Unravelling the phylogeography of the H.azurea *(Hypothymis azurea, Aves: Monarchidae)* across Southeast Asian rainforests

A Thesis

submitted to

Indian Institute of Science Education and Research Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme

by

Jaidhithyan G

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, INDIA.

Date: March 15, 2026

Under the guidance of

Supervisor: Dr Frank Rheindt,

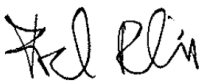
Affiliation of Supervisor

From May 2025 to Mar 2026

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

Certificate

This is to certify that this dissertation, entitled **Unravelling the phylogeography of the H.azurea *(Hypothymis azurea, Aves: Monarchidae)* across Southeast Asian rainforests** towards the partial fulfilment of the **BS-MS dual degree programme** at the I**ndian Institute of Science Education and Research, Pune,** represents study/work carried out by **Jaidhithyan G** at the **National University of Singapore, Singapore** under the supervision of **Dr. Frank Rheindt**, Associate Professor, Department of Biological Sciences, during the academic year 2025-2026.



Dr. Frank Rheindt

Committee:

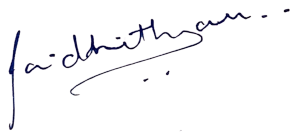
Dr. Frank Rheindt

Dr. Ramana Athreya

This thesis is dedicated to the birds of Asia.

Declaration

I hereby declare that the matter embodied in the report entitled "**Unravelling the phylogeography of the H.azurea *(Hypothymis azurea, Aves: Monarchidae)* across Southeast Asian rainforests**" are the results of the work carried out by me at the **Department of Biological Sciences, National University of Singapore(NUS)**, under the supervision of **Dr. Frank Rheindt**, and the same has not been submitted elsewhere for any other degree. Wherever others contribute, every effort is made to indicate this clearly, with due reference to the literature and acknowledgement of collaborative research and discussions.



Jaidhithyan G

20211079

Table of Contents

Abstract 8
Acknowledgments 9
Contributions 10
Chapter 1 Introduction 11
1.1 The importance of phylogenetic studies in Southeast Asia 11
1.2 Phylogenetic studies in Southeast Asian songbirds and their significance 11
1.3 Global distribution and subspecies of Hypothymis azurea 12
Chapter 2 Materials and Methods 12
2.1 Sampling and Data Collection: 13
2.2 DNA extraction and Sequencing: 14
2.3 Sequence processing and SNP-calling: 14
2.4 Phylogenetic analysis: 15
2.5 Population structure & gene flow analysis 15
Chapter 3 Results 17
3.1 Principal component analysis: 17
3.2 Phylogenetic trees: 18
3.3 Mito-nuclear discordance across taxa: 18
3.4 Species tree topology using Coalescent-Aware Species Tree Estimator: 20
3.5 Sliding window analysis reveals gene flow signals 21
Chapter 4 Discussion 24
References 27

List of Tables

Table A1. All samples of *H.azurea* and *H.puella* with subspecies and sampling locality information 33

List of Figures

Figure 1. Distribution and sampling localities of 49 *Hypothymis* samples (45 samples of *H.azurea* and 3 samples of *H.puella*) used in this study. 14

Figure 2. Overview summary of the population structure across Hypothymis azurea populations with labelled sampling localities. 18

Figure 3. Maximum-likelihood tree topology based on 2,871,412 concatenated SNPs from whole-genome samples and maximum-likelihood mitochondrial sequence tree based on 17,290 base pairs run with 1000 bootstraps, with Terpsiphone affinis as the outgroup. 20

Figure 4. Site-based species tree topology of all Hypothymis samples rooted to Terpsiphone affinis as the outgroup. 21

Figure 5. Normalised moving-average stacked plots for different population quartets (three ingroups and one outgroup)
23

Figure 6. Site-based species tree topology with inferred gene flow arrows from the sliding window analysis and sampling localities map with inferred gene flow arrows between approximate geographies.
24

Abstract

Biological diversification entails a dance between phenotypic divergence – as populations evolve morphological and behavioural differences from each other – and the underlying genetic differentiation. The two are seldom in perfect sync, as morphologically conservative lineages can be deeply diverged, whereas distinct populations may only exhibit shallow genetic divergence. To shed light on the interplay between phenotypic and genetic differentiation, we investigated the Black-naped Monarch (*Hypothymis azurea)* complex*,* distributed across mainland and insular tropical Asia, using whole-genome data sampled from populations across its distribution range. We found genetic cohesion in populations from across mainland Asia and shelf islands with a long history of continental connection, while some large island populations, especially from the Andamans and Taiwan, emerged as more deeply diverged. Phenotypically distinct populations from the Barusan island chain consistently formed a distinct clade embedded among the big continental-shelf clade, attesting to rapid morphological differentiation likely driven by founder effects and bottlenecks. Analyses of secondary gene flow revealed a recent full isolation of these Barusan populations, while other monarchs throughout the range, including those from distant islands like the Philippines, showed signs of recent admixture. Our results suggest peripatric founder effects can expedite phenotypic differentiation in short periods, insulating populations from secondary gene flow and driving differentiation that may result in speciation. This study demonstrates that geographic barriers and opportunities for gene flow play a central role in shaping the phylogeography within the *H.azurea* complex, emphasising the importance of using whole-genome approaches to understand evolutionary relationships and processes in widespread populations.

Acknowledgments

I owe the completion of this thesis to many people who have supported me along the way, and I am grateful to each of them.

First and foremost, I would like to thank my supervisor, Frank, for his support and mentorship throughout this study. He gave me the utmost freedom to shape my dissertation process, from conceptualising the project to performing the analyses. My eyes fell on his lab website and research work, and I immediately knew I would find myself working with him somewhere along the way. I am especially glad to have completed my thesis under his guidance. He not just provided opportunities related to my project, but also gave me opportunities to guide field assignments, join lab fieldwork for mist-netting birds and additionally join a birding trip where I understood a lot more about the biodiversity of this region.

Secondly, I want to extend my gratitude to my support system, those who have created a peaceful space around me to complete this thesis. Collin, Chen Xi, Charlotte and Yichen were an integral part of my learning and research journey here. From helping with troubleshooting to helping me find new birds around Singapore, they were always proactive and helpful, regardless of whether it was work-related or not. I also want to extend my gratitude to the remaining members, collaborators, and former members of the Avian Evolution lab, namely Keren, Arthur, Mayjean, Keita, Movin Samuel, Jeff, Joseph, Francesca, and Sijun, who made my time here much more memorable and enriching.

I express my gratitude to the previous members of the lab, especially Movin and Irham, and to our collaborators, Shashank Dalvi and members of the American Museum of Natural History, for their efforts in sample collection.

Similarly, I am grateful to the IISER Pune community, which has helped me develop numerous skills and introduced me to a rich network of peers who have pushed me

to learn and grow. A special thanks to Dr Ramana Athreya from IISER Pune, with whom I have had a fulfilling learning journey. He has provided me with direction and nurtured my scientific temperament.

I am forever beyond grateful to my IISER friends, who are literally my home away from home and have been by my side, through my highest of highs, my lowest of lows and everything in between. To Deep, Aparajita, Ishaan, Abhinav, Apoorvaa, Yash, Ameya, Sashank, Harsh, Autisha, Rhythm, Surana, Meghana, Shravan, Anantha, Saksham – thank you for helping me grow and shaping me into the person I am today. And an immense thank you to Avanthika- my dearest friend and companion, who supported and believed in me from the very beginning, right from applying for a thesis project till the completion. I cannot adequately express how grateful I am for her presence and support throughout.

Last but certainly not least, I thank my family: Mom, Dad, Ammuma, Tatha and my brother, for their constant support and belief in me to be able to pursue what I wanted to do ever since I was a kid. Thank you for all your blessings and sacrifices, and for inspiring me to dream big and to pursue what I have always been passionate about.

Contributions

Contributor name	Contributor role
Dr Frank Rheindt, Jaidhithyan G, Movin Nyanasengeran	Conceptualization Ideas
Jaidhithyan G, Dr Frank Rheindt, Collin Chua	Methodology
Jaidhithyan G	Software
Jaidhithyan G	Validation
Jaidhithyan G	Formal analysis
Jaidhithyan G	Investigation
Dr Frank Rheindt	Resources
Movin Nyanasengeran, Mohammed Irham, Pratibha Baveja	Data Curation
Li Sijun, Movin Nyanasengeran	DNA extraction and library preparation
Jaidhithyan G	Writing - original draft preparation
Jaidhithyan G, Dr Frank Rheindt, Collin Chua	Writing - review and editing
Jaidhithyan G	Visualization
Dr Frank Rheindt	Supervision
Jaidhithyan G, Dr Frank Rheindt	Project administration
Dr Frank Rheindt	Funding acquisition

This contributor syntax is based on the Journal of Cell Science CRediT Taxonomy[1].

[1] https://journals.biologists.com/jcs/pages/author-contributions

Chapter 1 Introduction

1.1 The importance of phylogenetic studies in Southeast Asia

Phylogeographical patterns have been used extensively to enhance our understanding of the evolutionary history of individual species or species complexes across various taxa (Gonzalez-Porter et al., 2011; Koblmüller et al., 2012; Den Tex & Leonard, 2014). Indo-Australian Archipelago, which comprises all of Southeast Asia, owing to its complex geography and dynamic geological history, is high in endemic biodiversity and turnover (Heaney, 1991). The complex biogeography of this archipelago was first brought to light by the work of Alfred Russell Wallace (Wallace, 1883).

Island dynamics and Pleistocene sea-level fluctuations have been known to shape genetic variation in faunal populations across the geographical mosaic of Southeast Asia. This is profound in insular systems like the Sunda Shelf. The Sunda Shelf is defined as the submerged extension of the continental shelf south of the Isthmus of Kra, which connects Sumatra, Borneo, and Java to mainland Asia. Here, climate fluctuations in the Pleistocene had complex interactions with the geology of the region, which resulted in high biodiversity and endemism (Sholihah et al. 2021). Over the last 2-3 million years, climate-induced fluctuations in sea level repeatedly exposed the Sunda shelf when the area of land was nearly two-fold the size of exposed land today, which was then fragmented during warm interglacial periods (Woodruff, 2010). During periods of low sea level, exposed land bridges linked regions such as Borneo, Sumatra, Java, Palawan, and Indochina, whereas during periods of high sea level, these regions were re-isolated (Heaney, 1986; Woodruff, 2010). With cyclic reconnection and fragmentation of terrestrial faunal populations, these sea-level oscillations often acted as a “species pump” promoting divergence in various taxa (Beck et al., 2017; Fengyuan Li et al., 2018; Othman et al., 2020).

1.2 Phylogenetic studies in Southeast Asian songbirds and their significance

Widespread songbirds in Southeast Asia span many environmental gradients, characterised by complex geology and repeated barriers to dispersal. Songbirds here occur across mainland and island systems- these landscapes are comprised of a mosaic of different habitats. This creates a natural laboratory where we can test how land-bridge shelves, dispersal capacity and habitat preferences shape gene flow and divergence in populations of widespread species.
Moreover, recent phylogenetic work on widespread passerines has revealed cryptic species-level diversity, especially in island and deep-forest taxa (Lohman et al., 2010; Cros et al., 2020). This becomes important in conservation action, as undetected endemics may receive low conservation priority or may lack appropriate protection (Lohman et al., 2010). A recent study highlighted a significant increase in the number of described bird species in the Indo-Australian Archipelago(IAA) since 2012, with nearly 78% of the newly recognised species-level taxa being island endemics from Sundaland and Australia (Prasetya et al., 2025). This adds to the importance of investigating the phylogeography of birds occurring across multiple islands in the Indo-Australian Archipelago to better understand their evolutionary history.

1.3 Global distribution and subspecies of *Hypothymis azurea*

The Black-naped Monarch (*Hypothymis azurea*) (Boddaert, 1783) is a widespread monarch flycatcher (Class: Aves, Family: *Monarchidae*) species found across the Southeast Asian tropical and subtropical rainforests of India, Sri Lanka, southern China, Taiwan, Thailand, the Malay Peninsula, the Philippines, and the Sundaland (Sumatra, Borneo, Java, and all the small islands of the Sunda Islands). The species complex of *Hypothymis* flycatchers consists of 23 subspecies, many of which are endemic to the islands of Southeast Asia and have been observed to be morphologically dissimilar to each other. For example, *H. a. abbotti,* a subspecies endemic to the island of Babi in Indonesia, was observed to be nearly twice the size of other populations. Another subspecies, *H.a.richmondi,* which is endemic to the island of Enggano, Indonesia, completely lacks the white belly that is found in the nominate race, is smaller than the nominate race and has been observed to wear a much bluer and darker plumage. This phenotypic variation emphasises the need to investigate molecular phylogeography in this species complex to elucidate potential

species boundaries and understand the evolutionary processes that shaped their current distribution, phenotypic variation and phylogeny, ultimately aiding more informed conservation actions (Lohman et al., 2010; Sheldon et al., 2015).

Here, we aim to determine the phylogenetic relationships and population structure among insular and continental populations of the Black-naped Monarch (*H. azurea*) and to test gene-flow scenarios that explain their observed phylogeography.

For this study, a comprehensive dataset of whole-genome sequence (WGS) data was obtained from fresh wild specimens sampled across the distribution range of the species, which includes most of the island endemic subspecies (Appendix Table A1).

Chapter 2 Materials and Methods

2.1 Sampling and Data Collection:

To build a comprehensive dataset of specimens, tissue, and blood samples from 47 individuals of the genus *Hypothymis*, 3 samples from *H. puella,* and 44 from *H.azurea* were obtained across their distribution range in Southeast Asia. 41 fresh samples were collected during field expeditions using mist nets from 2019 to 2022, and 6 fresh samples were obtained from the American Museum of Natural History's field collections in New York. To root the downstream phylogenetic analyses, a fresh sample of *Terpsiphone affinis* obtained from the previous field collections was further supplemented to the dataset.

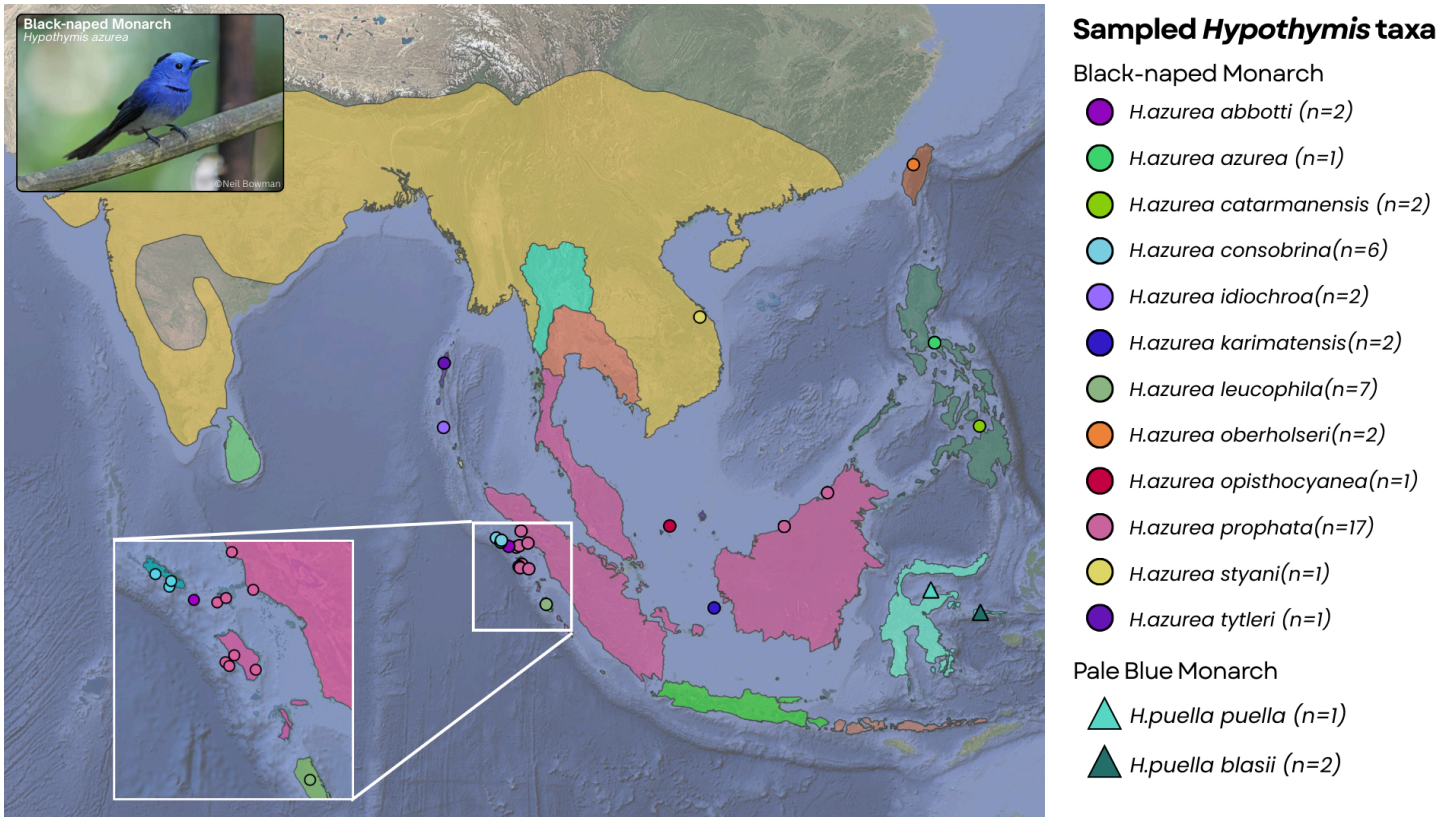


Figure 1. Distribution and sampling localities of 49 *Hypothymis* samples (45 samples of *H.azurea* and 3 samples of *H.puella*) used in this study. Samples are plotted with approximate localities. Samples from identical localities are overlaid and plotted with a single point. The ranges of unsampled subspecies are also included for reference. The map was constructed using the Quantum Geographic Information System v3.

2.2 DNA extraction and Sequencing:

DNA extraction was performed on both blood and tissue samples using the Qiagen DNeasy blood and tissue kit following the manufacturer's recommendations and protocol. DNA concentrations were measured using a Qubit 2.0 broad range DNA Assay Kit (Invitrogen, Waltham, MA, USA). Whole-genome libraries were prepared at BGI Genomics (Shanghai) using the DNBSEQ WGS library (BGI Genomics, Shenzhen, China) and sequenced at 10X coverage on the BGISEQ-500 platform (BGI Genomics, Shenzhen, China).

2.3 Sequence processing and SNP-calling:

For downstream analysis, the NUS IT department at the National University of Singapore offered access to their high-performance computing cluster (HPC), Atlas08. The quality of the sequenced reads was checked with FastQC v0.11.9 (Andrews, 2010). The adapter sequences from the sequenced reads were trimmed using cutadapt 4.2 (Martin, M. 2011) and then mapped to the genome of *Myiagra hebetior* (Genbank accession number GCA_013397015.1) using BWA mem 0.7.17 (Li and Durbin 2010). The SAMtools 1.16 (Li et al. 2009) program was used to combine multiple BAM files of each sample into one BAM file. Picard 2.16 (broadinstitute.github.io/picard) was then used to add the read groups to all the BAM files to aid in downstream analysis. The GATK 4.0.2.1(McKenna et al. 2010) was used to mark duplicate reads and sort each sample's BAM file.

A FASTA index file (.fai) of the *Myiagra hebetior* reference genome was generated using SAMtools 1.9 before indel realignment. Since the separate function for insertion-deletion realignment was discontinued in GATK 4.0.2.1, an older version, GATK 3.8 (McKenna et al. 2010), was used to correct for indels to avoid the pipeline from calling false variants during SNP calling. The tools RealignerTargetCreator and IndelRealigner in the GATK 3.8 (McKenna et al. 2010) program were used to generate interval files and perform the indel realignment (Van der Auwera & O'Connor 2020).

Qualimap (Fernando García-Alcalde et al. 2012) was used to generate and assess the alignment quality in each BAM file. Single-nucleotide polymorphisms(SNPs) were called using the ANGSD 0.935 program (Korneliussen, T.S. et al. 2014), where

parameter values were set based on the results of Qualimap (-geno_minDepth 5: SNPs will be called only when there's a minimum coverage of 5 reads at that position in the genome). The flag minInd was set to the total number of samples in the dataset, which ensured that SNPs were only included when present in all individuals within the dataset.

2.4 Phylogenetic analysis:

A maximum likelihood (ML) whole-genome SNP tree was generated by converting the VCF generated as part of the output SNPs dataset to PHYLIP format using vcf2phylip 2.8 (Ortiz, 2019). The phylip file was input into the IQ-TREE v2.2.2.6 to generate the concatenated phylogenetic SNP tree using the MFP+ASC model with 1000 bootstrap replicates. The resulting tree file was visualised in iTOL v6 (Letunic I and Bork P, 2024).

ML mitochondrial genome trees were constructed by first aligning the raw reads to the mitochondrial reference genome of *Myiagra hebetior* (Genbank accession number GCA_013397015.1) using BWA mem 0.7.17 (Li and Durbin 2010). The quality of the BAM files was assessed using Qualimap before converting them to FASTA format using ANGSD 0.935. The mitochondrial sequences of all the samples, including the outgroup, were aligned using MAFFT v7.407 to produce an aligned FASTA file. This file was then used as input for the IQ-TREE v2.2.2.6 to generate the mitochondrial genome tree using the MFP model with 1000 bootstrap replicates. The resulting tree file was visualised in iTOL v6.

A species tree was constructed using the Coalescent Aware Species Tree Estimator (CASTER) program (Zhang et al., 2025), which used the PHYLIP file, produced after SNP calling, as the input to generate a site-based species tree with 16 rounds of subsampling.

2.5 Population structure & gene flow analysis

A principal component analysis (PCA) is a preliminary analysis typically performed to infer the population structure of the taxon of interest. To perform PCA, output files from the SNP calling step were processed using PLINK 1.9 (www.cog-genomics.org/plink/1.9/, Chang et al. 2015) to generate eigenvalue and eigenvector files for Principal Component Analysis. An R script was used to plot and visualise the PCA. The PCA was performed on a dataset consisting of only *Hypothymis azurea* samples to better infer their summary population structure.
Sliding window analysis is typically performed to look for gene flow signals across different chromosomal regions among quartets (three ingroups and one outgroup). In the CASTER sliding window program (Zhang et al., 2025), gene flow signals are assessed across the genome by calculating the D statistic, which analyses frequencies of allelic patterns “ABBA/BABA” across the quartet. ABBA is when taxa 2 and 3 share a derived allele B at a certain locus, and BABA is when taxa 1 and 3 share a derived allele B at a certain locus in the chromosome sliding window. A CASTER sliding window was used to calculate the allele frequencies across 500kb sliding windows. These values were then used to plot stacked plots for multiple quartets to infer signals of gene flow between the samples. The quartets always included the outgroup *Corvus kubaryi*, along with triplet combinations chosen based on the resulting tree topology discordances and the PCA results.
To perform sliding-window analysis, a FASTA file for each sample was generated from the corresponding BAM file using doFasta in ANGSD 0.935 (Korneliussen, T.S. et al., 2014). The FASTA file of each sample consisted of reads with reference-aligned chromosome sequences of that sample. Each read corresponded to one out of 43 chromosomal regions in the sample’s sequence. These files were then used to filter out reads of chromosomal segments for specified samples. The ‘dstar’ command from the ASTER program (Zhang et al., 2025) was then used on these filtered reads to generate the sliding window output. The output values included the ABBA-BABA frequencies computed for each site within each 500-kilobase window and averaged across all sites in that window. These values were then used to plot the normalised moving average, which showed the normalised percentage of sites supporting a given topology for each subsequent 500kb window. Gene flow was determined to be present if the proportion of sites supporting any of the discordant tree topologies crossed the 50% mark on the stacked plot.

Chapter 3 Results

3.1 Principal component analysis:

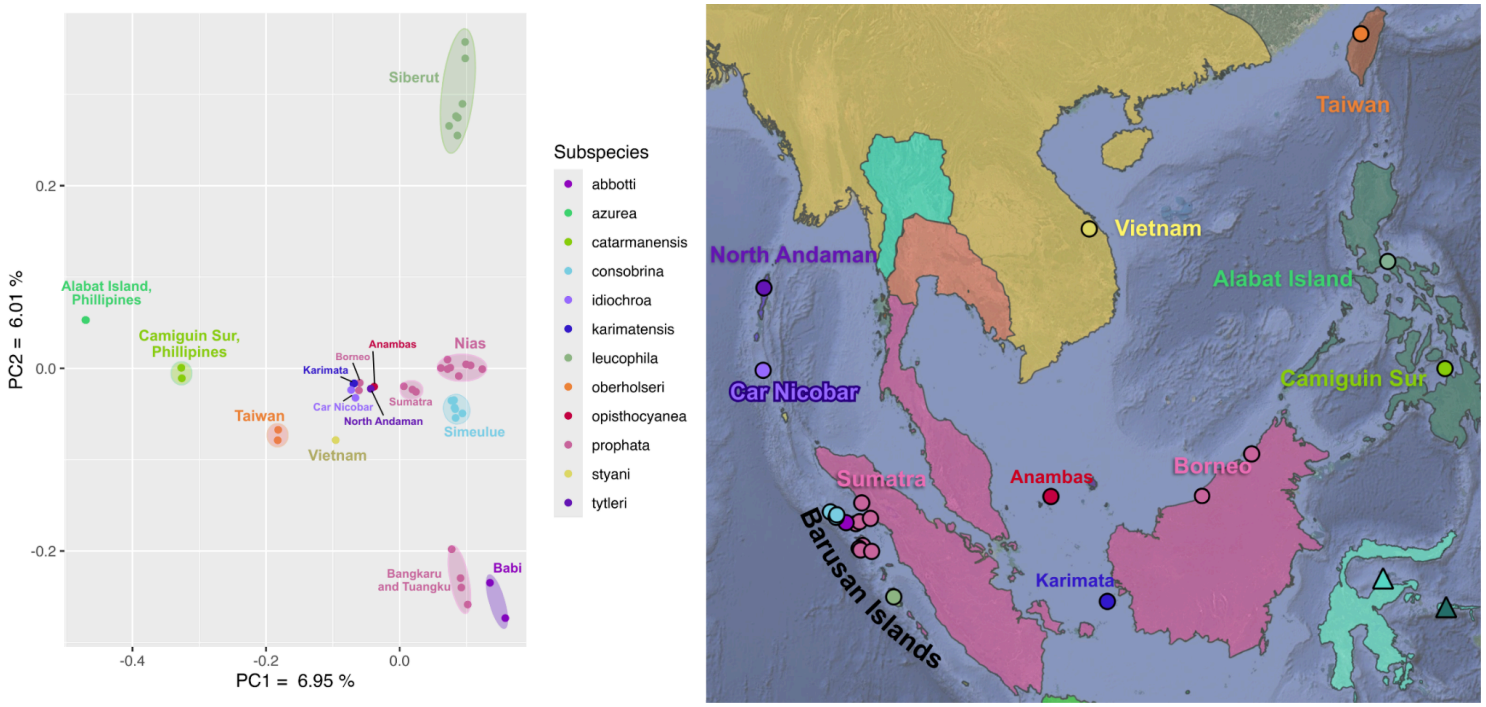


Figure 2. Overview summary of the population structure across *Hypothymis azurea* populations with labelled sampling localities.

In the PCA analysis based on 2,131,660 SNPs, many populations exhibit a clinal trend. The Barusan island forms (*H.a.prophata* from Bangkaru, Tuangku and Nias, *H.a.abbotti, H.a.consobrina* and *H.a.leucophila)* cluster following a clear geographical cline. The relatively northern populations (*H.a.styani, H.a.oberholseri, H.a.catarmanensis* and *H.a.azurea*) also follow a clinal distribution. However, the populations from the shelf islands of Sundaland (*H.a.karimatensis, H.a.prophata* from Sumatra, *H.a.opisthocyanea* and *H.a.prophata* from Borneo) and the Andaman and Nicobar islands (*H.a.tytleri* and *H.a.idiochroa*) seem to form a tight cluster nearer to the populations from Vietnam and Taiwan. Interestingly, these populations did not

exhibit clinal distribution despite being geographically distant from one another. The above results suggest that PCA should be complemented by phylogenetic and gene-flow analyses to understand evolutionary relationships and biogeography better.

3.2 Phylogenetic trees:

A whole-genome maximum likelihood SNP tree was constructed based on 2,871,412 SNPs. The ingroup *Hypothymis azurea* samples of the whole-genome SNP tree formed 8 main clades: (1) a Northern Barusan island clade comprising two subclades: the 'BBT clade' with samples from islands of Babi (*H.a.abbotti)* Bangkaru and Tuangku (*H.a.prophata)* and the Simeulue (*H.a.consobrina*) clade ; (2) a Southern Barusan island clade comprising two subclades: the Nias clade with *H.a.prophata* samples from Nias island and the Siberut clade with *H.a.leucophila* samples from Siberut island; (3) a clade with samples of *H.a.karimatensis* and a *H.a.prophata* sample from mainland Sumatra; (4) a clade with samples of *H.a.opisthocyanea*, *H.a.styani* and *H.a.prophata* (from Borneo and mainland Sumatra); (5) an Andaman island clade with *H.a.tytleri* and *H.a.idiohroa*; (6) a Philippines clade with *H.a.azurea* and *H.a.catarmanensis* samples emerging as sisters; (7) a clade with samples of *H.a.oberholseri*; and (8) a final clade comprising of samples of *Hypothymis puella.* The phylogenetic relationship among the Hypothymis samples reflects their geographic proximity. However, the third and fourth clades, with samples from Karimata, Vietnam, Anambas, Borneo, and mainland Sumatra, are not well resolved as a *H.a.prophata* sample from mainland Sumatra emerged as sister to *H.a.karimatensis* in a different clade from the other *H.a.prophata* monarchs of mainland Sumatra.

3.3 Mito-nuclear discordance across taxa:

There is very little congruence between the mitochondrial sequence tree and the whole-genome SNP tree, with topologies that differ in the phylogenetic positions of many taxa. The samples that emerged basal to the entire Barusan clade in the whole-genome SNP tree emerged as embedded within the Barusan clade in the mitochondrial sequence tree. For example, the Philippines clade (*H.a.azurea* and *H.a.catarmanensis*) that emerged basal to the Barusan clade in the whole-genome

SNP tree, emerged as paraphyletic in the mitochondrial sequence tree and was placed close to the Siberut island clade (*H.a.leucophila*) despite being geographically disjunct from each other. Another notable inconsistency is the *H.a.prophata* from the island of Tuangku emerging as a sister to the *H.a.karimatensis* clade in the mitochondrial sequence tree, whereas in the whole-genome tree, *H.a.karimatensis* was basal to the entire Barusan clade, and the *H.a.prophata* sample from Tuangku was within the Northern Barusan clade, embedded with the *H.a.prophata* from Bangkaru. Additionally, the sister taxon of many clades differed between the two trees, and a few monophyletic clades from the whole-genome tree emerged as paraphyletic in the mitochondrial sequence tree (the Andaman and Nicobar Islands clade, the Philippines clade and the Nias clade).

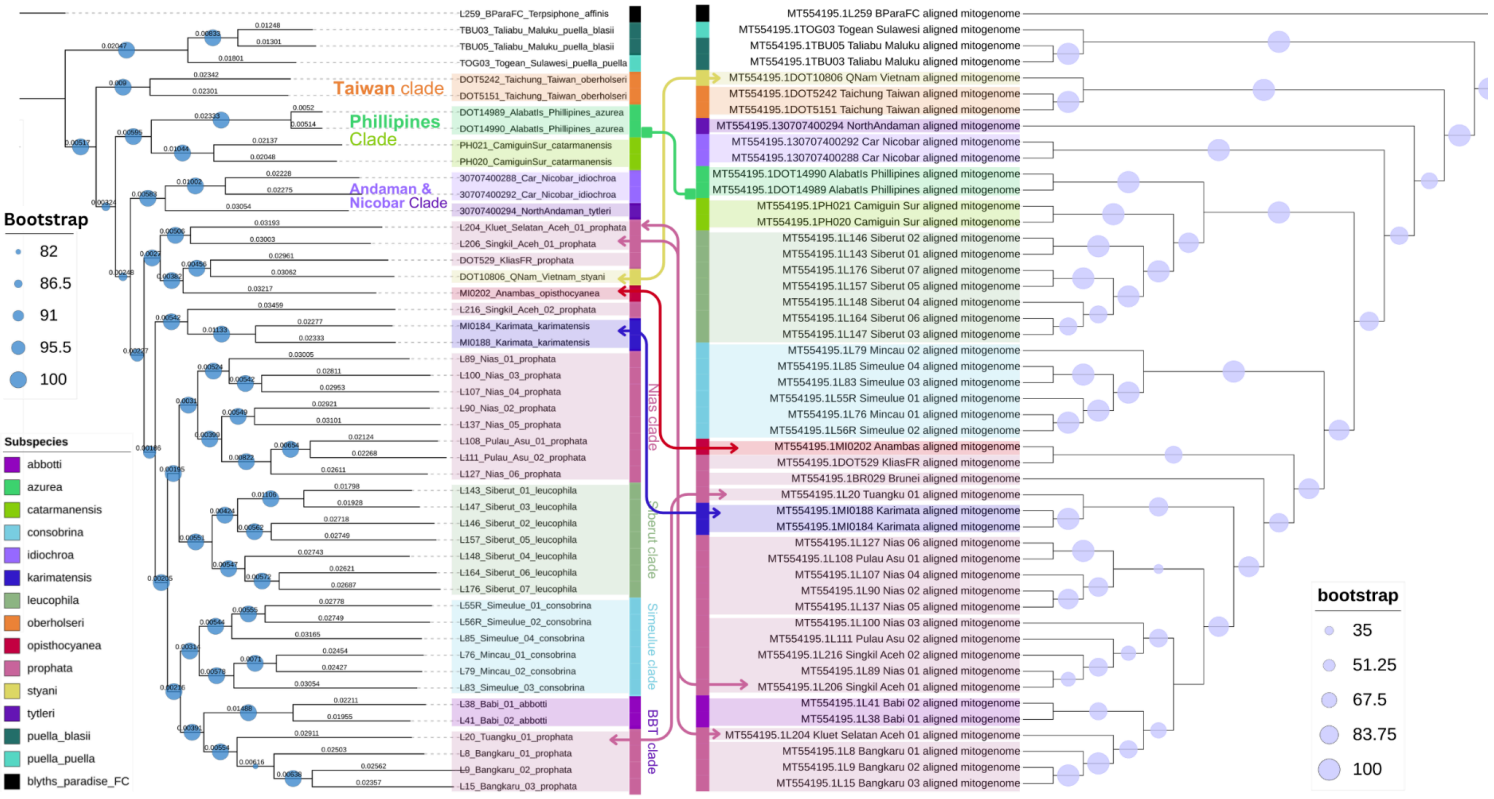


Figure 3. Maximum-likelihood tree topology based on 2,871,412 concatenated SNPs from whole-genome samples (left) and maximum-likelihood mitochondrial sequence tree based on 17,290 base pairs (right), run with 1000 bootstraps, with *Terpsiphone affinis* as the outgroup. All 49 samples are represented in the mitochondrial genome tree, while 48 samples are represented in the nuclear genome tree (BR029_Brunei is not included in the whole-genome concatenated SNP tree). The arrows depict clades exhibiting mito-nuclear discordance. The bootstrap legends were produced by iTOL v6 (Letunic I and Bork P, 2024). The arrows, taxon legends and clade labels were made using Canva Illustrator.

3.4 Species tree topology using Coalescent-Aware Species Tree Estimator:

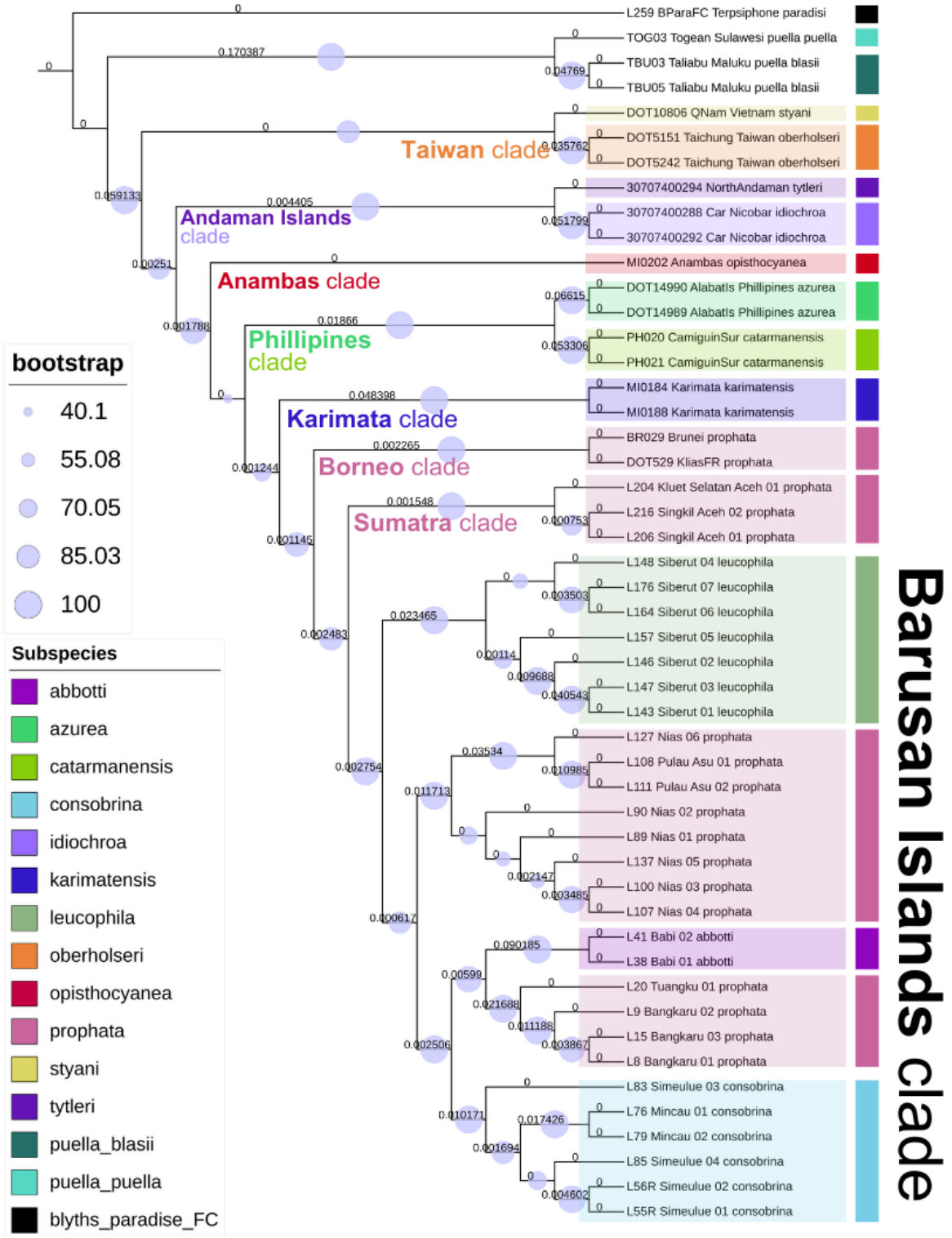


Figure 4. Site-based species tree topology of all *Hypothymis* samples rooted to *Terpsiphone affinis* as the outgroup. This site-based tree topology was generated using CASTER (Zhang et al., 2025). The bootstrap legend was produced by iTOL v6 (Letunic I and Bork P, 2024). The taxon legends and clade labels were made using Canva Illustrator.

The species tree constructed using the multispecies coalescent (MSC) model embeds multiple gene trees within a single species tree, allowing each locus to have

its own independent genealogy (B Rannala et al., 2020). On the other hand, a whole-genome concatenated SNP maximum-likelihood tree assumes a single genealogy for all sites, treating the sequences as a single long locus (Xiaodong Jian et al., 2019). Inferring from the SNP tree can be problematic when strong incomplete lineage sorting (ILS) is present. Therefore, a site-based species tree method, CASTER (Zhang et al., 2025), was employed to construct an alternative tree topology that accounts for incomplete lineage sorting.

The ingroup of the site-based species tree comprised 9 main clades: (1) The Barusan Islands clade with 4 subclades: the 'BBT' clade (*H.a.prophata* from Bangkaru islands and *H.a.abbotti*), the Simeulue clade (*H.a.consobrina*), the Nias clade (*H.a.prophata* from Nias) and the Siberut clade (*H.a.leucophila*). (2) The Sumatra clade (*H.a.prophata* from Sumatra); (3) The Borneo clade (*H.a.prophata* from Borneo); (4) The Karimata island clade (*H.a.karimatensis*); (5) The Philippines clade (*H.a.azurea* and *H.a.catarmanensis*); (6) The Anambas island clade (*H.a.opisthocyanea*); (7) The Andaman and Nicobar islands clade (*H.a.tytleri* and *H.a.idiochroa*); (8) the Taiwan-Vietnam clade (*H.a.oberholseri* and *H.a.styani*); and the final clade (9) the Pale Blue Monarch clade (*H.puella* and *H.puella blasii*). The phylogeny of the samples here more accurately reflects their geographic proximity than the whole-genome SNP tree. The samples that were embedded with geographically disjunct samples in the SNP tree emerged as distinct monophyletic clades in the site-based tree. Interestingly, the *H.a.styani* from Vietnam, which was previously embedded with the *Hypothymis* flycatchers from the Sundaland and the Andaman Islands in the SNP tree and was also observed to have very little genetic variation relative to these forms, emerged in the most basal clade, sister to the *H.a.oberholseri* from Taiwan. Additionally, across both the whole-genome SNP tree and the site-based species tree, the Barusan forms consistently formed a distinct clade. However, one instance of discordance was observed within the Barusan clade across both trees: The Siberut clade emerged basal to the rest of the Barusan forms in the site-based tree. In contrast, it emerged sister to the Nias clade in the whole-genome SNP tree. The Anambas clade, however, has very low bootstrap support as a basal clade and hence cannot be considered a strong result.

3.5 Sliding window analysis reveals gene flow signals between certain continental forms:

The sliding window analysis, based on the d* statistic, was visualised using the normalized moving -average stack plot. Each site in the genome has 3 possible topologies among three taxa A, B and C: ((A, B) C), ((A, C) B) and ((B, C) A). The y-axis represents the percentage of informative sites supporting each of the 3 topologies within a given window. The three topology frequency percentages sum up to 100%. The two discordant topologies (((A,C) B) and ((B,C) A)) are equal to each other under the expectation of incomplete lineage sorting and zero gene flow. Any deviation from this equality indicates a significant deviation from zero for the d*statistic values. This deviation is interpreted as a signal of gene flow.

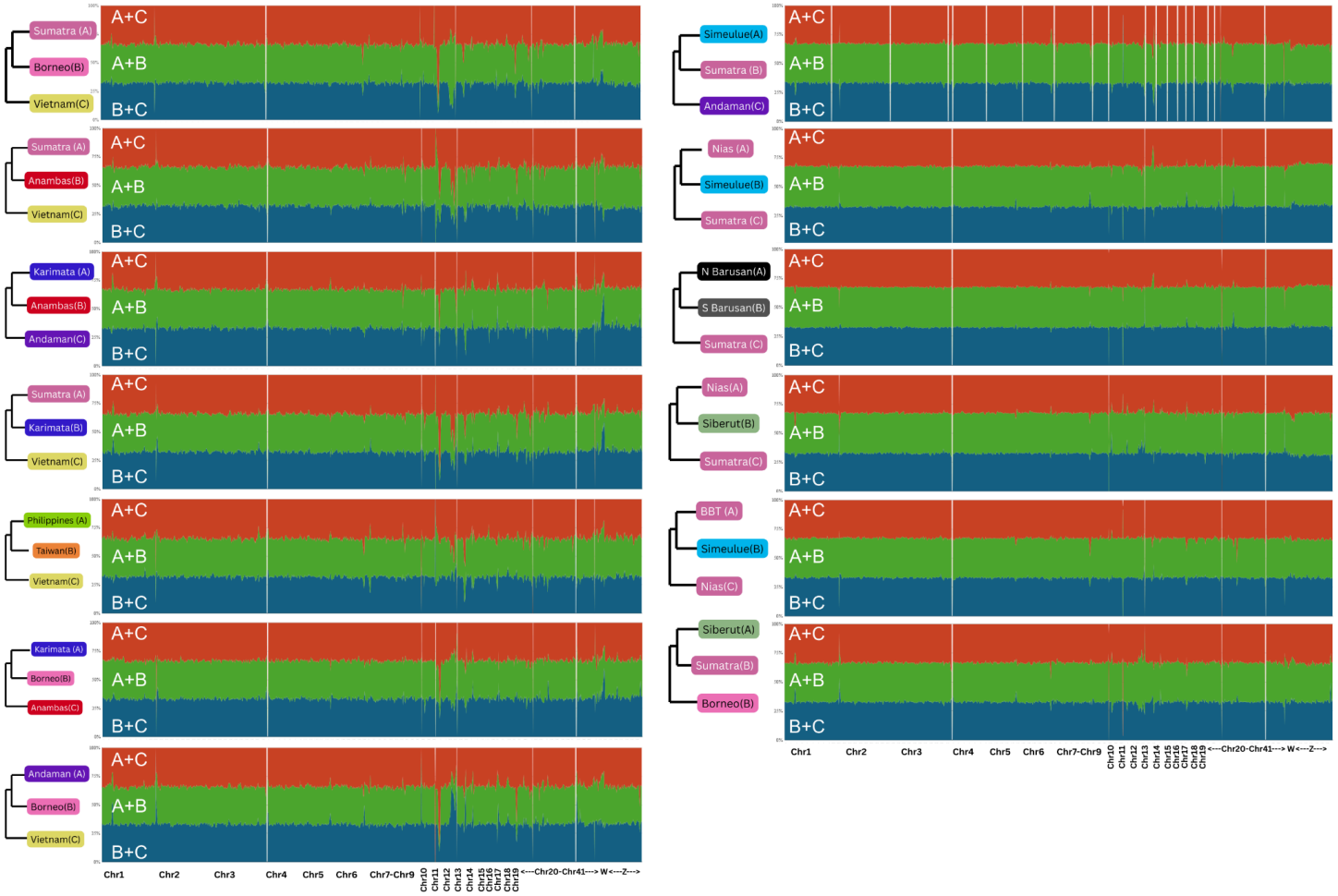


Figure 5. Normalised moving-average stacked plots for different population quartets (three ingroups and one outgroup). The red indicates the proportion of sites supporting the discordant topology ((A,C)B), the blue indicates the proportion of sites supporting the discordant topology ((B,C)A), and the green indicates the proportion of sites supporting the true topology ((A,B)C). The clade diagrams and chromosome labels were created using Canva Illustrator.

The stacked plots show that there exist strong gene flow signals between the *Hypothymis* forms from mainland Asia (*H.a.styani)*, Philippines (*H.a.azurea* and *H.a.catarmanensis*), Andaman and Nicobar islands(*H.a.tytleri* and *H.a.idiochroa*) and the Sunda Shelf (*H.a.prophata* from Sumatra and Borneo*, H.a.karimatensis, H.a.opisthocyanea).* These gene flow signals are mainly concentrated in chromosomes 11,12,13 and 14, with some gene flow signals also being detected in the sex chromosome W. Interestingly, significant gene flow signals remain absent among the deep-sea island forms (the Barusan clade). Gene flow between the deep-sea island and neighbouring Sunda Shelf forms (*H.a.prophata* from Sumatra and Borneo) is also inferred to be largely absent. However, the deep-sea island forms were observed to have very shallow divergence across all the tree topologies in our results.

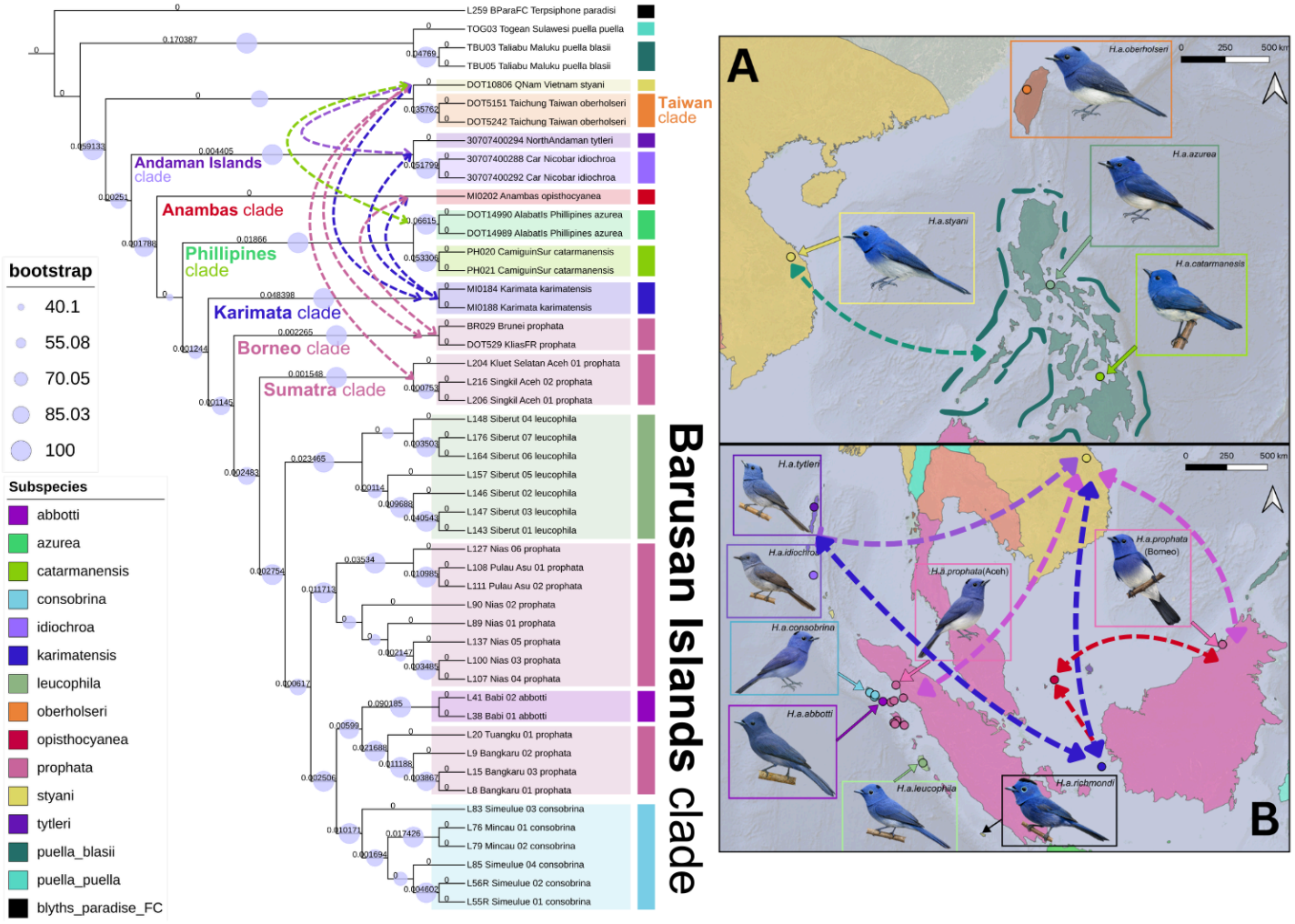


Figure 6. Site-based species tree topology with inferred gene flow arrows from the sliding window analysis (left) and sampling localities map with inferred gene flow arrows between approximate geographies (right). The inferred gene flow arrows were created in Canva Illustrator, and the map was constructed in Quantum Geographic Information System (QGIS) v3 (QGIS, 2024). The bird illustrations were created by ChatGPT using accurate image references for each subspecies and were then overlaid onto the map in Canva Illustrator.

Chapter 4 Discussion

Investigating the phylogeography of the pan-Asian songbird radiation of the *H.azurea* complex, our resulting tree topologies and population structure analysis show that geographical proximity and opportunities for gene flow drive the genomic structure of *H.azurea*. Genome-wide data reflect a mainland-shelf clade composed of numerous tropical Asian populations in the *H.azurea* complex, with a few island lineages (e.g., Andaman and Nicobar, Philippines, and Taiwan) basal to the clade with the Sunda shelf forms (e.g. Borneo, Sumatra, Karimata, Anambas). Within the mainland clade, we observe substantial gene flow among populations from mainland Asia and the Sunda Shelf, a collection of large islands that were regularly connected by land bridges during ice ages (episodes of global cooling) and Pleistocene sea-level fluctuations. As an exception, populations from the Barusan islands consistently emerged as a single clade embedded within this mainland clade; however, with no discernible gene flow to adjacent mainland/Sundaic populations.

The Barusan chain is viewed as uplifted arc fragments or oceanic islands that remained biogeographically distinct even during periods of low sea levels (L. Heaney, 1986; L. Husson et al, 2020). This explains why phenotypically dissimilar Barusan island populations consistently formed a cohesive clade despite being spatially adjacent, likely due to deep-sea barriers between these islands and the Sunda shelf, which limits opportunities for gene flow with the continental forms, as previous avian studies suggest (J. Pujolar et al.,2022; A. Radu et al.,2024). Today, the phenotypically most divergent subspecies of H. azurea reside on these deep-sea islands, where a lack of continental gene flow has allowed them to acquire these phenotypic differences. This finding is consistent with recent evolutionary studies showing that populations with shallow yet genome-wide divergence maintain pronounced phenotypic divergence (Amaral et al., 2018).

On the other hand, the observed gene flow in the continental and Sundaic forms suggests historical connectivity of these populations. This observation is likely a result of Pleistocene sea-level fluctuations that connected the Sunda Shelf to mainland Asia via exposed land bridges, thereby providing these currently disjunct

populations ample opportunities for gene flow (Cros et al., 2020; H.C. Lim et al., 2020; K. Garg et al., 2021). Continual gene flow in these populations explains their poor genetic structure and overall clinal distribution.

While deep-sea ridges form strong barriers to gene flow between insular avian populations, mainland songbird populations have also exhibited similar levels of divergence despite land connections. An increasing body of evidence suggests that such differentiations arose through ancient paleorivers, which functioned as geographical barriers comparable to open seas (Nascimento et al., 2013; Sholihah et al., 2021). These paleorivers, which were once spread across Sundaland during low sea levels in the Pleistocene, separated the landmass into distinct biogeographic regions and have been shown to act as barriers to dispersal and gene flow in a few rainforest songbirds (K. Garg et al., 2021). But in the case of the *H.azurea*, our results show that there is ample gene flow between Sunda Shelf populations that were once separated by Paleorivers, implying that the inland paleorivers did not form strong enough barriers and *H. azurea* were able to disperse across them. We speculate that, over evolutionary time, the Barusan island clades may diverge into distinct island-endemic species.

Gene flow signals in these continental and shelf-island populations have been observed to be concentrated in specific genomic regions, such as chromosomes 11-14. There could be a few reasons for this to be the case: (1) Selection pressure on genomic regions involved in local adaptation could have possibly removed the introgressed alleles, so gene flow signals in those regions become absent (R. Harrison et al., 2016); (2) Recombination rate, in most cases, tends to vary with chromosome size where smaller chromosomes tend to have higher recombination rates (Martin et al., 2018). In contrast, we observed no discernible gene flow signals in the smaller chromosomes (chromosomes 20- 41), which is unexpected. This could presumably be attributed to the previous reason, where positive selection acts as a post-mating genomic barrier to gene flow. (3) Demography and geography of the species can drive localised introgression on certain chromosomes even without strong selection, especially in cases of pronounced drift (Fraïsse et al., 2016; Seixas et al., 2018). Further investigation into the role of gene flow in shaping the genomic landscape of this species should be carried out to conclusively deduce the exact reason for the observed patterns of gene flow. Our analysis method used in this

study could only qualitatively determine gene flow signals, and hence, we aim to perform more rigorous quantitative analyses to quantify the amount of gene flow between the populations and, additionally, detect any gene flow from any unsampled *H.azurea* populations.

The strong mitochondrial-nuclear discordance observed in our results can be attributed to the fact that the mitochondrial genome is maternally inherited and hence does not reliably represent true phylogenetic relationships as apparent historical gene flow between the continental and shelf island populations could likely be leading to mitochondrial capture (Rheindt & Edwards, 2011, S. Perea et al., 2016; T. Bonnet et al., 2017, M. Andersen et al., 2021). Additionally, the literature suggests that mito-nuclear discordance may arise solely from incomplete lineage sorting rather than gene flow (D. DeRaad et al., 2023). Further investigation of this aspect is required to determine the exact cause of this discordance.

Furthermore, it is interesting to note that the discordance in the position of the Vietnamese *Hypothymis azurea* population (*H.a.styani*) between the site-based species tree and the whole-genome SNP tree can be well explained by the gene flow between *H.a.styani* and all other continental and shelf island populations, which likely led many SNP sites of the genome across these samples to share similarity with the continental and Sundaic forms. This discrepancy in phylogeny does not appear in the site-based tree, presumably because each gene is allowed to have its own genealogy, and the species tree is constructed from the consensus of many gene trees and genealogies to explain the topology of the study species best.

Our findings also show that the subspecies *H.a.prophata* is paraphyletic across our sampled populations, especially with *H.a.prophata* from Nias, Borneo and Sumatra emerging as distinct clades in our phylogenomic analyses. Hence, the taxonomic classification of the *prophata* subspecies needs to be revisited and revised accordingly.

Overall, our findings show that the *H.azurea* complex is an intricate avian radiation across insular and continental, tropical and subtropical Asia, branching into numerous phenotypically distinct forms. Our gene flow analyses demonstrate that

populations across the continent and the Sunda Shelf have remained relatively cohesive. In contrast, populations on deep-sea islands experienced a cessation of gene flow, leading to accelerated phenotypic differentiation. Some of these distinct phenotypic small-island forms are not necessarily older than other subspecies lineages from the Sunda Shelf and the mainland continent, but what sets them apart is a lack of continual gene flow with mainland populations.

References

Amaral, F., Maldonado-Coelho, M., Aleixo, A., Luna, L., Rêgo, P., Araripe, J., Souza, T., Silva, W., & Thom, G. (2018). Recent chapters of Neotropical history overlooked in phylogeography: Shallow divergence explains phenotype and genotype uncoupling in *Antilophia* manakins. *Molecular Ecology*, 27, 4108–4120. https://doi.org/10.1111/mec.14843

Andersen, M., McCullough, J., Gyllenhaal, E., Mapel, X., Haryoko, T., Jønsson, K., & Joseph, L. (2021). Complex histories of gene flow and a mitochondrial capture event in a nonsister pair of birds. *Molecular Ecology*, 30, 2087–2103. https://doi.org/10.1111/mec.15856

Beck, S., Carvalho, G., Barlow, A., Rüber, L., Tan, H., Nugroho, E., Wowor, D., Nor, S., Herder, F., Muchlisin, Z., & De Bruyn, M. (2017). Plio-Pleistocene phylogeography of the Southeast Asian Blue Panchax killifish, *Aplocheilus panchax*. *PLoS ONE*, 12. https://doi.org/10.1371/journal.pone.0179557

Bonnet, T., Leblois, R., Rousset, F., & Crochet, P. (2017). A reassessment of explanations for discordant introgressions of mitochondrial and nuclear genomes. *Evolution*, 71. https://doi.org/10.1111/evo.13296

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4(1). https://doi.org/10.1186/s13742-015-0047-8

Cros, E., Chattopadhyay, B., Garg, K., Ng, N., Tomassi, S., Benedick, S., Edwards, D., & Rheindt, F. (2020). Quaternary land bridges have not been universal conduits of gene flow. *Molecular Ecology*, 29, 2692–2706. https://doi.org/10.1111/mec.15509

DeRaad, D., McCullough, J., DeCicco, L., Hime, P., Joseph, L., Andersen, M., & Moyle, R. (2023). Mitonuclear discordance results from incomplete lineage sorting, with no detectable evidence for gene flow, in a rapid radiation of *Todiramphus* kingfishers. *Molecular Ecology*, 32, 4844–4862. https://doi.org/10.1111/mec.17080

den Tex, R.-J., & Leonard, J. A. (2014). The phylogeography of red and yellow coppersmith barbets (*Aves*: *Megalaima haemacephala*). *Frontiers in Ecology and Evolution*, 2:16. https://doi.org/10.3389/fevo.2014.00016

Fernando García-Alcalde et al. (2012). Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics*, 28(20), 2678–2679. https://doi.org/10.1093/bioinformatics/bts503

Fraïsse, C., Belkhir, K., Welch, J., & Bierne, N. (2016). Local interspecies introgression is the main cause of extreme levels of intraspecific differentiation in mussels. *Molecular Ecology*, 25. https://doi.org/10.1111/mec.13299

Garg, K., Chattopadhyay, B., Cros, E., Tomassi, S., Benedick, S., Edwards, D., & Rheindt, F. (2021). Island Biogeography Revisited: Museomics Reveals Affinities of Shelf Island Birds Determined by Bathymetry and Paleo-Rivers, Not by Distance to Mainland. *Molecular Biology and Evolution*, 39. https://doi.org/10.1093/molbev/msab340

González-Porter, G. P., Hailer, F., Flores-Villela, O., et al. (2011). Patterns of genetic diversity in the critically endangered Central American river turtle: human influence since the Mayan age? *Conservation Genetics*, 12, 1229–1242. https://doi.org/10.1007/s10592-011-0225-x

Harrison, R., & Larson, E. (2016). Heterogeneous genome divergence, differential introgression, and the origin and structure of hybrid zones. *Molecular Ecology*, 25. https://doi.org/10.1111/mec.13582

Heaney, L. (1986). Biogeography of mammals in SE Asia: estimates of rates of colonization, extinction and speciation. *Biological Journal of the Linnean Society*, 28, 127–165. https://doi.org/10.1111/j.1095-8312.1986.tb01752.x

Heaney, L. R. (1991). A synopsis of climatic and vegetational change in Southeast Asia. *Climatic Change*, 19, 53–61. https://doi.org/10.1007/BF00142213

Husson, L., Boucher, F., Sarr, A., Sepulchre, P., & Cahyarini, S. (2020). Evidence of Sundaland's subsidence requires revisiting its biogeography. *Journal of Biogeography*, 47, 843–853. https://doi.org/10.1111/jbi.13762

Ivica Letunic & Peer Bork (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, 52(W1), W78–W82. https://doi.org/10.1093/nar/gkae268

Koblmüller, S., Katongo, C., Phiri, H., & Sturmbauer, C. (2012). Past connection of the upper reaches of a Lake Tanganyika tributary with the upper Congo drainage suggested by genetic data of riverine cichlid fishes. *African Zoology*, 47(1), 182–186. https://doi.org/10.1080/15627020.2012.11407537

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, 15, 356. https://doi.org/10.1186/s12859-014-0356-4

Li, F., & Li, S. (2018). Paleocene–Eocene and Plio–Pleistocene sea-level changes as "species pumps" in Southeast Asia: Evidence from *Althepus* spiders. *Molecular Phylogenetics and Evolution*, 127, 545–555. https://doi.org/10.1016/j.ympev.2018.05.014

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, 26(5), 589–595. https://doi.org/10.1093/bioinformatics/btp698

Li, H., Handsaker, B., Wysoker, A., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lim, H., Shakya, S., Harvey, M., Moyle, R., Fleischer, R., Braun, M., & Sheldon, F. (2020). Opening the door to greater phylogeographic inference in Southeast Asia. *Ecology and Evolution*, 10, 3222–3247. https://doi.org/10.1002/ece3.5964

Lohman, D. J., Ingram, K. K., Prawiradilaga, D. M., Winker, K., Sheldon, F. H., Moyle, R. G., & Meier, R. (2010). Cryptic genetic diversity in "widespread" Southeast Asian bird species suggests that Philippine avian endemism is gravely underestimated. *Biological Conservation*, 143(8), 1885–1890.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1), 10–12.

Martin, S., Davey, J., Salazar, C., & Jiggins, C. (2018). Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biology*, 17. https://doi.org/10.1371/journal.pbio.2006288

McKenna, A., Hanna, M., Banks, E., et al. (2010). The genome analysis toolkit: a MapReduce framework for analysing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Nascimento, F., Lazar, A., Menezes, A., et al. (2013). The Role of Historical Barriers in the Diversification Processes in Open Vegetation Formations during the Miocene/Pliocene Using an Ancient Rodent Lineage as a Model. *PLoS ONE*, 8. https://doi.org/10.1371/journal.pone.0061924

NCBI (2020). *Myiagra hebetior* genome assembly (GCA_013397015.1_ASM1339701v1). https://www.ncbi.nlm.nih.gov/assembly/GCA_013397015.1

Ortiz, E. M. (2019). vcf2phylip v2.0. Zenodo. https://doi.org/10.5281/zenodo.2540861

Othman, S., Chen, Y., Chuang, M., Andersen, D., Jang, Y., & Borzée, A. (2020). Impact of the Mid-Pleistocene Revolution and Anthropogenic Factors on the Dispersion of Asian Black-Spined Toads (*Duttaphrynus melanostictus*). *Animals*, 10. https://doi.org/10.3390/ani10071157

Perea, S., Vukić, J., Šanda, R., & Doadrio, I. (2016). Ancient mitochondrial capture as factor promoting mitonuclear discordance. *PLoS ONE*, 11. https://doi.org/10.1371/journal.pone.0166292

Prasetya, A., Jaya, F., Moritz, C., Joseph, L., & Oliver, P. (2025). Supermatrix phylogenetic tree of passerine birds from the Indo-Australian Archipelago highlights contrasting histories of regional endemism. *Ecology and Evolution*, 15. https://doi.org/10.1002/ece3.71471

Pujolar, J., Blom, M., Reeve, A., et al. (2022). The formation of avian montane diversity across barriers and along elevational gradients. *Nature Communications*, 13. https://doi.org/10.1038/s41467-021-27858-5

Radu, A., Dudgeon, C., Clegg, S., Foster, Y., Levengood, A., Sendell-Price, A., Townsend, K., & Potvin, D. (2024). Genetic patterns reveal geographic drivers

of divergence in silvereyes (*Zosterops lateralis*). *Scientific Reports*, 14. https://doi.org/10.1038/s41598-024-71364-9

Rannala, B., Edwards, S., Leaché, A., & Yang, Z. (2020). The multi-species coalescent model and species tree inference.

Rheindt, F. E., & Edwards, S. V. (2011). Genetic introgression: An integral but neglected component of speciation in birds. *The Auk*, 128(4), 620–632. https://doi.org/10.1525/auk.2011.128.4.620

Seixas, F., Boursot, P., & Melo-Ferreira, J. (2018). The genomic impact of historical hybridisation with massive mitochondrial DNA introgression. *Genome Biology*, 19. https://doi.org/10.1186/s13059-018-1471-8

Sheldon, F. H., Lim, H. C., & Moyle, R. G. (2015). Return to the Malay Archipelago: the biogeography of Sundaic rainforest birds. *Journal of Ornithology*, 156(Suppl 1), 91–113. https://doi.org/10.1007/s10336-015-1188-3

Sholihah, A., Delrieu-Trottin, E., Condamine, F., et al. (2021). Impact of Pleistocene Eustatic Fluctuations on Evolutionary Dynamics in Southeast Asian Biodiversity Hotspots. *Systematic Biology*. https://doi.org/10.1093/sysbio/syab006

Van der Auwera, G. A., & O'Connor, B. D. (2020). *Genomics in the cloud: using Docker, GATK, and WDL in Terra*. O'Reilly Media

Woodruff, D. (2010). Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. *Biodiversity and Conservation*, 19, 919–941. https://doi.org/10.1007/s10531-010-9783-3

Zhang, C., Nielsen, R., & Mirarab, S. (2025). ASTER: A package for large-scale phylogenomic reconstructions. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msaf172

Appendix:

Table A1. All samples of *H.azurea* and *H.puella* with subspecies and sampling locality information

ID	Genus	Species	Subspecies	Locality	Source of tissue
DOT 529	Hypothymis	azurea	*prophata*	Klias Forest Reserve	AMNH field collections
DOT 5151	Hypothymis	azurea	*oberholseri*	Taichung, Taiwan	AMNH field collections
DOT 5242	Hypothymis	azurea	*oberholseri*	Taichung, Taiwan	AMNH field collections
DOT 10806	Hypothymis	azurea	*styani*	Quang Nam, Vietnam	AMNH field collections
DOT 14989	Hypothymis	azurea	*azurea*	Mt Camagong, Alabat Island, Philippines	AMNH field collections
DOT 14990	Hypothymis	azurea	*azurea*	Mt Camagong, Alabat Island, Philippines	AMNH field collections
TBU5	Hypothymis	puella	*blasii*	Desa Wahe, Taliabu, Maluku Utara	Avian Lab field collections
TBU3	Hypothymis	puella	*blasii*	Desa Wahe, Taliabu, Maluku Utara	Avian Lab field collections
TOG 03	Hypothymis	puella	*puella*	Togean, Sulawesi Tengah	Avian Lab field collections
BN-MONA01	Hypothymis	azurea	*prophata*	Bangkaru	Avian Lab field collections
BN-MONA02	Hypothymis	azurea	*prophata*	Bangkaru	Avian Lab field collections
BN-MONA03	Hypothymis	azurea	*prophata*	Bangkaru	Avian Lab field collections
BN-MONA04	Hypothymis	azurea	*prophata*	Tuangku	Avian Lab field collections
SIM12	Hypothymis	azurea	*abbotti*	Babi	Avian Lab field collections
SIM15	Hypothymis	azurea	*abbotti*	Babi	Avian Lab field collections
SIM35	Hypothymis	azurea	*consobrina*	Simeulue	Avian Lab field collections
SIM36	Hypothymis	azurea	*consobrina*	Simeulue	Avian Lab field collections
SIM55	Hypothymis	azurea	*consobrina*	Mincau	Avian Lab field collections
SIM58	Hypothymis	azurea	*consobrina*	Mincau	Avian Lab field collections
SIM65	Hypothymis	azurea	*consobrina*	Simeulue	Avian Lab field collections

SIM67	Hypothymis	azurea	*consobrina*	Simeulue	Avian Lab field collections
NIA05	Hypothymis	azurea	*prophata*	Onolimbu, Nias	Avian Lab field collections
NIA06	Hypothymis	azurea	*prophata*	Onolimbu, Nias	Avian Lab field collections
NIA27	Hypothymis	azurea	*prophata*	Onolimbu, Nias	Avian Lab field collections
NIA39	Hypothymis	azurea	*prophata*	Onolimbu, Nias	Avian Lab field collections
NIA41	Hypothymis	azurea	*prophata*	Pulau Asu	Avian Lab field collections
NIA44	Hypothymis	azurea	*prophata*	Pulau Asu	Avian Lab field collections
NIA59	Hypothymis	azurea	*prophata*	Bawa	Avian Lab field collections
NIA69	Hypothymis	azurea	*prophata*	Bawolato	Avian Lab field collection
SIB01	Hypothymis	azurea	*leucophila*	Siberut	Avian Lab field collections
SIB04	Hypothymis	azurea	*leucophila*	Siberut	Avian Lab field collections
SIB05	Hypothymis	azurea	*leucophila*	Siberut	Avian Lab field collections
SIB06	Hypothymis	azurea	*leucophila*	Siberut	Avian Lab field collections
SIB19	Hypothymis	azurea	*leucophila*	Siberut	Avian Lab field collections
SIB25	Hypothymis	azurea	*leucophila*	Siberut	Avian Lab field collections
SIB44	Hypothymis	azurea	*leucophila*	Siberut	Avian Lab field collections
KLT23	Hypothymis	azurea	*prophata*	Kluet Seletan, Sumatra	Avian Lab field collections
KBR02	Hypothymis	azurea	*prophata*	Singkil, Sumatra	Avian Lab field collections
KBR28	Hypothymis	azurea	*prophata*	Singkil, Sumatra	Avian Lab field collections
PH020	Hypothymis	azurea	*catarmanensis*	Camiguin Sur, Philippines	Avian Lab field collections
PH021	Hypothymis	azurea	*catarmanensis*	Camiguin Sur, Philippines	Avian Lab field collections
BR029	Hypothymis	azurea	*prophata*	Brunei	Avian Lab field collections
MI0184	Hypothymis	azurea	*karimatensis*	Karimata Island	Avian Lab field collections

MI0188	Hypothymis	azurea	*karimatensis*	Karimata Island	Avian Lab field collections
MI0202	Hypothymis	azurea	*opisthocyanea*	Anambas Island	Avian Lab field collections
30707400288_Car_Nicobar	Hypothymis	azurea	*idiochroa*	Car Nicobar Island	Shashank Dalvi field collections
30707400292_Car_Nicobar	Hypothymis	azurea	*idiochroa*	Car Nicobar Island	Shashank Dalvi field collections
30707400294_NorthAndaman	Hypothymis	azurea	*tytleri*	North Andaman	Shashank Dalvi field collections