DOI: 10.1111/jbi.13031

ORIGINAL ARTICLE

WILEY Journal of Biogeography

Sundaland's east–west rain forest population structure: variable manifestations in four polytypic bird species examined using RAD-Seq and plumage analyses

Haw Chuan Lim[1] | Dency F. Gawin[2,3] | Subir B. Shakya[3] | Michael G. Harvey[3] | Mustafa A. Rahman[2] | Frederick H. Sheldon[3]

[1]Department of Vertebrate Zoology, National Museum of Natural History & Center for Conservation Genomics, Smithsonian Institution, Washington, DC, USA

[2]Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia

[3]Museum of Natural Science and Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, USA

Correspondence
Haw C. Lim, Department of Vertebrate Zoology, National Museum of Natural History & Center for Conservation Genomics, Smithsonian Institution, Washington, DC, USA.
Email: limhc@si.edu

Funding information
Coypu Foundation; National Geographic Society, Grant/Award Number: 8753-10; National Science Foundation, Grant/Award Number: DEB-1241059; Smithsonian Institution Molecular Evolution Fellowship

Editor: Camila Ribas

Abstract

Aim: A current model of rain forest population diversification in Sundaland specifies east–west vicariance into refugia during the early Pleistocene. In some taxa, this division was followed by dispersal and apparent secondary contact on Borneo in the late Pleistocene. To investigate genetic, morphological, spatial and temporal characteristics of the model, we compared genomic population and plumage variation among four bird species with east–west mtDNA and plumage structure.

Location: Borneo and western Sundaland (Sumatra and the Malay Peninsula).

Methods: We quantified plumage patterns among populations of two muscicapids (*Copsychus saularis* and *Kittacincla malabarica*) and two timaliids (*Mixornis gularis* and *Trichastoma malaccense*), and compared them with population genetic patterns determined from (1) SNPs produced by RAD-Seq and (2) previously sequenced mtDNA.

Results: All four species exhibit east–west variation in morphological and some genetic characters, but patterns are idiosyncratic. *Copsychus saularis'* mtDNA and plumage change gradually across Borneo, but RAD-Seq comparisons indicate no population structure. In *K. malabarica*, all three characteristics change abruptly and concurrently on Borneo. In *M. gularis*, the main east–west break occurs between Borneo and western Sundaland, with marginal mtDNA, plumage and RAD-Seq structure on Borneo. *T. malaccense* exhibits two distinct mtDNA and genomic transitions, an early Pleistocene break between western Sundaland and Borneo, and a Pliocene break between the north-east and the rest of Borneo. Despite this deep genetic division, its plumage changes clinally across Borneo.

Main conclusions: MtDNA, plumage and RAD-Seq patterns may vary depending on such factors as pre-Pleistocene distribution, habitat requirements and dispersal propensity, differential introgression among the three character types, selection on plumage and phylogenetic relationships.

KEYWORDS
Borneo, ddRAD-Seq, phylogeography, Pleistocene refugia, secondary contact, subspecies, vicariance

1 | INTRODUCTION

Sundaland is the region of Southeast Asia encompassing the Malay Peninsula, Borneo, Sumatra, Java and other islands of the Sunda continental shelf. About 20 years ago, biogeographers started developing an explanation for a common faunal pattern in Sundaland: the subdivision of closely related rain forest taxa between eastern and western parts of the shelf. In some groups, the subdivision is extreme, e.g. disjunction between western Sumatra and eastern Borneo (Wilting, Sollmann, Meijaard, Helgen, & Fickel, 2012); in more dispersive taxa, such as birds, the subdivision is sometimes manifested as secondary contact on Borneo (Lim, Rahman, Lim, Moyle, & Sheldon, 2010). The east–west pattern is attributed to vicariance due to climatic changes during global-glacial events of the early Pleistocene, when falling sea levels caused the Sunda shelf to emerge as dry land. During glacial events, the interior of Sundaland cooled and dried to such an extent that the continental islands, although connected by land, may have remained isolated by habitat. Grassland or seasonal forest (Bird, Taylor, & Hunt, 2005) is believed to have divided the subcontinent's rain forest into eastern Bornean and western Javan/Sumatran refugia, which were sustained by orographic rainfall (Gathorne-Hardy, Syaukani, Davies, Eggleton, & Jones, 2002). The existence of Sundaland's dry interior and rain forest refugia is supported by several lines of evidence, including palynology, palaeontology of the Javan mammal megafauna and phylogeographical bifurcation of numerous rain forest groups (Morley, 2012; Sheldon, Lim, & Moyle, 2015; van den Bergh, de Vos, & Sondaar, 2001; Wilting et al., 2012).

Some recent evidence, however, has challenged this scenario, hereafter called the Sundaic Pleistocene Rain Forest Refuge (SPRR) model. Habitat modelling and botanical comparisons indicate that wet forest was widespread in Sundaland during the Last Glacial Maximum (LGM; *c*. 20,000 years ago; Cannon, Morley, & Bush, 2009; Raes et al., 2014), and thus would have precluded vicariance. These findings, however, can easily be reconciled with the SPRR model if the Pleistocene's numerous glacial events varied in their influence on habitat types and dispersion (Morley, 2012; Sheldon et al., 2015). Strong evidence exists that Sundaland's interior was dry during the early Pleistocene (2.6–1 Ma), causing vicariance. The existence of extensive wet forest during recent glacial events does not refute early Pleistocene drying. Rather, it helps explain population structure on Borneo, some of which is apparently caused by secondary contact of populations expanding across central Sundaland from the west within the last 0.5 Ma (Lim, Rahman et al., 2010).

Previously, we investigated population structure in several species of Bornean rain forest birds by comparing mtDNA and nuclear DNA sequences (Lim & Sheldon, 2011; Lim, Rahman et al., 2010; Lim, Zou et al., 2010). However, our analyses emphasized mainly a lineage-based rather than population genetic perspective and implicitly relied on the assumption that most variation among the contacting populations resulted from long-term isolation in refugia. We largely ignored other potentially important elements and causes of population variation, such as clinal change in response to environmental gradients, shape of contact zones, variability in genetic introgression at hybrid zones, phenotypic plasticity and effects of geographic distance. Now, with the help of next-generation sequencing (NGS), it is possible to bring substantially enhanced genetic data and improved analytical methods to bear on issues of population variation by assessing large numbers of loci and measuring rates of gene flow, population sizes, times of divergence and other factors necessary to understanding population dynamics.

Here, we compare phylogeographical patterns in four of our previously studied species using morphology and NGS to investigate variations in the SPRR model. The four species are (as classified by Dickinson & Christidis, 2014): white-rumped shama, *Kittacincla malabarica* (Muscicapidae); Oriental magpie-robin, *Copsychus saularis* (Muscicapidae); striped tit-babbler, *Mixornis gularis* (Timaliidae); and short-tailed babbler, *Trichastoma malaccense* (Pellorneidae). These species are polytypic, comprising a few to many subspecies, including plumage-defined subspecies distributed east–west on Borneo (Fig. 1). These species also display east–west mtDNA structure on Borneo. Thus, they are distributed generally in accordance with the SPRR model. However, they differ in plumage variation between subspecies, concordance between plumage and mtDNA structure and timing of population subdivision. These differences raise the likelihood that each species experienced a different evolutionary history. Their plumage and mtDNA differences may be explained by numerous alternative scenarios, including: (1) isolation and secondary contact (or juxtaposition) according to the SPRR model without introgression; (2) isolation and secondary contact with substantial introgression; (3) no isolation, but clinal variation caused by an ecological gradient and (4) multiple isolation and contact events. These scenarios depend on a range of species characteristics, including habitat requirements, dispersal capabilities, lineage ages and phylogenetic effects.

To evaluate which scenarios are consistent with diversity patterns in the four species, we quantified variation across Borneo by scoring plumages and examining population genetic patterns obtained through genome-wide sequencing using double digest restriction site-associated DNA sequencing (hereafter RAD-Seq). With the increased resolution afforded by genome-wide data, we were able to quantify interactions among populations, interpret morphological patterns, and reconstruct evolutionary history of the four species with greater accuracy than in our previous studies.

2 | MATERIALS AND METHODS

2.1 | Study taxa

Copsychus saularis is a widespread polytypic species in Southeast Asia with seven subspecies (Dickinson & Christidis, 2014). On Borneo, three subspecies are commonly recognized: *C. s. musicus* in the west, *C. s. adamsi* in the north-east and *C. s. pluto* in the south-east (Mees, 1986) (Fig. 1). *C. s. musicus*, which also occurs on the Malay Peninsula, Sumatra and Java, has a white belly and three white outer

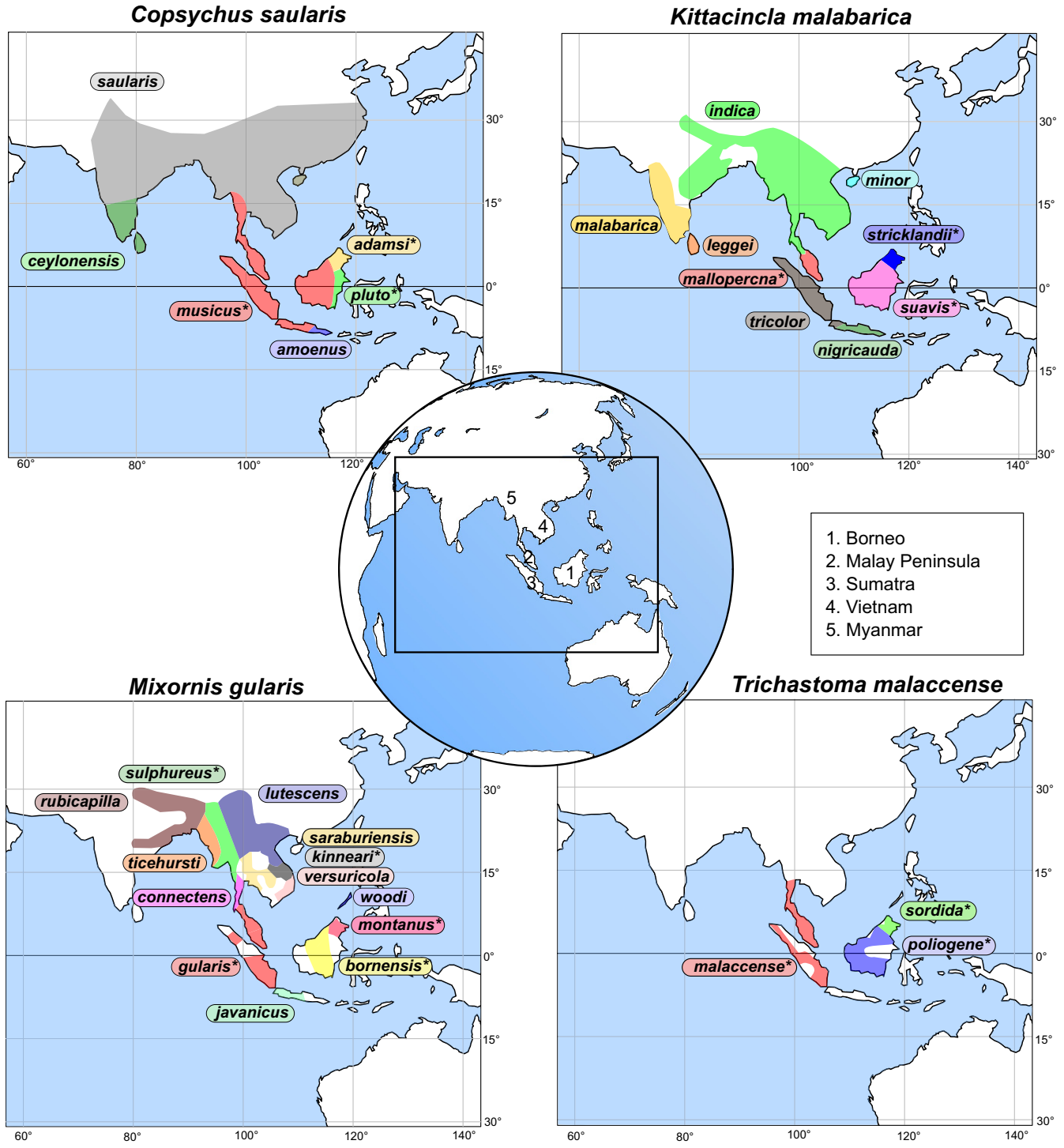

FIGURE 1 Range maps of the four study species, and approximate distributions of their main subspecies (Dickinson & Christidis, 2014). Asterisks next to subspecific names indicate that genetic samples were available for those subspecies. The central inserted map shows geographical regions from which genetic samples were obtained

tail feathers. *C. s. pluto* is all dark below with a mottled white vent and three white outer tail feathers. *C. s. adamsi* is all dark with no white underneath or in its tail. Two Pleistocene-aged mtDNA clades occur on Borneo, with haplotypes of one predominantly to the west (and most similar to haplotypes of birds from western Sundaland and mainland Asia), and the other predominantly to the east (Lim, Zou et al., 2010). Based on limited sampling, mtDNA haplotype and belly coloration are correlated: western haplotype–white belly, eastern haplotype–black belly.

Kittacincla malabarica is a widespread polytypic species in Southeast Asia with 17 subspecies (Dickinson & Christidis, 2014). Two subspecies occur on Borneo: *K. m. suavis*, with a dark crown and widespread distribution in lowland forests, and *K. m. stricklandii*, with a white crown and distribution restricted mainly to Sabah in

north-east Borneo (Fig. 1). MtDNA studies indicate an early Pleistocene divergence resulting in *K. m. stricklandii* as sister to the clade comprising *K. m. suavis* and all other *K. malabarica* subspecies, which occupy western Sundaland and south Asia. Plumage and mtDNA comparisons indicate that *K. m. suavis* and *K. m. stricklandii* overlap on Borneo in a contact zone *c*. 100–200 km wide (Lim, Zou et al., 2010).

Mixornis gularis is a widespread polytypic species in Southeast Asia with 21 subspecies (Dickinson & Christidis, 2014) (Fig. 1). On Borneo, two subspecies are recognized: *M. g. montanus* and *M. g. bornensis*. The Bornean subspecies are distinguishable at the extremes of their distributions: northern and eastern birds (*M. g. montanus*) have grey heads and mantles, whereas southern and western birds (*M. g. bornensis*) have reddish brown heads and mantles. MtDNA comparisons indicate a late Pleistocene break between Sabah and Sarawak populations, and an earlier Pleistocene break between Borneo and western Sundaland (Lim, Rahman et al., 2010).

Trichastoma malaccense is a Sunda endemic, usually divided into three subspecies: *T. m. poliogene* in most of Borneo, *T. m. sordida* in north-eastern Borneo and *T. m. malaccense* on the Malay Peninsula and Sumatra (Mees, 1986) (Fig. 1). *T. m. poliogene* has bright rufous sides, a dull greyish brown back and head (lacking contrast between the two), and a greyish brown tail. *T. m. sordida* is duller underneath, with a faint greyish band across the chest, greyish brown flanks, a rufous brown cap contrasting with a greyer back, and a rufous tail. Compared to *T. m. poliogene*, *T. m. malaccense* has a less rufous crown, more rufous wing margins, rump, and tail, and less brightly coloured underparts (Collar & Robson, 2007; Mees, 1986). On Borneo, plumages of the two subspecies grade into one another, resulting in combinations of colour patterns and leading Collar and Robson (2007) to combine them into a single subspecies, *M. m. poliogene*. MtDNA comparisons indicate deep phylogenetic breaks, Pliocene on Borneo and early Pleistocene between Borneo and western Sundaland (Lim & Sheldon, 2011).

2.2 | Plumage scoring

Specimen plumages were scored to produce hybrid indexes of geographical variation and plumage introgression across Borneo. DFG, HCL and FHS scored specimens by eye from the following collections: Louisiana State University Museum of Natural Science (LSUMNS), Delaware Museum of Natural History and National Museum of Natural History. All three authors studied the specimens extensively, and detailed notes and photographs were shared to ensure standardized scoring. Only adult birds with geographical coordinate information were compared. Specimens were scored geographically from west to east, so that 'pure' western plumages received the highest score for each character and sum of characters, and 'pure' eastern birds received the lowest character and total scores. Details of the characters and score values are provided in Tables 1 and S2–S5 in Appendix S2. Intraspecific scores, when added, maximize differentiation between the subspecies in each species (Fig. 2). Summed scores were plotted on a map using ArcGIS 10.2. We also conducted kriging analysis of summed scores using the Spatial Analyst extension of ArcGIS to interpolate plumage variation across Borneo. Kriging was conducted using the ordinary spherical semivariogram model, and the search radius was set to variable (8–12 nearest neighbours were used as the input sample points for interpolation).

2.3 | Molecular laboratory methods

We extracted DNA from preserved tissue or blood of the specimens listed in Table S1 in Appendix S2 using a DNeasy Blood and Tissue kit (Qiagen), following the manufacturer's protocol. Sampling focused on Malaysian Borneo, with a small number of additional samples from western Sundaland and mainland Asia. Some of these individuals were included in plumage scoring; others were not because they were molting, juvenile, or their study skins were not available (Tables S2–S5 in Appendix S2). After DNA quality and quantity verification, we used methods described in Parchman et al. (2012) to create amplified restriction fragment libraries (DOI of full protocol: https://doi.org/10.5061/dryad.m2271pf1). Briefly, this method involves a restriction digest of DNA, followed by adaptor ligation, PCR amplification, and size selection before the libraries are sequenced. Although this method uses a higher number of amplification PCR cycles (30) compared to other RAD-Seq methods (*c*. 18) (Baird et al., 2008), multiple studies have shown that this does not result in significant problems (e.g. Gompert et al., 2012; Parchman et al., 2012). We combined all cleaned libraries and sent them for approximately half a lane of Illumina Hiseq 2000 sequencing (101-bp single-end sequencing, SBS sequencing kit 3) at the University of Illinois Biotechnology Center (total number of samples per lane = 176).

2.4 | Bioinformatics

Programmes in STACKS 1.34 (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013) were used to demultiplex and filter raw sequencing reads, generate loci using *de novo* assembly, and determine the allelic state of each locus in each individual (see details in Appendix S1). We evaluated the use of different similarity threshold values in STACKS to assemble sequencing reads into loci (detailed information provided in Appendix S1). Testing similarity thresholds when conducting *de novo* assembly of sequencing reads into loci is important because overly stringent or liberal thresholds can result, respectively, in divergent alleles being divided into different loci (over-splitting) or paralogs forming nonsensical loci (under-splitting) (Harvey et al., 2015). Putatively paralogous loci (i.e. loci in which three or more alleles were present in an individual) were not used in any analyses.

2.5 | Population Genetics Analysis

Using STACKS' populations module, we exported SNP and haplotype data to conduct the following downstream analyses. Details related to STACKS commands and the level of data matrix completeness

TABLE 1 Scores assigned to plumage variation in individuals of the four Sundaic study species: *Kittacincla malabarica*, *Copsychus saularis*, *Mixornis gularis* and *Trichastoma malaccense*

K. malabarica		*C. saularis*						*M. gularis*		*T. malaccense*					
Crown		Belly		Vent		Tail[a]		Head and mantle		Crown to back		Breast and sides of breast		Tail	
Phenotype	Score	Phenotype	Score	Phenotype	Score	Phenotype	Score	Phenotype	Score	Phenotype	Score	Phenotype	Score	Phenotype	Score
Pure white	1	All dark	1	All dark	1	No white rectrices	0	Grey	1	Obvious chestnut to olive-brown break	1	Olive-brown	1	Chestnut	1
White with dark scalloping	2	Mottled	2	Mottled	2	Outermost rectrix white	1	Between grey and chestnut	2	Marginal break	2	Mixed	2	Slight chestnut	2
Dark	3	White	3	White	3	2 outer rectrices white	2	Rich chestnut	3	No difference (all olive-brown)	3	Pale tawny orange	3	Dull brown	3
						3 outer rectrices white	3								

[a]A rectrix is considered white if the amount of dark coloration on it < 10% of the surface area.

can be found in Table 2. First, using a matrix composed of concatenated SNPs and the program SplitsTree 4.13.1 (Huson & Bryant, 2006), we constructed an agglomerative NeighborNet phylogenetic network for each study species. A NeighborNet network is a graphical representation of a collection of splits, each representing one particular bipartition of the taxa. Each split is weighted based on how many SNPs support that particular taxon bipartitioning scheme. To remove splits with low weight (the threshold varies according to the species), we applied a filter on each species' network, removing 28–35% of the splits or 0.6–1.2% of the total split weight (Table 2).

For each species, we also ran structure 2.3.4 (Pritchard, Stephens, & Donnelly, 2000) on data matrices made up of one randomly selected SNP per RAD locus. We tested *k* (i.e. number of genetic cluster) values ranging from 1 to 8, and executed ten STRUCTURE runs per *k* value. The number of burnin steps and Markov chain Monte-Carlo (MCMC) steps after burnin was set to 100,000. All other run parameters were set to default, and no sampling location information was given as a prior. Upon the completion of STRUCTURE runs, based on metrics described by Evanno, Regnaut, and Goudet (2005), we used Structure Harvester 0.6.94 (Earl & vonHoldt, 2012) to estimate the optimal value of *k* (Fig. S1a and b in Appendix S3), in addition to using other information (e.g. from network analysis). After choosing an appropriate *k* value, we took STRUCTURE results from the ten relevant runs and used clumpp 1.1.2 (Jakobsson & Rosenberg, 2007) to generate a mean cluster membership coefficient for each of the samples. We used the R-function dudi.pca of the 'adegenet' 2 package to conduct principal component analysis on the SNP data (Jombart & Ahmed, 2011). We then plotted each sample along the first two principal component axes to visualize major population structure in each species (Patterson, Price, & Reich, 2006). For species lacking multiple divergent lineages on Borneo, we conducted isolation by distance analysis to test whether genetic distance among individuals was related to geographical distance (Slatkin, 1993). Euclidean genetic and geographical distances among individuals were calculated based on the STRUCTURE SNP data matrix and latitude–longitude information, respectively (using the dist function of 'R stats' package, R Development Core Team, 2014). We then performed a Mantel test between the two distance matrices using the mantel.randtest function of the R package 'ade4' (Dray & Dufour, 2007), with the number of Monte-Carlo permutations set to 999.

Finally, we applied G-PhoCS 1.2.3 (Gronau, Hubisz, Gulko, Danko, & Siepel, 2011) to haplotype data (number of loci/species: 175–1571) to estimate population divergence parameters through a coalescent-based approach. All individuals were used in the G-PhoCS analyses except for *M. gularis* individuals from Indo-Burma because they are distributed outside the study area. We assigned population membership of samples and topology of population relationships based on the phylogenetic network analyses (Table 2). For each species, we analysed the data using population divergence models that either permitted migration between pairs of modern populations (migration model) or did not (no-migration model). After removing burn-in iterations (variable length, based on inspection of

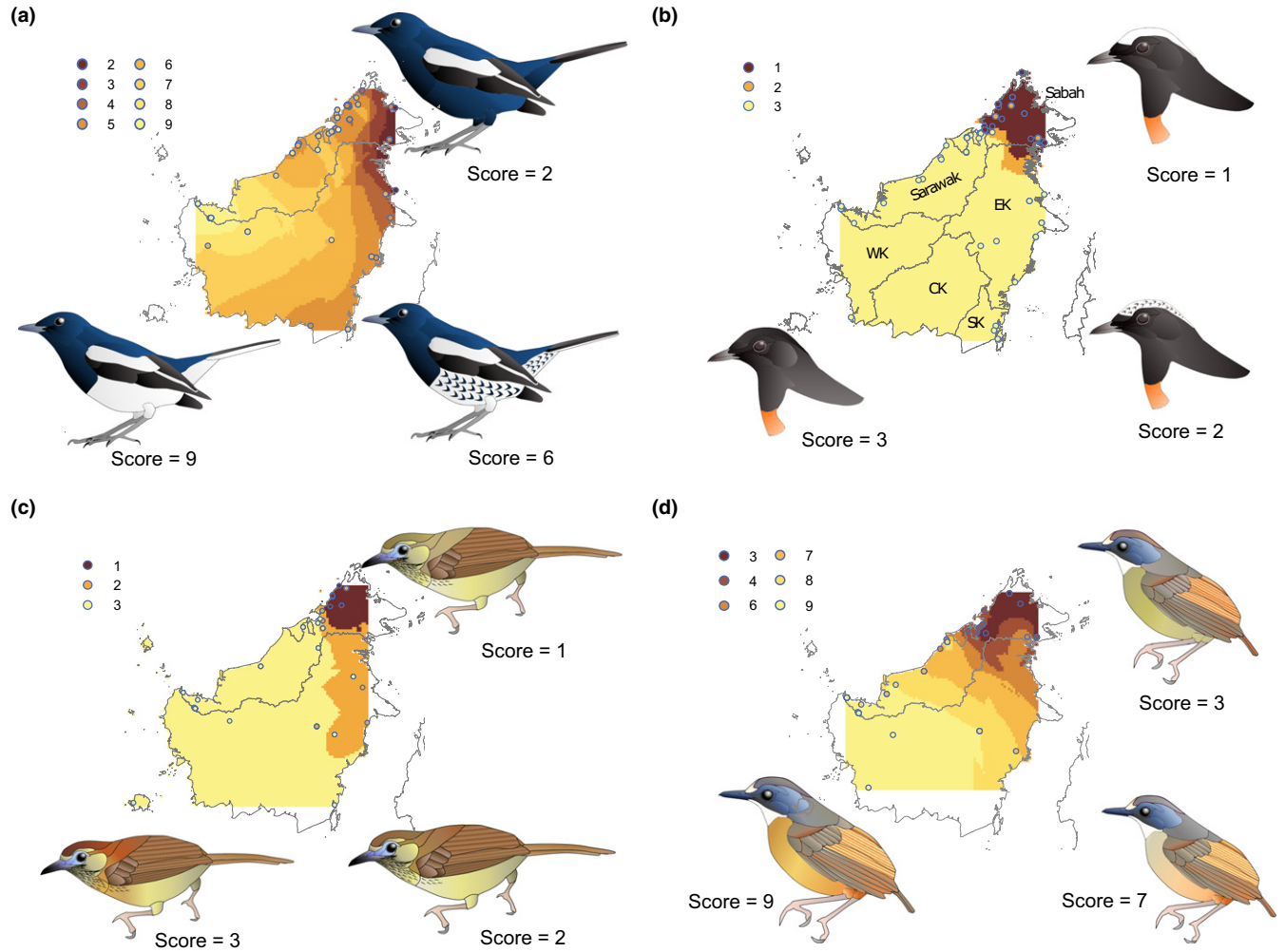

FIGURE 2 Plumage score totals (coloured circles) of (a) *Copsychus saularis*, (b) *Kittacincla malabarica*, (c) *Mixornis gularis* and (d) *Trichastoma malaccense* in Borneo. Interpolated colours result from krigging analysis. See Table 1 for information on how the plumage of each species was scored. Examples of birds showing plumages typical at each end of the spectra are also shown. EK = East Kalimantan; SK = South Kalimantan, CK = Central Kalimantan and WK = West Kalimantan

trace plots), we obtained estimates of the following population genetics parameters: θ (= $4N_e\mu$, where N_e is the effective population size and μ the mutation rate per nucleotide site per generation) for each population (current and ancestral), mutation rate-scaled divergence time τ (= $T\mu$, where T is the divergence time in generations) between populations, and per-generation migration rates between current populations, scaled by mutation rate (= M_{ST}/μ where M_{ST} is the per-generation proportion of individuals in receiving population T having immigrated from the source population S) (Gronau et al., 2011). Other details of G-PhoCS analyses can be found in Appendix S1.

3 | RESULTS

3.1 | Plumage distributions on Borneo

Plumage variation across Borneo is summarized by hybrid scores and kriging analysis patterns (Fig. 2, Tables S2–S5 in Appendix S2). In *Copsychus saularis*, plumage transition is gradual across Borneo, with whiter-bellied birds occurring to the west and darker-bellied birds occurring to the east. With two exceptions (LSU B73463 and USNM 472785), white-bellied individuals (belly score = 3) came no further east than 111.8°E. Pure *C. s. adamsi* individuals (with entirely dark underparts and tails) are uncommon in collections and the wild. Only two specimens we inspected possessed entirely black tails (Table S3 in Appendix S2: USNM 211572 and 182580). The transition from *C. s. adamsi* (north-east Borneo) to *C. s. pluto* (east Kalimantan: black belly, mottled vent and three white outer rectrices) is also gradual, with numerous birds possessing intermediate numbers of white outer rectrices (one or two). In contrast, *Kittacincla malabarica suavis* and *K. m. stricklandii* meet relatively abruptly near the Sarawak-Sabah border, and somewhat less abruptly near the Kalimantan-Sarawak border. *Mixornis gularis* plumage is fairly uniform over most of Borneo, except in the north and east, where the amount of grey on the head and mantle increases, particularly in Sabah. Plumage variation in *T. malaccense* is clinal and gradual on a south-west to north-east axis.

TABLE 2 Amount of RAD-Seq data generated and details of various population genetics analyses and polymorphism statistics for comparisons of populations within the four Sundaic species: *Kittacincla malabarica*, *Copsychus saularis*, *Mixornis gularis* and *Trichastoma malaccense*

Total number of individuals	*K. malabarica* 27	*C. saularis* 22	*M. gularis* 25	*T. malaccense* 19
Data amount				
Avg. number of filtered reads per sample (*SD*)	591,925 (106,091)	523,393 (112,121)	506,526 (81,834)	525,099 (142,230)
Avg. number of RAD loci per sample (*SD*)	58,551 (14,741)	32,955 (13,034)	46,875 (13,135)	51,653 (18,684)
NeighborNet Phylogenetic Network[a]				
Total of nucleotide sites	50,198	6937	18,103	37,896
Nucleotide substitution model used for calculating pairwise genetic distances	HKY	TrN	HKY	K80
Total network weight after filtering	0.854	0.839	0.932	0.831
Total number of splits after filtering	53	44	60	41
Principal component and structure analyses[b]				
Number of independent SNPs	20,703	4141	9247	16,092
Total variation explained by the first two PC axes (%)	14.4	15.5	20.8	35.5
Optimal *K* (no. of clusters)	2	2	3	3
G-PhoCS[c]				
Number of loci	1571	222	370	175
Topology of population relationships[a]	(SB, (SR, MY))	(BR (wSR, MY))	(BR, MY)	(SB, (SR, MY))

Polymorphism statistics[d]											
Populations and their sample sizes[e]	SB 10	SR 12	MY 5	BR 18	wSR 2	MY 2	BR 18	MY 5	SB 5	SR 7	MY 7
Exclusive polymorphisms[f]	15,528	16,293	4534	4293	484	475	10,934	2730	12,280	12,771	3714
Fixed differences[g]	32	2	83	40	46	106	364	364	2323	87	938
Shared polymorphisms[h]	10,933	12,606	7042	1299	892	579	1670	1670	2392	5315	1870

[a]Using data exported from STACKS' population module under the –phylip, –phylip_var command. SNPs are required to have non-missing data in ≥ 50% of the individuals. Distance matrix used by NeighborNet was calculated based on best fitting nucleotide substitution model estimated by jModeltest 2.1.3 (Posada, 2008).
[b]Using data exported from STACKS' population module under the –structure and –write_random_snp commands. SNPs are required to have non-missing data in ≥ 50% of the individuals.
[c]Using data exported from STACKS' population module under the –fasta_strict command. For each species, only RAD-Seq loci that were found in at least 75% of the individuals in every population were used.
[d]Codes for the populations are: MY = Malay Peninsula, BR = Sarawak and Sabah, SR =Sarawak, wSR = western Sarawak and SB = Sabah.
[e]Polymorphism statistics were calculated using the same dataset as NeighborNet Phylogenetic Network.
[f]Nucleotide polymorphisms exclusive to this population.
[g]Fixed nucleotide differences between this population and the other population(s).
[h]Nucleotide polymorphisms that are shared between this population and the other population(s).

3.2 | Data output and analysis of population structure

Sequencing produced 58 million reads when all samples were combined. After filtering, the number of reads decreased to 51 million. The average number of reads per sample was 536,104 (*SD* = 113,851), and the number of loci per sample averaged 47,948 (*SD* = 17,418) (Table 2). Across species, the average number of reads per locus per individual ranged from 3.9 to 4.3 (Table S7 in Appendix S2).

To construct NeighborNet phylogenetic networks, we used 6937 to 50,198 concatenated SNPs per species. Multiple SNPs per RAD-Seq locus were included if they met the data completeness criterion. In the phylogenetic networks, branch lengths represent the number of splits supporting the bipartitions specified by the branches. Principal component and STRUCTURE analyses were both based on one randomly selected SNP per locus (4141–20,703 SNPs per species; Table 2). In addition to the main dataset, we also assembled data with a higher coverage threshold (> 5 reads per locus per individual), and their analysis produced similar results as the following (Fig. S2 in Appendix S3).

In both network analysis and PCA, *Copsychus saularis* lacks distinct population genetic structure on Borneo, except for two individuals from western Sarawak that are more closely related to individuals from the Malay Peninsula than Borneo (Fig. 3a). STRUCTURE analysis indicated a similar pattern of weak population

structuring, with the same two western Sarawak individuals having membership coefficients closer to Malay Peninsula birds. *Copsychus saularis* from Borneo, excluding the two individuals falling in the Malayan clade, have a positive but marginally insignificant signature of isolation by distance (Mantel test: observed $r = .303$, $p = .069$, Fig. 4a).

Kittacincla malabarica stricklandii of north-east Borneo is distinct from the rest of *K. malabarica* in both the phylogenetic network and PCA plot (Fig. 3b). In the phylogenetic network, four *K. m. suavis* individuals from western Sarawak are slightly differentiated from other *K. m. suavis* in Borneo (15, 16, 19, and 21). Based on STRUCTURE metrics (Fig. S1a and b in Appendix S3) and other analyses, a *k* value of 2 is supported. Across Borneo, there is a change in genetic membership coefficient. The greatest rate of change occurred where Sarawak (*K. m. suavis*) and Sabah (*K. m. stricklandii*) birds meet. Across this zone, birds changed from having orange cluster membership of *c.* 11% to *c.* 50% (Fig. 3b). The zone encompassing birds with intermediate genotypic makeup (orange cluster membership *c.* 20–41%, birds 11, 13, 18 and 9, Fig. 3b) is *c.* 140 km wide, stretching from Mount Mulu, Sarawak (3.95° N, 114.78° E), to Mendolong, Sabah (4.93° N, 115.77° E).

Mixornis gularis exhibits little population genetic structure across Borneo, but substantial differentiation between Borneo and western Sundaland (Fig. 3c). Within Borneo, *M. gularis* exhibits a strong signature of isolation by distance (Mantel test: observed $r = .563$, $p = .001$, Fig. 4b). Birds from Vietnam and Myanmar are well differentiated from each other and from those from the Malay Peninsula.

Among the four study species, *T. malaccense* exhibits the greatest genetic differentiation among populations from different Sundaic regions (Fig. 3d). All three analyses demark three distinct populations (Sabah, Sarawak and Malay Peninsula/Sumatra) and negligible genetic admixture among any of these.

The total number of loci used in G-PhoCS analyses ranged from 175 to 1571 (Table 2). MCMC runs from which parameter estimates were obtained converged, and they produced sufficiently large numbers of independent samples of the demographic parameters (see Appendix S1 for details). In every species, values of θ in Bornean population(s) are generally larger than in western Sundaland (Fig. 5). For *C. saularis*, the divergence time between the Malay Peninsula and western Sarawak populations is similar to the divergence time between their ancestral population and *C. saularis* from the rest of Borneo (Fig. 5). In *K. malabarica*, the time of divergence (τ) between *K. m. stricklandii* (north-east Borneo) and *K. m. suavis/mallopercna* (elsewhere) is significantly earlier than the time of divergence between the two latter populations (Fig. 5). Qualitatively, divergence times among populations of *T. malaccense* are much earlier than those found in the other species, corroborating indications of deeply divergent populations revealed by other analyses (e.g. long branch lengths in the phylogenetic network). All marginal posterior distributions of inter-population gene flow in all species peaked at zero (data not shown). Moreover, 95% highest posterior density (HPD) intervals of all gene flow estimates encompassed zero. These results suggest that, for all species, post-divergence gene flow between populations has been minimal or that it was difficult to analytically tease apart the relative contributions of divergence time and gene flow (see Fig. S3 in Appendix S3 for θ and τ estimated under the no-migration model).

4 | DISCUSSION

4.1 | Geographical patterns

Each of the species we examined has a unique pattern of genetic and morphological variation. *Copsychus saularis* displays no genomic population structure on Borneo or between Borneo and western Sundaland, even though it exhibits gradual east–west plumage and mtDNA change on Borneo, and an mtDNA break between Borneo and Sumatra/Malay Peninsula. *Kittacincla malabarica* exhibits a sharp transition in plumage, mtDNA and genomic characters near the Sabah-Sarawak border. *Mixornis gularis* displays some plumage and mtDNA, but hardly any genomic, variation across Borneo. However, between Borneo and western Sundaland, a clear genomic, mtDNA and plumage break occurs. *Trichastoma malaccense* displays the greatest degree of genomic and mtDNA structure, both on Borneo and between Borneo and western Sundaland, and the genomic and mtDNA breaks concur near the Sabah-Sarawak border. However, its plumage variation on Borneo is clinal. A lack of concordance between plumage/mtDNA and nuclear variation also occurs in Borneo's endemic Mountain Blackeye (*Chlorocharis emiliae*), the only montane bird species in Sundaland for which thorough phylogeographical data are available (Manthey et al., 2017).

4.2 | Timing of genetic events

The SPRR model includes a timing component, with vicariance presumed to occur in the early Pleistocene. Unfortunately, RAD-Seq data are difficult to calibrate with respect to time. Thus, we compare ratios of mean RAD-seq G-PhoCS τ values (mutation-scaled divergence times) between populations to ratios of divergence times inferred from mitochondrial ND2 sequences. We use ND2 because data are available for all four species, and ND2's rate of substitution has been estimated (2.5% per million years; Smith & Klicka, 2010).

C. saularis divergence values between north-east Bornean and other Sundaic haplotypes, and between western Sarawak and the Malay Peninsula, are 2.0% and 0.8%, respectively (Lim, Zou et al., 2010). These percentages yield a ratio of 2.5 and an approximate within-Borneo divergence date of 0.8 Ma. The corresponding G-PhoCS τ values (1.74×10^{-4} and 1.72×10^{-4}) yield a much lower ratio (1.01). However, this discrepancy may be an analytical artefact, i.e. if G-PhoCS analysis interpreted shared variation between *C. saularis* individuals from eastern and western Borneo/western Sundaland as the result of recent divergence rather than post-divergence gene flow. We believe the low τ ratio is in fact due to ongoing gene flow in Borneo because of the presence of many individuals with intermediate plumages. Gronau et al. (2011:50 in Appendix S1) also cautioned that G-PhoCS's migration models should be viewed

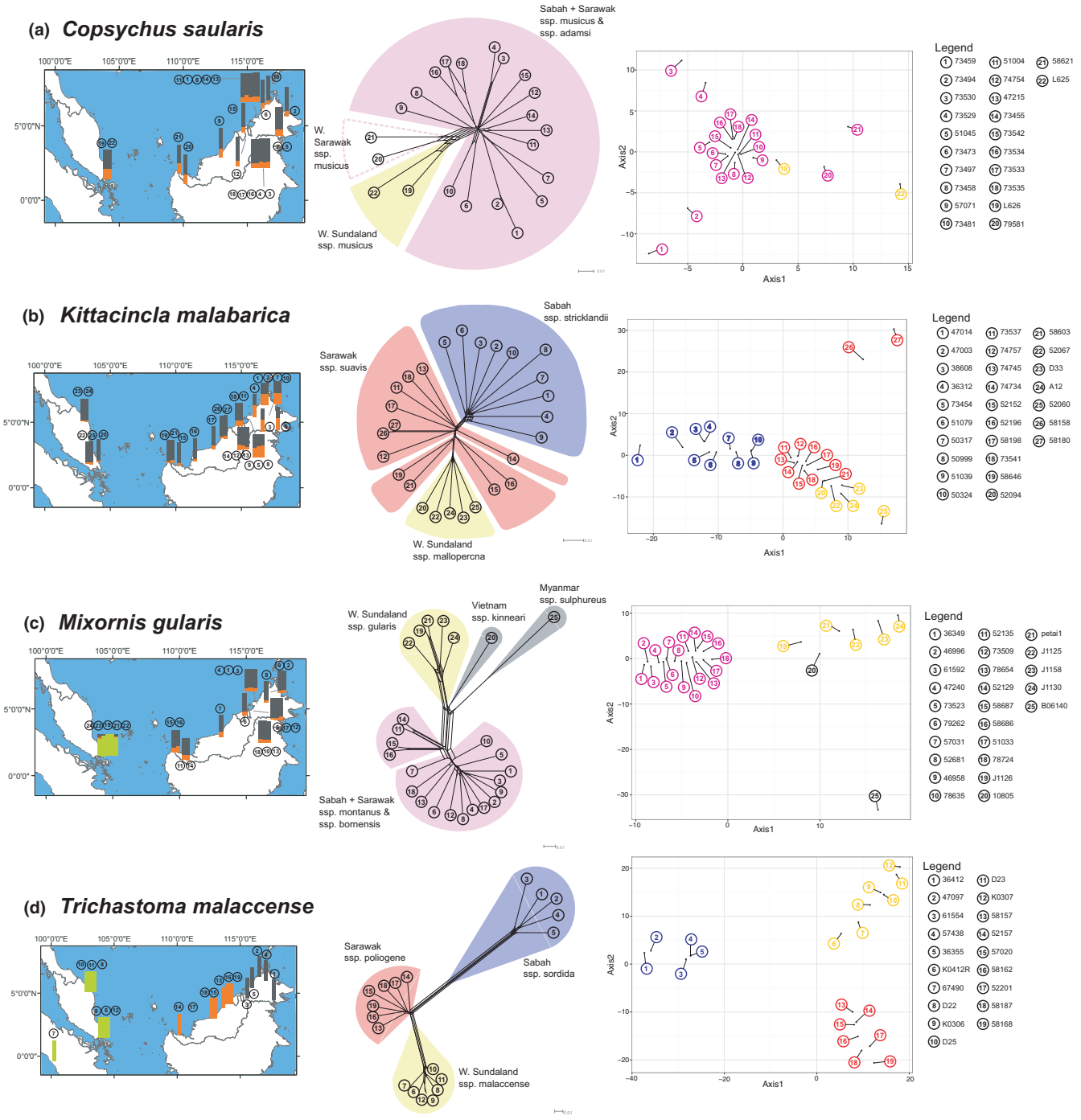

FIGURE 3 Patterns of population structure in *Copsychus saularis*, *Kittacincla malabarica*, *Mixornis gularis* and *Trichastoma malaccense*. On the left are STRUCTURE plots; columns representing individual samples are placed at approximate collection localities. Each STRUCTURE column shows the proportion of an individual's genomic makeup assigned to various genetic clusters. For *M. gularis*, STRUCTURE columns of individuals from Vietnam and Myanmar are not shown. Shading in the phylogenetic network (central) and PCA plot labels (right) use the following colour scheme indicating geographical localities of samples: blue – Sabah, red – Sarawak, pink – Sabah and Sarawak, yellow – western Sundaland (Malay Peninsula, Singapore and Sumatra), and grey – Vietnam and Myanmar. For *C. saularis*, samples indicated by dashed pink lines come from Borneo (western Sarawak), but are more closely related to samples from the Malay Peninsula. Scale bars represent the proportional split support for an edge

'primarily as a means for allowing for violations to a pure isolation model in the estimation of divergence times' rather than as tools for accurate measurement of gene flow.

Previous relaxed-clock phylogenetic analysis of *K. malabarica* using the ND2 calibration (Lim, Zou et al., 2010) found that *K. m. stricklandii* of north-east Borneo diverged from the rest of *K.*

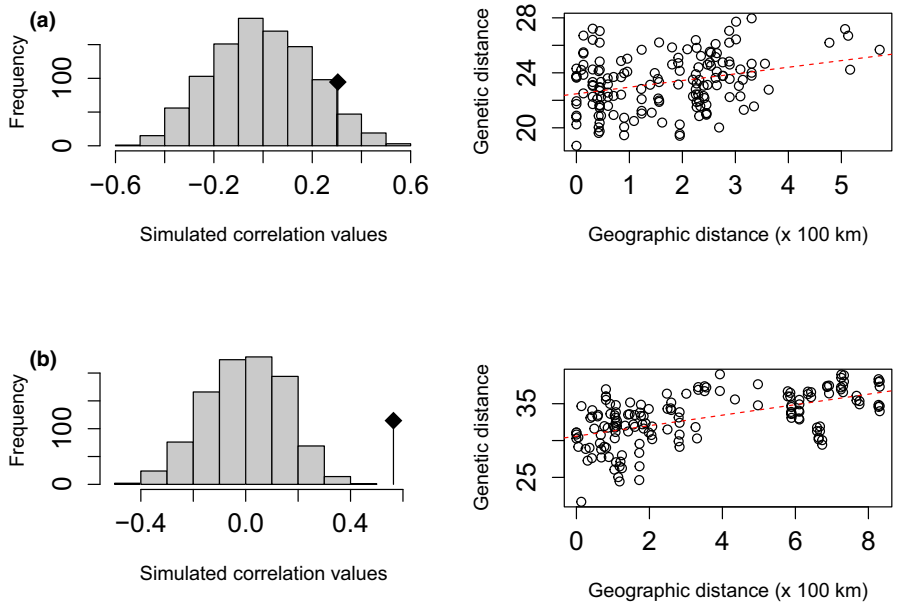

FIGURE 4 Results of isolation by distance analysis for (a) *Copsychus saularis* and (b) *Mixornis gularis* individuals on Borneo. Left panel: plots of observed correlation between genetic and geographical distance (diamonds) in a histogram of simulated values. Right panel: scatterplots of geographical and genetic distances (one circle = one pairwise comparison between individuals). The relationship between geographical distance and genetic distance is insignificant for *C. saularis* but significant for *M. gularis*

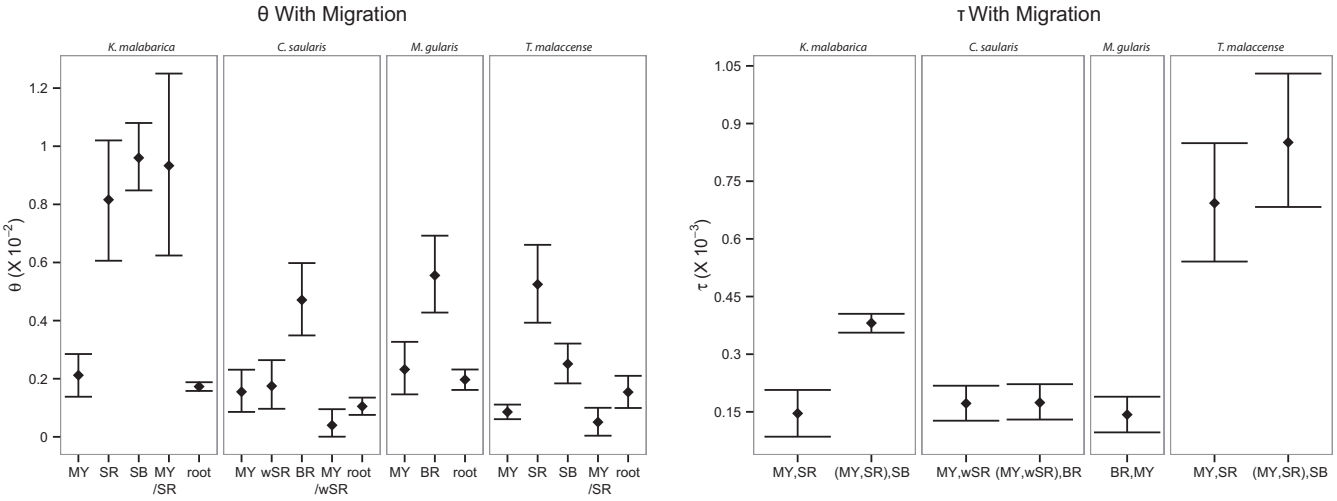

FIGURE 5 Estimates of θ and τ from G-PhoCS analyses under the migration model (diamonds = mean parameter estimate, error bars = 95% highest posterior density). Population abbreviations are: MY = Malay Peninsula, BR = Sarawak and Sabah, SR = Sarawak, wSR = western Sarawak and SB = Sabah. Estimates of θ include those for ancestral populations (e.g. MY/SR = ancestral population of MY and SR)

malabarica c. 1.5 Ma, and *K. m. suavis* of Sarawak diverged from *K. malabarica* on the Malay Peninsula and mainland Asia *c.* 1.1 Ma. The ratio between these two divergence events is 1.36, and the ratio of ND2 p-distances between the populations (2.5% and 1.7%) is 1.47. These ratios are smaller than the RAD-Seq ratio (2.61) of G-PhoCS τ values (3.81×10^{-4} and 1.46×10^{-4}) for these two divergences. Nonetheless, we consider them to be reasonably similar given the differences in genetic sample sizes and methods of analysis between the two studies.

Mixornis gularis exhibited no distinct genomic groups on Borneo, although mtDNA comparisons suggest a Pleistocene division on the island (Lim, Rahman et al., 2010). Thus, no τ-value ratio is computed. Between Borneo and the Malay Peninsula, the ND2 p-distance is 4.1%, indicating an east–west split at *c.* 1.6 Ma. The corresponding τ value is 1.43×10^{-4}.

Populations of *Trichastoma malaccense* are deeply divided on Borneo and between Sarawak and western Sundaland, with respective ND2 p-distances of 11.1% and 3.73% (Lim, Rahman et al., 2010)

and a ratio of 2.98. Corresponding τ values are 8.51×10^{-4} and 6.93×10^{-4}, yielding a ratio of 1.22. The ratio discrepancy may be due to deep coalescence of mitochondrial haplotypes separated by the older divergence, or mutation of restriction sites causing allele dropout in more divergent RAD loci. These factors would make a divergent Sabah population appear more similar genetically than it really is to the others (Gautier et al., 2013). Future studies using methods not affected by restriction site mutations (e.g. sequence capture, shotgun sequencing) will be useful for investigating this and other species with deep inter-population divergences.

4.3 | Variations in the SPRR model

All four species have subspecific and mtDNA structure consistent with Pleistocene isolation (plus Pliocene isolation in Sabah's population of *T. malaccense*) followed by more recent contact or juxtaposition. However, plumage, genomic and temporal structure in each species is idiosyncratic. To explain these differences requires some speculation. A simple model like the SPRR cannot account for the numerous geographical and biological possibilities responsible for the current distribution of rain forest birds in Sundaland. The geographical history of the region is complex. With each glacial event, rain forest types and dispersion may have changed substantially (Morley, 2012; Sheldon et al., 2015), and species' reactions to each shift in habitat are expected to differ depending on their distributions prior to and during the Pleistocene and on their ecological and behavioural characteristics. In addition, current NGS analysis, although an improvement over simple morphological and mtDNA comparisons, has limitations, e.g. in deciphering between recent population divergence and more ancient divergence followed by gene flow (Gronau et al., 2011).

The two muscicapids, *C. saularis* and *K. malabarica*, exemplify extremes in the SPRR model, most likely caused by differences in their habitat requirements and dispersal propensities. In both species, eastern populations appear to have been isolated from western populations in the Pleistocene (*C. saularis c.* 0.8 Ma, *K. malabarica c.* 1.5 Ma), followed by more recent expansion on Borneo. However, the degree of isolation and rate of post-isolation dispersal likely differed in the two species. *Copsychus saularis* occurs in open forest and coastal areas, and its broad Indian Ocean distribution suggests it is a good disperser (Lim, Zou et al., 2010; Sheldon et al., 2009). Its reduced genetic structure suggests either that *C. saularis* populations were connected periodically during the Pleistocene, or that they expanded quickly in the moderate climate of the late Pleistocene. *Kittacincla malabarica*, in contrast, is a species of well forested areas. It was presumably completely isolated for a relatively long period in the early Pleistocene, and it dispersed more widely only as tall forest expanded. When its eastern and western populations met on Borneo, they were distinct and largely failed to introgress.

Of the species examined, *M. gularis* presents the simplest manifestation of the SPRR model. Its Bornean and western Sundaic populations diverged from one another *c.* 1.6 Ma. Little population genetic structure occurs on Borneo, probably because this secondary scrub species is a good disperser.

Trichastoma malaccense adds a different dimension to interpreting the SPRR model: the influence of a deeper Sundaic history (Lim & Sheldon, 2011). Its Bornean populations — *T. m. sordida* in the north-east and *T. m. poliogene* in the west and south — are highly divergent. *T. m. sordida* was isolated, probably in eastern Borneo in the early Pliocene (*c.* 4.4 Ma). Such isolation was possible because eastern Borneo has supported rain forest continuously throughout the Cenozoic (Morley, 2012). More in accordance with the SPRR model, *T. m. poliogene* diverged from *T. m. malaccense* of western Sundaland *c.* 1.5 Ma. *T. m. poliogene* likely expanded from a Bornean refuge, but has been excluded from north-east Borneo by the older, genetically divergent *T. m. sordida*. The RAD-Seq data indicate the two taxa do not exchange genes. The layering of different aged lineages on Borneo is not surprising given the dynamism of the Sundaic region.

4.4 | Plumage versus molecules

Within the four species, concordance between plumage and genetic patterns varies widely, from general agreement in the muscicapids (*C. saularis* and *K. malabarica*) to disagreement in the timaliids (*T. malaccense* and *M. gularis*). Morphological-genetic discordance is common and has been observed since the advent of molecular systematics (King & Wilson, 1975). However, explaining it is difficult without studying genes responsible for plumage coloration and comparing enough taxa to identify phylogenetic effects on rates of morphological variation or the deterministic impact of species traits on population histories (Papadopoulou & Knowles, 2016).

Both timaliid species exhibit gradual plumage change across Borneo that is not reflected in their population genetic structure. Timaliid (as opposed to the muscicapid) plumage may be unusually susceptible to environmental influence. Of the climatic variables commonly quantified in habitat modelling (e.g. temperature and rainfall), annual precipitation provides a reasonably good match to plumage variation in both species on Borneo. Annual precipitation is generally lower in eastern Borneo (Lim, Zou, & Sheldon, 2015), corresponding with lower plumage scores in *M. gularis* and *T. malaccense* individuals. Colour differences in the two species involve changes in shades of brown, e.g. tawny orange to a darker olive-brown (breast band of *T. malaccense*), and from chestnut to grey (head and mantle of *M. gularis*). These changes appear consistent with colour changes associated with differences in melanin content in feathers or the ratio between eumelanin and phaeomelanin (McGraw, Safran, & Wakamatsu, 2005; Saino et al., 2013), and variation of these chemicals has in turn been shown to relate to rainfall or humidity gradients (Burtt & Ichida, 2004). Other potential explanations for these patterns, however, include differences in sexual selection pressures independent of the environment and neutral evolution of plumage associated with few genomic regions. Distinguishing among these hypotheses will require further studies involving

quantitative measurements of colour and pigment, better field sampling and explicit phenotype-environment analysis.

4.5 | Taxonomic issues

Copsychus saularis on Borneo exhibits no genomic structure. The eastern subspecies, *C. s. adamsi* and *C. s. pluto*, have essentially the same mtDNA, and they grade morphologically into one another and *C. s. musicus*. Populations on Borneo, therefore, do not merit treatment as separate taxonomic units.

Kittacincla malabarica stricklandii and *K. m. suavis* meet in a narrow contact zone corresponding roughly with the Sabah-Sarawak border. Despite the abruptness of this zone, STRUCTURE analysis indicates some gene flow between the two taxa. This finding conflicts with the G-PhoCS-based coalescent analysis, which indicates minimal gene flow. The conflict, however, may arise because of limitations in detecting gene flow using coalescent methods (Petit & Excoffier, 2009) and G-PhoCS analysis (see Gronau et al., 2011:50 in Appendix S1). Given *K. m. stricklandii's* relatively old divergence date, sister relationship with the rest of *K. malabarica*, distinctive plumage, and limited gene flow on Borneo, its treatment as a distinct species (e.g. by Gill & Donsker, 2016) is reasonable.

Mixornis gularis populations in eastern and western Borneo are only slightly differentiated. MtDNA comparisons indicated that *M. gularis* from western Sarawak and Sabah form reciprocally monophyletic groups, but with weak support (Lim, Rahman et al., 2010). The mitochondrial evidence, together with our current results, suggests that a western Sarawak versus the rest-of-Borneo phylogeographical break, albeit weak, may have contributed to the general pattern of isolation by distance on Borneo. Thus, whether to maintain two Bornean subspecies requires further study. The recent recognition of *M. bornensis* as a species distinct from *M. gularis* based on plumage (Collar & Robson, 2007) is supported by our genetic comparisons.

Trichastoma malaccense plumage and population genetic variation are surprisingly uncoupled. STRUCTURE, G-PhoCS and mtDNA analyses indicate the existence of deeply divergent lineages, one in western Sundaland and two on Borneo. In addition, SNP-based Fst values between population pairs range from 0.327 (Sabah-western Sundaland) to 0.169 (Sabah-Sarawak). These Fst values represent degrees of differentiation that are greater than commonly seen within bird species (Allendorf & Luikart, 2009), and together all data suggest that the three *T. malaccense* lineages may represent distinct species, despite plumage similarity.

ACKNOWLEDGEMENTS

We thank the Malaysian Prime Minister's Office and government agencies in Sabah and Sarawak for research permission and help. We also thank colleagues for collecting specimens and Vivien Chua for extracting DNA. Funding was provided by the Coypu Foundation, National Geographic Society (8753-10), National Science Foundation (DEB-1241059), and a Smithsonian Institution Molecular Evolution Fellowship. Zachary Cheviron kindly provided laboratory support. We thank the museums listed in Table S1 in Appendix S2 for use of specimens. Some computer work was conducted with support of the L.A.B. facilities of the National Museum of Natural History.

DATA ACCESSIBILITY

DNA sequences can be accessed through NCBI Short Read Archive under study number SRP103142 (BioProject number: PRJNA371350).

AUTHOR CONTRIBUTIONS

H.C.L., D.F.G. and F.H.S. conceived the research ideas; H.C.L., D.F.G., M.A.R. and F.H.S. collected the specimens, H.C.L, D.F.G. and F.H.S. collected the data; H.C.L., S.B.S. and M.G.H. analysed the data; and H.C.L. and F.H.S. wrote the manuscript.

REFERENCES

Allendorf, F.W. & Luikart, G. (2009). *Conservation and the genetics of populations*. New York: John Wiley & Sons.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., ... Johnson, E.A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, *3*, e3376.

Bird, M.I., Taylor, D. & Hunt, C. (2005). Environment of insular Southeast Asia during the last glacial period: A savanna corridor in Sundaland? *Quaternary Science Reviews*, *24*, 2228–2242.

Burtt, E.H. & Ichida, J.M. (2004). Gloger's rule, feather-degrading bacteria, and color variation among song sparrows. *Condor*, *106*, 681–686.

Cannon, C.H., Morley, R.J. & Bush, A.B.G. (2009). The current refugial rainforests of Sundaland are unrepresentative of their biogeographic past and highly vulnerable to disturbance. *Proceedings of the National Academy of Sciences USA*, *106*, 11188–11193.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A. & Cresko, W.A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140.

Collar, N.J. & Robson, C. (2007). Family Timaliidae (babblers). In J. Del Hoyo, A. Elliott & D. Christie(Eds.), *Handbook of the birds of the World* (Vol. 12, pp. 70–291). Barcelona: Lynx Edicions.

Dickinson, E.C. & Christidis, L. (2014). *The Howard & Moore complete checklist of the birds of the World* (Vol. 2, 4th ed.). Eastbourne, UK: Aves Press.

Dray, S. & Dufour, A.B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*, 1–20.

Earl, D.A. & vonHoldt, B.M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–361.

Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, *14*, 2611–2620.

Gathorne-Hardy, F.J., Syaukani, Davies, R.G., Eggleton, P. & Jones, D.T. (2002). Quaternary rainforest refugia in south-east Asia: Using termites (Isoptera) as indicators. *Biological Journal of the Linnean Society*, *75*, 453–466.

Gautier, M., Gharbi, K., Cezard, T., Foucaud, J., Kerdelhue, C., Pudlo, P., ... Estoup, A. (2013). The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Molecular Ecology*, *22*, 3165–3178.

Gill, F. & Donsker, D. (2016). *IOC World Bird List (v 6.1)*. Available at: http://www.worldbirdnames.org/ (accessed 7 July 2015).

Gompert, Z., Lucas, L.K., Nice, C.C., Fordyce, J.A., Forister, M.L. & Buerkle, C.A. (2012). Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. *Evolution*, *66*, 2167–2181.

Gronau, I., Hubisz, M.J., Gulko, B., Danko, C.G. & Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, *43*, 1031–1034.

Harvey, M.G., Judy, C.D., Seeholzer, G.F., Maley, J.M., Graves, G.R. & Brumfield, R.T. (2015). Similarity thresholds used in DNA sequence assembly from short reads can reduce the comparability of population histories across species. *PeerJ*, *3*, e895.

Huson, D.H. & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*, 254–267.

Jakobsson, M. & Rosenberg, N.A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*, 1801–1806.

Jombart, T. & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071.

King, M.-C. & Wilson, A.C. (1975). Evolution at two levels in humans and chimpanzees. *Science*, *188*, 107–116.

Lim, H.C., Rahman, M.A., Lim, S.L.H., Moyle, R.G. & Sheldon, F.H. (2010). Revisiting Wallace's haunt: Coalescent simulations and comparative niche modeling reveal historical mechanisms that promoted avian population divergence in the Malay Archipelago. *Evolution*, *65*, 321–334.

Lim, H.C. & Sheldon, F.H. (2011). Multilocus analysis of the evolutionary dynamics of rainforest bird populations in Southeast Asia. *Molecular Ecology*, *20*, 3414–3438.

Lim, H.C., Zou, F. & Sheldon, F.H. (2015). Genetic differentiation in two widespread, open-forest bird species of Southeast Asia (*Copsychus saularis* and *Megalaima haemacephala*): Insights from ecological niche modeling *Current*. *Zoology*, *61*, 922–934.

Lim, H.C., Zou, F., Taylor, S.S., Marks, B.D., Moyle, R.G., Voelker, G. & Sheldon, F.H. (2010). Phylogeny of magpie-Robins and shamas (Aves: Turdidae: *Copsychus* and *Trichixos*): Implications for island biogeography in Southeast Asia. *Journal of Biogeography*, *37*, 1894–1906.

Manthey, J.D., Moyle, R.G., Gawin, D.F., Rahman, M.A., Ramji, M.F.S. & Sheldon, F.H. (2017). Genomic phylogeography of the endemic Mountain Blackeye of Borneo (*Chlorocharis emiliae*): Montane and lowland populations differ in patterns of Pleistocene diversification. *Journal of Biogeography*, in press.

McGraw, K.J., Safran, R.J. & Wakamatsu, K. (2005). How feather colour reflects its melanin content. *Functional Ecology*, *19*, 816–821.

Mees, G.F. (1986). A list of birds recorded from Bangka Island, Indonesia. *Zoologische Verhandelingen Leiden*, *232*, 1–176.

Morley, R.J. (2012). A review of the Cenozoic palaeoclimate history of Southeast Asia. In D.J. Gower, K.G. Johnson, J.E. Richardson, B.R. Rosen, L. Rüber & S.T. Williams(Eds.), *Biotic evolution and environmental change in Southeast Asia* (pp. 79–114). Cambridge, UK: Cambridge University Press.

Papadopoulou, A. & Knowles, L.L. (2016). Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proceedings of the National Academy of Sciences USA*, *113*, 8018–8024.

Parchman, T.L., Gompert, Z., Mudge, J., Schilkey, F.D., Benkman, C.W. & Buerkle, C.A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, *21*, 2991–3005.

Patterson, N., Price, A.L. & Reich, D. (2006). Population Structure and Eigenanalysis. *PLoS Genetics*, *2*, e190.

Petit, R.J. & Excoffier, L. (2009). Gene flow and species delimitation. *Trends in Ecology and Evolution*, *24*, 386–393.

Posada, D. (2008). jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, *25*, 1253–1256.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

R Development Core Team (2014). *R: A language and environment for statistical computing*. https://www.r-project.org/

Raes, N., Cannon, C.H., Hijmans, R.J., Piessens, T., Saw, L.G., van Welzen, P.C. & Slik, J.F. (2014). Historical distribution of Sundaland's Dipterocarp rainforests at Quaternary glacial maxima. *Proceedings of the National Academy of Sciences USA*, *111*, 16790–16795.

Saino, N., Romano, M., Rubolini, D., Teplitsky, C., Ambrosini, R., Caprioli, M., ... Wakamatsu, K. (2013). Sexual dimorphism in melanin pigmentation, feather coloration and its heritability in the Barn Swallow (*Hirundo rustica*). *PLoS ONE*, *8*, e58024.

Sheldon, F., Lim, H. & Moyle, R. (2015). Return to the Malay Archipelago: The biogeography of Sundaic rainforest birds. *Journal of Ornithology*, *156*(Supplemental 1) 91–113.

Sheldon, F.H., Lohman, D.J., Lim, H.C., Zou, F., Goodman, S.M., Prawiradilaga, D.M., ... Moyle, R.G. (2009). Phylogeography of the magpie-robin species complex (Aves: Turdidae: *Copsychus*) reveals a Philippine species, an interesting isolating barrier and unusual dispersal patterns in the Indian Ocean and Southeast Asia. *Journal of Biogeography*, *36*, 1070–1083.

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, *47*, 264–279.

Smith, B.T. & Klicka, J. (2010). The profound influence of the Late Pliocene Panamanian uplift on the exchange, diversification, and distribution of New World birds. *Ecography*, *33*, 333–342.

van den Bergh, G.D., de Vos, J. & Sondaar, P.Y. (2001). The Late Quaternary palaeogeography of mammal evolution in the Indonesian Archipelago. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *171*, 385–408.

Wilting, A., Sollmann, R., Meijaard, E., Helgen, K.M. & Fickel, J. (2012). Mentawai's endemic, relictual fauna: Is it evidence for Pleistocene extinctions on Sumatra? *Journal of Biogeography*, *39*, 1608–1620.

BIOSKETCH

Haw Chuan Lim is a postdoctoral fellow at the Smithsonian Institution studying the evolution of Sundaic birds using next-generation DNA sequencing and niche modelling. The authors as a group study the phylogenetics, population genetics, systematics, ecology and conservation genetics of Southeast Asian birds.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Lim HC, Gawin DF, Shakya SB, Harvey MG, Rahman MA, Sheldon FH. Sundaland's east–west rain forest population structure: variable manifestations in four polytypic bird species examined using RAD-Seq and plumage analyses. *J Biogeogr*. 2017;00:1–13. https://doi.org/10.1111/jbi.13031