Multilocus phylogeny of Bornean Bent-Toed geckos (Gekkonidae: Cyrtodactylus) reveals hidden diversity, taxonomic disarray, and novel biogeographic patterns

Hayden R. Davis a,b, Kin Onn Chan c,*, Indraneil Das d, Ian G. Brennan e, Benjamin R. Karin f, Todd R. Jackman a, Rafe M. Brown g, Djoko T. Iskandar h, Izneil Nashriq c, L. Lee Grismer i, Aaron M. Bauer a

a Department of Biology and Center for Biodiversity and Ecosystem Stewardship, Villanova University, 800 Lancaster Avenue, Villanova, PA 19085, USA
b Department of Biology & Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195, USA
c Lee Kong Chian Natural History Museum, 2 Conservatory Drive, 117377, Singapore
d Institute of Biodiversity and Environmental Conservation, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia
e Division of Ecology and Evolution, Research School of Biology, ACT 2602, The Australian National University, Canberra, Australia
f Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, CA 94720, USA
g Biodiversity Institute and Department of Evolutionary Biology, 1345 Jayhawk Boulevard, University of Kansas, Lawrence, KS 66045, USA
h School of Life Sciences and Technology, Institut Teknologi, Bandung, Indonesia
i Department of Biology, La Sierra University, 4500 Riverwalk Parkway, Riverside, CA 92505, USA

ARTICLE INFO

Keywords:
Cryptic species
Species delimitation
Systematics
BioGeoBEARS
Ancestral range reconstruction
Sundaland

ABSTRACT

The gekkonid genus Cyrtodactylus is a highly diverse group of lizards (280 + species), which covers an expansive geographic range. Although this genus has been the focus of many taxonomic and molecular systematic studies, species on the Southeast Asian island of Borneo have remained understudied, leading to an unclear evolutionary history with cascading effects on taxonomy and biogeographic inferences. We assembled the most comprehensive multilocus Bornean dataset (one mitochondrial and three nuclear loci) that included 129 novel sequences and representatives from each known Cyrtodactylus species on the island to validate taxonomic status, assess species diversity, and elucidate biogeographic patterns. Our results uncovered a high proportion of cryptic diversity and revealed numerous taxonomic complications, especially within the C. consobrinus, C. malayanus, and C. pubisulcus groups. Comparisons of pairwise genetic distances and a preliminary species delimitation analysis using the Automatic Barcode Gap Discovery (ABGD) method demonstrated that some wide-ranging species on Borneo likely comprise multiple distinct and deeply divergent lineages, each with more restricted distributional ranges. We also tested the prevailing biogeographic hypothesis of a single invasion from Borneo into the Philippines. Our analyses revealed that Philippine taxa were not monophyletic, but were likely derived from multiple separate invasions into the geopolitical areas comprising the Philippines. Although our investigation of Bornean Cyrtodactylus is the most comprehensive to-date, it highlights the need for expanded taxonomic sampling and suggests that our knowledge of the evolutionary history, systematics, and biogeography of Bornean Cyrtodactylus is far from complete.

1. Introduction

The gekkonid genus Cyrtodactylus is a highly diverse group of geckos with over 280 recognized species, making it one of the most speciose lizard genera in the world (Uetz, 2019). The group spans a broad geographic range extending from South Asia to Melanesia, with a particularly concentrated diversity in Southeast Asia (Bauer et al., 2003, 2002; Brennan et al., 2017; Grismer et al., 2018, 2016a, 2016b, 2014, 2012; Luu et al., 2016; Murdoch et al., 2019; Nazarov et al., 2018, 2008; Pham et al., 2019; Tri, 2011, 2008; Tri et al., 2010).
Throughout their range, new species have been discovered and described at an unprecedented rate, largely due to the implementation of molecular phylogenetic methods (Agarwal et al., 2018; Davis et al., 2019; Grismer et al., 2019, 2018, 2016b, 2014, 2012; Luu et al., 2016; Murdoch et al., 2019; Nazarov et al., 2018; Oliver et al., 2018, 2012; Pham et al., 2019; Rijanto et al., 2015). However, despite the increased focus on Cyrtodactylus biodiversity, the island of Borneo has received relatively little attention (see references above), resulting in much lower species diversity and endemism compared to the surrounding regions (Davis et al., 2019).

Borneo is the third largest island in the world and is divided among three countries: Indonesia (Kalimantan provinces), Malaysia (states of Sabah and Sarawak, and the Federal Territory of Labuan), and the Nation of Brunei. The island is a biodiversity hotspot that has played a key role in the biogeographic history of many taxa (de Bruyn et al., 2014; Grismer et al., 2016c; Karin et al., 2016; Klaus et al., 2013; Matsui et al., 2010; O’Connell et al., 2018; Williams et al., 2017), primarily due to its large size and dynamic geologic history (de Bruyn et al., 2014; Hall, 2013; Myers et al., 2000). Its geologic history, which includes a recently protracted terrestrial connection to mainland Asia and the Indonesian Archipelago (Hall, 2013; Sarr et al., 2019) and close geographic proximity to the Philippines with possible intermittent terrestrial connections (Brown et al., 2013; Brown and Alcala, 1970; Diamond and Gilpin, 1983; Hall, 2013; Inger, 1954), has promoted high levels of in situ diversification, immigration, and emigration (de Bruyn et al., 2014; Evans et al., 2003; Grismer et al., 2016c; Karin et al., 2017, 2016; Klaus et al., 2013; Lohman et al., 2016; O’Connell et al., 2018; Sheldon et al., 2009; Siler et al., 2011; Tougard, 2001; Welton et al., 2014; Williams et al., 2017). Thus, the centralized location of Borneo, which serves as a biogeographic “stepping-stone,” as well as a center of radiation, makes it a crucial region for understanding the evolutionary history, systematics, and biogeographic patterns of Southeast Asian taxa.

Despite Borneo’s status as a biodiversity hotspot, only ten species of Cyrtodactylus are recognized from the island, nine of which are endemic (the range of C. consobrinus extends into Peninsular Malaysia, and Singkpe and Sumatra Islands in Indonesia). By comparison, Peninsular Malaysia, which shared a terrestrial connection with Borneo until the Pleistocene (Sarr et al., 2019), is the type locality for 32 species (Grismer and Quah, 2019; Quah et al., 2019) yet it encompasses a geographic area approximately one-fifth that of Borneo (130,395 km² versus 743,330 km²). A recent study on Sarawak Cyrtodactylus, which provided the first Bornean-focused molecular phylogeny, suggested that the ostensibly low species diversity was not representative the island’s true diversity and likely stemmed from poor geographic and genetic sampling (Davis et al., 2019). The accurate assessment of Bornean Cyrtodactylus diversity, however, is complicated by a high degree of intraspecific variability and interspecific phenotypic similarity across species. Additionally, prior to Davis et al. (2019), only one Bornean species (C. baluensis) had publicly available toptotypic genetic material. This lack of genetic and/or morphological clarity for the recognized Bornean species has thereby led to the misidentification of numerous vouchered specimens and published sequence data (personal observation).

A comprehensive Cyrtodactylus phylogeny published by Wood et al. (2012) provided a broader perspective of the species-level evolutionary relationships within the genus, and showed that Bornean species were not monophyletic; rather they were characterized as belonging to two separate clades [Clades “K” and “M”; Wood et al. (2012: Fig. 2)]. In that dataset, however, only six Bornean species were included with each being represented by just a single sequence. Further, all other subsequent studies that have included Bornean taxa have utilized similarly under-represented datasets (Brennan et al., 2017; O’Connell et al., 2019; Oliver et al., 2012, 2018; Siler et al., 2010; Welton et al., 2010b, 2010b). This limited sampling has been unable to provide a complete and accurate characterization of the biodiversity, biogeography, and evolutionary history on Borneo. For instance, the most recent biogeographic study of Philippine Cyrtodactylus inferred a single colonization event from Borneo (Siler et al., 2010). However, that study only included a handful of Bornean taxa, potentially indicating that the inclusion of additional Bornean representatives could provide a different perspective of biogeographic patterns in that region. Wider geographic sampling at the population-level is therefore necessary to develop a more comprehensive understanding of the evolutionary relationships among Bornean Cyrtodactylus and their surrounding congeners.

To address these issues, we infer the most comprehensive phylogeny of Bornean Cyrtodactylus to date using a multilocus dataset with toptypic genetic data for most taxa and unique populations from Sarawak, Sabah, and Kalimantan. We use our novel phylogenetic estimate to: (1) examine the taxonomic stability of current taxa; (2) assess genetic structure and variation to develop a better understanding of the diversity on the island; and (3) test and re-evaluate prevailing biogeographic hypotheses of how Cyrtodactylus colonized the Philippines.

2. Materials and methods

2.1. Data collection

We conducted field collection in Sarawak between the years 1998–2018 during the months of May through July. Additional tissue samples from Sabah and Kalimantan were obtained from tissue loans. Altogether, we obtained samples from 20 unique localities throughout Borneo, including toptotypic material for Cyrtodactylus cavernicolum, C. consobrinus, C. limajalur, C. muliensis, and C. pubisculus (Fig. 1). We euthanized specimens using a 1% MS-222 solution (IACUC: 1864; Conroy et al., 2009), and liver samples were extracted and stored in 95% ethanol. We preserved voucher specimens using a 10% formalin solution and subsequently transferred to 70% ethanol for long-term storage. Preserved specimens are either deposited at the California Academy of Sciences (CAS) or are pending deposition at the Sarawak Forestry Department (SFD; Table S1). Tissue samples are stored at Villanova University.

2.2. Molecular data

We isolated genomic DNA from our tissue samples using the extraction protocol described in Aljanabi and Martinez (1997). We amplified one mitochondrial (NADH dehydrogenase subunit 2 (ND2)); and three nuclear loci: [matrix remodeling associated 5 (MXRA5); recombination activating gene (RAG1); and the phosphoducin gene (PDC)] using a double-stranded Polymerase Chain Reaction (PCR) using the thermal profile outlined by Davis et al. (2019). The annealing temperature for ND2, RAG1, and PDC was 50 °C and was 55 °C for PDC, with an additional 186 published sequences obtained from GenBank included 16 outgroup taxa that were selected following Wood et al. (2012). All sequences used in this study and their associated GenBank accession numbers are presented in Table S1.

Sequences were assembled and aligned [MAFFT algorithm; (Kato and Standley, 2013)] using Geneious* v11.1.2 (Kearse et al., 2012). Each gene was aligned individually and subsequently concatenated for phylogenetic analyses. The final concatenated dataset included 146 taxa and 3944 bp (ND2: 1519 bp; RAG1: 1040 bp; MXRA5: 941 bp;
2.3. Phylogenetic and divergence time estimation

We performed concatenation analyses due to the relatively low number of genetic markers (1 mitochondrial and 3 nuclear genes) and incomplete taxonomic coverage for the nuclear genes. The concatenated sequence matrix was partitioned by gene and phylogenies were estimated using maximum likelihood (ML) and Bayesian inference. For ML, model selection was performed using ModelFinder (Kalyaanamoorthy et al., 2017) prior to tree inference using the program IQ-TREE (Nguyen et al., 2015). A consensus tree was obtained by summarizing 5000 bootstrap replicates using the ultrafast bootstrap method (UFBoot, Hoang et al., 2017). We considered nodes with UFB values > 95 highly-supported (Minh et al., 2013).

For Bayesian inference, we implemented the program BEAST2 v2.4 (Bouckaert et al., 2014) through the CIPRES Science Gateway (Miller et al., 2010). Model averaging was performed separately on each gene partition using the bModelTest plugin (Bouckaert and Drummond, 2017). We used a relaxed log-normal clock and Yule model as the molecular clock and tree priors respectively, with all other priors set to default values. Two independent MCMC runs were performed, each consisting of 50 million generations. Sampling parameters from both runs were combined using LogCombiner and convergence was assessed using the program Tracer v1.6 (Rambaut et al., 2007). Sampled trees that reached convergence from both runs were combined, with the first 10% discarded as burn-in. Finally, TreeAnnotator was used to construct a maximum clade credibility tree (MCC). We considered nodes with a posterior probability (PP) of 0.95 or above highly supported.

We also estimated a time-calibrated phylogeny using BEAST2 v2.4. Because no fossils are available for any of our samples, we used secondary calibration points derived from the broader study by Wood et al. (2012) that was calibrated using three fossils. We constructed the calibration prior using a normal distribution and adjusted the mean and sigma parameters so that the distribution matches the 95% credible intervals from the primary study. The three secondary calibration points used in this study include the divergence between [1] Cyrto- dactylus semenanjungensis and C. tiomanensis (mean = 24.1, sigma = 3); [2] C. batucolus and C. seribuatensis (mean = 11.4, sigma = 2.5); and [3] C. paradoxus and C. eisenmane + C. grismeri (mean = 20.2, sigma = 3). These points were selected because of the topological congruence of these clades in both this and Wood et al.’s study.

2.4. Genetic distance and species delimitation

Pairwise uncorrected genetic distances (p-distance) were calculated between sister lineages within focal clades. Calculations were based on the ND2 gene alignment and were performed using the program MEGA-X 10.1 (Kumar et al., 2018). We used the R package ‘gggridd’ to visualize and compare the distribution of genetic distances within and between putative species pairs. To obtain a preliminary hypothesis of species boundaries, we performed a species delimitation analysis based on the ND2 alignment using the sequence-based program Automatic Barcode Gap Discovery (ABGD; Puillandre et al., 2012). The relative gap width (X) was adjusted (0.5 < X < 1.5) until species delimitation partitions stabilized.

Table 1

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Design</th>
<th>Primer Sequence: 5′–3′</th>
</tr>
</thead>
<tbody>
<tr>
<td>ND2-METF1</td>
<td>Macey et al. (1997)</td>
<td><code>AAGCTTTCGGGCCCATACGCT</code></td>
</tr>
<tr>
<td>COI-R1</td>
<td>Macey et al. (1997)</td>
<td><code>AGGTGCCCGTCTTGCCTT</code></td>
</tr>
<tr>
<td>RAG1-SQAF396</td>
<td>Skipwith et al. (2016)</td>
<td><code>TTKCTGAAATGGGAATCGCATGCCTT</code></td>
</tr>
<tr>
<td>RAG1-397</td>
<td>Groth &amp; Barrowclough (1999)</td>
<td><code>GATGTGACCCTCCTAGGCGCAGATT</code></td>
</tr>
<tr>
<td>PDC-PHOF1</td>
<td>Bauer et al. (2007)</td>
<td><code>AAGTGAGCATGAGGAGATGA</code></td>
</tr>
<tr>
<td>PDC-PHOR1</td>
<td>Bauer et al. (2007)</td>
<td><code>TCCACATCCACAGCAAAAACCTCC</code></td>
</tr>
<tr>
<td>MXRA5-F2</td>
<td>Portik et al. (2012)</td>
<td><code>KGCTGAGCGCTCCTGATGTA</code></td>
</tr>
<tr>
<td>MXRA5-R2</td>
<td>Portik et al. (2012)</td>
<td><code>YCTMGCGCCTGCACATT</code></td>
</tr>
</tbody>
</table>

PDC: 443 bp. |
2.5. Ancestral range reconstruction

We used the R package ‘BioGeoBEARS’ (Matzke, 2014) to reconstruct ancestral ranges with the BEAST2 phylogeny as the input tree due to the higher proportion of well-supported nodes. To infer colonization patterns between Borneo and the Philippines, we considered Sundaland (mainland Southeast Asia and the Indonesian islands west of Wallace’s Line) as one biogeographic unit, and considered Borneo, Palawan, Philippines, and Wallacea (Sulawesi and the Lesser Sundas) as individual biogeographic units. Although Borneo is geologically part of Sundaland, its terrestrial connection with mainland Asia and the surrounding Sundaic islands have been decreasing since the Oligocene, and intermittently severed since the Pleistocene (Hall, 2013). Additionally, considering Borneo as a separate unit enabled us to parse out biogeographic patterns specific to Borneo and the Philippines. Similarly, although Sundaland and Indochina can be considered separate biogeographic regions (de Bruyn et al., 2014), we combined the two regions together to focus on testing biogeographic patterns between Borneo and the Philippines. While Palawan is geopolitically part of the Philippines, we coded it as a distinct unit because the island has been geographically distant from the remaining Philippine landmasses for much of its geological history (hereafter, references to the Philippines exclude Palawan, and vice versa, unless otherwise stated). Palawan is a crustal microcontinental block with a Cathaysian origin, whereas the remaining Philippine islands are oceanic and part of the Philippine Mobile Belt (Dimalanta et al., 2018; Padrones et al., 2017; Yumul et al., 2009a). We tested all biogeographic models in BioGeoBEARS, with and without the jump dispersal parameter (J). The biogeographic models tested were dispersal-vicariance (DIVA); dispersal-extinction-cladogenesis (DEC); and Bayesian analysis of biogeography when the number of areas is large (BayArea). Details on model assumptions can be obtained from Matzke (2013).

3. Results

3.1. Phylogenetic estimation and systematics

Our inferred ML consensus phylogeny was discordant at a number of nodes from our preferred Bayesian topology, largely towards the base of the trees and splits involving short internodes (Fig. 2). In both phylogenies, Bornean taxa were not monophyletic and formed a number of distinct clades of varying sizes. The phylogenetic placement of Bornean taxa/clades were largely congruent and highly supported across both ML and Bayesian trees with the exception of the Cyrtodactylus malayanus complex clade (Fig. 2), which was discordant and weakly supported in both analyses. Philippine and Palawan taxa were also recovered as non-monophyletic with high support (Fig. 2).

Three morphologically unidentified and genetically distinct lineages (Cyrtodactylus sp. 1, 2, 3) were congruent in their phylogenetic placement with high support across Bayesian and ML analyses (PP = 1.0; UFBoot = 100; Fig. 2). Cyrtodactylus sp. 1 from Ketapang, Kalimantan was recovered as the sister taxon to C. rosichonarie from Great Natuna Island, Indonesia (p-distance ~ 8%); C. sp. 2 from Kayong Utara, Kalimantan was sister to C. limajalur from Serian, Sarawak (p-distance ~ 13%); whereas C. sp. 3 from southeast Sabah was sister to C. redimiculos from Palawan (p-distance 17–18%). Our phylogenetic analyses also revealed numerous instances of incompatibility for many of the Bornean endemic taxa. Taxa identified on morphological grounds as C. pubisulcus were not monophyletic and are referred to herein as C. cf. pubisulcus. The true C. pubisulcus was determined based on sequence data from the type locality, Gunung Matang (Fig. 1). One clade was most closely related to the Philippine clade containing C. philippinicus, C. agusanensis, C. mamanwa, C. gubaot, and C. summori; a second clade was sister to C. maltensis; while a third clade formed a deeply divergent sister relationship to the true C. pubisulcus (p-distance 9–10%; Fig. 2). The C. maltensis complex, which includes C. ingeri, C. matsuii, and C. yoshii, formed a well-supported clade, but the accurate identification of each constituent species remains uncertain (discussed below).

3.2. Genetic distance and species delimitation

3.2.1. Cyrtodactylus pubisulcus complex

Samples morphologically identified as Cyrtodactylus pubisulcus were polyphyletic with multiple distinct and deeply divergent clades (Fig. 3A). In western Sarawak, true C. pubisulcus from Gunung Matang (and by implication Gunung Gading and Gunung Pueh) was 9–10% divergent from the adjacent populations at Bau and Borneo Highlands (Clade 3; Fig. 3B, C). In central Sarawak, the undescribed population from Pelagus (Clade 4) is 14–16% divergent from its sister lineage, C. maltensis, and 15% divergent from true C. pubisulcus (Fig. 3 B, C). Populations from Niah, Lambir Hills, Lawas, and Gunung Mulu (Clades 1, 2; Fig. 3B) form a distinct clade that is distantly related to true C. pubisulcus. Within this clade, populations from Niah and Lambir Hills (Clade 1) are 7–9% divergent from populations at Lawas and Gunung Mulu (Clade 2; Fig. 3C). Distinct and non-overlapping distributions of p-distances within C. maltensis and Clade 2 corresponded to populations on disjunct limestone formations (Davis et al., 2019) and distinct geographic populations respectively. The ABGD analysis inferred all focal clades (Clades 1–4, true C. pubisulcus and true C. maltensis) as candidate species (Supplementary material).

3.2.2. Cyrtodactylus malayanus complex

A single sample, Cyrtodactylus sp. 2 from Kayong Utara, Kalimantan, is 12–13% divergent from its sister taxon C. limajalur from Serian, Sarawak (Fig. 4). Large variances in genetic distance were also detected within C. consobrinus (up to 10%) and C. malayanus (up to 12%); Fig. 4C). The ABGD analysis delimited C. sp. 2 as a distinct species, while splitting C. consobrinus into four species (BRK 412; LSUHC 4062; CAS 262851 + CAS 262852; FRIM 11142 + LSUHC 6586), and C. malayanus into three species (FNMH 230113; BABA 24; RMBR 1205 + ZRC 2.7201 + ZRC 2.7188 + RMBR 847; Supplementary material).

3.3. Biogeographic reconstruction

3.3.1. Philippines

Model comparison for the BioGeoBEARS analysis preferred models that included the jump dispersal parameter (Table 2). The analysis favored the BAYAREALIKE + J model (LnL = -78.89; AIC = 155.7) and inferred numerous invasions from Borneo into the Philippines and Palawan (Fig. 5). One invasion gave rise to the species Cyrtodactylus philippinicus, C. agusanensis, C. mamanwa, C. gubaot, and C. summori during the late Oligocene to Early Miocene (95% HPD = 19–30.4 mya; Fig. 5; Supplementary material). Another invasion occurred during the same time period, which gave rise to the Palawan endemic C. tautbatorum and the Philippine endemics C. jambangan, and C. annulatus (95% HPD = 20.7–33.3 mya). However, the route and order of colonization remain uncertain as the ancestral range for the three species was reconstructed as either Palawan or the Philippines (Fig. 5). Palawan experienced a separate colonization event from Borneo during the Miocene (95% HPD = 10.8–22.9 mya) involving the ancestor of C. redimiculos and C. sp. 3 (Fig. 5).

Our analyses provided reconstructions for other Southeast Asian regions, but due to inadequate taxonomic coverage and discordant topologies pertaining to those clades, we refrain from drawing definitive conclusions. To infer more robust biogeographic patterns between Borneo and Sundaland comprehensive analyses beyond the scope of our study need to be conducted, thus we recommend caution in interpreting the reconstructed patterns (Fig. 5).
4. Discussion

4.1. Phylogenetic relationships

Our ML and Bayesian analyses are concordant across most branches but both trees lack support at deeper nodes, indicating a lack of phylogenetic information in our sampled loci, insufficient taxonomic sampling, and/or complex evolutionary processes such as incomplete lineage sorting (Alda et al., 2019; Alexander et al., 2017; Debiasse et al., 2014; Esselstyn et al., 2017; Knowles et al., 2018). The majority of discordant and poorly-supported nodes involve short internodes, alluding to the possibility of rapid diversification events driving incomplete lineage sorting (Whitlock and Lockhart, 2007). The sampling of additional loci and taxa will likely be necessary to provide a more robust and well-resolved phylogenetic estimate for future biogeographic and comparative analyses.

The discordant placement of the Cyrtodactylus malayanus complex clade with respect to the Indochina and Indonesia + Malaysia clades (Fig. 2) is the most notable difference between the ML and Bayesian topologies. With regards to the relationships among these clades, the Bayesian phylogeny is congruent with the ML phylogeny inferred by Wood et al. (2012), which included fewer Bornean taxa but more Indochinese and Australian taxa. However, due to the topological discordance and weak nodal support across multiple studies (and even between optimality criteria within this study), we consider it premature to assume high confidence in our preferred topology. Another distinction between our ML and Bayesian topologies is the relationship between the two Philippine endemics, C. philippinicus and C. agusanensis, despite strong support for both inference methods. Our Bayesian tree inferred C. philippinicus as the sister lineage to the subclade containing C. agusanensis, C. mamanwa, C. gubaot, and C. sumuroi, similar to the results from Welton et al. (2010b), whereas our ML analysis recovers C. agusanensis as the sister lineage, which agrees with the concatenated phylogeny of Siler et al. (2010). The phylogeny of Wood et al. (2012) inferred yet another relationship with C. philippinicus and C. agusanensis as sister taxa, although this topology was poorly supported and likely due to reduced taxon sampling in the subclade (C. mamanwa, C. gubaot, and C. sumuroi excluded). These uncertainties in the phylogenetic relationships may be due to mito-nuclear discordance, deep coalescence (Siler et al., 2010), hybridization, or reduced taxon sampling—all of which underscore the need for expanded research efforts focused on Philippine and Bornean Cyrtodactylus.

4.2. Cryptic diversity and phylogenetic structure

Our results revealed numerous instances of taxonomic ambiguity, potential undescribed lineages, and high levels of genetic divergence among populations that have traditionally been regarded as conspecific. This indicates that the relatively depauperate species diversity in Borneo may, in part, stem from the presence of unrecognized cryptic species. Our data suggest that many of the recognized Bornean species, long considered geographically widespread (e.g. Cyrtodactylus pubisulcus, C. consobrinus, C. malayanus), may comprise genetically distinct and locally endemic lineages. Among the ten recognized Bornean species included in our study, the ABGD analysis estimated as many as 19 putative species. Although the ABGD estimate should be viewed as a...
Fig. 3. A) Subset of the Bayesian tree depicting relationships of the Cyrtodactylus pubisulcus and C. baluensis complexes. Black circles on the nodes signify strong support (posterior probability $\geq 0.95$). B) Distribution of sampled localities with type localities represented by stars. Localities for the C. pubisulcus complex samples are coded by colors and represented by squares, whereas localities for the C. baluensis complex samples are coded in yellow and represented by triangle, circle, and diamond shapes. C) Distributions of pairwise uncorrected p-distances (ND2 gene) within populations and between closely-related putative species pairs. Color gradient corresponds to genetic distance.

Fig. 4. A) Subset of the Bayesian tree depicting relationships of the Cyrtodactylus consobrinus and C. malayanus complexes. Black circles on the nodes signify strong support (posterior probability $\geq 0.95$). Relevant clades are highlighted with colors that correspond to the inset distribution map. Asterisks (*) denote GenBank sequences that are likely misidentified. B) Distribution of sampled localities with type localities represented by stars. Localities are colored to correspond with the inset phylogeny. C) Distributions of pairwise uncorrected p-distances (ND2 gene) within populations and between closely-related putative species pairs. Color gradient corresponds to genetic distance.
preliminary assessment, it implies that the number of recognized species from Borneo is underestimated and suggests that some Bornean lineages may be speciating without accompanying phenotypic divergence (Barley et al., 2013, in press). Below, we provide detailed discussions of the systematics for each species complex.

4.2.1. Cyrtodactylus pubisulcus complex

Some of the most notable genetic variation was observed within Cyrtodactylus pubisulcus, which morphological analyses indicated to be a single, widespread lineage spanning from western Sarawak to southeast Sabah (Ahmad et al., 2019; Das, 2005; Hikida, 1990). The polyphyly of C. pubisulcus suggests that rather than being one wide-ranging species, C. pubisulcus sensu stricto (s.s.) is restricted to western Sarawak, at the type locality of Gunung Matang and the surrounding areas including Gunung Gading and Gunung Pueh (Fig. 3B).

Table 2
Model testing for the BioGeoBEARS analysis with and without found-event speciation (+ J). Models tested: dispersal-extinction-cladogenesis (DEC); Bayesian analysis of biogeography when the number of areas is large (BayArea); and dispersal-vicariance (DIVA).

<table>
<thead>
<tr>
<th>Null</th>
<th>Alt</th>
<th>LnL Null</th>
<th>LnL Alt</th>
<th>D-statistic</th>
<th>p-value</th>
<th>AIC (Null)</th>
<th>AIC (Alt)</th>
<th>AICc (Null)</th>
<th>AICc (Alt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEC</td>
<td>DEC + J</td>
<td>−89.57</td>
<td>−86</td>
<td>7.04</td>
<td>0.008</td>
<td>183.1</td>
<td>178.1</td>
<td>0.075</td>
<td>0.93</td>
</tr>
<tr>
<td>DIVALIKE</td>
<td>DIVALIKE + J</td>
<td>−96.1</td>
<td>−93.73</td>
<td>4.72</td>
<td>0.03</td>
<td>196.2</td>
<td>193.5</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>BAYAREALIKE</td>
<td>BAYAREALIKE + J</td>
<td>−83.89</td>
<td>−74.87</td>
<td>18.03</td>
<td>0</td>
<td>171.8</td>
<td>155.7</td>
<td>0.0003</td>
<td>1</td>
</tr>
</tbody>
</table>

Fig. 5. Time-calibrated biogeographic reconstruction of the best-fitting model from the BioGeoBEARS analysis. Pie charts represent relative probabilities instantaneously before (at nodes) and after (at corners) cladogenesis.
distinct and highly divergent clades compared to true *C. pubisulcus* + Clade 3 (16–21% divergent).

Comparisons of intra- and interpopulation genetic divergences show distinct and non-overlapping breaks, indicating that the genetic variation among the putative species do not correspond to intrapopulation level variation (Fig. 3C). Therefore, the phylogeographic structure and genetic divergences within the group support the distinction of Clade 4 and Clades 1 + 2 as separate species from *C. pubisulcus* s.s. Sister lineages within both the eastern and western clades exhibit comparatively lower genetic divergences (9–10%), with their respective geographic proximities and phylogenetic structures alluding to more recent and incipient diversification (Fig. 3C). Denser sampling in intervening areas among all populations will be necessary to determine whether the divergences among the eastern and western clades warrant specific recognition.

Despite the high levels of phylogeographic structure and genetic divergences within the *C. pubisulcus* complex, analyses of morphological data have not yet identified any potentially diagnostic characters which could be used to distinguish putative species (unpublished data). The absence of readily diagnosable character state differences between divergent, non-monophyletic groups suggests that this complex may be a promising study system to test hypotheses of convergent evolution versus cryptic speciation (Barley et al., 2013; Welton et al., 2017).

### 4.2.2. *Cyrtodactylus malayanus* complex

*Cyrtodactylus consobrinus* and *C. malayanus* demonstrate clear phylogeographic structure and high intraspecific genetic divergence (Fig. 4). The type locality of *C. consobrinus* is Gunung Matang (Fig. 1), but the species has been reported throughout Sarawak, Sabah (Das, 2005), Peninsular Malaysia (Grismer and Quah, 2019), and Singkеп and Sumatra Islands, Indonesia (De Rooij, 1915). Populations from Peninsular Malaysia (LSUC 6586; FRIM 1142) are 7–8% divergent from Bornean populations (Fig. 4C), whereas populations within Borneo are up to 11% divergent from one another. Similar levels of divergence are shown within *C. malayanus* (12%). However, the imprecise type locality for *C. malayanus* (“Indo-Australian Archipelago”; De Rooij, 1915) poses a problem for determining the population that represents true *C. malayanus*, especially considering the phenotypic similarities among the divergent lineages [specimens examined: ZRC 2.7188; ZRC 2.7201; ZRC 2.7198; ZRC 2.7200]. Therefore, a precise type locality needs to be determined, or designated, before taxonomic revisions for *C. malayanus* can commence.

Although our ABGD analysis and presence of relatively high genetic divergences support the recognition of multiple distinct species within the *C. consobrinus* and *C. malayanus* groups, we caution against recognizing species based on their phylogenetic structure and/or mitochondrial genetic distances, as they may reflect a sparsity of sampling. As such, we posit that with the widespread distributional ranges of both species, broader sampling of additional intervening populations may reveal subtler phylogenetic structure associated with population-level differentiation, as was shown in the *C. quadriangulatus* group (Johnson et al., 2012). This judicious approach is necessary to avoid describing populations as species due to the misinterpretation of high levels of genetic divergences that result from inadequate sampling and/or isolation by distance (Irwin et al., 2005; Wright, 1943).

### 4.2.3. *Cyrtodactylus baluensis* complex

The resolution of the *Cyrtodactylus baluensis* complex is challenging due to the lack of topotypic genetic material, overlapping distributions, Interspecific morphological similarity, and relatively shallow genetic divergences among the constituent species (*C. baluensis*, *C. ingeri*, *C. matsuii*, and *C. yoshii*). Moreover, *C. baluensis* and *C. matsuii* share the same type locality (Mount Kinabalu), whereas the type locality of *C. ingeri* is Poring Hot Springs, a mere 17 km from Mount Kinabalu (Fig. 3B). Based on available data, we are unable to identify the true *C. baluensis* with high confidence because voucher specimens associated with the three genetic sequences of *C. baluensis* from GenBank (SPM 6906, SPM 6891, SPM 6892) were not available for examination. One of those samples, SPM 6906 was highly divergent (12%) from the other two (SPM 6891, SPM 6892), which are instead more closely related (2.3–3.5% divergent) to a sequence assigned to the taxon *C. yoshii* from adjacent Poring Hot Springs (ZRC 2.4851). These data suggest that there are two distinct and highly divergent lineages that occur in the Kinabalu-Poring region. Our morphological examination of the *C. yoshii* specimen (ZRC 2.4851) closely matched the original description of the species (Hikida, 1990) and the type material. However, the identity of the specimen remains to be verified due to the lack of genetic data from the type locality at Sepilok, along the eastern coast of Sabah. Thus, pending confirmation from topotypic genetic data, available information suggests that sample SPM 6906 likely represents true *C. baluensis*, which is highly divergent and sympatric with *C. yoshii*. This also implies that the GenBank samples SPM 6891–92, identified as *C. baluensis*, are likely *C. yoshii* and that its distribution extends to western Sabah.

The accurate identification of *C. matsuii* and *C. ingeri* also remains dubious. The voucher specimen associated with the genetic sequence identified as *C. matsuii* (ID7167) collected from Crocker Range Park, Sabah was not available for examination. We examined three specimens of *C. cf. ingeri* (BRK 699, BRK 700, BRK 702) and found that they matched the original description and type material of *C. ingeri*. However, those samples were from Merapar, Gunung Murud, in northern Sarawak, implying that the distribution of *C. ingeri* could be wider than previously recorded (Fig. 3B). Unfortunately, poorly supported relationships and the absence of positively identified genetic material for *C. ingeri* precludes us from making confident taxonomic assessments for the *C. baluensis* complex.

### 4.3. Biogeography

The best-fit biogeographic models included the jump dispersal parameter (*j*), which considers founder-event speciation (Matzke, 2014). The use of this parameter, especially in conjunction with the DEC model, has been questioned because neither the model nor the parameter takes time into account, thus promoting cladogenetic over anagenetic events (Rie and Sanmartin, 2018). However, we found it prudent to consider models including the *j* parameter because dispersal over permanent marine barriers is a prevailing biological reality within the context of an archipelago model such as the Philippines (Brown, 2016; Brown et al., 2013; Brown and Gutman, 2002; Brown and Alcala, 1970; Evans et al., 2003; Inger, 1954; Linkem et al., 2013). Speciation due to founder effects is expected to have occurred in the Philippines due to the low probability of shared ancestral ranges across volcanic islands that arose *de novo*, with separation for most, if not all, of the region’s geological history (Brown et al., 2016, 2013; Brown and Alcala, 1970; Diamond and Gilpin, 1983; Hall, 2013; Inger, 1954; Tougaard, 2001). This was corroborated by the fact that models excluding the *j* parameter inferred less plausible biogeographic scenarios, which included shared ancestral ranges encompassing all of Sundaland and the Philippines (Supplementary material). Although most of Sundaland was land-locked for extended periods of time (Hall, 2013; Sarr et al., 2019), the Philippine islands have remained mostly separated from the Sunda Shelf by the Sulu and Western Philippine Seas with intermittent, temporary terrestrial connections (see references above). Therefore, the preferred models estimated from our biogeographic analyses, which inferred jump dispersals from Borneo into the Philippines, are more plausible than vicariance (as inferred by models without the *j* parameter; Supplementary material). Additionally, overseas dispersals are consistent with what is known about the geological history of the region (Hall, 2013; Yumul et al., 2009a, 2009b, 2003) and the biology of terrestrial vertebrates in oceanic archipelagos (Barley et al., 2013; Brown et al., 2016; Brown and Diesmos, 2009; Carlquist, 1966; de Queiroz, 2005; Queiroz, 2014). However, as researchers begin to build more robust phylogenies with expanded genetic and geographic
Sampling, we understand that new biogeographic hypotheses may arise.

4.3.1. Multiple independent invasions into the Philippines

Previous phylogenetic studies have included Bornean and Philippine Cyrtodactylus species, but the relationships among taxa from each region have been difficult to resolve, presumably due to an abundance of missing taxa (Brennan et al., 2017; Siler et al., 2010; Welton et al., 2010a, 2010b; Wood et al., 2012). Consequently, downstream biogeographic analyses predicated on an incomplete evolutionary framework could produce misleading results. Contrary to a single invasion hypothesis (Siler et al., 2010), we found strong support for a model with multiple independent invasions of the Philippines and Palawan, potentially providing support for the dual-invasion hypothesis (Brown and Guttman, 2002; Brown and Siler, 2014; Diamond and Gilpin, 1983; Inger, 1954; Jones and Kennedy, 2008).

Siler et al. (2010) found mito-nuclear discordance in the placement of the Palawan endemic, C. redimiculus. In that study, C. redimiculus was either sister to C. baluensis or the Philippine C. agusanensis clade depending on mitochondrial versus nuclear data, suggesting that the discordance could be due to incomplete lineage sorting of the nuclear genome (Debiasse et al., 2014). However, we attribute the inconsistent phylogenetic position of C. redimiculus to poor taxonomic sampling from Borneo. The inclusion of C. sp. 3 places C. redimiculus as sister to the unidentified C. sp. 3 species from Sabah, thus forming a clade distinct from the C. annulatus clade (including the Palawan endemic, C. taubatorium). This relationship, together with our biogeographic model (Fig. 5), indicates that the ancestor to C. redimiculus independently dispersed from Borneo into Palawan during the early Miocene. This coincides with the end of the subduction of the proto-South China sea, which significantly increased the land area of Borneo and Palawan (Hall, 2013).

One invasion into the Philippines occurred either directly from Borneo or through Palawan, eventually giving rise to the Philippine endemics C. jambangan and C. annulatus. Results from the biogeographic analyses inferred both dispersal routes with equal likelihood. The subduction of the Celebes Sea beneath south Sabah and the Sulu Arc could have produced ephemeral connections between Borneo and the Philippines via intermittently emergent volcanic islands during the Middle and Late Miocene (Hall, 2013), thereby facilitating the movement of Bornean lineages into the Philippines. On the other hand, the hypothesized break up of Zamoanga from the Palawan Continental Block and subsequent tectonic drift into its current position (Padrones et al., 2017; Yumul et al., 2004) indicates another potential scenario by which a Bornean ancestor colonized the Palawan microcontinent prior to the break off of Zamoanga. Upon subsequent contact with Mindanao during the Middle Miocene (Yumul et al., 2004), it is conceivable that this lineage could have expanded its range, colonized the remaining Mindanao Pleistocene Aggregate Island Complex (Brown and Diesmos, 2009), and dispersed overseas to a few islands to the north (Welton et al., 2010a). Our inferred timing of the diversification of Palawan (Cyrtodactylus taubatorium) and Philippine taxa (C. jambangan, C. annulatus) during the Middle Miocene (Fig. 5) fits both biogeographical scenarios.

Our analysis also shows another Philippine colonization event and possible radiation along the Sulu Archipelago, following the Eastern Island Arc dispersal route (Brown and Guttman, 2002; Brown and Siler, 2014; Diamond and Gilpin, 1983; Evans et al., 2003), which gave rise to the clade comprising Cyrtodactylus philippinicus, C. agusanensis, C. mawanwa, C. gubao, and C. sumoroi. Colonization by terrestrial vertebrates along the Eastern Island Arc have been shown with strong support (Brown et al., 2013); but we suspect that the long branch lengths in our topology between C. cf. pubissilcus and the clade containing C. philippinicus indicate that we may be missing critical taxa that could support a different biogeographic model. However, if our current model is supported with expanded taxonomic and genetic sampling, it would add additional support to the dual-invasion hypothesis, initially inferred in Philippine Hylarana (Brown and Guttman, 2002; Brown and Siler, 2014; Diamond and Gilpin, 1983) and later expanded to numerous terrestrial vertebrates (Brown et al., 2013). Expanded biodiversity studies on Borneo to further test our biogeographic hypotheses, especially regarding the island’s role in the evolutionary history of Southeast Asian fauna, is a promising and necessary area for future research.

CRediT authorship contribution statement

Hayden R. Davis: Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing, Visualization, Funding acquisition. Kin Onn Chan: Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing, Visualization. Indraeuil Das: Conceptualization, Resources, Writing - review & editing, Project administration. Ian G. Brennan: Resources, Writing - review & editing. Benjamin R. Karin: Resources, Writing - review & editing. Todd R. Jackman: Resources, Writing - review & editing. Rafe M. Brown: Writing - review & editing. Djoko T. Iskandar: Resources. Iznazil Nasir: Resources. L. Lee Grismer: Resources, Writing - review & editing. Aaron M. Bauer: Conceptualization, Resources, Writing - review & editing, Funding acquisition, Supervision.

Acknowledgements

We are grateful to the editor and anonymous reviewers for their comments, which greatly improved our manuscript. We thank the Sarawak Forestry Department, Malaysia for providing collections permits [NPW.907.4.4.(Jld.14)-79; (119)HJS/NCCD/600–7/2/107]. HRD and AMB were in part funded by the Gerald M. Lemole Endowed Chair funds and Villanova University, United States and HRD received additional funding from The Society for Integrative and Comparative Biology, United States (Fellowship of Graduate Student Travel Award), the Museum of Comparative Zoology, Harvard, United States (Ernst Mayr Grant), and the Lee Kong Chian Natural History Museum, Singapore (Collection Study Grant for Students). ID and IN were supported by a Niche Research Grant Scheme from the Ministry of Higher Education, Government of Malaysia (NRGS/1087/2013(01)). Fieldwork to collect Cyrtodactylus genetic material in the Philippines was funded in part by grants from the U.S. National Science Foundation to RMB and UK-based graduate students (KOC, C. D. Siler, and C.W. Linkem): DEB0743491, 0804115, 1418895, 0910341, 0640737, 0344430, 0073199, 1654388, 1702036; and EF-0334952). We thank Adam D. Leaché and his lab, and Simone Des Roches for their comments and suggestions on multiple versions of the manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2020.106785.

References
