

A Genomic Perspective on Cryptic Species Reveals Complex Evolutionary Dynamics in the Gray Zone of the Speciation Continuum

KIN ONN CHAN^{1,2,3,4,*}, DARIO N. NEOKLEOUS⁵, SHAHRUL ANUAR⁶, RAFE M. BROWN⁴,
 CARL R. HUTTER⁷, INDRANEIL DAS⁸, AND STEFAN T. HERTWIG^{5,9}

¹Department of Integrative Biology, Michigan State University, 288 Farm Lane, East Lansing, MI 48824, USA

²Ecology, Evolution, and Behavior Program, Michigan State University, 567 Wilson Road, East Lansing, MI 48824, USA

³MSU Museum, Michigan State University, 409 W Circle Drive, East Lansing, MI 48824, USA

⁴KU Biodiversity Institute and Natural History Museum, University of Kansas, 1345 Jayhawk Blvd, Lawrence, KS 66045, USA

⁵Naturhistorisches Museum der Burgergemeinde Bern, Bernastrasse 15, CH-3005 Bern, Switzerland

⁶School of Biological Sciences, Universiti Sains Malaysia, 1112 Persiaran Sains, 11800 Gelugor, Penang, Malaysia

⁷Virus and Prion Research Unit, National Animal Disease Center, USDA-ARS, 1920 Dayton Avenue, Ames, IA 50010, USA

⁸Institute of Biodiversity and Environmental Conservation, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

⁹Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, CH-3012 Bern, Switzerland

*Correspondence to be sent to: Department of Integrative Biology, Michigan State University, 288 Farm Lane, East Lansing, MI 48824, USA; E-mail: cko@msu.edu.

Received 4 June 2025; reviews returned 20 December 2025; accepted 30 December 2025

Associate Editor: Bryan Carstens

Abstract.—The evolutionary dynamics of cryptic species remain poorly understood, and their detection relies primarily on methods that quantify divergence, assuming that gene flow is absent. Here, we examine how gene flow shapes the evolutionary trajectories and species boundaries in Bornean Fanged Frogs, a renowned example of cryptic diversity where a single species has been split into 18 genetically divergent yet morphologically indistinguishable species. We employed target-capture data from over 13,000 loci to assess lineage independence of 14 nominal species distributed across Malaysian Borneo by evaluating both divergence and cohesion using network multispecies coalescent (NMSC) and MSC + migration approaches. Under the Unified Species Concept, only 6 of the 14 nominal species unambiguously form independently evolving lineages; the remainder represent cohesive metapopulation lineages nested within those six species. While mitochondrial *p*-distances varied substantially (up to 10%), genome-wide net divergences (D_a) were surprisingly consistent, ranging from 0.5 to 2%, placing all the hypothesized “cryptic species” within the empirical gray zone of the speciation continuum. We show that diversification in the gray zone is unpredictable and heavily impacted by gene flow, leading to two key phenomena that confound species delimitation: (1) the *artificial branch effect*, where admixed lineages are inferred as long, early-diverging branches, creating an illusion of deep divergence; and (2) the *species-definition anomaly zone*, where intraspecific pairwise sequence distances exceed interspecific ones. We further demonstrate that divergence in the gray zone varies substantially among metrics and genomic regions, reflecting heterogeneity in evolutionary dynamics across the genome. Different genomic markers also vary considerably in phylogenetic discordance and their ability to retain signatures of gene flow. Loci from anchored hybrid enrichment and ultraconserved elements produced less phylogenetic discordance and retained signals of older introgression but failed to detect recent migration, making them suitable for phylogenetic reconstruction and inferring ancient introgression, but not ongoing gene flow. Recognizing the central role of gene flow reframes our understanding of cryptic species; rather than being considered as genetically distinct units that failed to evolve morphological differentiation, they can be manifestations of continuous diversification in the gray zone. This shift in perspective offers a new and dynamic evolutionary framework for identifying and interpreting cryptic biodiversity across the Tree of Life. [Amphibians; artifactual branch effect; diversification; gene flow; reticulate evolution; species definition anomaly zone; species delimitation.]

Cryptic species are recognized to be prevalent across the Tree of Life, representing significant yet often overlooked components of biodiversity (Pfenninger and Schwenk 2007; Trontelj et al. 2009). Over the past two decades, molecular phylogenetic studies have alluded to a staggering proportion of undescribed cryptic diversity, leading to dramatic revisions of global biodiversity estimates. Older models, which did not account for cryptic species, suggested that Earth hosts between 2 and 9 million species (Mora et al. 2011). In stark contrast, more recent models incorporating undescribed cryp-

tic diversity estimate that the total number of species could surpass two billion (Li and Wiens 2023), potentially reshaping our understanding of Earth’s biodiversity. However, despite their purported prevalence, fundamental questions about cryptic species are unresolved. Their evolutionary history, the processes driving their diversification, the mechanisms by which they persist as separate evolutionary units, and the methodologies used to detect them are all subjects of ongoing debate (Trontelj et al. 2009; Fišer et al. 2018; Heethoff 2018; Struck et al. 2018; Chenuil et al. 2019).

This uncertainty raises important questions about the *raison d'être* of cryptic species, their roles in ecosystems, and their broader implications for biodiversity and conservation.

At its core, the definition of cryptic species remains ambiguous and contentious, with no universal consensus. Over time, numerous definitions have been proposed and used interchangeably (Shin and Allmon 2023). Nevertheless, cryptic species are generally understood as genetically divergent evolutionary lineages that are morphologically difficult to differentiate (Chenuil et al. 2019). This dual-criteria framework of genetic divergence and morphological similarity has fueled a surge of studies uncovering hidden cryptic diversity across many organisms (Adams et al. 2014; Struck et al. 2018; Cerca et al. 2020; Wattier et al. 2020). However, this simplistic definition introduces significant challenges, particularly with the inconsistent and subjective quantification of genetic and morphological differentiation. Which metrics should be used to measure genetic divergence? What level of genetic divergence separates populations from species? Which axes of morphological variation should be considered when assessing similarity? Adding to this complexity, speciation is usually viewed as a gradual and continuous process (Sukumaran and Knowles 2017; Stankowski and Ravinet 2021), and thus, cryptic species may also exist at various points along the speciation continuum, exhibiting varying degrees of “crypticity” in genetic and morphological divergence. Adopting this protracted view of speciation complicates the distinction between cryptic, pseudo-cryptic, and noncryptic species (Reis et al. 2020; Korshunova and Martynov 2024). Taken together, these issues underscore the need for a more robust and nuanced understanding of the evolutionary history and dynamics of cryptic species.

Within the framework of the Unified Species Concept (de Queiroz 2005), the definition of cryptic species implicitly assumes that genetically divergent clades represent independent evolutionary lineages, that is, species. However, in most studies, the hypothesis of evolutionary independence is rarely tested explicitly, leading to a disconnect between species concepts and operational methodologies (de Queiroz 2005). In practice, many cryptic species studies focus on quantifying genetic divergence to characterize lineage independence (Rannala and Yang 2020), often relying on proxies such as genetic distance (Puillandre et al. 2021), tree-based species delimitation algorithms (Fujisawa et al. 2016; Kapli et al. 2017; Cerca et al. 2020), or site-pattern differences (Jörger and Schrödl 2013). Diversification is rarely a discrete process, especially at the early stages of population differentiation. Instead, it is shaped by the interplay between two opposing forces: divergence and cohesion. These antagonistic processes jointly influence lineage splitting and fusion, generating a dynamic push-and-pull tension along the speciation continuum where species boundaries remain semiperme-

able (Seehausen et al. 2014; Stankowski and Ravinet 2021). This permeability can result in a chimeric assemblage of interconnected populations, where elements of divergence are intertwined with gene flow, introgression, and hybridization—hallmarks of the “gray zone” of speciation (Roux et al. 2016). As such, exclusively focusing on divergence produces a one-sided perspective on evolution, often leading to the overestimation of species numbers in empirical studies (Chan et al. 2022b). To overcome this bias, it is necessary to integrate assessments of both divergence and cohesion into species delimitation frameworks (Chan et al. 2022b; Jackson et al. 2017; Singhal et al. 2018, 2025; Prates et al. 2024). This approach not only mitigates the shortcomings of tree/distance-based methods but is also anchored to the Unified Concept of Species by explicitly evaluating lineage independence (de Queiroz 2005). By balancing the quantification of divergence with an assessment of cohesion, and recognizing species as dynamically evolving metapopulation lineages that could be at various stages of speciation (de Queiroz 2005), we can reduce taxonomic inflation and develop a more robust and nuanced understanding of cryptic species—and, potentially, their contribution to the broader diversity of life on Earth.

Assessing cohesion often conflicts with conventional approaches designed to quantify divergence. For instance, tree-based methods and classical multispecies coalescent models assume no gene flow following species divergence (Rannala and Yang 2003), an assumption directly at odds with the central aim of assessing cohesion. When this assumption is violated, phylogenies can exhibit the *artifactual branch effect*, where admixed lineages are erroneously inferred as distinct species due to their long branches at early-diverging positions of clades (Chan et al. 2020b; Pyron et al. 2022; Sánchez et al. 2024; Liu et al. 2025). Genetic distances within species can also exceed those between species when gene flow is present, a phenomenon known as the *species-definition anomaly zone* (Jiao and Yang 2020). These limitations are further exacerbated by the inconsistent application of genetic distance metrics, which are heavily influenced by the choice of calculation method, distance model, and data set (Chan et al. 2022a; Fregin et al. 2012). Therefore, if most cryptic species studies rely on tree/distance-based methods that are prone to overestimating species diversity, we find it reasonable to question whether quantitative extrapolations and sweeping claims regarding the prevalence of cryptic species across the Tree of Life are tenable. More importantly, if our fundamental understanding of cryptic species and the speciation process in general is clouded by incorrectly applied methodologies, unclear definitions, and inflated estimates, then it becomes imperative to reevaluate both the methods and assumptions driving cryptic species research. Without this critical reassessment, we risk perpetuating a distorted view of cryptic species, one that conflates methodological artifacts with biological reality.

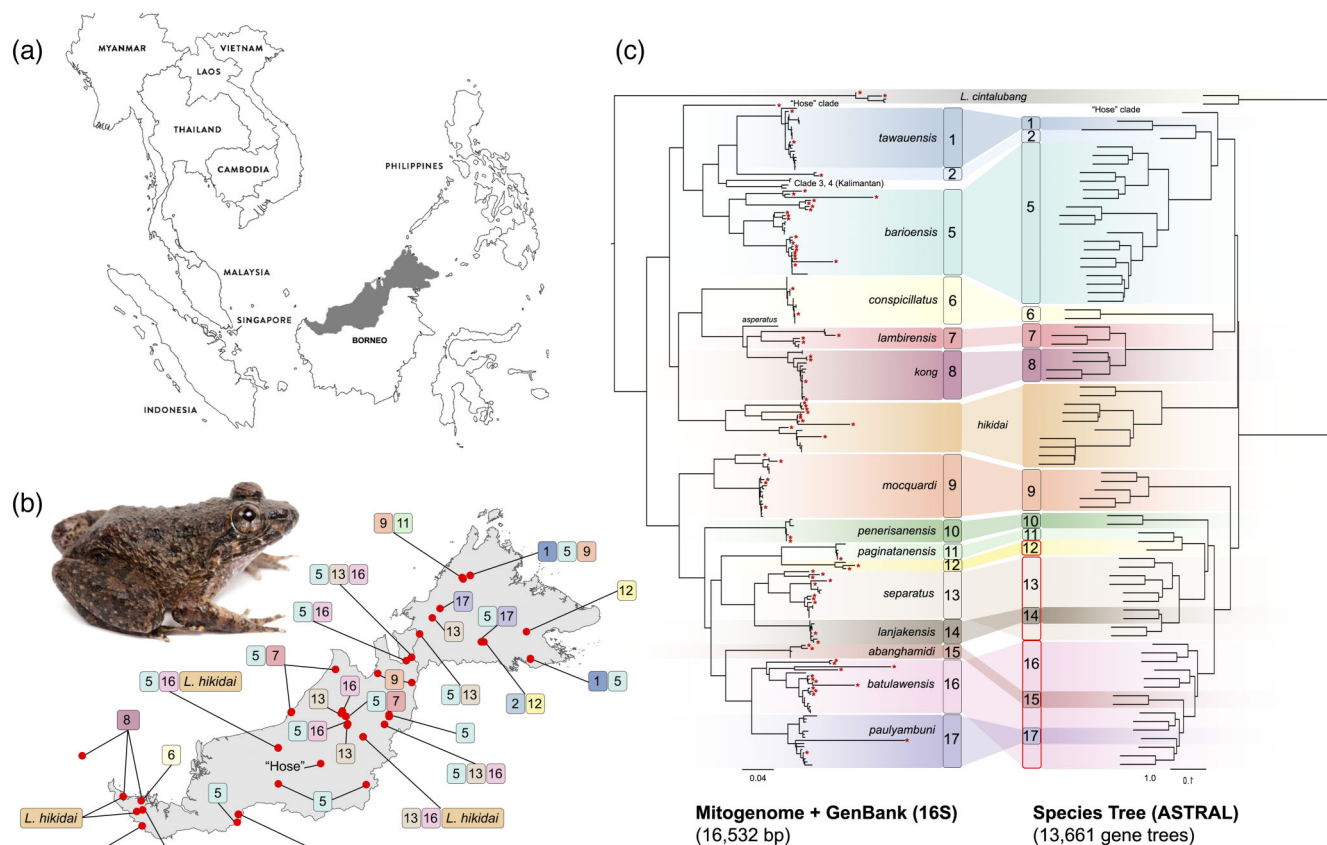


FIGURE 1. (a) Map of Southeast Asia with the study region (Malaysian states of Sabah and Sarawak) highlighted. (b) *Limnectes kuhlii* (unknown clade) and spatial distribution of samples labeled and colored according to mitochondrial clades (shown in panel C). Photo by S.T.H. (c) Left tree: Maximum Likelihood phylogeny based on mitogenomes (16,532 bps) and supplemented with additional 16S sequences obtained from GenBank. Names of species and their respective clade numbers are shown. *Limnectes asperatus*, Clade 3, and Clade 4 are from Indonesian Borneo and were unavailable for downstream analyses. Red asterisks indicated samples that were sequenced for genomic data. Right tree: Species tree inferred using ASTRAL-III (13,661 gene trees). All major clades along the backbone of both phylogenies were fully supported.

Fanged Frogs in the *Limnectes kuhlii* complex, renowned for their extraordinary genetic diversity but not accompanied by consistent morphological or ecological differentiation (Matsui et al. 2016), represent an ideal empirical system for studying evolutionary dynamics of cryptic species. Once considered a few widespread species distributed widely throughout Southeast Asia, molecular studies revealed extensive hidden diversity within the complex (Emerson et al. 2000; McLeod 2010; Matsui et al. 2016). These studies, primarily based on single-locus mitochondrial DNA phylogenies, sparked a wave of taxonomic revisions, leading to the description of at least 29 new species (Frost 2025). The populations of the *L. kuhlii* complex on the island of Borneo (hereafter referred to as Bornean Fanged Frogs; Fig. 1a) exemplify the seemingly remarkable cryptic diversity of these frogs, forming a monophyletic group encompassing 20 putative lineages, 18 of which have been formally described as distinct species (Matsui and Nishikawa 2014; Matsui et al. 2014, 2024; Dehling

and Dehling 2017; Gonggoli et al. 2025b). Moreover, some localities host up to four sympatric, non-sister “species” (Fig. 1b, c), a pattern that has been interpreted as evidence supporting lineage distinction (McLeod 2010; Matsui et al. 2016).

Bornean Fanged Frogs exhibit limited phenotypic and ecological disparity. Additionally, bioacoustics analysis of mating calls is intractable because they are either voiceless, produce barely audible calls, or call irregularly, making them difficult to record (Emerson and Berrigan 1993). As such, species delimitation of Bornean Fanged Frogs has solely relied on tree/distance-based methods (McLeod 2010; Matsui et al. 2016, 2024), and their species boundaries remain unvalidated by genomic data. Therefore, the purported but unvalidated cryptic species diversity of Bornean Fanged Frogs, combined with their complex, overlapping distributions, that may have created opportunities for gene flow, makes this an ideal system for studying the evolutionary drivers of cryptic species, while serving as a cautionary

case study for overdescribing cryptic species (Zachos et al. 2013; Leaché et al. 2019; Mahony et al. 2024; Chambers et al. 2025).

MATERIALS AND METHODS

Sampling and Genomic Data Collection

Genomic data were obtained from tissue samples of 79 vouchered specimens housed at the Naturhistorisches Museum Bern, Switzerland (Supplementary Table S1). The samples comprised populations and mitochondrial lineages of the *L. kuhlii* complex across the Malaysian states of Sabah and Sarawak on the island of Borneo. Whenever possible, multiple samples from every clade/nominal species were included, and their identities were verified by incorporating them in a phylogenetic analysis that included sequences of holotypes obtained from GenBank (Supplementary Table S1). Genomic data were obtained via target-capture sequencing using the FrogCap protocol. Our goal was to create a set of universal amphibian markers that could be targeted using a single set of probes. To initially select and design markers, we used the FrogCap marker set V3 (Hutter et al. 2022; <https://github.com/chutter/FrogCap-Sequence-Capture>). We selected the full FrogCap markers that were successfully captured across Anura, selecting markers captured from the Ranoidea V2 and Hyloidea V2 probe sets as well as the Metazoan Amphibian probe set from Dietz et al. (2023). We also added additional legacy markers, including those from the rapidly evolving long exon capture (RELEC) marker set (Karin et al. 2019) and anchored hybrid enrichment (AHE) markers (Hime et al. 2021) extracted from their respective publications by BLAST-ing to all the genomes in Supplementary Table S2 and extracting the markers from those genomes.

We aimed to add additional universal markers that could potentially be compared and included in phylogenetic analyses with other groups of organisms. We selected the Universal Single Copy Orthologs (USCO) set of markers, which are orthologous genes found across deep evolutionary scales (i.e., Kingdom, Phylum, Class). To identify USCO markers present across Anura, we used all the available amphibian genomes (Supplementary Table S2). We used the program BUSCO v3.0.2 (Seppey et al. 2019) to locate and identify the Vertebrate, Metazoan, and Tetrapoda USCO markers in these genomes. We included genes where at least half the genomes recovered USCO markers and included at least one Salamander and Caecilian sample.

To create baits from the FrogCap markers, a consensus sequence was created from the alignments of the selected markers, as this has been suggested to improve target capture success across large phylogenetic scales, and methods for bait design largely follow (Hutter et al. 2022). For all the markers, we used BLAST and

extracted each orthologous gene from each genome and separated each of the genes into exons based on the boundaries determined by each genome's annotations. We aligned each exon and created a consensus sequence for that exon from across all the genomes. The consensus sequences were used to design a MYbaits-2 (120,000 baits) custom bait library (Arbor Biosciences), using 120 mer baits to best capture sequences with greater than 5% divergence from the probes. The finalized set of markers was separated into probe sequences following a 2× tiling scheme, starting 20 bp behind the start codon of the exon and tiling 120 bp probes every 60 bp until 20 bp past stop codons. Individual probes were filtered using these criteria: (1) excluded probes that matched 70% or greater of the length to multiple locations in the genomes referenced above (Supplementary Table S2) using BLAST; (2) kept probes with a GC content between 30 and 50%; (3) retained probes if there were no repetitive sequences based on RepeatMasker annotations using the online server (Sun et al. 2015); and (4) kept probes if there were no matches with BLAST to other probes (using a 70% similarity criterion). This resulted in a final set of 34,114 markers from 5,050 genes covering a total of 9,268,697 bp. The total number of markers from each type of data set is shown in Supplementary Table S3. Raw FrogCap data were processed using the accompanying bioinformatic R package PhyloProcessR. Workflow details are available at GitHub: <https://github.com/chutter/PhyloProcessR>.

Phylogenomics and Network Analysis

Maximum likelihood (ML) consensus trees were inferred using IQ-TREE-2 (Minh et al. 2020) for near-complete mitogenomes and nuclear genomic loci separately. Loci were concatenated and a partitioned analysis by loci was performed. Due to the large number of nuclear loci, substitution model testing was intractable. Therefore, the most parameter-rich model (GTR + F + I + G4) was applied to every partition/locus, a viable strategy for large genomic data sets (Chan et al. 2020a). Branch support was assessed using 1000 Ultrafast Bootstrap replicates. IQ-TREE-2 was also used to infer gene trees for each locus using the best-fit substitution model inferred using ModelFinder. Individual gene trees were then used as input for a species tree analysis using ASTRAL-III (Zhang et al. 2018) and a quartet-based network analysis using the NANUQ algorithm (Allman et al. 2019) implemented in the R package MSCquartets (Rhodes et al. 2021). As recommended by the developers, we used high alpha ($1e-70$) and beta (0.1) values for the NANUQ quartet hypothesis test to ensure that inferred networks were highly significant. The program SplitsTree (Huson and Bryant 2024) was then used to visualize a splits network, minimum spanning network, and cluster network, all of which were additionally filtered at a weight threshold of 70, ensuring that

inferred reticulation was supported by a high number of quartets.

Population Structure

A data set comprising 13,659 unlinked single nucleotide polymorphisms (SNPs; one SNP per locus; exons from genes were combined into a single locus) was used to infer population structure using the program sNMF (sparse non-negative matrix factorizations) implemented in the R package LEA (Frichot and François 2015). Ancestry coefficients were estimated for 1–20 ancestral populations (K) using 100 replicates for each K . The cross-entropy criterion was then used to determine the best K based on the prediction of masked genotypes. To assess population clustering, we reduced the dimensionality of the SNP data set using the t -distributed stochastic neighbor embedding (t -SNE) algorithm that can capture local structure of high-dimensional data within two or three dimensions. This analysis was performed with the R package *Rtsne* (Krijthe 2015) using the following settings: perplexity = 11, max iterations = 1,000,000, and theta = 0.0.

Gene Flow and Species Delimitation

The program Dsuite (Malinsky et al. 2021) was used to perform a four-taxon ABBA–BABA test to detect introgression among clades. The *Dtrios* and *f*-branch functions were used to calculate the D , f_4 -ratio, and *f*-branch statistics. Migration rates were estimated using the MSC-M model implemented in the program BPP v 4.7.0 (Flouri et al. 2018) and performed through the Python pipeline HHSD (Kornai et al. 2024). For computational tractability, we performed this analysis on 100 randomly selected loci. The following settings were used: theta prior = 3, 0.001 e; tau prior = 3, 0.03; burn-in = 125,000; MCMC samples = 500,000. Other parameters were left at default settings. We also employed the hierarchical merge and split algorithms in HHSD to delimit species based on the genealogical divergence index (gdi) that considers population split times, population sizes, and migration rates (Jackson et al. 2017). We employed a threshold of $gdi \geq 0.7$, ≥ 0.5 for the split algorithm and $gdi \leq 0.3$, ≤ 1.0 for the merge algorithm.

Divergence Profile

Distance-based species delimitation approaches implicitly assume that lineages diverge in a predictable manner, where genetic distance increases linearly as they evolve from populations to species, and that intraspecific distances do not exceed interspecific ones. To test this assumption, we constructed genome-wide divergence profiles using several commonly applied metrics, including D_{xy} , F_{st} , D_a (net divergence), and uncorrected p -distances. We then compared the distributions

and relationships of these metrics with genomic marker type and between populations versus species to test two hypotheses: (1) Divergence patterns based on different metrics are consistent across the genome (using marker type as a proxy for genomic regions) and (2) divergence increases linearly from populations to species.

For Hypothesis 1, we used the program pixy to calculate unbiased estimates of nucleotide diversity (π), D_{xy} , and F_{st} by incorporating invariant sites and accounting for missing data (Korunes and Samuk 2021). Briefly, we generated a consensus alignment using EMBOSS to serve as a pseudo-reference. Raw reads that have been quality-filtered and trimmed were mapped to the pseudo-reference sequence using bwa-mem followed by variant-calling using bcftools (Danecek et al. 2021). A BED file was then created to map the starting and ending positions for each locus based on their sequence lengths to ensure that calculations were performed on a locus-by-locus basis as opposed to using arbitrary sliding windows. Net divergence (D_a) was calculated using the formula $D_{xy} - (\pi_x + \pi_y)/2$, where x and y represent the two populations compared. Correlations and variation in the distributions of D_a , D_{xy} , and F_{st} among different marker types were calculated to determine whether divergence patterns are consistent across the genome.

To test Hypothesis 2, we compared distributions of pairwise divergences between and within species across D_a , D_{xy} , F_{st} , and uncorrected p -distances (mitochondrial and nuclear). To test for linearity, we used the Akaike Information Criterion (AIC) to assess the fit of our data to the following models: linear, polynomial, generalized additive model (GAM), and spline. Finally, because the mitochondrial 16S rRNA marker is ubiquitously used as a proxy for species divergence, we tested whether genetic distances between the mitochondrial and nuclear genomes are correlated.

RESULTS

Our full genomic data set consists of 13,661 highly informative loci derived from several marker types: Frog-Cap (6453 loci), BUSCO (3626), UCE (3177), AHE (178), Functional (138), Legacy (69), and RELEC (20). After filtering sequences with <500 bps and <100 parsimony-informative sites (Hutter and Duellman 2023), the filtered data set retained 9,968 loci. Detailed summary statistics of the data are presented in Table 1. We refer to putative species by their respective clade numbers rather than species names (Fig. 1c), which are not central to this study and can be difficult to follow. These clade numbers have been well-established and consistently applied in the literature (Matsui et al. 2016, 2024), providing a coherent, standardized reference system. Only two species from Malaysian Borneo were described without an associated clade number: *Limnionectes citalubang* (outgroup) and *Limnionectes hikidai*. Species

TABLE 1. Summary statistics of the different marker types sequenced in this study

Marker	# of loci	Alignment length	% Missing	# Var. sites	% Var. sites	# PIS	% PIS
AHE	178	472–2681 1684.3	0.2–14.2 4.4	52–720 255.4	3.6–40.0 15.5	41–479 168.6	2.3–28.4 10.3
BUSCO	3626	139–53,790 1959.5	0.04–39.1 5.7	22–12,602 713.0	2.3–85.4 36.5	6–9394 520.0	1.5–68.5 26.5
FrogCap	6453	148–9411 860.9	0.01–43.5 5.3	11–4595 359.6	3.1–90.8 42.8	5–3474 266.8	1.3–85.2 31.7
Functional	138	183–10,705 1810.2	0.2–29.2 7.6	30–6330 752.7	9.6–80 41.7	16–4943 549.7	5.1–68.2 30.5
Legacy	69	474–4272 1357.1	0.4–20.0 4.5	38–1211 277.1	4.8–64.5 21.4	26–529 176	3.3–54.2 14.6
RELEC	20	869–5648 2408.3	0.2–22.3 5.1	107–778 370.1	5.3–32.8 16.4	73–482 240.6	3.9–17.5 10.5
UCE	3177	214–1666 880.1	0.05–18.6 3.4	7–781 180.5	0.8–77.3 20.5	2–739 103.7	0.2–73.1 11.9

AHE = anchored hybrid enrichment; BUSCO = benchmarking universal single copy orthologs; FrogCap = custom-designed probeset for frogs; Functional = known functional genes; Legacy = a set of commonly used genes in Sanger sequencing studies; RELEC = rapidly evolving long exon capture; UCE = ultraconserved elements; PIS = parsimony-information sites. Values are shown as min–max | mean.

names and their respective clade numbers and associated metadata are provided in [Supplementary Table S1](#).

Phylogenomics and Population Structure

ML phylogenies based on the mitogenome (16,532 bps), concatenated full genomic data set (14,992,489 bps), and the ASTRAL species tree phylogeny (13,661 gene trees) were fully supported at ancestral nodes along the backbone of the phylogenies but showed discordance at the clade and individual levels (Fig. 1c; [Supplementary Figs. S1 and S2](#)). Similar to previous studies ([Matsui et al. 2016, 2024](#)), *L. cintalubang* was consistently inferred as the first-branching lineage, serving as the outgroup to root our phylogenies. At the clade level, Clades 1–17 and *L. hikidai* were monophyletic in the mitogenomic phylogeny, but Clades 12, 13, and 16 were paraphyletic in the species tree (Fig. 1c). At the individual level, several errant samples from Clades 1, 13, 12, 16, and 17 were discordant and placed in other clades. Additionally, a new population from Mount Hose (designated “Hose” Clade) was inferred as the first-branching lineage of Clades 1, 2, and 5. Species trees inferred separately for each marker type were also highly discordant, though mostly well supported ([Supplementary Material](#)).

The sNMF population structure analysis yielded ambiguous results regarding the optimal number of ancestral populations, *K* (Fig. 2a). The cross-entropy criterion indicated that *K* values ranging from 6 to 14 had similar scores (Fig. 2b). The *t*-SNE analysis largely reflected the sNMF results, with several clades showing similar ancestries and forming tight, overlapping clusters (Fig. 2c). These include Clade 1 + 2, Clade 7 + 8, Clade 11 + 12, Clade 13 + 14, and Clade 15 + 16 + 17. Due to the lack of evidence supporting their separation, we merged some of these undifferentiated clusters to facilitate down-

stream gene flow and species delimitation analyses.

Phylogenetic Networks Analysis

The NANUQ analysis of both the full and filtered data sets revealed numerous quartets consistent with a 4-cycle network, even at a stringent threshold of $\alpha = 1e-70$ (Fig. 3a). A splits graph based on intertaxon distances from the NANUQ analysis showed extensive reticulation, especially among Clades 10–17 (Fig. 3b–d). A minimum spanning network identified six distinct clusters: *L. cintalubang*, *L. hikidai*, Clade 5, Clades 6 + 7 + 8, Clade 9, and Clades 10–17. Errant individuals that were discordant in the phylogenies were interspersed between Clade 5 and Clades 10–17 (Fig. 3c). A cluster network further revealed numerous reticulation edges across different evolutionary scales, suggesting both ancestral and contemporary gene flow. Reticulation within Clades 10–17 was more recent, whereas the other clades exhibited older reticulation events (Fig. 3d).

Gene Flow, Introgression, and Species Delimitation

The *D*suite and *F*-branch analyses detected extensive introgression across the major markers ([Supplementary Fig. S3](#)). Ancestral introgression was high between *L. hikidai* and the ancestral lineage leading to Clade 6–8 and Clades 10–17, as well as between the lineage leading to Clades 5–8 and that leading to Clades 10–17 ([Supplementary Fig. S3](#)). These findings are consistent with the NANUQ analysis, which also revealed reticulation between the ancestral branches of these lineages (Fig. 3d). The *F*-branch analysis further detected more contemporary introgression between Clade 5 and Clades 6–8; Clade 5 and Clade 13 + 14; Clade 7 + 8 and Clade 1 + 2 + 5; and Clade 13 + 14 and Clade 15 + 16 + 17 ([Supplementary Fig. S3](#)). Using 100 ran-

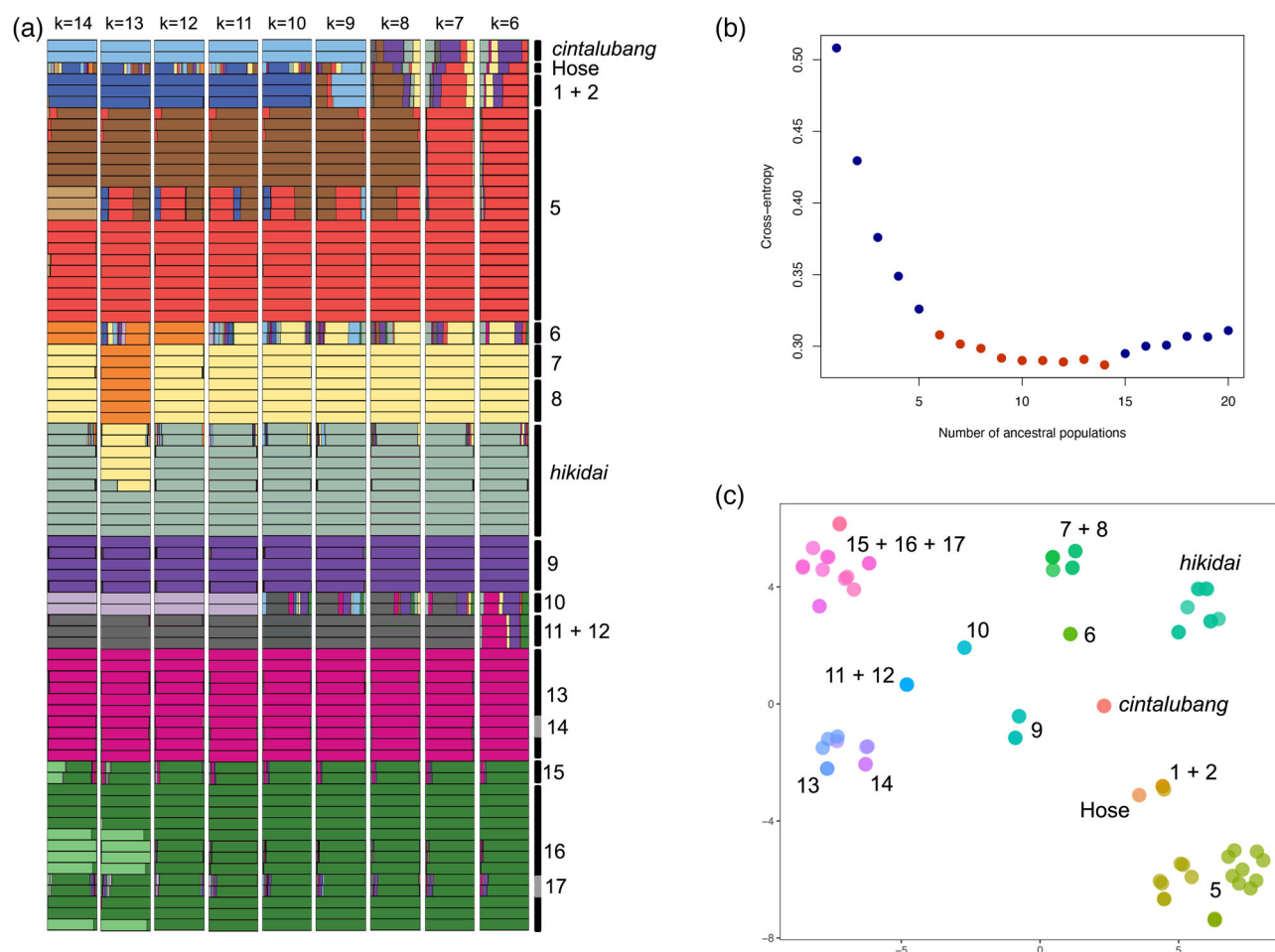


FIGURE 2. (a) Results of the sNMF population structure analysis for $K = 6-14$. (b) Cross-entropy scores for $K = 1-20$. The most plausible K 's range from 6–14. (c) Genomic clusters derived from a t -SNE analysis based on 100,000 iterations and perplexity = 20. Both sNMF and t -SNE analyses were based on 13,659 unlinked SNPs.

domly loci selected from all marker types, directional migration rates estimated from the MSC + migration model in the program BPP indicated high asymmetrical contemporary gene flow within and between major clades (Fig. 4). When migration rates, population split times, and effective population sizes were integrated into a species delimitation framework under the MSC + migration model, six species were supported, excluding the outgroup: Clade 1 + 2 + 5, Clade 6, Clade 7 + 8, Clade 9, Clade 10–17, and *L. hikidai* (Supplementary Fig. S4 and Table S4). We were unable to adequately assess the specific status of the “Hose” Clade because it comprises a single sample, while preliminary phylogenetic, population structure, and clustering analyses were unable to confidently assign it to any other population. However, the results clearly show that it is a highly admixed lineage most closely related to Clade 1 + 2 + 5; thus, we provisionally assign it to that clade. Due to this uncertainty and its highly admixed genetic

makeup, we excluded it from downstream analyses to avoid potential bias.

Consistency across Different Markers

We also performed network and migration analyses on AHE, BUSCO, FrogCap, and UCE markers separately to assess the consistency of results across different marker types (the other markers had <300 loci and were not assessed). The NANUQ analysis showed variable results. Notably, the UCE data set showed no evidence of reticulation, whereas the other marker types showed variable levels (Fig. 5a). Migration rates were also highly variable. No significant migration ($M < 0.1$) was detected from AHE and UCE markers, while BUSCO and FrogCap markers detected different levels of migration across different population pairs (Fig. 5b). Overall, networks, migration rates, split times (τ), and population sizes (θ) also varied widely across different marker types (Fig. 5; Supplementary Tables S5–S7).

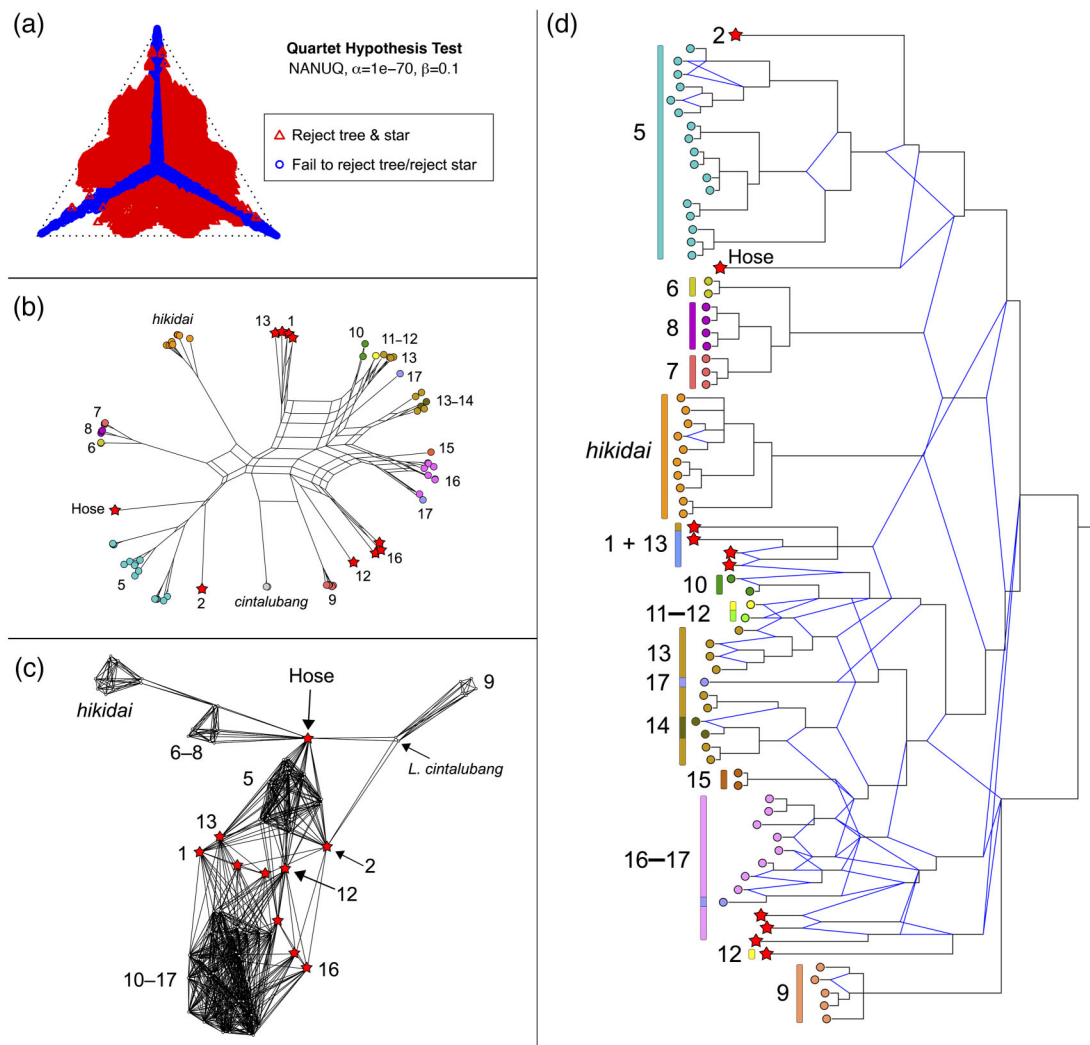


FIGURE 3. (a) Simplex plot from the NANUQ quartet hypothesis test showing rampant gene flow. Circle points correspond to quartets that form a bifurcating tree (no gene flow), whereas triangles correspond to quartets forming a 4-cycle network (gene flow). (b) A split graph based on intertaxon distances calculated from NANUQ. Stars at tips are putative hybrids that form paraphyletic clades and are highly discordant among trees. (c) A minimum spanning network based on NANUQ intertaxon distances. Stars represent putative hybrids. (d) A cluster network from NANUQ filtered at a weight threshold of 70 to reduce noise and weakly supported networks. Blue lines represent highly supported quartets (>70) that form a network, and colors represent mitochondrial clades.

Divergence Profile and the Speciation Continuum

Pairwise population comparisons of net divergence (D_a) revealed a three-stage gradient (Fig. 6a). The lowest net divergences ($D_a < 0.005$) were observed between Clades 7 and 8 and among Clades 13–17, followed by a gradual increase between 0.005 and 0.02, a range that corresponds to the gray zone of speciation (Roux et al. 2016). The third stage showed a sharper increase associated with comparisons involving the outgroup (*L. cintalubang*). Divergence patterns varied across metrics and marker types. UCE and AHE markers exhibited the lowest D_a and D_{xy} values, whereas FrogCap, BUSCO, and functional

markers were the highest. However, this pattern was not observed in F_{st} , which showed the opposite pattern, where FrogCap, BUSCO, and Functional markers had the highest values (Fig. 6a, b). Similarly, D_a and D_{xy} were highly correlated ($R = 0.919$), whereas F_{st} was the least correlated with D_{xy} ($R = 0.273$) and moderately correlated with D_a ($R = 0.572$) (Fig. 6b). Across all metrics, the only consistent pattern was that Clades 7 and 8, and Clades 13–17 had significantly lower divergences (Fig. 6a).

We determined the boundaries of species and populations based on the results of previous analyses performed in this study that supported a six-species de-

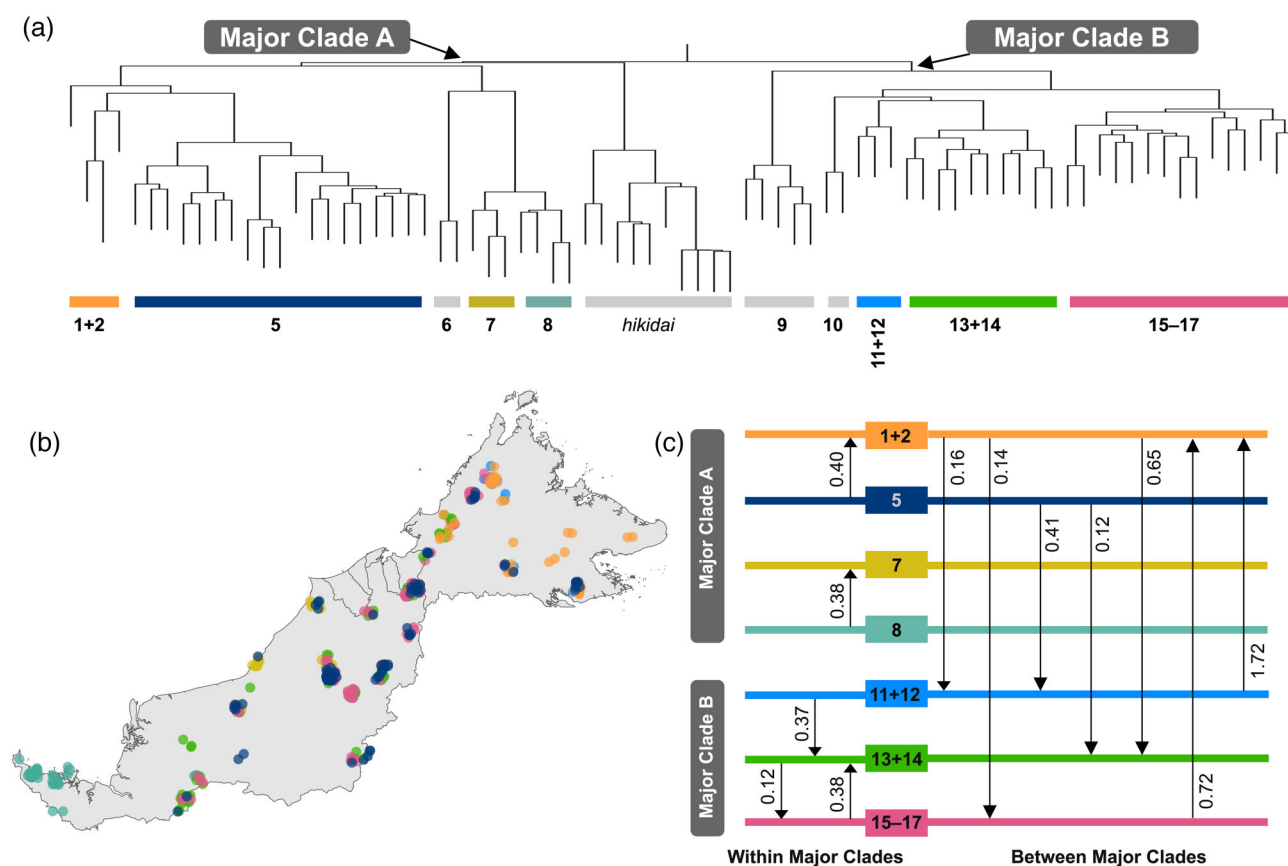


FIGURE 4. (a) Species tree inferred using ASTRAL. (b) Geographic distribution of populations among which significant ongoing gene flow (migration > 0.10) was inferred. For ease of interpretation, populations with low/no ongoing gene flow are not shown on the map. (c) Direction and magnitude of migration rates among populations with high gene flow (migration > 0.1) estimated using the MSC-M model in the program BPP. The diagram is split into the two major clades, A and B, to show migration within and between major clades.

limitation scheme (excluding the outgroup taxon *L. cintangubang*): Clade 1 + 2 + 5, Clade 6, Clade 7 + 8, Clade 9, Clade 10–17, and *L. hikidai*. Divergence patterns of species and populations were highly correlated between D_a and D_{xy} ($R = 0.855$) and least correlated between D_{xy} and F_{st} ($R = 0.195$; Fig. 6c). Notably, D_{xy} and F_{st} showed a weak correlation within species ($R = 0.267$) and no correlation between species ($R = 0.057$) (Fig. 6c). When divergences were aggregated and compared across the genome (all markers combined), genetic distances within several species exceeded distances between species, a phenomenon known as the *species-definition anomaly zone* (Jiao and Yang 2020) (Fig. 7). For D_a and p -distances, the species-definition anomaly zone was restricted to the earlier stages of the gray zone at the boundary between populations and species but was distributed throughout the entire speciation continuum for F_{st} (Fig. 7). Model testing showed that a linear model had the poorest fit to the data, whereas a generalized additive (GAM) or polynomial model had the best fit (Supplementary Table S8).

DISCUSSION

Insights into the Speciation Continuum

Empirical representations of the speciation continuum are often derived from model taxa (Seehausen et al. 2014; Supple et al. 2015) or reconstructed from unrelated taxon pairs at various stages of diversification (Edwards et al. 2016; Roux et al. 2016; McLaughlin et al. 2020). In this study, we demonstrate that Bornean Fanged Frogs encompass the full spectrum of the speciation continuum within a single species complex spanning a large continuous landscape. This system provides a unique opportunity to investigate the gradual process of speciation within a single continuously diverging clade. We identified three distinct stages within the speciation continuum: an initial stage, a gray zone, and a terminal stage. The initial stage comprised early-diverging populations that were clearly not evolving independently and consistently exhibited low divergence across all metrics (i.e., Clades 7 + 8 and 13–17) (Figs. 6a and 7). The terminal stage corresponds to complete speciation, exemplified by the most basal lineage, *L.*

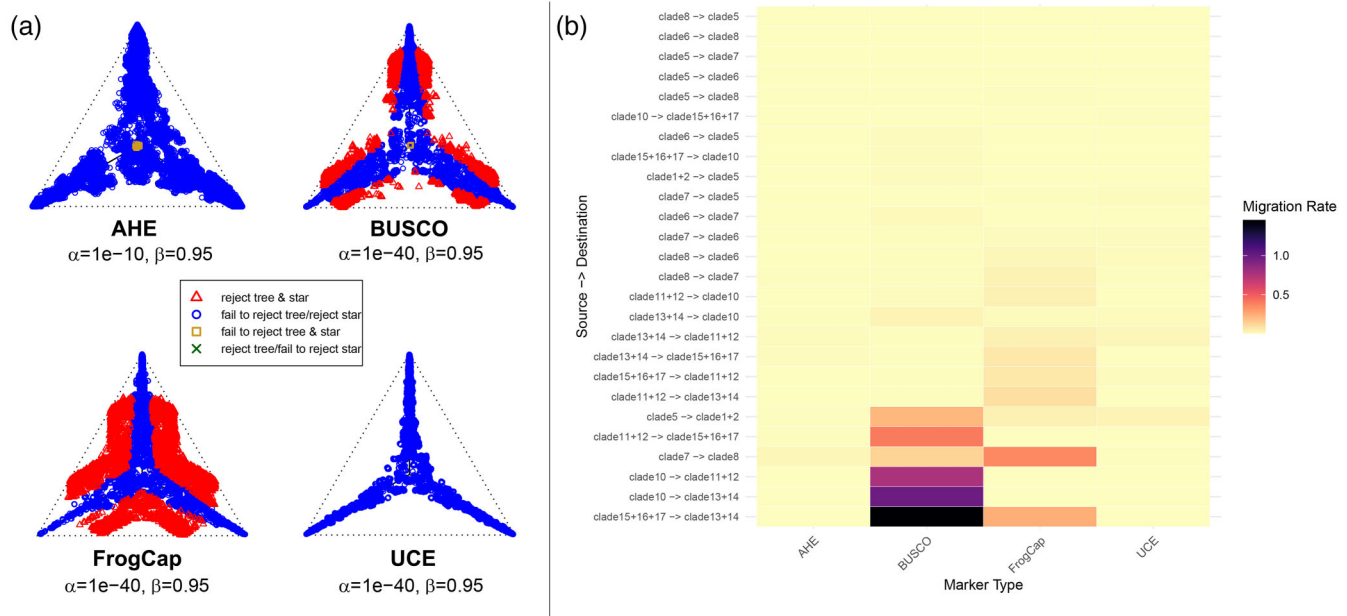


FIGURE 5. (a) Simplex plots from the NANUQ analysis performed separately on the four major marker types. Circle points correspond to quartets that form a bifurcating tree (no gene flow), whereas triangles correspond to quartets forming a 4-cycle network (gene flow). Other marker types contained less than 300 loci and were not assessed. (b) Heatmap of directional migration rates estimated from 100 randomly selected loci from each of the four major marker types.

cintalubang—the only uncontroversial species that exhibits clear phenotypic (significantly smaller size, orange hindlimbs, and presence of bluish-white spots) and ecological differentiation (burrow- vs. stream-dwelling) (Matsui et al. 2014). The gray zone lies between the initial and terminal stages of divergence and corresponds remarkably well with the 0.5–2% net divergence range (Fig. 6a) previously identified through comparative genomic analysis of 61 unrelated taxon pairs (Roux et al. 2016). Within this zone, populations and species exhibit unpredictable and nonlinear patterns of divergence (Figs. 6 and 7; Supplementary Table S8), creating considerable challenges for taxonomic inference. Furthermore, divergence profiles vary among metrics, marker types, and taxonomic scales (populations vs. species), reflecting differential evolutionary rates across genes and hierarchical levels (Romeiro-Brito et al. 2016).

Heterogeneity across Metrics, Methods, and Markers

As expected, D_a and D_{xy} were highly correlated because their calculations are related. However, D_a showed less variation across the genome than D_{xy} (Fig. 6a), likely because D_a corrects for within-population diversity while D_{xy} measures absolute divergence. F_{st} showed a weaker correlation with D_{xy} than with D_a (Fig. 6b), consistent with the fact that both F_{st} and D_a account for intrapopulation variation, whereas D_{xy} reflects total divergence, including both fixed differences and shared polymorphisms. Between-species compar-

isons showed that the relationship between D_{xy} and F_{st} becomes decoupled ($R = 0.057$; Fig. 6c). This pattern suggests that while absolute sequence divergence (D_{xy}) accumulates over time, population differentiation (F_{st}) saturates or increases more slowly, potentially due to ancestral polymorphisms shared through historical gene flow that inflate D_{xy} but contribute less to contemporary differentiation. Alternatively, if gene flow has ceased between species, divergence may accumulate primarily through lineage-specific mutations that increase D_{xy} while F_{st} reaches an asymptote apparently determined by effective population sizes. In contrast, within-species comparisons showed a positive correlation between D_{xy} and F_{st} ($R = 0.267$; Fig. 6c). This pattern is consistent with incomplete reproductive isolation and ongoing or recent gene flow, under which divergence accumulates through drift and mutation. However, gene flow remains insufficient to fully homogenize populations, allowing both D_{xy} and F_{st} to increase together. Notably, asymmetrical gene flow could create heterogeneous patterns where some population pairs diverge more rapidly than others. The differences in correlation strength and slope between population and species comparisons demonstrate that the dynamics of divergence differ fundamentally across these taxonomic scales. More importantly, we show that divergence in the gray zone is not linear (Fig. 6c, 7; Supplementary Table S8), reflecting the complex interplay of gene flow, drift, and selection during the early stages of diversification.

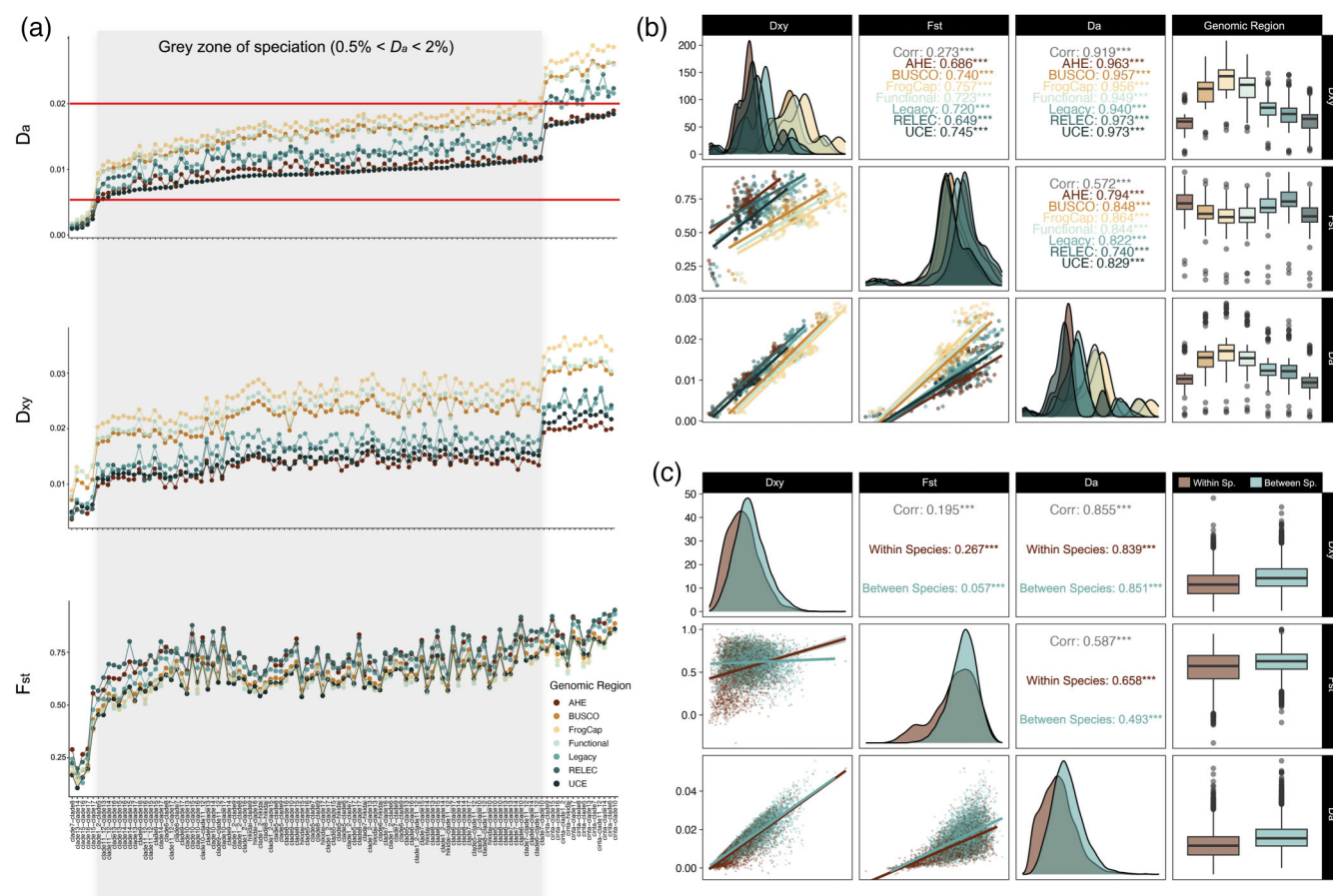


FIGURE 6. A genome-wide divergence profile of the speciation continuum. (a) Median values of D_a , D_{xy} , and F_{st} across different marker types. The horizontal lines demarcate the lower and upper bounds of the gray zone ($0.005 < D_a < 0.02$). All pairwise populations (x-axis) that fall within the gray zone are highlighted in a box. Plots are ordered according to ascending median D_a to compare consistency among marker types. (b) Distributions and correlations of D_a , D_{xy} , and F_{st} compared across different marker types and (c) between- and within-species comparisons. Species are grouped based on results from all other analyses: Clade (1 + 2 + 5), Clade 6, Clade (7 + 8), Clade 9, Clade 10–17, and *L. hikidai*.

Different analytical approaches and markers revealed contrasting patterns of gene flow. Quartet-based gene tree analysis (NANUQ) detected gene flow in AHE, BUSCO, and FrogCap markers, but not UCEs (Fig. 5a); site-based ABBA-BABA tests using SNPs showed variable introgression signals across all four marker types (Supplementary Fig. S3); whereas sequence-based migration rate estimates detected strong migration in BUSCO and FrogCap markers but not in AHE and UCEs (Fig. 5b; Supplementary Table S6). These discordant results demonstrate that analytical approach (quartet-, site-, or sequence-based), data type (gene trees, SNPs, or sequences), and genomic target (AHE, BUSCO, FrogCap, or UCE) vary considerably in their utility to detect gene flow across different time scales (historical vs. ongoing). These disparities likely reflect microevolutionary processes acting on different parts of the genome. Selection can generate false positives, leading to elevated estimates of migration rates in non-neutral loci (Smith and Hahn 2024), while the

magnitude and presence of gene flow could be associated with recombination rates (Nachman and Payseur 2012; Thom et al. 2024; Burbrink et al. 2025). Furthermore, gene flow itself involves many parameters, including timing, directionality, and rate, and model misspecifications can further lead to erroneous estimates (Thawornwattana et al. 2025). These complexities caution against relying on any single line of evidence, as apparent conflicts among methods may reflect genuine biological heterogeneity. It is thus not realistic to expect that a single marker type or analytical framework can fully capture the temporal and genomic complexity of diversification. Therefore, it is imperative for researchers to carefully match their choice of markers and methods to their specific research questions—whether investigating ancient introgression, ongoing migration, phylogenetic relationships, or species boundaries—while employing integrative approaches that account for methodological limitations and genome heterogeneity.

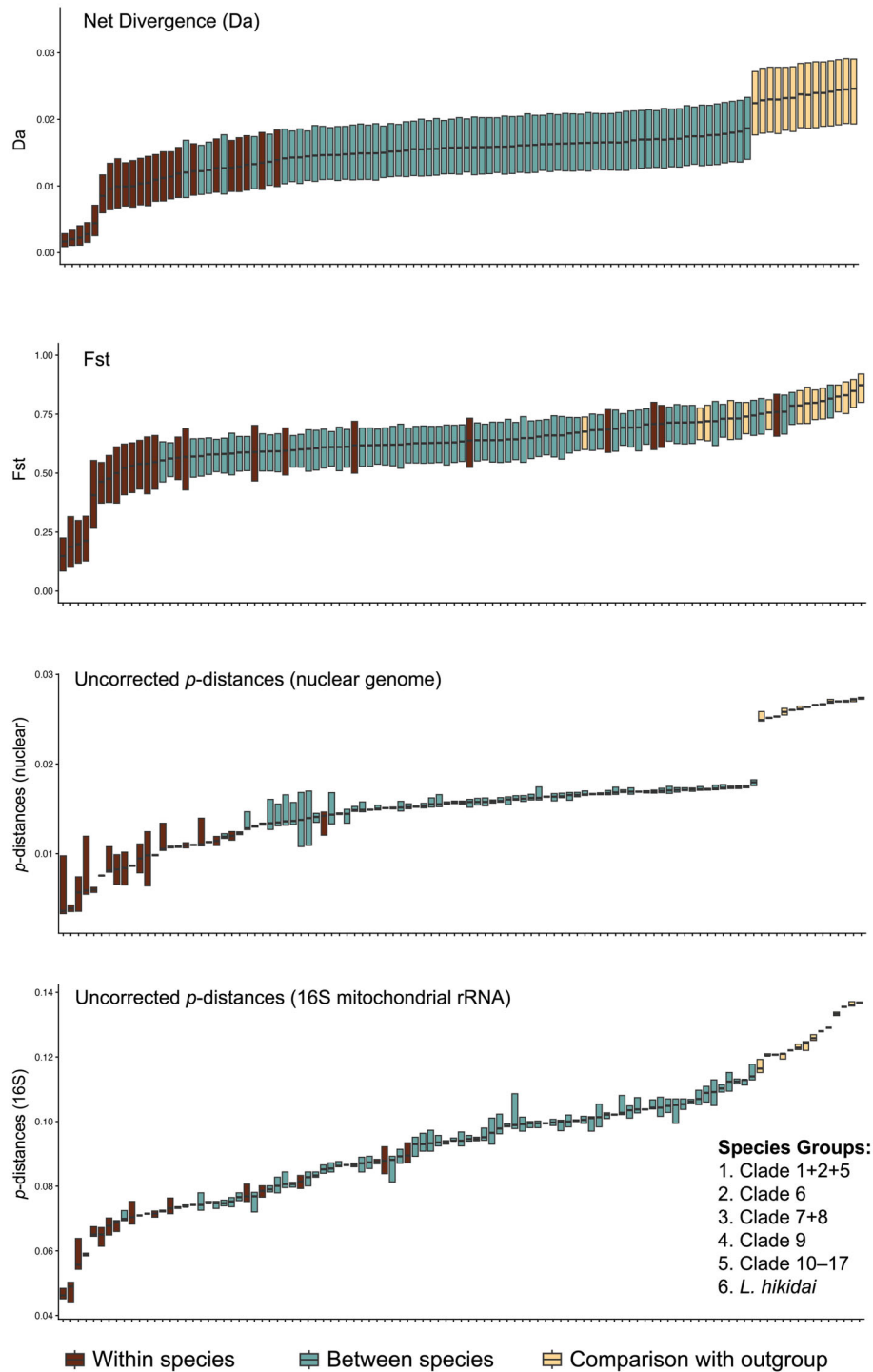


FIGURE 7. Boxplots of D_a , F_{st} , and uncorrected p -distances (nuclear and mitochondrial) with each pairwise population comparison aggregated across all marker types. Each plot is ordered from low to high median values and colored according to species groups to illustrate the inconsistent and nonlinear pattern of divergence across different metrics.

Notably, our findings demonstrate that UCEs, while highly effective for resolving phylogenetic relationships (Faircloth et al. 2012), are less appropriate for inferring recent demographic processes such as ongoing migration. UCE core regions are under strong purifying selection (Bejerano et al. 2004; Winker et al. 2018; Cummins et al. 2024; Thomas et al. 2025), which reduces genetic diversity not only in the core regions but also in nearby flanking regions through background selection (Charlesworth et al. 1993; Thomas et al. 2025). This limited variation makes UCEs poor targets for methods that rely on polymorphisms to infer recent demographic events but excellent for phylogenetics due to low gene tree discordance (Thomas et al. 2025). Conversely, FrogCap and BUSCO markers target protein-coding genes where synonymous sites can accumulate variation more freely (Simão et al. 2015). In such regions, deleterious mutations are likely removed by purifying selection, but synonymous variation or mildly beneficial alleles are allowed to persist (Charlesworth et al. 1993; Elyashiv et al. 2016). UCE markers may fail to detect ongoing migration using methods that rely on sequence variation such as BPP because little neutral variation is preserved in those regions due to background or linked selection.

The ABBA-BABA test differs from sequence-based methods by operating on biallelic SNP patterns rather than modeling mutation and drift processes. This allows it to detect introgression even in conserved loci like AHE and UCEs by leveraging patterns of shared derived alleles among taxa (Denton et al. 2018; Winker et al. 2018). Introgressed alleles, once established in a genome, can persist and create detectable asymmetries in allele sharing even when overall variability is low. However, ABBA-BABA cannot distinguish between recent and ancient introgression and may fail to detect ongoing migration if gene flow has not yet generated strong asymmetrical allele sharing patterns across the genome. Collectively, our results demonstrate that while AHE and UCEs can detect historical introgression in some cases, they may not be ideal for inferring recent or ongoing gene flow. For fine-scale demographic questions, markers with more polymorphic sites and less severe selection constraints can better capture recent evolutionary dynamics, though researchers must remain cognizant of potential biases introduced by selection at functional loci (Johri et al. 2021).

Recombination rates may also influence gene flow detection across different marker types. Regions with higher recombination rates facilitate detection because recombination breaks up linkage to deleterious alleles and prevents background selection from reducing genetic variation, allowing migrant alleles to persist and remain detectable (Thom et al. 2024; Burbrink et al. 2025). Conversely, low-recombination regions experience stronger effects of linked selection, where large haplotype blocks move together and variation

is more readily purged by selection against linked deleterious mutations (Schumer et al. 2018; Tigano et al. 2022). This reduction in variation can make recent gene flow more difficult to detect, though ancient introgression events may still be detectable if introgressed haplotypes became established before variation was lost. UCEs, which likely occur in regions with lower recombination rates and strong purifying selection, may lose variation through persistent background selection, reducing the genetic polymorphism available for detecting recent migration (Thom et al. 2024; Thomas et al. 2025).

Empirical Confirmation of the Species-Definition Anomaly Zone

This study presents the first empirical evidence of the *species-definition anomaly zone* that was theoretically posited but never conclusively demonstrated empirically. The *species-definition anomaly zone* refers to a scenario where different population sizes and asymmetrical migration rates cause within-species divergence to exceed between-species divergence, where divergence is measured by gene tree probability or average genetic distance (Jiao and Yang 2020). Our results provide evidence for asymmetrical migration rates (Fig. 5b; Supplementary Table S5), large differences in population sizes (Supplementary Table S7), phylogenetic discordance, and instances where within-species divergences are greater than between-species divergences across different metrics (Fig. 7). Although divergence estimates could be affected by the small sample sizes of certain populations, we do not believe that it has significantly biased the results. If systematic bias were present, we would expect to observe a correlation between sample size and estimations of metrics relevant to calculating divergence. For example, Clade 1 has a sample size of $n = 2$ but has the second-highest estimate of nucleotide diversity (π), whereas Clade 10, which also has a sample size of $n = 2$, has one of the lowest. *Limnonectes hikidai*, which has one of the largest sample sizes ($n = 10$) has a moderate π , whereas Clade 13 ($n = 6$), has the highest π among all populations (Supplementary Fig. S5). Therefore, although small sample size could affect the accuracy of estimates, it does not appear to systematically bias broader patterns of divergence.

Except for F_{st} , the *species-definition anomaly zone* appears to be restricted to the first half of the speciation continuum, where populations are in the earlier stages of divergence (Fig. 7). By estimating relevant demographic parameters (e.g., migration and population size), assessing gene tree discordance, and characterizing genome-wide divergence profiles, we demonstrate a mechanistic link among several confounding phenomena, such as phylogenetic discordance, the *artificial branch effect*, and the *species-definition anomaly*

zone, all of which are driven by gene flow within the gray zone.

Diversification Dynamics in the Gray Zone Create an Illusion of Cryptic Species

The complex diversification dynamics in the gray zone, which can manifest in confounding phenomena such as the *artifactual branch effect* and the *species-definition anomaly zone*, create an illusion of high species diversity. Meanwhile, historical and ongoing gene flow can result in continuous variation in morphological traits (Stefanello et al. 2021). This scenario provides a surprisingly straightforward explanation for the evolutionary origins, purported high diversity, and defining characteristics of cryptic species—genetic divergence coupled with morphological similarity. This gene-flow-driven process (herein referred to as the gene flow hypothesis) differs from other proposed explanatory models (Fišer et al. 2018), such as the recent divergence hypothesis: insufficient time for morphological differences to evolve (Egea et al. 2016); the phylogenetic niche conservatism hypothesis: selection maintains morphological stasis (Smith et al. 2011); or the morphological convergence hypothesis—morphological similarity driven by convergent evolution (Stayton 2006). The gene flow hypothesis expands our understanding of cryptic species as possible outcomes of continuous diversification in the gray zone, rather than simply genetically distinct units that failed to evolve morphological differentiation. This new perspective reconciles otherwise anomalous patterns, such as high genetic but low morphological divergence and the occurrence of non-sister lineages in sympatry. Nevertheless, the gene flow hypothesis complements rather than replaces other proposed mechanisms underlying the evolution of cryptic species and offers a novel, nuanced, and dynamic evolutionary framework for identifying and understanding cryptic biodiversity across the Tree of Life.

Cryptic Species Delimitation

Our results show that the evolutionary histories of cryptic species can be heavily shaped by gene flow, confounding divergence estimates and phylogenetic inferences, linking evolutionary trajectories of populations, and blurring species boundaries. Consequently, widely used methods that rely on guide trees and divergence thresholds while ignoring gene flow risk overestimating species diversity through two key mechanisms. First, gene flow creates lineages with mixed ancestry that are inferred as long, seemingly distinct branches, often positioned basally to the parental clade that contributes the most alleles—the *artifactual branch effect* (Chan et al. 2022b, 2023), which is increasingly being documented in the literature (Dolinay et al. 2021; Pyron et al. 2023; Barley et al. 2024; Sánchez et al. 2024; Waldron et al. 2024; Liu et al. 2025). In Bornean fanged frogs, Clades 1, 2, 10,

11, 12, and “Hose Clade” exemplify this phenomenon, highlighting the inability of conventional bifurcating phylogenies to accurately depict reticulate evolutionary history. When applied to downstream species delimitation approaches, artifactual branches are invariably misinterpreted as distinct species due to their long branch lengths and distinct topological placement, generating an illusion of deep divergence.

Second, we demonstrate that different distance metrics, such as D_a , F_{st} , and uncorrected p -distances (mitochondrial or nuclear), do not provide consistent measures of intra- and interspecific divergence (Fig. 7). We showed that D_a and nuclear p -distances were the most congruent (albeit not precise) with population and species boundaries, whereas mitochondrial p -distances and F_{st} were the least congruent and most affected by the *species-definition anomaly zone*, where multiple within-species divergences exceeded divergences between species. These discordant results demonstrate that widely used divergence metrics can produce erratic results in the face of gene flow and are thus poor proxies for species divergence, further undermining the static threshold approach for species delimitation.

This study illustrates the importance of extensive geographic sampling to obtain a more comprehensive understanding of divergence and cohesion. However, whereas we have sampled the Malaysian region of Borneo extensively, the Indonesian region of Kalimantan remains sparsely sampled. At least four nominal species occur in central and southeastern Kalimantan (Inger et al. 1996; Gonggoli et al. 2025a, 2025b) that were not available for our study. Of these, the mitochondrial phylogeny shows that Clade 3 + 4 from southeastern Borneo (Kalimantan) is sister to Clade 1 + 2 from northeastern Borneo. Given the highly admixed genotype of Clade 1 + 2 that readily exchanges genes with other sympatric lineages, gene flow between Clade 1 + 2 and Clade 3 + 4 is plausible. A tract of potentially suitable habitat along the entire eastern arc of Borneo connects their distributions, creating opportunities for gene exchange. However, the region between their ranges remains unsampled. The widely distributed Clade 5 may also occur there, potentially acting as a conduit for gene flow.

Broader Implications

Our findings reveal that reticulate evolution can play a major role in the evolutionary history of cryptic species, causing tree- and distance-based approaches to overestimate species diversity. Because these approaches are commonly used to detect cryptic and noncryptic species alike (Rannala and Yang 2020), reports on the prevalence of cryptic species across the Tree of Life could be considerably inflated, with profound ramifications for conservation. Misidentifying admixed lineages as distinct species atomizes taxonomy, artificially reducing distribution ranges and es-

timates of effective population sizes, while increasing perceived extinction risks for “range-restricted, micro-endemic” taxa (Cerca et al. 2020). Such misrepresentation elevates their conservation value, potentially misguiding resource allocation and undermining effective conservation (Prie et al. 2012; Zachos 2013; Duarte et al. 2014; Oates and Ting 2015; Conix 2019; Folt et al. 2019; Robuchon et al. 2019; Berrilli et al. 2024; Clavero et al. 2024; Mahony et al. 2024). The oversplitting of Bornean Fanged Frogs (Matsui et al. 2024) exemplifies this issue, where 14 named entities assessed in this study (eight of which are not distinct species) have been assigned excessively restricted ranges, several of which are erroneously defined outside protected areas, meeting the criteria for conservation threat status.

Ultimately, the prevalence of cryptic species reflects not just biological reality but also the limitations of certain methods when applied to organisms actively diversifying in the gray zone. By recognizing gene flow’s central role in shaping cryptic diversity and adopting integrative, hypothesis-driven approaches to species delimitation, we can move toward a more accurate understanding of biodiversity—one that acknowledges the dynamic, continuous nature of diversification while maintaining the taxonomic stability essential for effective conservation and comparative biology.

ACKNOWLEDGMENTS

I.D. and S.T. H. thank the Sarawak Forest Department for the issuance of collecting permits (147)JHS/NCCD/600-7/2/107/Jld.2 and Park Permit N0.74/2019, as well as the Institute of Biodiversity and Environmental Conservation, Universiti Malaysia Sarawak, and Niche Research Grant Scheme of the Ministry of Higher Education, Government of Malaysia: NRGS/1087/2013(01) and Volkswagen-Stiftung, Germany (1/79 405) for support of their research. We thank collaborators, guides, and field assistants who supported us during fieldwork over the years, specifically Alexander Haas, Lars Fehlandt, Pui Yong Min, Laurence Etter, Jana Flury, Natalie Reichen, Moric Toszeghi, Basil Minder, Samuel Cilloni, and Fabian Heussler. We also thank Gina Sommer for assistance with DNA extractions, Reto Hagmann for facilitating access to the specimens, and Pamela Nicholson and Judith Trüb for their assistance with specimen preparation.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Systematic Biology* online.

CONFLICT OF INTEREST STATEMENT

None declared.

FUNDING

FrogCap library preparation and sequencing were funded by a Universiti Sains Malaysia International Academic Fellowship awarded to KOC. Financial support to CRH and for initial FrogCap genomic resources development was provided by the U.S. National Sciences Foundation [GRFP 1540502, 1451148, and 0907996 to C.R.H.; and DEB 1654388 to R.M.B.] and to R.M.B. from the University of Kansas College of Liberal Arts and Sciences (Docking Faculty Scholar Fund) and KU’s Office of the Provost’s Research Investment Council (RIC-Level II Award No. 2300207).

DATA AVAILABILITY

Raw reads have been uploaded to GenBank SRA under BioProject accession PRJNA1210643. All code and data sets relevant to this study can be obtained at the Dryad Repository <https://doi.org/10.5061/dryad.c866t1gk6> and <https://github.com/chankinonn/kuhlii>

REFERENCES

- Adams M., Raadik T.A., Burridge C.P., Georges A. 2014. Global biodiversity assessment and hyper-cryptic species complexes: more than one species of elephant in the room? *Syst. Biol.* 63:518–533.
- Allman E.S., Baños H., Rhodes J.A. 2019. NANUQ: a method for inferring species networks from gene trees under the coalescent model. *Algorithms Mol Biol.* 14:1–25.
- Barley A.J., Nieto-Montes de Oca A., Manríquez-Morán N.L., Thomson R.C. 2024. Understanding species boundaries that arise from complex histories: gene flow across the speciation continuum in the Spotted Whiptail Lizards. *Syst. Biol.* 73:901–919.
- Bejerano G., Pheasant M., Makunin I., Stephen S., Kent W.J., Mattick J.S., Haussler D. 2004. Ultraconserved elements in the human genome. *Science* (1979). 304:1321–1325.
- Berrilli E., Gambioli B., Bombi P., Garzia M., Muraro M., Pardo C., Reale M., Sherpa S.t., Ficetola G.F., Vignoli L., Salvi D. 2024. Subspecies inflation hampers conservation efforts: a case study on wall lizards. *Biol. J. Linn. Soc.* 143:blae001.
- Burbrink F.T., Debaun D., Foley N.M., Murphy W.J. 2025. Recombination-aware phylogenomics. *Trends Ecol. Evol.* 3474:1–13.
- Cerca J., Meyer C., Purschke G., Struck T.H. 2020. Delimitation of cryptic species drastically reduces the geographical ranges of marine interstitial ghost-worms (*Stygocapitella*; Annelida, Sedentaria). *Mol. Phylogenet. Evol.* 143:106663.
- Chambers E.A., Lara-Tufiño J.É.D., Campillo-García G., Cisneros-Bernal A.Y., Dudek D.J., León-Règagnon V., Townsend J.H., Flores-Villela O., Hillis D.M. 2025. Distinguishing species boundaries from geographic variation. *Proc. Natl. Acad. Sci. USA* 122:e2423688122.
- Chan K.O., Hertwig S.T., Neokleous D.N., Flury J.M., Brown R.M. 2022a. Widely used, short 16S rRNA mitochondrial gene fragments yield poor and erratic results in phylogenetic estimation and species delimitation of amphibians. *BMC Ecol. Evol.* 22:37.
- Chan K.O., Hutter C.R., Wood P.L., Grismer L.L., Brown R.M. 2020a. Larger, unfiltered datasets are more effective at resolving phylogenetic conflict: introns, exons, and UCEs resolve ambiguities in Golden-backed frogs (Anura: Ranidae; genus *Hylarana*). *Mol. Phylogenet. Evol.* 151:106899.
- Chan K.O., Hutter C.R., Wood P.L., Grismer L.L., Das I., Brown R.M. 2020b. Gene flow creates a mirage of cryptic species in a Southeast Asian spotted stream frog complex. *Mol. Ecol.* 29:3970–3987.
- Chan K.O., Hutter C.R., Wood P.L.J., Su Y.-C.C., Brown R.M. 2022b. Gene flow increases phylogenetic structure and inflates cryptic

- species estimations: a case study on widespread Philippine Puddle Frogs (*Occidozyga laevis*). *Syst. Biol.* 71:40–57.
- Chan K.O., Mulcahy D.G., Anuar S. 2023. The artefactual branch effect and phylogenetic conflict: species delimitation with gene flow in mangrove pit vipers (*Trimeresurus purpureomaculatus-erythrurus* complex). *Syst. Biol.* 72:1209–1219.
- Charlesworth B., Morgan M.T., Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics*. 134:1289–1303.
- Chenuil A., Cahill A.E., Délémontey N., du Luc E., Fanton H. 2019. Problems and questions posed by cryptic species. A framework to guide future studies. In: Casetta E., da Silva J., Vecchi D., editors. From assessing to conserving biodiversity: conceptual and practical challenges. Cham: Springer International Publishing. p. 77–106.
- Clavero M., Naves J., Lucena-Perez M.I., Revilla E. 2024. Taxonomic inflation as a conservation trap for inbred populations. *Evol. Appl.* 17:1–7.
- Conix S. 2019. Taxonomy and conservation science: interdependent and value-laden. *HPLS*. 41:1–17.
- Cummins M., Watson C., Edwards R.J., Mattick J.S. 2024. The evolution of ultraconserved elements in vertebrates. *Mol. Biol. Evol.* 41:1–17.
- Danecek P., Bonfield J.K., Liddle J., Marshall J., Ohan V., Pollard M.O., Whitwham A., Keane T., McCarthy S.A., Davies R.M., Li H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10:1–4.
- Dehling J.M., Dehling D.M. 2017. A new wide-headed fanged frog of the *Limnodynastes* group (Anura: Dicroglossidae) from western Borneo with a redescription of *Rana conspicillata* Günther, 1872. *Zootaxa* 4317:291–309.
- Denton R.D., Morales A.E., Gibbs H.L. 2018. Genome-specific histories of divergence and introgression between an allopolyploid unisexual salamander lineage and two ancestral sexual species. *Evolution* 72:1689–1700.
- de Queiroz K. 2005. A unified concept of species and its consequences for the future of taxonomy. *Proc. California Acad. Sci.* 56:196–215.
- Dietz L., Eberle J., Mayer C., Kukowka S., Bohacz C., Baur H., Espeland M., Huber B.A., Hutter C., Mengual X., Peters R.S., Vences M., Wesener T., Willmott K., Misof B., Niehuis O., Ahrens D. 2023. Standardized nuclear markers improve and homogenize species delimitation in Metazoa. *Methods Ecol. Evol.* 14:543–555.
- Dolinay M., Nečas T., Zimkus B.M., Schmitz A., Fokam E.B., Lemmon E.M., Lemmon A.R., Gvoždík V. 2021. Gene flow in phylogenomics: sequence capture resolves species limits and biogeography of Afromontane forest endemic frogs from the Cameroon Highlands. *Mol. Phylogenet. Evol.* 163:107258.
- Duarte M., Guerrero P.C., Carvallo G., Bustamante R.O. 2014. Conservation network design for endemic cacti under taxonomic uncertainty. *Biol. Conserv.* 176:236–242.
- Edwards S.V., Potter S., Schmitt C.J., Bragg J.G., Moritz C. 2016. Reticulation, divergence, and the phylogeography–phylogenetics continuum. *Proc. Natl. Acad. Sci. USA* 113:8025–8032.
- Egea E., David B., Choné T., Laurin B., Feral J.P., Chenuil A. 2016. Morphological and genetic analyses reveal a cryptic species complex in the echinoid *Echinocardium cordatum* and rule out a stabilizing selection explanation. *Mol. Phylogenet. Evol.* 94:207–220.
- Elyashiv E., Sattath S., Hu T.T., Strutsofsky A., McVicker G., Andolfatto P., Coop G., Sella G. 2016. A genomic map of the effects of linked selection in *Drosophila*. *PLoS Genet.* 12:1–24.
- Emerson S.B., Berrigan D. 1993. Systematics of Southeast Asian Ranids: multiple origins of voicelessness in the subgenus *Limnodynastes* (Fitzinger). *Herpetologica* 49:22–31.
- Emerson S.B., Inger R.F., Iskandar D. 2000. Molecular systematics and biogeography of the fanged frogs of Southeast Asia. *Mol. Phylogenet. Evol.* 16:131–142.
- Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T., Glenn T.C. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* 61:717–726.
- Fišer C., Robinson C.T., Malard F. 2018. Cryptic species as a window into the paradigm shift of the species concept. *Mol. Ecol.* 27:613–635.
- Flouri T.A., Jiao X., Rannala B., Yang Z. 2018. Species tree inference with BPP using genomic sequences and the multispecies coalescent. *Mol. Biol. Evol.* 35:2585–2593.
- Folt B., Bauder J., Spear S., Stevenson D., Hoffman M., Oaks J.R., Wood P.L., Jenkins C., Steen D.A., Guyer C. 2019. Taxonomic and conservation implications of population genetic admixture, mito-nuclear discordance, and male-biased dispersal of a large endangered snake, *Drymarchon couperi*. *PLoS One* 14:1–21.
- Fregin S., Haase M., Olsson U., Alström P. 2012. Pitfalls in comparisons of genetic distances: a case study of the avian family Acrocephalidae. *Mol. Phylogenet. Evol.* 62:319–328.
- Frichot E., François O. 2015. LEA: an R package for landscape and ecological association studies. *Methods Ecol. Evol.* 6:925–929.
- Frost D.R. 2025. Amphibian species of the world: an online reference. New York (NY): American Museum of Natural History. <http://resea.arch.amnh.org/herpetology/amphibia/index.html>.
- Fujisawa T., Aswad A., Barraclough T.G. 2016. A rapid and scalable method for multilocus species delimitation using Bayesian model comparison and rooted triplets. *Syst. Biol.* 65:759–771.
- Gonggoli A.D., Shimada T., Nishikawa K., Matsui M., Larson T., Smith E.N., Hamidy A. 2025a. Unveiling a new species of Fanged Frog (Amphibia: Anura: Dicroglossidae) from Southeastern Borneo, Indonesia. *Spec. Div.* 30:177–191.
- Gonggoli A.D.E.D., Shimada T., Matsui M., Nishikawa K., Sidik I., Kadafi A.M., Farajallah A., Hamidy A. 2025b. Two new species of fanged frog from Southeastern Borneo, Indonesia (Amphibia: Anura: Dicroglossidae). *Zootaxa* 5575:387–408.
- Heethoff M. 2018. Cryptic species—conceptual or terminological chaos? A response to struck. *Trends Ecol. Evol.* 33:310.
- Hime P.M., Lemmon A.R., Lemmon E.C.M., Prendini E., Brown J.M., Thomson R.C., Kratochvil J.D., Noonan B.P., Pyron R.A., Peloso P.L.V., Kortyna M.L., Keogh J.S., Donnellan S.C., Mueller R.L., Raxworthy C.J., Kunte K., Ron S.R., Das S., Gaitonde N., Green D.M., Labisko J., Che J., Weisrock D.W. 2021. Phylogenomics reveals ancient gene tree discordance in the amphibian tree of life. *Syst. Biol.* 70:49–66.
- Huson D.H., Bryant D. 2024. The SplitsTree App: interactive analysis and visualization using phylogenetic trees and networks. *Nat. Methods* 21:1773–1774.
- Hutter C.R., Cobb K.A., Portik D.M., Travers S.L., Wood P.L., Brown R.M. 2022. FrogCap: a modular sequence capture probe set for phylogenomics and population genetics for all frogs, assessed across multiple phylogenetic scales. *Mol. Ecol. Resour.* 22:1100–1119.
- Hutter C.R., Duellman W.E. 2023. Filtration of gene trees from 9,000 exons, introns, and UCEs disentangles conflicting phylogenomic relationships in tree frogs (hylidae). *Genome Biol. Evol.* 15:1–18.
- Inger R.F., Boeadi, Taufik A. 1996. New species of ranid frogs (Amphibia: Anura) from central Kalimantan, Borneo. *Raffles Bull. Zool.* 44:363–369.
- Jackson N.D., Carstens B.C., Morales A.E., O'Meara B.C. 2017. Species delimitation with gene flow. *Syst. Biol.* 66:799–812.
- Jiao X., Yang Z. 2020. Defining species when there is gene flow. *Syst. Biol.* 70:108–119.
- Johri P., Riall K., Becher H., Excoffier L., Charlesworth B., Jensen J.D. 2021. The impact of purifying and background selection on the inference of population history: problems and prospects. *Mol. Biol. Evol.* 38:2986–3003.
- Jörger K.M., Schrödl M. 2013. How to describe a cryptic species? Practical challenges of molecular taxonomy. *Front Zool.* 10:59.
- Kapli P., Lutteropp S., Zhang J., Kobert K., Pavlidis P., Stamatakis A., Flouri T. 2017. Multi-rate poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics*. 33:1630–1638.
- Karin B.R., Gamble T., Jackman T.R., Vidal N. 2019. Optimizing phylogenomics with rapidly evolving long exons: comparison with anchored hybrid enrichment and ultraconserved elements. *Mol. Biol. Evol.* msz263:904–922.
- Kornai D., Jiao X., Ji J., Flouri T., Yang Z. 2024. Hierarchical heuristic species delimitation under the multispecies coalescent model with migration. *Syst. Biol.* 73:1015–1037.

- Korshunova T., Martynov A. 2024. The phyloperiodic approach removes the "Cryptic Species" and puts forward multilevel organismal diversity. *Diversity*. 16:220.
- Korunes K.L., Samuk K. 2021. pixy: unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Mol. Ecol. Resour.* 21:1359–1368.
- Krijthe J.H. 2015. Rtsne: distributed stochastic neighbor embedding using a Barnes-Hut implementation. <https://github.com/jkrijthe/Rtsne>.
- Leaché A.D., Zhu T., Rannala B., Yang Z. 2019. The spectre of too many species. *Syst. Biol.* 68:168–181.
- Li X., Wiens J.J. 2023. Estimating global biodiversity: the role of cryptic insect species. *Syst. Biol.* 72:391–403.
- Liu H., Han B., Mou H., Xiao Y., Jiang Y., Kong H., Xu G. 2025. Unraveling the extensive phylogenetic discordance and evolutionary history of spurless taxa within the *Aquilegia ecalcarata* complex. *New Phytol.* 246:1333–1349.
- Mahony S., Kamei R.G., Brown R.M., Chan K.O. 2024. Unnecessary splitting of genus-level clades reduces taxonomic stability in amphibians. *Vertebr. Zool.* 74:249–277.
- Malinsky M., Matschiner M., Svardal H. 2021. Dsuite—Fast D-statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* 21:584–595.
- Matsui M., Kuraishi N., Eto K., Hamidy A., Nishikawa K., Shimada T., Yambun P., Vairappan C.S., Hossman M.Y.B. 2016. Unusually high genetic diversity in the Bornean *Limnectes kuhlii*-like fanged frogs (Anura: Dicroglossidae). *Mol. Phylogenet. Evol.* 102:305–319.
- Matsui M., Nishikawa K. 2014. Description of a new species of *Limnectes* from Sarawak, Malaysian Borneo (Dicroglossidae, Anura). *Current Herpetology*. 33:135–147.
- Matsui M., Nishikawa K., Eto K. 2014. A new burrow-utilizing fanged frog from Sarawak, Malaysian Borneo (Anura: Dicroglossidae). *Raffles Bull. Zool.* 62:679–687.
- Matsui M., Nishikawa K., Shimada T., Eto K., Hamidy A., Sudin A., Hossman M.Y.B., Gumal M., Vairappan C.S. 2024. Revision of the *Limnectes kuhlii*-like fanged frogs from Malaysian Borneo (Amphibia: Anura: Dicroglossidae). *Curr. Herpetol.* 43:226–277.
- McLaughlin J.F., Faircloth B.C., Glenn T.C., Winker K. 2020. Divergence, gene flow, and speciation in eight lineages of trans-Beringian birds. *Mol. Ecol.* 29:3526–3542.
- McLeod D.S. 2010. Of least concern? Systematics of a cryptic species complex: *Limnectes kuhlii* (Amphibia: Anura: Dicroglossidae). *Mol. Phylogenet. Evol.* 56:991–1000.
- Minh B.Q., Schmidt H.A., Chernomor O., Schrempf D., Woodhams M.D., von Haeseler A., Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37:1530–1534.
- Mora C., Tittensor D.P., Adl S., Simpson A.G.B., Worm B. 2011. How many species are there on earth and in the ocean? *PLoS Biol.* 9:e1001127.
- Nachman M.W., Payseur B.A. 2012. Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Phil. Trans. R. Soc. B* 367:409–421.
- Oates J., Ting N. 2015. Conservation consequences of unstable taxonomies: the case of the red colobus monkeys. Canberra: ANU Press.
- Pfenninger M., Schwenk K. 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evol. Biol.* 7:1–6.
- Prates I., Hutchinson M.N., Singhal S., Moritz C., Rabosky D.L. 2024. Notes from the taxonomic disaster zone: evolutionary drivers of intractable species boundaries in an Australian lizard clade (Scincidae: *C. tenotus*). *Mol. Ecol.* 33:e17074.
- Prie V., Puillandre N., Bouchet P. 2012. Bad taxonomy can kill: molecular reevaluation of *Unio mancus* Lamarck, 1819 (Bivalvia: Unionidae) and its accepted subspecies. *Knowl. Managt. Aquatic Ecosyst.* 405:8.
- Puillandre N., Brouillet S., Achaz G. 2021. ASAP: assemble species by automatic partitioning. *Mol. Ecol. Resour.* 21:609–620.
- Pyron R.A., O'Connell K.A., Duncan S.C., Burbrink F.T., Beamer D.A. 2023. Speciation hypotheses from phylogeographic delimitation yield an integrative taxonomy for seal salamanders (*Desmognathus monticola*). *Syst. Biol.* 72:179–197.
- Pyron R.A., O'Connell K.A., Lemmon E.M., Lemmon A.R., Beamer D.A. 2022. Candidate-species delimitation in *Desmognathus* salamanders reveals gene flow across lineage boundaries, confounding phylogenetic estimation and clarifying hybrid zones. *Ecol. Evol.* 12:e8574.
- Rannala B., Yang Z. 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*. 164:1645–1656.
- Rannala B., Yang Z. 2020. Species delimitation. In: Scornavacca C., Del-suc F., Galtier N., editors. *Phylogenetics in the genomic era*. p. 5.5:1–5.5:18. <https://hal.inria.fr/PGE>.
- Reis V.J.C., dos Santos S.A., Britto M.R., de Assis Volpi T., de Pinna M.C.C. 2020. Iterative taxonomy reveals a new species of *Trichomycterus* Valenciennes 1832 (Siluriformes, Trichomycteridae) widespread in Rio Doce basin: a pseudocryptic of *T. immaculatus*. *J. Fish Biol.* 97:1607–1623.
- Rhodes J.A., Baños H., Mitchell J.D., Allman E.S. 2021. MSCquartets 1.0: quartet methods for species trees and networks under the multispecies coalescent model in R. *Bioinformatics* 37:1766–1768.
- Robuchon M., Faith D.P., Julliard R., Leroy B., Pellens R., Robert A., Thévenin C., Véron S., Pavoine S. 2019. Species splitting increases estimates of evolutionary history at risk. *Biol. Conserv.* 235:27–35.
- Romeiro-Brito M., Moraes E.M., Taylor N.P., Zappi D.C., Franco F.F. 2016. Lineage-specific evolutionary rate in plants: contributions of a screening for *Cereus* (Cactaceae). *Appl Plant Sci.* 4:1500074.
- Roux C., Fraisse C., Romiguier J., Anciaux Y., Galtier N., Bienne N. 2016. Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biol.* 14:1–22.
- Sánchez K.I., Recknagel H., Elmer K.R., Avila L.J., Morando M. 2024. Tracing evolutionary trajectories in the presence of gene flow in South American temperate lizards (Squamata: *Liolaemus kingii* group). *Evolution* 78:716–733.
- Schumer M., Xu C., Powell D.L., Durvasula A., Skov L., Holland C., Blazier J.C., Sankararaman S., Andolfatto P., Rosenthal G.G., Przeworski M. 2018. Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science* (1979) 360:656–660.
- Seehausen O., Butlin R.K., Keller I., Wagner C.E., Boughman J.W., Hohenlohe P.A., Peichel C.L., Saetre G.-P., Bank C., Brännström Å., Brelsford A., Clarkson C.S., Eroukhmanoff F., Feder J.L., Fischer M.C., Foote A.D., Franchini P., Jiggins C.D., Jones F.C., Lindholm A.K., Lucek K., Maan M.E., Marques D.A., Martin S.H., Matthews B., Meier J.I., Möst M., Nachman M.W., Nonaka E., Rennison D.J., Schwarzer J., Watson E.T., Westram A.M., Widmer A. 2014. Genomics and the origin of species. *Nat. Rev. Genet.* 15:176–192.
- Seppely M., Manni M., Zdobnov E.M. 2019. BUSCO: assessing genome assembly and annotation completeness. *Methods Mol. Biol.* 1962:227–245.
- Shin C.P., Allmon W.D. 2023. How we study cryptic species and their biological implications: a case study from marine shelled gastropods. *Ecol. Evol.* 13:1–18.
- Simão F.A., Waterhouse R.M., Ioannidis P., Kriventseva E.V., Zdobnov E.M. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212.
- Singhal S., Hoskin C.J., Couper P., Potter S., Moritz C. 2018. A framework for resolving cryptic species: a case study from the lizards of the Australian wet tropics. *Syst. Biol.* 67:1061–1075.
- Singhal S., Leaché A.D., Fujita M.K., Cadena C.D., Zapata F. 2025. A genomic perspective on species delimitation. *Annu. Rev. Ecol. Evol. Syst.* 56:467–489.
- Smith K.L., Harmon L.J., Shoo L.P., Melville J. 2011. Evidence of constrained phenotypic evolution in a cryptic species complex of agamid lizards. *Evolution* 65:976–992.
- Smith M.L., Hahn M.W. 2024. Selection leads to false inferences of introgression using popular methods. *Genetics* 227:1–13.
- Stankowski S., Ravinet M. 2021. Defining the speciation continuum. *Evolution* 75:1256–1273.

- Stayton C.T. 2006. Testing hypotheses of convergence with multivariate data: morphological and functional convergence among herbivorous lizards. *Evolution* 60:824.
- Stefanello F., Menezes R.S.T., Ribeiro J.R.I., Almeida E.A.B. 2021. Widespread gene flow model explains the genetic–morphological variation in a giant water bug species under fine-scale spatial sampling. *Ann. Entomol. Soc. Am.* 113:160–170.
- Struck T.H., Feder J.L., Bendiksy M., Birkeland S., Cerca J., Gusarov V.I., Kistenich S., Larsson K.H., Liow L.H., Nowak M.D., Stedje B., Bachmann L., Dimitrov D. 2018. Finding evolutionary processes hidden in cryptic species. *Trends Ecol. Evol.* 33:153–163.
- Sukumaran J., Knowles L.L. 2017. Multispecies coalescent delimits structure, not species. *Proc. Natl. Acad. Sci. USA* 114:1607–1612.
- Sun Y.-B., Xiong Z.-J., Xiang X.-Y., Liu S.-P., Zhou W.-W., Tu X.-L., Zhong L., Wang L., Wu D.-D., Zhang B.-L., Zhu C.-L., Yang M.-M., Chen H.-M., Li F., Zhou L., Feng S.-H., Huang C., Zhang G.-J., Irwin D., Hillis D.M., Murphy R.W., Yang H.-M., Che J., Wang J., Zhang Y.-P. 2015. Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative evolution of tetrapod genomes. *Proc. Natl. Acad. Sci. USA* 112:E1257–62.
- Supple M.A., Papa R., Hines H.M., McMillan W.O., Counterman B.A. 2015. Divergence with gene flow across a speciation continuum of *Heliconius* butterflies. *BMC Evol. Biol.* 15:1–12.
- Thawornwattana Y., Flouri T., Mallet J., Yang Z. 2025. Inference of gene flow between species from genomic data when the mode, direction, and lineages are misspecified. *Mol. Biol. Evol.* 42:1–18.
- Thom G., Moreira L.R., Batista R., Gehara M., Aleixo A., Smith B.T. 2024. Genomic architecture predicts tree topology, population structuring, and demographic history in Amazonian birds. *Genome Biol. Evol.* 16:1–25.
- Thomas G.W.C., Hughes J.J., Kumon T., Berv J.S., Nordgren C.E., Lampson M., Levine M., Searle J.B., Good J.M. 2025. The genomic landscape, causes, and consequences of extensive phylogenetic discordance in Murine rodents. *Genome Biol. Evol.* 17(2): evaf017.
- Tigano A., Khan R., Omer A.D., Weisz D., Dudchenko O., Multani A.S., Pathak S., Behringer R.R., Aiden E.L., Fisher H., MacManes M.D. 2022. Chromosome size affects sequence divergence between species through the interplay of recombination and selection. *Evolution* 76:782–798.
- Trontelj P., Fišer C. 2009. Cryptic species diversity should not be trivialised. *Syst. Biodivers.* 7:1–3.
- Waldron B.P., Watts E.F., Morgan D.J., Hantak M.M., Lemmon A.R., Lemmon E.M., Kuchta S.R. 2024. The limits of the metapopulation: lineage fragmentation in a widespread terrestrial salamander (*Plethodon cinereus*). *Syst. Biol.* 74:1–15.
- Wattier R., Mamos T., Copilaş-Ciocianu D., Jelić M., Ollivier A., Chaumot A., Danger M., Felten V., Piscart C., Žganec K., Rewicz T., Wysocka A., Rigaud T., Grabowski M. 2020. Continental-scale patterns of hyper-cryptic diversity within the freshwater model taxon *Gammarus fossarum* (Crustacea, Amphipoda). *Sci. Rep.* 10:1–16.
- Winker K., Glenn T.C., Faircloth B.C. 2018. Ultraconserved elements (UCEs) illuminate the population genomics of a recent, high-latitude avian speciation event. *PeerJ* 2018:1–17.
- Zachos F.E. 2013. Taxonomy: species splitting puts conservation at risk. *Nature* 494:35.
- Zachos F.E., Apollonio M., Bärmann E.V., Festa-Bianchet M., Göhlich U., Habel J.C., Haring E., Kruckenhauser L., Lovari S., McDevitt A.D., Pertoldi C., Rössner G.E., Sánchez-Villagra M.R., Scandura M., Suchentrunk F. 2013. Species inflation and taxonomic artefacts—a critical comment on recent trends in mammalian classification. *Mamm. Biol.* 78:1–6.
- Zhang C., Rabiee M., Sayyari E., Mirarab S. 2018. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinform.* 19(Suppl 6):153.