the geneticssociety

www.nature.com/hdy

ARTICLE

Check for updates

Deforestation-induced Hybridization in Philippine Frogs Creates a Distinct Phenotype With an Inviable Genotype

Kin Onn Chan[1]✉, Paul M. Hime[1,2] and Rafe M. Brown[1]

Hybridization plays a major role in the evolutionary history of many taxa and can generate confounding patterns affecting many downstream applications. In this study, we empirically demonstrate how hybridization obfuscates phylogenetic inference (via the artefactual branch effect), species boundaries, and taxonomy in an adaptive radiation of frogs. Philippine narrow-mouthed frogs of the genus *Kaloula* exhibit a wide range of phenotypic and ecological adaptations but their evolutionary history and taxonomy remain poorly understood. In particular, the *Kaloula conjuncta* complex contains numerous subspecies with unresolved taxonomic boundaries and unclear evolutionary relationships. Within this complex, *Kaloula conjuncta stickeli*, until now was considered a rare, enigmatic, and phenotypically distinct subspecies that had not been encountered since its original description nearly 80 years ago. Here, we show that *K. c. stickeli* shares alleles with *K. conjuncta meridionalis* and another species outside the *conjuncta* group, *K. picta*. Using target-capture sequencing and a robust analytical framework, we show that despite having a unique phenotype, *K. c. stickeli* is likely an inviable F1 hybrid between *K. c. meridionalis* and *K. picta* and thus, does not warrant taxonomic recognition. Our results show how industry-standard approaches in systematic inference and integrative taxonomy—morphological, phylogenomic, clustering, and distance-based methods—can generate misleading results for identifying and understanding affinities of hybrids. In contrast, we demonstrate how network multispecies coalescent and population genetic approaches are more effective at accurately inferring reticulated evolutionary history. We also propose a rare phenomenon of deforestation-induced hybridization, which could have important consequences in light of large-scale Southeast Asian forest destruction.

Heredity; https://doi.org/10.1038/s41437-025-00748-y

INTRODUCTION

Once considered to be rare, hybridization is now known to be prevalent in nature and to play a major role in the evolutionary history of many taxa across the Tree of Life (Payseur and Rieseberg 2016; Mallet et al. 2016; Taylor and Larson 2019; Edelman and Mallet 2021; Peñalba et al. 2024). However, hybridization can result in reticulated evolutionary histories that are difficult to disentangle, and when unaccounted for, can generate a multitude of confounding patterns (Abbott et al. 2016; Moran et al. 2021). In phylogenetic analyses that assume lineage splitting as a strictly bifurcating process, hybridization obfuscates species boundaries by creating artefactual patterns of lineage distinctiveness and spurious affinities (Chan et al. 2020b, 2022), and/or produces unclear signals indicative of incomplete lineage separation—often characterized as the "Grey Zone" of speciation (Burbrink and Guiher 2015; Roux et al. 2016; Coates et al. 2018; Elgetany et al. 2020; Chan et al. 2023; Sánchez et al. 2023). Blurred species boundaries in the grey zone of the speciation continuum often lead to taxonomic uncertainty and one way practitioners have tried to categorize modestly divergent and incompletely separated lineages is to invoke the subspecies rank.

The struggle to characterize species versus subspecies is not a new problem. In the seminal book On the Origin of Species, Darwin noted: "Certainly no clear line of demarcation has as yet been drawn between species and sub-species" (Darwin, 1859). This species versus subspecies conundrum continues to persist today and has divided scientists who have argued for (Durrant 1955; Mayr 1982; Coates et al. 2018; Hillis 2020; Vences et al. 2024) and against (Burbrink et al. 2022) the continued usage of subspecies. Typically, the subspecies rank is applied to characterize diagnosable (phenotypically and genotypically) or geographically circumscribed populations which are not reproductively isolated from other such populations, although specific criteria on when or how to apply the subspecies rank remain inconsistent (Sackett et al. 2014; Patten 2015). However, despite the availability of genome-scale data and sophisticated methods that have provided unprecedented insights into the process of species divergence in the grey zone, (Huang and Knowles 2016; Schiaffini 2020; Burgon et al. 2021; Chambers et al. 2023), philosophical opinions regarding the ontology and epistemology of subspecies induce a certain level of subjectivity and, rightly, remain open to interpretation (De Queiroz 2020; Burbrink et al. 2022).

In this study, we investigated the role of hybridization in a species-subspecies complex of narrow-mouthed frogs of the genus *Kaloula*. The genus *Kaloula* is relatively small, containing 20 species ranging from India, Bangladesh, parts of China, Korea, and throughout southeast Asia, with almost half of the species

[1]Biodiversity Institute and Natural History Museum, University of Kansas, Lawrence, KS, USA. [2]Present address: McDonnell Genome Institute and Department of Genetics, Washington University in St. Louis, St. Louis, MO, USA. Associate editor: Ben Evans. ✉email: chanko@ku.edu

Received: 16 September 2024 Revised: 31 January 2025 Accepted: 3 February 2025
Published online: 16 February 2025

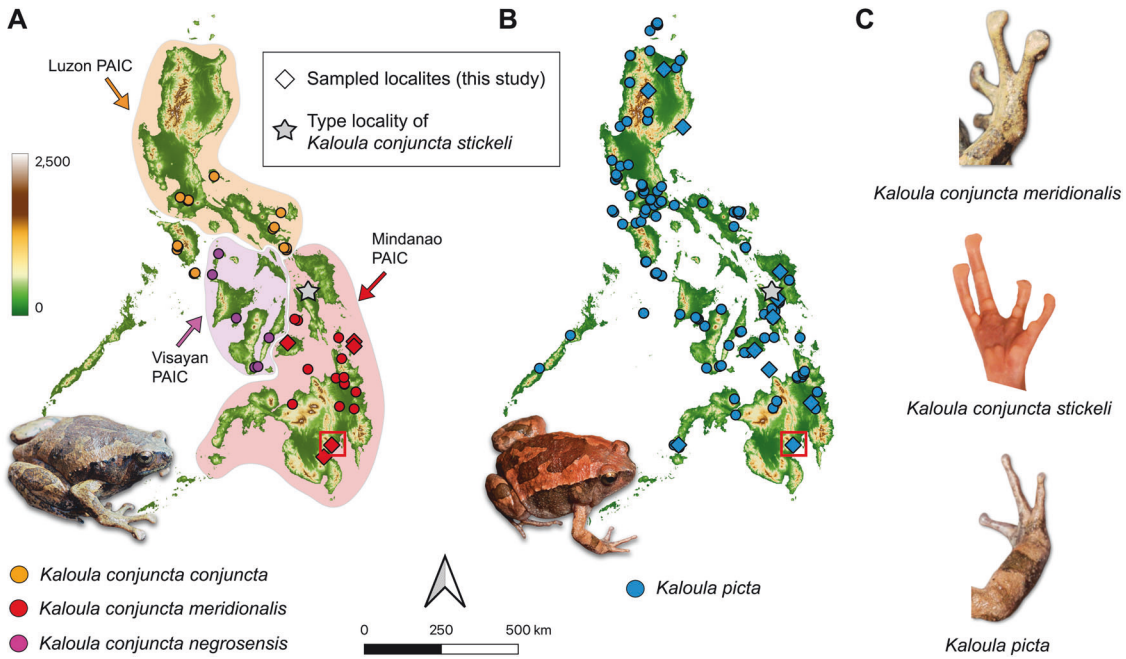

Fig. 1 Distribution of the focal taxa and examples of their finger disc widths. Distribution of the *Kaloula conjuncta* complex **A** and **B** *K. picta*. Diamond = location of samples sequenced for this study; star = type locality of *K. c. stickeli*; Red box = location of the putative hybrid zone where *K. c. meridionalis*, *K. c. stickeli*, and *K. picta* co-occur. **C** Differences in widths of the finger discs in the focal taxa.

diversity occurring in the Philippines (Frost 2024). In the Philippines, *Kaloula* represents an adaptive radiation exhibiting a range of structural habitat specialists (ecotypes) including terrestrial ground frogs, scansorial shrub frogs, and arboreal tree-hole specialists (Blackburn et al. 2013). Ground frogs have narrow finger and toe discs and an enlarged and sharpened metatarsal tubercle (or digging "spade") under the heel, while scansorial shrub and arboreal frogs have substantially wider digital discs, interpreted as functionally significant for arboreal microhabitats and adhesion while climbing (Inger 1954; Blackburn et al. 2013). However, despite its ecological and evolutionary significance, Philippine *Kaloula* remain poorly understood with multiple undescribed species, phylogenetic uncertainty, and unresolved species complexes (Blackburn et al. 2013; Alexander et al. 2017). One such complex is the *Kaloula conjuncta* group which comprises four geographically circumscribed and phenotypically distinct subspecies (*K. c. conjuncta, K. c. meridionalis, K. c. negrosensis*, and *K. c. stickeli*). *Kaloula conjuncta conjuncta* is restricted to the northern Philippines' Luzon Pleistocene Aggregate Island Complex (PAIC); *K. c. meridionalis* to the southern Mindanao PAIC; and *K. c. negrosensis* to the central West Visayan PAIC (Fig. 1). The only exception to this near-perfect matching of subspecies ranges with the archipelago's geological island bank platforms (Brown and Diesmos 2009; Brown et al. 2013) is the curious case of the Leyte Chorus Frog, *K. c. stickeli* that co-occurs with *K. c. meridionalis* on the Mindanao PAIC islands (Mindanao, Samar, Leyte, and Bohol). Described from a single collecting event in 1944, comprising three specimens from northern Leyte Island on the Mindanao PAIC (Inger 1954), *K. c. stickeli* is reported to possess an intermediate phenotype, bearing moderately expanded finger and toe discs; this taxon was never again recorded over the second half of the last century. Despite the existence of a sustained field research program on Leyte Island and the other four major landmasses of the Mindanao PAIC faunal region, plus numerous visits to the precise locality from where the original specimens were collected, this phenotypically unique but elusive subspecies remained surrounded by confusion, concern, and conjecture for 80 years —particularly in the conservation community that began to consider the possibility of habitat destruction and extinction (Brown and Alcala 1970; Alcala 1986; Alcala and Brown 1998; Diesmos et al. 2015; Sanguila et al. 2016).

On the Mindanao PAIC, another species of *Kaloula* outside of the *conjuncta* complex co-occurs with *K. c. meridionalis* and *K. c. stickeli*. The ground frog *K. picta* has narrow digital discs and is a widespread species that occurs throughout the Philippine archipelago (Fig. 1). A previous phylogenomic study of the *Kaloula* adaptive radiation, based on 1867 ultraconserved elements (UCEs), clarified the phylogenetic affinity of *K. c. stickeli* and *K. picta* and further showed that a specimen matching the phenotype of *K. c. stickeli* shared alleles with *K. c. meridionalis* and *K. picta* (Alexander et al. 2017). These results, especially when considered in tandem with the overlapping ranges and intermediate phenotype of *K. c. stickeli* (moderately expanded digital discs vs. narrow in *K. picta* and wide in *K. c. meridionalis*) suggest that *K. c. stickeli* could potentially be a hybrid between *K. picta* and *K. c. meridionalis*. However, the Alexander et al. (2017) study only included one taxon per island, which was not sufficient to elucidate fine-scale population genetic structure. Therefore, the hybridization hypothesis is still untested, and the 80-year-old mystery of the Leyte Chorus Frog remains unresolved.

In this study, we used denser population-level sampling from the Mindanao PAIC (Fig. 1) and implemented a target-capture approach using the FrogCap protocol (Hutter et al. 2022) to determine whether *Kaloula conjuncta stickeli* could be a hybrid between *K. c. meridionalis* and *K. picta*. We included sampling from all large islands of the Mindanao faunal region, and all localities from which genetic material of the three focal taxa has been sampled (Fig. 1). Finally, to provide a novel continuous phenotypic perspective, we collected additional quantitative morphometric data from all known specimens representing the "*K. c. stickeli*-like" intermediate phenotype from Mindanao (Blackburn et al. 2013) and re-examined the morphology of the original type-series specimens of *K. c. stickeli* from Leyte (Inger 1954). Here, we ask if true *K. c. stickeli* (based on type material) has a distinct phenotype, whether that phenotype is shared with other "*stickeli*-like intermediate" populations on Mindanao, and, ultimately, whether we can successfully use genomic data to infer the evolutionary

origins of the mysterious Leyte Chorus Frog and clarify its taxonomic status.

MATERIALS AND METHODS

Sampling

Genomic data were obtained for 45 samples including five outgroup and 40 ingroup samples (Table S1). The ingroup comprised taxa which were morphologically and/or genetically identified as *Kaloula conjuncta meridionalis* (wide toepads; $n = 20$), *K. c. stickeli* (intermediate toepads; $n = 3$), or *K. picta* (narrow toepads; $n = 17$) based on morphological examination and verified by genetic results from previous studies (Blackburn et al. 2013; Alexander et al. 2017). Because fresh or archival samples from the type series were not available, we used genetic material from a new population of putative *K. c. stickeli* (Fig. 1) that was determined based on morphological comparison with the type material (see below). New morphological data were also collected from 37 specimens of the focal taxa. These data were combined with morphological data from Blackburn et al. (2013) to boost sampling size and to capture more variation. Details on morphological data collection and analyses are provided below.

Morphological Data and Analyses

Due to the small number of female specimens per taxon, only male specimens were analyzed. A total of 53 adult male specimens were assessed including the taxa *Kaloula conjuncta conjuncta* ($n = 10$), *K. c. meridionalis* ($n = 11$), *K. c. stickeli* ($n = 7$), and *K. picta* ($n = 25$). To verify the identity of putative *K. c. stickeli* collected from a different locality on Mindanao ($n = 5$), we compared them with two specimens from the type locality on which the subspecies description was based. We measured and analyzed nine morphological characters that were determined to be informative through previous analyses (Blackburn et al. 2013): snout-vent-length (SVL), tip of snout to cloaca; head width (HW), measured across the jaw articulations; snout length (SNL), tip of snout of anterior corner of orbit; forearm length (ForearmL), inflection of wrist to inflection of elbow; 3rd finger disc width (3FDW), maximum width of 3rd finger disc; femur length (FemurL), cloaca to inflection of knee; tibia length (TibiaL), inflection of knee to inflection of ankle; 3rd toe disc width (3TDW), maximum width of 3rd toe; extent of 3rd toe web (3ToeWeb), tip of 3rd toe to point of web attachment measured on the postaxial side of toe. Specimens were obtained from the museum holdings of the University of Kansas Natural History Museum (KU), Texas Natural History Collections, The University of Texas at Austin (TNHC), and the Field Museum of Natural History, Chicago (FMNH). Raw measurements of morphological data are included in Table S2.

To eliminate bias stemming from ontogenetic variation (Chan and Grismer 2021), we performed allometric body-size correction using the Thorpe method (Thorpe 1983) implemented in the *GroupStruct* R package (Chan and Grismer 2022). We then used principal components analysis (PCA) to find the best low-dimensional representation of variation in the data to determine whether morphological variation could form the basis of detectable group structure. ANOVA was used to determine whether differences in character means were significantly different. The allometric body-size correction, PCA, and ANOVA analyses were performed using the *morpho_struct* function in the *GroupStruct* R package.

Genomic Data

We designed a set of markers for population genetics based on exons from the main FrogCap Reduced-Ranoidea marker set (Hutter et al. 2022). We selected markers found in the genus *Kaloula* which were not part of the same gene and which were separated by at least 100,000 base pairs (bp) to reduce the chances of linkage due to genomic proximity. We also included 47 ultra-long exons (> 5000 bp) not previously included in the Ranoidea V1 set. The Reduced-Ranoidea set was designed after the 40 K Ranoidea V1 set that had been successfully tested across the entire superfamily Ranoidea (Hutter et al. 2022). We began with alignments from the 24-sample Ranoidea set and reduced the probe set to markers successfully captured across 75% or greater of the samples, resulting in 10,274 markers. Next, we filtered the markers as follows: (1) UCEs were excluded; (2) the largest exon within 100,000 bp of another exon on the *Nanorana* genome was retained to reduce potential genetic linkage due to recombination; and (3) all exons greater than 500 bp were included. This final candidate set included 4312 exons. To accommodate the probe limit, exons were randomly deleted until the 20,020 bait limit was reached, resulting in 3161 exons retained. Our final Reduced-Ranoidea marker set included 3247 markers targeting 1,519,233 bp of the genome.

The bioinformatics pipeline for filtering adapter contamination, assembling contigs, trimming, and alignment is scripted in R (R Core Team 2014), available at https://github.com/chutter/FrogCap-Sequence-Capture (bioinformatics-pipeline_stable-v1). Our final set of matching markers was aligned on a marker-by-marker basis using MAFFT (Katoh and Standley 2013) local pair alignment. We screened each alignment for samples that were ≥40% divergent from consensus sequences, which were almost always incorrectly assigned contigs. Alignments were retained if they included four or more taxa, had length ≥100 bp, and mean sample specificities (i.e., the "breadth of coverage" of the sample; see below) ≥50% across the alignment to prevent non-overlapping segments of the alignment. We then separated alignments into two datasets: (1) "Exons-only," which included only exonic contigs (with intronic regions trimmed from each alignment using the *Nanorana* genome sequence reference exon as a guide). Each Exons-only alignment was adjusted to be in an open-reading frame in multiples of three bases and trimmed to the largest reading frame that accommodated >90% of the sequences; (2) "All-markers," which included the entire matching contig to the reference marker.

Variant Discovery.—variant calling for single nucleotide polymorphisms (SNPs) was conducted through a custom pipeline in R which is available at https://github.com/chutter/FrogCap-Sequence-Capture (variant-pipeline_stable-v1). We used GATK v4.1 (McKenna et al. 2010) following developer best practices recommendations for discovering and calling variants (Van der Auwera et al. 2013). To discover potential variant data (e.g. SNPs, indels), we used a consensus sequence from each alignment from the target group as a reference and mapped cleaned reads back to reference markers from each sample. We used BWA (Li 2013) to map cleaned reads to our reference markers, adding the read group information (e.g. Flowcell, Lane, Library) obtained from the fastq header files. Next, we used SAMTOOLS (Li et al. 2009) to convert the mapped reads' SAM file to a cleaned BAM file, and then merged BAM files with our unmapped reads, as required for downstream analyses. We used the program PICARD to mark exact duplicate reads that may have resulted from optical and PCR artifacts and reformatted each dataset for variant calling. To locate variant and invariant sites, we used GATK to generate a preliminary variant dataset using the GATK program *HaplotypeCaller* to call haplotypes in GVCF format for each sample individually.

After processing each sample, we used the GATK *GenomicsDBImport* program to aggregate samples from separate datasets into their own combined database. Using these databases, we used the *GenotypeGVCF* function to genotype the combined sample datasets and output separate vcf-formatted files for each marker containing variant data from all samples for final filtration. The preliminary variant set was filtered into a final dataset as follows: (1) All variants were kept after moderate filtering to remove probable errors filtered at a PHRED quality score >5; (2) High-quality variants were kept including SNPs, MNPs (multi-nucleotide polymorphisms), and indels filtered at a quality >20; (3) SNPs were chosen after high-quality filtering (quality >20); and (4) our final dataset consisted of one high-quality SNP from each exon which was most variable across samples.

Phylogenomics

Phylogenetic trees were inferred separately for the Exons-only and All-markers datasets using maximum likelihood (ML) and summary coalescent methods. Individual exon alignments were concatenated and summarized using the program SEGUL (Handika and Esselstyn 2024). A partitioned (by marker) ML tree was inferred using the program IQ-TREE 2 (Minh et al. 2020) with the GTR + F + I + G4 substitution model applied to each partition following recommendations by Chan et al. (2020a). Branch support was assessed using 1000 ultrafast bootstrap replicates (Hoang et al. 2017). IQ-TREE was also used to estimate individual gene trees for a summary coalescent species tree analysis using weighted-ASTRAL-III (Zhang et al. 2018), weighted by gene tree uncertainty (Zhang and Mirarab 2022).

Population Structure

Population structure/ancestry was estimated using the program sNMF (Frichot et al. 2014) through the R package LEA (Frichot and François 2015). Ancestry coefficients were calculated for 1–7 ancestral populations (K) with 100 replicates for each K. The cross-entropy criterion was used to determine the best K based on the prediction of masked genotypes.

Table 1. Summary statistics of the two genomic datasets analyzed in this study.

	All loci	Exons only
General summary:		
Total taxa	45	45
Total loci	2532	1077
Total sites	1,636,027	457,443
Total characters	67,969,852	19,038,711
Missing data	4,160,706	0
% Missing data	6.12%	0%
Characters	67,969,852	19,038,711
Nucleotides	63,633,409	18,652,746
Alignment and taxon summary:		
Min length	160 bp	111 bp
Max length	7037 bp	5301 bp
Mean length	646.14 bp	424.74 bp
Min taxa	22	22
Max taxa	45	45
Mean taxa	40.81	40.9
Matrix completeness:		
100% taxa	927	389
70% taxa	2288	981
50% taxa	2532	1077
Variable and PIS sites:		
Variable sites	272,498	36,157
% Variable sites	16.66%	7.90%
PIS sites	140,613	17,624
% PIS sites	8.59%	3.85%

Because sNMF can produce inaccurate results when sampling is unbalanced, we also inferred population structure using the program PopCluster, which implements a scaling scheme to account for highly unbalanced sampling (Wang 2022). To estimate the most likely *K*, PopCluster implements two estimators. The D_{LK2} estimator is based on the second-order rate of change of cluster log-likelihoods, and is similar in principle to the Evanno method (Evanno et al. 2005) except it does not use the mean and standard deviation of log-likelihood values among replicate runs for a given *K*. The second estimator, F_{STIS} is based on Wright's *F*-statistics in which the best *K* has the "strongest" population structure, characterized by high differentiation (F_{ST}) of each inferred cluster and low deviation from Hardy-Weinberg equilibrium (F_{IS}) within each inferred cluster (Wang 2022). Outgroups were removed before conducting the sNMF and PopCluster analyses.

We also used dimension reduction to uncover underlying group structure/clustering. We implemented the *t*-distributed stochastic neighbor embedding algorithm (*t*-SNE) that can capture local structure of high-dimensional data within 2–3 dimensions (van der Maaten and Hinton 2008; Li et al. 2017; Derkarabetian et al. 2019). This analysis was performed with the R package *Rtsne* (Krijthe 2015) using the following settings: perplexity = 10, max iterations = 1,00,000, theta = 0.0.

Hybridization and Population Genetics

We used the NANUQ algorithm to test for hybridization (Allman et al. 2019). NANUQ was implemented through the R package MSCQuartets v1.1.0 using very stringent alpha (1*e*-40) and large beta (0.95) values for hypothesis testing, as recommended by the developers. Alpha and beta determine the *p*-value significance levels for testing the null hypothesis of no hybridization and a polytomy (star tree), respectively, for each quartet. Therefore, a smaller alpha requires stronger evidence to support hybridization (weighted for fewer reticulations), whereas a larger beta applies a more conservative test for a star tree (weighted for less polytomies). Test results on quartet counts were used to calculate a network quartet distance matrix between taxa and these results were visualized as a splits graph under the Neighbor-Net algorithm using the software SplitsTree4 (Huson and Bryant 2006). We then performed bootstrap resampling in the program HyDe (Blischak et al. 2018) to estimate the degree of admixture within the putative hybrid population.

To further characterize hybridization and patterns of ancestry, we used a triangle plot to visualize hybrid index and interclass heterozygosity derived from ancestry-informative markers. SNPs were filtered using allele frequency difference thresholds to identify SNPs that are differentially fixed in the two putative parental species. The triangle plot analysis was implemented through the R package triangulaR (Wiens and Colella 2024). SNPs were similarly filtered and used in an ancestry painting analysis to investigate patterns of homozygosity and heterozygosity of putative hybrids using custom scripts available at https://github.com/millanek/tutorials/tree/master/analysis_of_introgression_with_snp_data. Finally, we used the R package PopGenome (Pfeifer et al. 2014) to calculate and compare nucleotide diversity and F_{ST} among populations. A sliding window approach was implemented with a window size of 100 bp with 25 bp between windows to accommodate the relatively low number of SNPs.

RESULTS

Genomic Data and Phylogenetic Inference

The All-markers dataset contained 2532 loci (272,498 variable sites; 140,613 parsimony informative sites; 6.12% missing data) whereas the Exons-only dataset contained 1077 loci (36,157 variable sites; 17,624 parsimony informative sites; 0% missing data). Additional summary statistics for both datasets are presented in Table 1. Because both datasets produced similar and strongly supported topologies at the clade level, and the All-markers dataset contained significantly more loci and informative sites, we focused subsequent analyses on the All-markers dataset. For variant calling, we obtained 2529 unlinked SNPs after filtering for one high-quality SNP per locus.

All phylogenetic analyses (weighted ASTRAL-III and IQ-TREE) inferred *Kaloula conjuncta meridionalis* and *K. picta* as highly supported clades, whereas the three *K. c. stickeli* samples were positioned between the *K. c. meridionalis* and *K. picta* clades in a stepwise, successively-branching, non-monophyletic pattern. The phylogenetic analyses also revealed geographic substructure within *K. c. meridionalis* (Fig. 2A; left).

Genetic Structure and Morphological Structure

The sNMF population structure analysis inferred $K = 3$–5 as the optimal number of ancestral populations—$K = 3$, $K = 4$, and $K = 5$ had similar cross-entropy values (Supplementary material Fig. S1). At $K = 3$, *Kaloula conjuncta meridionalis*, *K. picta*, and *K. c. stickeli* were inferred as distinct, non-admixed populations, whereas at $K = 4$, *K. c. meridionalis* was further partitioned into two separate populations consistent with the geographic subclades with small amounts of admixture in two samples. At $K = 5$, *K. picta* was further separated into two non-admixed populations. At $K = 3$–5, *K. c. stickeli* was inferred as a distinct, non-admixed population. (Supplementary material Fig. S2). The PopCluster analysis inferred $K = 2$ (F_{STIS}) and $K = 3$ (D_{LK2}) as the best *K*'s. At $K = 2$, *K. c. meridionalis* and *K. picta* were inferred as distinct, non-admixed populations, while *K. c. stickeli* was inferred to be ~50% admixed with alleles from *K. c. meridionalis* and *K. picta* (Fig. 2A). At K = 3, *K. c. meridionalis*, *K. picta*, and *K. c. stickeli* were inferred as distinct, non-admixed populations (Supplementary material Fig. S2).

The PCA on morphological data (Fig. 2B) and t-SNE analysis on SNP data (Fig. 2C) clearly separated *Kaloula conjuncta meridionalis*, *K. picta*, and *K. c. stickeli* into three distinct clusters. For morphological data, *K. c. stickeli* from Mindanao Island (sequenced in this study) clustered with *K. c. stickeli* from the type locality (Leyte Island; not sequenced), confirming that populations from both islands are morphologically similar and represent a distinct phenotype that is distinguishable from *K. c. meridionalis* and *K. picta* (Fig. 2B). The t-SNE analysis on SNP data, further separated the geographic subclades within *K. c. meridionalis* but this

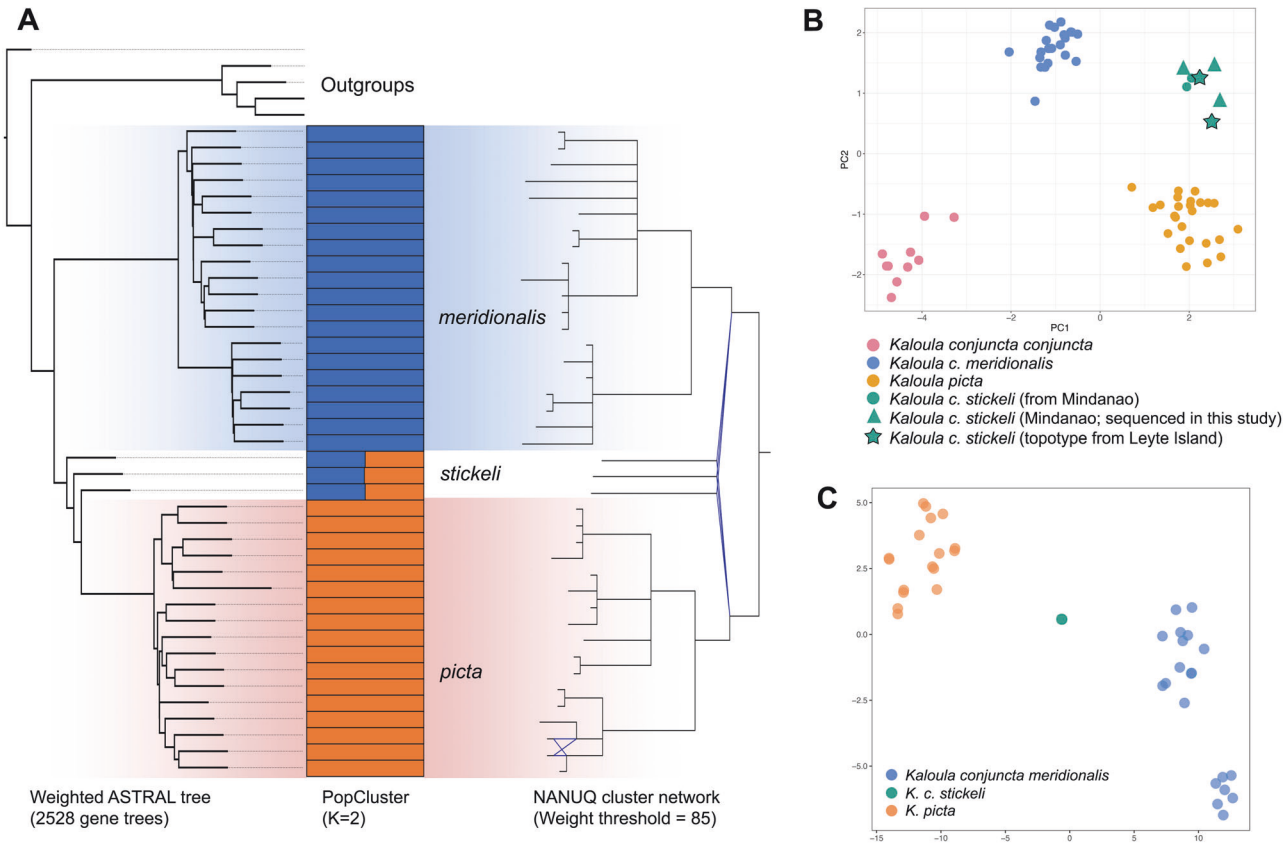

Fig. 2 Results of the phylogenetic, network, and dimension reduction analyses. A Weighted ASTRAL tree and the cluster network inferred from the NANUQ analysis where blue lines show reticulation events. A barplot of ancestry coefficients from the PopCluster analysis at the optimal $K = 2$ (middle) is shown between the ASTRAL tree and NANUQ cluster network. **B** PCA plot based on nine morphological characters. **C** t-SNE plot based on 2529 SNPs.

separation was not reflected in the morphological PCA. The ANOVA on morphological data showed that *K. c. stickeli* differed significantly from *K. picta* across all examined characters and differed significantly from *K. c. meridionalis* in hindlimb length (femur and tibia), finger and toe tip width, and degree of foot webbing (Supplementary material Fig. S3). The genetic divergence between *K. c. stickeli* and both *K. c. meridionalis* and *K. picta* was high ($F_{ST} = 0.717$ and 0.746, respectively) and slightly lower compared to *K. c. meridionalis* vs. *K. picta* ($F_{ST} = 0.864$). The sliding window F_{ST} analysis corroborated these results and showed that all three taxa were genetically highly differentiated from each other (median $F_{ST} > 0.7$; Supplementary material Fig. S4).

Hybridization

The NANUQ quartet hypothesis test showed that a substantial number of quartets were consistent with a quartet network containing a 4-cycle (red triangles; Fig. 3), indicating that many quartets showed signatures of hybridization, even at a very stringent threshold (alpha = 1e-40). These 4-cycle quartets significantly deviated from the expectations of the network multispecies coalescent and exhibited distinct quartet frequency patterns. Quartet topologies that inferred *Kaloula conjuncta stickeli* as the closest relative with either *K. c. meridionalis* or *K. picta* were the dominant topologies and occurred in statistically similar frequencies, while other topological groupings had very low frequencies (Fig. 3). A complete table summarizing the frequencies and *p*-value of all possible quartet combinations is presented in Table S3. The results of NANUQ were then converted into a cluster network in SplitsTree, applying a weight filtering threshold of 85 to filter out weakly supported splits and hybridization edges. The cluster network clearly showed that *K. c. stickeli* is a hybrid between *K. c. meridionalis* and *K. picta* (Fig. 2A; right).

For the triangle plot and ancestry painting analyses, a total of 834 SNPs were retained after applying a stringent allele frequency threshold of 1.0. The triangle plot showed that *Kaloula conjuncta stickeli* had an intermediate hybrid index and high heterozygosity (Fig. 4A), while the ancestry painting analysis showed that *K. c. stickeli* was heterozygous at most SNPs (Fig. 4B). The HyDe analysis corroborated these results and showed that admixture was high and within a narrow range of 0.51–0.535 (Fig. 4C). Taken together, these analyses provide strong evidence that *K. c. stickeli* is an F1 hybrid between *K. c. meridionalis and K. picta*. This is further supported by the nucleotide diversity of *K. c. stickeli*, which is significantly higher than *K. c. meridionalis* and *K. picta* (Fig. 4D).

DISCUSSION

The Misleading Features of Hybrids

Our results demonstrate that the subspecies *K. conjuncta stickeli* is an F1 hybrid between *K. c. meridionalis* and *K. picta*. Even with robust genomic data, F1 hybrids can still be challenging to characterize because they are phenotypically and genetically divergent from the parentals ($F_{ST} > 0.7$), they appear as a distinct group in clustering-based analyses (Fig. 2C), and they manifest the Artefactual Branch Effect where hybrids are inferred as distinct, sometimes non-monophyletic lineages (Chan et al. 2020b, 2022, 2023; Dolinay et al. 2021; Pyron et al. 2023) (Fig. 1A). Moreover, we showed that population structure methods can also struggle to identify hybrids. The optimal number of ancestral populations (*K*) inferred by the sNMF algorithm ($K = 3$–5) identified the hybrids as having a distinct ancestral genotype. At

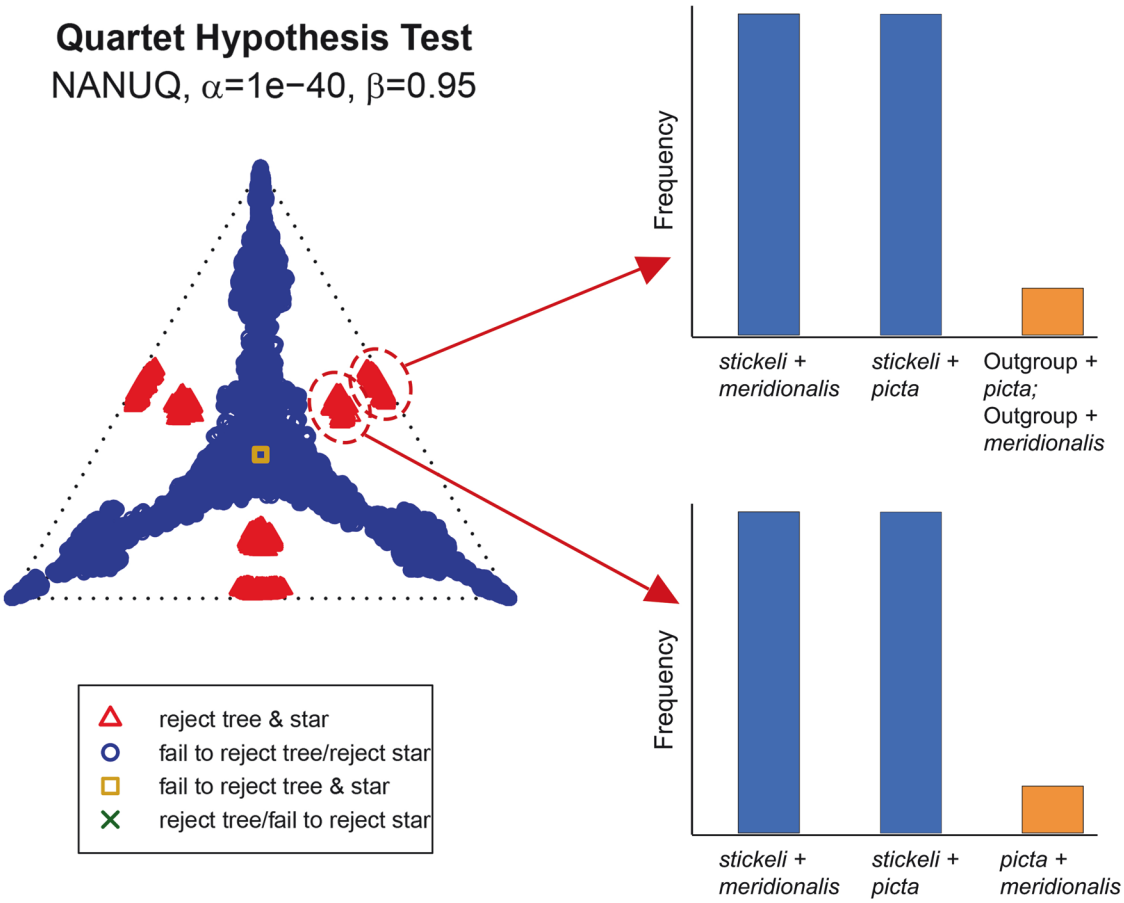

Fig. 3 A simplex plot showing the results of the NANUQ quartet hypothesis test at a stringent alpha of 1e-40. Blue circles indicate quartets that are tree-like while red triangles indicate quartets that represent a 4-cycle network (not tree-like). A histogram illustrating the frequency patterns of non tree-like quartets are shown on the right. The blue bars show that the frequency of quartets that inferred *Kaloula conjuncta stickeli* as the closest relative with either *K. c. meridionalis* or *K. picta* are the dominant topologies and occur in statistically equal frequency. The complete table of all possible quartet frequencies and their associated *p*-value is presented in Supplementary Material Table S3.

$K = 2$, the hybrids were correctly inferred but this K had a significantly poorer cross-entropy score compared to $K = 3$–5. Interestingly, a different algorithm that adjusts for highly imbalanced sampling implemented in the program PopCluster inferred $K = 2$ as the optimal model and showed that the hybrid genomes are derived from a distinct combination of ancestry components. This discrepancy could indicate that imbalanced sampling can bias results in population structure analyses (Wang 2022). Overall, we showed that integrating network multispecies coalescent (Figs. 2A, 3) and population genetic approaches (Fig. 4) are more effective at inferring hybridization because they can assess the connectance of lineages, in contrast to conventional tree- and distance-based methods that focus on quantifying divergence, while assuming that gene flow is absent (Chan et al. 2020b, 2022, 2023).

What is '*Kaloula conjuncta stickeli*?'. We hypothesize that *K. conjuncta stickeli* is unlikely to reproduce beyond the F1 generation or to backcross with either parental due to its rarity in nature and the absence of any other genetic intermediates. These hybrids are also unlikely to be confused with other sympatric species due to their distinct and easily-distinguishable phenotype. Despite decades of field surveys across the Mindanao PAIC (Brown et al. 2013; Diesmos et al. 2015), only seven hybrid specimens have ever been collected, including the three specimens on which the description of the subspecies was originally based (Inger 1954). In contrast, the two parental species *K. c. meridionalis* and *K. picta* are widely distributed and relatively abundant throughout their ranges (Fig. 1). Moreover, both parental species represent highly specialized and divergent ecotypes which occupy distinct niches (structural microhabitats). *Kaloula conjuncta meridionalis* is a scansorial shrub and understorey tree dweller that breeds in ephemeral pools within undisturbed, or relatively undisturbed forest, whereas *K. picta* is a semifossorial burrowing frog that breeds in puddles in highly disturbed open habitats such as rice fields and ditches. Under natural circumstances, these two species would not come into contact. However, deforestation for agriculture has created at least two artificial contact zones between these species which enabled them to hybridize. Field observations by RMB noted that females of the scansorial shrub ecotype *K. c. meridionalis* were frequently ambushed by males of the ground species (*K. picta*) as they crossed open ground (cleared locally due to deforestation), attracted to a mixed chorus of *K. c. meridionalis* and *K. picta* males calling in temporary pools. Within undisturbed forested habitats, we suspect that the shrub ecotype (*K. c. meridionalis*) traverses vegetation during the breeding season and is likewise unlikely to encounter the ground ecotype (*K. picta*).

We consider this hybridization event to be an unnatural, human-mediated phenomenon which may only be made possible by clearing forests for agriculture, thus facilitating contact between the scansorial forest (*K. c. meridionalis*) and ground, open habitat (*K. picta*) species. Although the West Visayan PAIC islands (Negros, Panay, Guimaras, and Cebu) were heavily deforested during Spanish colonization, it was the onset of the Industrial Revolution in the early 20th Century that saw the rapid wholesale removal of vast forests of the Mindanao PAIC islands (Mindanao, Leyte, Samar), beginning with

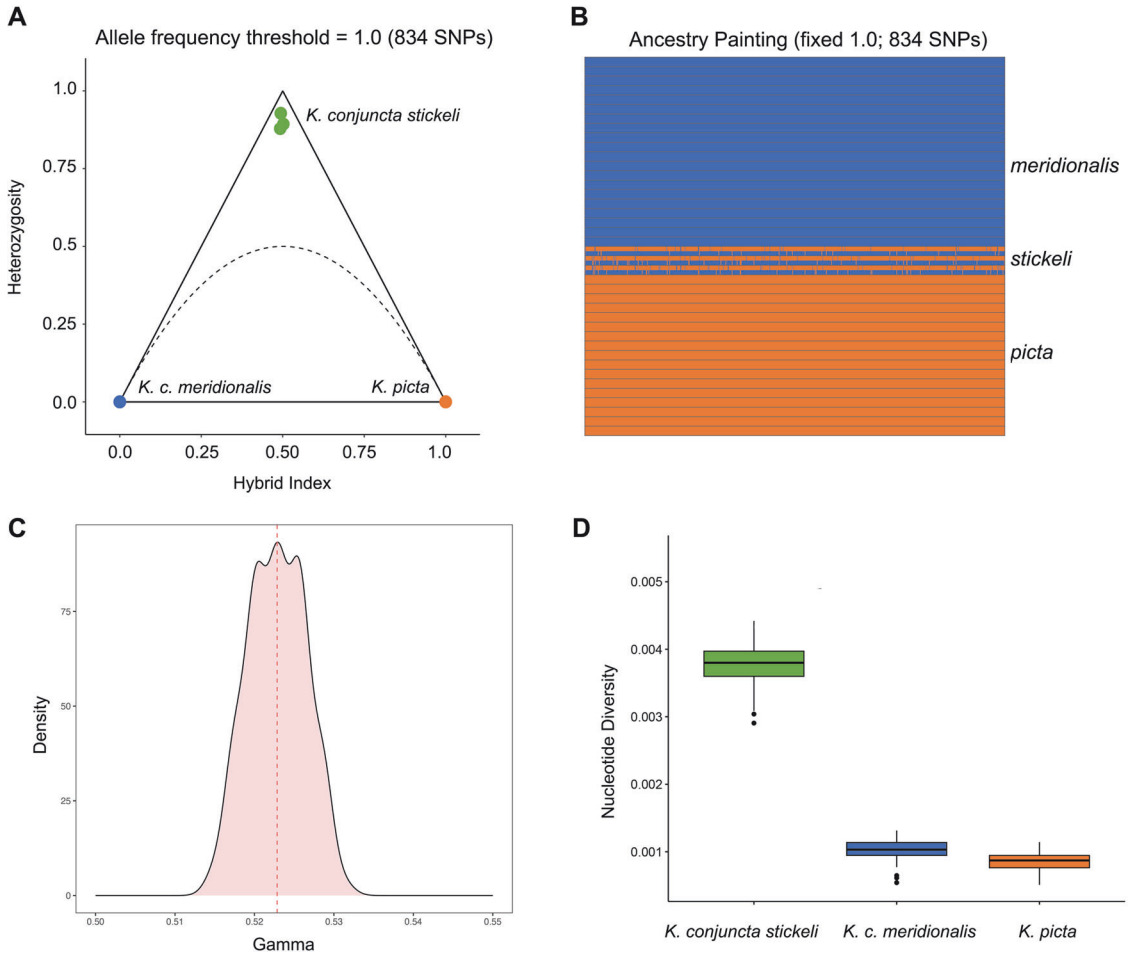

Fig. 4 Population genetic results confirming the F1 status of *Kaloula conjuncta stickeli*. **A** Triangle plot of hybrid index and heterozygosity at an allele frequency threshold of 1.0 (834 SNPs). *Kaloula conjuncta stickeli* is inferred as an F1 hybrid based on high heterozygosity (~ 1.0) and intermediate hybrid index (0.5). **B** Bar plot from the ancestry painting analysis showing that *K. c. stickeli* is heterozygous across most SNPs that were differentially fixed across *K. c. meridionalis* and *K. picta*. **C** Results of the HyDE bootstrap resampling analysis showing that admixture (Gamma) was high and within a narrow range of 0.51–0.535. **D** Nucleotide diversity calculated from a sliding window analysis.

the low-elevation coastal plains of these landmasses (Environmental Center of the Philippine Foundation 1998; Catibog-Sinha and Heaney 2006). During the 1900s, during periods of Japanese and U.S. occupation, it is estimated that as much as 50–70% of original lowland Philippie forersts were removed from areas accessible to coasts (Kummer 1992; Brown and Diesmos 2009)—likely including the type locality of *K. c. stickeli* (Carigara, Leyte). We note that the original collector's notes from November 1944 indicate that the type specimens were taken from a deforested, grassy area in a mixed chorus, with *K. picta*, at Carigara Town (Inger 1954), a coastal lowland area, adjacent to the northern-most extent of the central spine of mountains that run south-to-north through Leyte Island. Within the larger Municipality of Carrara, montane foothills extend to approximately 1000 m above sea level, suggesting that forest patches or fragments of suitable habitats for *K. c. meridionalis* may have existed until the date of collection—all of which is consistent with our own observation of mixed choruses and apparent hybridization (collection of a the "*K. c. stickeli*-like" intermediate) between *K. c. meridionalis* and *K. picta* in the late 1990s, in patchy forest framents at mid-elevation on Mindanao Island's Mt. Apo (Baracatan, Calinan, Davao, Mindanao). Although the association between hybridization and deforestation remains circumstantial, this study sets the stage for future studies to explore the consequences of anthropogenic disturbances on the evolutionary trajectories of species.

To date, only two such hybrid zones have been detected, one from the Island of Leyte where the hybrids were originally collected, and a second from the southern region of Mindanao in the Davao Del Sur Province (evinced by one of the author's field observations from the early 1990s, and reinforced by genomic data presented in this study). Taken together, the results from this study, the paucity of other hybrid samples (Blackburn et al. 2013; Diesmos et al. 2015; Alexander et al. 2017) despite decades of targeted fieldwork in the region (Brown et al. 2013), the absence of different recombinant genotypes representing other hybrid classes, and our natural history observations of reproductive behavior provide strong circumstantial evidence that *K. c. stickeli* is an F1 hybrid which is unable to reproduce. Despite its unique phenotype and genotype, *K. c. stickeli* appears to be inviable, and thus, should not be assigned a species or subspecies rank, nor can it be considered conspecific with either *K. c. meridionalis* or *K. picta*. Instead, we propose that *K. conjuncta stickeli* be taxonomically and nomenclaturally considered a hybrid, i.e.: "*K. c. meridionalis* x *K. picta*."

Broader Implications

This study is the first demonstration of hybridization in narrow-mouthed frogs. Although rare, human-mediated hybridization by way of habitat modification has been documented before. At a site in Auburn, Alabama, USA, the mowing and removal of vegetation surrounding ponds facilitated the hybridization between two species of frogs of the family Hylidae: *Dryophytes cinereus* and *D. gratiosus* (Mecham 1960; Lamb and Avise 1986). Males of *D. gratiosus* typically call from the water surface of

ponds, while *D. cinereus* calls from elevated perches on vegetation surrounding ponds. However, in this instance, regular removal of vegetation around ponds forced males of *D. cinereus* to call from the ground. Females of *D. gratiosus* that were attracted to conspecific calls from the pond had to encounter males of *D. cinereus*, that amplexed and hybridized with female *D. gratiosus* as they approached the study-area pond.

Given the wide array of adaptive ecotypes and the presence of other subspecies in Asian narrow mouth frogs, it is possible that there is hybridization in other species/subspecies of *Kaloula*. The study by Alexander et al. (2017) suggested the presence of gene flow among several other microhylid taxa, although this remains to be tested rigorously with more robust data and denser taxon sampling. Although our study showed that the "*stickeli*" (*K. c. meridionalis* x *K. picta*) hybrid is likely inviable, it remains possible that other hybridization events between different species could produce viable offspring and could generate unique phenotypic and genotypic lineages, especially given the overlapping distribution range of many microhylid taxa (Fig. 1). It is thus plausible that hybridization played a role in the adaptive radiation of *Kaloula* (Blackburn et al. 2013; Brown et al. 2013). This is known as the hybrid adaptive radiation hypothesis—whereby hybridization can generate exceptional genetic variation, which may facilitate the exploitation of new ecological opportunities, sometimes resulting in an adaptive radiation (Seehausen 2004). This phenomenon is theoretically possible (Kagawa and Seehausen 2020) and has been demonstrated to occur in nature (Mallet 2009; Meier et al. 2017; Grant and Grant 2019). Furthermore, although the width of the finger and toe discs is a clear intermediate between the wide (*K. c. meridionalis*) and narrow (*K. picta*) widths of the parentals, other traits associated with body size (e.g., snout-vent-length and limb length) appear to be additive, wherein the hybrid is larger than both parental species (Supplementary material Fig. S3). This contrasting pattern suggests a complicated dynamic of inheritance involving incomplete dominance, polygenic inheritance, or heterosis that could potentially lead to an array of adaptive traits. Overall, our results indicate that the evolutionary history and systematics of Philippine *Kaloula* are still poorly understood, but also suggest that the mechanisms underlying adaptive radiation in the genus are recoverable by, and demand, nuanced and in-depth investigation drawing on suites of analyses beyond the current industry-standard.

DATA AVAILABILITY

Raw sequence reads have been deposited in the GenBank SRA as BioProject accession PRJNA1133022.

REFERENCES

Abbott RJ, Barton NH, Good JM (2016) Genomics of hybridization and its evolutionary consequences. Mol Ecol 25:2325–2332

Alcala AC (1986) Guide to Philippine Flora and Fauna. Vol X, Amphibians and Reptiles. Natural Resource Management Center, Ministry of Natural Resource Management Center, Ministry of Natural Resources and the University of the Philippines, Manila, Philippines

Alcala AC, Brown RM (1998) Philippine Amphibians: An Illustrated Field Guide. Bookmark Press, Makati City, Philippines

Alexander AM, Su YC, Oliveros CH, Olson KV, Travers SL, Brown RM (2017) Genomic data reveals potential for hybridization, introgression, and incomplete lineage sorting to confound phylogenetic relationships in an adaptive radiation of narrow-mouth frogs. Evolution 71:475–488

Allman ES, Baños H, Rhodes JA (2019) NANUQ: A method for inferring species networks from gene trees under the coalescent model. Algorithms Mol Biol 14:1–25

Blackburn DC, Siler CD, Diesmos AC, Mcguire JA, Cannatella DC, Brown RM (2013) An adaptive radiation of frogs in a southeast Asian island archipelago. Evolution 67:2631–2646

Blischak PD, Chifman J, Wolfe AD, Kubatko LS (2018) HyDe: A python package for genome-scale hybridization detection. Syst Biol 67:821–829

Brown RM, Siler CD, Oliveros CH, Esselstyn JA, Diesmos AC, Hosner PA et al. (2013) Evolutionary processes of diversification in a model island archipelago. Annu Rev Ecol Evol Syst 44:411–435

Brown WC, Alcala AC (1970) The zoogeography of the herpetofauna of the Philippine Islands, a fringing archipelago. Proc Calif Acad Sci 38:105–130

Brown RM, Diesmos AC (2009). Philippines, Biology. In: Gillespie R, Clague D (eds) *Encyclopedia of Islands*, University of California Press: Berkeley, pp 723–732.

Burbrink FT, Guiher TJ (2015) Considering gene flow when using coalescent methods to delimit lineages of North American pitvipers of the genus *Agkistrodon*. Zool J Linn Soc 173:505–526

Burbrink FT, Crother BI, Murray CM, Smith BT, Ruane S, Myers EA et al. (2022) Empirical and philosophical problems with the subspecies rank. Ecol Evol 12:1–17

Burgon JD, Vences M, Steinfartz S, Bogaerts S, Bonato L, Donaire-Barroso D et al. (2021) Phylogenomic inference of species and subspecies diversity in the Palearctic salamander genus Salamandra. Mol Phylogenet Evol 157:107063

Catibog-Sinha CS, Heaney LR (2006) Philippine Biodiversity: principles and practice. Quezon City, Philippines, Haribon Foundation for Conservation of Natural Resources, pp 495

Chambers EA, Marshall TL, Hillis DM (2023) The Importance of Contact Zones for Distinguishing Interspecific from Intraspecific Geographic Variation. Syst Biol 72:357–371

Chan KO, Grismer LL (2021) A standardized and statistically defensible framework for quantitative morphological analyses in taxonomic studies. Zootaxa 5023:293–300

Chan KO, Grismer LL (2022) GroupStruct: An R package for allometric size correction. Zootaxa 5124:471–482

Chan KO, Mulcahy DG, 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

Chan KO, Hutter CR, Wood PL, Grismer LL, Brown RM (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 KO, Hutter CR, Wood PLJ, Su Y-CC, Brown RM (2022) 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 KO, Hutter CR, Wood PL, Grismer LL, Das I, Brown RM (2020b) Gene flow creates a mirage of cryptic species in a Southeast Asian spotted stream frog complex. Mol Ecol 29:3970–3987

Coates DJ, Byrne M, Moritz C (2018) Genetic diversity and conservation units: Dealing with the species-population continuum in the age of genomics. Front Ecol Evol 6:1–13

Darwin C (1859) On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life. John Murray, London

De Queiroz K (2020) An updated concept of subspecies resolves a dispute about the taxonomy of incompletely separated lineages. Herpetol Rev 51:459–461

Derkarabetian S, Castillo S, Koo PK, Ovchinnikov S, Hedin M (2019) A demonstration of unsupervised machine learning in species delimitation. Mol Phylogenet Evol 139:106562

Diesmos A, Watters J, Huron N, Davis D, Alcala A, Crombie R et al. (2015) Amphibians of the Philippines, part I: checklist of the species. Proc Calif Acad Sci 62:457–539

Dolinay M, Nečas T, Zimkus BM, Schmitz A, Fokam EB, Lemmon EM et al. (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

Durrant SD (1955) In defense of the subspecies. Syst Zool 4:186–190

Edelman NB, Mallet J (2021) Prevalence and Adaptive Impact of Introgression. Annu Rev Genet 55:265–283

Elgetany AH, van Rensburg H, Hektoen M, Matthee C, Budaeva N, Simon CA et al. (2020) Species delineation in the speciation grey zone—The case of Diopatra (Annelida, Onuphidae). Zool Scr 49:516–534

Environmental Center of the Philippines Foundation (1998) Environment and Natural Resources Atlas of the Philippines. Environmental Center of the Philippines Foundation, Manila, pp 394

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Frichot E, François O (2015) LEA: An R package for landscape and ecological association studies. Methods Ecol Evol 6:925–929

Frichot E, Mathieu F, Trouillon T, Bouchard G, François O (2014) Fast and efficient estimation of individual ancestry coefficients. Genetics 196:973–983

Frost DR (2024). Amphibian Species of the World: an Online Reference. Version 6.2 (last accessed 6 Sep 2024). *Electronic Database accessible at* http://research.amnh.org/herpetology/amphibia/index.html American Museum of Natural History, New York, USA.

Grant PR, Grant BR (2019) Hybridization increases population variation during adaptive radiation. Proc Natl Acad Sci USA 116:23216–23224

Handika H, Esselstyn JA (2024) SEGUL: Ultrafast, memory-efficient and mobile-friendly software for manipulating and summarizing phylogenomic datasets Mol Ecol Resour 24:e13964
Hillis D (2020) The detection and naming of geographic variation within species. Herpetol Rev 51(1):52–56
Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Le SV (2017) UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol 35:518–522
Huang JP, Knowles LL (2016) The species versus subspecies conundrum: Quantitative delimitation from integrating multiple data types within a single Bayesian approach in Hercules beetles. Syst Biol 65:685–699
Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23:254–267
Hutter CR, Cobb KA, Portik DM, Travers SL, Brown RM, Wood PL et al. (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
Inger RF (1954) Systematics and zoogeography of Philippine amphibia. Fieldiana 33:183–531
Kagawa K, Seehausen O (2020) The propagation of admixture-derived adaptive radiation potential. Proc Royal Soc B Biol Sci 287:20200941
Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol 30:772–780
Krijthe JH (2015) Rtsne: T-Distributed Stochastic Neighbor Embedding using a Barnes-Hut Implementation. *URL*: https://github.com/jkrijthe/Rtsne.
Kummer DM (1992) Deforestation in the Postwar Philippines. Ateneo De Manila University Press, Manila, Philippines. pp 178
Lamb T, Avise JC (1986) Directional introgression of mitochondrial DNA in a hybrid population of tree frogs: The influence of mating behavior. Proc Natl Acad Sci USA 83:2526–2530
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al. (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079
Li W, Cerise JE, Yang Y, Han H (2017) Application of t-SNE to human genetic data. J Bioinf Comput Biol 15:1–14
Li H (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at https://doi.org/10.48550/arXiv.1303.3997.
Mallet J, Besansky N, Hahn MW (2016) How reticulated are species? BioEssays 38:140–149
Mallet J (2009) Rapid speciation, hybridization and adaptive radiation in the *Heliconius melpomene* group. In: Butlin RK, Bridle JR, Schluter D (eds) *Speciation and Patterns of Diversity*, Cambridge University Press, pp 177–194.
Mayr E (1982) Of what use are subspecies? Auk 99:593–595
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A et al. (2010) The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303
Mecham JS (1960) Introgressive hybridization between two southeastern treefrogs. Evolution 14:445–457
Meier JI, Marques DA, Mwaiko S, Wagner CE, Excoffier L, Seehausen O (2017) Ancient hybridization fuels rapid cichlid fish adaptive radiations. Nat Commun 8:14363
Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A et al. (2020) IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534
Moran BM, Payne C, Langdon Q, Powell DL, Brandvain Y, Schumer M (2021) The genomic consequences of hybridization. Elife 10:1–33
Patten MA (2015) Subspecies and the philosophy of science. Auk 132:481–485
Payseur BA, Rieseberg LH (2016) A genomic perspective on hybridization and speciation. Mol Ecol 25:2337–2360
Peñalba JV, Runemark A, Meier JI, Singh P, Wogan GOU, Sánchez-guillén R et al. (2024) The role of hybridization in species formation and persistence. Cold Spring Harb Perspect Biol 16:a041445
Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ (2014) PopGenome: An efficient Swiss army knife for population genomic analyses in R. Mol Biol Evol 31:1929–1936
Pyron RA, O'Connell KA, Duncan SC, Burbrink FT, Beamer DA (2023) Speciation hypotheses from phylogeographic delimitation yield an integrative taxonomy for seal salamanders (Desmognathus monticola). Syst Biol 72:179–197
R Core Team (2014) A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
Roux C, Fraïsse C, Romiguier J, Anciaux Y, Galtier N, Bierne N (2016) Shedding light on the grey zone of speciation along a continuum of genomic divergence. PLoS Biol 14:1–22
Sackett LC, Seglund A, Guralnick RP, Mazzella MN, Wagner DM, Busch JD et al. (2014) Evidence for two subspecies of Gunnison's prairie dogs (*Cynomys gunnisoni*), and the general importance of the subspecies concept. Biol Conserv 174:1–11
Sánchez KI, Huesa EGD, Breitman MF, Avila LJ, Sites JW, Morando M (2023) Complex patterns of diversification in the gray zone of speciation: model-based approaches applied to Patagonian liolaemid lizards (Squamata: *Liolaemus kingii* clade). Syst Biol 72:739–752
Sanguila MB, Cobb KA, Siler CD, Diesmos AC, Alcala AC, Brown RM (2016) The amphibians and reptiles of Mindanao Island, southern Philippines, II: The herpetofauna of northeast Mindanao and adjacent islands. Zookeys 624:1–132
Schiaffini MI (2020) Are subspecies (of *Eira barbara*) real? J Mammal 101:1410–1425
Seehausen O (2004) Hybridization and adaptive radiation. Trends Ecol Evol 19:198–207
Taylor SA, Larson EL (2019) Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. Nat Ecol Evol 3:170–177
Thorpe RS (1983) A review of the numerical methods for recognising and analysing racial differentiation. In: Felsenstein J (ed) *NATO ASI Series G1*, Springer-Verlag: Berlin, Heidelberg, pp 404–423.
Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A et al. (2013) From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. Curr Protoc Bioinf 43:11.10.1–11.10.33
van der Maaten L, Hinton G (2008) Visualizing data using t-SNE. J Mach Learn Res 9:2579–2605
Vences M, Miralles A, Dufresnes C (2024) Next-generation species delimitation and taxonomy: Implications for biogeography. J Biogeogr 51:1709–1722
Wang J (2022) Fast and accurate population admixture inference from genotype data from a few microsatellites to millions of SNPs. Heredity 129:79–92
Wiens BJ, Colella JP (2024) triangulaR: an R package for identifying AIMs and building triangle plots using SNP data from hybrid zones. *bioRxiv* 03.28.5871.
Zhang C, Mirarab S (2022) Weighting by gene tree uncertainty improves accuracy of quartet-based species trees. Mol Biol Evol 39:1–22
Zhang C, Rabiee M, Sayyari E, Mirarab S (2018) ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinf 19:15–30

ACKNOWLEDGEMENTS

RMB acknowledges funding from the U. S. National Sciences Foundation for fieldwork (DEB 0073199 & 0743491) and development of genomic resources (DEB 1654388), and thanks the University of Kansas Office of the Provost Research Investment Council for the funding (2013–2014 R.I.C. Level II award 2300207) that initiated conversion of the KU Biodiversity Institute molecular genetics laboratory to collection & analysis of genome-scale data. KOC and PMH thank the University of Kansas Biodiversity Institute and Natural History Museum for postdoctoral fellowship support. RMB thanks Jim McGuire, Chris Austin, Isagani Bulalacao, Susing Babao, and Viscente Yngente for enthusiastic companionship during 1996 fieldwork on Mindanao Island, for sharing their own observations of frog behavior in multiple-species breeding aggregations, and for stimulating discussion regarding the possibility of interspecific hybridization among species of Philippine Kaloula.

AUTHOR CONTRIBUTIONS

KOC, RMB, and PMH conceived the study and helped with writing. PMH collected genomic data and performed preliminary analyses. KOC conducted analyses and wrote the paper's initial draft; all authors approved and contributed to the writing.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41437-025-00748-y.

Correspondence and requests for materials should be addressed to Kin Onn Chan.

Reprints and permission information is available at http://www.nature.com/reprints

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