SK. Evolutionary cytogenetics in salamanders. Chromosome Res. 2008; 16: 183–201. https://doi.org/10.1007/s10577-007-1205-3 PMID: 18293112

53. Nakatani Y, Takeda H, Kohara Y, Morishita S. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. Genome Res. 2007; 17: 1254–1265. https://doi.org/10.1101/gr.6316407 PMID: 17652425

54. Voss SR, Kump DK, Putta S, Pauly N, Reynolds A, Henry RJ, et al. Origin of amphibian and avian chromosomes by fission, fusion, and retention of ancestral chromosomes. Genome Res. 2011; 21: 1306–1312. https://doi.org/10.1101/gr.116491.110 PMID: 21482624

55. Uno Y, Nishida C, Tarui H, Ishishita S, Takagi C, Nishimura O, et al. Inference of the protokaryotypes of amniotes and tetrapods and the evolutionary processes of microchromosomes from comparative gene mapping. PLOS ONE. 2012; 7: e53027. https://doi.org/10.1371/journal.pone.0053027 PMID: 23300852

56. Romanov MN, Farré M, Lithgow PE, Fowler KE, Skinner BM, O'Connor R, et al. Reconstruction of gross avian genome structure, organization and evolution suggests that the chicken lineage most closely resembles the dinosaur avian ancestor. BMC Genomics. 2014; 15: 1060. https://doi.org/10.1186/1471-2164-15-1060 PMID: 25496766

57. O'Connor RE, Romanov MN, Kiazim LG, Barrett PM, Farré M, Damas J, et al. Reconstruction of the diapsid ancestral genome permits chromosome evolution tracing in avian and non-avian dinosaurs. Nat Commun. 2018; 9: 1883. https://doi.org/10.1038/s41467-018-04267-9 PMID: 29784931

58. Damas J, Kim J, Farré M, Griffin DK, Larkin DM. Reconstruction of avian ancestral karyotypes reveals differences in the evolutionary history of macro- and microchromosomes. Genome Biol. 2018; 19: 155. https://doi.org/10.1186/s13059-018-1544-8 PMID: 30290830

59. Kuraku S, Ishijima J, Nishida-Umehara C, Agata K, Kuratani S, Matsuda Y. cDNA-based gene mapping and GC_3 profiling in the soft-shelled turtle suggest a chromosomal size-dependent GC bias shared by sauropsids. Chromosome Res. 2006; 14: 187–202. https://doi.org/10.1007/s10577-006-1035-8 PMID: 16544192

60. Matsubara K, Kuraku S, Tarui H, Nishimura O, Nishida C, Agata K, et al. Intra-genomic GC heterogeneity in sauropsids: evolutionary insights from cDNA mapping and GC_3 profiling in snake. BMC Genomics. 2012; 13: 604. https://doi.org/10.1186/1471-2164-13-604 PMID: 23140509

61. Young MJ, O'Meally D, Sarre SD, Georges A, Ezaz T. Molecular cytogenetic map of the central bearded dragon, *Pogona vitticeps* (Squamata: Agamidae). Chromosome Res. 2013; 21: 361–374. https://doi.org/10.1007/s10577-013-9362-z PMID: 23703235

62. Figuet E, Ballenghien M, Romiguier J, Galtier N. Biased gene conversion and GC-content evolution in the coding sequences of reptiles and vertebrates. Genome Biol Evol. 2014; 7: 240–250. https://doi.org/10.1093/gbe/evu277 PMID: 25527834