Unveiling new perspective of phylogeography, genetic diversity, and population dynamics of Southeast Asian and Pacific chickens

Cyrill John P. Godinez 1,2,8*, John King N. Layos 3,5, Yoshio Yamamoto 1, Tetsuo Kunieda 4, Monchai Duangjinda 5, Lawrence M. Liao 6, Xun-He Huang 7 & Masahide Nishibori 1,2,8*

The complex geographic and temporal origins of chicken domestication have attracted wide interest in molecular phylogeny and phylogeographic studies as they continue to be debated up to this day. In particular, the population dynamics and lineage-specific divergence time estimates of chickens in Southeast Asia (SEA) and the Pacific region are not well studied. Here, we analyzed 519 complete mitochondrial DNA control region sequences and identified 133 haplotypes with 70 variable sites. We documented 82.7% geographically unique haplotypes distributed across major haplogroups except for haplogroup C, suggesting high polymorphism among studied individuals. Mainland SEA (MSEA) chickens have higher overall genetic diversity than island SEA (ISEA) chickens. Phylogenetic trees and median-joining network revealed evidence of a new divergent matrilineage (i.e., haplogroup V) as a sister-clade of haplogroup C. The maximum clade credibility tree estimated the earlier coalescence age of ancestral D-lineage (i.e., sub-haplogroup D2) of continental chickens (3.7 kya; 95% HPD 1985–4835 years) while island populations diverged later at 2.1 kya (95% HPD 1467–2815 years). This evidence of earlier coalescence age of haplogroup D ancestral matriline exemplified dispersal patterns to the ISEA, and thereafter the island clade diversified as a distinct group.

The domestication of animals has led to important shifts in human demographics that helped shape early human societies. Domestic chicken (Gallus gallus domesticus) is one of the world’s most widely distributed domestic animal species. It plays a key role in human societies as the largest source of animal protein1,2 and as a significant factor in socio-cultural development3. Since domestication, chickens have been distributed throughout various countries and continents, resulting in a wide range of chicken breeds today4,5. Despite their global distribution, studies on the chicken domestication process and translocation history remain obscure. Modern biological and zooarchaeological approaches suggest that chicken domestication probably occurred across southwest China and Southeast Asia, involving one or more wild progenitors across their native geographical range6–12. Subsequently, domestic chickens have been translocated out of their domestication centers to every inhabited region by human migration and trade expansion. This led to the evolution of subpopulations of chickens in response to natural selection pressure and selective breeding for adaptation to the variety of agro-ecological conditions13.

Southeast Asia (SEA), being the most geographically complex tropical region on Earth, has given rise to a diverse and highly endemic avifauna14,15. The emergence of agricultural societies harboring domesticated animals allowed a remarkable expansion of genetically divergent domesticated populations, a case seen in chickens that

1Laboratory of Animal Genetics, Graduate School of Integrated Sciences for Life, Hiroshima University, Higashi-Hiroshima 739-8528, Japan. 2Department of Animal Science, College of Agriculture and Food Science, Visayas State University, Visca, Baybay City, Leyte 6521, Philippines. 3College of Agriculture and Forestry, Capiz State University, Burias, Mambusao, Capiz 5807, Philippines. 4Faculty of Veterinary Medicine, Okayama University of Science, Imabari, Ehime 794-8555, Japan. 5Department of Animal Science, Faculty of Agriculture, Khon Kaen University, Khon Kaen 40002, Thailand. 6Laboratory of Aquatic Botany, Graduate School of Integrated Sciences for Life, Hiroshima University, Higashi-Hiroshima 739-8528, Japan. 7School of Life Sciences, Jiaying University, Meizhou 514015, China. 8These authors contributed equally: Cyrill John P. Godinez and Masahide Nishibori. 9email: cyrilljohn.godinez@vsu.edu.ph; nishibo@hiroshima-u.ac.jp
Results

Haplotype variation and genetic diversity. We analyzed complete mtDNA control region sequences of chickens from Cambodia (n = 173), Laos (n = 63), Thailand (n = 25), Myanmar (n = 78), the Philippines (n = 6), and Fiji (n = 24) generated in this study and including previously published sequences from the Philippines (n = 129), Indonesia (n = 10), and Pacific (n = 11). A total of 133 haplotypes were identified, with 70 variable sites among the studied individuals. Island populations (i.e., Philippine and Pacific chickens) accounted for 28% of haplotypes, while 17.3% of haplotypes were shared transregionally across SEA, suggesting high polymorphism among the studied individuals. Island populations (i.e., Philippine and Pacific chickens) accounted for 28% of all unique haplotypes identified, while 72% were unique to continental populations. Summary of observed polymorphic sites and haplotype variations are presented in Supplementary Tables S1–S2.

The indices of genetic diversity for each geographic population are shown in Table 1. Undoubtedly, haplotypic diversity was very high in SEA chicken populations. The mainland Southeast Asia (MSEA) chickens had higher total haplotypic diversity (Hd = 0.963 ± 0.005) and nucleotide diversity (π = 0.00782 ± 0.00398) than the Island Southeast Asia (ISEA) chickens (Hd = 0.942 ± 0.009; π = 0.00466 ± 0.00249), although no major
Phylogeography and genetic affinities of continental and island SEA chickens. The sequences generated in the present study and the reference sequences that represent chicken mtDNA control region-based haplogroup nomenclatures were used to reconstruct the matrilineal phylogeny (Supplementary Tables S1, S4). Pioneering molecular phylogenetic studies based on mtDNA control region and mitogenomes revealed fourteen haplogroups (A-I and V-Z) of chicken worldwide. Divergent haplogroups D and V showed enigmatic phylogeny resolution and previously claimed to have been distributed in ISEA and Thailand, respectively. In this study, model-based maximum likelihood and Bayesian phylogenetic analyses produced concordant topologies and comparable branch lengths of the tree (Fig. 1a; Supplementary Figs. S2–S3). Major clades have strong SH-aLRT and UFB support for the ML tree and significant posterior probability support for the Bayesian tree. Minor differences involved only some rearrangements of terminals for haplotypes: Hap_60, Hap_61, Hap_62, Hap_66, and Hap_122, as they clustered with haplogroup D1 (i.e., sub-haplogroup D1a) in the ML tree (Supplementary Fig. S2), while grouped with haplogroup D2 in the Bayesian tree (Supplementary Fig. S3).

Phylogenetic analyses grouped the MSEA chickens into major haplogroups A, B, D, E, and F, with the evidence of newfound haplogroup V as a sister-clade of haplogroup C (Fig. 1a; Supplementary Fig. S2). Haplogroup V, classified by ancestral mutation motifs A281G, T355C, and C363T, was further subdivided into two sub-haplogroups (Fig. 1b). Here, we documented evidence of sub-haplogroup V2 (clade) with unique mutational signatures: C228T, A237G, C391T only identified in Cambodian and Laotian domestic chickens, and one haplotype of Thai RJFs (G. g. gallus (Hap_103) at the basal position of the sub-clade (Supplementary Fig. S3; Supplementary Table S1). Evidence of sub-haplogroup V1 observed predominantly in Thai RJFs (Hap_107, Hap_108, and Hap_110) shared commonality to the reclassified haplogroup V of RJFs in Thailand and Cambodia (Supplementary Fig. S3). Interestingly, both model-based phylogenetic trees revealed ancestral lineage of haplogroup D2 from MSEA chickens, mostly observed in Cambodian chickens (38.7%) and some low frequency of Laotian (7.9%) and Thai chickens (8.0%), while remaining undetected in Myanmar chickens (Fig. 1a; Supplementary Fig. S2; Supplementary Table S1). Haplogroups A and B have wide geographical distribution all over SEA, while haplogroup F was prevalent among Myanmar chickens (34.6%), with some low frequency detected in Thai chickens (12.0%). Likewise, ISEA chickens (i.e., the Philippines and Pacific) have a shared genetic affinity for predominant haplogroup D1. Godinez et al. previously characterized this island sub-group as the “Philippine-Pacific sub-clade.” This sub-clade is defined by five unique mutational motifs, A281G, C296T, T306C, A342G, and G686A, and includes diagnostic motifs from the downstream region of the complete mtDNA control region sequence (Fig. 1b; Supplementary Table S1). These findings also correspond to the diagnostic motifs (SNPs: A281G, C296T, T306C, A342G) of Polynesian chicken ancient DNA.

Consistent classification of the major mitochondrial lineages of SEA chickens was also depicted in the median-joining (MJ) network analysis (Fig. 2). Notably, haplogroup V lineage was separated from haplogroup D and haplogroup C with nine and seven mutational sites, respectively. Within the haplogroup V lineage, newly classified sub-haplogroup V2 was separated from sub-haplogroup V1 with four mutational signatures. The geographical-specific MJ network analyses exhibited a close transregional evolutionary relationship of MSEA chickens in major haplogroups except for haplogroup F, which was predominated in Myanmar chickens (Supplementary Fig. S4a–d). Similarly, the Philippine and Pacific chickens also shared closely related haplotypes classified under sub-haplogroup D1b (Supplementary Fig. S4e–f).

Population structure and genetic differentiation. We carried out a multivariate approach to complement the phylogenetic analyses to assess further the relationships among and between geographical populations, including database sequences of East Asia, South Asia, and Middle East chickens (Supplementary Table S5). The result of the PCoA distinguished population substructure between mainland and island SEA chickens along the first two axes, which accounted for 52.09 or 52.53% variation (Fig. 3a–b). A homogenous subgroup was observed within island populations, particularly among the Philippine and Pacific chickens ($F_{ST}$ = 0.06936), while MSEA populations showed more a diverse assemblage, consistent with the phylogenetic analyses and haplogroup variations. In addition, we documented closer relationships between Myanmar chickens and Yunnan chickens than any other Chinese chicken population. The pairwise $F_{ST}$ value confirmed that Myanmar and Yunnan chickens were not differentiated from each other ($F_{ST} = 0.00816; p value < 0.01$). Meanwhile, within MSEA chickens, trans-regional population substructures were observed ranging from 0.06895 between Laos and Thailand to 0.19202 between Cambodia and Myanmar (Supplementary Table S6). Interestingly, Cambodian chickens were situated halfway between other continental populations and ISEA chickens, supporting the basal affiliations of identified ancestral matriline (i.e., sub-haplogroup D2) depicted in both ML and BI phylogenetic trees. The PCoA plot also indicates a significant genetic differentiation and substructure between East Asian chickens and South Asian-Middle Eastern chickens, ranging from 0.14938 to 0.77115 (Fig. 3a; Supplementary Table S6). Similarly, we observed a close genetic affinity of Japanese and Korean chickens to the Chinese and MSEA chicken populations after removing South Asian and Middle Eastern chickens from the dataset (Fig. 3b).
Hierarchical AMOVA revealed that the majority of the variations (i.e., 79.21% between ISEA and MSEA chickens and 79.74% between MSEA and EA chickens) could be attributed to within-population differentiation, specifically chickens distributed across Southeast and East Asia (Table 2). Higher within-population variation was also observed within ISEA and Cambodian chickens. Likewise, no significant population genetic differentiation was found among groups of the island and mainland SEA chickens and groups of MSEA and East Asian chickens. These observed patterns of genetic differentiation from the partitioned variances among hierarchical groups reflect consistency established in the previous phylogenetic and PCoA analyses.

**Demographic history and divergence time estimate.** The simulations for neutrality tests indicated both MSEA and ISEA chickens deviated from neutrality (Table 1), which supported a demographic expansion. The negative and significant Tajima’s $D$ and Fu’s $Fs$ statistical values of MSEA chickens and significantly negative Fu’s $Fs$ value of ISEA chickens provided evidence for population growth signatures in the Asia–Pacific region. To obtain a better inference of the demographic history of MSEA and ISEA chickens, we evaluated the changes in maternal effective population sizes ($N_e$) at the different points along with the genealogical timescale. The Bayesian Skyline Plot (BSP) showed evidence of MSEA chicken populations experiencing an episode of popu-
lation stasis during the early Holocene period, but \( N_e \) started to increase around 4000 years BP, and imminent population growth commenced about 3000–3500 years BP (Supplementary Fig. 5a–c). On the other hand, the Philippine and Pacific chickens later started to increase their \( N_e \) around 2500 years BP and 1500 years BP, respectively (Supplementary Fig. 5d–e). Looking into the individual geographical population, BSP indicated earlier population growth of Myanmar chickens (~ 4.0 kya) than the Cambodian and Laotian chickens (~ 3.0 kya). We did not run for demic demographic inference for the Thai chicken population because it violates the sample size parameter. Similarly, among within-island populations, Philippine chickens were observed to show increased \( N_e \) around 2.5 kya, while Pacific chickens have a much recent population growth expansion estimated at 1.0–1.5 kya (Supplementary Fig. S5).

The maximum clade credibility (MCC) tree estimating the divergence time using a calibration method under an uncorrelated lognormal relaxed clock model revealed age estimates for biogeographically important nodes of haplogroups D and V in our dataset (Fig. 4c). The node age of macrohaplogroup CDV was estimated to be 6.67 kya with credibility intervals of 4235–7996 years (95% HPD). The coalescence age of sub-haplogroup D1b (PP = 1) was dated back to 2.1 kya (95% HPD 1467–2815 years) while diverging from the ancestral D-lineage approximately 3.7 kya (95% HPD 1985–4835 years). Haplogroup CV (PP = 0.80) diverged much earlier from macrohaplogroup CDV and coalesced around 5.5 kya (95% HPD 3116–7275 years) while succeeding divergence of haplogroup V (PP = 0.96) occurred around 3.9 kya (95% HPD 2125–5880 years). Newfound evidence of

Figure 2. Median-joining network of the complete mtDNA D-loop region (1232 bp) depicting the evolutionary relationship of MSEA and ISEA chicken populations. The area of each circle is proportional to the frequency of the corresponding haplotypes. The length of the branch connecting to other haplotypes corresponds to mutational positions. The figure was created using PopArt v1.7 software (http://popart.otago.ac.nz/).
sub-haplogroup V2 (PP = 1) has a more recent coalescence age dated back to 1.5 kya (95% HPD 690–2788 years), while sub-haplogroup V1 (PP = 0.97) diversified earlier (2.3 kya; 95% 1005–3815 years).

**Discussion**

The timing and location of chicken domestication have been the subject of protracted debate worldwide and have stimulated several molecular studies using modern biological and zooarchaeological data. The consensus among researchers and several molecular studies confirmed that domestic chickens evolved from red junglefowl somewhere in South and Southeast Asia, but identifying their exact geographic center of origin has been challenging. Here, we present a comprehensive resolution of mitochondrial lineage diversity and phylogenetic analyses, population differentiation, demographic inference, and divergence time estimates of chickens in Southeast Asia and the Pacific region. Patterns of sequence variation indicated that chickens in the MSEA region have higher intrapopulation genetic diversity than island populations. The average genetic diversity...
more, zooarchaeological DNA analysis can further clarify the evolutionary history of chickens in this region\textsuperscript{12,47}.

Modern chicken genetic information with increased data sources spanning Southeast Asia and Oceania. Further, haplogroup V as a sister group to haplogroup C only (Figs. 1, 2; Supplementary Fig. S2). This resulted in a clearer haplogroup CD. However, because of the expanded sample distribution and increase in samples, we characterized previously identified haplogroup C samples to haplogroup V and linked them as a sister clade to the macrohaplogroup. Selection, effective management of genetic resources, and future conservation efforts\textsuperscript{1,4,44,45}.

We also observed in these lineages, likely due to geographic proximity to the center of domestication. These diversity as they reflect ancestral genetic variations. Divergent sub-haplogroups that retained ancestral variations were also observed in these lineages, likely due to geographic proximity to the center of domestication. These defined indices of biodiversity offer great opportunities for developing genetic improvement strategies, trait selection, effective management of genetic resources, and future conservation efforts\textsuperscript{1,4,44,45}.

Pioneering molecular studies and DNA sources based on the hypervariable region (partial sequence)\textsuperscript{7}, complete mtDNA control region sequences\textsuperscript{8,39,40}, mitogenome\textsuperscript{31,34}, and whole-genome data\textsuperscript{20,22} provided important insights in resolving the chicken phylogeny. In addition, recent genome-wide phylogenetic inferences provided a new perspective of wild species ancestry (i.e., G. g. spadiceus) of domestic chickens in southwestern China and Southeast Asia\textsuperscript{11}. However, topological discrepancies have also been documented in genome-wide data, often explained by differences in data sources and taxon sampling\textsuperscript{20,22,46}. The scope of the present study defines new evidence for the high mitochondrial genetic variation documented in the major haplogroups, particularly haplogroup D with a large number of divergent haplotypes and haplogroup V, which has been detected only in Thailand, Cambodia, and Laos (Supplementary Tables S2, S3). However, we cannot invalidate the influence of RJFs samples on the overall genetic diversity as they reflect ancestral genetic variations. Divergent sub-haplogroups that retained ancestral variations were also observed in these lineages, likely due to geographic proximity to the center of domestication. These defined indices of biodiversity offer great opportunities for developing genetic improvement strategies, trait selection, effective management of genetic resources, and future conservation efforts\textsuperscript{1,4,44,45}.

Population genetic and phylogenetic analyses of more than 500 complete mtDNA control region sequences unveiled new perspectives on the population dynamics of SEA and Pacific chickens. Consistent with reports from various population genetic analyses, haplogroups A and B were widely distributed in East and Southeast Asia, while haplogroup E had the widest global distribution\textsuperscript{7,19,31,46}. Haplogroup F was primarily represented in Myanmar chickens and shared this matriline with chicken populations in adjacent Yunnan Province, China\textsuperscript{7,21,34}. Consistent with the phylogenetic analyses, the pairwise F\textsubscript{ST} value of Myanmar chickens was not genetically different from those of Yunnan chicken populations (Fig. 3a–b; Supplementary Table S6). This can be explained by the geographic proximity and the course of the Burma Road, which connects Myanmar and Yunnan Province\textsuperscript{48}. Genetic differentiation of populations and PCoA analyses revealed genetic substructure between geographically isolated populations, i.e., between MSEA and ISEA chickens, South Asian and East Asian chickens, and South Asian and ISEA chickens (Fig. 3a; Supplementary Table S6). Transregional population substructure was also observed within Southeast Asian chickens, reflecting deep phylogeographic diversification. Strong topological supports consistently define major haplogroup nomenclatures and provide evidence for the presence of a haplogroup D ancestral lineage (i.e., sub-haplogroup D2) from MSEA populations. A new matriline (i.e., sub-haplogroup V2) gave rise to the population of domestic chickens sampled in Cambodia, Laos, and Thailand, whereas their ancestral lineage (i.e., sub-haplogroup V1) was represented in Thai red junglefowl (i.e., G. g. gal.lus). The previously reconstructed mtDNA phylogenetic tree described by Huang et al.\textsuperscript{31} assigned some of the previously identified haplogroup C samples to haplogroup V and linked them as a sister clade to the macrohaplogroup CD. However, because of the expanded sample distribution and increase in samples, we characterized haplogroup V as a sister group to haplogroup C only (Figs. 1, 2; Supplementary Fig. S2). This resulted in a clearer

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline Group & N & No. of population & No. of groups & Source of variation (%) & \\
\hline No grouping & 526 & 7 & 1 & Among groups & Within populations \\
\hline A (MSEA-1,2,3 vs. MSEA-4,5,6,7) & 526 & 7 & 2 & 9.06 & 11.73** \ & 79.21** \\
\hline a.1 (1,2,3 vs. 4) & 360 & 4 & 2 & 5.92 & 7.38** \ & 86.70** \\
\hline a.2 (1,2,3 vs. 5) & 250 & 4 & 2 & 24.37 & 6.06** \ & 69.57** \\
\hline a.3 (1,2,3 vs. 6) & 212 & 4 & 2 & 22.59 & 6.32** \ & 71.09** \\
\hline a.4 (1,2,3 vs. 7) & 265 & 4 & 2 & 26.12 & 5.20** \ & 68.68** \\
\hline B (MSEA vs. EA) & 677 & 15 & 2 & 15.66* & 18.63** \ & 65.71** \\
\hline C (MSEA vs. SA) & 259 & 6 & 2 & 18.53* & 8.14** \ & 73.33** \\
\hline D (MSEA vs. ME) & 244 & 5 & 2 & 34.25* & 5.63** \ & 60.12** \\
\hline E (MSEA vs. EA\textsuperscript{a}) & 829 & 16 & 2 & 1.17 & 19.09** \ & 79.24** \\
\hline F (MSEA vs. SA\textsuperscript{a}) & 411 & 7 & 2 & 10.86* & 11.39** \ & 77.25** \\
\hline G (MSEA vs. ME\textsuperscript{a}) & 396 & 6 & 2 & 19.93* & 10.57** \ & 69.51** \\
\hline
\end{tabular}
\caption{Population genetic structure estimated from the AMOVA based on complete mtDNA D-loop sequences from (1) Philippine chickens, (2) Pacific chickens, (3) Indonesian chickens, (4) Cambodian chickens, (5) Laotian chickens, (6) Thailand chickens, (7) Myanmar chickens, and database sequences from East Asia (EA), South Asia (SA), and Middle East (ME). *Database sequences retrieved from GenBank. Significant fixation indices at * \textit{p}<0.05; ** \textit{p}<0.01.}
\end{table}
reclassification of macrohaplogroup CDV (Fig. 1b). Interestingly, the ancestral matrilines classified under subhaplogroup D2 and haplogroup V were identified in sampling areas along the Lower Mekong subregion, for example, in Champasak and Bolikhamsai provinces in Laos, in Kampong Cham, Mondulkiri, Stung Treng, and

---

**Figure 4.** Time tree phylogeny depicting divergence time estimate based on primary and secondary calibration using BEAST2 v2.6.6. Red dots indicate the nodes with calibrations. Node labels indicate the median estimated divergence time, blue bars indicate the 95% HPDs. Tip labels highlighted in red indicate red junglefowl. Tree file was visualized and edited in FigTree v1.4.4. (http://tree.bio.ed.ac.uk/software/figtree/).
to this region, suggesting a human-mediated scenario of its phylogeography. This may be due to the dispersal of animals (e.g., pig, cattle, dog, deer, and chicken) and clay animal figurines were excavated in the human burial sites, suggesting that animals were part of the ritual practices during prehistoric inhumation. It was well documented that agriculture and animal-raising were among the subsistence activities of domestic communities during prehistoric settlements in the broad valleys of the Lower Mekong. In addition, ancient DNA of Thai chicken remains from other known archaeological sites in Thailand as early as 4000 BP were dated. On the other hand, the demographic expansion pattern of the island chicken population seems to suggest the timeline of Austronesian settlement in the region.

The time tree phylogeny in the coalescent framework allowed us to estimate nodal ages of haplogroups relevant to this study. We combined primary calibration (i.e., fossil) from ancient Pacific chickens and secondary calibration from previous estimations by Lawal et al. The latter calibration can provide close derived estimates of macrohaplogroup CDV is estimated around 6.67 kya (95% HPD: 4235–7996 years). A similar age estimate was reported by Huang et al. under a relaxed molecular clock model using the same molecular rate. The evidence of earlier coalescence age of haplogroup D ancestral matriline (i.e., sub-haplogroup D2) from MSEA chickens, while island populations showed more recent demographic growth signatures. Although our BSP results consider relevant sampling schemes with high sample sizes per demes, we still carefully acknowledge the potential impact of population structure on demographic estimates. The timing of the demographic expansion of MSEA chickens observed in this study can be explained by the cultural importance of stock-raising in the archaeological sites of Non Nok Tha and Ban Chiang in Thailand around 4000–3000 BP. Bones of animals (e.g., pig, cat, dog, deer, and chicken) and clay animal figurines were excavated in the human burial sites, suggesting that animals were part of the ritual practices during prehistoric inhumation. It was well documented that agriculture and animal-raising were among the subsistence activities of domestic communities during prehistoric settlements in the broad valleys of the Lower Mekong. In addition, ancient DNA of Thai chickens that were recovered in Ban Non Wat dated back to around 2500 BP, also supported the demographic expansion of MSEA chickens. Recent morphological bone identification further documented the existence of chicken remains from other known archaeological sites in Thailand as early as 4000 BP. On the other hand, the demographic expansion pattern of the island chicken population seems to suggest the timeline of Austronesian settlement in the region.

Our coalescent-based Bayesian demographic analyses detected earlier effective population size expansion in MSEA chickens, while island populations showed more recent demographic growth signatures. Although our BSP results consider relevant sampling schemes with high sample sizes per demes, we still carefully acknowledge the potential impact of population structure on demographic estimates. The timing of the demographic expansion of MSEA chickens observed in this study can be explained by the cultural importance of stock-raising in the archaeological sites of Non Nok Tha and Ban Chiang in Thailand around 4000–3000 BP. Bones of animals (e.g., pig, cat, dog, deer, and chicken) and clay animal figurines were excavated in the human burial sites, suggesting that animals were part of the ritual practices during prehistoric inhumation. It was well documented that agriculture and animal-raising were among the subsistence activities of domestic communities during prehistoric settlements in the broad valleys of the Lower Mekong. In addition, ancient DNA of Thai chickens that were recovered in Ban Non Wat dated back to around 2500 BP, also supported the demographic expansion of MSEA chickens. Recent morphological bone identification further documented the existence of chicken remains from other known archaeological sites in Thailand as early as 4000 BP. On the other hand, the demographic expansion pattern of the island chicken population seems to suggest the timeline of Austronesian settlement in the region.
coalescence age (3.9 kya; 95% HPD: 2125–5880 years) than the previous estimate of this reclassified haplogroup. Caution is warranted for this interpretation because the coalescence age estimate of gene copies in ancestral populations is not equivalent to a population split, nor does it represent the actual onset of domestication. In conclusion, this study provides a comprehensive insight into the genetic diversity and unique population dynamics of Southeast Asian chickens. High-resolution maternal phylogenies shed new light on the evolutionary history of globally acknowledged haplogroups of SEA and Pacific chickens. It provides evidence of a new divergent maternal lineage (i.e., haplogroup V) distributed across its native range in the Lower Mekong subregion. The phylogeographic and time tree phylogeny suggests human-mediated translocation of the haplogroup D ancestral matriline (i.e., haplogroup D2) from MSEA, which later diversified, forming a divergent sub-haplogroup D1b distinct to the island populations (i.e., Philippine-Pacific subclade). Future integrated genome-wide and environmental adaptation studies are required to unravel new elements of genomic evolution of SEA chickens for sustainable genetic improvement, effective management strategies, and future conservation endeavors.

Materials and methods

Ethics statement. Animal care and experimental procedures were approved by the Institutional Animal Care and Used Committee Guidelines of Hiroshima University as established by the Laboratory of Animal Genetics, Graduate School of Integrated Sciences for Life (Approval No. 015A170426). All blood sample collections were conducted following the fundamental Guidelines on the Use of Experimental Animals of the Laboratory of Animal Genetics, Hiroshima University, Japan.

Sampling and DNA extraction. Blood samples were collected from a total of 369 individuals from Cambodia (n = 173, domestic chickens), Laos (n = 63, domestic chickens), Myanmar (n = 75, domestic chickens; n = 3, red junglefowls), Thailand (n = 18, red junglefowls; n = 7, domestic chickens), Philippines (n = 6, red junglefowls), and Fiji, Melanesia (n = 24, domestic chickens) (Supplementary Fig. S1). Details of the sampled animals and their geographical distribution are listed in Supplementary Table S1. Sampling was carried out from the unrelated individuals (e.g., sampling from a different known family and different sites within every province) to avoid lineage contamination during the later analyses. Genomic DNA was extracted from stored whole blood samples using the phenol–chloroform method.

The final dataset was supplemented with previously published sequences of Philippine chickens (n = 129) and directly submitted sequences of Indonesian (n = 10) and Pacific chickens (n = 11) retrieved from GenBank (Supplementary Table S1).

PCR amplification and sequencing. The target complete mtDNA control region (1232 bp) was amplified in two procedures. First, about 5.0 kb mtDNA D-loop fragments were amplified using a long and accurate—PCR (LA-PCR) kit (KOD-FX Neo Polymerase, TOYOBO, Osaka, Japan) with chicken DNA as a template and LA-PCR primer sets: Cytb-Forward: 5’-TACACGATCCAGGCTCATTAAACCCTTTGAACTGC-3’, 16S-Reverse: 5’-TGACAATGATCGATTCTGCTGATCTAATCATCAGGTG-3’ recommended by Nishibori et al. The reaction began with a preliminary denaturation at 94 °C for 2 min, followed by 30 cycles of DNA denaturation at 98 °C for 10 s, annealing of primers at 57 °C for 30 s, and primer extension at 68 °C for 2 min and 30 s, using a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, USA). Second, the amplified fragments were used for segmental amplification of the complete mtDNA D-loop region (1.3 kb) following the primer sets: GallIF 5’-AGGACATACGGGTAAAGCCCATG-3’ and GallIR 5’-GCTGATCCTGGTGGGTGCTG-3’ in 20 μl reaction volume containing 2 × PCR buffer, 0.4 mM dNTPs, 0.3 μM concentrations of each primer, 0.4 U of KOD-FX Neo DNA Polymerase, and 15–25 ng of amplified fragment DNA as template. The PCR cycling condition began with a preliminary denaturation at 94 °C for 2 min, followed by 30 cycles of DNA denaturation at 98 °C for 10 s, annealing of primers at 59 °C for 30 s, and primer extension at 68 °C for 30 s, using a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, USA). The DNA fragments obtained from the segmental amplification were cleaned and purified using Exonuclease I (Exol) and Shrimp Alkaline Phosphatase (SAP) to degrade the residual PCR primers and dephosphorylate the remaining dNTPs, respectively. The two PCR primers and one internal primer: GallIF 5’-TGCACCTACAGAGAGATCGCAACCC-3’ and GallIR 5’-TGCACCTACAGAGAGATCGCAACCC-3’ were used for the sequencing reaction. Subsequently, the mtDNA D-loop fragments were directly sequenced using 3130/3130xl Genetic Analyzers (Applied Biosystems, Foster City, USA).

Sequence alignment. Three hundred sixty-nine complete mtDNA control region sequences generated in this study were initially edited using GeneStudio Pro tool (GeneStudio, Inc., http://genestudio.com/). Ambiguous sites were trimmed, and cleaned sequences were aligned in MEGAX with ClustalW. Aligned nucleotide sequences were viewed using BioEdit 7.2.5 software. All newly generated sequences were deposited in the GenBank database with accession numbers OM240181-OM240549 (Supplementary Table S1).

Genetic diversity and phylogenetic inference. Intrapopulation level and intracline genetic diversity indices such as the number of haplotypes (Ht), haplotype diversity (Hd), and nucleotide diversity (π) were estimated using the DnaSP v6.0 software.

Phylogenetic analyses were inferred using two different model-based approaches: maximum-likelihood (ML) and Bayesian inference (BI). Maximum-likelihood analysis was performed in IQ-TREE with the best-fit substitution model, TIM2 + F + I + G4, based on the Bayesian Information Criterion (BIC) determined by Modelfinder. Statistical node support was calculated using ultrafast bootstrap support and SH-aLRT with 1,000 replicates. Bayesian inference was performed using BEAST v2.6.6 under uncorrelated relaxed clock
log-normal distribution setting a clock rate $3.13 \times 10^{-7}$ mutations/site/year rate$^{86}$. We used a general time reversible (GTR) nucleotide substitution site model with assumed rate heterogeneity among sites modeled under gamma distribution and a coalescent-based model as a tree prior. The second-best model in BIC (GTR model) was implemented because the TIM2 model is not available in BEAST2 v2.6.6. We estimated posterior distributions of parameters via Markov chain Monte Carlo (MCMC) with duplicate runs of 50 million generations, sampling every 10,000 steps, and the initial 10% trees of each MCMC run were discarded as burn-in. Convergence of MCMC chains was assessed using Tracer v1.7.1 and sufficient sampling was verified with all estimated parameters exceeding 200 ESS values. A maximum clade credibility (MCC) tree (target tree) was obtained from a sample of trees using TreeAnnotator v2.6.3$^{85}$.

Phylogenetic trees were visualized and edited in FigTree v1.4.4. (http://tree.bio.ed.ac.uk/software/figtree/).

Median-joining network was constructed to infer the evolutionary relationships between haplotypes using PopArt v1.7 software$^{87}$. The number and assignment of haplotypes were determined using DnaSP v6.0 software. The definition of haplogroups was employed in DomeTree (http://dometree.kiz.ac.cn/) and MitoToolPy (http://mitotool.kiz.ac.cn/)$^{88}$.

### Population genetic structure and demographic inference

The population pairwise net genetic distance based on population pairwise $F_{ST}$ (significant values were accepted at $p<0.05$) was estimated using Arlequin v3.5.2.2 software (with 10,000 permutations)$^{89}$. Population pairwise $F_{ST}$ values were plotted into the principal coordinate analysis (PCoA) using GenAlEx v6.503$^{90}$ to visualize the pattern of genetic relationships between geographical populations. Estimation of the genetic structures was calculated by the analysis of molecular variance (AMOVA) as implemented by Arlequin v3.5.2.2 software. The level of significance was evaluated based on 1,000 random permutations.

Inference for the population growth model was initially estimated by statistical neutrality tests, such as Tajima's $D$$^{91}$ and Fu's $F_{S}$ statistics$^{92}$. These population expansion tests measure haplotype frequencies under neutrality. Statistical tests and confidence intervals were based on a coalescent simulation algorithm under a neutral infinite-site model. To further support the inference for the population expansion signal, a coalescent-based Bayesian Skyline Plot (BSP) was cautiously used to quantify the relationship between genealogies and the demographic history of the population$^{93}$. BSP was simulated to infer deeper insights into the demographic history of Southeast Asian and Pacific chicken populations as implemented in BEAST v2.6.3$^{85}$. BSP was generated with a relaxed molecular clock model and setting with $3.13 \times 10^{-7}$ mutations/site/year rate$^{86}$. The piecewise constant function and HKY + G4 nucleotide substitution model as determined by BIC in jModelTest v2.1.1$^{94}$ was used for the analysis. The MCMC chain was run for $5 \times 10^7$ generations, with a sampling of parameters every 5000 steps and $5 \times 10^6$ generations served as burn-in. Convergence of the posterior estimates of the effective population size ($N_e$) to the likelihood stationary distribution was evaluated using Tracer v1.7.1 software$^{95}$.

### Divergence time estimate

Bayesian analyses were performed to estimate divergence times using the program BEAST2 v2.6.6. We employed a relaxed molecular clock model, which allows substitution rates to vary across branches setting with $3.13 \times 10^{-7}$ mutations/site/year rate$^{86}$ under uncorrelated lognormal distribution and GTR + G4 substitution model as determined by BIC in jModelTest v2.1.1. We set a coalescent-based constant population to model the tree prior. The ancient DNA records of Polynesian chickens were used to calibrate the crown node of sub-haplogroup D1b (Philippine-Pacific sub-clade) (Supplementary Table S1). For this calibration point, we used a lognormal prior (mean: 2.5, SD: 0.20, offset: 0) with the maximum age of the archaeological record set as the minimum bound for the crown calibration$^{93-95,96}$. For the calibration of the root node of the tree, we used the established divergence time between red junglefowl and domestic chickens (8093 years CI: 7014–8768)$^{94}$ as a secondary calibration. We used a lognormal prior (mean: 8.09, SD: 0.05, offset: 0) covering the confidence interval range of the divergence time estimate$^{93-95}$. Time tree analysis was run for 50 million generations, sampling every 50,000 generations, and the initial 10% trees of each MCMC run were discarded as burn-in. The resulting log files were examined in Tracer v1.7.1 software$^{86}$ to confirm acceptable mixing and convergence of all parameters in the independent runs and adequate effective sample sizes (ESS > 200). The MCC tree was created from the tree file using TreeAnnotator v2.6.3$^{85}$ with the posterior probability set to 0.5 and common ancestor node heights summarized. These results were visualized as a single tree in FigTree v1.4.4. (http://tree.bio.ed.ac.uk/software/figtree/).

### Data availability

The complete mtDNA D-loop sequences are deposited and available in GenBank database (Accession Numbers: OM240181–OM240549).

Received: 18 January 2022; Accepted: 22 August 2022
Published online: 26 August 2022

### References

1. Food and Agriculture Organization. The state of the world's animal genetic resources for food and agriculture (Rome, 2007).
2. Lawler, A. Dawn of the chicken revealed in Southeast Asia. Science 368, 1411 (2020).
3. Sykes, N. A social perspective on the introduction of exotic animals: the case of the chicken. World Archaeol. 44, 158–169 (2012).
4. Groeneveld, L. F. et al. Genetic diversity in farm animals—a review. Annu. Genet. 44, 6–31 (2010).
5. Malomane, D. K. et al. The SYNBRREED chicken diversity panel: A global resource to assess chicken diversity at high genomic resolution. BMC Genomics 20, 345 (2019).
6. West, B. & Zhou, B. X. Did chickens go North? New evidence for domestication. J. Archaeol. Sci. 15, 515–533 (1988).
7. Liu, Y. P. et al. Multiple maternal origins of chickens: Out of the Asian jungles. Mol. Phylogenet. Evol. 38, 12–19 (2006).

---

https://doi.org/10.1038/s41598-022-18904-3
8. Miao, Y. W. et al. Chicken domestication: An updated perspective based on mitochondrial genomes. Heredity 110, 277–282 (2013).
9. Eda, M. et al. Reevaluation of early Holocene chicken domestication in northern China. J. Archaeol. Sci. 67, 25–31 (2016).
10. Peters, J., Lebrasseur, O., Deng, H. & Larson, G. Holocene cultural history of red jungle fowl (Gallus gallus) and its domestic descendant in East Asia. Quat. Sci. Rev. 142, 102–119 (2016).
11. Wang, M. S. et al. 863 genomes reveal the origin and domestication of chicken. Cell Res. 30, 693–701 (2020).
12. Eda, M. Origin of the domestic chicken from modern biological and zooarchaeological approaches. Anim. Front. 11, 52–61 (2021).
13. Lawal, R. A. & Hanotte, O. Domestic chicken diversity: Origin, distribution, and adaptation. Anim. Genet. 52, 385–394 (2021).
14. Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. Biodiversity hotspots for conservation priorities. Nature 403, 853–858 (2000).
15. Lohman, D. J. et al. Biogeography of the Indo-Australian archipelago. Annu. Rev. Ecol. Evol. Syst. 42, 205–231 (2011).
16. Larson, G. & Burger, J. A population genetics view of animal domestication. Trends Genet. 29, 197–205 (2013).
17. Mittermeier, R. A. et al. Hotspots revisited: Earth’s biologically richest and most endangered terrestrial ecoregions. CEMEX (2004).
18. Pitt, J., Gillingham, P. K., Maltby, M. & Stewart, J. R. New perspectives on the ecology of early domestic fowl: An interdisciplinary approach. J. Archaeol. Sci. 74, 1–10 (2016).
19. Herrera, M. B. et al. European and Asian contribution to the genetic diversity of mainland South American chickens. R. Soc. Open Sci. 7, 191558 (2020).
20. Lawal, R. A. et al. The wild species genome ancestry of domestic chickens. BMC Biol. 18, 13 (2020).
21. Tiley, G. P., Pandey, A., Kimball, R. T., Braun, E. L. & Burleigh, J. G. Whole genome phylogeny of Gallus: Introgression and data-type effects. Avian Res. 11, 7 (2020).
22. Mariadassou, M. et al. Unraveling the history of the genus Gallus through whole genome sequencing. Mol. Phylogenet. Evol. 158, 107044 (2021).
23. Hata, A. et al. Origin and evolutionary history of domestic chickens inferred from a large population study of Thai red junglefowl and indigenous chickens. Sci. Rep. 11, 2035 (2021).
24. Storey, A. A. et al. Investigating the global dispersal of chickens in prehistory using ancient mitochondrial DNA signatures. PLoS ONE 7, e39171. https://doi.org/10.1371/journal.pone.0039171 (2012).
25. Eda, M., Shoooczenghe, R., Auertrakulit, P. & Kachajiva, J. The history of chicken and other bird exploitation in Thailand: Preliminary analysis of bird remains from four archaeological sites. Int. J. Osteoarchaeol. 29, 231–237 (2019).
26. Furumido, H. et al. One subspecies of the red junglefowl (Gallus gallus) suﬃces as the matrilinear ancestor of all domestic breeds. Proc. Natl. Acad. Sci. U. S. A. 91, 12505–12509 (1994).
27. Furumido, H. et al. Monophyletic origin and unique dispersal patterns of domestic fowls. Proc. Natl. Acad. Sci. U. S. A. 93, 6792–6795 (1996).
28. Kanginakudru, S., Metta, M., Jakati, R. D. & Nagaraju, J. Genetic evidence from Indian red jungle fowl corroborates multiple domestication of modern day chicken. BMC Evol. Biol. 8, 174 (2008).
29. Nishibori, M., Shimogiri, T., Hayashi, T. & Yasue, H. Molecular evidence for hybridization of species in the genus Gallus except for Gallus varius. Anim. Genet. 36, 367–373 (2005).
30. Eriksson, J. et al. Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. PLoS Genet. 4, e1000010. https://doi.org/10.1371/journal.pgen.1000010 (2008).
31. Huang, X. H. et al. Was chicken domesticated in northern China? New evidence from mitochondrial genomes. Sci. Bull. 63, 743–746 (2018).
32. Mwacharo, J. M. et al. Mitochondrial DNA reveals multiple introductions of domestic chicken in East Africa. Mol. Phylogenet. Evol. 58, 374–382 (2011).
33. Al-Jumaili, A. S. et al. The maternal origin of indigenous domestic chicken from the Middle East, the north and the horn of Africa. BMC Genet. 21, 30 (2020).
34. Mon, S. L. Y. et al. Phylogeographic analysis of Myanmar indigenous chickens using mitochondrial D-loop sequence reveals their characteristics as a genetic resource. Anim. Sci. J. 92, e13647 (2021).
35. Dancause, K. N., Vilar, M. G., Steffy, R. & Lum, J. K. Characterizing genetic diversity of contemporary pacific chickens using mitochondrial DNA analyses. Anim. Genet. 50, 425–434 (2019).
36. Thomson, V. A. et al. Using ancient DNA to study the origins and dispersal of ancestral Polynesian chickens across the Pacific. Proc. Natl. Acad. Sci. U. S. A. 111, 4826–4831 (2014).
37. Godinez, C. J. P., Dadios, P. D. J., Espina, D. M., Matsunaga, M. & Nishibori, M. Population genetic structure and contribution of Philippine chickens to the Pacific chicken diversity inferred from mitochondrial DNA. Front. Genet. 12, 698401 (2021).
38. Gao, Y. S. et al. The genetic diversity of chicken breeds from Jiangxi, assessed with BCD02 and the complete mitochondrial DNA D-loop region. PLoS ONE 12, e0173192. https://doi.org/10.1371/journal.pone.0173192 (2017).
39. Oka, T. et al. Analysis of mtDNA sequences shows Japanese native chickens have multiple origins. Anim. Genet. 38, 287–293 (2007).
40. Hata, A. et al. Geographic origin and genetic characteristics of Japanese indigenous chickens inferred from mitochondrial D-loop region and microsatellite DNA markers. Animals 10, 2074 (2020).
41. Islam, M. A., Osman, S. A. M. & Nishibori, M. Genetic diversity of Bangladeshi native chickens based on complete sequence of mitochondrial DNA D-loop region. Br. Poult. Sci. 60, 628–639 (2019).
42. Nisar, A., Waheed, A., Khan, S., Feng, X. & Shah, A. H. Population structure, genetic diversity and phylogenetic analysis of different rural and commercial chickens of Pakistan using complete sequence of mtDNA D-loop. Mitochondrial DNA Part A 30, 273–280 (2018).
43. Osman, S. A. M., Yonezawa, T. & Nishibori, M. Origin and genetic diversity of Egyptian native chickens based on complete sequence of mitochondrial DNA D-loop region. Poult. Sci. 95, 1248–1256 (2016).
44. Toro, M. A., Fernández, J. & Caballero, A. Molecular characterization of breeds and its use in conservation. Livest. Sci. 120, 174–195 (2009).
45. Boettcher, P. J. et al. Objectives, criteria and methods for using molecular genetic data in priority setting for conservation of animal genetic resources. Anim. Genet. 41, 64–77 (2010).
46. Reddy, S. et al. Why do phylogenomic data sets yield conflicting trees? Data type influences the avian tree of life more than taxon sampling. Syst. Biol. 66, 857–879 (2017).
47. Frantz, L. A. F., Bradley, D. G., Larson, G. & Orlando, L. Animal domestication in the era of ancient genomics. Nat. Rev. Genet. 21, 449–460 (2020).
48. le Bail, H. & Tourneur, A. From Kumming to Mandalay: The new “Burma Road”. Asie Vis. 25, 1–46 (2010).
49. Higham, C. The Archaeology of Mainland Southeast Asia: From 10,000 BC to the Fall of Angkor (Cambridge University Press, 1989).
50. Collias, N. E. & Saichua, P. Ecology of the red junglefowl in Thailand and Malay with reference to the origin of domestication. Nat. Hist. Bull. Siam Soc. 22, 189–209 (1967).
51. Giles, F. H. Miscellaneous notes. No. III. Migration of Jungle-fowl. Nat. Hist. Bull. Siam. Soc. 8, 333–334 (1932).
52. Hung, H. C. et al. The first settlement of remote oceania: The Philppines to the Marianas. Antiquity 85, 909–926 (2011).
53. Bellwood, P. First Islanders: Prehistory and Human Migration in Island Southeast Asia (Wiley Blackwell, 2017).
54. Piper, P. The origins and arrival of the earliest domestic animals in Mainland and Island Southeast Asia: A developing story of complexity. In New Perspectives in Southeast Asian and Pacific Prehistory (eds Piper, P. et al.) 251–273 (ANU Press, 2017).
Acknowledgements
We wish to thank Sweet Charish Goriding-Godinez for making the base maps. This work was supported by the Grant for Animal Research Overseas from the Institute of Animal Science and the Monbukagakusho Scholarship (to C.J.P.G.—No. 19372) of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

Author contributions
C.J.P.G., J.K.N.L., L.M.L., and M.N. conceived the study and designed the experiment. C.J.P.G., Y.Y., T.K., M.D., X.H., and M.N. collected the samples. C.J.P.G., J.K.N.L., and M.N. performed the experiment and genetic analyses. C.J.P.G. and M.N. wrote the manuscript. All authors reviewed and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-18904-3.

Correspondence and requests for materials should be addressed to C.J.P.G. or M.N.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022