ZooKeys 757: 1–67 (2018)
doi: 10.3897/zookeys.757.10965
http://zookeys.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
ZooKeys
Launched to accelerate biodiversity research

What is a species? A new universal method to measure differentiation and assess the taxonomic rank of allopatric populations, using continuous variables

Thomas M. Donegan[1]

1 *Unaffiliated, London, UK*

Corresponding author: *Thomas M. Donegan* (thomasdonegan@yahoo.co.uk)

Academic editor: *Y. Mutafchiev* | Received 28 October 2016 | Accepted 3 April 2018 | Published 10 May 2018

http://zoobank.org/BC105D91-2B58-418E-ADD8-F13E2B52429F

Citation: Donegan TM (2018) What is a species? A new universal method to measure differentiation and assess the taxonomic rank of allopatric populations, using continuous variables. ZooKeys 757: 1–67. https://doi.org/10.3897/zookeys.757.10965

Abstract

Existing models for assigning species, subspecies, or no taxonomic rank to populations which are geographically separated from one another were analyzed. This was done by subjecting over 3,000 pairwise comparisons of vocal or biometric data based on birds to a variety of statistical tests that have been proposed as measures of differentiation. One current model which aims to test diagnosability (Isler et al. 1998) is highly conservative, applying a hard cut-off, which excludes from consideration differentiation below diagnosis. It also includes non-overlap as a requirement, a measure which penalizes increases to sample size. The "species scoring" model of Tobias et al. (2010) involves less drastic cut-offs, but unlike Isler et al. (1998), does not control adequately for sample size and attributes scores in many cases to differentiation which is not statistically significant. Four different models of assessing effect sizes were analyzed: using both pooled and unpooled standard deviations and controlling for sample size using *t*-distributions or omitting to do so. Pooled standard deviations produced more conservative effect sizes when uncontrolled for sample size but less conservative effect sizes when so controlled. Pooled models require assumptions to be made that are typically elusive or unsupported for taxonomic studies. Modifications to improving these frameworks are proposed, including: (i) introducing statistical significance as a gateway to attributing any weighting to findings of differentiation; (ii) abandoning non-overlap as a test; (iii) recalibrating Tobias et al. (2010) scores based on effect sizes controlled for sample size using *t*-distributions. A new universal method is proposed for measuring differentiation in taxonomy using continuous variables and a formula is proposed for ranking allopatric populations. This is based first on calculating effect sizes using unpooled standard deviations, controlled for sample size using *t*-distributions, for a series of

different variables. All non-significant results are excluded by scoring them as zero. Distance between any two populations is calculated using Euclidian summation of non-zeroed effect size scores. If the score of an allopatric pair exceeds that of a related sympatric pair, then the allopatric population can be ranked as species and, if not, then at most subspecies rank should be assigned. A spreadsheet has been programmed and is being made available which allows this and other tests of differentiation and rank studied in this paper to be rapidly analyzed.

Keywords
diagnosis, species limits, species scoring, statistics, subspecies limits, taxonomy

Introduction

This paper aims to help address the "allopatric problem" when determining species rank in taxonomic science. Humans have categorized populations into named groups since the dawn of known civilization (Aristotle c. 350 B.C.) and these were first referred to as "species" over 300 years ago (Willughby 1676, 1678). As defined by Ray (1686): "no matter what variations occur in the individuals or the species, if they spring from the seed of one and the same plant, they are accidental variations and not such as to distinguish a species... Animals likewise that differ specifically preserve their distinct species permanently; one species never springs from the seed of another nor vice versa."

Sympatric species, which occur together in the same place during the breeding season but do not successfully interbreed to any material extent, are demonstrably real. With enough data and persistence, it is usually possible to determine whether or not sympatric populations interbreed regularly and whether they produce fertile offspring (Mayr 1940) and therefore whether or not the two populations are reproductively isolated. Where hybridization is rare or occurs in narrow zones, this can cause difficulties in delimiting species and may need judgment to be applied.

A traditionally more difficult problem, and the focus of this paper, is that of "allopatric" (Mayr 1942) populations (referred to as "asympatric" by Poulton 1904, 1908, who originally identified this problem), i.e., those which do not occur together in the same geographical place during the breeding season. Allopatric populations can be recognized either as subspecies of polytypic species or as monotypic species under Mayr (1940, 1942)'s scheme. However, allopatric populations should only be ranked as species where they are as distinctive as sympatric species (Helbig et al. 2002). This is not an artificial test. Over a period of time, two disjunct populations facing different selection pressures may differentiate from one another, and at some point, they will attain sufficient differentiation that this can be observed to attain or exceed that shown between sympatric species. At such a point, but not otherwise, it is reasonable to assume that they have speciated.

The subjectivity involved in comparing allopatric species and the rise of molecular science have doubtless encouraged the development of a multitude of different species criteria or concepts. As noted by De Queiroz (1998, 1999), many of these are

simply different ways of finding out what a species is, as opposed to being based on different ideas of what species are. However, proponents of these concepts challenge the "comparative approach" to assessing the rank of allopatric populations (e.g., Halley et al. 2017). Under phylogenetic and related species concepts (PSC), diagnosability and monophyly ("clusters of individuals with a pattern of ancestry and descent": Cracraft 1983) are the hallmarks of species rank. Such "clusters" can be ascertained using molecular biology, a discipline that does not need to be informed by real-world differentiation in morphology, animal sounds, or biometrics. Because all diagnosable units under this model are called species, some PSC proponents have argued for the subspecies rank to be abandoned (Zink 2003). However, whilst molecular research has revolutionized higher-level taxonomy, it is less useful at addressing questions of species rank, since sympatric species show variable intraspecific DNA differentiation, ranging from between 0% to at least 8% (Sorenson et al. 2003, Marks et al. 2002). Many modern ornithological taxonomists seek to take into account the results of both molecular and traditional analyses where possible in assessing rank. Biological species concepts, often integrating "lineage"-based concept thinking (De Queiroz 1998, 1999) remain in prevailing usage among leading checklist committees (e.g., AOU 1998, Helbig et al. 2002, Remsen et al. 2018) and in taxonomic reference works (e.g., Dickinson and Christidis 2014), albeit often informed by molecular data and diagnosability (Sangster 2014).

Whilst statistical and mathematical techniques to analyze molecular data have been a rich field for methodological advancement, the same cannot be said for the study of real world variables. Supportable statistical schemes for assessing between-population differentiation are noteworthy principally by their absence. Those schemes which have been proposed are either widely criticized, only applicable to particular taxonomic groups or vague.

Helbig et al. (2002) developed a set of guidelines for taxonomic committees to assess species and subspecies rank, in the context of de Queiroz (1998, 1999)'s lineage concept. In relation to allopatric populations, these authors recommended that: "The likelihood that allopatric taxa will remain distinct can only be judged by the degree of their divergence, preferably in comparison with taxa that are closely related to the group under investigation and that are known to coexist in sympatry". They recommended that, in order to be ranked as species, allopatric populations should usually be diagnosable by several discrete or continuously varying characters related to different functional contexts, e.g., structural features (often related to foraging strategy), plumage colors, vocalizations (both often related to mate recognition) or DNA sequences, and the sum of the character differences should correspond to or exceed the level of divergence seen in related species that coexist in sympatry.

This paper will concentrate on the traditional currency of taxonomy: continuous variables such as those based on measurement of specimens, whether in the museum or in the field. Many researchers and advanced amateurs do not have a molecular laboratory available and few genera have been exhaustively sampled in a way that includes multiple individuals at population level. In contrast, vocal and biometric data

are easy to collate, accessible to many and cheaper to analyze. A wide variety of other 'real world' organism characters are capable of measurement as continuous variables. For vocalizations, lengths or acoustic frequencies of notes can be measured using sonograms, for example. Coloration can be measured using spectrometry. Non-continuous or discrete variables, e.g., presence or absence of a particular character and molecular markers, can be analyzed best using cladistics and other phylogenetic tools and are not covered here in detail.

Hubbs and Perlmutter (1942) proposed that, in order to assess diagnosability using continuous variables, taxonomists should calculate the distance between the means of the two populations for a particular character and measure that distance in terms of standard deviations (SDs), a measure referred to in statistics as "effect size". Where the means of two populations differ by four average SDs, then under a normal distribution with infinite sample size, there is no overlap between data to 95% confidence and the populations can be considered "diagnosable" for the character in question. As noted by McKitrick and Zink (1988) and Remsen (2010), aiming for 100% diagnosability is conceptually and methodologically unreasonable. 95% is the standard confidence internal in science, the benchmark for assessing diagnosability using discrete characters (Wiens and Servedio 2000, Walsh 2000) and the benchmark for testing diagnosis using continuous variables (Isler et al. 1998). Hubbs and Perlmutter (1942) further proposed a "50% diagnosis" test that might be used for assessing subspecies rank, where populations differ by two SDs: effectively denoting differentiation of a character half-way towards diagnosability. Later, a 75%/99+% diagnosis test for subspecies (e.g., Amadon 1949, Patten and Unitt 2002) was developed and became more widely used. It has more recently been proposed that full (95% statistical) diagnosability in a single character should be the benchmark for subspecies, which is synonymous with a PSC species definition (Remsen 2010).

Isler et al. (1998) modified Hubbs and Perlmutter (1942)'s tests by taking into account sample sizes using student *t*-distribution values rather than bare SDs, to measure the difference between population means (detailed below under Methods: Level 5). This resulted in a model for measuring differentiation and assessing species rank that effectively requires an elevated distance between means of two populations, with greater distances for data using smaller sample sizes. Based on studies of closely related sympatric birds in a particular bird family, the antbirds (Thamnophilidae), Isler et al. (1998) concluded that three diagnostic vocal differences between songs or calls was typical of the differentiation observed between sympatric but related species. As a result, the benchmark of three diagnosable differences was considered a good "point of reference" for assigning species rank to allopatric populations in the same family. Diagnostically distinct populations not meeting this standard are ranked as subspecies under this model (Remsen 2010). Donegan and Avendaño (2008) applied this method to the tapaculos (Rhinocryptidae) and found examples of sympatric species that differed by only one, not three, diagnosably distinct vocal characters. This suggested that vocal benchmarks cannot be applied universally to all birds, even those in quite closely related families.

When species rank is assessed across a taxonomic group as a whole, consistency is a virtue. Under a biological species concept-based approach, attaining such consistency will require a determination of which allopatric populations have differentiated to the same extent as related sympatrics and which have not. Those that have so differentiated are species; those that have not are, at most, subspecies. Unfortunately, consistency is not attained in current classifications, especially as regards more diverse tropical faunas. This is generally due to discrepancies in available data, the regularity of different genera being revised and differences in approaches by regional committees or textbook authorities to studies using different taxonomic methods (e.g., molecular vs. morphological) (Sangster 2014, Donegan et al. 2015, Collar et al. 2016). Even in a popular group such as birds, in the tropics there are many more species and subspecies than there are taxonomists, meaning that only a small number of groups have been subject to modern studies. However, inconsistencies and stasis are compounded by biases of some taxonomic committees towards keeping "status quo" treatments of previous authorities, ahead of reflecting the results of modern reviews in certain publications (e.g., the field guide literature or less-prestigious journals) (Donegan et al. 2015). Large numbers of allopatric populations inhabiting different mountain ranges, lowland regions or islands lack modern studies to assess their rank, or studies may exist which have been ignored, and taxonomies as a whole are often based on tradition more than rationality.

Helbig et al. (2002)'s scheme for the comparative assessment of sympatric species has been applied by some taxonomic committees in Europe as the basis for splitting of a number of questionably valid species (e.g., Carrion Crow *Corvus corone* from Hooded Crow *Corvus cornix*: Parkin et al. 2003; American Herring Gull *Larus smithsonianus* from European Herring Gull *Larus argentatus*: Collinson et al. 2008). The former two crows are well-known to hybridize and establish relatively narrow contact zones where intermediate plumages prevail. Their split relies in part on a marginal bias towards non-crossing mate choice in such zones (Parkin et al. 2003). The latter two gulls have been considered diagnosable in immature plumages and mtDNA, but they have yet been found to be fully diagnosable in any adult plumage character and infrequent hybridization between allopatric related species obscures any interpretation of molecular results (Lonergan and Mullarney 2004, Sonsthagen et al. 2016), whilst voice has not yet been subject to detailed statistical analyses demonstrating diagnosability.

Neither of these two splits is problematic from a phylogenetic species concept or "enthusiastic splitter" perspective in isolation; and further studies could give stronger support to these treatments. However, based on my experience of working with birds in the Neotropics, the benchmark applied to these situations would result in the specific recognition of probably several thousands of current subspecies or unnamed taxa occurring in that region. Barrowclough et al. (2016) estimated that the number of recognized bird species globally would almost double, were phylogenetic species concepts to be applied. That factor would increase further under models that treat populations with non-diagnosable adults, such as the Herring Gulls referred to above, as species. Discrepancies arise because, at the same time as Europe's leading taxonomic committees embarked on a program of enthusiastic splitting, countless diagnosable allopatric

populations in the tropics that exhibit more considerable vocal or morphological differentiation (some of which have been shown by molecular studies not to be sister taxa or which barely resemble one another in voice or morphology) remain lumped by the more conservative taxonomic authorities addressing those regions. The current status of global bird taxonomies is, therefore, highly irrational and subject to regional bias.

Tobias et al. (2010) highlighted the internal inconsistency of avian taxonomies on a global scale and the lack of a universal framework for species delimitation. They proposed a universal "species scoring" test for assessing the taxonomic rank of birds. This takes into account not just vocal characters (as is broadly the case under the Isler et al. 1998 model) but also plumage, biometrics, sympatry/parapatry, hybridization, habitat, and ecology. Their system is based upon a series of scores of 0–4 for a maximum number of characters in particular categories. Differences are classified as minor (1) medium (2), major (3) or exceptional (4). For plumage, various guidelines were proposed for a judgement-based assessment. For continuous variables, Tobias et al. (2010) measured pooled effect sizes without controlling for sample sizes using *t*-distributions. In their system, populations showing 0.2–2 effect size difference (minor to below 50% diagnosability) score 1 point, 2–5 effect sizes (equivalent to 50% to >95+% diagnosability depending on sample size) score 2 points, those at 5–10 effect sizes score 3 and >10 score 4. This system was developed based on a study of 58 pairs of closely related sympatric species from 29 families. Del Hoyo and Collar (2014, 2016) applied the Tobias et al. (2010) system to all birds in a major book series, proposing over 400 splits and 20 lumps in the first edition alone.

The Tobias et al. (2010) method and outcomes of Del Hoyo and Collar (2014, 2016)'s new taxonomy have been criticized, on conceptual and organizational grounds (Remsen 2015, 2016, Bakker 2015, Garnett and Christidis 2017). Many of Del Hoyo and Collar (2014, 2016)'s South American splits were however supported by a critical review, although not in Toucans, a group that shows extraordinary intra-specific variation where species scoring produced unsupportable outcomes (Donegan et al. 2015). Although there have been calls for proposed new taxonomies in the work to be rejected (Remsen 2015) or restricted to situations where significant data gaps exist (Remsen 2016), some authors have reviewed the proposals and accepted or rejected them on a case-by-case basis (e.g., Donegan et al. 2015, Gill and Donsker 2018).

Garnett and Christidis (2017) criticized the "anarchy" in current taxonomy, citing the large number of splits by Del Hoyo and Collar (2014, 2016) and calling for the regulation by committee of splitting and lumping in taxonomy and moves to "restrict the freedom of taxonomic action". This proposal has itself been widely criticized (e.g., Thomson et al. 2018, Collar 2018), some authors commenting that it "conflict[s] with some basic and indisputable principles underpinning the philosophy of science" (Raposo et al. 2017). There appears to be broad disagreement as to whether existing taxonomies are either (i) well-developed, only to be changed following review of the scientific literature by appropriately appointed persons; or (ii) irrational and in need of expeditious root-branch review. Those in both camps have claimed that the needs of conservation support their approach (Garnett and Christidis 2017, Collar et al. 2016). I have argued elsewhere that we are "fiddling while Rome burns, if being closed-mind-

ed to new findings that may challenge preconceptions or requiring perfect data sets for change", in this era of extinctions (Donegan et al. 2015). Regardless of who has the best ideas about the politics of how taxonomy is organized, it can be said that *all* these modern controversies have a single underlying cause, namely the "allopatric problem" of species: how assessments are to be made, whether it matters that this is considered consistently, how urgent any reassessment is, what the right benchmark is and which persons or bodies are properly qualified to make the decisions.

In light of the difficulties with scoring "systems" and other developments, Halley et al. (2017) have argued for a return to monophyly and essentially Cracraft (1983)'s scheme as the basis for determining species rank for allopatrics. They cite the lack of a broadly supported universal benchmark test, the difficulty of finding sympatric sister groups for study and inconsistencies in existing taxonomies but also did not regard it as a problem that recognized allopatric versus sympatric species might show different levels of differentiation. Under such an approach, many named and unnamed subspecies occurring on different mountain ranges and islands in the tropics would be afforded species rank. Difficulties as to the appropriate setting of a benchmark in difficult cases are transferred from the "equivalent to a species" benchmark to a different point which distinguishes other borderline situations: i.e., claimed barely monophyletic versus claimed non-monophyletic groupings. Gill (2014) separately proposed that a null hypothesis of species rank should apply to some allopatric populations, but this proposal was criticized by Toews (2015). Such methods and approaches are not considered further here since, in the words of Halley et al. (2017), I am "philosophically tied to a yardstick approach".

Over the last 20 years, I have been studying the taxonomy of birds in Colombia using biometric data (from mist-netting and museums) and using sound recordings. This resulted in the production of a large amount of data relevant to studying differentiation. It has become transparent to me that steps might be taken towards resolving some of these seemingly intractable fundamental disagreements, by developing an objective and agreeable basis, grounded in scientific method, statistics, the analysis of large data and based on traditional biological species concept thinking, that could be used better, more consistently and more rationally to assess the rank of allopatric populations. Ultimately, the aim of this study is to attempt definitively to provide a robust, objective and universal method to address the centuries-old question (unresolved since Poulton 1904), "What is a species?", in the context of the allopatric problem and using real world data rather than molecular data.

Materials and methods

In the present study, I took a large data set that had been developed for purposes of various particular taxonomic studies of birds (citations below) and used this to road-test proposed and possible alternative statistical tests for measuring differentiation or diagnosis, with the intention of studying outcomes of tests in order to inform recommendations.

I compiled vocal and biometric data from multiple studies, including of representatives of the three major assemblages of birds: non-passerines (three families), suboscine passerines (four families), and oscine passerines (two families) (citations in Tables 1–2). In all of these studies, an exhaustive approach was applied to obtaining relevant sound recordings from the world's two largest avian sound recording repositories (as such databases stood prior to the point of publication): the xeno-canto.org collection and Macaulay Library, as well as commercially available CDs and DVDs and private sound recordings of the authors and other contacts. In relation to biometrics, most studies involved a relatively comprehensive set of available Colombian museum specimens, typically with over five and often more museums studied, including most of the main museums in Colombia, the USA, the UK, and France. For some studies, the largest Venezuelan collection was also studied. Full details of methods can be read in each relevant paper.

Vocal variables always included measures of maximum acoustic frequency, length, number of notes and speed. In some studies, change in pace, minimum frequencies, frequencies of particular notes, note bandwidth, changes in acoustic frequency and position of peaks or troughs of frequency within a vocalization, or any of the same measures for particular parts of vocalizations, were also measured. In each study, the variables under study were designed so as to document as fully as possible observed subjective differences between populations. Biometric variables were in all cases wing, tail, tarsus and bill length and mass, except for Trochilidae (no tarsus length) and Grallariidae (where bill width was additionally measured). Note shape and other subjective vocal characters were also studied, as were plumages. However, information on non-continuous variables was discarded for purposes of this present study.

Pairwise comparisons were undertaken on a matrix basis of each population against each other population. Some pairwise tests were omitted due to lack of data for a particular population, i.e., where there were $n < 2$ recordings of a particular type of vocalization (which could represent either a sampling gap or genuine lack of delivery of such vocalization by the population in question); or $n < 2$ specimens of the population available in museums that were studied. In such cases, where $n < 2$, standard deviations could not be calculated and *t*-tests could not be run, so the comparison was excluded to ensure full comparability between all tests applied.

The data set was not designed for the study of statistical tests used in taxonomy, since this study had not been conceived at the time of data collection. The choice of taxonomic groups was not based only on studies which include among their components sympatric pairs (cf. Tobias et al. 2010). Necessarily, in those studies involving more than two populations, not all the populations undergoing pairwise comparisons are sisters of one another and, in some instances, subsequent molecular studies have demonstrated other (unstudied) taxa to be sister to some of the populations in the group under study. The distribution of study species is highly localized to northwestern South America. The non-passerines part of the study set is much smaller than the passerines part. All studies involve situations where diversity appeared to have been previously underestimated at either species or subspecies level or both or followed a discovery of a new taxon whose taxonomic rank was investigated, resulting

Table 1. Summary information on the vocal studies used in the analysis.

Order: Family	Genus	No. taxa / populations	No. spp. before review	No. spp. after review	No. continuous vocal variables	No. Pairwise tests omitted	Pairwise comparisons	Sample sizes (mean ± s.d.) (min–max)	Reference
Columbiformes: Columbidae	*Geotrygon*	2	1	2	2	0	2	22.0 ± 4.6 (18–26)	Donegan and Salaman (2012)
Apodiformes: Trochilidae	*Adelomyia*	2	1	1	10	0	10	15.7 ± 1.5 (14–18)	Donegan and Avendaño (2015)
Piciformes: Bucconidae	*Hypnelus*	2	1	2	5	0	5	5.5 ± 1.6 (4–7)	Donegan et al. (2015)
Passeriformes: Thamnophilidae	*Myrmeciza*	8	4	5	26	114	614	42.7 ± 49.2 (3–179)	Donegan (2012)
Passeriformes: Grallariidae	*Grallaricula*	10	1	2	14	224	406	18.2 ± 12.9 (3–63)	Donegan (2008)
Passeriformes: Rhinocryptidae	*Scytalopus 1*	8	3	3	12	0	336	23.0 ± 13.8 (4–57)	Donegan and Avendaño (2008)
Passeriformes: Rhinocryptidae	*Scytalopus 2*	2	1	1	7	0	7	14.9 ± 2.2 (12–17)	Donegan et al. (2013)
Passeriformes: Tyrannidae	*Sirystes*	4	1	4	18	64	44	39.1 ± 41.5 (3–146)	Donegan (2013)
Passeriformes: Parulidae	*Basileuterus*	13	3	6	19	558	924	25.5 ± 19.0 (2–78)	Donegan (2014)
	TOTALS	**51**	**16**	**26**	**113**	**960**	**2348**	**29.0 ± 32.5 (2–179)**	

in a detailed study being undertaken. Most of the taxon pairs under comparison are subspecies/subspecies situations, and many of them involve populations that are not taxonomically recognized at all. For several populations, vocal studies were concluded without biometric data. For one study (*Anisognathus*), only biometric data were analyzed but not vocal data.

Several statistical tests were applied multiple times on a pairwise basis using a Microsoft Excel spreadsheet devised by the author for rapid assessment of multiple pairwise statistical tests across multiple populations. This spreadsheet is being published on the author's researchgate.net page, and should assist authors in better and more swiftly analyzing diagnosability in future studies. Calculations, described below, were undertaken to measure inter-population differences in the context of various species and subspecies concepts.

First, the entire data set was subjected to various proposed tests of species or subspecies rank. In the formulae used below, $\bar{x}_1$ and s_1 are the sample mean and standard deviations of Population 1; $\bar{x}_2$ and s_2 refer to the same parameters in Population 2; and the t value uses a one-sided confidence interval at the percentage specified for the

Table 2. Summary information on the biometric studies used in the analysis.

Order: Family	Genus	No. taxa / populations	No. spp. before review	No. spp. after review	No. continuous biometric variables	No. Pairwise tests omitted	Pairwise comparisons	Sample sizes (mean ± s.d.) (min.–max.)	Reference
Apodiformes: Trochilidae	*Adelomyia*	2	1	1	4	0	4	9.6 ± 3.2 (6–13)	Donegan and Avendaño (2015)
Passeriformes: Thamnophilidae	*Myrmeciza*	7	4	5	5	18	87	21.1 ± 19.5 (2–65)	Donegan (2012)
Passeriformes: Grallariidae	*Grallaricula*	11	1	3	6	49	281	12.4 ± 9.3 (3–37)	Donegan (2008)
Passeriformes: Rhinocryptidae	*Scytalopus 1*	8	4	3	5	31	109	8.7 ± 7.4 (2–24)	Donegan and Avendaño (2008)
Passeriformes: Rhinocryptidae	*Scytalopus 2*	2	1	1	5	0	5	4.9 ± 2.3 (3–9)	Donegan et al. (2013)
Passeriformes: Thraupidae	*Anisognathus*	10	2	2	5	39	186	25.1 ± 34.7 (4–214)	Donegan and Avendaño (2010)
Passeriformes: Parulidae	*Basileuterus*	9	3	5	5	30	150	15.4 ± 10.9 (2–42)	Donegan (2014)
	TOTALS	**49**	**16**	**20**	**35**	**167**	**822**	**15.5 ± 19.1 (2–214)**	

relevant population and variable, with t_1 referring to Population 1 and t_2 referring to Population 2.

LEVEL 1: Welch's *t*-test at $p<0.05/n_v$, i.e., applying a Bonferroni correction. An unequal variance (Welch's) *t*-test was used. This is preferable to other *t*-tests in that it makes no assumptions about whether the SD of one population differs from that of the other. For vocal data potentially based on ratios, such as song speed, a two-sample Kolmogorov-Smirnov test can be applied instead to account for the possibility of a non-normal distribution. However, in order to standardize the study outputs, only Welch's *t*-test was applied here.

When applying tests of statistical significance across multiple variables for the same pair, there is a risk of so-called "type 1" errors occurring. If testing for $p < 0.05$ for 100 independent variables of the same two populations, it would be expected that 5 variables would meet the requirements of the relevant test at this level of confidence. Various methods were tested which purport to reduce the risk of "type 1" errors. First, Bonferroni corrections were applied based on each of: (i) the total number of variables studied for the pair as a whole; (ii) separately for two "families" of vocal versus

biometric variables; and (iii) separately for each different kind of vocalization, where applicable. Applying Bonferroni correction for a study involving five variables, $p < (0.05/5) = 0.01$ is the corrected confidence interval. Dunn-Šidák is a widely used but less conservative alternative to Bonferroni and was applied also to all three of the same situations as above in order to examine the impacts and outcomes using alternative corrections.

LEVEL 2: a '50%/95%' test, following one of Hubbs and Perlmutter (1942)'s subspecies proposals but modified to control for sample size by using Isler et al. (1998)'s framework (see under Level 5) based on *t*-distribution. This test is passed if sample means are two average SDs or more apart controlling for sample size, i.e., the sample mean of each population falls outside the range of 95% of the other population:

$$|(\bar{x}_1 - \bar{x}_2)| > (s_1(t_{1\ @\ 97.5\%}) + s_2(t_{2\ @\ 97.5\%}))/2$$

LEVEL 3: The traditional '75% / 99+%' test for subspecies (Amadon 1949, Patten and Unitt 2002), modified to control for sample size, which requires both the following tests to be passed:

$$|(\bar{x}_1 - \bar{x}_2)| > s_1(t_{1\ @\ 99\%}) + s_2(t_{2\ @\ 75\%}) \text{ and}$$
$$|(\bar{x}_2 - \bar{x}_1)| > s_2(t_{2\ @\ 99\%}) + s_1(t_{1\ @\ 75\%})$$

LEVEL 4: diagnosability based on non-overlap of recorded values (the first part of Isler et al.'s 1998 diagnosability test).

LEVEL 5: 'Full' diagnosability (where sample means are four average SDs apart at the 95% level, controlling for sample size) the second part of Isler et al.'s (1998) diagnosability test:

$$|(\bar{x}_1 - \bar{x}_2)| > s_1(t_{1\ @\ 97.5\%}) + s_2(t_{2\ @\ 97.5\%})$$

Figure 1 illustrates how each of the Level 1 to Level 5 statistical tests measures differentiation.

These five tests were applied to 2348 population/variable combinations for voice and 822 population/variable combinations for biometrics. A population/variable combination is one comparison between two populations for a single variable. For example, in the *Grallaricula* study, a comparison of the main East Andes population against the Central Andes population for song length would constitute a single population/variable combination. With five diagnosability tests (Levels 1–5 above) conducted per population/variable combination, this means that a total of 15,610 pairwise statistical tests were run in this part of the study. (A further four tests conducted in later sections bring that total to over 28,000 separate statistical tests in this study.) Each population/variable combination was placed in a category summarizing which diagnosability tests it satisfied. The total number of population/variable combinations meeting particular tests was then summed for the biometric and vocal data sets separately, and then similar kinds of outcomes were grouped using the framework set out in Table 3. In order

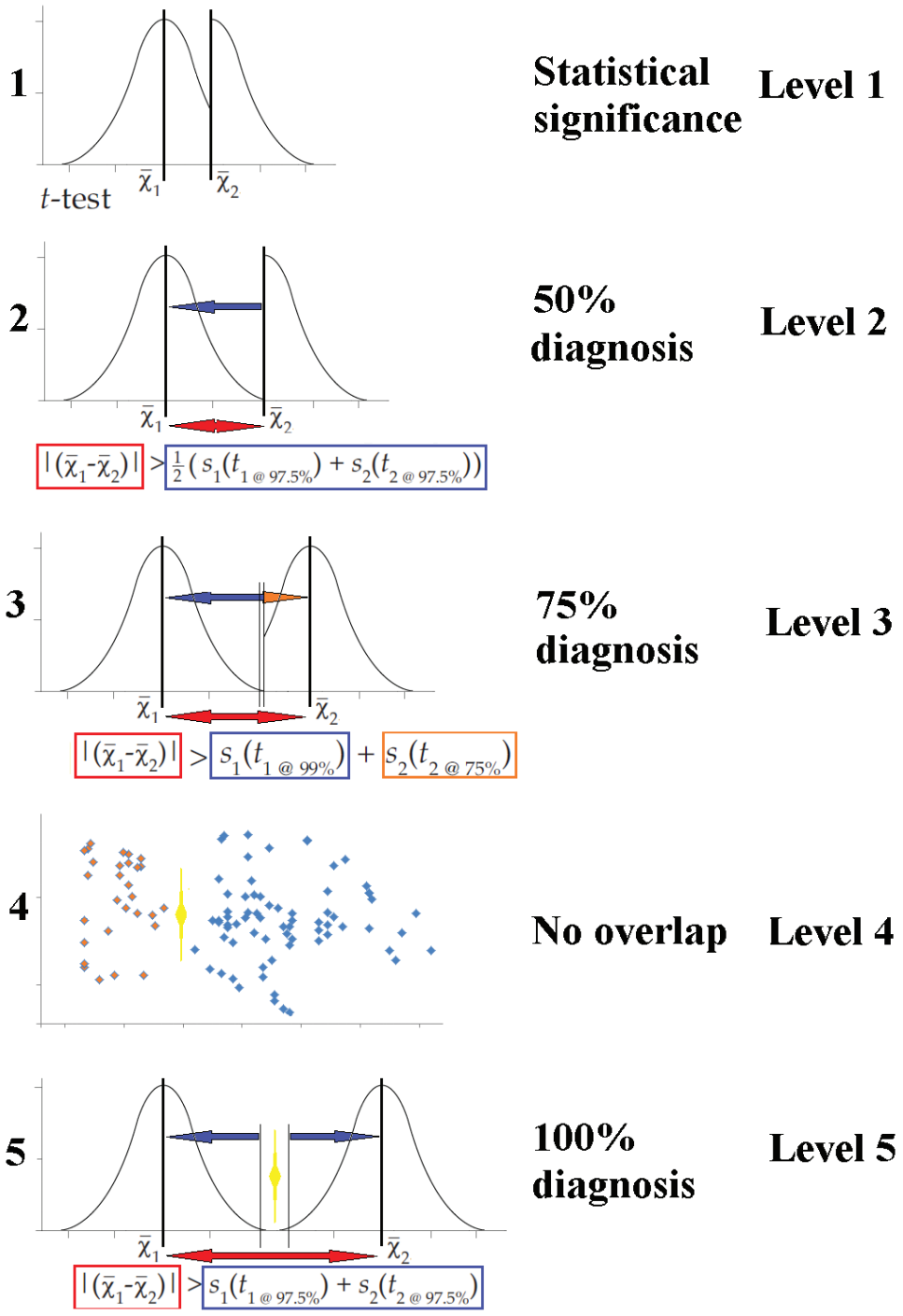

Figure 1. Graphical depiction of datasets which satisfy the Level 1–5 statistical tests addressed in this study.

Table 3. Recorded test satisfaction outcomes and the mapping of such outcomes to diagnosis groupings.

Outcome	Meaning	Grouping
0	None of the tests are met.	No diagnosis
1	Statistically significant difference between means but no tests of diagnosis are met and data overlap.	Statistical significance
14	Statistically significant difference between means and data show no overlap but no tests of diagnosis are met.	
12	Statistically significant difference between means but diagnosis only up to 50% and data overlap.	50% differentiation
124	Statistically significant difference between means, diagnosis up to 50% and data show no overlap.	
123	Statistically significant difference between means and diagnosis at both 50% and 75% levels but data overlap.	75% differentiation
1234	Statistically significant difference between means and diagnosis at both 50% and 75% levels and data do not overlap	
12345	Statistically significant difference between means and diagnosis at 50%, 75% and 95% levels and data do not overlap	95% differentiation
1235	Statistically significant difference between means and diagnosis at 50%, 75% and 95% levels but data overlap.	
1245	Statistically significant difference between means and diagnosis at 50% and 95% levels and data overlap but 75% test is not met.	
125	Statistically significant difference between means and diagnosis at 50% and 95% levels and data overlap but 75% test is not met and data overlap.	
2	No statistically significant difference between means, but 50% diagnosis test is met.	Possible false results
2345	No statistically significant difference between means, but 50%, 75% and 95% diagnosis tests are met.	
24	No statistically significant difference between means, but 50% diagnosis test is met and data do not overlap.	
245	No statistically significant difference between means, but 50% and 95% diagnosis tests are met and data do not overlap.	
25	No statistically significant difference between means and data overlap, but 50% and 95% diagnosis tests are met.	
4	Data do not overlap but no other statistical tests are met	

to consider taxonomic differences between the vocal and biometric data sets, data for studies involving the same taxonomic groups only are also presented.

Certain minor methodological changes were undertaken here as compared to some of the underlying studies on which this paper is based: (i) where a single population had only one data point, it was excluded here from analyses, since only "Level 4" tests can be applied where degrees of freedom are 0 and this paper sought to compare outcomes for all comparisons; (ii) for the number of notes in the call for *Grallaricula*, several populations had uniformly one note in their calls, with standard deviation of zero, producing "divide by zero" errors for several tests, and so pairwise comparisons between such populations for that variable were excluded; (iii) some underlying studies presented biometric data for either males or females or all specimens or both; here, one or other of the "male" or "all specimens" data sets was selected, depending on

whether material sexual differences in biometrics were observed and on sample size (generally, for studies with larger samples, using male only data is preferable, whilst in those studies with fewer specimens available, a combined data set was used here); (iv) for the main *Scytalopus* data set (Donegan and Avendaño 2008), whose study was accepted for publication during a formative stage of the development of the methods used here, the data sets needed amendment to apply some of the methods set out below; (v) Bonferroni correction for purposes of "Level 1" pass/fail analysis was applied based on number of vocal variables as a whole and not partitioned for different kinds of vocalization (this method ultimately being selected for reasons discussed later on); and (vi) only Welch's *t*-tests (and no other "Level 1" tests used in the underlying studies) were applied, to promote comparability of outcomes. As a result, the results here differ in some instances from those found in the appendices to some of the papers it is based upon. Overall, these methodological changes result in differing numbers of positive outcomes at Levels 1 and 4 in particular, compared to those presented in the original publications. Also as a result of these changes, the entire data set was re-analyzed using Excel spreadsheets in order to produce comparisons and ensure reliable counting, with no reliance on previously published analyses of the same data.

Effect sizes

The second part of this study aimed to measure effect sizes four different ways, in order to inform appropriate benchmarks for measuring or scoring differentiation. The impacts of using pooled standard deviations (as per Tobias et al. 2010), unpooled standard deviations (as per Isler et al. 1998) and of controlling for sample size using *t*-distribution (as per Isler et al. 1998) or not (as per Tobias et al. 2010) were compared.

Bare unpooled effect sizes

Effect sizes were first calculated using the following formula:

$$|(\bar{x}_1-\bar{x}_2)| / [(s_1+s_2)/2]$$

This uses an arithmetic mean of the standard deviations of the two populations to measure the difference between the means of the same two populations.

Controlled unpooled effect sizes

A control was applied using *t*-distribution values, following Isler et al. (1998), to produce a further set of effect size measurements:

$$|(\bar{x}_1 - \bar{x}_2)| \,/\, ¼[s_1(t_{1\ @\ 97.5\%}) + s_2(t_{2\ @\ 97.5\%})]$$

This measures the distance between the means of two populations in terms of numbers of SDs, but controlling for sample size using a *t* distribution. The factor of ¼ is included to maintain parity with bare unpooled effect sizes and other measures studied in this section, i.e., where mean differences are measured with the equivalent of a single standard deviation for their denominator. For a normal distribution, as *n* tends to infinity, *t* tends to *c*.2 (actually nearer to 1.98), capturing essentially the whole sample within 2 standard deviations. As a result, $s_1(t_{1\ @\ 97.5\%}) + s_2(t_{2\ @\ 97.5\%})$ is equivalent to $2s_1 + 2s_2$, or 4*s*, calling for division by 4 to retain parity with 1*s*.

To illustrate the impact of this correction versus the results from using bare unpooled effect sizes, the maximum acoustic frequency in the "slow song" in Santa Marta Warbler *Basileuterus basilicus* differs from that in the East Andes population of Three-striped Warbler *B. tristriatus* by 4.087 SDs, using bare unpooled effect sizes. The controlled unpooled effect size for this variable is lower at 3.910 SDs. This is because $n = 9$ for *basicilus* and $n = 53$ for East Andes *tristriatus*; one-sided *t*-distribution values at 97.5% are 2.306 and 2.007 respectively, effectively reflecting that an average SD of 2.157 using these sample sizes is equivalent to an SD of *c*.2 with infinite data points). This particular population/variable comparison therefore moved from being in a diagnosable category (> 4 SDs' difference) to *not* being diagnosable (< 4 SDs' difference and failing Isler et al. 1998's diagnosability test), when controlling for sample size.

Bare pooled effect sizes

Effect sizes using a pooled standard deviation, or Cohen's *d*, were calculated. First, the pooled standard deviation was calculated:

$$s_p = \sqrt{[((n_1-1)s_2^2 + (n_2-1)s_2^2))/(n_1+n_2-2)]}$$

Cohen's *d* was then calculated as:

$$|(\bar{x}_1 - \bar{x}_2)| \,/\, s_p.$$

or, in full:

$$|(\bar{x}_1 - \bar{x}_2)| \,/\, \sqrt{[((n_1-1)s_2^2 + (n_2-1)s_2^2))/(n_1+n_2-2)]}$$

This was the measure of effect size used by Tobias et al. (2010) and is used widely in science.

Controlled pooled effect sizes

Bare pooled effect sizes were subjected to an equivalent control for sample size (as for bare unpooled effect sizes), but using *t*-values at the degrees of freedom of the pooled standard deviation:

Cohen's d / (($t_{\text{pooled @ 97.5}}$)/2),

Where t_{pooled} is based on the degrees of freedom for the pooled standard deviation: d.f.=n_1+n_2–2.

or, in full:

$$|(\bar{x}_1-\bar{x}_2)| \,/\, (\sqrt{[((n_1-1)s_2^2 + (n_2-1)s_2^2))/(n_1+n_2-2)]}) \,/\, ((t_{\text{pooled@97.5\%}})/2).$$

Effect size buckets

Thee four measures of effect sizes were calculated for each population/variable combination and each outcome was then placed into two sets of buckets. First, in order to obtain a general resolution on effect sizes magnitude in taxonomic studies, population/variable combinations were placed into a set of buckets divided at 2 effect sizes (i.e., at approximately 50% differentiation) intervals: 0–2, 2–4, 4–6, etc. A second set of buckets was based on Tobias et al. (2010)'s scheme categories: character differences with an effect size of 0–0.2, 0.2–2; 2–5; 5–10 and >10.

Plots and correlations

To compare the outcomes achieved using the four different measures of effect size and analyses of Levels 1–5, plots were produced between several of the outcomes. Spearman's rank correlation coefficient was calculated as between statistical significance and effect size outcomes, based on the entire vocal and biometric data sets, so as to examine the inter-relation between the outcomes of applying different measures of differentiation.

Results

Type 1 correction analysis

Tables 4–5 illustrate the impacts on positive outcomes for statistical significance tests when applying different kinds of "type 1 correction". The data also provide more de-

Table 4. Effect of applying different "Type 1 error" corrections on the vocal data set. The tests are ordered (A-G) from least to most conservative corrections. *Sirystes*, *Geotrygon* and *Hypnelus* data are presented outside the totals, since there were no biometric data set on which more conservative cumulative corrections could be applied. In the case of the latter two genera, *Adelomyia* and *Scytalopus 2*, only one kind of vocalization was studied.

	Adelomyia	***Myrmeciza***	***Grallaricula***	***Scytalopus 1***	***Scytalopus 2***	***Basileuterus***	**TOTALS**	***[Sirystes]***	***[Geotrygon]***	***[Hypnelus]***
No. of vocal variables	10	26	14	12	7	19		18	2	5
A. No correction										
p<	0.05	0.05	0.05	0.05	0.05	0.05		0.05	0.05	0.05
Passed	2	406	323	236	4	240	1211	32	2	2
Total	10	614	406	336	7	343	1716	44	2	5
% passed	*20%*	*66.1%*	*79.6%*	*70.2%*	*57.1%*	*70.0%*	***70.6%***	*72.7%*	*100%*	*40%*
B. Dunn-Šidák with each kind of vocalisation separately										
p<	0.00512	0.00639	0.00730	0.00851	0.00730	0.00730		0	0.0253	0.0102
Passed	2	357	293	197	3	200	1052	25	2	1
Total	10	614	406	336	7	343	1716	44	2	5
%	*20%*	*58.1%*	*72.2%*	*58.6%*	*42.9%*	*58.3%*	***61.3%***	*56.8%*	*100%*	*20%*
C. Bonferroni with each kind of vocalisation separately										
p<	0.005	0.00625	0.00714	0.00833	0.00714	0.00714286		0.01	0.025	0.01
Passed	2	357	293	197	3	200	1052	25	2	1
Total	10	614	406	336	7	343	1716	44	2	5
%	*20%*	*58.1%*	*72.2%*	*58.6*	*42.9%*	*58.3%*	***61.3%***	*56.8%*	*100%*	*20%*
D. Dunn-Šidák with voice and biometrics separately										
p<	0.00512	0.00197	0.00366	0.00427	0.00730	0.00730		0.00285	0.0253	0.0102
Passed	2	321	267	186	3	189	968	20	2	1
Total	10	614	406	336	7	343	1716	44	2	5
%	*20%*	*52.2%*	*65.8%*	*55.3%*	*42.9%*	*55.1%*	***56.4%***	*45.4%*	*100%*	*20%*
E. Bonferroni with voice and biometrics separately										
p<	0.005	0.00192	0.00357	0.00417	0.00714	0.00714		0.00278	0.025	0.01
Passed	2	321	266	185	3	188	965	20	2	1
Total	10	614	406	336	7	343	1716	44	2	5
%	*20%*	*52.2%*	*65.5%*	*55.1%*	*42.9%*	*54.8%*	***56.2%***	*45.5%*	*100%*	*20%*
F. Dunn-Šidák: biometrics plus voice										
p<	0.00366	0.00165	0.00256	0.00301	0.00427	0.00427				
Passed	2	317	260	179	3	182	943			
Total	10	614	406	336	7	343	1716			
%	*20%*	*51.6%*	*64.0%*	*53.3%*	*42.9%*	*53.1%*	***55.0%***			
G. Bonferroni: biometrics plus voice										
p<	0.00357	0.00161	0.0025	0.00294	0.00417	0.00417				
Passed	2	317	260	178	3	181	941			
Total	10	614	406	336	7	343	1716			
%	*20%*	*51.6%*	*64.0%*	*53.0%*	*42.9%*	*52.8%*	***54.8%***			

Table 5. Effect of applying different "Type 1 error" corrections on the biometric data set. The tests are ordered (A–E) from least to most conservative corrections. *Anisognathus* data are presented outside the totals, since there was no vocal data set on which more conservative cumulative corrections could be applied.

	Adelomyia	*Myrmeciza*	*Grallaricula*	*Scytalopus 1*	*Scytalopus 2*	*Basileuterus*	TOTALS	[*Anisognathus*]
No. of biometric variables	4	5	6	5	5	5		5
A. No correction								
p<	0.05	0.05	0.05	0.05	0.05	0.05		0.05
Passed	0	46	142	45	3	66	**302**	88
Total	4	87	281	109	5	150	**636**	186
% passed	*0%*	*52.9%*	*51.8%*	*41.3%*	*60%*	*44%*	***47.5%***	*47.3%*
B. Dunn-Šidák with biometrics and voice separately								
p<	0.0127	0.0102	0.00851	0.0102	0.0102	0.0102		0.0102
Passed	0	31	108	35	2	50	**226**	66
Total	4	87	281	109	5	150	**636**	186
%	*0%*	*35.6%*	*38.4%*	*32.1%*	*40%*	*33.3%*	***35.5%***	*35.5%*
C. Bonferroni with biometrics and voice separately								
p<	0.0125	0.01	0.00833	0.01	0.01	0.01		0.01
Passed	0	31	108	35	2	50	**226**	66
Total	4	87	281	109	5	150	**636**	186
%	*0%*	*35.6%*	*38.4%*	*32.1%*	*40%*	*33.3%*	***35.5%***	*35.5%*
D. Dunn-Šidák: biometrics plus voice								
p<	0.00366	0.00165	0.00256	0.00301	0.00427	0.00213		
Passed	0	33	92	31	2	40	**198**	
Total	4	87	281	109	5	150	**636**	
%	*0%*	*37.9%*	*32.7%*	*28.4%*	*40%*	*26.7%*	***31.1%***	
E. Bonferroni: biometrics plus voice								
p<	0.00357	0.00161	0.0025	0.00294	0.00417	0.00208		
Passed	0	33	92	31	2	40	**198**	
Total	4	87	281	109	5	150	**636**	
%	*0%*	*37.9%*	*32.7%*	*28.4%*	*40%*	*26.7%*	***31.1%***	

tailed information on "Level 1" diagnosis (on which see further Tables 6–9 and "Levels analysis" below). For vocal variables (Table 4), over 70% of pairwise comparisons passed the Level 1 test using $p<0.05$. However, almost 15% of positive outcomes were eliminated when using the most conservative correction. The greatest impact among the cascade of tested corrections was to correct for sample size at all, which eliminated over 9% of positive outcomes. Fewer than 5% of outcomes were eliminated by treating all vocal variables as linked. The final impact on vocal data, treating all biometrics and voice as part of the same family of variables, affected <1.3% of outcomes. Generally speaking, Dunn-Šidák corrections had virtually nil impact compared to Bonferroni, with only five individual movements (<0.2%) from significant to non-significant categories across the entire set of vocal comparisons.

In the biometrics study, lower levels of statistically significant differentiation were found than for voice. More comparisons were non-significant (52.5%) than significant, even prior to applying any type 1 corrections. Applying type 1 corrections eliminated a further 12–16% of outcomes. Fewer than 5% of these eliminations result from treating voice and biometrics together; the bulk resulted from applying Bonferroni on the biometric data set itself. Dunn-Šidák corrections had no impact compared to using Bonferroni.

Levels analysis

Tables 6–9 summarize the outcomes of diagnosis tests using the "Levels 1–5" model. After grouping the data, three main categories of positive diagnosis were revealed across the two studies, for both biometrics and voice: statistically significant, 50% differentiation and 95% differentiation. The category for 75% differentiation represented < 2.5% of outcomes in both studies.

As foreshadowed in the type 1 error analysis (Tables 4–5), "no diagnosis" was the largest segment in the voice study, albeit a minority overall. For biometrics, "no diagnosis" exceeded all other outcomes combined. Possible false results were < 3% for the vocal sample but rose to 8.4% for the biometric data set, mostly relating to instances of non-overlap (Level 4). Such outcomes are more frequent when dealing with the smaller sample sizes that are more regularly presented by studies of specimens (see Tables 1–2). Experience from the process of collecting data during the course of these studies and re-running analyses is that Level 4 differences in initial analyses will ultimately often convert into Level 1, 2, 3 or even 5 differences with a greater sample, whilst others will erode to nothing and will simply have reflected a clustering of data points.

Isler et al. (1998)'s gold standard of diagnosis was met by 14.5% of vocal pairwise comparisons and 6.3% of biometrics comparisons. Approximately triple this number of outcomes, a total of 36% (voice) and 29% (biometrics) of outcomes, involved non-diagnosable but statistically significant differentiation.

Levels 1–5 were generally ordered by least to most exacting in terms of difficulty to pass. However, several examples of "outliers" were uncovered, where more liberal test outcomes were apparently "skipped", e.g.: (i) only statistical significance and non-overlap (1&4); (ii) statistical significance with 50% and non-overlap but not 75% diagnosis (124); (iii) all tests being passed except non-overlap (123&5); (iv) all tests including 95% diagnosis being passed, but excluding 75% diagnosis (124&5); (v) full statistical diagnosis and 50% and 95% diagnosis being met but neither 75% nor non-overlap (12&5); and (vi) combinations skipping statistical significance altogether, but passing other tests (all outcomes starting with 2 or 4). These outcomes are all statistically plausible, including as a result of the values of t at particular sample sizes for different confidence limits, even if in some cases they are logically counterintuitive.

Table 6. Outcomes of pairwise comparisons for vocal characters, placed into the different categories recovered by testing statistical tests of Levels 1 through 5. 1 = statistically significant, 2 = 50% diagnosis, 3=75% diagnosis, 4 = actual value diagnosis, 5 = 95% diagnosis (as detailed further in methods). See Table 3 for further information on meaning of codes used here.

Voice: levels passed	**None**	**1**	**14**	**12**	**124**	**123**	**1234**	**12345**	**1235**	**1245**	**125**	**2**	**2345**	**24**	**245**	**4**
Geotrygon		1			1											
Adelomyia	8	2														
Hypnelus	3				1											1
Myrmeciza	284	38	53	35	23	9	4	109	11	2	37	2		5	2	
Grallaricula	118	92	1	47	27	7	5	71	3	12	1	2	12			8
Scytalopus 1	148	82		37	17	5	7	36	1			1		2		
Scytalopus 2	4				3											
Sirystes	22	7		7	3		1	2				1			1	
Basileuterus	504	188	3	73	53	7	6	47	7	2	0	1		11	0	22
TOTAL	**1091**	**410**	**57**	**199**	**128**	**28**	**23**	**265**	**22**	**16**	**38**	**7**	**12**	**18**	**3**	**31**
Percentages	*46.5%*	*17.5%*	*2.4%*	*8.5%*	*5.5%*	*1.2%*	*1.0%*	*11.3%*	*0.9%*	*0.7%*	*1.6%*	*0.3%*	*0.5%*	*0.8%*	*0.1%*	*1.3%*

Table 7. Outcomes of pairwise comparisons for biometric characters, placed into the different categories recovered by testing statistical tests of Levels 1 through 5. 1 = statistically significant, 2 = 50% diagnosis, 3 = 75% diagnosis, 4 = actual value diagnosis, 5 = 95% diagnosis (as detailed further in methods). See Table 3 for further information on meaning of codes used here.

Biometrics: levels passed	**None**	**1**	**14**	**12**	**124**	**123**	**1234**	**12345**	**1235**	**1245**	**125**	**2**	**2345**	**24**	**245**	**4**
Adelomyia	4															
Myrmeciza	25	24		9	1			4								24
Grallaricula	166	17	1	15	23	1	5	35		3				7		8
Scytalopus 1	47	6	4	8	12		2	3								27
Scytalopus 2	3				1			1								
Anisognathus	120	43		12	6		1	4								
Basileuterus	97	27	1	10	9		1	2						1		2
TOTAL	**462**	**117**	**6**	**54**	**52**	**1**	**9**	**49**	**0**	**3**	**0**	**0**	**0**	**8**	**0**	**61**
Percentages	*56.2%*	*14.2%*	*0.7%*	*6.6%*	*6.3%*	*0.1%*	*1.1%*	*6.0%*	*0.0%*	*0.4%*	*0.0%*	*0.0%*	*0.0%*	*1.0%*	*0.0%*	*7.4%*

Table 8. Outcomes of pairwise comparisons using Levels analysis, for voice, by grouping. See Table 3 for information on Levels groupings used for column labels. % comparable data includes only those data sets in which both biometrics and voice were studied.

Voice: Taxon	Pairwise statistical tests (/5)	No diff.	Poss. false results	Signif. Only	50%	75%	95%
Geotrygon	2	0	0	1	1	0	0
Adelomyia	10	8	0	2	0	0	0
Hypnelus	5	3	1	0	1	0	0
Myrmeciza	614	284	9	91	58	13	159
Grallaricula	406	118	22	93	74	12	87
Scytalopus 1	336	148	3	82	54	12	37
Scytalopus 2	7	4	0	0	3	0	0
Sirystes	44	22	2	7	10	1	2
Basileuterus	924	504	34	191	126	13	56
TOTALS	**2348**	**1091**	**71**	**467**	**327**	**51**	**341**
OVERALL %		*46.5%*	*3.0%*	*19.9%*	*13.9%*	*2.2%*	*14.5%*
% (comparable)		*46.4%*	*2.9%*	*20.0%*	*13.7%*	*2.1%*	*14.7%*

In terms of specific findings for birds, biometric data were less informative than vocal data with "possibly false results" also being more frequent for biometric comparisons. Tobias et al. (2010) also found that vocal characters exhibit greater measured differentiation than biometric variables. The biometric data set rarely attained higher levels of diagnosability, with 75% and 95% outcomes around half those for vocal data. This pattern remains after controlling for taxonomy.

Effect sizes by 2d

Results for effect sizes divided into buckets of 2*d* are set out in Tables 10–11 for each of the four effect size measures used in the study, in each case for both voice and biometrics. In all data sets, a predominance of low differentiation (0–2 effect sizes, or less than 50% differentiation) is evident. A considerable 63–74% of outcomes fell into this lowest category, with a gradual tailing off of outcomes at increasing levels of differentiation.

A good portion (15–22%) of outcomes fell into the 2–4 effect sizes category, which, when using controlled unpooled effect sizes, corresponds to Level 2 in Tables 5–6. Outcomes in this bucket exceeded the total number of outcomes across all higher diagnosability categories. Even after applying the most conservative effect size calculations, very large effect sizes of over 20 were recorded in a handful of instances. Outcomes in all categories above 4–6 (inclusive) using controlled unpooled effect sizes correspond to the number of outcomes meeting Isler et al. (1998)'s diagnosis test (Level 5 in Tables 6–7), which is based on a score of 4*d* or more.

Table 9. Outcomes of pairwise comparisons using Levels analysis, for biometrics, by grouping. See Table 3 for information on Levels groupings used for column labels. % comparable data includes only those data sets in which both biometrics and voice were studied.

Biometrics: Taxon	Pairwise statistical tests (/5)	No diff.	Poss. false results	Signif. only	50%	75%	95%
Adelomyia	4	4	0	0	0	0	0
Myrmeciza	87	25	24	24	10	0	4
Grallaricula	281	166	15	18	38	6	38
Scytalopus 1	109	47	27	10	20	2	3
Scytalopus 2	5	3	0	0	1	0	1
Anisognathus	186	120	0	43	18	1	4
Basileuterus	150	97	3	28	19	1	2
TOTALS	**822**	**462**	**69**	**123**	**106**	**10**	**52**
OVERALL %		*56.2%*	*8.4%*	*15.0%*	*12.9%*	*1.2%*	*6.3%*
% (comparable)		*53.8%*	*10.8%*	*12.6%*	*13.8%*	*1.4%*	*7.5%*

Table 12 illustrates the impact of applying increasingly more conservative tests of effect size using the 2*d* analysis, which is discussed further under "Pooled versus unpooled and bare versus controlled effect sizes" below.

Effect sizes by Tobias et al. (2010) categories

Results for effect sizes divided into Tobias et al. (2010) bucket categories are set out in Tables 13–14 for each of the four effect size measures used in the study, in each case for both voice and biometrics. In contrast to the Levels 1–5 analysis, where "no diagnosis" predominated, or the 2*d* buckets, where the lowest category was largest, an overwhelming proportion (87–91%) of outcomes scored 1 or more under this system. The 3-point threshold of 5 or more effect sizes returned fewer positive scores (3.8–11.9%) than the Level 5 (or Isler et al. 1998) test of diagnosability or the total of elements in 2*d* buckets over 4 effect sizes. Overall, 77–85% of outcomes scored 1 or 2 points.

Changes between category (Table 15) were reduced here compared to the 2*d* effect size analysis, reflecting the smaller number of diagnosability categories studied and their greater effect size ranges.

Pooled versus unpooled and bare versus controlled effect sizes

Tables 12 and 15 summarize the impacts of applying different tests of effect size to the 2*d* and Tobias et al. (2010) categories studies. In both the biometric and vocal studies, the least to most conservative ways of calculating effect size were: (i) bare unpooled effect sizes; (ii) bare pooled effect sizes; (iii) controlled pooled effect sizes; and finally

Table 10. Results of the effects size study for voice, partitioning the data into 2 effect size intervals. The top two tables are based upon actual standard deviations for each set of data subjected to pairwise comparison. The lower two tables are based on pooled standard deviation data pooling. In each case, "bare" effect sizes are shown first (above). The second and fourth tables use "controlled effect sizes" for the relevant pooling approach, calculated by taking into account *t*-distribution values for the relevant sample size (or pooled sample size).

Bare Unpooled Effect Sizes (Voice)											
Taxon	**0–2**	**2–4**	**4–6**	**6–8**	**8–10**	**10–12**	**12–14**	**14–16**	**16–18**	**18–20**	**20+**
Geotrygon	1	1									
Adelomyia	10										
Hypnelus	3	2									
Myrmeciza	361	119	55	33	11	14	8	6	6	1	0
Grallaricula	208	91	47	16	10	7	3	14	6	3	1
Scytalopus 1	227	66	30	7	6						
Scytalopus 2	4	3									
Sirystes	27	14	2	0	1						
Basileuterus	654	168	47	29	20	4	1	1	0	0	0
TOTAL	**1497**	**462**	**181**	**85**	**48**	**25**	**12**	**21**	**12**	**4**	**1**
Percentage	*63.7%*	*19.8%*	*7.7%*	*3.6%*	*2.0%*	*1.1%*	*0.5%*	*0.9%*	*0.5%*	*0.2%*	*0.0%*
Controlled Unpooled Effect Sizes (Voice)											
Taxon	**0–2**	**2–4**	**4–6**	**6–8**	**8–10**	**10–12**	**12–14**	**14–16**	**16–18**	**18–20**	**20+**
Geotrygon	1	1									
Adelomyia	10										
Hypnelus	4	1									
Myrmeciza	375	114	53	32	14	15	0	7	3	1	
Grallaricula	219	88	43	25	14	5	6	4	1	0	1
Scytalopus 1	230	69	27	7	3						
Scytalopus 2	4	3									
Sirystes	29	12	3								
Basileuterus	717	151	34	14	6	2					
TOTAL	**1589**	**439**	**160**	**78**	**37**	**22**	**6**	**11**	**4**	**1**	**1**
Percentage	*67.7%*	*18.7%*	*6.8%*	*3.3%*	*1.6%*	*0.9%*	*0.3%*	*0.5%*	*0.2%*	*0.0%*	*0.0%*
Bare Pooled Effect Sizes (Voice)											
Taxon	**0–2**	**2–4**	**4–6**	**6–8**	**8–10**	**10–12**	**12–14**	**14–16**	**16–18**	**18–20**	**20+**
Geotrygon	1	1									
Adelomyia	10										
Hypnelus	3	2									
Myrmeciza	371	130	48	19	17	10	9	4	6	0	0
Grallaricula	222	98	33	16	6	9	7	6	3	2	4
Scytalopus 1	232	58	36	5	5						
Scytalopus 2	4	3									
Sirystes	28	13	2	0	0	0	0	1			
Basileuterus	687	151	40	16	11	8	5	2	1	1	2
TOTAL	**1558**	**456**	**159**	**56**	**39**	**27**	**21**	**13**	**10**	**3**	**6**
Percentage	*66.4%*	*19.4%*	*6.8%*	*2.4%*	*1.7%*	*1.1%*	*0.9%*	*0.6%*	*0.4%*	*0.1%*	*0.3%*

Bare Unpooled Effect Sizes (Voice)											
Controlled Pooled Effect Sizes (Voice)											
Taxon	**0–2**	**2–4**	**4–6**	**6–8**	**8–10**	**10–12**	**12–14**	**14–16**	**16–18**	**18–20**	**20+**
Geotrygon	1	1									
Adelomyia	10										
Hypnelus	3	2									
Myrmeciza	374	127	49	21	19	6	8	4	6		
Grallaricula	226	99	31	17	8	8	6	5	2	2	2
Scytalopus 1	233	58	35	5	5						
Scytalopus 2	4	3									
Sirystes	28	13	2	0	0	0	1				
Basileuterus	694	153	31	22	8	7	4	1	1	1	2
TOTAL	**1573**	**456**	**148**	**65**	**40**	**21**	**19**	**10**	**9**	**3**	**4**
Percentage	*67.0%*	*19.4%*	*6.3%*	*2.8%*	*1.7%*	*0.9%*	*0.8%*	*0.4%*	*0.4%*	*0.1%*	*0.2%*

(iv) controlled unpooled effect sizes. However, this shift was in no ways uniform, as is illustrated in Figures 2–3 and Tables 16–19.

The largest shift was observed between bare unpooled effect sizes versus controlled unpooled effect sizes, where a 3.9% (voice) or 9.1% (biometrics) increase in the number of outcomes in the lowest category of differentiation (0–2) was observed.

The impact of using bare pooled versus bare unpooled effect sizes is illustrated in Figures 2–3. Bare pooled effect sizes were overall more conservative than bare unpooled effect sizes. Several outcomes increased in effect size under a pooled method, which will occur where the population with the larger sample had a smaller standard deviation than the population with the smaller sample. Using pooled standard deviations generally had the result of reducing the magnitude of effect sizes compared to using unpooled standard deviations, which must relate to instances of smaller standard deviations in data sets with smaller sample sizes. This may be a natural phenomenon for highly localized and specialized populations or could result from clustering.

The overall magnitude of reduction of effect size measurements between bare pooled effect sizes and controlled pooled effect sizes was moderate. Degrees of freedom for pooled standard deviation are higher (the sum of the two samples' sample sizes minus 2) than when using unpooled methods (where each sample is treated separately), resulting in lower *t*-values when using pooled standard deviations. Application of *t*-distribution corrections on effect sizes using unpooled standard deviations resulted in the most conservative of all measures of effect sizes, linked to overall lowest degrees of freedom in corrections and overall higher *t*-values.

Although these overall trends were observed, the impact of applying differing methods of measurement of effect sizes on actual pairwise comparisons was not uniform (see Figures 2–3). The movement to lower categories in more conservative tests was merely an overall trend, with >97% correlation according to Spearman's ranking correlation coefficients (Tables 16–17). Even the application of "corrections" for

Table 11. Results of the effects size study for biometrics, partitioning the data into 2 effect size intervals. The top two tables are based upon actual standard deviations for each set of data subjected to pairwise comparison. The lower two tables are based on pooled standard deviation data. In each case, "bare" effect sizes are shown first (above). The second and fourth tables use "controlled effect sizes" for the relevant pooling approach, calculated by taking into account *t*-distribution values for the relevant sample size (or pooled sample size).

Bare Unpooled Effect Sizes (Biometrics)											
Taxon	**0–2**	**2–4**	**4–6**	**6–8**	**8–10**	**10–12**	**12–14**	**14–16**	**16–18**	**18–20**	**20+**
Adelomyia	4										
Myrmeciza	58	25	4								
Grallaricula	170	63	21	14	5	3	1	2	0	2	
Scytalopus 1	67	33	7	1	1						
Scytalopus 2	2	2	0	1							
Anisognathus	154	26	5	0	1						
Basileuterus	116	27	7								
TOTAL	**571**	**176**	**44**	**16**	**7**	**3**	**1**	**2**	**0**	**2**	**0**
Percentage	*69.5%*	*21.4%*	*5.4%*	*1.9%*	*0.9%*	*0.4%*	*0.1%*	*0.2%*	*0.0%*	*0.2%*	*0.0%*
Controlled Unpooled Effect Sizes (Biometrics)											
Taxon	**0–2**	**2–4**	**4–6**	**6–8**	**8–10**	**10–12**	**12–14**	**14–16**	**16–18**	**18–20**	**20+**
Adelomyia	4										
Myrmeciza	73	10	4								
Grallaricula	192	51	20	12	3	0	0	1	2		
Scytalopus 1	84	22	2	1							
Scytalopus 2	3	1	1								
Anisognathus	163	19	3	1							
Basileuterus	127	21	2								
TOTAL	**646**	**124**	**32**	**14**	**3**	**0**	**0**	**1**	**2**	**0**	**0**
Percentage	*78.6%*	*15.1%*	*3.9%*	*1.7%*	*0.4%*	*0.0%*	*0.0%*	*0.1%*	*0.2%*	*0.0%*	*0.0%*
Bare Pooled Effect Sizes (Biometrics)											
Taxon	**0–2**	**2–4**	**4–6**	**6–8**	**8–10**	**10–12**	**12–14**	**14–16**	**16–18**	**18–20**	**20+**
Adelomyia	4										
Myrmeciza	57	28	2								
Grallaricula	177	54	23	14	3	5	2	2	1		
Scytalopus 1	67	36	6								
Scytalopus 2	2	2	1								
Anisognathus	158	23	4	0	1						
Basileuterus	127	17	6								
TOTAL	**592**	**160**	**42**	**14**	**4**	**5**	**2**	**2**	**1**	**0**	**0**
Percentage	*72.0%*	*19.5%*	*5.1%*	*1.7%*	*0.5%*	*0.6%*	*0.2%*	*0.2%*	*0.1%*	*0.0%*	*0.0%*
Controlled Pooled Effect Sizes (Biometrics)											
Taxon	**0–2**	**2–4**	**4–6**	**6–8**	**8–10**	**10–12**	**12–14**	**14–16**	**16–18**	**18–20**	**20+**
Adelomyia	4										
Myrmeciza	58	27	2								
Grallaricula	181	54	21	11	8	3	1	1	1		
Scytalopus 1	75	30	4								
Scytalopus 2	3	2									
Anisognathus	161	21	3	0	1						
Basileuterus	128	17	5								
TOTAL	**610**	**151**	**35**	**11**	**9**	**3**	**1**	**1**	**1**	**0**	**0**
Percentage	*74.2%*	*18.4%*	*4.3%*	*1.3%*	*1.1%*	*0.4%*	*0.1%*	*0.1%*	*0.1%*	*0.0%*	*0.0%*

Table 12. Changes in effect size categories resulting from increasingly more conservative tests of effect size being applied. This table is based upon changes between the categories in Tables 10–11.

Voice: changes into effect size category	**0–2**	**2–4**	**4–6**	**6–8**	**8–10**	**10–12**	**12–14**	**14–16**	**16–18**	**18–20**	**20+**
Bare Unpooled –> Bare Pooled	+63	-8	-22	-29	-9	+2	+9	-8	-2	-1	+5
Bare Pooled –> Controlled Pooled	+15	0	-11	+9	+1	-6	-2	-3	-1	0	-2
Controlled Pooled –> Controlled Unpooled	+16	-17	+12	+13	-3	+1	-13	+1	-5	-2	-3
Total change from Bare Unpooled –> Controlled Unpooled	**+94**	**-25**	**-21**	**-7**	**-11**	**-3**	**-6**	**-10**	**-8**	**-3**	**0**
As percentage of total	*+4.0%*	*-1.1%*	*-0.9%*	*-0.3%*	*-0.5%*	*-0.1%*	*-0.3%*	*-0.4%*	*-0.3%*	*-0.1%*	*0.0%*
Biometrics: changes into effect size category	**0–2**	**2–4**	**4–6**	**6–8**	**8–10**	**10–12**	**12–14**	**14–16**	**16–18**	**18–20**	**20+**
Bare Unpooled –> Bare Pooled	+21	-16	-2	-2	-3	+2	+1	0	+1	-2	0
Bare Pooled –> Controlled Pooled	+18	-9	-7	-3	+5	-2	-1	-1	0	0	0
Controlled Pooled –> Controlled Unpooled	+36	-27	-3	+3	-6	-3	-1	0	+1	0	0
Total change from Bare Unpooled –> Controlled Unpooled	**+75**	**-52**	**-12**	**-2**	**-4**	**-3**	**-1**	**-1**	**+2**	**-2**	**0**
As percentage of total	*+9.1%*	*-6.3%*	*-1.5%*	*-0.2%*	*-0.5%*	*-0.4%*	*-0.1%*	*-0.1%*	*+0.2%*	*-0.2%*	*0.0%*

Table 13. Results of the effects size study for voice using Tobias et al. (2010) categories. The top two tables are based upon actual standard deviations for each set of data subjected to pairwise comparison. The lower two tables are based on pooled standard deviation data. In each case, "bare" effect sizes are shown first (above). The second and fourth tables use "controlled effect sizes" for the relevant pooling approach, calculated by taking into account *t*–distribution values for the relevant sample size (or pooled sample size).

Bare Unpooled Effect Sizes (Voice)					
Taxon	**0–0.2**	**0.2–2**	**2–5**	**5–10**	**10+**
Geotrygon	0	1	1		
Adelomyia	3	7			
Hypnelus	0	3	2		
Myrmeciza	70	291	152	66	35
Grallaricula	34	174	119	45	34
Scytalopus 1	31	196	83	26	
Scytalopus 2	0	4	3		
Sirystes	4	23	14	3	
Basileuterus	96	558	200	64	6
TOTAL	**239**	**1257**	**574**	**204**	**75**
Percentage	*10.1%*	*53.5%*	*24.4%*	*8.7%*	*3.2%*

Controlled Unpooled Effect Sizes (Voice)					
Taxon	**0–0.2**	**0.2–2**	**2–5**	**5–10**	**10+**
Geotrygon	0	1	1		
Adelomyia	3	7			
Hypnelus	0	4	1		
Myrmeciza	73	302	144	69	26
Grallaricula	37	182	121	49	17
Scytalopus 1	33	197	85	21	
Scytalopus 2	0	4	3		
Sirystes	4	25	12	3	
Basileuterus	113	604	171	34	2
TOTAL	**263**	**1326**	**538**	**176**	**45**
Percentage	*11.2%*	*56.5%*	*22.9%*	*7.5%*	*1.9%*
Bare Pooled Effect Sizes (Voice)					
Taxon	**0–0.2**	**0.2–2**	**2–5**	**5–10**	**10+**
Geotrygon	0	1	1		
Adelomyia	3	7			
Hypnelus	0	3	2		
Myrmeciza	71	300	157	57	29
Grallaricula	35	187	118	35	31
Scytalopus 1	31	201	78	26	
Scytalopus 2	0	4	3		
Sirystes	4	24	14	2	
Basileuterus	105	582	181	37	19
TOTAL	**249**	**1309**	**554**	**157**	**79**
Percentage	*10.6%*	*55.7%*	*23.6%*	*6.7%*	*3.4%*
Controlled Pooled Effect Sizes (Voice)					
Taxon	**0–0.2**	**0.2–2**	**2–5**	**5–10**	**10+**
Geotrygon	0	1	1		
Adelomyia	3	7			
Hypnelus	0	3	2		
Myrmeciza	71	303	158	58	24
Grallaricula	36	190	115	40	25
Scytalopus 1	30	203	79	24	
Scytalopus 2	0	4	3		
Sirystes	4	24	14	2	
Basileuterus	109	585	177	37	16
TOTAL	**253**	**1320**	**549**	**161**	**65**
Percentage	*10.8%*	*56.2%*	*23.4%*	*6.9%*	*2.8%*

sample size resulted in some increases in effect size measures for other sets using large samples, since *t* tends to 1.98 rather than 2 for sample sizes of over 100.

Statistical significance presented a weak negative correlation with most effect size measurements, but being most closely correlated with controlled unpooled effect sizes. In the case of biometrics, there was a strong negative correlation with controlled unpooled effect sizes (Tables 16–17). The strongest correlations were between the two effect size measurements using pooled standard deviations, which is consistent with the relatively modest correction resulting from the control for sample size, discussed above.

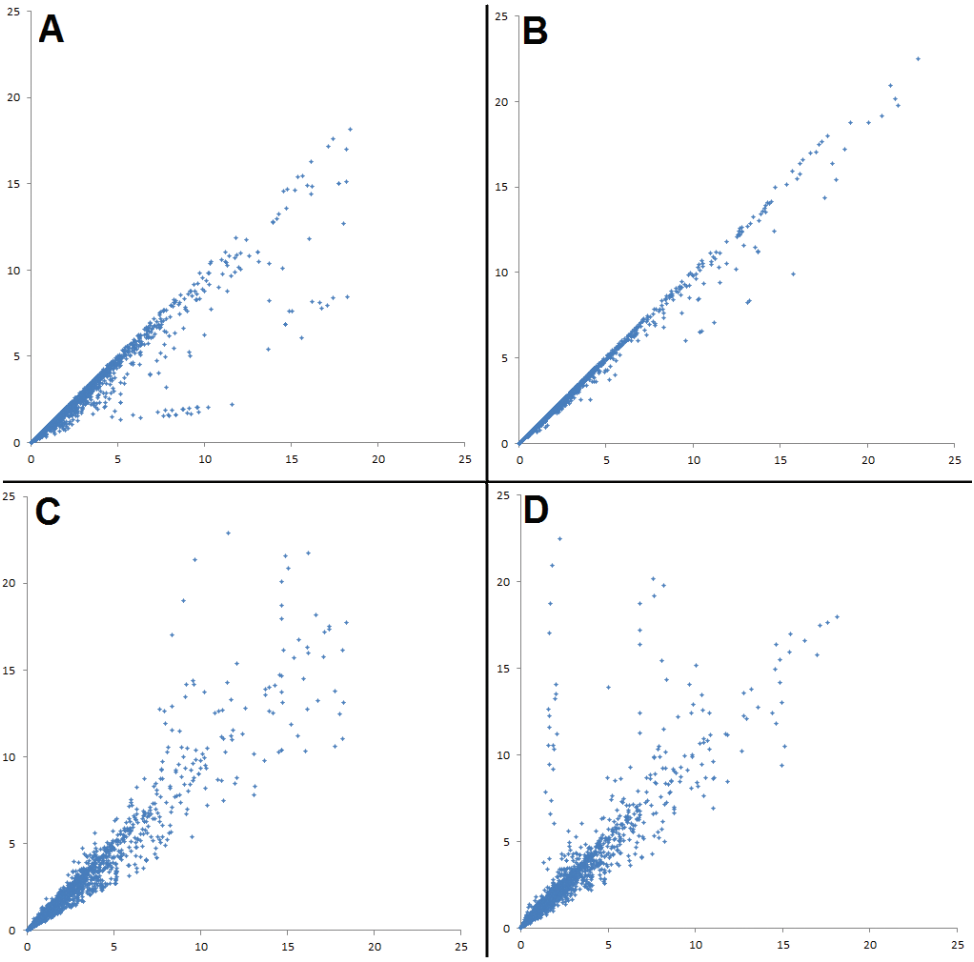

Figure 2. Scatter-graphs showing the effects of applying different corrections of effect size on the entire vocal data set. Each axis shows effect size, measured in a different way. **A** Controlling for sample size using unpooled data – *x*-axis: bare unpooled effect size; *y*-axis: controlled unpooled effect size **B** Controlling for sample size using pooled data – *x*-axis: bare pooled effect size; *y*-axis: controlled pooled effect size **C** Using pooled versus unpooled effect sizes without controlling for sample size – *x*-axis: bare unpooled effect size; *y*-axis: bare pooled effect size **D** Using pooled versus unpooled effect sizes and controlling for sample size – *x*-axis: controlled unpooled effect size; *y*-axis: controlled pooled effect size. A single data point of greater than 25 effect sizes was excluded to improve presentation of the results.

The variability between particular scores using different effect size measures are defined further in Tables 18–19, where positive values for the mean indicate that the test named in the column was broadly more conservative, whilst negative numbers for the mean indicate that the test named in the column was broadly less conservative. Where negative numbers are observed among the observed range of outcomes in a cell with a positive mean, this signifies cases where particular outcomes increased in measured effect size despite the application of an overall more conservative method. Up to 0.45

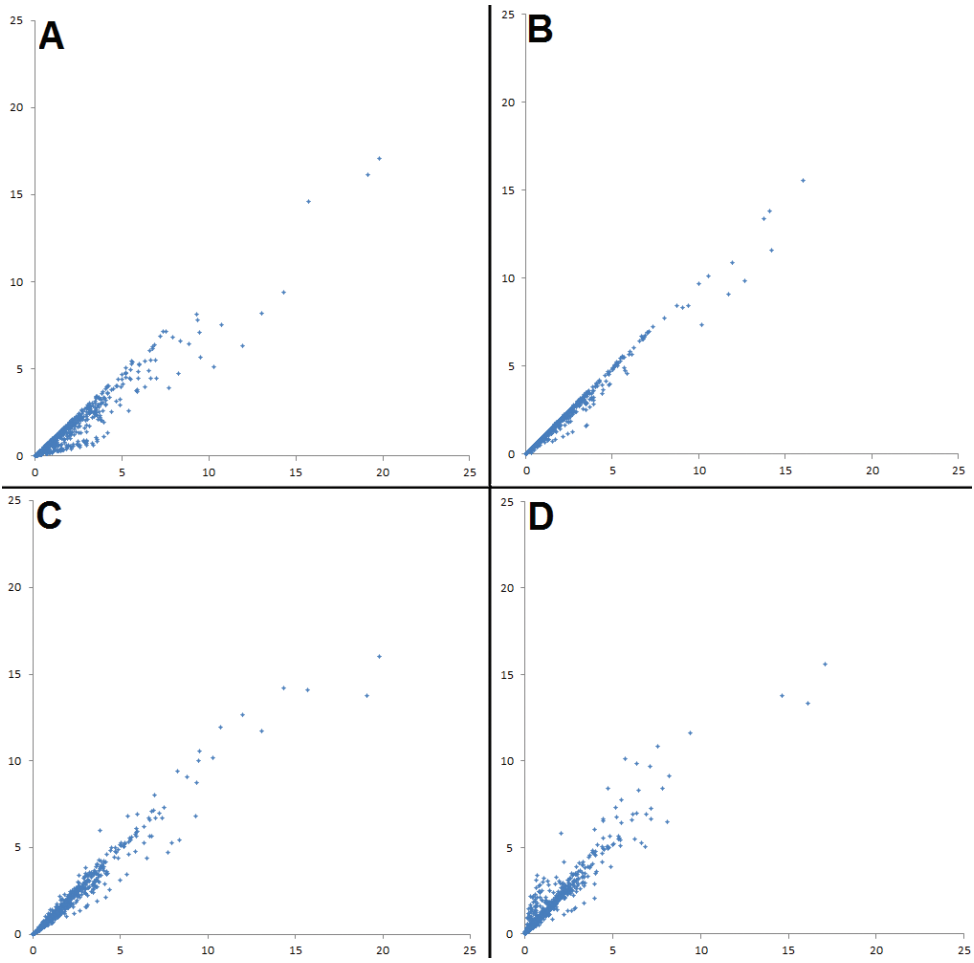

Figure 3. Scatter-graphs showing the effects of applying different corrections of effect size on the entire biometric data set. Scatter-graphs showing the effects of applying different corrections of effect size on the entire vocal data set. Each axis shows effect size, measured in a different way. **A** Controlling for sample size using unpooled data – *x*-axis: bare unpooled effect size; *y*-axis: controlled unpooled effect size **B** Controlling for sample size using pooled data – *x*-axis: bare pooled effect size; *y*-axis: controlled pooled effect size **C** Using pooled versus unpooled effect sizes without controlling for sample size – *x*-axis: bare unpooled effect size; *y*-axis: bare pooled effect size **D** Using pooled versus unpooled effect sizes and controlling for sample size – *x*-axis: controlled unpooled effect size; *y*-axis: controlled pooled effect size.

average magnitude of effect size change can be observed simply by applying a different method to measure effect sizes, which is a figure over double in magnitude that of the minimum effect size limit for scoring in Tobias et al. (2010)'s system. Reductions of up to 24 effect sizes magnitude were observed by controlling for sample size.

The relationship between each measurement of effect size and statistical significance is explored in Tables 20–21. Higher levels of confidence (lower values of *p*) correspond broadly to higher effect sizes in each case. However, the variation in ef-

Table 14. Results of the effects size study for biometrics using Tobias et al. (2010) categories. The top two tables are based upon actual standard deviations for each set of data subjected to pairwise comparison. The lower two tables are based on pooled standard deviation data. In each case, "bare" effect sizes are shown first (above). The second and fourth tables use "controlled effect sizes" for the relevant pooling approach, calculated by taking into account *t*-distribution values for the relevant sample size (or pooled sample size).

Bare Unpooled Effect Sizes (Biometrics)					
Taxon	**0–0.2**	**0.2–2**	**2–5**	**5–10**	**10+**
Adelomyia	1	3			
Myrmeciza	5	53	27	2	
Grallaricula	24	146	70	33	8
Scytalopus 1	7	60	40	2	
Scytalopus 2	0	2	2	1	
Anisognathus	22	132	28	4	
Basileuterus	19	97	32	2	
TOTAL	**78**	**493**	**199**	**44**	**8**
Percentage	*9.5%*	*60.0%*	*24.2%*	*5.4%*	*1.0%*
Controlled Unpooled Effect Sizes (Biometrics)					
Taxon	**0–0.2**	**0.2–2**	**2–5**	**5–10**	**10+**
Adelomyia	1	3			
Myrmeciza	8	65	14		
Grallaricula	29	163	61	25	3
Scytalopus 1	17	67	24	1	
Scytalopus 2	0	3	2		
Anisognathus	24	139	21	2	
Basileuterus	28	99	23		
TOTAL	**107**	**539**	**145**	**28**	**3**
Percentage	*13.0%*	*65.6%*	*17.6%*	*3.4%*	*0.4%*
Bare Pooled Effect Sizes (Biometrics)					
Taxon	**0–0.2**	**0.2–2**	**2–5**	**5–10**	**10+**
Adelomyia	1	3			
Myrmeciza	6	51	28	2	
Grallaricula	26	151	63	31	10
Scytalopus 1	8	59	40	2	
Scytalopus 2	0	2	3		
Anisognathus	23	135	24	4	
Basileuterus	20	97	21	12	
TOTAL	**84**	**498**	**179**	**51**	**10**
Percentage	*10.2%*	*60.6%*	*21.8%*	*6.2%*	*1.2%*
Controlled Pooled Effect Sizes (Biometrics)					
Taxon	**0–0.2**	**0.2–2**	**2–5**	**5–10**	**10+**
Adelomyia	1	3			
Myrmeciza	7	51	28	1	
Grallaricula	30	151	63	31	6
Scytalopus 1	11	64	32	2	
Scytalopus 2	0	2	3		
Anisognathus	23	138	22	3	
Basileuterus	21	107	22		
TOTAL	**93**	**516**	**170**	**37**	**6**
Percentage	*11.3%*	*62.8%*	*20.7%*	*4.5%*	*0.7%*

Table 15. Changes in Tobias et al. (2010) category resulting from increasingly more conservative tests of effect size being applied. This table is based upon changes between the categories in Tables 13–14.

Voice: changes into effect size category	**0–0.2**	**0.2–2**	**2–5**	**5–10**	**10+**
Bare Unpooled –> Bare Pooled	+11	+52	-20	-47	+4
Bare Pooled –> Controlled Pooled	+4	+11	-5	+4	-14
Controlled Pooled –> Controlled Unpooled	+10	+6	-11	+15	-20
Total change from Bare Unpooled –> Controlled Unpooled	**+25**	**+69**	**-36**	**-28**	**-30**
As percentage of total	*1.1%*	*2.9%*	*-1.5%*	*-1.2%*	*-1.3%*
Biometrics: changes into effect size category	**0–0.2**	**0.2–2**	**2–5**	**5–10**	**10+**
Bare Unpooled –> Bare Pooled	+6	+5	-20	+7	+2
Bare Pooled –> Controlled Pooled	+9	+18	-9	-14	-4
Controlled Pooled –> Controlled Unpooled	+14	+23	-25	-9	-3
Total change from Bare Unpooled –> Controlled Unpooled	**+29**	**+46**	**-54**	**-16**	**-5**
As percentage of total	*3.5%*	*5.6%*	*-6.6%*	*-1.9%*	*-0.6%*

Table 16. Spearman's rank correlation coefficients as between the results of five statistical tests carried out on the vocal data set. Confidence for the values stated were given in PAST as zero or less than $p<1x10^{100}$ for all tests.

Tests conducted on vocal data set	**Bare Unpooled Effect Sizes**	**Controlled Unpooled Effect Sizes**	**Bare Pooled Effect Sizes**	**Controlled Pooled Effect Sizes**
Statistical significance (student's *t*)	-0.863	-0.910	-0.849	-0.859
Bare Unpooled Effect Sizes		0.987	0.988	0.908
Controlled Unpooled Effect Sizes			0.894	0.980
Bare Pooled Effect Sizes				0.999

Table 17. Spearman's rank correlation coefficients as between the results of five statistical tests carried out on the biometric data set. Confidence for the values stated were given in PAST as zero or less than $p<1\times10^{100}$ for all tests.

Tests conducted on biometric data set	**Bare Unpooled Effect Sizes**	**Controlled Unpooled Effect Sizes**	**Bare Pooled Effect Sizes**	**Controlled Pooled Effect Sizes**
Statistical significance (student's *t*)	-0.851	-0.951	-0.831	-0.857
Bare Unpooled Effect Sizes		0.936	0.990	0.983
Controlled Unpooled Effect Sizes			0.920	0.936
Bare Pooled Effect Sizes				0.994

fect size scores within each category of significance was high: some effect sizes of up to 18 were non-significant, whilst some effect sizes as low as 0.36 were significant. All scores for effect sizes falling in statistically significant categories exceeded the 0.2 limit for scoring suggested by Tobias et al. (2010), whilst many non-significant effect sizes were in excess of 0.2.

Table 18. Changes in actual effect sizes resulting from changes between different methods of measuring effect sizes in the format actual mean ± standard deviation (minimum–maximum) [mean using absolute values ± standard deviation using absolute values], for voice. The figure in each cell demonstrates the outcomes of subtracting the effect size in the columns from the effect sizes in the rows, for each data point studied.

Tests conducted on vocal data set	Controlled Unpooled Effect Sizes	Bare Pooled Effect Sizes	Controlled Pooled Effect Sizes
Bare Unpooled Effect Sizes	0.35 ± 1.15 (-0.20–24.26) [0.35 ± 1.15]	0.15 ± 0.91 (-11.73–7.13) [0.36 ± 0.84]	0.24 ± 1.07 (-11.32–19.28) [0.41 ± 1.02]
Controlled Unpooled Effect Sizes		-0.19 ± 1.55 (-20.70–4.40) [0.45 ± 1.50]	-0.11 ± 1.36 (-20.26–5.63) [0.40 ± 1.31]
Bare Pooled Effect Sizes			0.08 ± 0.46 (-0.27–15.40) [0.09 ± 0.46]

Table 19. Changes in actual effect sizes resulting from changes between different methods of measuring effect sizes in the format actual mean ± standard deviation (minimum–maximum) [mean using absolute values ± standard deviation using absolute values] for biometrics. The figure in each cell demonstrates the outcomes of subtracting the effect size in the columns from the effect sizes in the rows, for each data point studied.

Tests conducted on biometric data set	Controlled Unpooled Effect Sizes	Bare Pooled Effect Sizes	Controlled Pooled Effect Sizes
Bare Unpooled Effect Sizes	0.41 ± 0.68 (-0.01–5.61) [0.41 ± 0.68]	0.09 ± 0.44 (-2.17–5.34) [0.19 ± 0.40]	0.21 ± 0.53 (-1.97–5.77) [0.26 ± 0.51]
Controlled Unpooled Effect Sizes		-0.32 ± 0.73 (-6.31–2.36) [0.39 ± 0.69]	-0.20 ± 0.57 (-4.4–2.78) [0.30 ± 0.53]
Bare Pooled Effect Sizes			0.12 ± 0.28 (-0.01–2.84) [0.12 ± 0.28]

Table 20. Effect sizes under the four models studied here, grouped into the three "zones" of statistical significance illustrated in Figures 4–5, for vocal data, in the format: mean ± standard deviation (minimum–maximum). $p>0.05$ refers to non-significant results and corresponds to the red rhombuses in Figures 4–5. $0.5/n_v<p<0.5$ refers to possibly significant results which are excluded after applying Bonferroni correction and corresponds to the yellow squares in Figures 4–5. $p<0.05/n_v$ refers to statistically significant results and corresponds to the green triangles in Figures 4–5.

Statistical significance	Bare unpooled effect sizes	Bare pooled effect sizes	Controlled pooled effect sizes	Controlled unpooled effect sizes
$p>0.05$	0.60 ± 1.31 (0.00–11.56)	0.38 ± 0.33 (0.00–2.21)	0.71 ± 2.13 (0.00–22.92)	0.67 ± 2.02 (0.00–22.47)
$0.05/n_v<p<0.05$	1.82 ± 2.69 (0.33–18.22)	1.29 ± 1.24 (0.33–8.47)	1.82 ± 3.00 (0.26–21.60)	1.67 ± 2.61 (0.26–20.14)
$p<0.05/n_v$	3.69 ± 3.30 (0.36–45.33)	3.32 ± 2.73 (0.37–21.07)	3.32 ± 3.05 (0.37–41.45)	3.22 ± 2.81 (0.37–26.05)

Table 21. Effect sizes under the four models studied here, grouped into the three "zones" of statistical significance illustrated in Figures 4–5, for biometric data, in the format: mean ± standard deviation (minimum–maximum). $p>0.05$ refers to non-significant results and corresponds to the red rhombuses in Figures 4–5. $0.5/n_v<p<0.5$ refers to possibly significant results which are excluded after applying Bonferroni correction and corresponds to the yellow squares in Figures 4–5. $p<0.05/n_v$ refers to statistically significant results and corresponds to the green triangles in Figures 4–5.

Statistical significance	Bare unpooled effect sizes	Bare pooled effect sizes	Controlled pooled effect sizes	Controlled unpooled effect sizes
$p>0.05$	0.72 ± 0.70 (0.00–3.97)	0.42 ± 0.29 (0.00–1.95)	0.71 ± 0.72 (0.00–3.92)	0.63 ± 0.60 (0.00–3.41)
$0.05/n_v<p<0.05$	1.45 ± 0.86 (0.44–3.98)	1.06 ± 0.41 (0.44–2.44)	1.37 ± 0.96 (0.40–6.01)	1.26 ± 0.84 (0.40–5.81)
$p<0.05/n_v$	3.37 ± 2.64 (0.61–19.80)	2.79 ± 2.07 (0.61–17.09)	3.15 ± 2.44 (0.58–16.04)	2.97 ± 2.23 (0.58–15.57)

Conclusions and discussion

The dataset studied here exhibits comparable overall levels of variation to Tobias et al. (2010)'s data set. The latter was developed using sympatric species pairs on a global basis. However, this data set involves comparisons of many populations that are currently recognized as subspecies and several of which are unnamed (Table 1). Here, 55.7% and 60.6% scored in the *minor* (score 1) category for voice and biometrics respectively (using bare pooled standard deviations), versus 58% and 63% for the Tobias et al. (2010) data set. At a score of 2 (2–5 effect sizes), this study produced scored 23.6% and 21.8% of the sample for voice and biometrics versus Tobias et al. (2010)'s 26% and 24%. This similar set of outcomes means that comparisons between outcomes of other tests are likely to be a reasonable proxy for how Tobias et al. (2010)'s database would perform, under other tests.

Several broader aspects of the results can be explained by considering the number of standard deviations' difference required to satisfy various models (Figure 6). The lack of a control for sample size using *t*-distributions in the Tobias et al. (2010) effect size calculation resulted in a liberal approach, which may over-score differentiation at low sample sizes. However, at the 5 effect sizes level (3 points), their model low-scores differentiation for sample sizes of greater than 7 (compared to using a controlled unpooled effect size of 4) (Fig. 6). The 1-point test of Tobias et al. (2010) at 0.2 effect sizes is very liberal indeed, set at almost half the lowest recorded effect size measurement in this study that was statistically significant (0.36 effect sizes: Table 20). This inconsistency in treatment of outcomes showing very low levels of variation explains the large number of "1" scores in the Tobias et al. (2010) analysis, compared to the much smaller number of pairwise comparisons achieving Level 1 or greater variation under the Levels study.

The overall lower differentiation levels in biometrics can in part be explained due to lower sample size (see Tables 1–2) but likely also reflects lower variability of these kinds of variables.

Pooled versus unpooled and bare versus controlled effect sizes

The outcomes of using pooled versus unpooled and bare versus controlled effect sizes are substantial across the data set as a whole and can be drastic in individual cases (Tables 18–19).

The distinction between using pooled versus unpooled standard deviations in taxonomy has passed by barely without discussion in ornithological taxonomic literature. Isler et al. (1998)'s test applies *t*-distribution data on an unpooled basis to two data sets under comparison, but Tobias et al. (2010) and Hubbs and Perlmutter (1942) used effect sizes based on Cohen's *d* statistic, which calls for a pooled standard deviation without controlling for sample size using *t*; neither commented on their selection. Nakagawa and Cuthill (2007) recommended presenting confidence interval data and Tobias et al. (2010) refer to this, but it is unclear how this was built into their framework nor whether any cut-off based on low confidence intervals was applied.

Usage of pooled standard deviations, as a matter of statistical methodology, should only be undertaken where the standard deviations of the two populations under comparison can be assumed to be equal. This does not necessarily mean that measured standard deviations of the two populations must be equal, or even close to one another, since these will usually differ for two measured populations as a result of the sampling. However, it must be reasonable to make this assumption in order to apply this method. The pooling formula attributes greater weight to the standard deviation of the population with higher sample size and produces a "weighted average" standard deviation which is closer to that of the population of which there is a larger sample. Degrees of freedom for the pooled standard deviation are greater due to summing those of the two separate populations. In practice, in taxonomy, we will usually have no idea as to whether or not the standard deviations of two populations under comparison are equal or not. Special care should be adopted in using pooled standard deviations where estimated population sizes, molecular or geographical attributes of the two populations vary greatly. For example, comparing an isolated, very small montane population with low intra-population molecular variation versus a very widespread lowland population which is known to exhibit substantial clinal variation and has higher intra-population molecular variation would be inappropriate, since assumptions underlying the usage of pooled standard deviations are likely not just to be unknown but incorrect. The greater correlation between statistical significance and controlled unpooled effect sizes (Tables 16–17) is also noteworthy. In summary, the unpooled / Isler et al. (1998) model, which does not make unnecessary assumptions and is overall more conservative (Tables 10–15), is methodologically more supportable among the four methods for measuring effect size analyzed here.

There are still likely to be "use cases" for pooled standard deviations to measure effect sizes in taxonomy. Salaman et al. (2009) compared two populations, one being undescribed and probably extinct, known only from a single specimen. With d.f. = 0, unpooled effect size calculations produce "divide by zero" errors. In that publication, a modified version of the Isler et al. (1998) formula was developed as an indicator

of diagnosis, using the standard deviation of the better-known population for both populations. In other situations, particularly those involving very small sample sizes, appropriate usage of pooled standard deviations could be considered. A particular risk when studying smaller populations with unpooled standard deviations is that sampled data points may "cluster", resulting in small recorded standard deviations, which could exaggerate measured differentiation. Usage of pooled standard deviations can be a hedge against such outcomes, even if the underlying assumptions for using pooling are not met. However, using *t*-distributions also provides such a hedge and moreover involves a statistical test specifically designed to cater for the risk of clustering. Use case scenarios for pooled effect sizes in taxonomy are unlikely to be the norm and may not in any event justify adopting *t*-distribution-based corrections using inflated degrees of freedom.

Statistical significance

Most papers concerning the application of statistical tests for determining the taxonomic rank of allopatric populations have noted that statistical significance is not a good measure, due to its potential for liberal satisfaction by increasing sample size, its failure to indicate higher levels of differentiation or false positives when sampling from different parts of a geographical cline; and then move quickly on to discuss better tests (Patten and Unitt 2002, Nakagawa and Cuthill 2007, Remsen 2010, Tobias et al. 2010). For example, Tobias et al. (2010) noted that: "The fact that it is easy to achieve statistically significant differences merely by increasing sample size may lead to inappropriate taxonomic decisions." Bizarrely then, many modern taxonomic papers, including in some of the field's most prestigious journals, erroneously claim "diagnosis" on the basis of overlapping data sets that are presented as satisfying tests of statistical significance, such as *t*-test, Tukey-Kramer, Wilks' Lamda, ANOVA or Kruskall-Wallis tests (e.g., Benkman et al. 2009, Lara et al. 2012, Freitas et al. 2012). With this background, it is worth dwelling more on the usage of statistical significance.

Although large samples sizes are cited as the basis to reject statistical significance as a useful measure in taxonomy, such problems are rarely faced by taxonomists. Having too few specimens (whether in museums or measured in the field) or sound recordings is likely a more material problem. For new species descriptions in birds published between 1935–2009, 332 of 477 (70%) were based on 0–5 specimens, only one was based on >100 specimens and the mean number of specimens was 6 (Sangster and Luksenburg 2015). One approach to avoid false positives would be to introduce minimum criteria for sample size. Tobias et al. (2010) called, where possible, for sample sizes of at least 10. Walsh (2000) proposed that taxonomic studies using discrete data should have sample sizes in excess of 50 to exclude the possibility of polymorphisms occurring at $p < 0.05$ that cause incorrect interpretations. However, minimum limits such as these would be blunt and arbitrary and could prejudice against taxonomic recognition of highly distinct but very rare populations where only small samples are

available (see also Lim et al. 2012 and Sangster and Luksenburg 2015). With continuous data, we can instead apply *t*-distribution corrections to address such concerns.

The classic test of statistical significance between two populations of data is the Student's *t*-test. This evaluates the probability of whether two normally-distributed data sets relate to two different populations, by considering whether or not their mean averages are likely to differ from one another. Various other similar tests can assess differences between mean, median, or modal averages, such as *F*, Mann-Whitney *U*, Kolmorov-Smirnov, Wilks' Lamda, ANOVA, Kruskal-Wallis, and Tukey-Kramer. Some of these tests are better suited to continuous variables which are non-normally distributed, such as ratios or products of raw data.

Although the *t*-test will evaluate the likelihood that two populations are different, it tells us little about the extent of differences between the two populations. With a large enough sample, the two sample means may be very close to one another. Here, the lowest distance between statistically significant outcomes was 0.36 effect sizes. Tests of statistical significance can also be failed on data showing effects sizes as high as 18 (Table 20), where sample sizes are small. An example of data with close means passing a test is the following, based on a large sample with almost complete overlap of variables:

Donegan (2012, Appendix 3A and Appendix 4): Maximum acoustic frequency of last note of male song (kHz). Data are in the form *average ± SD (lower bound – upper bound) (n = sample size)*:

Myrmeciza melanoceps: *2.42 ± 0.12 (1.98–2.62) (n = 143)*
Myrmeciza goeldii: *2.29 ± 0.08 (2.01–2.49) (n = 173)*

The *t*-test was passed at $p<0.0002$, yet these data reveal small differences between means and substantial overlaps in recorded values. The *t*-test result suggests that the two populations in question have begun to diverge from one another, which is interesting and makes it valid to discuss their relationship and possible isolation mechanisms. However, identification of a sound recording to one or the other species on the basis of these data would be impossible. The effect size here was 1.31, considerably in excess of the lowest (0.36) score, but fewer than 50% of individuals could be identified based on this variable and it would be useless for identification. Regularly, diagnosis is incorrectly asserted in the taxonomic literature based on data like those in this example (see citations above).

The *t*-test and similar tests demonstrate *statistical significance* of differences between means. Such differences may have some evolutionary significance. However, a positive *t*-test is not necessarily of much *taxonomic significance* (Fig. 1): we must also consider how much the two populations have differentiated.

In the field of medicine, the outcome of tests of statistical significance is widely understood and accepted to be just a first phase in demonstrating an interesting result. A variety of different approaches exist in medical science which must also be passed to show *clinical significance*, which, for example, would support the usage of drugs. In any

taxonomic study, it is similarly important to move on from the ecology class, beyond *statistical significance* to consider the *taxonomic significance* of any results.

That all said, statistical significance can be a tougher one than some proposed measures of differentiation. Instances were found here of pairwise comparisons passing Isler et al. (1998)'s gold standard of diagnosability but failing tests of statistical significance (Table 6: 15 outcomes, or 0.6%, in categories 2, 3, 4 & 5 or 2, 4 & 5). Instances of differentiation being scored under the Tobias et al. (2010) system in statistically insignificant situations were widespread (Figure 4). Under the Tobias et al. (2010) system, up to 87–91% of outcomes studied here scored 1 or more points (Tables 13–14) but on the same data set 49.5% (voice) to 64.6% (biometrics) of outcomes failed tests of statistical significance (Tables 8–9), suggesting false attribution of scores to non-significant situations under this system in at least 36% of cases (see Figure 4). This present study therefore highlights the importance of considering both significance and effects-based differentiation in taxonomy. Demonstrating that the means of two populations have actually diverged at all (using statistical significance) should be a baseline requirement for any assertion of diagnosis between two populations, an omission in both Isler et al. (1998)'s and Tobias et al. (2010)'s systems. To avoid "false positive" assessments of diagnosability, the *t*-test or another test of statistical significance between means should be introduced as a gateway and *additional* requirement to tests of diagnosis. It also flows from this study (Figure 4, Tables 20–21) that an effect size of 0.2 cannot be supported as a basis for attributing taxonomic significance when using continuous variables.

Type 1 error corrections

Introducing the Dunn-Šidák correction (as opposed to the simpler but overly conservative Bonferroni correction) had a virtually negligible effect (Tables 4–5). As a result, for taxonomic data sets involving similar numbers of variables to those addressed here (fewer than 30), it will probably not be worth the trouble of applying more complex corrections.

Bonferroni (and Dunn-Šidák) corrections are appropriately applied to "families" of variables. A middle-ground of treating voice and biometrics as separate "families" is recommended based on this study. This could be criticized, since certain aspects of voice and biometrics can be linked (e.g., Podos et al. 2004, Phillips and Derryberry 2017). However, Bonferroni is inherently conservative: for example, biometric measures are likely to be correlated with one another and so may not be truly independent: for example, longer-winged birds might be more likely to have longer tails as well, due to environmental or hereditary conditions affecting feather growth or size. The proposed liberalism here of treating voice separately from biometrics counterbalances the over-compensation inherent in Bonferroni between data sets in variables that may show some correlations. Treating these variable families separately can be justified since many taxonomic studies (including some included here) address only one or the other

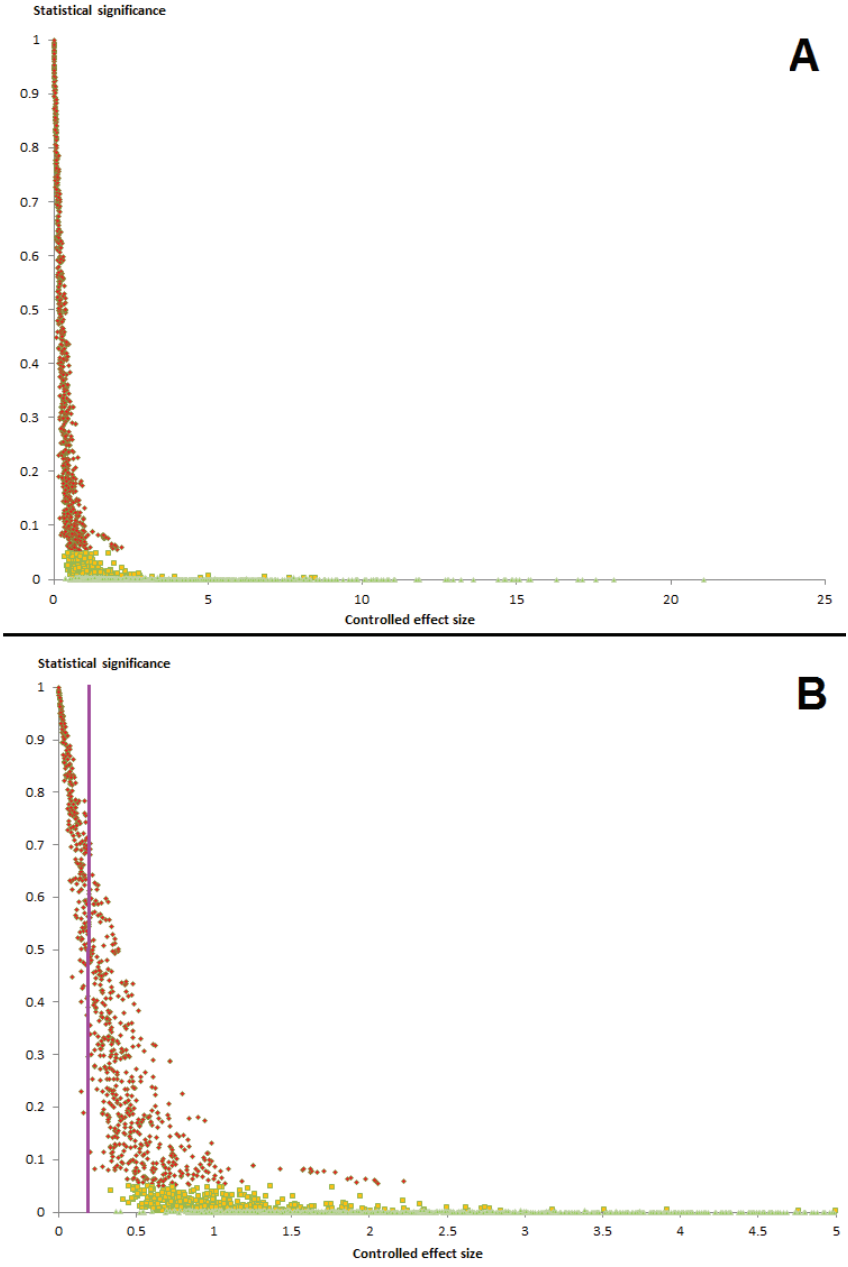

Figure 4. Scatter-graphs of controlled unpooled effect size (*x*-axis) versus statistical significance (p<x) (*y*-axis). All outcomes to the right of the two purple lines represent pairwise comparisons which are given scores of at least 1 under the Tobias et al. (2010) system. Note the statistically insignificant outcomes (red rhombuses, to the right of the purple line) which are given scores of 1 under the Tobias et al. (2010) system and the small numbers of such outcomes scoring 2 (two effect sizes or more) and 3 (five effect sizes or more). Figure 4A, B use all 2348 pairwise comparisons for voice. Figure 4C, D show all 822 pairwise comparisons for biometrics. Figure 4B, D are close-ups of Figure 4A, C respectively, showing the area below 5 effect sizes.

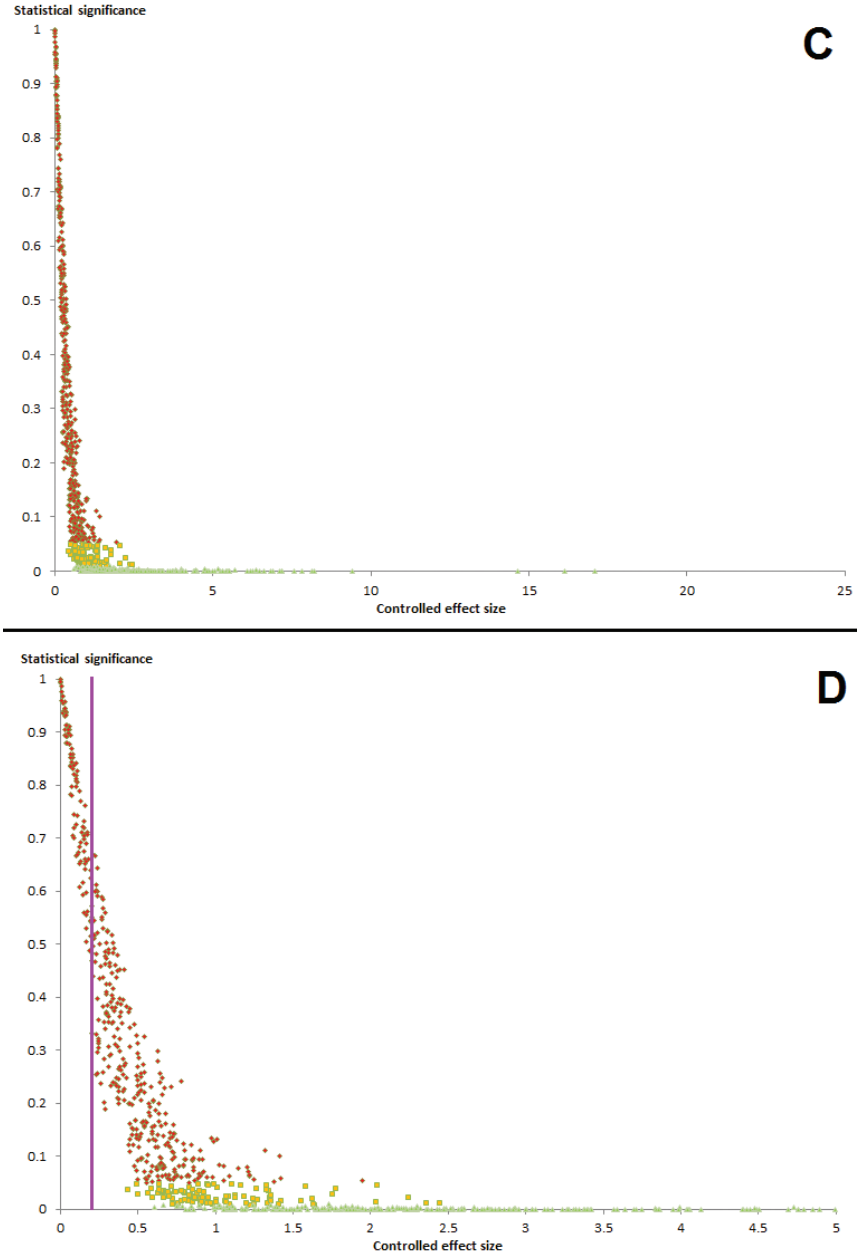

Figure 4. Continued: Red rhombuses represent a lack of statistical significance. Yellow squares represent possible statistical significance, less than $p<0.05$ but failing tests of significance after applying Bonferroni correction. Green triangles, which are almost continuous along the *x*-axis, represent statistically significant outcomes. The lower two graphs show greater resolution of the same graphs at effect sizes of up to 5 (where 4 signifies full diagnosis). Note: the graphical representation is generous to Tobias et al. (2010) in applying the most conservative effect size definition from this study on the *x*-axis; using bare pooled effect sizes as applied by these actions would generally shift outcomes to the right, as illustrated in Figures 2–3.

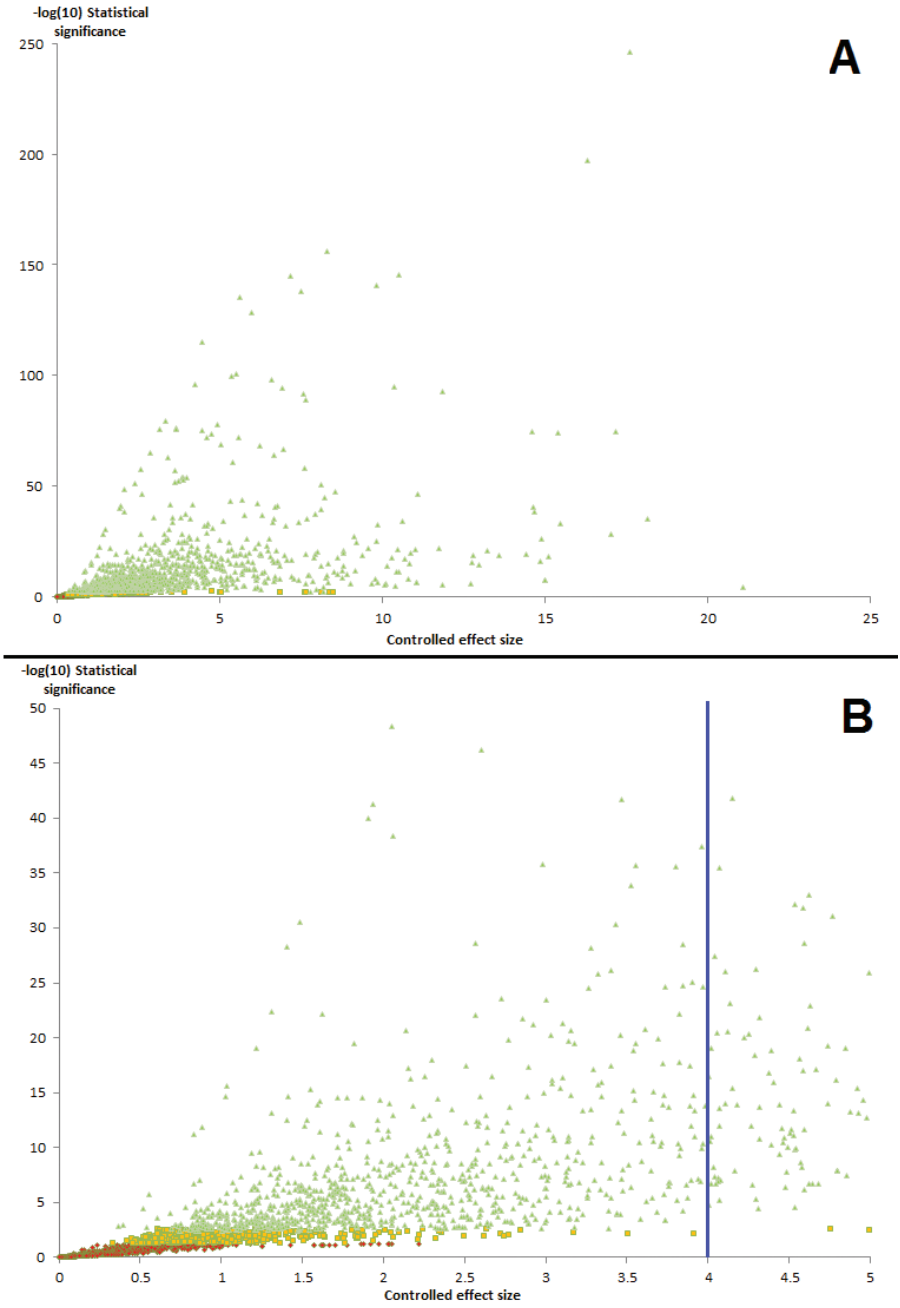

Figure 5. Logarithmic plot of the same data as in Figure 4, showing controlled unpooled effect size (*x*-axis) versus a logarithm of statistical significance (p<x) (*y*-axis). Note the number of instances of statistically significant variation ignored under the Isler et al. (1998) approach (green triangles, to the left of the purple line). Figures 5A, B use all 2348 pairwise comparisons for voice. Figure 5C, D show all 822 pairwise comparisons for biometrics. Figures 5B, D are close-ups of Figures 5A, C respectively, showing the area below 5 effect sizes.

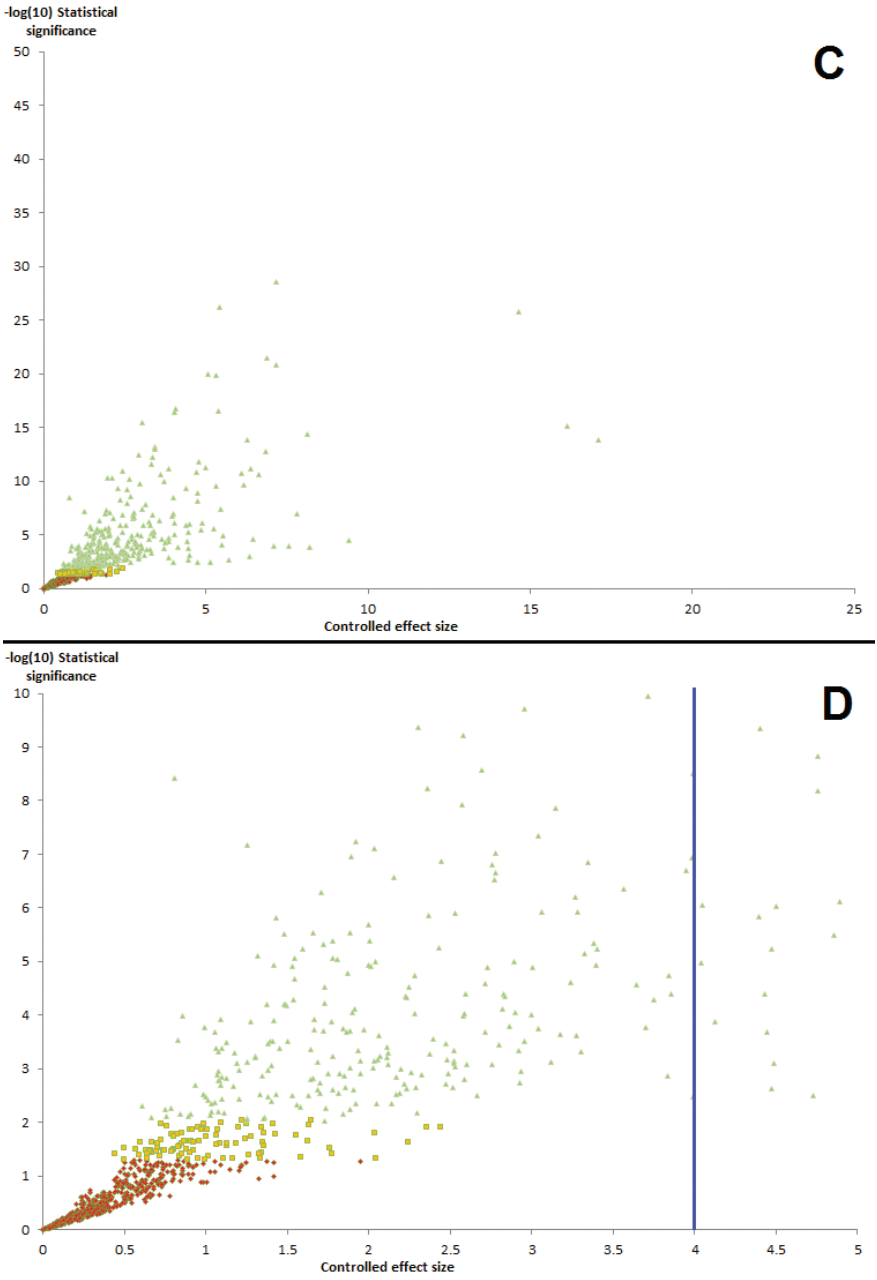

Figure 5. Continued: Red rhombuses, which are almost continuous along the *x*-axis on the left hand side of the two upper graphs, represent failure of tests of statistical significance. Yellow squares represent possible statistical significance, less than $p<0.05$ but failing tests of significance after applying Bonferroni correction. Green triangles represent statistically significant outcomes. Those outcomes to the right of the purple line are given credit as diagnosable under the Isler et al. (1998) test of species rank. Note also in the upper graphs the handful of yellow squares at effect sizes of greater than 4, which represent statistically insignificant outcomes which are nonetheless given credit under the Isler et al. (1998) model.

kind of variables. Applying statistical corrections based on the full number of variables studied makes results less comparable and penalizes more holistic studies. In contrast, it would seem less justifiable to apply a more liberal standard to, say, one type of call versus another type of call. It was found when re-analyzing these data that such an approach resulted in a more liberal approach to "less variable-rich" vocalizations, such as single-note calls. This seems inappropriate given that such calls are likely less relevant to mate choice than male songs, which tend to be more complex and require more variables to properly analyze. (This proposed family-treatment of variables based on all vocalization-types for purposes of Bonferroni correction presents a change from the methodology underlying several of the studies listed in Table 1.)

Diagnosability

Diagnosability was considered to be the most frequently applied criterion to assess rank in a review of over 1000 taxonomic revisions (Sangster 2014). It is an important concept, not least because the International Code of Zoological Nomenclature (ICZN 1999) recommends that: "When describing a new nominal taxon, an author should make clear his or her purpose to differentiate the taxon by including with it a diagnosis, that is to say, a summary of the characters that differentiate the new nominal taxon from related or similar taxa" (Recommendation 13A). Under Article 13.1.1 of the Code, a newly described species name is not "available" unless it is "accompanied by a description or definition that states in words characters that are purported to differentiate the taxon" (or includes reference to a text which does so). Moreover, diagnosability allows for identification, which enables users of names to label populations.

Dubois (2017) has provided some interesting insights into diagnosability as a concept in taxonomy and nomenclature. However, he did not directly address its measurement and statistical evaluation using data sets based on continuous variables, which are the focus of this paper. The two models analyzed here that address diagnosability are: (i) that of Isler et al. (1998), which is based on controlled unpooled effect sizes of 4; and (ii) that of Tobias et al. (2010) which attributes a score of 3 to situations where the means differ by 5 bare pooled effect sizes. There are clear advantages to controlling for sample size in assessing whether any overlap exists, in that diagnosability tests would then not bias against studies using different sample sizes. The Isler et al. (1998) test also benefits from a statistical and conceptual purity in measuring a meaningful mathematical and statistical situation of diagnosis to 95%, compared to an arbitrary number of 5 effect sizes which often exceeds what is needed to demonstrate statistical diagnosis.

In this study, 14.5% of the vocal data set and 6.3% of the biometric data set passed the Isler et al. (1998) diagnosability test ("Level 5", Tables 8–9). Tobias et al. (2010)'s equivalent test of "5 effect sizes" (including scores of over 10) was met by fewer population/variable comparisons for voice, but more comparisons for biometrics: 10.1% and 7.4% respectively (Tables 13–14, data for bare pooled standard deviations). It is expected that an effect size of 5 would result in fewer positive outcomes than an ef-

fect size of 4. The larger number of positive outcomes for biometrics at 5 effect sizes is due to using pooled standard deviations and consequent impact of controlling for sample sizes using lower values of *t*. A rationale for excluding the 4.4% of tests in the vocal sample which met a 4 standard deviations standard (controlled for sample size) but failed a test of 5 bare pooled standard deviations as diagnosably distinct is elusive. It would be prudent for the 3 point score of Tobias et al. (2010) to be recast so as to align to the Isler et al. (1998) test (as modified here), as was done in Donegan and Avendaño (2014). Tobias et al. (2010) adopted a model which is a conservative proxy for diagnosis and simple to calculate. However, whilst overly liberal in the trigger for assigning 1 point, they were unnecessarily conservative in setting a higher trigger for assigning 3 points.

Differentiation below diagnosability, subspecies, and the 50% and 75% tests

There is no consensus as to whether any differentiation below diagnosability for particular characters ought to be recognized in taxonomy. Remsen (2010) essentially rejected this proposition, requiring a valid subspecies to show 95% diagnosability in "one or more phenotypic traits" and "not to multiple simultaneous comparisons". This can be equated to a single character meeting the Isler et al. (1998) standard of >4 controlled standard deviations in a single character. There are however conceptual difficulties with models that depend on diagnosability of a single character. Two taxa may show considerable overlap for two or more variables if these characters are analyzed separately, but may be diagnosable in multivariate space. Quantitative criteria based on a series of univariate datasets are liable to overlook significant diagnosis below full differentiation and could result in taxonomic non-recognition of diagnosable populations. Elucidating diagnosable characters is important to satisfy the requirements of the ICZN code and for identification, but any morphological or vocal character identified by taxonomists is simply a measurable slice of multivariate character space differentiating two populations.

Originally, 75% diagnosis tests for subspecies used one of two approaches: (i) a 75%/75% test (i.e., 75% of population 1 is diagnosable from 75% of population 2; or (ii) a 75%/99+% test (i.e., 75% of population 1 is diagnosable from essentially all of population 2). Amadon (1949) opined that any measure below 75%/99%+ diagnosability "does not seem to set a high enough standard". Patten and Unitt (2002) also reaffirmed use of this measure, but added to Hubbs and Perlmutter (1942) and Amadon (1947)'s framework by controlling for sample size using *t*-distributions, based on a similar method of adjustment to that of Isler et al. (1998). They noted that, where this 75% method is applied controlling for sample sizes, it achieves outcomes close to those obtained for full diagnosability, as also commented by Remsen (2010). Hubbs & Perlmutter (1942)'s conceptual alternative of a 50%/100% test was effectively adopted in part as a benchmark by Tobias et al. (2010) by giving two points in their scoring system to populations which are over 2 bare pooled standard deviations apart.

The application of sharp, seemingly arbitrary, tests such as these to classify normally distributed data into segments to which scores are attributed is a situation not unique to taxonomy. Similar hard boundaries are also rife in most education and examination systems. In UK universities, a student scoring 60.1% or 69.9% in an examination will be given the same award (an upper second class degree) but a student attaining 70.1% will get a different award, a first class degree. This is despite the students scoring 69.9% and 70.1% having attained more similar levels of achievement to one another. Whilst any cut-off may be criticized as arbitrarily generous or harsh to outcomes falling close to the line on either side, the application of cut-offs is something that humans tend to do in their quest to categorize things. Where cut-offs are applied, a test of whether the cut-off is a valid one should best be based upon: (a) differentiation of a meaningful number of outcomes; and (b) the setting of boundaries at statistically-, mathematically-, or biologically-meaningful positions.

In this data set, with very large numbers of pairwise comparisons, necessarily many individual cases fall very close to each of the cut-off boundaries proposed by previous models for attributing taxonomic significance, whether at or below diagnosability. Two populations differing by 95% using Isler et al. (1998)'s diagnosis test will be given credit for that character, but two populations differing by 94.9% will not. However, this seemingly arbitrary distinction is supportable because 95% is the standard confidence interval used in science. In contrast, the 75% test of subspecies flunks both requirements for a supportable test of differentiation. Depending on the sample size, the 99/75% concept equates to an average diagnosability of over 90%+ per population (Patten and Unitt 2002, Remsen 2010), and so gets close to a species test. Because *t* at 99% can exceed *t* at 97.5% with very small samples, it is even sometimes the case that *t* at 99% + *t* at 75% exceeds 2*t* at 97.5%, explaining the 2.3% of outcomes in Table 6 of pairwise comparisons passing the Level 5 test of 95% diagnosis but counterintuitively failing the Level 3 test of 75% diagnosis (Levels 1245 or 125 therein). In this study, the 75% category was narrow to the extent of being almost negligible for both vocal (2.2% of sample) and biometric (1.2%) data (Tables 8–9). Neither does 75% differentiate any biologically meaningful or statistically meaningful delineation of which I am aware. It merely provides a marginally more liberal subspecies test than that proposed by Remsen et al. (2010) for most sample sizes. This 75% test should be abandoned altogether, as was also proposed by Remsen (2010).

In contrast to the 75% (Level 3) test, 50%/95% differentiation (Level 2) measures a mathematically relevant point of differentiation, when the mean of one population moves outside the normal distribution of the other. It also signifies the point at which a population has moved half way towards diagnosability. The number of pairwise comparisons meeting the Level 2 test (but not falling in other buckets) was material but not enormous. Only 30.6% for voice and 20.6% for biometrics of outcomes passed this test at all (Tables 8–9). Its outcomes compare to Tobias et al. (2010)'s scoring of over 70% of outcomes and the number of outcomes meeting tests of statistical significance. Only 13.9% and 12.9% of the sample respectively, for voice and biometrics, fell into a category where the 50% Level 2 test (but not others) were met, totals which rise to

16.1% and 14.1% respectively if the results of the (broadly useless) 75% category are aggregated. Once sample size is controlled for using *t*-distributions, 50% differentiation arguably becomes the closest defendable proxy to the traditional 75% test of subspecies on which most avian taxonomy was built, if one takes into account the example studies of Hubbs and Perlmutter (1942) and Amadon (1949) and the sample sizes that they used. When controlling for sample size, a 50% test can be quite an exacting standard to pass and gives considerable comfort that material differentiation has taken place. In contrast, adopting a 75%/99%+ diagnosis (near-diagnosability) as a test for subspecies rank would place many currently recognized and largely identifiable geographic variants of birds occurring on different mountain ranges into synonymy, which is undesirable because many of these populations have been shown to merit taxonomic recognition based on studies of plumage or molecular characters. The most extreme example of low differentiation calling for taxonomic recognition in the study here relates to the Yariguíes and northern East Andes population of Speckled Hummingbird *Adelomyia melanogenys*, which reached only up to "Level 2" differentiation in voice, no differentiation in biometrics and are near-diagnosable (but non-diagnosable) in plumage. However, they exhibit *c*.5.8% mtDNA differention (Chaves and Smith 2011). Such populations should arguably be recognized taxonomically, at least as subspecies.

Diagnosis based on actual data

In addition to the diagnosis formula for Level 5, Isler et al. (1998) require satisfaction of the Level 4 test of non-overlap. Although such considerations can help identify situations requiring further investigation, the Level 4 test biases towards positive outcomes in studies using small samples (Fig. 6). For pairwise comparisons which narrowly meet Isler et al. (1998)'s 95% diagnosis test, it would be expected for 5 sampled measures out of 100 actually to overlap. For such a data set, satisfaction of the Level 4 test could be random in that a $p<0.05$ result might arise at any point of data collection between $n = 1$ and $n = 100$, rendering Level 4 unsatisfied and denying credit for observed differentiation. The likelihood of an outlier existing in the sample increases on a linear basis with sample size. There were 22 instances in the vocal study (0.9% of outcomes) most of which included populations with $n > 100$ sample size, in which levels 1, 2, 3 & 5 were passed (Table 6), i.e., all statistical tests were passed except actual non-overlap. A rationale for denying significance to these outcomes is elusive, but retaining this criterion penalizes studies using large sample sizes. Usage of this test as a gateway to affording weighting to observed differentiation in taxonomy should be abandoned.

Adapting Tobias et al. (2010)

Tobias et al. (2010)'s standard of 0.2 effect sizes as a starting point for attributing taxonomic significance was probably based upon Cohen (1998)'s original scheme for inter-

preting effect sizes, as embellished by Sawilowsky (2009), in which effect sizes are categorized as "small" above 0.2, "medium" above 0.5, "large" above 0.8, "very large" above 1.2 and "huge" above 2. This study shows the supposedly unusual "huge" category actually to be fairly standard in taxonomy, with 33–37% of vocal comparisons achieving this benchmark (see further the discussion above on 50% differentiation). Isler et al. (1998) and Remsen (2010) only value effect sizes of 4 (double, "huge") and no less. Here, the highest effect size recorded between relevant taxa, all of which were considered congeners at the time of the study, was over 40. Effect sizes of over 10 represented over 3% of the vocal sample. Tobias et al. (2010)'s higher scores also attribute second and third points at 2 ("huge") and 5 (more than double "huge"), such that some acknowledgement of the inappropriateness of traditional effect size interpretations is evident in their system.

Traditional interpretations of effect sizes may be appropriately used in other fields but are inappropriate for taxonomic study. It should be borne in mind that the traditional subjective descriptors for effect sizes starting at 0.2 have been developed largely in the fields of social and behavioral science (Cohen 1998, Sawilowsky 2003). Such fields by definition only consider intra-specific differences (typically, within *Homo sapiens*) and not between-species differences. In taxonomy, we are primarily interested in diagnosability and identification, which look to greater levels of differentiation.

Overall, this study suggests that: (i) Tobias et al. (2010)'s score of 1 for minor differences is set at too liberal a level, attributing taxonomic value to differences which in many situations are of no statistical significance; (ii) their scores of 1, 2, 3, and 4 are all based on bare pooled effect sizes, which involve a certain degree of "hedging" against measured standard deviation error, but do not include a sufficient control for sample size and are based on inappropriate assumptions; (iii) the score of 3 is based on 5 standard deviations' difference, which is an arbitrary value set unnecessarily high when 4 effect sizes equates to diagnosability, and which is overly conservative for any study with a sample size greater than 7 but overly liberal otherwise (Fig. 6); (iv) the score of 4 is based on 10 SDs' difference, a very high standard to meet for any data set, and also set arbitrarily; (v) total measured variation in effect sizes may theoretically vary by a factor of up to 3, for different situations which attain the species benchmark score of 7 points (Table 25); and (vi) the overall scheme results in a homogeneous scoring system where almost every comparison attains 1–2 points and few comparisons get no or higher scores. Several challenges arise in recalibrating the model. First, the only supportable measures below diagnosability studied here are considered those of statistical significance (Level 1) and 50% diagnosability (Level 2), yet the Tobias et al. (2010) system calls for two scores (1 and 2) below the level of diagnosability and a score of 3 which seems intended to approximate to diagnosis. Secondly, the comparison with sympatric "good" species in Tobias et al. (2010) would need to be re-run entirely to check whether the benchmark score of 7 for species rank requires modifying if their model is modified. I propose here two possible approaches which should be explored further to improve the scoring system's application to continuous variables:

Solution A:

1 point: Level 1 statistical significance only.
2 points: Level 1 plus Level 2 50% diagnosability.
3 points: Level 1 plus Level 5 full diagnosability (3 points).
4 points: Level 1 plus a new measure of a "species and a half" worth of diagnosability (equivalent to 6 controlled effect sizes).

Solution B: would use more proportionate scoring, eliminate the weighting for statistical significance and allow only three scores:

1.5 points: Level 1 plus Level 2 (2 controlled effect sizes).
3 points: Level 1 plus Level 5 (4 controlled effect sizes).
4.5 points: Level 1 plus 6 controlled effect sizes.

Solution C: would abandon these various cut-offs and instead use controlled unpooled effect sizes, calibrated by a scale factor such that no difference = 0 and full diagnosability = 3 and capped at a score of 4.

As has been argued elsewhere (e.g., Remsen 2015, Donegan et al. 2015), the Tobias et al. (2010) system should be restricted to situations of allopatry and their positive scorings for hybridization should be removed from the model. Despite the above conclusions, Tobias et al. (2010)'s proposals represent an important step forwards towards an holistic measure of species rank (incorporating plumage, habitat, voice and biometrics data) and so have several notable benefits and important objectives in light of rationality issues affecting modern taxonomy. This holistic approach also has benefits over systems which consider only continuous data. It also seems that Del Hoyo and Collar (2014, 2016) in practice did not attribute vocal and biometric scores to situations of trifling differentiation and, as a result, the recommendations in those works are likely to be more closely aligned to outcomes under the amended basis for attributing scores discussed above.

Amendments to the Isler et al. (1998) method

For the reasons above, the Level 4 non-overlap test should be abandoned from this framework in order to positively score the 2.5% of vocal outcomes which were diagnosable to 95% but actually overlapped due to very large sample sizes (Table 6). This 2.5% of overall outcomes represented 16.4% of those with positive "Level 5" tests. Secondly, statistical significance using a *t*-test should be assessed as a gateway to concluding any positive outcome of diagnosability, in order to avoid counting false positives. These made up 0.6% of overall outcomes including 4% of outcomes with

positive "Level 5" tests. These amendments are lower in their impact compared to those proposed here to the Tobias et al. (2010) system.

A disadvantage of the Isler et al. (1998) method remains its general exclusion from consideration of situations which were statistically significant but non-diagnosable, making for conservative interpretations. However, conservative interpretations are conceptually less supportable than interpretations based on a "best view" of available data, give precedence to history or tradition over rationality, reinforce geographical biases in status quo taxonomies and ultimately misinform biodiversity conservation and other users. I have been particularly frustrated in the past at being required by peer reviewers at the same time to (i) show satisfaction of all species or subspecies concepts in order to describe or recognize a new or synonymized taxon, but (ii) at the same time being asked to disprove satisfaction of all recognized concepts in order to lump a taxon (e.g., see Donegan and Avendaño 2008, 2010). The result of such requirements is the non-description of equally diagnosable taxa as those recognized in current taxonomies or the non-lumping of presently recognized but dubious taxa. Moreover, as illustrated in Figure 5, outcomes from comparisons showing differentiation below diagnosability represented a rich source of potentially useful information, which can be taken into account using other methods developed below.

"Hard cut-offs" in existing models of species rank and their elimination

Isler et al. (1998) and Tobias et al. (2010)'s models both suffer from a common shortcoming. Testing whether two allopatric populations are or are not as differentiated as two sympatric species (Helbig et al. 2002) will mean that, in a large data set, some will just pass and some will just fail. However, a difficulty embedded in existing systems of assessing rank is that they create a series of further examinations, all of which have their own inflection points, in order to come to an answer (Figure 7).

One could adapt the Isler et al. (1998) test into a universal test, as follows (using the definitions set out in the formula in the next following section):

$$\sum[n{=}1 \text{ IF } (\min s_1 > \max s_2 \text{ OR } \min s_2 > \max s_1) \text{ AND } (|(\bar{x}_1 - \bar{x}_2)| > s_1(t_{1\,@\,97.5\%}) + s_2(t_{2\,@\,97.5\%}))] \leq \sum[n{=}1 \text{ IF } (\min s_3 > \max s_4 \text{ OR } \min s_3 > \max s_4) \text{ AND } (|(\bar{x}_3 - \bar{x}_4)| > s_3(t_{3\,@\,97.5\%}) + s_4(t_{4\,@\,97.5\%}))].$$

Such a hard-edged statistical framework would go beyond the recommendations of Isler et al. (1998) test, who presented their method as a "point of reference" and not a requirement, and take other considerations such as plumage and not-quite-diagnosable characters into account. In practice, those using this method also identify characters that barely failed the test or for which there was an outlier that may cause overlap and consider morphological and other evidence that might be relevant to species recognition (M. Isler in litt. 2016). However, this is only so necessary because the statistical method itself suffers from a shortcoming. A drawback of the system, if applied rigidly, is that those pairs which pass two vocal diagnosability tests very easily will fail to meet

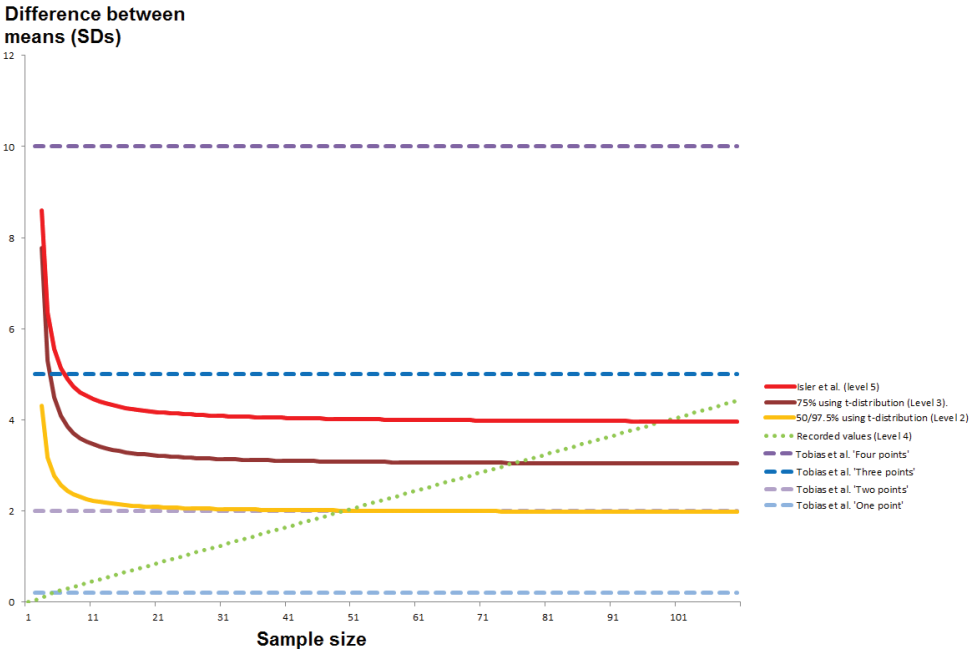

Figure 6. Graph showing relationship between sample size (*x*-axis) and numbers of effective SD differences between means or effect sizes (*y*-axis) required in order to pass a test of diagnosability shown in the legend. Dashed lines represent the four boundaries for affording scores under Tobias et al. (2010). Solid lines represent the Levels 2, 4 and 5 (Isler et al. 1998) tests of diagnosis. The dotted line is based on diagnosability using actual values, for a pairwise comparison of two populations which marginally meet the Level 5 test, where results falling outside a 95% distribution are averaged out in their linear occurrence in the data set. In reality, a data point outside of the 95% distribution could occur randomly at any point along this line, including as the first data point or as data point numbers 96–100. Differences arising from usage of unpooled versus pooled standard deviations are ignored for purposes of simplicity.

the requirement of species rank, even if a third, fourth, fifth and sixth vocal variable fail the test by only a tiny margin.

The Tobias et al. (2010) test is less severely impacted by cut-offs (Figure 7) because it takes data from a broader variety of sources and partitions scores into different marks of up to 4, rather than applying a single cut-off scored on a 0/1 basis. This is in principle a step forwards compared to the Isler et al. (1998) model. However, the Tobias et al. (2010) model still uses hard cut-offs.

There is a relatively simple solution to this shortcoming: with continuous data, to move away from models which attribute cut-offs and instead to apply precise scoring under a system which only uses a hard cut-off at the very final point of determining species or subspecies rank. Such an approach was effectively attempted in a recent study (e.g., Freeman and Montgomery 2017 discussed below) and is also applied through methods such as multivariate statistics. However, multivariate techniques fail to illustrate the full range of variation in multidimensional space, do not test particular characters for diagnosability, and are often presented in a way that affords little assis-

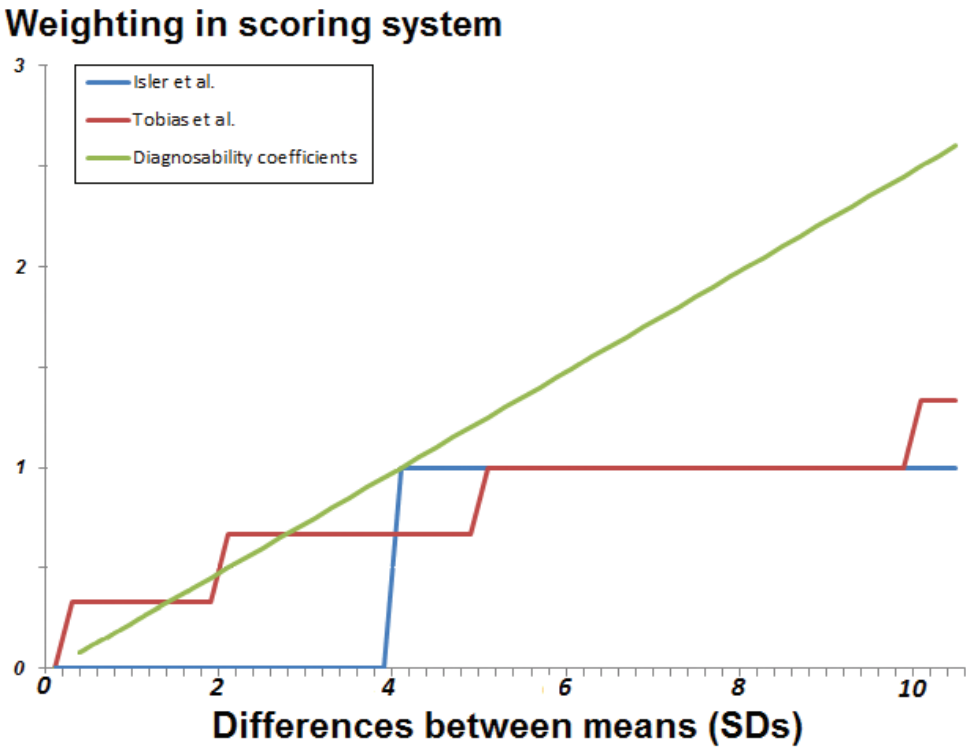

Figure 7. Graph illustrating the "hard cut-off" approaches of Isler et al. (1998) and Tobias et al. (2010). The *y*-axis shows the score attributed under the relevant system, weighted for 1 = diagnosability. The *x*-axis shows effect sizes. In addition to division by three, the Tobias et al. (2010) scores are treated more conservatively by assuming that bare pooled effect sizes are equivalent to controlled unpooled effect sizes. The scores for effect sizes at the lower end of the graph are somewhat artificial with a starting score of 0.34. This is based on the lowest recorded controlled unpooled effect size which passed a statistical significance test for biometrics (see Table 20). In reality, some lower differentiation with larger samples will be scored and some higher variation with lower samples will not be scored at all: see Tables 20–21.

tance to identification. Multivariate tests also require a "complete" set of variables for each individual data "row" which precludes applying the technique to an holistic set of data based, for example, on both in-the-field sound recordings and museum specimen measurements of the same population.

Note on Freeman and Montgomery (2017)

Immediately prior to going to press, Freeman and Montgomery (2017) compared measured differentiation in voice between pairs of allopatric birds against their own bespoke measures of responses to playback using field studies. They measured differentiation by analyzing bare pooled effect sizes, applying a conversion to standardize the data set for a universal mean of 0 and standard deviation of 1, ran principal compo-

nent analysis of the modified data set and then took the measure on the PC1 axis as a surrogate for between-population differentiation. The PC1 axis was found to measure 48% of observed variation in multidimensional space. This method shares a number of close parallels with the methods that will be set out in the next section, in that it attempts to take all measured variation into account and avoids the usage of any hard cut-offs or scoring system except at the point of final diagnosis. Their conversion to SD of uniformly 1 has the same result as the effect size measures used here. Their method does however share several non-optimal aspects of previous studies, in particular: (i) failing to exclude statistically insignificant comparisons; (ii) using bare pooled effect sizes, when controlled unpooled effect sizes are recommended here; and (iii) discarding 52% of observed variation at the last stage, by relying on a single principal component value. Freeman and Montgomery (2017)'s method could be improved in variation capture by measuring centroid distance between PC1 and PC2 and eliminating non-significant outcomes. Other drawbacks of multivariate methods referred to above apply equally here. These authors highlight the importance of playback studies to assess the "allopatric problem" in birds. Considering the results of the analyses proposed here together with the results of molecular studies, playback studies and studies of discrete characters in an holistic manner is of course important in coming to a more informed view of particular taxonomic questions.

A new universal system for measuring differentiation

In this and the next section, a new, universal measure of differentiation is developed. It is potentially usable in any taxonomic group where continuous variables are studied and in other contexts to measure effect sizes.

Step 1: identify a comparison group.
For an assessment of the rank of allopatric populations, this method compares: (i) two sympatric and closely related populations which are demonstrably good species and broadly accepted as such (Species 1 and Species 2) as well as (ii) two allopatric populations under study (Population 3 and Population 4). Ideally, Species 1 and Species 2 should also be sister taxa or be known or suspected to be very closely related through molecular studies, such that they represent a good benchmark. However, this may not always be known for certain. Preferably, Species 1 and 2 and Populations 3 and 4 should all be congeneric, but this might not be possible and they might be merely a good example from the same family or order, depending on how speciose the relevant higher-level taxonomy is. Either (but not both) of Population 3 or Population 4 might be the same as Species 1 or Species 2 or they may all be different populations.

Step 2: collect data for relevant variables using continuous measurements.
It is critical to ensure a fair identification of variables, which adequately and honestly document the maximum possible observed variations between all populations (i.e., not

just the allopatric pair, but also the sympatric pair). Variables differentiating sympatric Species 1 and 2 should not be overlooked, even if more time is spent studying allopatric Populations 3 and 4. Returning to the theme of taxonomic significance and not simply statistical significance, it is important that the variables under study are likely to be taxonomically relevant. Field experience or knowledge of the organisms concerned is important to avoid splits or lumps being published based on statistical tests applied to inappropriately selected variables.

Unlike in multivariate statistics, the technique presented here will not require each data set to have the same measures from the same individuals. This means that a biometric data set based on museum specimens and a vocal data set based on a different set of individuals and with different sampling can be combined, so data from all possible sources can be collated and combined. The broadest possible geographical and numerical sampling is important (e.g., Isler et al. 1998, Tobias et al. 2010).

Step 3: undertake pairwise comparisons using controlled unpooled effect sizes.
The following formula should be applied to measure controlled unpooled effect sizes on a pairwise basis, separately for each population/variable combination under study, e.g., for Species 1 and Species 2:

$$\left|(\bar{x}_1 - \bar{x}_2)\right| / \tfrac{1}{4}[s_1(t_{1\ @\ 97.5\%}) + s_2(t_{2\ @\ 97.5\%})]$$

Step 4: exclude all the statistically insignificant data.
Comparisons showing no statistical significance should be eliminated and scored as 0. This process needs conducting separately for each population/variable combination under study: a variable might be scored as zero as between Species 1 and Species 2, but may be scored positively as between Population 3 and Population 4. Bonferroni correction is applied here, in order to keep the formula simple and due to the near-nil impact of using less conservative "type 1" error corrections. It is recommended that different sets or "families" of data (biometric, vocal, colorimetric) are treated separately for purposes of determining the appropriate Bonferroni correction. Other more complex "type 1 error" corrections such as Dunn-Šidák should be considered for situations where very large numbers of variables are compared. The exclusion of statistically insignificant data results in the following modification to the effect size formula above, e.g., for Species 1 and Species 2:

$$p<0.05/n_v \rightarrow \left|(\bar{x}_1 - \bar{x}_2)\right| / \tfrac{1}{4}[s_1(t_{1\ @\ 97.5\%}) + s_2(t_{2\ @\ 97.5\%})]$$

Step 5: add up all the results of the above calculations (using a Euclidian approach).
It would be simple then to add up all the effect sizes, as follows, and see whether Species 1 vs Species 2 or Population 3 vs Population 4 had the better score. This would apply the formula:

$$\sum [p<0.05/n_v \rightarrow |(\bar{x}_1-\bar{x}_2)| \,/\, ¼[s_1(t_{1\ @\ 97.5\%}) + s_2(t_{2\ @\ 97.5\%})]]$$
$$\leq$$
$$\sum [p<0.05/n_v \rightarrow |(\bar{x}_1-\bar{x}_2)| \,/\, ¼[s_3(t_{3\ @\ 97.5\%}) + s_4(t_{4\ @\ 97.5\%})]]$$

However, this would be sub-optimal statistically. Applying such a formula would reflect the underlying conceptual approach of existing systems to rank allopatric populations (including Tobias et al. 2010), which afford weighting to distances in multiple variables by simple addition. However, in bivariate or multivariate space, a distance based on simple addition of mean differences is overly liberal. In Figure 8, the two circles represent two populations of different standard deviation, with each ring representing one controlled unpooled effect size from the centroid for a relevant variable and each population just being diagnosable from the other. In univariate space, when Var2 does not vary, then the difference between the two populations in Var1 (*x*-axis) is equal to the total variation. However, simple summation of the differentiation between Var1 and Var2 over-estimates the actual distance between the centroids in bivariate space. With two data variables, the distance between the data set (a_1, a_2) and (b_1, b_2) is not $|\ (a_1\text{-}b_1)\ | + |\ (a_2\text{-}b_2)\ |$ but $\sqrt{[(a_1\text{-}b_1)^2 + (a_2\text{-}b_2)^2]}$ (Fig. 8). When analyzing the multi-dimensional points as follows:

$(a_1, a_2, a_3, a_4 \ldots. a_n)$ and $(b_1, b_2, b_3, b_4 \ldots b_n)$,

then Pythagorian principles result in the following calculation of distance between points *a* and *b* in multi-dimensional space:

$$\sqrt{[(a_1-b_1)^2 + (a_2-b_2)^2 + (a_3-b_3)^2 \ldots + (a_n-b_n)^2]}$$

And this can be simplified to:

$$\sqrt{(\sum (a_n-b_n)^2)}$$

This approach cannot perfectly be applied to a series of effect size measures based on multiple pairwise comparisons, in that such data are not necessarily linked to one another as a set of corresponding coordinates. However, assuming that the variables studied are independent, it is valid to measure distance this way. Independence of variables can be verified through correlation tests and promoted in variable selection by seeking to capture the maximum possible observed variation efficiently.

Each controlled unpooled effect size (that has not been eliminated to zero using the statistical significance filter) can be considered to represent the equivalent of a distance $|a_n\text{-}b_n|$. The distance in multi-dimensional space between the two populations is better approximated than through simple addition by taking the square of each controlled unpooled effect size (which has not been excluded due to non-significance), adding those up, and then calculating the square root of the sum of all of them.

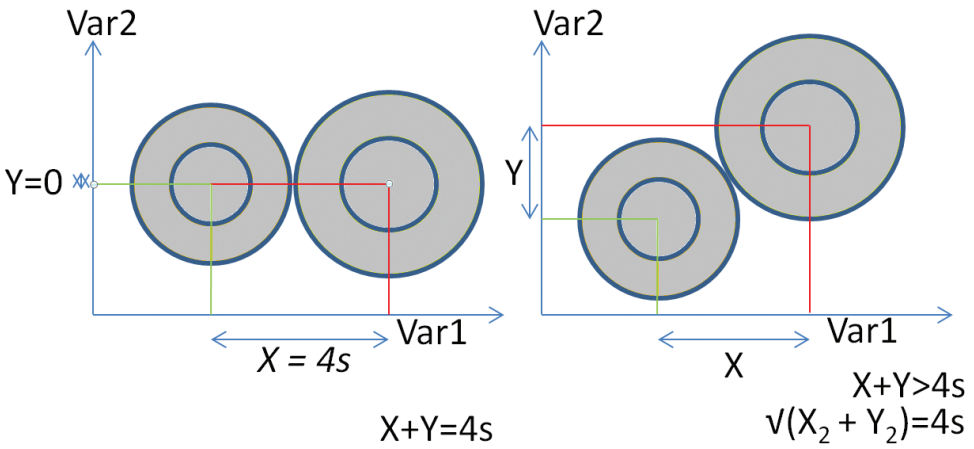

Figure 8. Justification for Euclidian summation, using a univariate/bivariate example.

A new universal formula to determine taxonomic rank of allopatric populations using continuous variables

In studies using continuous variables, allopatric populations should be ranked as species if they show equal to or greater variation than that shown between closely related sympatric species (Helbig et al. 2002). This can be measured by carrying out pairwise comparisons to calculate effect sizes for all variables under study, controlling all these effect sizes for sample size using *t*-distributions, excluding all statistically insignificant outcomes and applying Euclidian summation.

Viz, an allopatric population will be a candidate for species rank if:

$$\sqrt{(\sum [p<0.05/n_v \rightarrow |(\bar{x}_1-\bar{x}_2)| / ¼[s_1(t_{1\ @\ 97.5\%}) + s_2(t_{2\ @\ 97.5\%})]]^{\ 2})}$$
$$\leq$$
$$\sqrt{(\sum [p<0.05/n_v \rightarrow |(\bar{x}_3-\bar{x}_4)| / ¼[s_3(t_{3\ @\ 97.5\%}) + s_4(t_{4\ @\ 97.5\%})]]^{2})}$$

Where:

Species 1 and Species 2 are two sympatric species that are closely related to one another (preferably known to be sisters) and which are related to Population 3 and Population 4.

Population 3 and Population 4 are two allopatric populations whose rank is being determined.

p: the probability using Welch's unequal variance *t*-test (or other similar technique for non-normally distributed data), as set out under Level 1 in Methods that the means of the populations differ.

n_v: the number of continuous variables of a particular "family" considered in the study, so as to apply a Bonferroni correction.

1, 2, 3, and 4 refer to relevant data for Species 1, Species 2, Population 3 and Population 4 respectively.

$\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$, and $\bar{x}_4$ are the sample means of a relevant data set for a particular variable for Species 1, Species 2, Population 3 and Population 4, respectively.

s_1, s_2, s_3, and s_4 are the standard deviations for a relevant data set for a particular variable for Species 1, Species 2, Population 3 and Population 4, respectively.

t refers to the *t*-value (based on *t*-distribution) using a one-sided confidence interval at the percentage specified for the relevant population and variable, with t_1, t_2, t_3 and t_4 referring to such value for Species 1, Species 2, Population 3 and Population 4 respectively.

Because this formula is not simple to calculate, a spreadsheet is being published alongside this paper on the author's researchgate.net site, which facilitates rapid calculations.

In some ways, when one looks carefully at the outcomes of different statistical tests undertaken here, the formula is a statement of the obvious. The statistics underlying it are basic. It merely relies upon the good aspects of long-established statistical methods for comparing continuous variables of previous authors (e.g., Hubbs and Perlmutter 1942, Amadon 1949, Isler et al. 1998, Tobias et al. 2010) and stands on their shoulders considerably. However, no one has to my knowledge or to date (at least in ornithology) proposed such a universal test of species rank of allopatric populations based on continuous variables into a single statistical framework.

As regards Isler et al. (1998)'s contributions, this formula borrows from their concept in using controlled unpooled effect size measures as its basis. However, their approach is genericized in a way potentially applicable to all taxonomic groups, given that different groups can show differing levels of intra-species variation when subjected to universal scoring systems (Donegan et al. 2015). Isler et al. (1998)'s additional test of non-overlap is discarded but a different gateway test, namely statistical significance, is added. Finally, effect sizes are applied in such a way as to take into account all validly measured variation, not just those variables which are diagnosable. This test also borrows from certain aspects of the Tobias et al. (2010) species-scoring technique – attributing value to a broader range of different effect sizes, including those below diagnosability, but unlike those authors discarding statistically insignificant data, using unpooled standard deviations and controlling better for sample size. Contrary to all previous models, a proportionate score based on a sliding scale is attributed, depending on how large the measured differences are. The test also has further benefits which are not delivered under any other systems to evaluate differentiation and taxonomic rank. First, it eliminates the possibility of non-statistically significant results affecting taxonomy at all. Secondly, it eliminates "hard cut-offs" entirely, other than in respect of the final determination of species rank. Thirdly, it is extensible. A study of two variables or of 100 variables can be applied equally into this framework. Fourthly, it

is taxonomically neutral and can in principle be applied to any genus, family, order or class of any organism when continuous variables are studied, or indeed for any kind of statistical study using continuous variables in which diagnosability or comparability of differences is tested.

Some important recommendations should be borne in mind when using this method, in addition to those set out above under "Steps":

(i) *Continuous versus non-continuous variables*: Some issues arose in the case studies here due to data gaps. In the studies of *Sirystes* and *Basileuterus*, non-homologous vocalizations were not compared with one another. There, populations with different measures for the same sorts of variables are recovered as more differentiated than those populations whose variables cannot validly be compared at all. Diagnosability based on the comparison of non-homologous vocalizations can be important, but it can also have pitfalls (notably, Chaves et al. 2000 claimed discrete variation in calls to claim sufficient differentiation under the Isler et al. (1998) model to support a split in antbirds, but homologous vocalizations existed that were ignored). Where populations differ principally by non-continuous rather than continuous variables and sample sizes are sufficient (Walsh 2000), then the scoring system of Tobias et al. (2010) (as modified here) is better applied, since that incorporates the study of both kinds of variables. In general, where populations are so different that variables cannot effectively or meaningfully be compared at all using continuous data, then this is likely of itself to be indicative of species rank. See further paragraph (iv) below.

(ii) *Sample sizes*: If either Species 1 or Species 2 have very low sample sizes, then: (i) for many variables under study, data may not meet the threshold test of statistical significance; and (ii) those which do could be affected by low standard deviations and inflated effect sizes caused by clustering. These issues apply in reverse where Population 3 or Population 4 suffers from such constraints. Although the test above is in principle sample size-neutral, caution should be exercised in interpreting results based on smaller samples (see *Myrmeciza* biometrics discussion below and Table 24). That said, small sample sizes are often an inescapable fact and, if this is the case, then we should feel comfortable about applying statistical methods such as these (which seek to address sample sizes to a particular confidence interval) and acting on the basis of their outcomes. We should also be enthusiastic about revising taxonomies and conclusions when further data implies a need to do so, rather than affording undue weight to *status quo* taxonomies or to older studies, which are usually based on even fewer data or even lower sample sizes.

(iii) *Scale factoring and manipulation through overloading*: Where there are 15 vocal measures and 5 biometric measures, it should be considered whether to weight scoring on a 50:50 basis, following Tobias et al. (2010)'s recommendations of equal weighting for different sources of data. There is a potential risk of misuse linked to

scale factors. For example, if tail length varies between Population 3 and Population 4 but is equal between Species 1 and Species 2, it would seem inappropriate to increase the impact of this observation though ten separate tail feather measurements. With bill length, similarly, one could measure separately from the skull, nostril, and feathering to create three variables out of one; or biometric weightings could be tripled in impact with male, female and combined data or duplicated by using both museum specimen and live specimen data separately. In some cases, where all the wing feathers are also all measured, then measuring all the tail feathers might be appropriate, but such variables may then exceed the number of vocal variables in a study and require scale factoring. In some species, male and female songs are often very similar to one another; female songs showing similar patterns of between-population differentiation to male songs may be best excluded to avoid bias and doubling-up of scores. In some groups, biometrics may be more likely to be informative to taxonomy than voice or *vice versa*. Based on the case studies included here, which are largely of forest species where vocal characters are important, I will suggest that scale factoring is best addressed simply by an honest attempt to encapsulate the maximum possible extent of observed differences through the smallest possible number variables, with no more, even if (as in many studies here) this means giving equal weightings to tens of vocal measures and only five biometric measures. However, this suggestion should not discourage more in-depth and detailed studies involving weighting to avoid particular components becoming dominant. Isler et al. (1998) applied correlation tests to eliminate related variables, which should also be considered to improve the robustness of variable sets and is effectively a form of weighting. Remsen (2016) suggested studying different families for taxonomically informative characters, which could easily be incorporated into this framework.

(iv) *Not going over the top*: The formula presented here is proposed for usage in more difficult, borderline, or complicated cases. Where simpler studies can show allopatric populations or newly discovered populations to be very different indeed from one another, then there should be no need for a litany of statistical analyses to be undertaken. It should be appropriate in some cases simply for an author to publish photographs of specimens or sonograms or a brief subjective text to describe the differences observed. A good example of a situation in this category would be the allopatric Western and Eastern Woodhaunters *Automolus virgatus* and *A. subulatus*, whose vocalizations resemble one another not one iota (Ridgely and Greenfield 2001, Donegan et al. 2011) and show no mutual playback response (Freeman and Montgomery 2017), but whose split has not been universally accepted to date. Notably, Remsen et al. (2018) rejected this, one committee member considering that "there is value in requiring some minimum standards of published data for making taxonomic changes" and in a further ongoing attempt at promoting this change, another committee member has proposed rejection "out of principle". Such approaches waste limited human taxonomic resources on simple situations and reinforce irrational taxonomies.

(v) *Possible usage of controlled pooled effect size*. The formula above does not use pooled standard deviations and so makes no assumptions about the comparability of the variances of the different populations under study. As discussed above, there may be use cases for controlled pooled effect size, especially as a hedge for small sample sizes, but this should be applied only with caution. In any cases where assumptions of equal SD may be made among all four populations 1, 2, 3 and 4, then a more complicated formula using controlled pooled standard deviations might be used instead (see Materials and methods for details of equations that may be substituted in).

Example of using the test: *Myrmeciza* antbirds

Myrmeciza was chosen here as an example because the recommendations of the relevant paper (Donegan 2012) have been accepted by all relevant authorities (Remsen et al. 2018, Gill and Donsker 2018, Del Hoyo and Collar 2016) and are justified based on both the Isler et al. (1998) and Tobias et al. (2010) models. Moreover, the species rank candidates (Blue-lored Antbird *M. immaculata* versus Zeledon's Antbird *M. zeledoni)* are sisters with respect to one another, as are the comparator sympatric pair (Goeldi's Antbird *M. goeldii* versus White-shouldered Antbird *M. melanoceps*) (Isler et al. 2013). Vocal data is considered here primarily, since only two specimens of *M. goeldii* were found in the study, resulting in no statistical significance being recovered in any biometric comparisons and scores of 0 across the board. A simplistic comparison is first shown, of the Central Andes population *M. i. concepcion* versus the proximate Chocó population of *M. z. macrorhyncha*. In reality, the relevant allopatric species both each comprise two allopatric subspecies. Measures for each of the vocal variables in question can be inspected in Donegan (2012). Bonferroni correction at $p<0.05/26$ vocal variables produces $p<0.0019$ for voice.

Table 22 includes a work-through of the calculation under the new methodology proposed here. Scores, of *immaculata/zeledoni*: 13.75 > 7.13: *melanoceps/goeldii*, imply that differences in voice between the allopatric pair are greater than those between related sympatric pair. The requirement of the new formula is satisfied and *zeledoni* and *immaculata*, which were treated as subspecies under traditional taxonomies, are therefore valid species with respect to one another.

Table 23 shows vocal scores for cross-comparisons of the entire study group in *Myrmeciza.* The two scores in italics are for those achieving only subspecies rank under this system, i.e., those populations failing to attain the 7.14 suggested benchmark for species rank in this genus. Other allopatric populations concerned have all speciated with respect to one another and were previously ranked in separate species under traditional taxonomies. Sooty Antbird *Myrmeciza fortis* is also sympatric with respect to both *M. goeldii* and *M. melanoceps*.

Biometric scores are shown in Table 24. Even if a 3.52 uplift could be applied to biometric scores for the sympatric pair, which would raise the overall benchmark to 10.66, *M. zeledoni* and *M. immaculata* still meet the required benchmark for species with respect to one another.

What sorts of scores are good enough for assessing species and subspecies rank?

As above, although a universal *formula* is proposed here, no universal *score* is proposed here for ranking species, since the differentiation required to rank a species is likely to vary depending on the number of variables studied and by taxonomic group (Donegan and Avendaño 2008, Donegan et al. 2015). Simply, the score given to the allopatrics under study must exceed that of the related sympatrics.

There are however some parameters and examples available from the case studies (Table 25). The range of scores here for sympatrics may or may not be typical. Some presently recognized allopatric species scored less than these scores in some studies. A score of 4 under this system should be regarded as a *very bare minimum* for any proposal to rank a species based solely on continuous data. At that point, the two populations are differentiated in multi-dimensional space to 95% confidence. However, the actual value (being greater than 4) that it is necessary so as to afford species rank will depend on the data set and intra-specific variation in the group under study.

As regards subspecies or PSC species, any score of 4 or more (i.e., allowing full diagnosis in multidimensional space) would be a supportable benchmark. There may however be cases of valid subspecies which achieve lower scores than this, such as in the *Adelomyia melanogenys* study where the pair discussed above scored only 2.10 for voice and 0 for biometrics, based on a fairly exhaustive attempt at measuring biometric and vocal variables. However, this is probably an exceptionally low-scoring example.

Among the un-named populations in the study group, only the "Apurímac south" population of *Basileuterus tristriatus* in Peru was recovered as diagnosable versus all proximate subspecies (6.33 versus Marañon to Apurímac population, 14.79 versus Bolivia) and therefore requires formal description. Other notable unnamed populations include the Tamá population of *Grallaricula nana* (scores 3.32 versus Mérida) and the West Andes populations of the same species (scores 2.86 against Central Andes). The two new *Grallaricula* taxa described in Donegan (2008) each scored over 4 compared to proximate populations. A noteworthy split proposed by Donegan (2008) but rejected by all relevant learned taxonomic committees (Remsen et al. 2018, Gill and Donsker 2018, Del Hoyo and Collar 2016) is that of *Grallaricula nana kukenamensis*, which scored 6.24–11.19 on biometric data alone compared to all other populations in the *nana* group with which it is purportedly lumped – although only a single tentative sound recording remains available. Such biometric differentiation exceeds that of almost all sympatric pairs studied here (including many that are not sister taxa). Taking into account plumage differences, it also exceeds the scoring benchmark of Tobias et al. (2010) for species rank.

Probably the most difficult taxonomic decision in this series of papers was that of how to rank *Scytalopus rodriguezi*, whose allopatric subspecies scored 5.40 for biometrics and 5.01 for voice, total 10.41 and so was more diagnosable than some sympatric tapaculos. A large component of this score (compared to sympatric pairs) was in biometrics and the two populations were found to respond to one another's playback. In borderline cases such as this, where different kinds of variables differ

Table 22. Worked-through example of the new formula proposed herein, assessing the rank of two allopatric *Myrmeciza* antbirds (*M. immaculata* vs *M. zeledoni*) by comparison to a pair of sympatric sister taxa in the same genus (*M. goeldii* vs. *M. melanoceps*), using vocal data only.

Variable	***M. goeldii*** vs. ***M. melanoceps***			***M. immaculata*** vs. ***M. zeledoni***		
	Controlled unpooled effect size	***p* value**	**Score**	**Controlled unpooled effect size**	***p* value**	**Score**
Male song						
No. of notes	1.41	5.2×10^{-29}	1.41	4.54	6.92×10^{-33}	4.54
Song length	0.37	0.0015	0.37	0.71	0.0022	0
Song speed	2.05	3.9×10^{-49}	2.05	7.65	1.03×10^{-89}	7.65
Max. acoustic frequency second note	1.03	2.7×10^{-16}	1.03	3.30	7.20×10^{-18}	3.30
Max. acoustic frequency of last note	1.31	3.9×10^{-23}	1.31	2.88	2.03×10^{-15}	2.88
Change in acoustic frequency	0.52	3.7×10^{-5}	0.52	1.36	2.62×10^{-8}	1.36
Position of peak of frequency	4.75	4.5×10^{-74}	4.75	0.08	0.76	0
Position of trough in frequency	3.83	9.5×10^{-55}	3.83	0.03	0.91	0
Single note call						
Call length	0.91	0.00270	0	0.99	0.040	0
Maximum acoustic frequency	0.99	0.00103	0.99	0.37	0.16	0
Multi-note call						
No. of notes	0.11	0.770	0	0.01	0.99	0
Song length	0.06	0.880	0	0.22	0.57	0
Song speed	0.21	0.494	0	0.21	0.56	0
Max. acoustic frequency	0.31	0.371	0	1.71	0.00049	1.71
Min. acoustic frequency	0.19	0.529	0	2.29	0.00013	2.29
Change in acoustic frequency	0.24	0.447	0	0.45	0.25	0
Position of peak of frequency	0.49	0.396	0	0.65	0.12	0
Position of trough in frequency	0.02	0.946	0	0.12	0.74	0
Female song						
No. of notes	0.78	0.0073	0	5.52	2.41×10^{-15}	5.52
Song length	0.60	0.034	0	1.00	0.023	0
Song speed	1.18	7.07×10^{-5}	1.18	7.00	1.50×10^{-19}	7.00
Max. acoustic frequency second note	0.61	0.031	0	1.62	0.0034	0
Max. acoustic frequency of last note	1.14	0.00020	1.14	2.61	3.86×10^{-7}	2.61
Change in acoustic frequency	0.23	0.40	0	0.71	0.18	0
Position of peak of frequency	0.04	0.89	0	0.96	0.11	0
Position of trough in frequency	0.27	0.33	0	0.47	0.41	0
Euclidian distance (square root of sum of the squares)			**7.09** *(7.14 using data to more s.f.)*			**13.95** *(13.75 using data to more s.f.)*

Table 23. Full scores across the *Myrmeciza* data set for vocal data only. Bold denotes a sympatric pair of sister taxa. Bold italics denote other sympatric pairs. Denote pairs (with asterisk) which are subspecies based on overall scoring and discrete characters (see also Table 24). All other comparisons are between allopatric populations ranked with respect to one another as species.

	M. i. concepcion	***M. z. macrorhyncha***	***M. z. zeledoni***	***M. fortis***	***M. goeldii***	***M. melanoceps***
M. i. immaculata	3.56*	12.22	11.23	28.00	18.15	14.60
M. i. concepcion		13.75	15.64	19.51	21.49	15.50
M. z. macrorhyncha			4.08*	23.74	21.13	17.85
M. z. zeledoni				28.45	28.15	24.39
M. fortis					***22.81***	***12.73***
M. goeldii						**7.14**

Table 24. Scores across the *Myrmeciza* data set for biometrics. All *goeldii* scores (sample size *n*=2 specimens) were actually zero. Square bracketed figures showing alongside *M. goeldii* are based on controlled effect sizes without deleting insignificant data and are presented for reference only Bold denotes a sympatric pair of sister taxa. Bold italics denote other sympatric pairs. Denote subspecies (with asterisk) based on overall scoring (see also Table 23). All other comparisons are between allopatric populations ranked with respect to one another as species.

Taxon	***M. i. concepcion***	***M. z. macrorhyncha***	***M. z. zeledoni***	***M. fortis***	***M. goeldii***	***M. melanoceps***
M. i. immaculata	0*	3.02	1.72	3.35	[5.36]	5.96
M. i. concepcion		2.49	0	3.01	[5.13]	5.72
M. z. macrorhyncha			1.64*	2.22	[3.01]	5.07
M. z. zeledoni				2.54	[4.59]	5.35
M. fortis					***[3.47]***	***3.58***
M. goeldii						**[3.52]**

between the sympatric pair and allopatric pair, then scale factoring may be appropriate. Voice is a very important character for tapaculos and in this case, the vocal score, whilst showing full diagnosis, did not attain the differentiation shown between known sympatric comparators.

The Tobias et al. (2010) system produces some scores which are consistent with the differences between sympatric species studied here. However, species can be justified under that scoring system based on wildly differing measured variation (Table 25). The quest for a universal scoring system requires a more exacting basis for calculations in order to produce the rational taxonomy that it seeks. It was most surprising in this study that the measured differentiation here between sympatric *Scytalopus* exceeded that between sympatric *Myrmeciza*. In contrast, as mentioned in the introduction, Donegan and Avendaño (2008) found the same sympatric tapaculo pairs to differ to a lesser extent than sympatric antbirds under Isler et al. (1998)'s framework. The present study shows that when below-diagnosis differentiation in other characters is taken into

Table 25. Scores of examples from the data set which are both (i) sister species (or relevant sympatric subspecies of sister species) as shown by molecular studies; and (ii) sympatric, to show ranges of scores. Also presented are examples of scores passing other authors' species tests. Note the Isler et al. (1998) score is likely an underestimate these since it does not take into account any non-diagnosable but significant other variation. Note that Tobias et al. (2010) allows only a maximum of four continuous variables (2 biometrics, 2 vocal) to be counted, so should be similarly interpreted as conservative scores.

Sympatric pair or proposed score for species rank	Type of data	Score
Myrmeciza goeldii vs *Myrmeciza melanoceps*	Voice	7.14
Scytalopus griseicollis griseicollis vs. *Scytalopus spillmanni* undescribed East Andes population	Voice + biometrics = total	9.16 + 0 = 9.16
Scytalopus griseicollis gilesi vs. *Scytalopus spillmanni* undescribed East Andes population	Voice + biometrics = total	10.59 + 0 = 10.59
Scytalopus griseicollis morenoi vs. *Scytalopus spillmanni* undescribed East Andes population	Voice + biometrics = total	8.79 + 0 = 8.79
Grallaricula ferrugineipectus Venezuela vs *G. nana nanitaea* Merida Andes	Voice + biometrics = total	7.90 + 8.01 = 15.91
Average ± s.d. ***(min.–max.)*** ***(n=sample number)***		**10.32 ± 3.36** **(7.14–15.91)** **(*n*=5)**
Basis for Isler et al. (1998) model	Voice: diagnosability of three characters = 3 × 4 SD	6.92
A basis for Tobias et al. (2010) score of 7 for species rank	Voice or biometrics: 1 × 10 SD (score 4), 1 × 5 SD (score 3)	11.18
A basis for Tobias et al. (2010) score of 7 for species rank	Voice and biometrics: 1 × 10 SD (score 4), 1 × 2 SD (score 2), 1 x 0.2 SD (score 1)	10.20
A basis for Tobias et al. (2010) score of 7 for species rank	Voice and biometrics: 1 × 10 SD (score 4), 3 × 0.2 SD (score 1 each)	10.01
A basis for Tobias et al. (2010) score of 7 for species rank	Voice and biometrics: 2 × 5 SD (score 3 each), 1 × 0.2 SD (score 1)	7.07
A basis for Tobias et al. (2010) score of 7 for species rank	Voice and biometrics: 1 × 5 SD (score 3), 2 × 2 SD (score 2 each)	5.74
A basis for Tobias et al. (2010) score of 7 for species rank	Voice and biometrics: 1 × 5 SD (score 3), 1 × 2 SD (score 2), 2 × 0.2 SD (score 1 each)	5.39
A basis for Tobias et al. (2010) score of 7 for species rank	Voice and biometrics: 3 × 2 SD (2 each), 1 × 0.2 SD (1 each)	3.47

account (see Figure 5), sympatric tapaculos are much closer in their vocal differentiation to sympatric antbirds.

The method proposed here involves no universal score for species rank. However, it would still be interesting to see how other pairwise situations involving sympatric sister species measure up under this system, and then possibly to revisit the philosophy underlying Tobias et al. (2010)'s methods accordingly.

Acknowledgments

Thanks to Michael Patten, Mort Isler, and George Sangster for helpful comments on this paper. The inclusion of a Bonferroni correction in the Level 1 test was suggested by F. Gary Stiles (in litt. 2007). Michael Patten made various insightful comments which resulted in addition of all the analyses and methods involving pooled standard deviation data and the various additional "type 1" corrections. Some of the methods underlying this paper were developed in papers published together with Jorge Avendaño, Paul Salaman, and others. Thanks as always to Blanca and Lucas for their love, good cheer, and encouragement.

References

Amadon D (1949) The seventy-five per cent rule for subspecies. Condor 51: 250–258. https://doi.org/10.2307/1364805

American Ornithologists' Union (1998) Checklist of North American Birds (7th edn). American Ornithologists' Union, Washington, D.C.

Aristotle (c. 350 B.C.) The history of animals.

Bakker G (2015) Review of: HBW and BirdLife International Illustrated Checklist of the Birds of the World Volume 1: Non-passerines. Dutch Birding online. https://www.dutchbirding.nl/recensies/1028/hbw_and_birdlife_international_illustrated_checklist_of_the_birds_of_the_world_volume_1_non-passerines

Barrowclough GF, Cracraft J, Klicka J, Zink RM (2016) How many kinds of bird are there and why does it matter? PLoS ONE 11: e0166307. https://doi.org/10.1371/journal.pone.0166307

Benkman C, Smith JW, Keenan PC, Parchman TL (2009) A new species of the Red Crossbill (Fringillidae: *Loxia*) from Idaho. *Condor* 111: 169–176. https://doi.org/10.1525/cond.2009.080042

Chaves JA, Smith TB (2011) Evolutionary patterns of diversification in the Andean hummingbird genus *Adelomyia*. Molecular Phylogenetics and Evolution 60: 207–218. https://doi.org/10.1016/j.ympev.2011.04.007

Chaves JC, Cuervo AM, Miller MJ, Cadena CD (2010) Revising species limits in a group of *Myrmeciza* antbirds reveals a cryptic species within *M. laemosticta*. Condor 112: 718–730. https://doi.org/10.1525/cond.2010.100098

Cohen J (1988) Statistical Power Analysis for the Behavioral Sciences (2nd edn). Routledge, 400 pp.

Collar NJ (2018) Taxonomy as tyranny. Ibis 160 (2): 481–484. https://doi.org/10.1111/ibi.12569

Collar NJ, Fishpool LDC, del Hoyo J, Pilgrim JD, Seddon N, Spottiswoode CN, Tobias JA (2016) Toward a scoring system for species delimitation: a response to Remsen. Journal of Field Ornithology 87(1): 104–115. https://doi.org/10.1111/jofo.12141

Collinson JM, Parkin DT, Knox AG, Sangster G, Svensson L (2008) Species boundaries in the Herring and Lesser Black-backed Gull complex. British Birds 101: 340–363.

Cracraft J (1983) Species concepts and speciation analysis. Current Ornithology 1: 159–187. https://doi.org/10.1007/978-1-4615-6781-3_6

de Queiroz K (1998) The general lineage concept of species, species criteria,and the process of speciation: A conceptual unification and terminological recommendations. In: Howard DJ, Berlocher SH (Eds) Endless Forms: Species and Speciation. Oxford University Press, Oxford, 57–75.

de Queiroz K (1999) The general lineage concept of species and the defining properties of the species category. In: Wilson RA (Ed.) Species: New Interdisciplinary Essays. MIT Press, Cambridge, Massachusetts, 49–89.

Del Hoyo J, Collar NJ (2014) Illustrated Checklist of the Birds of the World, Volume 1: Non-passerines. Lynx Edicions, Barcelona & BirdLife International, Cambridge, 904 pp.

Del Hoyo, Collar NJ (2016) Illustrated Checklist of the Birds of the World, Volume 2: Passerines. Lynx Edicions, Barcelona & BirdLife International, Cambridge, 1013 pp.

Dickinson EC, Christidis L (2014) The Howard and Moore Complete Checklist of the Birds of the World: Passerines. Vol 2. Aves Press Ltd., Eastbourne, 752 pp.

Donegan TM (2008) Geographical variation in Slate-crowned Antpitta Grallaricula nana, with two new subspecies from Colombia and Venezuela. Bulletin of the British Ornithologists' Club 128: 150–178.

Donegan TM (2012) Geographical variation in Immaculate Antbird *Myrmeciza immaculata*, with a new subspecies from the Central Andes of Colombia. Bulletin of the British Ornithologists' Club 132: 3–40.

Donegan TM (2013) Vocal variation and species limits in the genus *Sirystes* (Tyrannidae). Conservación Colombiana 19: 11–30.

Donegan TM (2014) Geographical variation in morphology and voice of Three-striped Warbler *Basileuterus tristriatus*. Bulletin of the British Ornithologists' Club 134: 79–109.

Donegan TM, Avendaño JE (2008) Notes on tapaculos (Passeriformes: Rhinocryptidae) of the Eastern Andes of Colombia and Venezuelan Andes, with a new subspecies of *Scytalopus griseicollis* from Colombia. Ornitología Colombiana 6: 24–65.

Donegan TM, Avendaño J (2010) A new subspecies of mountain tanager in the *Anisognathus lacrymosus* complex from the Yariguíes Mountains of Colombia. Bulletin of the British Ornithologists' Club 130: 13–32.

Donegan TM, Avendaño JE (2014) Bogotá' type specimens of the hummingbird genus *Adelomyia*, with diagnosis of an overlooked subspecies from the East Andes of Colombia. Bulletin of the British Ornithologists' Club 135(3):195–215.

Donegan TM, Salaman PGW (2012) Vocal differentiation and conservation of Indigo-crowned Quail-Dove *Geotrygon purpurata*. Conservación Colombiana 17: 15–19.

Donegan TM, Avendano JE, Lambert F (2013) A new Tapaculo related to *Scytalopus rodriguezi* from Serranía de los Yariguíes, Colombia. Bulletin of the British Ornithologists' Club 133: 256–271.

Donegan T, Quevedo A, Salaman P, McMullan M (2012) Revision of the status of bird species occurring or reported in Colombia 2012. Conservación Colombiana 17: 4–14.

Donegan T, Quevedo A, Verhelst JC, Cortés-Herrera O, Ellery T, Salaman P (2015) Revision of the status of bird species occurring or reported in Colombia 2015, with discussion of BirdLife International's new taxonomy. Conservación Colombiana 23: 3–48.

Dubois A (2017) Diagnoses in zoological taxonomy and nomenclature. Bionomina 12: 63–85. https://doi.org/10.11646/bionomina.12.1.8

Freeman BF, Montgomery GA (2017) Using song playback experiments to measure species recognition between geographically isolated populations: a comparison with acoustic trait analyses. Auk 134(4): 857–870. https://doi.org/10.1642/AUK-17-63.1

Freitas GHS, Chaves AV, Costa LM, Santos FR, Rodrigues M (2012) A new species of *Cinclodes* from the Espinhaço Range, southeastern Brazil: insights into the biogeographical history of the South American highlands. Ibis 154: 738–755. https://doi.org/10.1111/j.1474-919X.2012.01268.x

Garnett ST, Christidis L (2017) Taxonomy anarchy hampers conservation. Nature 546: 25–27. https://doi.org/10.1038/546025a

Gill FB (2014) Species taxonomy of birds: Which null hypothesis? The Auk: Ornithological Advances 131: 150–161. https://doi.org/10.1642/AUK-13-206.1

Gill F, Donsker D (Eds) (2018) IOC World Bird List (v8.1). https://doi.org/10.14344/IOC.ML.8.1.

Helbig AJ, Knox AG, Parkin DT, Sangster G, Collinson M (2002) Guidelines for assigning species rank. Ibis 144: 518–525. https://doi.org/10.1046/j.1474-919X.2002.00091.x

Hubbs CL, Perlmutter A (1942) Biometric comparison of several samples with particular reference to racial investigations. American Naturalist 76: 582–592. https://doi.org/10.1086/281082

International Commission on Zoological Nomenclature (ICZN) (1999) International code of zoological nomenclature (4th edn). The International Trust for Zoological Nomenclature, Natural History Museum, London, 106 pp.

Isler ML, Isler PR, Whitney BM (1998) Use of vocalizations to establish species limits in antbirds (Passeriformes; Thamnophilidae). Auk 115: 577–590. https://doi.org/10.2307/4089407

Isler ML, Bravo GA, Brumfield RT (2013) Taxonomic revision of Myrmeciza (Aves: Passeriformes: Thamnophilidae) into 12 genera based on phylogenetic, morphological, behavioral, and ecological data. Zootaxa 3717: 469–497. https://doi.org/10.11646/zootaxa.3717.4.3

Lara CE, Cuervo AM, Valderrama SV, Calderón-Franco D, Cadena CD (2012) A new species of wren (Troglodytidae) from the dry Cauca river canyon, northwestern Colombia. Auk 129(3): 537–550. https://doi.org/10.1525/auk.2012.12028

Lim G, Balke M, Meier R (2012) Determining species boundaries in a world full of rarity: singletons, species delimitation methods. Systematic Biology 61: 165–169. https://doi.org/10.1093/sysbio/syr030

Lonergan P, Mullarney K (2004) Identification of American Herring Gull in a western European context. Dutch Birding 26: 1–35.

Marks BD, Hackett SJ, Capparella AP (2002) Historical relationships among Neotropical lowland forest areas of endemism as determined by mitochondrial DNA sequence variation within the Wedge-billed Woodcreeper (Aves: Dendrocolaptidae: *Glyphorynchus spirurus*). Molecular Phylogenetics and Evolution 24: 153–16. https://doi.org/10.1016/S1055-7903(02)00233-6

Mayr E (1940) Speciation phenomena in birds. American Naturalist 74, 249–278. https://doi.org/10.1086/280892

Mayr E (1942) Systematics and Origin of Species. Columbia University Press, New York.

Mayr E, Linsley EG, Usinger RL (1953) Methods and Principles of Systematic Zoology. McGraw-Hill, New York, 328 pp.

McKitrick MC, Zink RM (1988) Species concepts in ornithology. Condor 90: 1–14. https://doi.org/10.2307/1368426

Nakagawa S, Cuthill IC (2007) Effect size, confidence intervals and statistical significance: a practical guide for biologists. Biological Reviews of the Cambridge Philosophical Society 82: 591–605. https://doi.org/10.1111/j.1469-185X.2007.00027.x

Parkin DT, Collinson M, Helbig AJ, Knox AG, Sangster G (2003) The taxonomic status of Carrion and Hooded Crows. British Birds 96: 274–290.

Patané JSL, Weckstein JD, Aleixo A, Bates JM (2009) Evolutionary history of *Ramphastos* toucans: molecular phylogenetics, temporal diversification, and biogeography. Molecular Phylogenetics & Evolution 53: 923–934. https://doi.org/10.1016/j.ympev.2009.08.017

Patten MA, Unitt P (2002) Diagnosability versus mean differences of Sage Sparrows subspecies. Auk 119: 26–35. https://doi.org/10.1642/0004-8038(2002)119[0026:DVMDOS]2.0.CO;2

Phillips JN, Derryberry EP (2017) Equivalent effects of bandwidth and trill rate: support for a performance constraint as a competitive signal. Animal Behaviour 132: 209–215. https://doi.org/10.1016/j.anbehav.2017.08.012

Podos J, Southall JA, Rossi-Santos MR (2004) Vocal mechanics in Darwin's finches: correlation of beak gape and song frequency. Journal of Experimental Biology 207: 607–619. https://doi.org/10.1242/jeb.00770

Poulton EB (1904) What is a species? Proceedings of the Entomological Society of London 1903: 77–96.

Poulton EB (1908) What is a species? (The Presidential Address read at the Annual Meeting of the Entomological Society of London, January 20 1904. Reprinted from the Proceedings of the Society, 1903, p. lxxvii. Revised and modified: several paragraphs added and several rewritten: additional footnotes). In: Poulton EB (Ed.) Essays on Evolution. 1889–1907. Clarendon Press, Oxford, 46–94

Raposo MA, Stopiglia R, Brito GRR, Bockmann FA, Kirwan GM, Gayon J, Dubois A (2017) What really hampers taxonomy and conservation? A riposte to Garnett and Christidis (2017). Zootaxa 4317: 179–184. https://doi.org/10.11646/zootaxa.4317.1.10

Ray J (1686) Historia plantarum. Vol 1. Clark, London, 983 pp.

Remsen JV (2010) Subspecies as a meaningful taxonomic rank in avian classification. Ornithological Monographs 67: 62–78. https://doi.org/10.1525/om.2010.67.1.62

Remsen JV (2015) Review of: HBW and BirdLife International Illustrated Checklist of the Birds of the World Volume 1: Non-passerines, Josep del Hoyo and Nigel J. Collar 2014. Lynx Edicions, Barcelona. 903 pages, hundreds of color plates. ISBN 9788496553941. $209 (Hardcover). Journal of Field Ornithology 86: 182–187. https://doi.org/10.1111/jofo.12102

Remsen JV (2016) A "rapid assessment program" for assigning species rank? Journal of Field Ornithology 87(1):110–115. https://doi.org/10.1111/jofo.12142

Remsen JV, Areta JI, Cadena CD, Claramunt S, Jaramillo A, Pacheco JF, Pérez-Éman J, Robbins MB, Stiles FG, Stotz DF, Zimmer KJ (2018) A classification of the bird species of South America (version 10 March 2018).

Ridgely RS, Greenfield PJ (2001) The birds of Ecuador (Volume I) – Status, Distribution, and Taxonomy. Cornell University Press, Ithaca, New York, 880 pp.

Salaman P, Donegan TM, Prŷs-Jones R (2009) A new subspecies of Brown-banded Antpitta (Grallariidae: *Grallaria milleri*) from Antioquia, Colombia. Bulletin of the British Ornithologists' Club 129(1): 5–17.

Sangster G (2014) The application of species criteria in avian taxonomy and its implications for the debate over species concepts. Biological Reviews 89: 199–214. https://doi.org/10.1111/brv.12051

Sangster G, Luksenburg JA (2015) Declining rates of species described per taxonomist: slowdown of progress or a side-effect of improved quality in taxonomy? Systematic Biology 64: 144–151. https://doi.org/10.1093/sysbio/syu069

Sawilowsky SS (2003) A different future for social and behavioral science research. Journal of Modern Applied Statistical Methods 2: 128–132. https://doi.org/10.22237/jmasm/1051747860

Sonsthagen SA, Wilson RE, Chesser RT, Pons J-M, Crochet P-A, Driskell A, Dove C (2016) Recurrent hybridization and recent origin obscure phylogenetic relationships within the 'white-headed' gull (*Larus* sp.) complex. Molecular Phylogenetics and Evolution 103: 41–54. https://doi.org/10.1016/j.ympev.2016.06.008

Sorenson MD, Sefc KM, Payne RB (2003) Speciation by host switch in brood parasitic indigobirds. Nature 242: 928–931. https://doi.org/10.1038/nature01863

Thomson S and 183 other authors (2018) Taxonomy based on science is necessary for global conservation. PLoS Biology 16(3): e2005075. https://doi.org/10.1371/journal.pbio.2005075

Tobias JA, Seddon N, Spottiswoode CN, Pilgrim JD, Fishpool LDC, Collar NJ (2010) Quantitative criteria for species delimitation. Ibis 152: 724–746. https://doi.org/10.1111/j.1474-919X.2010.01051.x

Toews DPL (2015) Biological species and taxonomic species: Will a new null hypothesis help? (A comment on Gill 2014) Auk 132: 78–81. https://doi.org/10.1642/AUK-14-138.1

Walsh PD (2000) Sample size for the diagnosis of conservation units. Conservation Biology 14: 1533–1537. https://doi.org/10.1046/j.1523-1739.2000.98149.x

Wheeler QD, Meier R (1999) Species Concepts and Phylogenetic Theory: A Debate. Columbia University Press, New York, 256 pp.

Wiens JJ, Servedio MR (2000) Species delimitation in systematics: inferring diagnostic differences between species. Proceedings of the Royal Society of London B 267: 631–636. https://doi.org/10.1098/rspb.2000.1049

Willughby F (1676) Ornithologiae libri tres. J. Martyn, London, 307 pp.

Willughby F (1678) The ornithology of F. Willughby of Middleton. J. Martyn, London, 441 pp.

Zink RM (2003) The role of subspecies in obscuring avian biological diversity and misleading conservation policy. Proceedings of the Royal Society of London B 27: 561–564.

ZooKeys 757: 69–83 (2018)
doi: 10.3897/zookeys.757.21900
http://zookeys.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
ZooKeys
Launched to accelerate biodiversity research

Influence of habitat heterogeneity on anuran diversity in Restinga landscapes of the Parnaíba River delta, northeastern Brazil

Kássio C. Araújo[1], Anderson Guzzi[3], Robson W. Ávila[1,2]

1 *Programa de Pós-Graduação em Ecologia e Recursos Naturais, Bloco 902, Centro de Ciências, Universidade Federal do Ceará – UFC, Campus do PICI, Av. Humberto Monte, s/n, 60455-760 Fortaleza, Ceará, Brazil* **2** *Programa de Pós-Graduação em Bioprospecção Molecular, Universidade Regional do Cariri, Campus do Pimenta, Rua Cel. Antônio Luiz, 1161, Bairro do Pimenta, 63105-000, Crato, Ceará, Brazil* **3** *Curso de Ciências Biológicas, Centro de Ciências do Mar, Universidade Federal do Piauí – UFPI, Campus Ministro Reis Veloso, Av. São Sebastião, 2819, Planalto Horizonte, 64202-020, Parnaíba, Piauí, Brazil*

Corresponding author: *Kássio C. Araújo* (kassio.ufpi@gmail.com)

Academic editor: *A. Crottini* | Received 28 October 2017 | Accepted 9 April 2018 | Published 10 May 2018

http://zoobank.org/C0F6C60D-24BD-4FCE-9533-7E76A2D0FB69

Citation: Araújo KC, Guzzi A, Ávila RW (2018) Influence of habitat heterogeneity on anuran diversity in Restinga landscapes of the Parnaíba River delta, northeastern Brazil. ZooKeys 757: 69–83. https://doi.org/10.3897/zookeys.757.21900

Abstract

Anurans have close associations with environmental conditions and therefore represent an interesting vertebrate group for examining how resource availability and environmental variables influence species diversity. Associations between habitat heterogeneity and anuran species diversity were tested in the Restinga landscapes of the Parnaíba River delta in northeastern Brazil. Twenty-one anuran species were sampled in the rainy season during monthly excursions (December 2015 to June 2016) into areas of Restinga on two islands in the Parnaíba River delta. The fourth highest anuran diversity was found in this type of environment in Brazil and is the third in northeastern Brazil. Microenvironments, characterized by a combination of vernal pools with different vegetational and physical structures, better explained anuran species composition in the Parnaíba River delta.

Keywords

Amphibians, heterogeneity, Parnaíba River delta

Introduction

Scientists have long attempted to explain species distribution patterns and species richness worldwide, and several ecological hypotheses and theories have been proposed (e.g., Hutchinson 1959, Pianka 1966, MacArthur and Wilson 1967, Huston 1979, Hubbell 2001, Tjørve et al. 2008), including the habitat heterogeneity hypothesis of MacArthur and MacArthur (1961), which proposed that heterogeneous environments improve species richness by allowing species coexistence.

The habitat heterogeneity hypothesis has since been used to explain distribution patterns and species richness throughout the world (e.g., Atauri and Lucio 2001, Tews et al. 2004, Bastazini et al. 2007, González-Megias et al. 2007, Vasconcelos et al. 2009, Silva et al. 2010, Jimenez-Alfaro et al. 2016). Several studies in Brazil have shown a close relationship between environmental heterogeneity and amphibian diversity, although those studies have been largely concentrated in the Amazon rain forest and Atlantic Forest (Keller et al. 2009, Vasconcelos et al. 2009, Silva et al. 2011). Studies in open formations in Brazil, such as in the morphoclimatic domains Tropical Atlantic, Caatingas, and Cerrados (see Ab'Sáber 1977 for definition of morphoclimatic domains), have been scarce (e.g., Bastazini et al. 2007, Xavier and Napoli 2011, Dória et al. 2015, respectively).

Although the habitat heterogeneity hypothesis of MacArthur and MacArthur (1961) is well understood, the measurement of this heterogeneity is difficult due to the close connection with resources variety and availability. Thus, resources such as size of water pond may be important to amphibian richness, as predicted by species-area relationship in the Islands Biogeography theory of MacArthur and Wilson (1967). In addition, duration and depth of water pond is important for amphibian reproductive success especially in regions with irregular rainfall (Becker et al. 2007).

Vegetation structure in and around water bodies is an important resource for local diversity of anurans (Bastazini et al. 2007, Dória et al. 2015) by providing conditions of more reproductive modes (Andrade et al. 2016). Amphibians are strongly influenced by environmental conditions (Duellman and Trueb 1994) and, therefore, represent an interesting vertebrate group to investigate how resource availability can influence species diversity.

The Parnaíba River delta in northeastern Brazil is dominated by Restinga coastal vegetation with sandy soils and open herbaceous, shrubby, and arboreal plant formations (Silva and Britez 2005, Santos-Filho et al. 2010, Santos-Filho et al. 2015, Serra et al. 2016) with approximately 363 known plant species belonging to 74 families (Santos-Filho et al. 2015) – indicating high local heterogeneity. The relationship of this presumed heterogeneity with anuran diversity in the Parnaiba River Delta, however, remains unknown (Andrade et al. 2016, Andrade et al. 2014, Andrade et al. 2012, Loebmann and Mai 2008). The present study aimed to test the influence of habitat heterogeneity on anurans diversity in the Restinga landscapes of the Parnaíba River Delta.

Materials and methods

Study area: The Parnaíba River Delta is contained within an Environmental Protection Area (EPA) created in August 1966, covering approximately 313,800 ha in the Brazilian states of Piauí, Maranhão, and Ceará (Fig. 1) (Brasil 2002). The region is composed of a transitional area between Caatinga and Cerrado formations and marine systems (Brasil 2002). The predominant physiognomy is the Restinga environment, quaternary habitats characterized by sandy soils with high salt concentrations covered predominantly by herbaceous and shrubby xerophytic vegetation (see Xavier et al. 2015 for the definition of a Restinga). Rainfall is concentrated mainly from January through May (IBAMA 1998).

Sampling: Amphibians were collected in areas of Restinga from two islands in the Parnaíba River Delta: Ilha Grande de Santa Isabel Island in the state of Piauí (2°52'27"S, 41°47'20"W, WGS84 datum, 5 m a.s.l.) and Canárias Island in the state of Maranhão (2°48'09"S, 41°52'19"W, WGS84 datum, 8 m a.s.l.). First we selected the areas of Restinga in Parnaíba River Delta according to the soil type (Embrapa Solos UEP Recife 2006). We then selected three Restinga landscapes covering approximately 10 km² from these areas. Using the ArcToolbox (Create Random Points) function from the software ARCGIS, version 9.3 (ESRI 2008), two plots of 1 km² were ran-

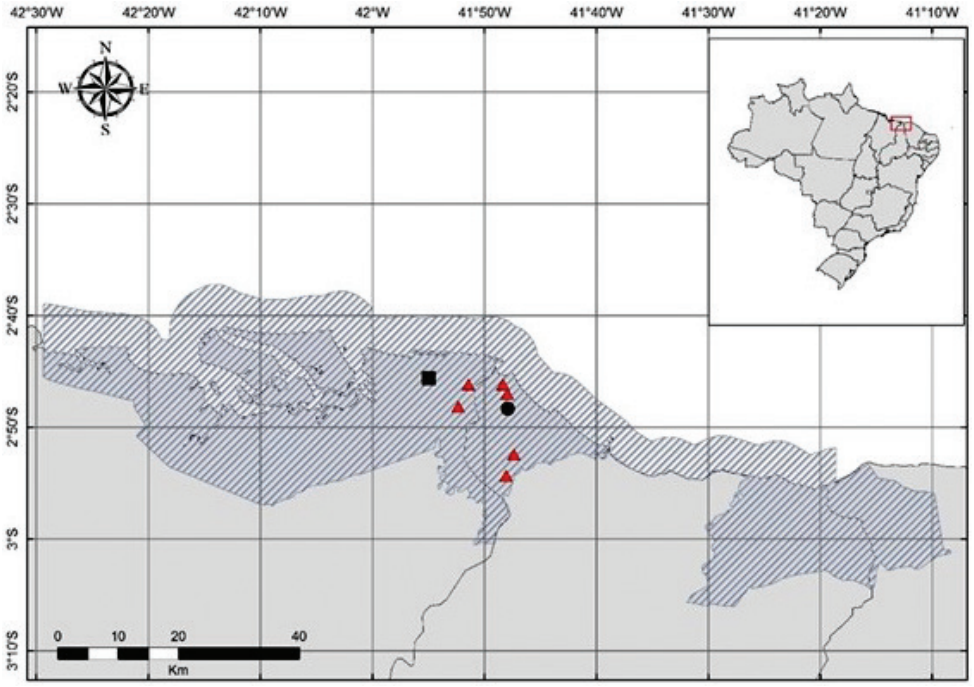

Figure 1. Map of the Environmental Protection Area of Parnaíba River Delta (shaded area), northeastern Brazil, with the location of the study area featuring six sampling points (red triangles). Key: black square, Canárias Island, state of Maranhão; black circle, Ilha Grande de Santa Isabel Island, state of Piauí.

domly chosen in each landscape as sampling points. The three Restinga landscapes and the sampling points were marked using a C7 GPS, version 1.0.

Anuran sampling was undertaken monthly on consecutive days during the rainy season (from December 2015 to June 2016) employing visual searches (Crump and Scott Jr 1994). Our sampling effort was approximately 336 hours/4 researchers. Vouchered specimens were deposited in the amphibians' collection of the Universidade Regional do Cariri (URCA) and Universidade Federal do Piauí (UFPI) (Appendix 1). Anuran nomenclature follows Frost (2017). The species were identified according to literature and comparisons of specimens deposited in the amphibians' collection of URCA and UFPI.

Habitat heterogeneity was quantified using seven environmental descriptors adapted from Santos et al. (2007). Values from 1 to 4 were ascribed for each environmental descriptor, being 4 the highest heterogeneity local indicator. The habitat heterogeneity of each sampling point was then quantified using the mean values of environmental descriptors (Table 1). he mean value was used to give the same importance for each environmental descriptor.

Species distributions and associations with Brazilian morphoclimatic domains (Ab'Sáber 1977) were obtained from literature records (Bastazini et al. 2007, Valdujo et al. 2012, Roberto et al. 2013, Gondim-Silva et al. 2016). Species that occurs in the four Brazilian morphoclimatic domains were considered of wide distribution (Appendix 1).

Statistical analyses: the SHANNON-WIENER diversity index and EQUITY OF PIELOU (Krebs 2000) were used to measure anuran diversity; the estimator CHAO 1, which uses the number of rare species to estimate species richness of a community (Chao 1984, Colwell and Coddington 1994), was used to estimate the expected richness of amphibians. The BERGER-PARKER index (d) was used as a measure of species dominance, using Vegan package (Oksanen et al. 2016). We then produced sample-

Table 1. Main characteristics of the six sampling points in the Parnaíba River Delta: duration (in months) of the water pond (MWP), size (in meters) of water pond (SWP), depth (in centimeters) of water pond (DWP), approximate percentage of vegetation cover on water surface (PVC), types of vegetation within water (TVI), number of types of marginal vegetation (TMV) and types of margin (TM). Types of vegetation: herbaceous and macrophytes (HM), shrub (SH), and arboreal (AB). Types of margin: plans (MP), inclined (MI), and plan and inclined (MPI). Locality (LC) of the sampling points: Ilha Grande de Santa Isabel Island (ILG) and Canárias Island (ILC). In parentheses, the value of each environmental descriptors (1–4). Mean (Mean values of environmental descriptors).

	Point I	**Point II**	**Point III**	**Point IV**	**Point V**	**Point VI**
LC	ILG	ILG	ILG	ILG	ILC	ILC
MWP	5–8 (2)	5–8 (2)	5–8 (2)	1–5 (1)	1–5 (1)	1–5 (1)
SWP	300 (2)	700 (3)	400 (2)	300 (2)	300 (2)	400 (2)
DWP	> 61 (3)	> 61 (3)	> 61 (3)	31–50 (2)	31–50 (2)	> 61 (3)
PVC	76–100 (4)	76–100 (4)	31–50 (2)	31–50 (2)	31–50 (2)	31–50 (2)
TVI	HM (2)	HM (2)	HM (2)	HM (2)	HM (2)	HM (2)
TMV	AB (3)	AB (3)	AB (3)	AB (3)	AB (3)	AB (3)
TM	MPI (2)	MPI (2)	MPI (2)	MP (1)	MPI (2)	MPI (2)
Mean	2.57	2.71	2.28	1.85	2	2.14

based accumulation curves with 1000 sampling randomizations, using ESTIMATE S VERSION 9.1 software (Colwell 2013) to verify if the sampling effort was sufficient to adequately represent the species community.

The normal distribution assumption was tested for both diversity and habitat heterogeneity data using the SHAPIRO-WILK test (Shapiro and Wilk 1965), at each sampling point, and was not rejected (diversity p-value = 0.5653 and habitat heterogeneity p-value = 0.8006). A linear regression analysis was used to test the influence of habitat heterogeneity (independent variable) on anurans diversity (dependent variable) (null hypothesis of no association between anuran diversity and habitat heterogeneity). All statistical analyses were performed in R software (R Development Core Team 2011), using Vegan package (Oksanen et al. 2016).

Results

1822 anuran specimens were recorded, belonging to six families (Bufonidae, Hylidae, Leptodactylidae, Microhylidae, Odontophrynidae, and Phyllomedusidae), 12 genera, and 21 species (see Appendix 1 and 2).

The most abundant species belonged to the families Leptodactylidae and Hylidae (Fig. 2), and they also showed the highest BERGER-PARKER dominance values (d): *Pseudopaludicola mystacalis* (d = 0.14), *Leptodactylus macrosternum* (d = 0.13), *Pleurodema diplolister* (d = 0.12), *Leptodactylus fuscus* (d = 0.11), and *Dendropsophus nanus* (d = 0.10). The CHAO 1 species richness estimator was 21.5 ± 3 species in the Parnaíba River Delta; 18 ± 1 species in Ilha Grande de Santa Isabel Island and 14 ± 3 species in Canárias Island (Tab. 2).

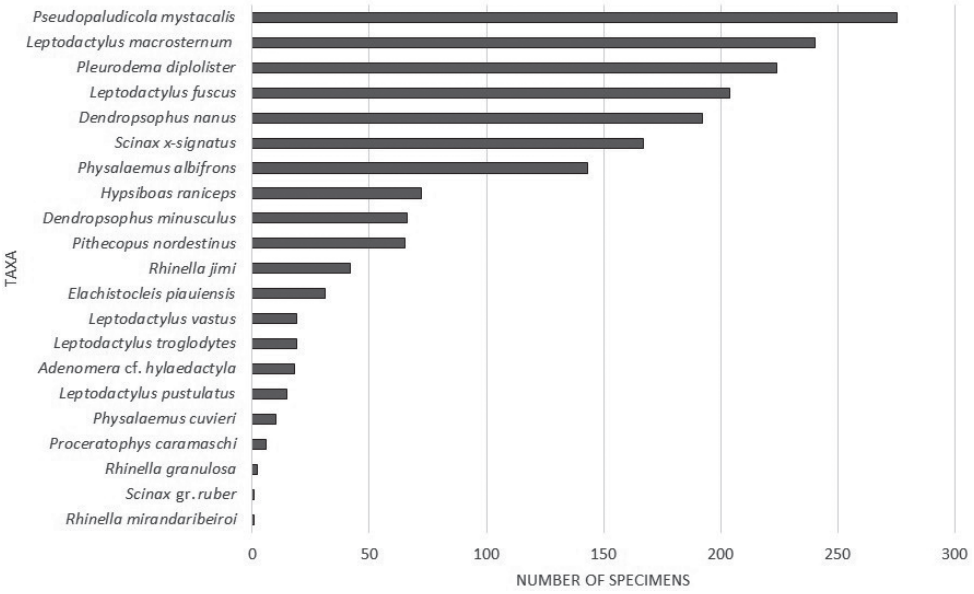

Figure 2. Abundance of anurans species obtained in Ilha Grande de Santa Isabel Island and Canárias Island, Parnaíba River Delta, Northeastern Brazil.

Table 2. Anuran diversity in the Parnaíba River Delta (PRD), Ilha Grande de Santa Isabel Island (ILG) and Canárias Island (ILC), with data on species richness (CHAO 1 species richness estimator), dominant species (BERGER-PARKER index) and evenness (PIELOU's index J').

	PRD	ILG	ILC
Number of individuals	1822	1465	357
Species richness (observed)	21	18	14
Species richness (estimated)	21.5 ± 3	18 ± 1	14 ± 3
Dominant species	*P. mystacalis*	*P. mystacalis*	*L. fuscus*
Dominance observed	14%	14%	19%
Shannon - Wiener (H')	2.485	2.476	2.185
Pielou's index J'	0.8165	0.8569	0.8282

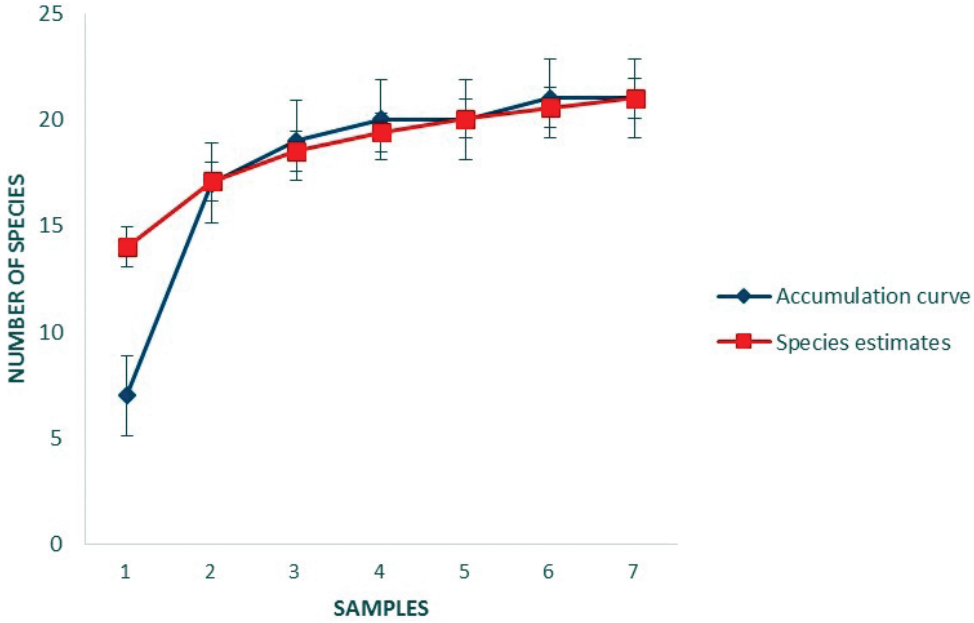

Figure 3. Accumulation curve for anurans sampled in the Parnaíba River Delta, northeastern Brazil, constructed from 1000 randomizations on the order of samplings. Species estimates (Chao 1 estimator).

The sample-based accumulation curve tended asymptote (Fig. 3), which suggest that the sampling effort was sufficient to adequately represent the species community in Parnaíba River Delta, northeastern Brazil.

The species richness at the six sampling points varied from 8 to 17 (Tab. 2). The highest values of species diversity were recorded at points II, I and III, respectively, while point IV had the lowest diversity value. The highest values of habitat heterogeneity were observed at points II, I, and III, respectively, all located in Ilha Grande de Santa Isabel Island. Points V and VI showed intermediated values, while point IV had the lowest habitat heterogeneity value (Tab. 3). The combination of all environmental descriptors is the reason for different heterogeneity indexes in present study.

Table 3. Habitat heterogeneity, SHANNON-WIENER diversity index, evenness (Pielou's index J') and habitat heterogeneity value for each sampling point in the Parnaíba River Delta.

Sampled points	Diversity index	Pielou's index J'	Heterogeneity
Point I	H' = 2.279	J' = 0.8637	He = 2.57
Point II	H' = 2.467	J' = 0.8708	He = 2.71
Point III	H' = 2.220	J' = 0.8935	He = 2.28
Point IV	H' = 1.768	J' = 0.8502	He = 1.85
Point V	H' = 1.815	J' = 0.7881	He = 2
Point VI	H' = 2.052	J' = 0.8557	He = 2.14

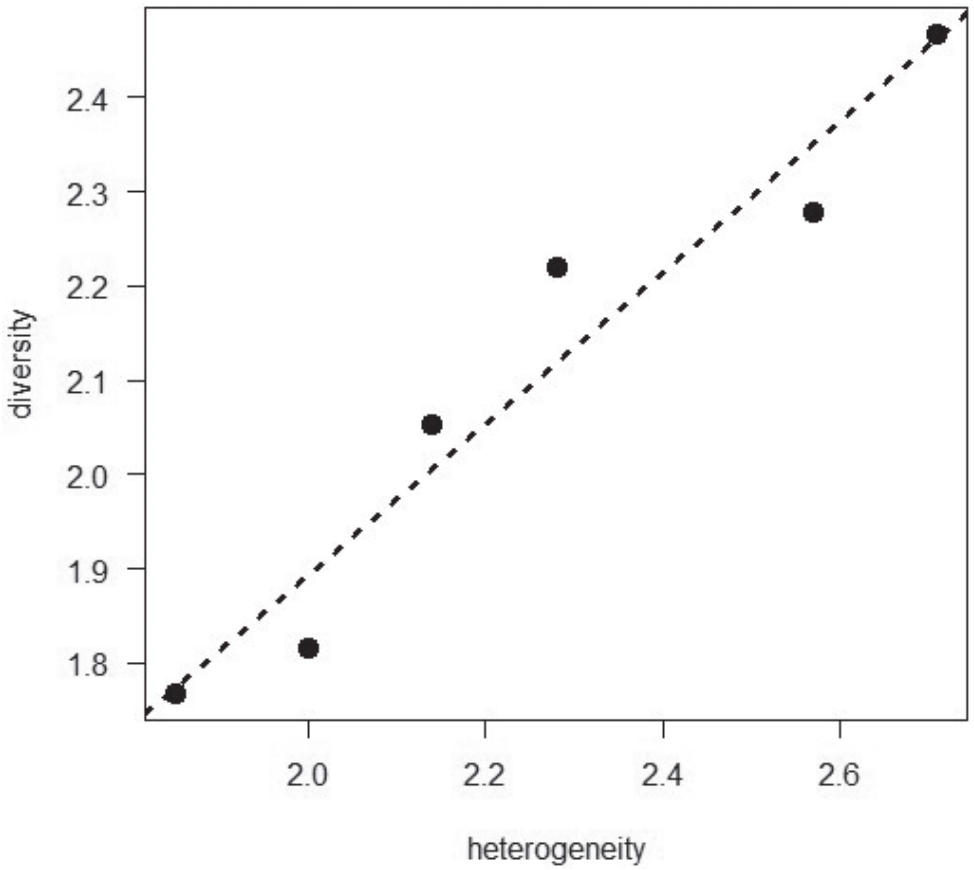

Figure 4. Association between anurans' species diversity (SHANNON-WIENER diversity index) and habitat heterogeneity in the Parnaíba River Delta, Northeastern Brazil (R^2 = 0.9204, p-value = 0.0015). Computation of the habitat heterogeneity index is explained in Material and methods.

The linear regression analysis evidenced that the habitat heterogeneity of the Restinga environment in the Parnaíba River Delta is able to explain the anuran diversity (R^2 = 0.9204, p = 0.0015) (Fig. 4).

Discussion

The Restinga of the Parnaíba River Delta have the fourth highest anuran richness in this type of environment in Brazil and the third in northeastern Brazil (21 species). The Restinga areas with the highest anuran diversity were encountered in the municipalities of Mata de São João (34 species; Bastazini et al. 2007, Oliveira and Rocha 2015, Xavier et al. 2015) and Conde (33 species; Gondim-Silva et al. 2016), both in the state of Bahia, and in the municipality of Grumari, Rio de Janeiro state (22 species; Telles et al. 2012).

The anuran species composition of the Parnaíba River Delta was similar to that reported by Borges-Leite et al. (2014) and Gondim-Silva et al. (2016) for the municipalities of São Gonçalo do Amarante and Conde, in the Brazilian states of Ceará and Bahia, respectively. The aforementioned study in the Ceará state was carried out in an ecotonal environment with floristic elements of Caatinga, Cerrado and Restinga (Borges-Leite et al. 2014) and the study in Bahia state included only "Open Restinga" (Gondim-Silva et al. 2016). The similarity between the present work and these studies could be explained by the presence of floristic elements of Caatinga, Cerrado and Restinga in our open Restinga area studied.

Nevertheless the Restinga of the Parnaíba River Delta differed greatly from Restinga sites in southeastern Brazil (states of São Paulo, Rio de Janeiro and Espírito Santo; Rocha et al. 2008, Silva et al. 2008, Vilela et al. 2011, Telles et al. 2012) and other regions of the state of Bahia (municipalities of Prado, Trancoso and Mata de São João; Bastazini et al. 2007, Rocha et al. 2008, Narvaes et al. 2009). These studies included lowland forests that can be very important for explaining the differences in anuran composition between them. The high habitat heterogeneity in Restinga environments (Gomes et al. 2016), however, could also account for those differences.

Increased habitat structural complexity results in greater species diversity (MacArthur and MacArthur 1961), with homogeneous areas showing less microhabitat availability, which hampers species coexistence and resource partitioning (Macarthur and Levins 1967). Highly heterogeneous environments promote higher species richness by promoting the coexistence, persistence, and diversification of species at different spatial and temporal scales (Stein and Kreft 2014).

Positive relationships between habitat heterogeneity and anuran diversity have been recorded in different morphoclimatic domains in Brazil, as well in the present study. Habitat heterogeneity has been shown to influence anuran diversity in Restinga areas in northeastern Brazil (Bastazini et al. 2007), in "Campo rupestre" vegetation in the Caatinga (Xavier and Napoli 2011), and in Cerrado vegetation with a predominance of semi-deciduous seasonal forest (Dória et al. 2015). A clear relationship between habitat heterogeneity and anuran diversity was recorded in the Atlantic Forest (Lop et al. 2012, Santos et al. 2012) as did Silva et al. (2011) in pasture areas, both in southeastern Brazil. Some studies, however, could not identify relationships between habitat heterogeneity and anuran diversity (Eterovick 2003, Vasconcelos and Rossa-Feres 2005, Santos et al. 2007), and more studies will consequently be necessary to elucidate the importance of environment heterogeneity to species diversity.

Anuran populations from the Restinga of the Parnaíba River Delta are influenced by habitat complexity and the variety of available microhabitats, in agreement with Bastazini et al. (2007) who highlighted the importance of shrub formations and bromeliad densities to explain changes in anuran composition in Restinga environments.

Earlier studies highlighted the importance of pond size and edge vegetation to anuran diversity (Parris and McCarthy 1999, Burne and Griffin 2005, Bastazini et al. 2007, Vieira et al. 2007, Xavier and Napoli 2011, Dória et al. 2015, Gonçalves et al. 2015). Furthermore, microenvironments composed of vernal pools with different edge vegetation structures and percentage of vegetation cover on water surface better explained the different compositions of anuran communities in Parnaíba River Delta.

Acknowledgments

We are grateful to the Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq for the research grant awarded to R.W.A (PQ # 303622/2015-6); Tássia G. P. Lima, Ocivana A. Pereira, and Arthur S. N. Ribeiro for their assistance during fieldwork and Marcelo F. Napoli for the valuable comments on the manuscript. We are grateful to the CNPq/FUNCAP/CAPES for financial support (PROTAX – Processes 440511/2015-1; 5574685/2017; 88882.156872/2016-01); and ICMBio for the collection license (Permit: ICMBio 7770-1, process: 29613). This study was approved by the Animal Ethics Committee of the Universidade Regional do Cariri (Comissão de experimentação e Uso de Animais – CEUA/URCA: process 00026/2015.2).

References

Ab'Sáber AN (1977) Os domínios morfoclimáticos da América do Sul. Primeira Aproximação. Geomorfologia 53(1): 1–52.

Andrade EB, Leite JRSA, Andrade GV, Macedo RG (2012) Diversidade de Anfíbios do Delta do Parnaíba. In: Guzzi A (Org.). Biodiversidade do Delta do Parnaíba. EDUFPI, Parnaíba, 209–269.

Andrade EB, Leite JRSA, Andrade GV (2014) Anurans from the municipality of Ilha Grande, Parnaíba River Delta, Piauí, northeastern Brazil. Herpetology Notes 7: 219–226.

Andrade EB, Leite JRSA, Andrade GV (2016) Diversity and distribution of anuran in two islands of Parnaíba River Delta, Northeastern Brazil. Journal of Biodiversity and Environmental Sciences 8(2): 74–86.

Atauri JA, Lucio JV (2001) The role of landscape structure in species richness distribution of birds, amphibians, reptiles and lepidopterans in Mediterranean landscapes. Landscape Ecology 16: 147–159. https://doi.org/10.1023/A:1011115921050

Bastazini CV, Munduruca JFV, Rocha PLB, Napoli MF (2007) Which environmental variables better explain changes in anuran community composition? A case study in the restinga of Mata de São João, Bahia, Brazil. Herpetologica 63(4): 459–471. https://doi.org/10.1655/0018-0831(2007)63[459:WEVBEC]2.0.CO;2

Becker CG, Fonseca CR, Haddad CFB, Batista RF, Prado PI (2007) Habitat split and the global decline of amphibians. Science 318(1): 1775–1777. https://doi.org/10.1126/science.1149374

Brasil (2002) Zoneamento Ecológico-Econômico do Baixo Parnaíba: primeira aproximação. MMA, Brasília. [CD-ROM]

Borges-Leite ML, Rodrigues JFM, Borges-Nojosa DM (2014) Herpetofauna of a coastal region of northeastern Brazil. Herpetology notes 7(1): 405–413.

Burne MR, Griffin CR (2005) Habitat associations of poolbreeding amphibians in eastern Massachusetts, USA. Wetlands Ecology Management 13(3): 247–259. https://doi.org/10.1007/s11273-004-7519-4

Chao A (1984) Nonparametric estimation of the number of classes in a population. Scandinavian Journal of Statistics 11(4): 265–270.

Colwell RK, Coddington JA (1994) Estimating terrestrial biodiversity through extrapolation. Philosophical Transactions of the Royal Society of London. Series B 345(1): 101–118. https://doi.org/10.1098/rstb.1994.0091

Colwell RK (2013) Estimates: statistical estimation of species richness and shared species from samples. Version 9.3. User's Guide and application published at: http://viceroy.eeb.uconn.edu/estimates

Crump ML, Scott Jr NJ (1994) Visual encounter surveys. In: Measuring and Monitoring Biological Diversity – Standard Methods for Amphibians. Heyer, WR, Donnelly MAR, McDiarmid W, Hayek LAC, Foster MS (Eds.), Smithsonian Institution Press, Washington, 84–91.

Dória TAF, Klein W, Abreu RO, Santos DC, Cordeiro MC, Silva LM, Bonfim VMG, Napoli MF (2015) Environmental variables influence the composition of frog communities in riparian and semi-deciduous forests of the Brazilian Cerrado. South American Journal of Herpetology 10(2): 90–103. https://doi.org/10.2994/SAJH-D-14-00029.1

Duellman WE, Trueb L (1994) Biology of Amphibians. The Johns Hopkins University Press, Baltimore, 670 pp.

Embrapa Solos UEP Recife (2006) Electronic Database accessible at: http://www.uep.cnps.embrapa.br/solos/index.html [Access on: 10 Jan of 2017]

ESRI (2008) ArcGIS Desktop: version 9.3. Redlands, CA: Environmental Systems Research Institute. Electronic Database accessible at: http://www.esri.com/software/arcgis/index.html [Access on: 10 Jan of 2017]

Eterovick PC (2003) Distribution of anuran species among montane streams in southeastern Brazil. Journal of Tropical Ecology 19: 219–228. https://doi.org/10.1017/S0266467403003250

Frost DR (2017) Amphibian Species of the World: an Online Reference. Version 6.0 (10 Jan of 1017). Electronic Database accessible at http://research.amnh.org/herpetology/amphibia/index.html. American Museum of Natural History, New York, USA.

Gomes VSM, Vecchi MB, Loiselle BA, Missagia CCC, Alves MAS (2016) Heterogeneity of bird communities in a mosaic of habitats on a resting ecosystem in southeast Brazil. Zoologia 33(5): 1–13. https://doi.org/10.1590/S1984-4689zool-20150205

Gonçalves DS, Crivellari LB, Conte CE (2015) Linking environmental drivers with amphibian species diversity in ponds from subtropical grasslands. Annals of the Brazilian Academy of Sciences 87(3): 1751–1762. https://doi.org/10.1590/0001-3765201520140471

Gondim-Silva FAT, Andrade ARS, Abreu RO, Nascimento JS, Corrêa GP. Menezes L, Trevisan CC, Camargo SS, Napoli MF (2016) Composition and diversity of anurans in the Restinga of the Conde municipality, northern coast of the state of Bahia, northeastern Brazil. Biota Neotropica 16(3): 1–16. https://doi.org/10.1590/1676-0611-BN-2016-0157

González-Megias A, Gómes JG, Sánchez-Piñero F (2007) Diversity-habitat heterogeneity relationship at different spatial and temporal scales. Ecography 30: 31–41. https://doi.org/10.1111/j.0906-7590.2007.04867.x

Hubbell SP (2001) The unified neutral theory of biodiversity and biogeography. Princeton University Press, New Jersey, 396 pp.

Huston MA (1979) General hypothesis of species diversity. The American Naturalist 113(1): 81–101. https://doi.org/10.1086/283366

Hutchinson GE (1959) Homage to Santa Rosalia or why are there so many kind of animals? The American Naturalist 93(870): 145–159. https://doi.org/10.1086/282070

IBAMA (1998) Proteção e controle de ecossistemas costeiros: manguezais da Baía de Babitonga/ Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis. Coleção meio ambiente. Série estudos – pesca, 25. Brasília, 145 pp.

Jimenez-Alfaro B, Chytrý M, Mucina L, Grace JB, Rejmánek M (2016) Disentangling vegetation diversity from climate-energy and habitat heterogeneity for explaining animal geographic patterns. Ecology and Evolution 6(5): 1515–1526. https://doi.org/10.1002/ece3.1972

Keller A, Rodel MO, Linsenmair KE, Grafe TU (2009) The importance of environmental heterogeneity for species diversity and assemblage structure in Bornean stream frogs, Journal of Animal Ecology 78: 305–314. https://doi.org/10.1111/j.1365-2656.2008.01457.x

Krebs CJ (2000) Ecological Methodology. 2 ed. Harper and Row Publishers, New York.

Loebmann D, Mai ACG (2008) Amphibia, Anura, Coastal Zone, Piauí state, Northeastern Brazil. Check List 4: 161–170. https://doi.org/10.15560/4.2.161

Lop S, Caldart VM, Santos TG, Cechin SZ (2012) What is the role of heterogeneity and spatial autocorrelation of ponds in the organization of frog communities in Southern Brazil? Zoological Studies 51(7): 1094–1104.

Macarthur RH, Macarthur JW (1961) On bird species diversity. Ecology 42: 594–598. https://doi.org/10.2307/1932254

Macarthur RH, Levins R (1967) The limiting similarity, convergence, and divergence of coexisting species. The American Naturalist 101(921): 377–385. https://doi.org/10.1086/282505

Macarthur RH, Wilson EO (1967) The Theory of Island Biogeography. Princeton University Press, Princeton, NJ, 215 pp.

Narvaes P, Bertoluci J, Rodrigues MT (2009) Species composition, habitat use and breeding seasons of anurans of the resting forest of the Estação Ecológica Juréia-Itatins, Southeastern Brazil. Biota Neotropica 9(2): 117–123. https://doi.org/10.1590/S1676-06032009000200011

Oliveira JCF, Rocha CFD (2015) A review on the anurofauna of Brazil's sandy coastal plains. How much do we know about it? Journal of Coastal Conservation 19: 35–49. https://doi.org/10.1007/s11852-014-0354-8

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2016) Vegan: Community Ecology Package. R package version 2.3-3. https://CRAN.R-project.org/package=vegan [Last access in 15/12/2017]

Parris KM, Mccarthy MA (1999) What influences the structure of frog assemblages at forest streams? Austral Journal of Ecology 24(5): 495–502. https://doi.org/10.1046/j.1442-9993.1999.00989.x

Pianka ER (1966) Latitudinal gradients in species diversity: A review of concepts. The American Naturalist 100(910): 33–46. https://doi.org/10.1086/282398

R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/

Roberto IJ, Ribeiro SC, Loebmann D (2013) Amphibians of the state of Piauí, Northeastern Brazil: a preliminary assessment. Biota Neotropica 13(1): 322–330. https://doi.org/10.1590/S1676-06032013000100031

Rocha CFD, Hatano FH, Vrcibradic D, Van Sluys M (2008) Frog species richness, composition β-diversity in coastal Brazilian restinga habitats. Brazilian Journal of Biology 68(1): 101–107. https://doi.org/10.1590/S1519-69842008000100014

Santos TG, Rossa-Feres DC, Casatti L (2007) Diversidade e distribuição espaço-temporal de anuros em região com pronunciada estação seca no sudeste do Brasil. Iheringia Série Zoologia 97(1): 37–49. https://doi.org/10.1590/S0073-47212007000100007

Santos-Filho FS, Almeida Jr EB, Soares CJRS, Zickel CS (2010) Fisionomias das restingas do Delta do Parnaíba, Nordeste, Brasil. Revista Brasileira de Geografia Física 3: 218–227. https://doi.org/10.26848/rbgf.v3i3.232605

Santos TG, Vasconcelos TS, Haddad CFB (2012) The role of environmental heterogeneity in maintenance of anuran amphibian diversity of the Brazilian Mesophytic Semideciduous Forest. Tropical Forests 1: 119–138. https://doi.org/10.5772/31241

Santos-Filho FS, Almeida Jr EB, Lima PB, Soares CJRS (2015) Checklist of the flora of the restingas of Piauí state, Northeast Brazil. Check List 11(2): 1–10. https://doi.org/10.15560/11.2.1598

Serra FCV, Lima PB, Almeida Jr EB (2016) Species richness in restinga vegetation on the eastern Maranhão State, Northeastern Brazil. Acta Amazonica 46(3): 271–280. https://doi.org/10.1590/1809-4392201504704

Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52(3–4): 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Silva SM, Britez RM (2005) A vegetação da planície costeira. In: Marques MCM, Britez RM (orgs.). História Natural e Conservação da Ilha do Mel. Curitiba: Ed. UFPR, 49–84.

Silva HR, Carvalho ALG, Bittencourt-Silva GB (2008) Frogs of Marambaia: a naturally isolated Restinga and Atlantic Forest remnant of southeastern Brazil. Biota Neotropica 8(4): 167–174. https://doi.org/10.1590/S1676-06032008000400017

Silva RJ, Diniz S, Vaz-de-melo FZ (2010) Heterogeneidade do habitat, riqueza e estrutura da assembleia de besouros rola-bostas (Scarabaeidae: Scarabaeinae) em áreas de Cerrado na Chapada dos Parecis, MT. Neotropical Entomology 39(6): 934–940. https://doi.org/10.1590/S1519-566X2010000600014

Silva RA, Martins IA, Rossa-Feres DC (2011) Environmental heterogeneity: anuran diversity in homogeneous environments. Zoologia 28(5): 610–618. https://doi.org/10.1590/S1984-46702011000500009

Telles FBS, Menezes VA, Maia-Carneiro T, Dorigo TA, Winck GR, Rocha CFD (2012) Anurans from the "Restinga" of Parque Natural Municipal de Grumari, state of Rio de Janeiro, southeastern Brazil. Check List 8(6): 1267–1273. https://doi.org/10.15560/8.6.1267

Tews J, Brose U, Grimm V, Tielbörger K, Wichmann MC, Schwager M, Jeltsch F (2004) Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. Journal of Biogeography 31: 79–92. https://doi.org/10.1046/j.0305-0270.2003.00994.x

Tjørve E, Kunin WE, Polce C, Tjørve KMC (2008) Species-area relationship: separating the effects of species abundance and spatial distribution. Journal of Ecology 96: 1141–1151. https://doi.org/10.1111/j.1365-2745.2008.01433.x

Valdujo PH, Silvano DL, Colli G, Martins M (2012) Anuran species composition and distribution patterns in Brazilian Cerrado, a Neotropical hotspot. South American Journal of Herpetology 7(2): 63–78. https://doi.org/10.2994/057.007.0209

Vasconcelos TS, Rossa-Feres DC (2005) Diversidade, distribuição espacial e temporal de anfíbios anuros (Amphibia, Anura) na região noroeste do estado de São Paulo, Brasil. Biota Neotropica 5(2): 137–150. https://doi.org/10.1590/S1676-06032005000300010

Vasconcelos TS, Santos TG, Rossa-Feres DC, Haddad CFB (2009) Influence of the environmental heterogeneity of breeding ponds on anuran assemblages from southeastern Brazil. Canadian Journal of Zoology 87(8): 699–707. https://doi.org/10.1139/Z09-058

Vieira WLS, Arzabe C, Santana GG (2007) Composição e distribuição espaço-temporal de anuros no Cariri paraibano, Nordeste do Brasil. Oecologia Brasiliensis 11(3): 383–396. https://doi.org/10.4257/oeco.2007.1103.08

Vilela VMFN, Brassaloti RA, Bertoluci J (2011) Anuran fauna of the restinga forest of the Parque Estadual da Ilha do Cardoso, Southeastern Brazil: species composition and breeding site utilization. Biota Neotropica 11(1): 83–93. https://doi.org/10.1590/S1676-06032011000100008

Xavier AL, Napoli MF (2011) Contribution of environmental variables to anuran community structure in the Caatinga Domain of Brazil. Phyllomedusa 10(1): 45–64. https://doi.org/10.11606/issn.2316-9079.v10i1p45-64

Xavier AL, Guedes TB, Napoli MF (2015) Biogeography of anurans from the poorly known and threatened coastal sandplains of eastern Brazil. PloS One 10(6): 1–18. https://doi.org/10.1371/journal.pone.0128268

Appendix 1

Anuran species obtained at the Restinga of the Parnaíba River delta, northeastern Brazil. Morphoclimatic domains (Ab'Sáber 1977): Caatingas (CA), Cerrados (CE), Tropical Atlantic (AT), and Equatorial Amazonian (AM). Species that occur in the four Brazilian morphoclimatic domains were considered as having a wide distribution (WD). Species without voucher specimens were represented by photography.

APPTABLE CAPTION

Taxon	Voucher specimens	Sampling points (I–VI)	Morphoclimatic domains
Bufonidae			
Rhinella granulosa (Spix, 1824)	Photographed	I, II	CA, AT
Rhinella jimi (Stevaux, 2002)	Photographed	All sampling points	CA, AT
Rhinella mirandaribeiroi (Gallardo, 1965)	Photographed	V	CE
Hylidae			
Dendropsophus minusculus (Rivero, 1971)	URCA-H12120	I, II, V, VI	WD
Dendropsophus nanus (Boulenger, 1889)	CZDP473	I, II, III, V, VI	WD
Hypsiboas raniceps (Cope, 1862)	URCA-H12115	I, II, III, V	WD
Scinax sp. (gr. *ruber*)	URCA-H12123	II	No information
Scinax x-signatus (Spix, 1824)	Photographed	All sampling points	WD
Leptodactylidae			
Adenomera cf. *hylaedactyla* (Cope, 1868)	URCA-H12125	V	AM, CE, AT
Leptodactylus fuscus (Schneider, 1799)	Photographed	All sampling points	WD
Leptodactylus macrosternum Miranda-Ribeiro, 1926	Photographed	All sampling points	WD
Leptodactylus pustulatus (Peters, 1870)	URCA-H12126	II	CE
Leptodactylus troglodytes Lutz, 1926	CZDP485	II, IV	CA, CE, AT
Leptodactylus vastus Lutz, 1930	Photographed	I, II, III	CA, CE, AT
Physalaemus albifrons Spix, 1824	Photographed	I, II, III, IV, VI	CA, CE, AT
Physalaemus cuvieri Fitzinger, 1826	CZDP470	II	WD
Pleurodema diplolister Peters, 1870	Photographed	I, II, III, IV, VI	CA, CE, AT
Pseudopaludicola mystacalis (Cope, 1887)	URCA-H12118	I, II, III, V	WD
Microhylidae			
Elachistocleis piauiensis Caramaschi and Jim, 1983	URCA-H12124	I, III, VI	CA, CE
Odontophrynidae			
Proceratophrys caramaschii Cruz, Nunes and Juncá, 2012	Photographed	VI	CA
Phyllomedusidae			
Pithecopus nordestinus (Caramaschi, 2006)	Photographed	I, II, III, IV	CA, CE, AT

Appendix 2

Anurans recorded at the Restinga of the Parnaíba River Delta, Northeastern Brazil. In brackets, the vouchered specimen with the acronym of the scientific collection followed by the respective institutional registration number and specimen snout-vent length (SVL) in millimeters. Some species only have photographic records. (A) *Rhinella granulosa*, (B) *R. jimi*, (C) *R. mirandaribeiroi*, (D) *Dendropsophus minusculus* (URCA-H12120, SVL 18.4), (E) *D. nanus* (CZDP473, SVL 19.2), (F) *Hypsiboas raniceps* (URCA-H12115, SVL 62.6), (G) *Scinax* sp. (gr. *ruber*) (URCA-H12123, SVL 20.1), (H) *S. x-signatus*, (I) *Adenomera* cf. *hylaedactyla* (URCA-H12125, SVL 15.8), (J) *Leptodactylus fuscus*, (K) *L. macrosternum*, (L) *L. pustulatus* (URCA-H12126, SVL 41.2), (M) *L. troglodytes* (CZDP485, SVL 43.3), (N) *L. vastus*, (O) *Physalaemus albifrons*, (P) *P. cuvieri* (CZDP470, SVL 24.5), (Q) *Pleurodema diplolister*, (R) *Pseudopaludicola mystacalis* (URCA-H12118, SVL 11.9), (S) *Elachistocleis piauiensis* (URCA-H12124, SVL 30.2), (T) *Proceratophrys caramaschii*, (U) *Pithecopus nordestinus*. Photographs: Kássio C. Araújo and Ocivana A. Pereira.

ZooKeys 757: 85–152 (2018)
doi: 10.3897/zookeys.757.24453
http://zookeys.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
ZooKeys
Launched to accelerate biodiversity research

Filling the BINs of life: Report of an amphibian and reptile survey of the Tanintharyi (Tenasserim) Region of Myanmar, with DNA barcode data

Daniel G. Mulcahy[1], Justin L. Lee[2], Aryeh H. Miller[3], Mia Chand[4], Myint Kyaw Thura[5], George R. Zug[6]

1 *Global Genome Initiative, National Museum of Natural History, Smithsonian Institution, 10th & Constitution Ave., Washington, DC, 20013 USA* **2** *College of Computer, Mathematical and Natural Sciences, University of Maryland, College Park Maryland, 20742 USA* **3** *Department of Biology, University of North Carolina Asheville, Asheville, NC 28804 USA* **4** *College of William & Mary, Williamsburg, Virginia, 23187 USA* **5** *Myanmar Environment Sustainable Conservation (MESC), Yangon, Myanmar* **6** *Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, 20013 USA*

Corresponding author: *Daniel G. Mulcahy* (MulcahyD@si.edu)

Academic editor: *J. Penner* | Received 13 February 2018 | Accepted 9 April 2018 | Published 10 May 2018

http://zoobank.org/559E4F4F-7C35-4380-89D5-BA42A5D38004

Citation: Mulcahy DG, Lee JL, Miller AH, Chand M, Thura MK, Zug GR (2018) Filling the BINs of life: Report of an amphibian and reptile survey of the Tanintharyi (Tenasserim) Region of Myanmar, with DNA barcode data. ZooKeys 757: 85–152. https://doi.org/10.3897/zookeys.757.24453

Abstract

Despite threats of species extinctions, taxonomic crises, and technological advances in genomics and natural history database informatics, we are still distant from cataloguing all of the species of life on earth. Amphibians and reptiles are no exceptions; in fact new species are described nearly every day and many species face possible extinction. The number of described species continues to climb as new areas of the world are explored and as species complexes are examined more thoroughly. The use of DNA barcoding provides a mechanism for rapidly estimating the number of species at a given site and has the potential to record all of the species of life on Earth. Though DNA barcoding has its caveats, it can be useful to estimate the number of species in a more systematic and efficient manner, to be followed in combination with more traditional, morphology-based identifications and species descriptions. Herein, we report the results of a voucher-based herpetological expedition to the Tanintharyi (Tenasserim) Region of Myanmar, enhanced with DNA barcode data. Our main surveys took place in the currently proposed Tanintharyi National Park. We combine our results with photographs and observational data from the Chaung-nauk-pyan forest reserve. Additionally, we provide the first checklist of amphibians and reptiles of the region,

with species based on the literature and museum. Amphibians, anurans in particular, are one of the most poorly known groups of vertebrates in terms of taxonomy and the number of known species, particularly in Southeast Asia. Our rapid-assessment program combined with DNA barcoding and use of Barcode Index Numbers (BINs) of voucher specimens reveals the depth of taxonomic diversity in the southern Tanintharyi herpetofauna even though only a third of the potential amphibians and reptiles were seen. A total of 51 putative species (one caecilian, 25 frogs, 13 lizards, 10 snakes, and two turtles) were detected, several of which represent potentially undescribed species. Several of these species were detected by DNA barcode data alone. Furthermore, five species were recorded for the first time in Myanmar, two amphibians (*Ichthyophis* cf. *kohtaoensis* and *Chalcorana eschatia*) and three snakes (*Ahaetulla mycterizans*, *Boiga dendrophila*, and *Boiga drapiezii*).

Keywords

Anura, biodiversity, Gymnophiona, Thai-Malay Peninsula, natural history, Southeast Asia, species diversity, Squamata, Testudines, Thailand

Introduction

Despite advances in technologies, warnings of taxonomic crises, and increased interest in taxonomy (Mallet and Willmott 2003, Tautz et al. 2003), scientists are still trying to provide an accurate measure of global biodiversity in terms of absolute numbers of extant species of life on Earth (e.g. Costello et al. 2013, Caley et al. 2014, Wilson 2017). Amphibians and reptiles are no exceptions to groups with accurate estimates of extant species because new species are described nearly every day and estimates suggest ca. 30% of amphibians (Stuart et al. 2004) and ca. 20% of reptiles (Böhm et al. 2013) may be threatened. The benefits of knowing and understanding global biodiversity are enormous and span fields of human interest from agriculture, pest management, disease control, natural products, conservation, and wildlife management. Many of the discrepancies in estimates are the result of non-statistical calculations (i.e. "simple best guesses"), statistical calculation that contain wide ranges of error (e.g. "+/- three million"), and estimates that do not build on one another, and overlap with previous analyses (Giller 2014). Furthermore, without careful comparisons of known material (voucher specimens), estimates of unknown species may contain significant overlap with currently recognized species (synonymies). Certainly, in this age of genomics and bioinformatics, we have the ability to accurately measure and record global species diversity with resources like the Encyclopedia of Life (EOL), Tree of Life projects (e.g. http://www.tolweb.org/tree/), the Barcode of Life Database (BOLD), GenBank, and taxonomic specific databases such as Amphibian Species of the World 6.0 (ASoW, Frost 2017) and the Reptile Database (Uetz et al. 2018). To confound matters, there have been recent requests to regulate and standardize practices in taxonomy (Garnett and Christidis 2017), which has stirred debate on the theoretical aspects of species and the practicality of regulating 'taxonomic freedom' (Raposo et al. 2017). While issues of regulating taxonomic actions remain unresolved, DNA barcoding (Hebert et al. 2003a) offers a standardized mechanism for measuring biodiversity at the spe-

cies level and a database to manage it (BOLD). However, DNA taxonomy has many caveats and limitations, such as proposed thresholds of percent sequence divergence vary among groups, and it is less effective among recently diverged groups (Lipscomb et al. 2003; Tautz et al. 2003), and we are still a long way from obtaining a complete DNA barcode library of life of Earth. Currently, the BOLD database contains DNA barcodes for approximately 275,000 formally described species of fungi, plants, and animals (http://www.boldsystems.org/index.php, queried 11 February 2018), of the approximately 1.5 million catalogued (only ~18%), and of the ~7 million estimated (< 3%) species (Caley et al. 2014).

The use of DNA barcoding offers a starting point for recording the number of species of life on Earth (Hebert et al. 2003a, 2000b). The concept of a database containing representatives of every species with a common molecular marker, shared among all living organisms (though different for plants and animals) for comparison is attractive. Once a reference library is established, it provides researchers with resources of numerous possibilities, ranging from agriculture, ecological and environmental studies, biodiversity surveys, conservation, food and drug administration, and the prevention of wildlife trafficking. In animals, the DNA barcode is a portion of the mitochondrial DNA gene cytochrome oxidase subunit 1 (COI); different markers are used for plants (chloroplast DNA), fungi, etc. Furthermore, a system of Barcode Index Numbers (BINs) has been developed to assist in specimen identifications, by using several algorithms to compare COI data, combining sequences into operational taxonomic units (OTUs), which likely correspond to biological species (Ratnasingham and Hebert 2013). Investigators can quickly compare COI sequences in a database (BOLD) and rapidly determine whether their samples are unique, or similar to described, or even un-described species, thus eliminating or reducing the number of synonymies in species estimates. For example, if one identifies a species as "sp. A" and if they, or someone else identifies other individuals of the same species and refers to them as "sp. B," this creates a synonymy. This is why it is important to compare newly acquired material with all available data, and to maintain current usage of placeholding names. For example, a researcher might identify a specimen to be the same species as "sp. A" of another study (e.g. Diechmann et al. 2017). By DNA barcoding newly discovered species, one can quickly verify its degree of difference via the BINs, only if other closely related species are also barcoded. The BINs are automatically generated if the sequences are deposited in BOLD, and new sequences will be placed in existing BINs if within ~2%, or new BINs will be created; discordant BINs (e.g. a single BIN with specimens bearing different names) are flagged and easily identified. However, there are several caveats to consider when evaluating whether a BIN represents a legitimate species, or whether a BIN is discordant because of disagreements on higher-level classification (e.g. constantly changing generic names). Additionally, some wide-ranging, genetically variable species may occupy multiple BINs. A point worth noting is that the BINs are not formally named (i.e., they do not bear species names), thus they provide an objective, standardized measure of comparison for evaluating species boundaries.

Presently, there are approximately 518,000 BINs, representing ~180,000 formally described animal species currently in BOLD. This indicates that there are currently only DNA barcodes for a fraction of formally described animal species (< 10%), and BINs for ~338,000 un-described animal species (granted legitimate species may occupy several BINs, thus reducing the estimated number of BINs of undescribed species). Contributions of DNA barcodes for known taxa (identified to species by traditional morphological characters – and complemented with molecular DNA data) are appreciated and can be provided in terms of "data release papers" (e.g. Zuniga et al. 2017). However, current researchers conducting biodiversity surveys, particularly of poorly known taxonomic groups, and groups also poorly represented by DNA barcode data, are left with the challenge of identifying cryptic species diversity using whatever molecular data is available (e.g. Stuart et al. 2006a). This identification must be done with some level of taxonomic expertise, where the specimens in hand are compared with species descriptions, and sequence data aligned with known reference material (e.g. GenBank). When multiple OTUs, or clades, are discovered among specimens identified with the same name, careful comparisons must be made to the original type descriptions, geographic distributions, and genetic data. As cryptic species are revealed, original descriptions of species and their geographic ranges must be modified to account for current taxonomic understanding. Nevertheless, there is a pressing need for biodiversity surveys in many parts of the world, and especially including groups in taxonomic disarray such as amphibians.

Here, we provide an example by incorporating DNA barcode data with biodiversity inventory survey data of amphibians and reptiles collected in a poorly known region of the world, the Tanintharyi Region of Myanmar (the 'Tenasserim'). This includes one of the most poorly known vertebrate groups in terms of taxonomy – anurans in Southeast Asia. Prior to this study, there were only 1259 anuran species with DNA barcodes in BOLD for the approximately 7727 currently recognized species of anurans known globally (Frost 2017). Specifically, we set out to determine how many species occur at our study site. We surveyed a region within the proposed Tanintharyi National Park near the village of Yeybu (Fig. 1), conducting day and night surveys, collecting representative voucher specimens with tissue samples for genetic analyses. We used DNA barcoding in conjunction with traditional methods to assist in our specimen identifications, not to delimit species (Collins and Cruickshank 2013). In the process, we discovered what likely represent new, undescribed species, 'species discovery' (Collins and Cruickshank 2013). We recommend the use of additional data and analyses to formally evaluate, describe, and recognize potential species identified with the barcode data. Here, we combine our survey results with a shorter survey in an area to the south, near the village of Chaung-nauk-pyan (Fig. 1), consisting of reserve forest and a recently slash and burned area, where only photographs and other observations were made. Our results show how the use of DNA barcode data can augment and increase the accuracy of biodiversity inventory surveys and suggest caution should be taken when identifications are made solely on morphological identifications, particularly for some of the more cryptic species complexes of anurans in this region. We offer our protocol and results as a model for oth-

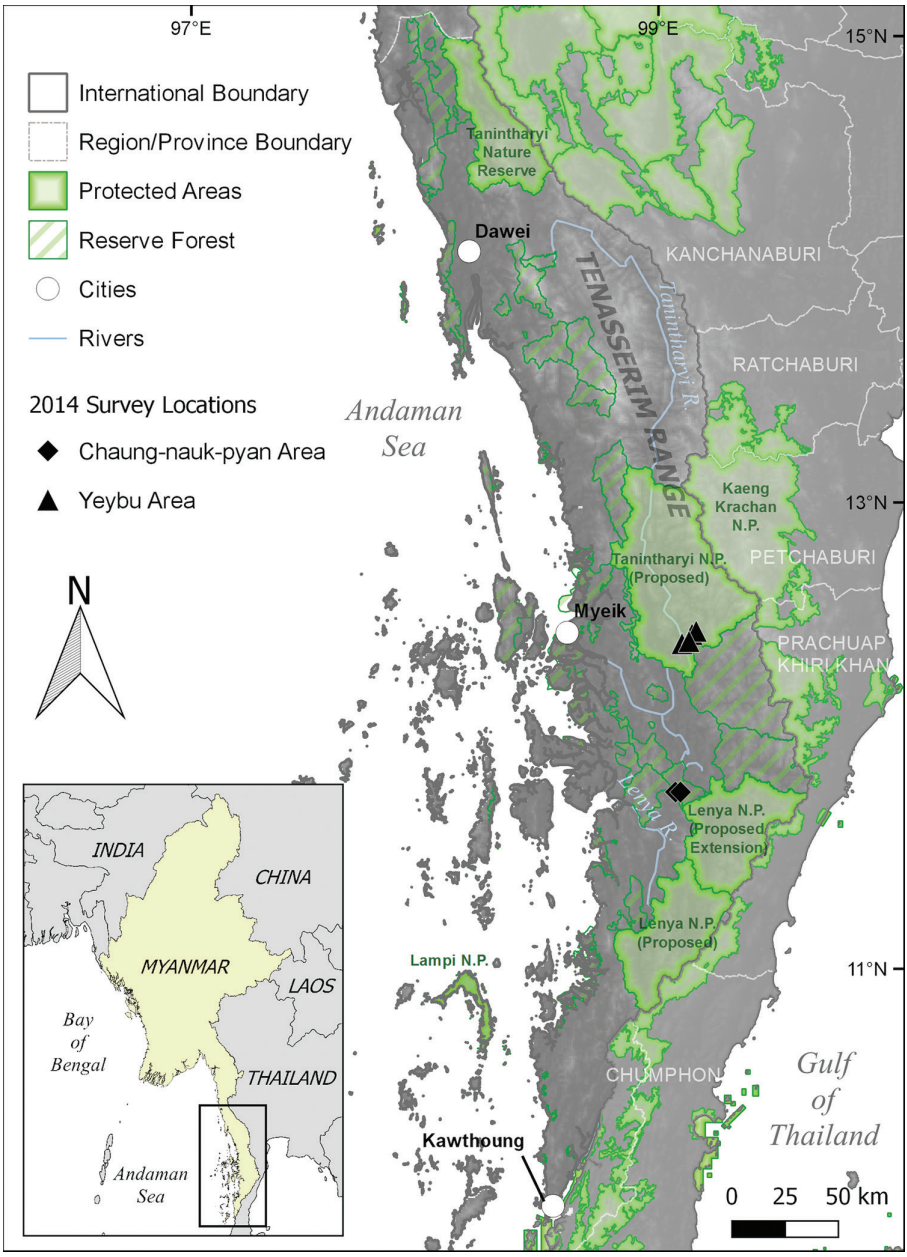

Figure 1. Map of Tanintharyi Region, Myanmar. The Tanintharyi Nature Reserve, north of Dawei, and the Lampi National Park, island northwest of Kawthoung, are officially designated as national parks in Myanmar. The other areas in Myanmar are proposed as national parks (Protected Areas) or being considered for future protection (Reserve Forests). The main survey reported here was conducted in the Yeybu area of the proposed Tanintharyi National Park (triangles: Forest 1–2 and Gardens sites). A shorter survey, with fewer people, was conducted in the Reserve Forest near the Chaung-nauk-pyan area (diamonds: Forest 3 and slash & burnt sites) and is also included in this report. Map provided by Grant M. Connette of the Smithsonian Conservation Biology Institute (SCBI).

ers working with groups in taxonomic disarray. Finally, we provide the first comprehensive checklist of amphibians and reptiles of the region (excluding marine species) based on our results, distinguished by DNA barcode verification versus observation only, and other expected species based on museum records and the literature.

Methods

The Tanintharyi

Tanintharyi is the southern-most political division of Myanmar, now formally known as the Tanintharyi Region. This region occupies about the southern two-thirds of the former colonial British administrative unit of Tenasserim; the northern portion is now Mon State. Biological surveys of Tanintharyi have been limited in postcolonial times owing to political disagreements and military activities. The last herpetofaunal summary of Tanintharyi is Theobald's 1868 report. The Myanmar Herpetological Survey (MHS) was permitted access to southernmost Tanintharyi (Kawthaung area) in 2002 and again to north-central Tanintharyi (Dawei area) in 2009 and 2010. Since then political change has allowed broader access. The Tenasserim, or Tanintharyi, contains type localities for at least seven amphibian and seven reptile species. Some of these, referenced near "Moulmein" (= Mawlamyine) are in present-day Mon State, while others, referenced near Dawei, "Mergui" (= Myeik), and the "Valley of the Tenasserim" are in the Tanintharyi Region. Theobald (1868) provided the first and last report of the amphibians and reptiles of the Tenasserim. Other researchers have reported on the occurrence of individual species or sets of species but no single attempt has been made to review the herpetofauna of the entire region. We joined Fauna & Flora International's biodiversity survey team in June 2014 to provide a preliminary assessment of the amphibian and reptile biodiversity survey in the proposed Tanintharyi National Park, of southern Myanmar. We realize that a rapid assessment survey would sample at best only a quarter to a third of the herpetofauna (Zug 2011) and only the species active in the early monsoon. The timing of this survey emphasizes amphibian species. The details of the sites are presented below in the survey itineraries. We supplement our morphological identifications with DNA barcoding. We provide COI data to build upon the taxonomic representation in BOLD and the Barcode Index Number system (BINs). We also include 16S data that can be directly compared with currently available published sequences in GenBank to provide better molecular identifications of our specimens. Many sequences in GenBank are incorrectly identified, which subsequently pollutes the database, especially so among Southeast Asian anurans. Tracking voucher specimen information can sometimes be difficult or nearly impossible if the information was not appropriately provided or was lost. Therefore, we attempt to exemplify how to efficiently review the taxonomy on a species-group basis, compare specimen morphology to species descriptions, genetic data with GenBank and BOLD records, evaluate those records, and how to interpret proper taxonomic nomenclatural assignments.

Lacking a recent review for the Tanintharyi, we relied on reptile and amphibian checklists and studies of Thailand (Nabhitabhata et al. 2004, Pauwels et al. 2002, Pauwels and Chan-ard 2006, Matsui et al. 2005a, Vogel et al. 2009, Grismer et al. 2010, Chan-ard and Makchai 2011, Ohler et al. 2011, Grismer et al. 2012, Sumontha et al. 2012, Grosjean et al. 2015a, Grismer et al. 2016, 2017, Wood et al. 2017, Pauwels et al. 2000, 2009), reports from other regions in Myanmar (Dowling and Jenner 1988, Zug et al. 2006a, b, 2007), and recent work in the Tanintharyi (e.g. Wilkinson et al. 2012, Connette et al. 2017, Mulcahy et al. 2017; Zug et al. 2017; Lee et al. 2018). An earlier report by Pauwels et al. (2002) for Phang-nga Province, Thailand, was south of the southern tip of Tanintharyi and below the Isthmus of Kra, which is useful for comparison of species that may cross the isthmus and occur in the southernmost Tanintharyi. We compiled a comprehensive species list of amphibians and reptiles documented in the Tanintharyi, either by our collections, observations, or specimens at the California Academy of Sciences (CAS). We also generated some DNA barcode data for specimens previously collected in Myanmar from the National Museum of Natural History, Smithsonian Institution (USNM) and CAS tissue collections. Some of the specimens we barcoded remain in the Myanmar Biodiversity Museum (MBM) in Hlawga National Park, north of Yangon. These specimens have yet to be catalogued at the MBM; therefore, we refer to them as MBM-Collector Number (e.g. MBM-JBS 19825). Comments on the occurrence and biological aspects of single species or groups of related species are included in the individual species accounts below. The Red List Status for each species was taken from www.iucnredlist.org on 9 March 2018.

Team members

The survey team for the proposed Tanintharyi National Park area comprised Myint Kyaw Thura (ENCA), Daniel G. Mulcahy (NMNH-SI) and Thaw Zin, for the Reserve Forest area Myint Kyaw Thura and Thaw Zin.

Survey itineraries and sites

1) Proposed Tanintharyi National Park – Yeybu area (Fig. 1). The survey team traveled to Myeik (12.4359°N, 98.5941°E, 7 m ASL; all latitude and longitude coordinates were taken with WGS84 datum) on 4 June 2014, hired a 4wd vehicle and drove to Tegu, then hired a boat to Yeybu on 5 June. Yeybu village (12.3927°N, 99.1044°E) is about 168 km northeast of Myeik. The village is 500 m east of the Tanintharyi River. On 6 June, the team with porters and cooks walked in and established the first camp ("Forest 1": 12.4345°N, 99.1442°E, 93 m ASL) alongside Yeybuchaung-ngal (ngal = stream). Our searching for amphibians and reptiles was conducted within a 500 m radius of the camp, principally along the creek and its smaller feeder streams owing to the absence of trails through the dense forest. The team moved upstream to a second camp ("Forest

2": 12.4478°N, 99.1621°E, 116 m ASL) on 10 June. Exceptionally heavy monsoonal rain on 12 June and rapidly rising stream level forced the team's return to the eastern edge of Yeybu village, where they used the cook's house as the third camp ("Gardens": 12.4039°N, 99.1312°E, 30m ASL) and searched for amphibians and reptiles in this area through the morning of 16 June, and then returned by boat to Tegu, whence by 4wd vehicle to Myeik. The first and second camps were within the evergreen forest. At both sites bamboo was a dominant feature of the vegetation and the canopy was closed, or nearly so. The third survey site was open agricultural land, principally of small gardens, orchards, and numerous small temporary ponds along the floodplain of the Tanintharyi River. Total survey time was 10 days, voucher specimens were taken.

2) Reserve forest – Chaung-nauk-pyan area (Fig. 1). A smaller survey team (over a shorter period) traveled by road from Myeik to the village of Chaung-nauk-pyan on 4 July 2014 and whence by foot the following day to a degraded evergreen forest site, approximately 4.25 km southwest of the village. Surveys at this site ("Forest 3": 11.7574°N, 99.0730°E, 49m ASL) occurred from 5 July through the morning of 7 July when they shifted to a recently cleared secondary forest site ("Slash & burnt": 11.7573°N, 99.0945°E, 67m ASL) and searched for amphibians and reptiles for the next 24 hours, returning to Chaung-nauk-pyan on the afternoon of 8 July and returned to Myeik on 10 July. Total survey time was four days. Captured frogs and reptiles were photographed and released at the site of capture.

Collections of amphibians and reptiles were made at four sites: hotel in Myeik (commensal); Yeybuchaung-ngal stream Camp 1 (Forest 1); Yeybuchaung stream Camp 2 (Forest 2); and vicinity of Yeybu village (Gardens). Only observations and photographs were taken at two sites: near the village of Chaung-nauk-pyan in a Reserve Forest (Forest 3) and nearby in a recently cleared site (Slash & burnt). Dates and latitude and longitude coordinates are identified above in the itinerary. The survey protocol was visual searching along Yeybuchang-ngal stream and its smaller side-branches. The stream was searched both during the day and at night (with flash-lights and head lamps). All amphibians were captured by hand; reptiles by hand, rubber-bands, slingshots (catapults), and snake tongs for large or dangerous species. Transport of specimens from the field to the camp was done in plastic bags for amphibians and cloth bags for reptiles. Unique individuals were usually photographed. All specimens to be retained as vouchers were euthanized following IUCAC protocols, with a drop of 5% benzocaine on the head (amphibians) or into the oral cavity (reptiles). Genomic tissue samples (piece of liver and/or muscle) were taken from all specimens. The genomic samples were harvested prior to formalin preservation; each sample was place in individual 1.5 ml tube with salt-saturated ethylene-diamine-tetracetic acid/ Dimethyl sulfoxide (EDTA/DMSO) buffer for long-term storage and future genetic analyses modified from Seutin et al. (1991) with 25% DMSO instead of 20% (Mulcahy et al. 2016). Specimens were individually tagged with a unique field number and preserved in 10% formalin. Voucher specimens and tissues were deposited at the National Museum of Natural History, Smithsonian Institution (USNM) collection.

Molecular data

We attempted several rounds of PCR and sequencing for each specimen collected, with the exception of four *Odorrana hosii* Boulenger and three *Ansonia thinthinae* Wilkinson, Sellas, & Vindum. In addition to our samples from our expedition, we DNA barcoded 108 additional specimens of amphibians from previous USNM collections in Myanmar, mostly northern states, to verify if these were the same species that we collected in the Tanintharyi. Tissue of these specimens are from the USNM tissue collection and were initially collected into 95% EtOH and subsequently stored at -80 °C. Extractions of genomic DNA from all specimens were performed on an AutoGenprep 965 (2011 AutoGen, Inc.), using standard phenol manufacturer protocols. Genomic DNA was eluted in 100 µl of AutoGen R9 re-suspension buffer. Polymerase chain reactions (PCR) were conducted for the mtDNA large ribosomal subunit (rrnL: 16S) and cytochrome oxidase subunit I (COI) using the primers: 16Sar 5' CGCCTGTTTATCAAAAACAT 3' and 16Sbr 5' CCGGTCTGAACTCAGATCACGT 3' (Palumbi et al. 1991) and COI-ReptBCF 5' TCAACAAACCAYAAAGAYATYGG 3' and COI-ReptBCR 5' TAAACTTCAGGGTGGCCRAARAATCA 3' (Castañeda and de Queiroz 2011). For some specimens, we also sequenced either part of the ND2 gene using the primers L4437–H5934 (Macey et al. 1997) or 12S (12SI: 5' TGCCAGCAGYCGCGGTTA 3' and 12SIII 5' AGAGYGRCGGGCGATGTGT 3'; Puillandre et al. 2009) in order to compare with sequences available for these, or closely related species in GenBank. The PCRs were performed in 96-well plates, in 10 µl reactions, following protocols "3.6 PCR Methods: Amplification" and "3.8 PCR Purifications: EXOSAP-IT" of Weigt et al. (2012), with annealing temperatures of 54 °C for 16S and 12S, 48 °C for COI, and 52 °C for ND2. Sequence reactions were performed in 96-well plates with the PCR primers using BigDye® Terminator v3.1 Cycle Sequencing Kit's in 0.25 × 10 µl reactions and run on an Automated ABI3730 Sequencer (2011 Life Technologies). Raw chromatograms were edited in Sequencher v5.1 (2012 Gene Codes Corp.), complementary strands were aligned, and COI was inspected for proper translation, alignments were done using the MUSCLE option in Sequencher. Neighbor-joining (NJ) trees were generated in PAUP* v4.0b10 (Swofford, 2002) for the 16S and COI data separately, and of the combined data. Scale bars at bottom of each tree represent uncorrected p-distances.

Specimen identification

Sequences for uncertain taxa were further assessed by multiple methods. First, we considered specimens placed in the same COI BINS (Ratnasingham and Hebert 2013) to represent the same species. Specimens placed in separate BINs from different geographic localities, which grouped together in the NJ trees, and that were indistinguishable based on morphology, were considered the same species with genetic

variation associated with geography. Specimens that were placed in separate BINs that were either different morphologically or did not group together (i.e. grouped with other taxa) in the NJ trees, were considered different species. For specimens that we could not identify based on COI BINs, we created alignments with material from GenBank representing the same genera with as many species as possible. Neighbor-joining trees were estimated from these alignments at the family-level. This was mostly done for amphibians using 16S sequence data from GenBank. The 16S locus is known to evolve much slower than protein-encoding mitochondrial loci. Therefore, our assessments of specimen identification based on 16S data were done on a case-by-case basis, considering the geographic distance between specimens being compared and whether or not our specimens met the morphological description of the species they clustered with.

Results

Our compiled list of species documented in the Tanintharyi contains 46 amphibians and 110 reptiles, including one caecilian, 45 anurans, 100 squamates (42 lizards and 58 snakes), and 10 turtles (Table 1). Results from our surveys in the Tanintharyi represent total observations of 51 species, 43 species (24 amphibians and 19 reptiles) in the proposed Tanintharyi National Park area, and eight additional species (one amphibian and seven reptiles) in the lowland areas (Table 1). We produced COI DNA barcode data for 297 specimens (GenBank MG935416–MG935712) and 16S data for 292 specimens (GenBank MG935713–MG936004), representing 72 species (55 amphibians and 17 reptiles), including 25 amphibians and 17 reptiles observed in the Tanintharyi (Table 1) and an additional 30 species of amphibians from northern Myanmar from our reference material (Table 2). Eleven of the species barcoded from the northern Myanmar material were also discovered in the Tanintharyi. Our COI sequences were placed into 93 BINs, of which 18 already existed. The BIN results are only mentioned in the text below if sequences went into pre-existing BINs, or if specimens of the same species were placed in separate BINs. In total, we provide sequence data for 81 species of amphibians and reptiles (Fig. 2). We provide accounts for each species observed in the Tanintharyi below, with additional comments on the reference material from northern Myanmar. The following descriptions offer brief characterization of the specimens vouchered and examined; the general distributions contain condensed and abstracted geographic data derived from ASoW (Frost 2017) and the Reptile Database (Uetz et al. 2018); both accessed 20–21 January 2018. All species were recorded in the Yeybu area unless otherwise noted. Species only observed in the Forest Reserve (Chaung-naukpyan area) are noted in the Natural History Notes. See Table 1 for a complete list of species observed at each site.

Table 1. The herpetofauna of southern Tanintharyi. Marine or estuarine species are excluded. Occurrence data is derived from the California Academy of Sciences (CAS) or our observations; only species identification for the Tanintharyi Proposed National Park (PNP) area and a few of the Reserve Forest (RF) species were confirmed by specimen examination and DNA barcode data. Abbreviations: D = Dawei area, K = Kawthaung area, M = Myeik, ? = have specimens but not identified to species and/or DNA barcoding needed to confirm identification; + species present; – species absent; √ DNA barcoded.

Order	Family – Subfamily/Species	CAS	PNP Area‡	RF
ANURANS	**Bufonidae**			
	Ansonia thinthinae	D	√	–
	Duttaphrynus melanostictus	DK	–	–
	Ingerophrynus parvus	DK	√	+
	Phrynoidis asper	DK	√	+
	Ceratobatrachidae			
	Alcalus tasanae	K	–	–
	Dicroglossidae – Dicroglossinae			
	Fejervarya sp. (hp2)	?	√	–
	Fejervarya sp. (hp3)	?	√	+
	Hoplobatrachus rugulosus	D	–	–
	Limnonectes blythii	DK	√	+
	Limnonectes doriae	DK	√	+
	Limnonectes hascheanus	DK	–	–
	Limnonectes kohchangae	D	–	–
	Limnonectes laticeps	K	–	–
	Limnonectes limborgi	D	√	–
	Limnonectes macrognathus	D	–	–
	Dicroglossidae – Occidozyginae			
	Ingerana tenasserimensis	D	√	–
	Occidozyga lima	?	√	–
	Occidozyga martensii	DK	√	+
	Megophryidae			
	Leptobrachium smithi	DK	–	–
	Megophrys sp.	D	–	–
	Microhylidae – Kalophryninae			
	Kalophrynus interlineatus	D	–	–
	Microhylidae – Microhylinae			
	Kaloula latidisca	D	√	–
	Kaloula pulchra	DK	–	M√
	Microhyla berdmorei	DK	–	–
	Microhyla butleri	?	√	–
	Microhyla heymonsi	DK	√	–
	Microhyla fissipes	?	√	+
	Microhyla pulchra	D	–	–
	Micryletta inornata	DK	–	–
	Ranidae			
	Amolops marmoratus	D	–	–
	Amolops panhai	D	√	–
	Chalcorana eschatia	?	√	–
	Clinotarsus alticola	DK	–	–

Order	Family – Subfamily/Species	CAS	PNP Area‡	RF
ANURANS	*Hydrophylax leptoglossa*	D	–	–
	Hylarana erythraea	DK	√	–
	Odorrana hosii	K	√	–
	Odorrana livida	D	√	–
	Sylvirana malayana	?	√	–
	Sylvirana nigrovittata	DK	–	–
	Rhacophoridae – Rhacophorinae			
	Chirixalus vittatus	D	–	–
	Nyctixalus pictus	K	–	–
	Polypedates mutus	?	√	–
	Polypedates leucomystax	DK	√	+
	Rhacophorus cyanopunctatus	D	–	–
	Rhacophorus verrucosus	D	–	–
	Theloderma phyrnoderma	D	–	–
CAECILIANS	**Ichthyophiidae**			
	Ichthyophis cf. kohtaoensis	DK	√	–
TESTUDINES	**Geoemydidae**			
	Cyclemys dentate	D	–	–
	Testudinidae			
	Indotestudo elongata	–	+	–
	Trionychidae – Trionychinae			
	Dogania subplana	–	+	–
	Nilssonia formosa	D	–	–
SQUAMATES – LIZARDS	**Agamidae**			
	Acanthosaura crucigera	DK	√	–
	Bronchocela burmana	K√	–	–
	Calotes emma	DK	√	+
	Calotes "versicolor"	DK	–	–
	Draco blanfordii	DK	√	–
	Draco maculatus	DK	–	–
	Draco taeniopterus	DK	–	+
	Gekkonidae			
	Cyrtodactylus brevipalmatus	D	–	–
	Cyrtodactylus lenya	√	–	–
	Cyrtodactylus oldhami	DK	–	–
	Cyrtodactylus payarhtanensis	√	–	–
	Gehyra mutilata	DK	–	–
	Gekko gecko	DK	√	+
	Hemidactylus frenatus	DK	–	M√
	Hemidactylus garnotii	DK	√	–
	Hemidactylus karenorum	?	–	–
	Hemidactylus platyurus	K	–	–
	Hemidactylus tenkatei	K	–	M√
	Lacertidae			
	Takydromus sexlineatus	–	√	–
	Scincidae			
	Dasia olivacea	K√	–	–
	Eutropis longicauda	?	–	–

Order	Family – Subfamily/Species	CAS	PNP Area‡	RF
SQUAMATES – LIZARDS	*Eutropis macularia*	DK	√	–
	Eutropis multifasciata	DK	√	+
	Lipinia vittigera	DK	–	–
	Lygosoma bowringii	K	–	–
	Scincella reevesi	D	–	–
	Sphenomorphus maculatus	DK	√	+
	Tropidophorus robinsoni	DK	–	–
	Varanidae			
	Varanus rudicollis	–	+	–
SQUAMATES – SNAKES	**Acrochordidae**			
	Acrochordus granulatus	D	–	–
	Colubridae – Colubrinae			
	Ahaetulla mycterizans	K	√	–
	Ahaetulla fronticincta	D	–	–
	Ahaetulla prasina	DK	–	+
	Boiga cyanea	DK	–	–
	Boiga dendrophila	–	√	–
	Boiga drapiezii	K	√	–
	Boiga multomaculata	D	–	–
	Boiga siamensis	D	–	–
	Coelognathus radiatus	M	–	–
	Dendrelaphis formosanus	K	–	–
	Dendrelaphis haasi	D	–	–
	Dendrelaphis pictus	D	–	+
	Dendrelaphis striatus	DK	–	+
	Dryocalamus subannulatus†	–	–	–
	Gonyosoma oxycephalum	D	–	–
	Lycodon aulicus	D	–	–
	Lycodon subcinctus	D	–	–
	Colubridae – Natricinae			
	Rhabdophis chrysargos	DK	√	–
	Rhabdophis nigrocinctus	D	–	+
	Xenochrophis piscator	D	√	+
	Xenochrophis punctulatus	D	–	–
	Xenochrophis trianguligerus	DK	√	–
	Elapidae			
	Bungarus sp.	K	–	–
	Calliophis maculiceps	D	–	–
	Hydrophis schistosus	D	–	–
	Naja kaouthia	–	–	+
	Homalopsidae			
	Cantoria violacea	D	–	–
	Homalopsis semizonata	D	–	–
	Cerberus rynchops	D	–	–
	Fordonia leucobalia	D	–	–
	Lamprophiidae – Pseudaspidinae			
	Psammodynastes pulverulentus	DK	–	–

Order	Family – Subfamily/Species	CAS	PNP Area‡	RF
SQUAMATES – SNAKES	**Pareatidae**			
	Pareas carinatus	DK	–	–
	Pareas macularius	D	–	–
	Pareas margaritophorus	?	–	–
	Pythonidae			
	Malayopython reticulatus	–	+	–
	Typhlopidae			
	Indotyphlops braminus	D	–	–
	Uropeltidae			
	Cylindrophis burmanus	K	–	–
	Viperidae – Crotalinae			
	Trimeresurus purpureomaculatus	DK	–	–
	Trimeresurus stejnegeri	K	–	–
	Trimeresurus sp.§	√	–	–
	Xenopeltidae			
	Xenopeltis unicolor	D	–	–

‡Most of the species observed in the proposed National Park were vouchered and barcoded. Four taxa (*Dogania cartilaginea*, *Indotestudo elongata*, *Varanus rudicollis*, and *Malayopython reticulatus*) were seen in possession of locals and only photos were taken. At the Reserve Forest sites, specimens were identified in field and released after being photographed.

§See Mulcahy et al. (2017) for a taxonomic treatment of the *Trimeresurus* in the Tanintharyi.

†See Lee et al. (2018).

Table 2. Additional species for which DNA barcodes were generated from the reference material from northern Myanmar. Asterisks indicate species also detected in the Tanintharyi Region.

Family	Species	State/Province
Ichthyophiidae	*Ichthyophis multicolor*	Ayeyawady
Bufonidae	*Duttaphrynus melanostictus*	Sagaing
Dicroglossidae	*Fejervarya* sp. (BFL2007)	Sagaing
		Magway
	Fejervarya sp. A (DGM2018)	Sagaing
		Mandalay
	Fejervarya sp. (hp2)*	Yangon
		Bago
	Fejervarya sp. (hp3)*	Bago
	Hoplobatrachus litoralis	Yangon
		Bago
	Hoplobatrachus rugulosus	Sagaing
	*Ingerana tenasserimensis**	Mon State
	*Limnonectes doriae**	Yangon
		Bago
		Mon State
	*Limnonectes limborgi**	Bago
		Mon State

Family	Species	State/Province
Dicroglossidae	*Limnonectes longchuanensis*	Mandalay
	*Occidozyga lima**	Mandalay
		Sagaing
	Occidozyga sp. A (DGM2018)	Yangon
	Occidozyga sp. B (DGM2018)	Yangon
	Occidozyga sp. C (DGM2018)	Bago
	Sphaerotheca breviceps	Sagaing
Megophryidae	*Leptobrachium smithi*	Mon State
	Leptolalax sp. A (DGM2018)	Mandalay
Microhylidae	*Glyphoglossus molossus*	Sagaing
	Kalophrynus anya	Sagaing
	*Kaloula pulchra**	Yangon
		Bago
		Sagaing
		Mandalay
	Microhyla berdmorei	Yangon
		Bago
	*Microhyla fissipes**	Yangon
		Sagaing
		Bago
		Mandalay
		Magway
	*Microhyla heymonsi**	Bago
		Mandalay
	Microhyla sp. A (DGM2018)	Sagaing
	Microhyla sp. B (DGM2018)	Magway
Ranidae	*Amolops marmoratus*	Mon State
	Humerana humeralis	Bago
	*Hylarana erythraea**	Yangon
	Hylarana lateralis	Yangon
		Sagaing
	Hylarana sp. A	Sagaing
	Hylarana tytleri	Bago
	Odorrana cf. *chloronota*	Mandalay
	*Odorrana livida**	Mon State
	Sylvirana nigrovittata	Mon State
		Mandalay
	Sylvirana lacrima	Mandalay
Rhacophoridae	*Chiromantis* sp. A (DGM2018)	Sagaing
	Chiromantis doriae	Mandalay
	Polypedates teraiensis	Yangon
		Sagaing
		Bago
	Rhacophorus rhodopus	Mandalay

Caecilians
Ichthyophiidae – Asian caecilians (Suppl. material 1: Fig. 1)

***Ichthyophis* cf. *kohtaoensis* (Nishikawa et al., 2012)**
Striped Caecilian

Description. A single individual was found. Not dissected, sex and maturity unknown, likely juvenile; 144 mm SVL, 2 mm TailL. This individual had a bright yellow ventrolateral stripe in life (white in preservation) on each side ending below eye, ~273 primary annuli, 3 caudal annuli, eye visible, and tentacle opening much nearer eye than external choana.

Natural history notes. This individual was discovered on the forest floor, immediately following a heavy rain.

General distribution. Tanintharyi and peninsular Thailand.

Molecular data. The 16S sequence is 98% identical to several sequences in GenBank, including GB AB686168, *Ichthyophis* cf. *supachaii* UKMHC 877 and KUHE 23189 from Malaysia and Thailand, respectively. However, our specimen fell outside of the 16S clade containing *I.* cf. *supachaii*, *I.* cf. *hypocyaneus,* and *I.* cf. *kohtaoensis* in a neighbor-joining tree of *Ichthyophis* 16S sequences in GenBank. Additional 12S data for our specimen (GenBank MG944814) placed it in the "*Ichthyophis* cf. *kohtaoensis"* clade (Suppl. material 1: Fig. 1) sister to KUHE 19615, 19617, and 19659 (GenBank AB686107–9), from Ko Samui Island, Thailand (Nishikawa et al. 2012). We note a fourth specimen identified as "*I.* cf. *kohtaoensis*" (GenBank AB686146) by Nishikawa et al. (2012), from the southern Malaysian Peninsula, is placed in a "*I.* cf. *supachaii*" + *I.* cf. *hypocyaneus* + *Ichthyophis* sp. 1" clade, consistent with their study (Suppl. material 1: Fig. 1).

Specimens examined. USNM 586851.

Red List status. *Ichthyophis kohtaoensis* listed as LC (Least Concern).

Additional *Ichthyophis*. We included one caecilian from the legacy collection, an *Ichthyophis multicolor* from Ayeyarwady Region, Myanmar (USNM 576283). This specimen is 14% different (COI) and 8.3% different (16S) from our *I.* cf. *kohtaoensis* specimen, and is identical to GenBank FR716007, *I. multicolor*, CAS 212254, a paratype (Wilkinson et al. 2014) also collected from Ayeyarwady Region, Myanmar (Suppl. material 1: Fig. 1). Additionally, we sequenced four individuals from the California Academy of Sciences, two "*Ichthyophis* sp." from Bago Division (CAS 239657; 12S GenBank MG944807; and CAS 239722; 12S and 16S GenBank MG944808–9) and they were placed in the *I. multicolor* clade, thus extending the known distribution of this species. The other two specimens were from near Dawei (CAS 247969; 12S and 16S GenBank MG944812–13) and near Kawthaung (CAS 247466; 12S and 16S GenBank MG944810–11), both in the Tanintharyi Region, and were placed in our *I.* cf. *kohtaoensis* clade, expanding the range of this clade from the southern Tanintharyi Region and the Thai-Malay Peninsula into the northern Tanintharyi.

Anurans
Bufonidae – toads (Suppl. material 1: Fig. 2)

***Ansonia thinthinae* Wilkinson, Sellas & Vindum, 2012**
Thin Thin's Stream Toad

Description. Sample of two immature males 22.0, 22.3 mm SVL, mature males 19.5–23.4 mm (n = 5), immature females 21.1–22.8 (n = 6) and two mature females 23.3–25.6 mm SVL.

Natural history notes. All individuals were on rocks in and alongside small cascades in full canopied areas of forest streams.

General distribution. Known only from Tanintharyi, Myanmar.

Molecular data. Our specimens form a single clade with 99–100% similarities based on 16S data, and are 96–97% similar to the type series from northern Tanintharyi, Myanmar (Wilkinson et al. 2012). We note that the type series forms a clade with our samples, and that clade is sister to *A. kraensis* (AB435250–52) to the exclusion of other peninsular species (Grismer et al. 2017). The long branch between our samples and the type series may represent genetic variation associated with geography in a low-dispersal group, or it suggests this may represent a species complex (see Suppl. material 1: Fig. 2).

Comments. The sample appears to represent a single reproductive-season cohort amid maturation. If our assessment of maturity is correct, this population has slightly smaller adults than the more northern topotypic population where adult males were 22–28 mm SVL and a single adult female was 31.8 mm.

Specimens examined. USNM 586852–866.

Red List status. EN (Endangered).

***Ingerophrynus parvus* (Boulenger, 1887)**
Dwarf Toad

Description. Adult male 37.7 mm and adult female, 45.2 mm SVL.

Natural history notes. Both individuals were found in the leaf-litter of forest sites 1–2, and also observed in the slash & burnt area.

General Distribution. Southern Myanmar and southwestern Thailand through Malay Peninsula into Greater Sunda Islands.

Molecular Data. Our specimens are genetically similar to one another (99.6% identical) and, based on 16S data are placed in a clade with other *I. parvus,* though showing substantial genetic differences (91–94% identical) from GenBank material (AB746455 and AB530649–51). GenBank specimens are from Malaysia, suggesting either this may represent a species complex, or this represents a single species that shows high genetic diversity, possibly attributed to a low dispersal rate of a leaf-litter species.

Specimens examined. USNM 586867–868.

Red List status. LC.

Figure 2. Combined (16S and COI) Neighbor-joining tree for all specimens sequenced in this study.

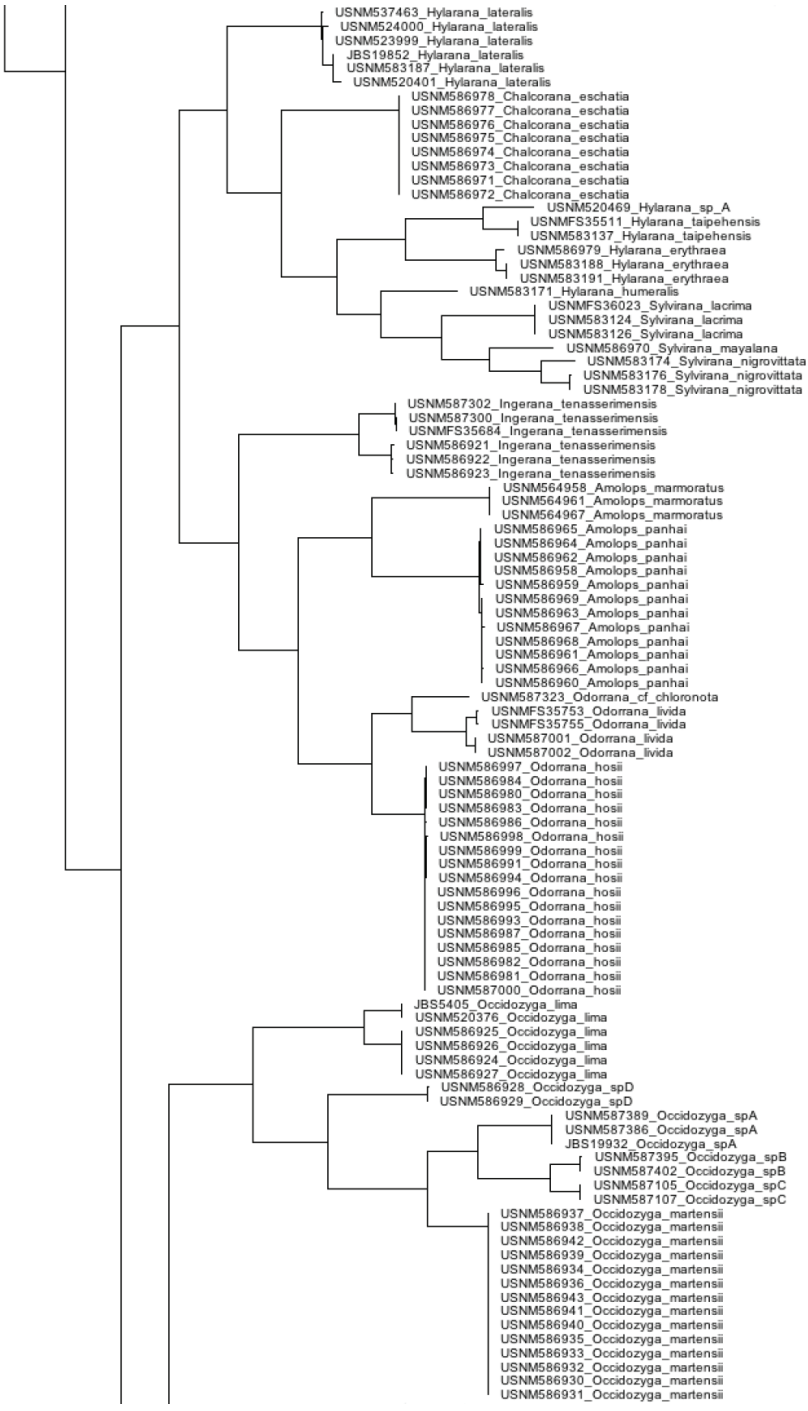

Figure 2. Continued.

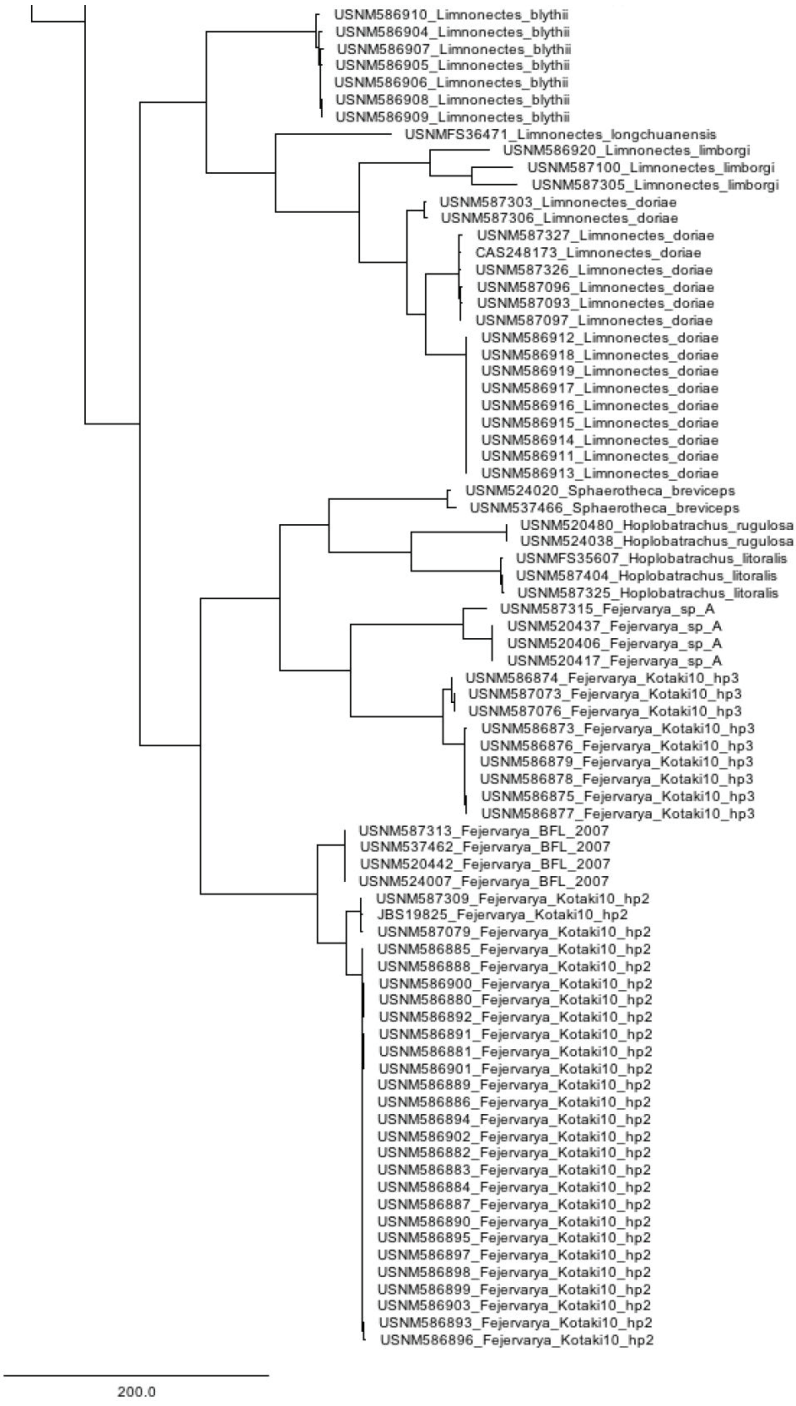

Figure 2. Continued.

***Phrynoidis asper* (Gravenhorst, 1829)**
Asian Giant Toad
Fig. 3A

Description. Three adult males 103.0, 104.3, 104.5 mm SVL, and juvenile 34.5 mm SVL.

Natural history notes. This riverine species occurs along stream borders but is principally a terrestrial species.

General Distribution. Tanintharyi, peninsular Thailand and Malaysia to Sumatra, Java, and Borneo.

Molecular Data. Our specimens are genetically nearly identical to one another (99–100% identical) and some are identical to one specimen in GenBank (DQ158432; FMNH 248148) from Brunei, suggesting low genetic diversity in a potentially high rate of dispersal species. These specimens form a clade with other *P. asper* from GenBank (Suppl. material 1: Fig. 2).

Specimens examined. USNM 586969–972.

Red List status. LC.

Additional bufonids. We also sequenced two specimens of *Duttaphrynus melanostictus* from Sagaing, Myanmar (USNM 523959 and USNM 520316), for genetic comparisons. The *D. melanostictus* species complex is in need of taxonomic revision (e.g. Wogan et al. 2016). Our samples were nested among other *D. melanostictus* specimens in GenBank (not shown), identical to one (KF665340) specimen (CAS 247174), also from Sagaing but a different locality. We refer to these specimens as *D. melanostictus* until the species complex is revised.

Dicroglossidae (Suppl. material 1: Fig. 3)
Dicroglossinae – grass and fanged frogs

***Fejervarya* sp. 'hp2' (Clade 21 of Kotaki et al. 2010)**
Grassfrog

Description. Medium-sized morph, adult females (7) 41.1–55.3 mm, adult males (16) 38.5–45.0 mm SVL.

Natural history notes. These frogs occur in a variety of human-modified habitats from drainage ditch to rice fields. All females are gravid and bear a mix of pigmented ova and small developing follicles, although only one had a full complement of pigmented ova. Presumably the other females had bred and deposited about half of their mature ova.

General Distribution. Western Thailand, Bangkok to Mae Hong Son and Three Pagoda Pass, to Yangon, Bago, and Tanintharyi, Myanmar.

Molecular Data. In addition to the specimens collected in Tanintharyi, we sequenced three other individuals from Magway (USNM 587309), Yangon (MBM-JBS 19825), and Bago (USNM 587079). Our specimens were placed into two COI BINs, one containing all of the Tanintharyi specimens, and one containing the rest. The

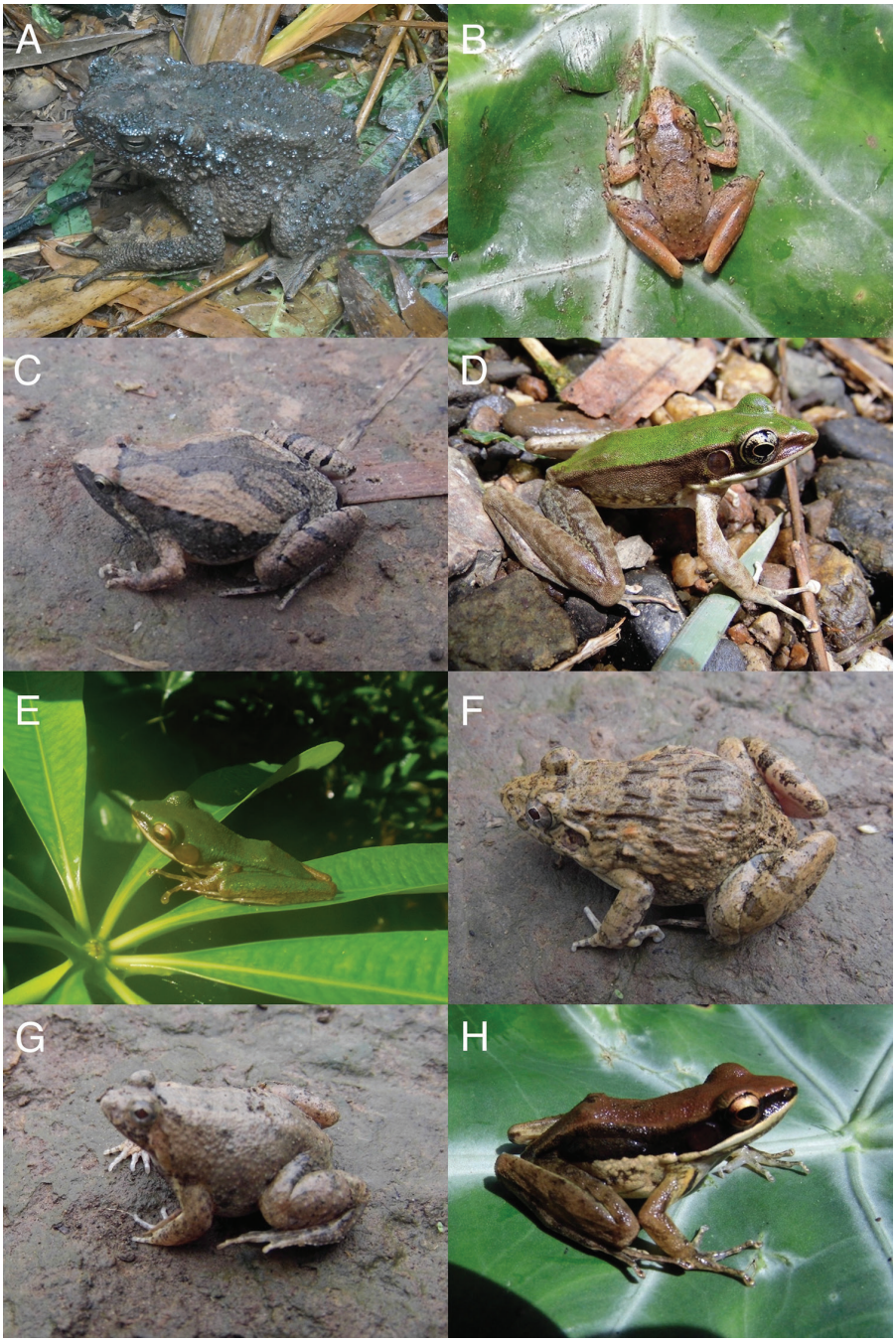

Figure 3. Selected amphibians found during this study's expedition. **A** *Phrynoidis aspera* (USNM 586871) **B** *Limnonectes doriae* (USNM 586911) **C** *Microhyla fissipes* (USNM 586949) **D** *Odorrana hosii* (USNM 586980) **E** *Chalcorana eschatia* (USNM 586971) **F** *Fejervarya sp.* (USNM 586881) **G** *Occidozyga martensii* (USNM 586930) **H** *Sylvirana malayana* (USNM 586970). Photos by Myint Kyaw Thura and Daniel G. Mulcahy.

Tanintharyi specimens were placed in a pre-existing BIN (ACT3129) identified as *F. triora*. The specimens in that BIN do not appear to be publicly available, though they appear to be from Grosjean et al. (2015b). The other BIN (ADG3054) comprising our Magway, Yangon, and Bago specimens is novel. Our specimens were placed in a 16S clade with a specimen identified as *Fejervarya limnocharis* hp2, Clade 21 from Thailand (AB277299, Kotaki et al. 2010), as well as another specimen identified as *Fejervarya limnocharis* (AB162444, Sumida et al. 2007), and specimens identified as *F. triora* (Grosjean et al. 2015b, KR827756–61), all from Thailand. One of the paratypes of *F. triora* (FMNH 266160), an additional specimen (FMNH 266337) from the type description (DQ860094–95, Stuart et al. 2006b), and a third individual (AB488883) identified as *F. triora* are placed elsewhere in the phylogeny (Suppl. material 1: Fig. 3). Thus, it appears the specimens from Grosjean et al. (2015b) appear to be mis-identified, including the COI BIN BOLD:ACT3129. We refer to this clade as "*Fejervarya* sp. hp2," belonging to the 'hp2,' Clade 21 of Kotaki et al. (2010).

Comments. Our material extends the range of this clade from Mae Hong Son and Bangkok, Thailand to the Tanintharyi, Myanmar.

Specimens examined. MBM-JBS 19825, USNM 587079, USNM 587309, USNM 587313, USNM 586880–903.

Red List status. NE (Not Evaluated).

***Fejervarya* sp. 'hp3' (Clade 11 of Kotaki et al. 2010)**
Grass Frog

Description. Adult females (n = 3) 32.3–38.5 mm, adult males (n = 4) 31.0–33.5 mm SVL.

Natural history notes. These frogs occurred in a variety of human-modified habitats from drainage ditch to rice fields.

General Distribution. Pilok, western Thailand, to Bago and Tanintharyi, Myanmar.

Molecular Data. We included two individuals from Bago (USNM 587073, USNM 587076) that were related to our specimens. Our specimens were placed into two COI BINs, one for the Bago specimens (ADG3052) and one for the Tanintharyi specimens (ADG2768). We did not obtain COI sequence from one of our Tanintharyi specimens (USNM 586874), yet it was placed sister to the Bago specimens in our combined tree. All of our specimens were placed in a 16S clade with a specimen (AB277300) from GenBank identified as *Fejervarya* sp. 'hp3' Clade 11 of Kotaki et al. (2010). Thus, we refer to our specimens and this clade as *Fejervarya* sp. belonging to the 'hp3,' Clade 11 of Kotaki et al. (2010). This entire clade was placed sister to a specimen (AB488889) from the Andaman Islands, India identified as *Fejervarya* sp. 'hp6.' Clade 12 (Kotaki et al. 2010).

Comments. Our material extends the range of this clade from Bangkok, Thailand to Bago and the Tanintharyi, Myanmar. The Grassfrogs, *Fejervarya limnocharis* complex, has gone from a single species of widespread tropical Asian frog in the early 1950s to twenty plus species in 2008 (Zug 2011) to double that number now.

The number will likely continue to increase over the next decade. In Myanmar, two species commonly occur together as represented by our Tanintharyi vouchers. The sympatric species display non-overlapping size ranges between males and females of the two species, although the males of the larger species may overlap in size with that of the females of smaller species.

Specimens examined. USNM 587073, USNM 587076, USNM 586873–879,

Red List status. NE.

Additional *Fejervarya*. We sequenced three specimens from Sagaing (USNM 520442, USNM 524007, USNM 537462) and one from Magway (USNM 587313) that were placed in one COI BIN, and were placed in a 16S clade with material in GenBank identified as *Fejervarya* sp. BFL 2007, large types 1–2 from Bangladesh (Islam et al. 2008, Hasan et al. 2012a). These were placed sister to three *F. orissaensis*, from Odisha, India and are over 8% sequence divergence from the *Fejervarya* sp. hp2 clade for COI. To be consistent, we refer to these specimens as *Fejervarya* sp. BFL 2007, which extends this Bangladesh clade into Myanmar. We sequenced three additional specimens from Sagaing (USNM 520406, USNM 520417, USNM 520437) and one from Mandalay (USNM 587315) that were placed in two COI BINs, respectively. These specimens were placed in a 16S clade with a sequence in GenBank (AF206466) of a specimen (USNM 520407) collected from the same locality in Sagaing. This clade was placed sister to a clade consisting of two new species (*F. dhaka* and *F. asmati*) recently described from Bangladesh (Howlader et al. 2016). We refer to our specimens as *Fejervarya* sp. A.

***Limnonectes blythii* (Boulenger, 1920)**

Blyth's Fanged Frog

Description. Adult females (2) 114.8, 127.6 mm SVL, immature female 89.1 mm, presumed adult males 88.1 mm SVL. The vouchers, also including three juveniles, range from 30.2 to 56.3 mm SVL. All individuals have dark or black soles of hindfeet.

Natural history notes. These frogs occurred in or at the edge of the forest streams.

General Distribution. Southern Myanmar and western Thailand southward to Sumatra and Borneo.

Molecular Data. Our specimens were placed in a 16S clade with material from GenBank identified as *L. blythii* from neighboring Thailand (GU934328) and elsewhere (RBU55270, RBU66127, RBU66131, RBU66133, and RBU66135; no locality data provided). There are several other sequences from GenBank identified as "*L. blythii*" elsewhere in the tree (e.g. RBU55269, RBU66115). The type locality for *L. blythii* is "Tenasserim valley", Myanmar. Therefore, we consider our specimens to represent *L. blythii* based on material closest to the type locality (and fitting the description), and the other specimens in GenBank are misidentified.

Comments. These semi-aquatic frogs are the largest anurans in this area in both mass and length.

Specimens examined. USNM 586904–910

Red List status. NT (Near Threatened).

***Limnonectes doriae* (Boulenger, 1887)**
Doria's Fanged Frog

Description. Juveniles (*n* = 4) 27.6–33.1 mm, subadult female 48.6 mm, adult males (4) 48.7–53.5 mm SVL.

Natural history notes. The collection contains two distinct size classes. The gonads of the adults were sexually quiescent; this information and the "half-grown" juveniles indicate an end of the monsoon- early dry season breeding cycle.

General Distribution. Myanmar and western and peninsular Thailand.

Molecular Data. We included other individuals from Yangon (USNM 587326–27, and CAS 248173), Bago (USNM 587093 and 587097), and Mon (USNM 587303 and 587306) states. These were placed into three COI BINs, one for the Yangon and Bago specimens (ADG3667), one for the Mon State specimens (ADG3666), and one for the Tanintharyi specimens (AAB2123). All of these specimens formed a 16S clade with a specimen from GenBank (GU934330) identified as *L. doriae* from Myanmar, Pegu (CAS 208425) and another specimen identified as "*L. nitidus*" from peninsular Thailand (Grosjean et al. 2015b; KR827897), a species known only from the Cameron Highlands and Fraser's Hills, of Peninsular Malaysia. The type locality for *L. doriae* is northern Tenasserim, near Mawlamyine, between the Tanintharyi and northern individuals. We note that Grosjean et al. (2015b) did not report having any *L. doriae* in their study, a species that extends onto the Thai-Malay Peninsula. Therefore, we consider the Tanintharyi specimens, the CAS specimens, and the Grosjean et al. (2015b) *L. nitidus* to be *L. doriae*.

Specimens examined. USNM 586911–919, USNM 587326–27, USNM 587093, USNM 587097, USNM 587303, USNM 587306, CAS 248173.

Red List status. LC

***Limnonectes limborgi* (Sclater, 1892)**
Limborg's Fanged Frog

Description. A single juvenile, 24.8 mm SVL.

Natural history notes. Found along forest stream.

General Distribution. Northwestern Myanmar and adjacent Thailand.

Molecular Data. We included two individuals, one each from Bago (USNM 587100) and Mon (USNM 587305) states. Each specimen was placed in its own COI BIN. These all form a single clade in our combined analysis (Fig. 2). These specimens all form a 16S clade with a large number of *L. limborgi* from GenBank (GU934334–36, GU934339–48, GU934353–55, and GU934357–65). Our Tanintharyi specimen (USNM 586920) is at the base of this *L. limborgi* clade, with a sequence in GenBank (AB981417) from Malaysia and the two are quite different from the rest of the clade that contains specimens from Malaysia and northern Myanmar. The type locality for *L. limborgi* is the "Tenasserim", Myanmar. Thus, this clade likely represents multiple species that needs further investigation. Our Tanintharyi specimen represents the closest sampled to the type locality.

Specimens examined. USNM 586920, USNM 587100, USNM 587305

Red List status. LC

Additional *Limnonectes*. We sequenced an additional specimen (MBM-USNM-FS 36471) from Mandalay Region and it was placed in a 16S clade with specimens representing a newly described species, *L. longchuanensis,* from China and northern (Kachin, Chin, and Sagaing) Myanmar (Suwannapoom et al. 2016).

Occidozyginae – puddle frogs

***Ingerana tenasserimensis* (Sclater, 1892)**
Tenasserim Trickle Frog

Description. Adult females (*n* =3) 17.6–19.7 mm SVL.

Natural history notes. These frogs were found along the edges of streams in areas of low falls caused by flat rocky outcrops. All were gravid with two or three unpigmented ova.

General Distribution. Eastern Myanmar and adjacent Thailand to northern West Malaysia.

Molecular Data. We included three individuals from Mon State (USNM 587300, USNM 587302, and MBM USNM-FS 35684). These and our specimens were placed in two separate COI BINS (ADG3231 and ADG3230) that were 7.04% sequence divergence, and they all formed a 16S clade with another individual in GenBank from near Dawei (KF991266; CAS 246787) also identified as *Ingerana tenasserimensis*.

Comments. Two other specimens in GenBank are placed sister to our clade based on 16S, one (KR827831) identified as *I.* te*nasserimensis* from Thailand (Grosjean et al. 2015b) and another (KU589219) from India identified as "*I.* sp. SB2016". These two individuals each probably represent a different species. The type locality for this species is Tenasserim, southern Myanmar. Therefore, if this species is split into multiple species, our specimens likely represent true *I. tenasserimensis*. Another sequence in GenBank (AY322302) is identical to another (DQ283235) identified as *I. borealis*.

Specimens examined. MBM-USNM-FS 35684, USNM 587300, USNM 587302, USNM 587306, USNM 586921–923.

Red List status. LC.

***Occidozyga lima* (Gravenhorst, 1829)**
Gray-green Puddle Frog

Description. Adult females (*n* = 3) 26.5–32.3 mm SVL, adult male (*n* = 1) 26.5 mm SVL. All had strongly tuberculate skin dorsally on trunk, bold black horizontal stripe on rear of thighs, and strongly patterned venter with pair of dark chin stripes.

Natural history notes. These frogs occurred in human-modified habitats. All females were gravid.

General Distribution. Widespread, eastern India to southern China southward through Southeast Asia to Java.

Molecular Data. We included one individual from Sagaing (USNM 520376) and one from Mandalay (MBM-JBS 5405). These two were placed in the same COI BIN and the Tanintharyi specimens formed a separate BIN. These were sister to each other in our combined analysis (Fig. 2). These specimens formed a 16S clade with specimens from GenBank identified as *O. lima* from Java (AB530619), Myanmar (DQ283224), Thailand, Cambodia, and Laos, (KR827958–60, respectively). We note other specimens identified as *O. lima* in GenBank are placed elsewhere in the tree but are misidentified, such as AF215398 placed with *O. laevis*, and AB488903 placed with *O. martensii* specimens.

Comments. The Common Puddlefrog in Myanmar or the frogs that have been identified as *O. lima* contain at least three distinct morphotypes. The taxa vary in size and coloration. The southern Tanintharyi "*O. lima*" is smaller and has a bold black and white ventral pattern lacking in the "*O. lima*" from northern Mon State and adjacent Bago, but it does share the bold, dark thigh stripe of the northern frogs.

Specimens examined. MBM-JBS 5405, USNM 520376, USNM 586924–927.

Red List status. LC.

***Occidozyga martensii* (Peters, 1867)**
Malay Puddle Frog

Description. Adult females (n = 7) 26.4–28.4 mm SVL, adult males (n = 7) 19.0–24.3 mm SVL. Dorsal skin lightly rugose; dorsum dusky brown and few individuals with a faint pattern of mid-dorsal dark stripe bordered by lighter parasagittal stripe on each side; posterior thigh with faint and narrow dark horizontal stripe; venter immaculate from chest to pubis, chin and throat dark in males, dusky to immaculate in females.

Natural history notes. Found in flooded fields and other human-modified habitats.

General Distribution. Tanintharyi to northern West Malaysia, Thailand to southern China.

Molecular Data. The Tanintharyi specimens were nearly identical to each other (<1% sequence divergence COI and 16S) and formed a 16S clade with specimens in GenBank identified as *O. martensii* (AB530610 and KP318725) from Thailand. Other sequences in GenBank identified as *O. martensii* from Vietnam (AF285214, DQ283357) and Yunnan China (DQ458255–56) form a separate clade sister to ours, indicating this may represent a species complex.

Specimens examined. USNM 586930–943.

Red List status. LC.

***Occidozyga* sp. A–D**
Puddle frogs

Description. We were unable to identify two of the Tanintharyi specimens to species, an adult female 20.8 mm, adult male 22.1 mm SVL. Dorsal skin lightly rugose; broad mid-dorsal brown stripe bordered by broad parasagittal tan stripes on dorsum; posterior thigh without dark horizontal bar although with sharp delineation between dorsal brown and ventral white; venter immaculate white.

Natural history notes. Occurred in the same area as the previous two *Occidozyga* species.

General Distribution. *Occidozyga* sp. A–B are known from Yangon, *Occidozyga* sp. C is known from Bago, and *Occidozyga* sp. D is known from Tanintharyi.

Molecular Data. We included other *Occidozyga* from the legacy collection for comparative purposes. These individuals were very different genetically from the Tanintharyi specimens, and some from each other, including individuals from the same geographic regions forming different clades. This is likely a cryptic species complex; therefore, we treat each of these clades as separate, unidentified species, each was placed in its own COI BIN: *O. sp.* A from Yangon (USNM 587386, USNM 587389, and MBM JBS 19932; ADG1328), *O. sp.* B from Yangon (USNM 587395 and USNM 587402; ADG1330), *O. sp.* C from Bago (USNM 587105 and USNM 587107; ADG2685), and our specimens from the Tanintharyi Region as *O. sp.* D (USNM 586928–29; ADG1329).

Specimens examined. spA-DGM2018: USNM 587386, USNM 587389, JBS 19932; spB-DGM2018: USNM 587395, USNM 587402; spC-DGM2018: USNM 587105, USNM 587107; spD-DGM2018: USNM 586928–29.

Red List status. NE.

Additional dicroglossids. We sequenced several other dicroglossids for comparison, including three specimens of *Hoplobatrachus tigerinus* from Yangon (USNM 587325, USNM 587404) and Bago (MBM-USNM-FS 35607). These were placed in a 16S clade with specimens from GenBank labeled as *H.* cf. *tigerinus* MS 2009 (AB530502 and AB543600) and MS 2011 (AB671173–81). These specimens are now considered to be *H. litoralis*, a recently described species from Cox's Bazar district of Bangladesh (Hasan et al. 2012b), which extends the range of this species into Myanmar. Our 16S sequences range from 2.0–3.5% sequence divergence from the Bangladesh sequences, including one of the paratypes (AB671174). Two specimens of *Hoplobatrachus rugulosus* from Sagaing (USNM 520480, USNM 524038) were sequenced and placed in a 16S clade with other individuals in GenBank identified as *H. rugulosus*. We sequenced two individuals of *Sphaerotheca breviceps* from Sagaing (USNM 524020, USNM 537466) that were placed in a 16S clade with individuals in GenBank identified as *S. breviceps.*

Megophryidae

Though we did not encounter any megophryid frogs during our surveys, we sequenced one *Leptobrachium smithi* (USNM 572047) from Mon State and one *Leptolalax* (USNM 572048) from Mandalay. There are three species of *Leptolalax* known to occur in Myanmar (fide Frost 2017): *L. lateralis*, *L. melanoleucus,* and *L. pelodytoides*. The 16S data from our specimen is 90% similar to four species in GenBank (*L. bourreti*, *L. fuliginosus*, *L. petrops*, and *L. tengchongensis*), while it ranges from 85–89% similar to *L. melanoleucus* and *L. pelodytoides*. No *L. lateralis* genetic data are available for comparison; however, this species is known only from northern Myanmar, from Bhamò to Nagaland, northeastern India. We tentatively refer to our specimen as *Leptolalax* sp. A.

Microhylidae (Suppl. material 1: Fig. 4)

***Kaloula latidisca* Chan, Grismer & Brown, 2014**
Malay Painted Sticky Frog

Description. Immature female 41.6 mm SVL.

Natural history notes. This single individual was found in a field near the village.

General Distribution. Tanintharyi, Myanmar to northern peninsular Malaysia.

Molecular Data. Our individual was placed at the base of a 16S clade containing many other specimens in GenBank, some labeled *K. baleata* (AB634687, KC822570, KM509153) and many others labeled *K.* sp. from Palawan, Peninsular Malaysia, Sulawesi, and Vietnam. Two other individuals in GenBank identified as *K. baleata* (KC179969, KC180032) were placed elsewhere in the tree with other specimens identified as *K.* sp. from Vietnam. These sequences are from a study focused on the Philippine Archipelago (Blackburn et al. 2013), in which these new species in the *K. baleata* complex were identified, each from Vietnam, Peninsular Malaysia, Palawan, and Sulawesi, and *K. baleata* was restricted to Java. The Tanintharyi specimen is at the base of the 16S clade (with the addition of 12S data, GenBank MG944815) containing the Palawan, Peninsular Malaysia, Sulawesi specimens, and the Vietnam specimens are elsewhere in the tree. Chan et al. (2013) described the Vietnam specimens as *K. indochinensis*, and Chan et al. (2014) described the Peninsular Malaysian specimens as *K. latidisca.* Given the geographic proximity, the Tanintharyi specimen likely represents *K. latidisca*, which could be confirmed with additional sequence data. We note one individual from the Blackburn et al. (2013) study (TNHC 67086) identified as *K.* sp. nov. Vietnam, but is here placed in the *K. baleata* sensu stricto clade. This specimen is actually from Java, thus incorrectly labelled in the 2013 study.

Specimen examined. USNM 586944.

Red List status. NE.

Figure 4. Selected turtles and lizards found during this study's expedition. **A** *Indotestudo elongata* (USNM HerpImage 2896) **B** *Dogania subplana* (USNM HerpImage 2897) **C** *Acanthosaura crucigera* (USNM 587019) **D** *Eutropis multifasciata* (USNM 587035) **E** *Takydromus sexlineatus* (USNM 587034) **F** *Calotes emma* (USNM 587022). Photos A & C by Daniel G. Mulcahy, all others by Myint Kyaw Thura.

Kaloula pulchra (Gray, 1831)

Asian Painted Sticky Frog

Description. Immature female 52.8, immature male 52.9 mm SVL.

Natural history notes. Both individuals were collected from the border of a hotel's parking lot in Myeik.

General Distribution. Widespread, northeast India and Bangladesh to southern China and Thailand southward through Thai-Malay Peninsula to Greater Sunda Islands.

Molecular Data. We sequenced other individuals from Sagaing (USNM 520322, USNM 520326, USNM 523967), Mandalay (MBM-USNM-FS 36482), Bago (MBM-USNM-FS 35512), and Yangon (MBM-JBS 19849). Our specimens were placed into three COI BINS, one for the Bago specimen, one for the Sagaing and Mandalay specimens, and one of our Tanintharyi specimens (USNM 586946) was placed in a BIN with the Sagaing and Mandalay specimens, and the other (USNM 586945) was placed a BIN with the Yangon specimen. This BIN contains nine individuals from Vietnam, Cambodia, Thailand, and Myanmar. Other specimens in BOLD identified as *K. pulchra* are placed in a different BIN, but these records are not publicly available. This variable placement suggests significant genetic variation in this group and likely indicates a species complex. Our specimens were all similar to one another based on 16S data and were placed in a 16S clade with many other individuals in GenBank identified as *K. pulchra*.

Specimens examined. USNM 586945–46, USNM 520322, USNM 520326, USNM 523967

Red List status. LC.

Microhyla butleri Boulenger, 1900

Butler's Narrow-mouthed Frog

Description. Immature males 22.9, 24.2 mm SVL.

Natural history notes. All three species of *Microhyla* were captured in the same flooded fields. Data were not taken on which species were calling or difference in vocalization.

General Distribution. Northeast India to southern China and Taiwan southward through Myanmar and Southeast Asia to Singapore.

Molecular Data. We sequenced one individual from Yangon (MBM-JBS 2952). It was placed in a separate COI BIN from our Tanintharyi specimen, and they were placed in a 16S clade with many other *M. butleri* sequences in GenBank.

Comments. Of the three species in the voucher collection, the two *M. butleri* do not display expanded vocal sacs and internally the testes appear immature.

Specimens examined. USNM 586947–948.

Red List status. LC.

Microhyla fissipes Boulenger, 1884

Oriental Ornate Narrow-mouthed Frog

Description. Adult females (n = 3) 24.6–28.3 mm, adult males (n = 3) 25.9–27.0 mm SVL.

Natural history notes. Based on our limited sampling, this species appears to have been the most abundant of the breeding *Microhyla* in the flooded fields.

General Distribution. Southern and central China, Myanmar and Southeast Asia to Singapore.

Molecular Data. We sequenced other individuals from Sagaing (USNM 520349), Mandalay (USNM 587159), Magway (USNM 587166), Bago (USNM 587110), and Yangon (MBM-JBS 19916). These latter specimens were all placed in a single COI BIN, our Tanintharyi specimens were placed in a separate COI BIN. Based on 16S data, the northern Myanmar samples were similar to our Tanintharyi specimens; all were place in a 16S clade with other specimens in GenBank identified as *M. fissipes*, *M. ornata*, *M. mukhlesuri*, and *M. mymensinghensis*. The latter species formed a clade nested within the greater *M. fissipes* clade, which may be an artifact of limited data (only 16S). *Microhyla ornata* sequences from GenBank were nested throughout this clade (see comment below).

Comments. *Microhyla fissipes* was recently resurrected (Matsui et al. 2011) for populations ranging in Myanmar, Indochina, and China previously recognized as *M. ornata*. *Microhyla ornata* is now restricted to southern India and Sri Lanka. While we are using the name *M. fissipes*, we recognize that Myanmar and Indo-China populations are a different species than the eastern China ones from which the holotype of *M. fissipes* derives; however, no systematist has yet sorted out the taxonomy of these more western Southeast Asian populations.

Specimens examined. USNM 586949–954, USNM 520349, USNM 587159, USNM 587166, USNM 587110.

Red List status. LC.

Microhyla heymonsi Vogt, 1911

Black-sided Narrow-mouthed Frog

Description. Adult female (*n* = 1) 25.1 mm, adult males (2) 20.1, 20.6 mm SVL.

Natural history notes. From flooded fields.

General Distribution. Northeast India through southern China to Taiwan, southward through Southeast Asia to Sumatra.

Molecular Data. We sequenced four other individuals, two from Bago (USNM 587130, MBM-USNMFS 35509) and two from Mandalay (USNM 587138, USNM 587140). These were each placed in their own COI BIN, as were our specimens from the Tanintharyi. These were similar to our Tanintharyi specimens based on 16S data and were all placed in a 16S clade with specimen in GenBank identified as *M. heymonsi*. There were two distinct clades within the GenBank *M. heymonsi* material; ours were placed in one with other specimens from Myanmar (e.g. KC179993), and Singapore (e.g. HM359093). Sheridan et al. (2010) identified three clades within *M. heymonsi*. These data indicate this represents a species complex in need of further revision.

Specimens examined. USNM 586955–957, USNM 587130, USNM 587138, USNM 587140.

Red List status. LC.

Additional *Microhyla*. We sequenced nine additional specimens of *Microhyla* from northern Myanmar. Four (USNM523975, USNM523976, USNM523979,

USNM 537450) from Sagaing did not match any species description, were placed in their own COI BIN, were placed in their own clade in the 16S tree, and likely represent a new species (*M.* sp. A.). One identified as *M. rubra* was placed in a COI BIN with another individual identified as *M. rubra* from Myanmar and the 16S sequence is identical to a specimen in GenBank (KM509166) from Magway, Myanmar. Other specimens identified as *M. rubra* in GenBank were placed elsewhere in the tree. However, six of these (KU214856–61) represent a recently described species (*M. mihintalei,* Wijayathilaka et al. 2016), while the other two (AB201192 and KU214855) represent *M. rubra* from Sri Lanka and India (Karnataka). The type locality for *M. rubra* is "in the Carnatic near rivers, in sandy banks… also Ceylon" India and Sri Lanka (fide Frost 2017). Therefore, our specimen, the KM509166, and the other specimen in the BOLD BIN (BOLD:ACW0810) likely represent a new species. We refer to our specimen as *M.* sp. B. Two others identified as *M. berdmorei* from Yangon (MBM-JBS19917 MBM-JBS 19929), and two *M. berdmorei* from Bago (MBM-USNM-FS 35556, USNM 587407) were placed in the same COI BIN, and were placed in a 16S clade identical to one from Sagaing, Myanmar (KC179981; de Sá et al. 2012) and similar to three others identified as "*M. sp*. B MS-2009" from Bangladesh (Hasan et al. 2012a). This clade was placed sister to two other clades identified as *M. berdmorei* (and one *M. fowleri*), which suggests a species complex in need of revision. We sequenced a paratype (USNM 523965) of *Kalophrynus anya* (Zug 2015) and one *Glyphoglossus molossus* (USNM 523961), both from Sagaing, Myanmar.

Ranidae – true frogs (Suppl. material 1: Figs 5–6)

Because a considerable amount of COI barcode data are available for ranid frogs, we also conducted a similar comparison using a neighbor-joining tree with material from GenBank.

***Amolops panhai* Matsui & Nabhitabhata, 2006**
Panha Torrent Frog

Description. All juveniles (n = 12) 29.4–35.2 SVL.

Natural history notes. Moderately common in the areas of large rocks and splash zones.

General Distribution. Tanintharyi, Myanmar and western/central peninsular Thailand.

Molecular Data. Our specimens were less than 1% (COI and 16S) sequence divergence from each other. These were placed in a COI BIN with two individuals of *A. panhai* from Thailand (Grosjean et al. 2015b). Our specimens were placed in a 16S clade with other individuals from GenBank identified as *A. panhai* from Dawei, Tanintharyi, Myanmar (JF794451, Dever et al. 2012) and Thailand (AB211487–8,

Matsui et al. 2006; KR827705–6, Grosjean et al. 2015b). Our specimens were placed sister to two COI barcodes for *A. panhai* (KR087620–1), two of the same individuals for which 16S data were available (Grosjean et al. 2015b).

Specimens examined. USNM 586958–969.

Red List status. LC.

Comments. We included three individuals (USNM 564958, USNM 564961, USNM 564967) of *Amolops marmoratus* from Mon State, Myanmar. These were placed in a COI BIN with another individual from Thailand identified as *A. marmoratus*, and in a 16S clade with other individuals identified as *A. marmoratus* from Myanmar (Dever et al. 2012) in the Tanintharyi (JF794450), Mon State (JF794452–56), and Shan State (JF794470) and Thailand (AB211486, Matsui et al. 2006). Our specimens were placed sister to one COI barcode for *A. marmoratus* (KR087617), one of the same individuals for which 16S data were available (Grosjean et al. 2015b).

Chalcorana eschatia Inger, Stuart & Iskandar, 2009

Peninsular Copper-cheeked Frog
Fig. 3E

Description. An immature female, 35.4 mm SVL, two adult females, 47.3–48.4 mm and five adult males, 30.5–34.6 mm SVL. Both adult females are gravid.

Natural history notes. Streamside in the primary forest.

General Distribution. Tanintharyi Myanmar to southern Thailand.

Molecular Data. Our specimens were placed in a single COI BIN with four specimens from Thailand identified as *Hylarana eschatia*. Our specimens were placed in a 16S clade with material from GenBank, with all specimens of *C. eschatia* from neighboring Thailand, including type material (e.g. FMNH 268523–30, Inger et al. 2009). Our specimens ranged from 0–1.4% sequence divergence from these individuals in GenBank. Our specimens were placed sister to a clade of COI barcodes for *C. eschatia* (KR087702–5), some of the same individuals for which 16S date were available (Grosjean et al. 2015b).

Comments. The Malayan populations were formerly included in *Hylarana chalconota*, which is now restricted to southern Sumatra, Java, and Bali. Our findings represent new country records for this species.

Specimens examined. USNM 586971–979.

Red List status. NE.

Hylarana erythraea (Schlegel, 1837)

Green Paddy Frog

Description. Adult female, 71.8 mm SVL.

Natural history notes. This gravid female was found in a flooded field.

General Distribution. Eastern Myanmar and Southeast Asia southward to Borneo.

Molecular Data. In addition to our single specimen, we sequenced two individuals (USNM 583188 and USNM 583191) from the Yangon region. Our specimen was 2.3% sequence divergence from the Yangon specimens for COI and they were placed in separate BINs, and ranged from 0–0.2% sequence from each other for 16S, and were placed in a 16S clade with other individuals identified as *H. erythraea* from the Yangon area (KR264118–19; USNM 583188 and USNM 583190), and Tanintharyi (KR264061, KR264066; CAS 229614 and CAS 247465), Myanmar (Oliver et al. 2015), and Thailand (KR827786, Grosjean et al. 2015b). We note there are several other clades of *H. erythraea* in our 16S tree, and though they all represent a monophyletic group, there is great molecular divergence among them, indicating another species complex in need of revision. Our specimens were placed sister to a COI barcode for *H. erythraea* (KR087693), one of the same individuals for which 16S date were available (Grosjean et al. 2015b), and this clade was sister to another clade of *H. erythraea*, similar to the 16S results.

Specimens examined. USNM 586979, USNM 583188–91.

Red List status. LC.

Additional *Hylarana*. When we began the barcode analysis of our Tanintharyi *Hylarana*, the genus contained more than three dozen species. *Hylarana* sensu lato was clearly not a monophyletic group, as Oliver et al. (2015) subsequently demonstrated by restricting it to four species (*H. erythraea*, *H. macrodactyla*, *H. tytleri*, and *H. taipehensis*) of which the first three species likely have populations in Myanmar. Because of the larger content of *Hylarana* s. l., we sequenced several other individuals from elsewhere in Myanmar, including six *H. lateralis*, two from Yangon (USNM 583187 and MBM-JBS 19852), four from Sagaing (USNM 520401, USNM 523999, USNM 524000, and USNM 537463), two *H. macrodactyla* from Bago (USNM 583137 and MBM-USNM-FS 35511), one *H. macrodactyla* from Sagaing (USNM 520469), and one *Humerana* cf. *humeralis* (USNM 583171) from Bago. Our *H. lateralis* were placed in a single COI BIN, and in a 16S clade with specimens identified as *Humerana lateralis* (see Oliver et al. 2015 for new generic allocations), which is sister to *Humerana miopus*. Our COI data placed our *H. lateralis* sister to two *H. lateralis*, the same specimens as in the 16S tree (Grosjean et al. 2015b). The *H. macrodactyla* from Sagaing was placed in its own COI BIN, and the two from Bago were placed in a separate COI BIN, and in a 16S clade with several '*H.* cf. *taipehensis*' (AB530522–5, AB543603; we note that these identifications are almost certainly incorrect) from Bangladesh (Hasan et al. 2012a), and a *H.* cf. *tytleri* (KM069012) from Tripura, India (Biju et al. 2014). This clade was sister to a clade of *H. macrodactyla* from Myanmar and Laos. Our *H. macrodactyla* specimen (USNM 520469) from Sagaing was placed sister to this *H. macrodactyla* + *H.* cf. *taipehensis* clade in our 16S tree. Our COI data placed our three specimens in a clade with the "*H.* cf. *tytleri*" specimen from Tripura, India, though with considerable sequence variation. This latter clade was sister to an *H. macrodactyla* COI clade. The type locality for *H. tytleri* is Bangladesh, whereas the type localities for *H. taipehensis* and *H. macrodactyla* are Taiwan and Hong Kong, respectively. Oliver et

al. (2015) identified four specimens as *H. tytleri*, all from Myanmar (their materials examined in Appendix A), but mistakenly labelled them in GenBank as *H. erythraea*; these specimens are placed in the *H. erythraea* clade of the 16S tree. Clearly, *H. erythraea* is another group in need of revision. We tentatively refer to our Bago specimens as *H.* cf. *tytleri* because the *H.* cf. *tytleri* Tripura, India (Biju et al. 2014) specimen is the closest geographically to the type locality of *H. tytleri* and we refer to our Sagaing specimen (USNM 520469) as *H.* sp. A. Our *Humerana humeralis* is identical to a specimen (USNM 583170, collected contemporaneously and already in GenBank, KR264113) identified as *Humerana* sp., and the two were placed sister to other specimens identified as *Humerana* cf. *humeralis* (KM069010) and *Humerana humeralis* (KU589217, KU589223–4); though with considerable genetic differences (6–15% sequence divergence) this clade ranges from Assam, India to Bago, Myanmar and the type locality is Bhamò, Kachin State, Myanmar. The COI data for our specimen are considerably different (>18%) from the *Humerana* cf. *humeralis* from Assam, India.

***Odorrana hosii* (Boulenger, 1891)**
Green Odor Frog
Fig. 3D

Description. A total of 17 individuals were collected in the Tanintharyi. Two adult females, 78.3, 87.8 mm SVL, adult males (n = 13) 52.5–60.4 mm SVL (measurements for adults only). All individuals share dark lores, a white upper lip with white stripe extending to above axilla, and an immaculate (nearly white) venter from chin to pubic area.

Natural history notes. The unpigmented follicles and enlarging oviducts in females and the modest ductus deferens and no external visible vocal sacs of males suggest that breeding had not yet begun in this population. All were found on branches over and adjacent to forest streams.

General distribution. Peninsular Myanmar, Thailand, West Malaysia to Sumatra and Borneo.

Molecular data. Our specimens ranged from 0–0.5% sequence divergence from each other based on COI data and were placed in one BIN. These were also placed in a 16S clade identical to other individuals (e.g. DQ650595–604) identified as *O. hosii* from neighboring Thailand (Stuart et al. 2006a). This clade also contained two individuals identified as *O. livida* (KR827970–1) from Thailand (Grosjean et al. 2015b), which are presumably misidentified. Our specimens were placed sister to the same two specimens identified as *O. livida* (KR087841–2) in the 16S tree (Grosjean et al. 2015b), which are misidentified. Two other clades of *O. hosii* from GenBank were recovered in the 16S tree, sister to each other, and that clade is sister to the one containing our specimens.

Specimens examined. USNM 586981–87, 586991, 586993–587000.

Red List status. Least Concern.

***Odorrana livida* (Blyth, 1856)**
Cascade Odor Frog

Description. Single adult female 86.8 mm SVL, adult male 73.0 mm SVL. These individuals have strongly dusky colored chins and anterior chests.

Natural history notes. N/A

General Distribution. Northeast India to peninsular Myanmar and Thailand.

Molecular Data. We included two individuals from Mon (MBM-USNM-FS 35753, MBM-USNM-FS 35755). These and our Tanintharyi specimens ranged from 0–2.5% sequence divergence from each other based on COI data and were placed in two COI BINs (Mon and Tanintharyi) and were placed in a 16S clade with other individuals (DQ650612–615) identified as *O. livida* from neighboring Thailand and Myanmar (Stuart et al. 2006a). Two other specimens from GenBank (AB200949–50) identified as *O. supranarina*, from the Ryukyu Islands, Japan, were also in this clade. However, these specimens may be misidentified, because a sequence identified as *O. livida* (AB200955) from the same study (Matsui et al. 2005b) was placed in the *O. chloronota* clade, and the *O. chloronota* sequence (AB200954) from that study (Matsui et al. 2005b) was placed in the *O. graminea* clade. There were no other *O. livida* COI sequences available, except for the two mis-identified specimens (see *O. hosii* above).

Specimens examined. USNM 587001–02.

Red List status. DD (Data Defficient).

Additional *Odorrana*. We sequenced one individual (USNM 587323) of *Odorrana* from Mandalay, Myanmar. This individual was placed in its own COI BIN and in a 16S clade with specimens identified as *Odorrana graminea* (KR827967–8) and *O.* cf. *chloronota* (DQ650605–11) from Thailand (Stuart et al. 2006a). Therefore, we tentatively identify this specimen as *O.* cf. *chloronota* (sensu Stuart et al. 2006a).

***Sylvirana malayana* Sheriden & Stuart, 2018**
Black-sided Forest Frog
Fig. 3H

Description. Immature female, 39.9 mm SVL.

Natural history notes. Specimen was found adjacent to the forest stream.

General Distribution. Once thought to be widespread, Nepal, northern peninsular and Northeast India to southwest China and Southeast Asia. However, several recent studies based on molecular data suggest "*S. nigrovittata*" represents a multiple-species complex, with this species (*S. malayana*) being recently described from the Thai-Malay Peninsula our our specimen extends the range into central Tanintharyi, ostensibly overlapping with *S. nigrovittata* sensu strico (see below).

Molecular Data. One specimen was collected in 2014, it was placed in its own COI BIN and in a 16S clade with other individuals identified as *S. nigrovittata* in GenBank from Phang-Nga, Thailand (KR827826, Grosjean et al. 2015b), and other unpublished sequences that lack specimen information (KF738999–9002, EU604197).

Our specimen was placed sister to the COI barcode for the same individual for which 16S date were available (Grosjean et al. 2015b). A very recent paper (Sheriden and Stuart 2018), published during revision of this manuscript, describes four new species in this complex. Our specimen falls (not shown) within one of their newly described peninsular-Malaysian species (*S. malayana*), extending it into Tanintharyi. Dubois (1992) designated a lectotype of *Limnodytes nigrovittatus* Blyth 1856 and restricted the type locality to "Mergui and the valley of the Tenasserim River." Mergui is present day Myeik, adjacent to the mouth of the Tanintharyi River. We find it peculiar that our specimen, from a tributary of the Tanintharyi River, represents this newly described species and not *S. nigrovittata* as it was collected in between their *S. nigrovittata* genetic samples (from western Thailand) and the type locality (Myeik). Nevertheless, two clades appear to extend across the Isthmus of Kra on the eastern (*S. nigrovittata*) and western (*S. malayana*) sides. The type of *S. nigrovittata* is a female specimen for which distinguishing characteristics are lacking. Until sequence data can be obtained from the type specimen, or additional material can be collected from the Myeik area proper, we remain skeptical that the newly described *S. malayana* may represent *S. nigrovittata* sensu stricto and populations sampled by Sheriden and Stuart (2018) to the north, and east may represent a new species.

Specimens examined. USNM 586970.

Red List status. LC.

Additional *Sylvirana*/*Hylarana*. We sequenced several other individuals from northern Myanmar identified as *Sylvirana*/*Hylarana* sp. We found several clades of "*S. nigrovittata*" in our 16S tree. We sequenced two individuals from Mon State (USNM 583176 and USNM 583178) and one from Mandalay (USNM 583174) that were identified as *S. menglaensis* and were placed in two COI BINs (Mon and Mandalay). The Mandalay specimen was placed in a BIN with nine other individuals, seven identified as *H. menglaensis*, and two as *S. nigrovittata*. The COI data placed these individuals in a clade with other *S. menglaensis* for which COI data were available, with the BINs forming clades, with other clades of specimens identified as *S. menglaensis*. These were all placed in a 16S clade containing other specimens identified as *S. menglaensis* in GenBank (KR827810–22, Grosjean et al. 2015b). Sheriden and Stuart (2018) placed *S. menglaensis* in synonymy with *S. nigrovittata*. We sequenced three additional specimens from Mandalay (USNM 583124, USNM 583126, and MBM-USNM-FS 36020). These specimens were placed in their own COI BIN and in a 16S clade, nearly identical to two individuals identified as *Hylarana* sp. C MS-2010 (AB543604–5, Hasan et al. 2012a) from Bangladesh, and two sequences in GenBank (KR264116–7), from the same series as ours (USNM 583124–5) identified as *Sylvirana* cf. *nigrovittata* (Oliver et al. 2015). Our sequences differ by two base-pairs from the ones in GenBank, including USNM 583124 (Oliver et al. 2015). We sequenced this specimen twice and provide the raw trace files in BOLD. There are many other 16S sequences in GenBank identified as *Sylvirana nigrovittata* elsewhere in the tree, but no other COI sequences to compare. Sheriden and Stuart (2018) described specimens of this clades as a new species *S. lacrima*, and our specimens fall out within this clade (not shown).

Rhacophoridae – Whipping Frogs (Suppl. material 1: Fig. 7)

***Polypedates* cf. *leucomystax* (Gravenhorst, 1829)**
White-lipped Tree Frog

Description. Three immature females 61.7, 62.8, 69.7 mm, adult female 82.0, and two adult males 45.7, 45.8 mm SVL.

Natural history notes. Two species of *Polypedates* were found in a rural landscape during heavy rains. At the time of collection, all specimens (six adult males and ten females) were assumed to represent a single species and the location of individual specimens was not noted, though all were collected from the same flooded fields. The results from the barcode analysis revealed that two genetic lineages were present in the total *Polypedates* sample, and one of lineages was represented by only males and the other by only females. The males (45.3–50.2 mm SVL) have the vocal sacs open although there is no indication externally (i.e., stretched throat skin and pigmented) and the testes are enlarged. The majority (n = 9) of the females range from 60.7–70.0 mm SVL; their oviducts have only begun to enlarge and the follicles within the ovaries are small and presumably pre-vitellogenic or in early vitellogenesis; a single large female 82.0 mm SVL has mature oviducts and ovarian follicles are well yolked but not pigmented. All females have distinct dark brown longitudinal stripes (commonly broken) on the dorsum; stripes are absent or reduced on most of the males. Additionally, the lower lip of the females is black bordered and immaculate in the males. It is notable that without the barcode data, we would have interpreted the vouchers as a single species with distinctly smaller males and larger females. The reproductive data suggest that the smaller species breeds early in the monsoon and the larger one in the late monsoon or early dry season.

General Distribution. Widespread in South Asia, eastern India to southwestern China through Southeast Asia to the Greater Sunda Islands and Philippine Islands.

Molecular Data. Our specimens were placed in their own COI BIN, and were placed at the base of a large 16S clade of *P. leucomystax*, from GenBank (sensu Kuraishi et al. 2012). These specimens may represent a new species, closely related to *P. leucomystax* (see comments below). For now, we refer to them as *P.* cf. *leucomystax.*

Comments. Initially, the specimens collected were considered to represent a single species, but the DNA barcoding revealed two distinct lineages. The *P. leucomystax* complex of frogs remains contentious. Several recent studies have produced 16S (e.g. Kuraishi et al. 2012, Pan et al. 2013) and COI (Buddhachat and Suwannpoom 2018) sequence data, resolving some of the issues within this group. Our Tanintharyi frog loosely fits the morphological description of the *P. leucomystax.* Our clade was placed sister to the COI *Polypedates* cf. *leucomystax* clade of Buddhachat and Suwannpoom (2018; not shown) However, a detailed morphological comparison and additional sequence data are supporting our lineage represents a new species that occurs from northern Tanintharyi, and further to the north in Myanmar (Wilkinson, Mulcahy, Zug, in prep.).

Specimens examined. USNM 587003–7008

Red List status. *Polypedates leucomystax* is listed as LC.

***Polypedates mutus* (Smith, 1940) Burmese**
Whipping Frog

Description. All individuals are immature; six females 59.9–66.6 mm, three males 45.3–50.3, and a sex indeterminate specimen 45.2 mm SVL (*n*=10).

Natural history notes. See preceding species account.

General Distribution. Tanintharyi Myanmar, Thailand, Yunnan and Guangxi, China.

Molecular Data. The second clade of our *Polypedates*, and an additional specimen from Mandalay (USNM 587059) were placed in a COI BIN with four individuals from Thailand identified as *Polypedates* sp. Our specimens were placed in a 16S clade with individuals identified as *P. impresus* (Pan et al. 2013) and *Polypedates* cf. *mutus* 2 of Kuraishi et al. (2012). Note, older specimens in GenBank in this clade are labeled as *P. leucomystax, P. megacephalus,* and *P. sp.* The very recently published paper examining Thailand species of *Polypedates* with COI data (Buddhachat and Suwannpoom 2018) identified five major clades in the *P. leucomystax* complex. Our specimens were placed in their "Northern A *Polypedates* sp." clade (not shown). Our ongoing work (Wilkinson, Mulcahy, Zug, in prep.) suggests that this clade represents *P. mutus* senus stricto (the *P. mutus* 1 clade of Kuraishi et al. 2012).

Specimens examined. USNM 587009–018, USNM 587059.

Red List status. LC.

Additional rhacophorids. We sequenced three individuals initially identified as *Polypedates teraiensis* from Sagaing (USNM 524030), Yangon (USNM 587048), and Bago (587049). These were all placed in their own COI BIN and in a 16S clade with specimens from GenBank identified as *P. teraiensis* (AB530512–21) and two individuals (AB728167–8) labeled *P. leucomystax*, presumably misidentified. Additionally, we sequenced three additional specimens initially identified as *Chiromantis* spp. Two specimens (USNM 560923, USNM 560927) from Mandalay initially identified as *C. hansenae*, were placed in their own COI BIN and in a 16S clade with other individuals from GenBank identified as *Chiromantis doriae*. There is considerable genetic variation among the *C. doriae* specimens in GenBank, indicating that *C. doriae* as currently used is a species complex in need of revision. The third specimen (USNM 524023) from Sagaing was initially identified as *C. nongkhorensis*, but was placed at the base of the 16S clade containing *C. nongkhorensis* and *C. doriae* specimens from GenBank (Aowphol et al. 2013). This specimen may represent a new species; however, we treat it as *Chiromantis* sp. A for now.

Testudines
Testudinidae – tortoises

***Indotestudo elongata* (Blyth, 1853)**
Elongate Tortoise
Fig. 4A

Description. A shell of this species was seen in Yeybu village. Carapace length (straight) was approximately 22 cm; sex indeterminate owing to absence of a plastron. Nine distinct growth annuli were visible on the second right pleural scute.

Natural history notes. The tortoise from which the shell was derived was presumably from the adjacent forest.

General Distribution. Widespread, Nepal, northern peninsular and Northeast India to Southeast Asia into northernmost West Malaysia.

Molecular Data. No molecular data available.

Specimens examined. The specimen was found in a camp and photo vouchered. USNM Herp Image 2896

Red List status. EN (Endangered); CITES Appendix II.

Trionychidae – softshell turtles

***Dogania subplana* (Saint Hilaire, 1809)**
Hillstream Softshell Turtle
Fig. 4B

Description. No measurements were taken.

Natural history notes. Uncertain origin, but presumably from the nearby river.

General Distribution. Myanmar and Thailand through Peninsular Malaysia, south to Sumatra, Java, and Borneo.

Molecular Data. No molecular data available.

Comments. This individual was seen in Yeybu village. No measurements were taken.

Specimens examined. USNM Herp Image 2897

Red List status. LR/LC (Low Risk/Least Concern; needs updating); CITES Appendix II.

Squamata
Lizards
Agamidae – spiny lizards (Suppl. material 1: Fig. 8)

***Acanthosaura crucigera* Boulenger, 1885**
Masked Prickly-naped Lizard
Fig. 4C

Description. Adult male 110.4 mm SVL, 222 mm TailL, 27.3 mm HeadL; 42% TrunkL/SVL, 30% Forarm/CrusL, 25% HeadL/SVL, 82% HeadW/HeadL, 61% HeadH/HeadL, 61% SnEye/HeadL, 23% EyeEar/HeadL.

Acanthosaura crucigera is a striking lizard with its postorbital and nape spines, highlighted by dark brown face mask bordered below by white lips and jowl. This specimen has seven cervical spines, 20 fourth finger lamellae and 26 fourth toe lamellae.

Natural history notes. Only one individual was seen during the six days of survey days in the forest. This individual was on a branch overhanging the stream.

General Distribution. Tanintharyi Myanmar through southern Thailand and Cambodia, southward to northern West Malaysia.

Molecular Data. No COI sequences are currently available for *A. crucigera*, and our specimen was placed in its own COI BIN and differs from other *Acanthosaura* species by 16–18%. Our specimen is 97% similar to an *A. crucigera* in GenBank (AB031980) from Koh Chang Island, Thailand (Honda et al. 2000a) based on 16S data and is placed sister to this specimen in our 16S tree.

Specimens examined. USNM 587019.

Red List status. NE.

***Calotes emma* (Gray, 1845)**
Barred Forest Lizard
Fig. 4F

Description. Adult females (*n* = 2) 101.7, 106.1 mm SVL, 279, 284 mm TailL, adult males (*n* = 2) 84.5, 102.2 mm SVL, 236, 279 mm TailL; 16.4 , 17.4 mm & 12.2, 16.7 mm, respectively HeadL; 46–50% TrunkL/SVL,14–17% HeadL/SVL, 68–79% HeadW/HeadL, 33–46% HeadH/HeadL, 43–44% SnEye/HeadL, 25–29% EyeEar/HeadL. Dorsal spines range from 39 to 52, fourth finger lamellae 22 to 25, and fourth toe lamellae 25 to 31.

Natural history notes. Although *C. emma* was found within the forest, these individuals were in an area of open canopy. It appears to be mainly a forest-edge and woody fencerow denizen.

General Distribution. Northeast India to southwestern China southward to northern West Malaysia.

Molecular Data. Our specimens were placed in their own COI BIN. No other COI sequences are currently available for *C. emma*, and our sequences differ from other *Calotes* species by 10–12%. Based on 16S data, our specimens were placed sister to all other sequences of *Calotes* currently in GenBank.

Specimens examined. USNM 587020–023.

Red List status. NE.

***Draco blanfordii* Boulenger, 1885**
Blandford's Draco

Description. Adult females (*n* = 2) 101.4, 107.9 mm SVL, adult male (*n* = 1) 96.4 mm SVL, 183–190, 178 mm TailL respectively; 175, 178 mm & 174 mm HeadL; 46–50% TrunkL/SVL, 16–18% HeadL/SVL, 68–74% HeadW/HeadL, 42–51% HeadH/HeadL, 11% SnEye/HeadL, 25–29% EyeEar/HeadL. Fourth finger lamellae range from 26 to 28 and fourth toe lamellae 26 to 30.

Natural history notes. As for the preceding *C. emma, Draco* occurs in open-canopied situations and along forest edges. Each of the females was gravid; the smaller female bears four eggs, two on each side; the larger one has seven eggs, three on right, four on left. The shelled eggs are ~10–12 mm in length.

General Distribution. Bangladesh through southern Myanmar and Thailand to northern West Malaysia, also Vietnam.

Molecular Data. Our specimens were placed in their own COI BIN, no other COI sequences are currently available for *D. blanfordii*. Our 16S sequences are 100% identical to *D. blanfordii* specimens in GenBank (AB023751) from Thailand and Peninsular Malaysia (Honda et al. 1999).

Specimens examined. USNM 587024–026.

Red List status. NE.

Gekkonidae – geckos (Suppl. material 1: Figs 9–10)

***Gekko gecko* (Linnaeus, 1758)**
Tokay Gecko

Description. Adult male 179 mm SVL, 111 mm TailL (half regenerated), 81.7 mm TrunkL, 19.7 ForeaL, 24.3 mm CrusL, 43.9 mm HeadL, 35.4 mm HeadW, 17.9 mm SnEye, 13.7 mm NarEye, 16.4 mm SnW. Adult proportions 46% TrunkL/SVL, 34% SnFor/SVL, 41% ForeaL/CrusL, 18% CrusL/SVL, 24% HeadL/SVL (23.0±1.0), 81% HeadW/HeadL, 41% SnEye/HeadL, 43% EyeEar/HeadL.

Head and trunk scalation predominantly granular, enlarged scales bordering the mouth, 13 supralabials to rictus, 10 infralabials; 96 scales at midbody, 148 scales ventrally from mental to vent border; 21 enlarged lamellae on 4th finger, 20 on fourth toe.

Natural history notes. This specimen derived from Yeybu village; others were heard in the residual forest at the July slash and burnt site.

General Distribution. Widespread, Nepal and Northeast India to South China southward into Lesser Sundas.

Molecular Data. Our specimen was placed in its own COI BIN and was between 94.5–96.3% similar to other *G. gecko* COI in GenBank. Our sequence was placed in a clade at the base of the *G. gecko* clade; these basal diverging clades show substantial sequence divergence, indicating this may represent a species complex.

Specimens examined. USNM 587027.

Red List status. NE.

Hemidactylus berdmorei (Blyth, 1853)

Tanintharyi Smooth Gecko

Comments. No species of the *Hemidactylus bowringii* group were seen at any of the sites visited. *Hemidactylus berdmorei* is known only from a disintegrating holotype collected in 1853 in Mergui (= Myeik), hence vouchers of the Smooth Gecko from this area are essential for resolving the taxonomic status of this named taxon (McMahan and Zug 2007). If it is an autochthonous species, this gecko is likely a valid species; however, because Mergui was an active seaport in the mid 19th century, it is possible that the specimen was a recent arrival from elsewhere in Asia and did not become established.

Red List status. NE.

Hemidactylus frenatus Duméril & Bibron, 1836

Indo-Pacific House Gecko

Description. Adult females (n = 2) 52.1–53.2 mm SVL, 51–37 mm TailL; both regenerated; adult male (n = 1) 54.9 mm SVL; 25.3 mm TailL; 23.8 mm HeadL; 43–49% TrunkL/SVL, 13–15% CrusL/SVL, 25–26% HeadL/SVL, 38–39% HeadW/HeadL, 64–66% HeadH/HeadL, 42–44% SnEye/HeadL, 29–33% EyeEar/HeadL.

Natural history notes. A synanthrope. Collected on the outside wall of the hotel in Myeik.

General Distribution. Widespread human commensal, worldwide in subtropics and tropics.

Molecular Data. Three specimens of *H. frenatus* were collected in Myeik. They were placed in two COI BINs, two (USNM 587030 and 587032) in their own BIN, and USNM 587031 was placed in the same BIN with 18 other *H. frenatus* from Honduras. Our specimens were placed with other specimens in GenBank from the "Myanmar clade" of *H. frenatus* from Tonione et al. (2011). Two other specimens were placed together, on a long branch, sister to all other *H. frenatus*. These specimens were subsequently identified as *H. tenkatei* (see below).

Natural history notes. Individuals collected from the outside wall of a hotel.

Comments. The House Gecko is a widespread and invasive species. It seldom occurs on vegetation away from human buildings. The tissue gathered from this small sample and barcode analyzed reveals that the Myeik population contains two genetic lineages.

Specimens examined. USNM 587030–032.

Red List status. LC.

Hemidactylus garnotii Duméril & Bibron, 1836

Fox Gecko

Description. Adult female 59.0 mm SVL. 14.6 mm HeadL; 47% TrunkL/SVL, 14% CrusL/SVL, 25% HeadL/SVL, 79% HeadW/HeadL, 34% HeadH/HeadL, 47% SnEye/HeadL, 28% EyeEar/HeadL, 14% SnW/HeadL.

Natural history notes. A synanthrope.

General Distribution. Widespread human commensal, native to South Asia and Pacific islands.

Molecular Data. Our specimen was placed in a COI BIN with three other *H. garnotii*, and two *H. stejnegeri.* There are currently no COI nor 16S sequences available in GenBank for *H. garnotii*. Ours is 100% identical to two other sequences in BOLD, not publicly available (from New Caledonia). However, these are also identical to two *H. stejnegeri* in BOLD, also not publicly available (from Vietnam and the United States).

Comments. The Fox Gecko is an all-female species with a broad distribution in Asia and the Pacific. Of all invasive *Hemidactylus*, it regularly occurs in the vegetation of disturbed habitats rather than on human buildings.

Specimens examined. USNM 587033.

Red List status. NE.

Hemidactylus tenkatei Lidth de Jeude, 1895

Southeast Asian Spiny Gecko

Description. Adult males (n = 2) 60.8–61.3 mm SVL, 43–30 mm TailL both regenerated; 15.7–15.8 mm HeadL; 44–45% TrunkL/SVL, 80–87% Forearm/CrusL, 26% HeadL/SVL, 75–76% HeadW/HeadL, 36–39% HeadH/HeadL, 13% SnEye/HeadL, 28–31% EyeEar/HeadL. .

Natural history notes. Collected on the outside wall of the hotel in Myeik.

General Distribution. Myanmar, West Malaysia, Timor; although likely more widespread in South Asia.

Molecular Data. Two specimens initially thought to be *H. frenatus* were placed at the base of the *H. frenatus* COI tree (see above), these were each placed in their own COI BIN. We then sequenced the ND2 locus (GenBank MG948675 and MG944816) for these individuals to align with the sequences from Kathriner et al.

(2014). Our specimens were each placed in one of the *H. tenkatei* clades of Kathriner et al. (2014). See Suppl. material 1: Fig. 10 for the ND2 tree.

Comments. Kathriner et al. (2014) have demonstrated that the Burmese specimens of this taxon from Yangon and Tanintharyi associate genetically with *H. tenkatei* from Timor and other Sundan areas.

Specimens examined. USNM 587028–29.

Red List status. NE.

Lacertidae – grass lizards

***Takydromus sexlineatus* Daudin, 1802**
Long-tailed Grass Lizard
Fig. 4E

Description. Small adult female 48.2 mm SVL, 94 mm TailL tip regenerated; 10.8 mm HeadL; 44% TrunkL/SVL, 12% CrusL/SVL, 22% HeadL/SVL, 52% HeadW/HeadL, 38% HeadH/HeadL, 45% SnEye/HeadL, 29% SnW/HeadL.

Head with 7 supralabial scales, 5th very large and beneath eye, 6 infralabials; 34 dorsal scales at nape, midbody, and above vent; 17 enlarged lamellae on 4th finger, 20 on fourth toe. In preservative, dark above and white below, dorsal ground color medium olive with lighter olive dorsolateral stripe from snout onto base of tail, medium brown loreal stripe from snout through eye to inguina and border below by white stripe from snout tip across supralabials onto trunk, fading at two-third of trunk.

Natural history notes. This species was seen in the weedy fencerows of the June gardens survey and subsequently at the July slash & burnt site. The June voucher specimen is gravid with large yolked but unshelled follicles.

General Distribution. Widespread, Northeast India through southern China to Taiwan and southward into Greater Sunda Islands.

Molecular Data. Our specimen was placed in its own COI BIN, it is 3.0–7.7% sequence divergence (COI) from four specimens in BOLD. Three of those were mined from GenBank (AY248546–48), with no locality data available.

Specimens examined. USNM 587034

Red List status. LC

Scincidae – skinks

***Eutropis macularia* (Blyth, 1853)**
Side-spotted Sun Skink

Description. Two adult females 59.4 (crushed), 61.8 mm SVL, incomplete, 83 mm regenerated TailL; NA, 12.6 mm HeadL; 45–53% TrunkL/SVL, 37–42% HindlL/SVL, NA, 20% HeadL/SVL, NA, 84% HeadW/HeadL, 51% HeadH/HeadL, NA,

47% SnEye/HeadL, NA, 32% EyeEar/HeadL. Supralabials 7, 5th largest and beneath eye, 7 or 8 infralabials; 34, 36 dorsal scale rows from nape to above vent, dorsal scales 5 to 7 keeled, predominately 7 keels; 30, 32 scales around midbody; 11 fourth finger lamellae, 14 fourth toe lamellae. In preservative, dark above and dusky below; dorsum medium reddish brown from snout onto tail, laterally lighter reddish brown broad stripe from snout to hind limbs, bordered above by tannish stripe from snout to mid trunk and below by white stripe from snout across supralabials to anterior trunk.

Natural history notes. Seen on the banks of a forest stream.

General Distribution. Widespread, Pakistan through northern peninsular India to Southeast Asia and northern West Malaysia.

Molecular Data. Our specimens were placed in their own COI BIN and are 15–16% divergent from six *E. macularia* in BOLD from Vietnam (not public). The 16S sequences are 98–99% similar to two specimens in GenBank (AY159078, KX231450) from Myanmar, Ayeyarwady Region (CAS 212475) and Tanintharyi, Dawei (CAS 247949) and were placed in the same clade as these individuals. It is likely that the Myanmar and Vietnam populations represent different species.

Comments. This taxon likely contains multiple cryptic species (Barley et al. 2015).

Specimens examined. USNM 587035–036

Red List status. NE.

***Eutropis multifasciata* (Kuhl, 1820)**
Common Sun Skink
Fig. 4D

Description. Immature male 60.8 mm SVL, 82 mm regenerated TailL; 27.2 mm HeadL; 45% TrunkL/SVL, 41% HindlL/SVL, 23% HeadL/SVL, 65% HeadW/HeadL, 42% HeadH/HeadL, 40% SnEye/HeadL, 30% EyeEar/HeadL. Supralabials 7, 5th largest and beneath eye, 7 infralabials; 47 dorsal scale rows from nape to above vent, dorsal scales tricarinate; 32 scales around midbody; 13 fourth finger lamellae, 19 fourth toe lamellae. In preservative, dark above and dusky white below; scattered small white spots laterally between ear and forelimb.

Natural history notes. Specimens were seen in both primary and secondary forest.

General Distribution. Widespread, Northeast India through southern China to Taiwan southward into Sundanese Indonesia and Philippines.

Molecular Data. Our specimen was placed in its own COI BIN and is 3.65–7.9% sequence divergence from many other specimens in BOLD, ranging from Vietnam to Indonesia. Our 16S sequence is 98–99% similar to specimens in GenBank, including CAS 212916 from Ayeyarwady Region, Myanmar (erroneously reported as "CAS 2120916" in GenBank), and was placed in a 16S clade with other *E. multifasciata*. There are several clades in the 16S tree, with deep divergences (~8%), indicating that this represents a species complex in need of revision.

Specimens examined. USNM 587037.

Red List status. NE.

***Sphenomorphus indicus* (Gray, 1853)**
Indian Forest Skink

Natural history notes. Seen in the slash & burnt site.

General Distribution. Widespread, northern peninsular India and Nepal eastward to southern China southward into Myanmar and Southeast Asia.

Molecular Data. No molecular data available.

Comments. There are no vouchers to confirm this field identification. Because these specimens were observed in the slash & burnt site versus the forest and lack vouchers, we tentatively accept the field identification.

Specimens examined. Field observation by Myint Kyaw Thura; captured for confirmation of identification and then released.

Red List status. NE

***Sphenomorphus maculatus* (Blyth, 1853)**
Asian Spotted Forest Skink

Description. Two adult males 49.9–54.7 mm SVL, 97–98 regenerated mm SVL; 11.5–12.2 mm HeadL; 45–48% TrunkL/SVL, 54–55% HindlL/SVL, 22–23% HeadL/SVL, 62–63% HeadW/HeadL, 44–46% HeadH/HeadL, 39–40% SnEye/HeadL, 33–34% EyeEar/HeadL.

Natural history notes. Seen and captured among forest leaf litter.

General Distribution. Widespread, Nepal to western China southward through Myanmar and Southeast Asia into Peninsular Malaysia.

Molecular Data. Our COI sequences are 7.1–7.5% divergent from specimens in BOLD from Vietnam (not publicly available). Our 16S sequences are 100% identical to a specimen in GenBank (AB028821) from Kaeng Krachan, Thailand (Honda et al. 2000b) and were placed with this individual and other *S. maculatus* in GenBank.

Specimens examined. USNM 587038–039.

Red List status. NE.

Varanidae – monitor lizards

***Varanus rudicollis* (Gray, 1845)**
Rough-necked Monitor

Description. Sex unknown, 57.0 cm SVL, 70.0 cm TailL.

General Distribution. Southern Myanmar and Thailand through Malay Peninsula into Greater Sunda Islands.

Molecular Data. No molecular data available.

Comments. Captured by a villager (Yeybu); photographed; likely subadult or adult.

Specimens examined. USNM Herp Image 2891
Red List status. NE; CITES II.

Snakes
Pythonidae – pythons

***Malayopython reticulatus* (Schneider, 1801) Reticulated Python**
Fig. 5A

Description. Female, maturity uncertain, 119 cm SVL, 18.5 cm TailL.

Natural history notes. The specimen was collected in a fisherman's net, in the stream, near Camp 1.

General Distribution. Widespread, Bangladesh through Southeast Asia to Philippines and the Maluku Islands.

Molecular Data. No molecular data available.

Comments. We note that the genus name *Broghammerus* is not nomenclaturally available for the Reticulated Python, because it was promulgated in a herpetological blog, which did not and does not meet the International Nomenclatural Code's criteria for the valid establishment of formal taxonomic names. A recent study (Reynolds et al. 2014) has proposed the name *Malayopython* for the clade containing the *M. reticulatus* and *M. timoriensis*.

Specimens examined. Specimen remained in FFI field office in Yangon, but it was subsequently destroyed.

Specimens examined. USNM Herp Image 2892
Red List status. NE; CITES II.

Colubridae – colubrid snakes

***Ahaetulla mycterizans* (Linnaeus, 1758)**
Malayan Vinesnake
Fig. 5B

Description. Adult female 745 mm SVL, 385 TailL in life; dorsal scales in 15-15-13 rows, ventrals 190, 148 paired subcaudals with undamaged tip; single precloacal (anal) scale. Snout is blunt, rostral scale truncate anteriorly; 2 loreal scales on right, 1 left, each a small lanceolate scale, isolated in suture between the supralabials and internasal and prefrontal scales; internasal (dorsal surface of snout) flat anteriorly and convex posteriorly. In preservative, the head and nuchal area are bright green dorsally and laterally; trunk gradually become darker green and at about one-third length is olive to tip of tail; laterally trunk pale green to narrow white ventrolateral edge of upturned ventral scales (forming longitudinal stripe), then bordered medially by narrow dark green longitudinal stripe; white stripe becomes yellow by midbody and continues yellow onto base of tail;

Figure 5. Selected snakes found during this study's expedition. **A** *Malayopython reticulatus* (USNM HerpImage 2892) **B** *Ahaetulla mycterizans* (USNM 587040) **C** *Dendrelaphis pictus* (USNM HerpImage 2893) **D** *Boiga dendrophila* (USNM 587041) **E** *Xenochrophis trianguligerus* (USNM 587045) **F** *Rhabdophis chrysargos* (USNM 587044) **G** *Rhabdophis nigrocinctus* (USNM HerpImage 2894) **H** *Naja kaouthia* (USNM HerpImage 2895). Photos **A–B, D–F** by Daniel G. Mulcahy, **C, H–I** by Myint Kyaw Thura.

green stripe disappears 15–20 ventrals anterior to the vent; remainder of venter is white from tip of chin to about midbody then becoming greenish yellow continuing onto tail.

Natural history notes. Collected at 1000 hours near Yeybuchaung-ngal stream, approximately 100 meters downstream from Camp 1.

General Distribution. Tanintharyi southward through Malaya Peninsula to Java and Sumatra.

Molecular Data. There are no other COI sequences in BOLD for *A. mycterizans*. Our specimen was placed in its own COI BIN and the 16S is 99% similar to other *A. mycterizans* in GenBank (KX660161, KX660205; no locality data provided) and was placed in a 16S clade with these specimens and an "*A. prasina*" (FMNH 269042 from Borneo) that is likely misidentified (KX660195).

Comments. This species was an unsuspected find, because it had not been reported previously from Myanmar and the closest Thai records are in southern Thailand south of the Isthmus of Kra. This specimen represents the northernmost record of this species (see Lee et al. 2015). A second specimen (CAS 247859) was confirmed (JLL) further south from Kawthaung, Tanintharyi Region, Myanmar previously identified as *A. prasina*.

Specimens examined. USNM 587040.

Red List status. LC.

***Boiga dendrophila* (Boie, 1827)**
Mangrove Cat Snake
Fig. 5D

Description. Adult female 1450 mm SVL, 362 mm TailL in life; dorsal scales in 21-21-17 rows, ventrals 225, subcaudals 96 with single precloacal scale; 8 supralabials, 3rd, 4th & 5th touch eye; 44 lateral white bars on trunk from neck to vent, venter become entirely black at ventral 126. Adult male 1450 mm SVL, 350 mm TailL in life; dorsal scales in 21-23-17 rows, ventrals 211, subcaudals 91 with single precloacal scale; 8 supralabials, 3rd, 4th & 5th touch eye; 42 lateral white bars on trunk from neck to vent, venter becomes entirely black at ventral 84.

Natural history notes. Collected in primary rainforest along a small tributary to Yeybuchaung-ngal downstream from Camp 1.

General Distribution. Widespread, Tanintharyi to Greater Sunda and Philippine Islands.

Molecular Data. Our COI sequences are 1% different (and placed in the same BIN) from two specimens in BOLD (currently private), from Nakhon Si Thammarat Province, Thailand, south of the Isthmus of Kra. There are currently no other 16S sequences available for *B. dendrophila*.

Comments. This species was also an unsuspected find owing to the absence of previous records for Myanmar; the closest Thai records are from southern Thailand south of the Isthmus of Kra (see Lee et al. 2015). The morphology of these specimens matches that of the subspecies *B. dendrophila melanota* (Boulenger, 1896).

Specimens examined. Adult female, adult male (USNM 587041–042, respectively).

Red List status. NE.

***Boiga drapiezii* (Boie, 1827)**
White-spotted Cat Snake

Description. Adult female 1340 mm SVL, 380 mm TailL in life; dorsal scales in 19-19-15 rows, ventrals 279, subcaudals 144 with single precloacal scale; 8 supralabials, 3rd, 4th & 5th touch eye. 48 dark brown lateral blotches from neck to above vent, venter tan heavily dusted with brownish gray.

Natural history notes. Collected at 1900–2000 hours in primary rainforest near Yeybuchaung-ngal stream, approximately 100 meters upstream from Camp 1 (N 12.43413', E 99.14505') at an elevation of 179 meters ASL The female is gravid, bearing four shelled eggs, each ~34 mm long, 19 mm diameter.

General Distribution. Tanintharyi and Thai-Malay Peninsula through Greater Sundas to Philippines.

Molecular Data. There are currently no other available COI sequences for *B. drapiezii,* it was placed in its own BIN. Our 16S sequence is 99% similar to two *B. drapiezii* in GenBank (KX660209–10; no locality data provided).

Comments. This species was also an unsuspected find owing to the absence of previous records for Myanmar and the closest Thai records in southern Thailand south of the Isthmus of Kra. Two unreported specimens (CAS 247770, CAS 247864) were collected further south in Tanintharyi 40–60 km north of Kawthung (see Lee et al. 2015).

Specimens examined. USNM 587043.

Red List status. LC.

***Dendrelaphis pictus* (Gmelin, 1789)**
Painted Bronzeback
Fig. 5C

Description. Not sexed or measured; field identification.

Natural history notes. From disturbed areas.

General Distribution. Bangladesh and Northeast India through southern Myanmar to northern West Malaysia.

Molecular Data. No molecular data available.

Comments. Field observation by Myint Kyaw Thura; captured for confirmation of identification, photographed and then released.

Specimens examined. USNM Herp Image 2893

Red List status. NE.

Natricidae – water snakes

***Rhabdophis chrysargos* (Schlegel, 1837)**
Speckle-bellied Keelback
Fig. 5F

Description. Juvenile, not sexed, 258 mm SVL, 88 mm TailL; dorsal scales in 17-15-15 rows, ventrals 168, subcaudals 112 with divided precloacal scale; 9 supralabials, 4th, 5th & 6th touch eye. 12–14 anterior maxillary teeth, 2 or 3 short ones, then 2 to 3 slightly enlarge posterior maxillary teeth. White nuchal chevron with arms extending onto supralabials to beneath eye; dorsum and sides of trunk medium olive brown with lateral series of small, faint light spots; venter immaculate white.

Natural history notes. Collected at Camp 1 along the Yeybuchaung-ngal stream.

General Distribution. Southern Southeast Asia through Sundas to Philippines.

Molecular Data. Our COI sequence was placed in its own BIN and is 5.3% divergent from a specimen in BOLD (private) from Vietnam. There are currently no other COI sequences are available at this time. There are currently no other 16S sequences with which to compare.

Comments. *Rhabdophis chrysargos* occurs widely through the southern half of Southeast Asia including the Greater Sunda Islands, the Philippines and Peninsular Malaysia. All records from Myanmar are from the Tanintharyi (Dowling and Jenner 1988).

Specimens examined. USNM 587044.

Red List status. LC.

***Rhabdophis nigrocinctus* (Blyth, 1856)**
Banded Green Keelback
Fig. 5G

Description. Not sexed or measured; field identification.

Natural history notes. Seen in the slash & burnt site.

General Distribution. Southeast Asia, Myanmar and Yunnan, China.

Molecular Data. No molecular data available.

Comments. Field observation by Myint Kyaw Thura; captured for confirmation of identification, photographed, and then released.

Specimens examined. USNM Herp Image 2894

Red List status. LC.

***Xenochrophis piscator* (Schneider, 1799)**
Common Checkered Keelback

Description. Juvenile 225 mm SVL, 101 mm TailL; 131 ventrals, 89 subcaudals, 19-19-17 dorsals; adult male 515 mm SVL, 261 mm TailL; 130 ventrals, 93 subcaudals,

19-19-17 dorsals. In both specimens a single rectangular loreal, single preocular, 3 postoculars, 9 supralabials with the 4th–5th touching eye, 10 infralabials. No data recorded for the individual observed in the slash & burnt area. Dorsum olive-brown, gray-brown with rows of darker rectangular checkered blotches, "V" nuchal mark on the head. Venter plain, lighter, dark fringes on ventrals absent, but some speckling along the margins.

Natural history notes. Occurs in ponds and streams in human-impacted areas.

General Distribution. Widespread, Pakistan to Southeast Asia and southern China.

Molecular Data. There are currently no other COI sequence for *X. piscator* available for comparison. The closest COI sequences available are from *X. flavipunctatus* and they are more than 10% divergent. Our 16S sequences are 4% different from a *X. piscator* in GenBank (KX277271; no locality data provided).

Comments. Vogel & David (2012) did not record this species from the Tanintharyi. However, recent records deposited from CAS exist from Dawei. These two specimens extend the distribution of this species ~175 km due south. *Xenochrophis flavipunctatus* should also occur here, but as yet its presence has not been confirmed. *Xenochrophis piscator* may represent a species complex, as significant morphological variation occurs in different populations, especially those in India and Sri Lanka (Vogel and David 2012).

Specimens examined. USNM 587046 (adult)–587047 (juvenile).

Red List status. NE.

***Xenochrophis trianguligerus* (Boie, 1827)**
Red-sided Keelback (Fig. 5E)

Description. Not sexed, juvenile specimen 245 mm SVL, 95 TailL; 225 ventrals, 96 subcaudals, 19-19-15 dorsals; single loreal rectangular, preocular single, 3 postoculars. Dorsum olive brown, dark rectangular blotches laterally with the anterior portion of the sides yellow and red, fading in coloration towards the midbody. Head hued with blue, dark sutures on some of the supralabials.

Natural history notes. Discovered streamside in the primary forest.

General Distribution. Northeast India to Southeast Asia through Malay Peninsula into Greater Sunda Islands.

Molecular Data. There are currently no other COI nor 16S sequences available for *X. trianguligerus* with which to compare.

Comments. Earlier country-wide herpetofaunal surveys of Myanmar indicate that *X. trianguligerus* occurs only in Tanintharyi.

Specimens examined. USNM 587045.

Red List status. LC.

Elapidae – cobras and kraits

***Naja kaouthia* (Lesson, 1831)**
Monocled Cobra (Fig. 5H)

Description. Juvenile, not sexed or measured; field identification.

Natural history notes. Found in human-disturbed habitats. Field observation at Forest 3 site by Myint Kyaw Thura; photographed for confirmation of identification and then released.

General Distribution. Widespread, Nepal and northern peninsular India to Southeast Asia and southern China.

Molecular Data. No molecular data available.

Specimens examined. USNM Herp Image 2895

Red List status. LC; CITES II.

Discussion

DNA Barcode data

The DNA barcode data greatly improved our estimate of species numbers in the Tanintharyi, within anurans in particular, similar to another recent study using DNA barcode data for estimating species diversity (Diechmann et al. 2017). The placement of sequences from individual specimens of the same genus into multiple COI BINs indicated it was likely that multiple species were collected. For instance, collections at the edge of the forest and on the Tanintharyi River floodplain, at the eastern edge of Yebu Village, yielded multiple anuran specimens identified initially as four morphospecies in four genera (*Polypedates*, *Microhyla*, *Fejervarya*, and *Occidozyga*). However, DNA barcoding and comparisons with our northern Myanmar reference material revealed each genus was likely represented by two to three species. We determined this based on the fact that at least one of the Tanintharyi clades (within each genus) grouped with specimens from the north, rather than with the other clades (of the respective genus) in the Tanintharyi. This increased our total number of species from four to ten. We note that the short sequence data from DNA barcode data, while useful for determining relationships among closely related groups, such as these cases of populations within genera, falls short at resolving higher-level relationships. For instance, even our combined dataset (COI + 16S) fails to recover several families as monophyletic, such as Microhylidae and Dicroglossidae, and even fails to recover some genera as monophyletic, such as *Xenochrophis* (Fig. 2). Whereas analyses with more complete taxonomic sampling, more loci, and more robust analyses recover these families as monophyletic (e.g. de Sá et al. 2012).

The DNA barcode data allowed us to identify several cryptic species of anurans in multiple genera and families, including Rhacophoridae: *Polypedates* – two species

in the Tanintharyi, and a third in the north; Microhylidae: *Microhyla* – three species in the Tanintharyi, all of those and a three more species in the north; Dicroglossidae: *Fejervarya* – two species in the Tanintharyi, and two in the north; and *Occidozyga* – three species in the Tanintharyi, one of the same and three additional species in the north. Likewise, we were able to determine the number of species present at the forest sites based on the COI barcode data, in the family Dicroglossidae: *Limnonectes* – three species in the Tanintharyi, two of the same and a fourth in the north; the family Microhylidae: *Kaloula* – two species in the Tanintharyi, one also occurring in the north; Ranidae: *Hylarana* – one species in the Tanintharyi and the north, and three additional species in the north; *Odorrana* – two species in the Tanintharyi, one of these and a third in the north. Because a comprehensive COI barcode library is lacking for southeast Asian anurans, we relied on 16S sequence data as a supplemental barcode marker to help identify specimens to species based on comparisons with known material published in GenBank. The inclusion of the 16S data allowed us to compare our specimens with published material in GenBank, and enabled us to identify several specimens to named species (e.g. *Limnonectes blythii*, *L. doriae*, *L. limborgi*, and *L. longchuanensis*, *Microhyla berdmorei*, *M. butleri*, *M. fissipes*, and *M. heymonsi*), including several recently described species (*Hoplobatrachus litoralis*, *Kaloula latidisca*), and some recently identified, but not formally described, species such as *Fejervarya* sp. 'hp2–3' (Kotaki et al. 2010), *Fejervarya* sp. BFL 2007 (Islam et al. 2008, Hasan et al. 2012a), and *Sylvirana* sp. C 'MS-2010' (Hasan et al. 2012a), as well as some of our own new discoveries (*Fejervarya* sp. A, *Occidozyga* spp. A–D, *Leptolalax* sp. A, *Microhyla* spp. A–B, and *Chiromantis* sp. A). The use of the DNA barcode database (BOLD) allows us to "BIN" these un-named species, such that researchers conducting future expeditions can compare their specimens to ours to determine if they are the same un-named species, "known unknowns", or yet newly discovered un-named species "unknown unknowns" (Collins and Cruickshank 2014). Traditional, morphological species descriptions need to follow in order to properly assess the biodiversity of this region. However, this process can be slow, requires taxonomic expertise, and is less supported by many academic institutions. Given the high rate of putative cryptic species, particularly among anuran genera, we issue caution when using morphological identifications alone, as in the Reserve Forest areas presented in this study. The fact remains that guide books of the region (SE Asia) that attempt to include morphological identifications contain overlapping character descriptions (e.g. Grismer 2011), or lack identifications altogether (e.g. Koch 2012). Once more robust taxonomic treatments of each group are conducted, more reliable morphological diagnoses may become available.

A recent study suggests that species delimitations based solely on mtDNA may be misleading, and over-estimating species in biodiversity studies (Chan et al. 2017). We maintain that our method, preliminary as it is because it acts as a triage assessment, is still a valid method in rapid biodiversity surveys for a number of reasons. First of all, it may be better to over-estimate species diversity in rapid biodiversity surveys, rather than under-estimating in order to secure proper protection of the area. Secondly, this method may be valid because estimates made using this method may be correct and documenting the accurate number of species. Thirdly, even if species numbers are

over-estimated, recognizing several lineages (or operational taxonomic units, OTUs) at a minimum recognizes the genetic diversity in a lineage (if they are later determined to represent the same species), and is thus beneficial from a conservation-genetics point of view. And fourthly, rapid assessments identifying potential cryptic species can direct future research to taxonomic groups in need of further investigation. It is largely for this reason that we recommend the use of place-holder names, such as "sp. A" until more in-depth investigations can be conducted, including additional taxonomic and geographic sampling, and additional markers (e.g. nuclear). In reality, the forests may disappear before such in-depth analyses can be conducted, as each group (e.g. genus) may require essentially a dissertation chapter's-worth of work (e.g. Chan et al. 2017), and for the number of groups covered in this report, for example, may take over a decade to complete, by which time the forests could be gone (Connette et al. 2017).

Important absences and presences in the Tenasserim

Our brief surveys are inadequate to address the presence or absence of all potential members of the southern Tanintharyi herpetofauna. Studies of the herpetofauna of the Myanmar Central Dry Zone at the Chatthin Wildlife Sanctuary (Zug et al. 1998, Zug 2011) had a team of four regular members with the assistance of the entire Sanctuary staff. Even with all these eyes and hands and with weekly transect surveys and monthly drift-fence trapping, 40–41 weeks were required to record 90% of the Chatthin herpetofauna, and previously undocumented species were still being discovered in the third and final year of that survey. We especially note that nine species of turtles have been recorded on the Thailand side of the Tanintharyi mountain range. Crocodilians were not expected owing to shallowness of the streams in the immediate area of the survey.

At this stage of our inventory of the Tanintharyi proposed National Park and its environs, we wish to emphasize the discovery of taxa previously undocumented for Myanmar. Surprisingly these undocumented taxa include only two amphibians (*Ichthyophis* cf. *kohtaoensis* and *Chalcorana eschatia*) and three species of snakes (*Ahaetulla mycterizans*, *Boiga dendrophila*, and *Boiga drapiezii*; see Lee et al. 2015). None of these taxa were reported from Phetchaburi Province, the Thailand province immediately east of our survey site. They all represent species whose primary distribution is south of the Isthmus of Kra; the two *Boiga* were reported in Pauwels et al. (2000) survey of Phang-Nga Province, Thailand.

Some recently described taxa were also detected in Myanmar based on our study, including the likely occurrence of *Kaloula latidisca*, a species recently described from Peninsular Malaysia (Chan et al. 2014). As we included the reference material from more northern Myanmar, we identified species that have not been previously detected in Myanmar. This includes the dicroglossid *Hoplobatrachus litoralis,* a recently described species from Cox's Bazar district of Bangladesh (Hasan et al. 2012b), which now includes two specimens from the Yangon area and one from Mon State, extending the range of this species from Bangladesh to south-central Myanmar. In total, the formally threatened species according to the IUCN Red List encountered include two anurans

(*Ansonia thinthinae*, *Limnonectes blythii)*, and the following represent CITES II species: two turtles (*Indotestudo elongata*), a lizard (*Varanus rudiocolis*), and two snakes (*Naja kaouthia* and *Malayopython reticulatus*).

Conclusion

The use of DNA barcoding allowed us to determine how many species were present at the site of our biodiversity inventory survey. The inclusion of the supplementary marker 16S allowed us to assign several individuals to named species for which 16S data were available for comparison (whereas the COI reference library is less complete), and to identify others as previously identified, undescribed species. The use of the DNA barcode database (BOLD) allows us to "BIN" these un-named species, such that researchers conducting future expeditions can compare their specimens to ours to determine if they are the same un-named species. Biodiversity research needs more "boots on the ground," because an incomplete taxonomy hinders our ability to protect biodiversity and guide conservation (Wilson 2017). As we proceed to fill the "BINs of life," we will eventually be able to record and catalogue all species of life on Earth. We encourage researchers to continue to add to these databases, and most importantly, to update existing records as our knowledge increases.

Acknowledgements

The herpetofaunal surveys were proposed and directed by Flora & Fauna International–Myanmar Programme (FFI). In country funding was provided by this organization; travel to and from Myanmar was supported by the National Museum of Natural History, Smithsonian Institution (NMNH-SI) Grand Challenges Consortia Award (to Melissa Songer). Laboratory expensed were provided by NMNH-SI Global Genome Initiative (GGI). All of the laboratory and/or computer work were conducted in and with the support of the L.A.B. facilities of the NMNH. We thank the NMNH-SI Internship program for hosting authors JLL and AHM to collect measurements of many of the amphibian and reptile voucher specimens, and the NMNH-SI Youth Engagement through Science (YES!) Program for hosting M. Chand for DNA barcoding the specimens. We appreciate and thank the director and staff of the Nature and Wildlife Conservation Division, Forestry and Environmental Conservation Ministry, for permission to conduct the survey and to allow the export specimens to permit accurate identification. Thaw Zin provided exceptional assistance in the collection and preparation of voucher specimens. The cook, U Win, allowed us to share his home and access to the garden habitats thereabout at camp 3. We thank Zin Win Htun (Nature and Wildlife Conservation Division, The Republic of the Union of Myanmar) for his escort service into the proposed park area, and Aung Lin (FFI), for his guidance and logistical support. We also thank Frank Momberg and Mark Grindley (FFI), for organizing the survey, arranging for the various government

authorizations that allowed the survey team to concentrate on the survey and to allow study of specimens at the NMNH-SI for accurate identification of the vouchers. We thank Grant M. Connette (SCBI) for providing the map for Figure 1. We thank L.L. Grismer, T. Nguyen, and O.S.G. Pauwels for constructive comments and corrections on the manuscript. JLL thanks D. Lee and P. Lee for their constant support. GRZ's research into the biology and evolution of Asian herpetofauna has received support from numerous sources: The Smithsonian Scholarly Studies Program, and NMNH-Research Opportunities Fund, Biodiversity. Additionally, GRZ's home institution, the Smithsonian Institution and the NMNH, have provided him with decades of research support through its many bureaus, offices, branches, and especially from its collections management staff. GRZ thanks P. Zug for her continuing support in data collection and entry.

References

Aowphol A, Rujirawan A, Taksintum W, Arsirapot S, McLeod DS (2013) Re-evaluating the taxonomic status of *Chiromantis* in Thailand using multiple lines of evidence (Amphibia: Anura: Rhacophoridae). Zootaxa 3702: 101–123. http://dx.doi.org/10.11646/zootaxa.3702.2.1

Barley AJ, Datta-Roy A, Karanth KP, Brown RM (2015) Sun skink diversification across the Indian–Southeast Asian biogeographical interface. Journal of Biogeography 42: 292–304. https://doi.org/10.1111/jbi.12397

Biju SD, Garg S, Mahony S, Wijayathilaka N, Senevirathne G, Meegaskumbura M (2014) DNA barcoding, phylogeny and systematics of Golden-backed frogs (*Hylarana*, Ranidae) of the Western Ghats-Sri Lanka biodiversity hotspot, with the description of seven new species. Contributions to Zoology 83: 269–335.

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. https://doi.org/10.1111/evo.12145

Blyth E (1853) Notices and descriptions of various reptiles, new or little-known. Part I. J. Asiatic Society of Bengal 22: 639–655.

Boie F (1827) Bemerkungen über Merrem's Versuch eines Systems der Amphibien, 1. Lieferung: Ophidier. Isis van Oken 20: 508–566.

Böhm M, Collen B, Baillie JEM, Bowles P, Chanson J et al. (2013) The conservation status of the world's reptiles. Biological Conservation 157: 372–385. https://doi.org/10.1016/j.biocon.2012.07.015

Boulenger GA (1884) Descriptions of new species of reptiles and batrachians in the British Museum. Part. II. Annals and Magazine of Natural History, Series 5, 13: 396–398. https://doi.org/10.1080/00222938409459259

Boulenger GA (1885) Catalogue of the Lizards in the British Museum I. Geckonidae, Eublepharidae, Uroplatidae, Pygopodidae, Agamidae. London, 450 pp.

Boulenger GA (1887) An account of the reptiles and batrachians obtained in Tenasserim by M. L. Fea of the Genoa Civic Museum. Annali del Museo Civico di Storia Naturale di Genova. Serie 2, 5: 474–486.

Boulenger GA (1891) On new or little-known Indian and Malayan reptiles and batrachians. Annals and Magazine of Natural History, Series 6, 8: 288–292. https://doi.org/10.1080/00222939109460437

Boulenger GA (1900) Descriptions of new batrachians and reptiles from the Larut Hills, Perak. Annals and Magazine of Natural History, Series 7, 6: 186–193. https://doi.org/10.1080/00222930008678356

Boulenger GA (1920) A monograph of the South Asian, Papuan, Melanesian and Australian frogs of the genus *Rana*. Records of the Indian Museum 20: 1–226. https://doi.org/10.5962/bhl.title.12471

Blyth E (1856) Report for October Meeting, 1855. Journal of the Asiatic Society of Bengal 24: 711–723.

Caley MJ, Fisher R, Mengersen K (2014) Global species richness estimates have not converged. Trends in Ecology & Evolution 29: 187–188. https://doi.org/10.1016/j.tree.2014.02.002

Castañeda MdR, de Queiroz K (2011) Phylogenetic relationships of the *Dactyloa* clade of *Anolis* lizards based on nuclear and mitochondrial DNA sequence data. Molecular phylogenetics and evolution. 61: 784–800. https://doi.org/10.1016/j.ympev.2011.07.004

Chan KO, Blackburn DC, Murphy RW, Stuart BL, Emmett DA, Ho CT, Brown RM (2013) A new species of narrow-mouthed frog of the genus *Kaloula* from eastern Indochina. Herpetologica 69: 329–341. https://doi.org/10.1655/HERPETOLOGICA-D-12-00094

Chan KO, Grismer LL, Brown RM (2014) Reappraisal of the Javanese Bullfrog complex, *Kaloula baleata* (Müller, 1836) (Amphibia: Anura: Microhylidae), reveals a new species from Peninsular Malaysia. Zootaxa 3900: 569–580. http://dx.doi.org/10.11646/zootaxa.3900.4.7

Chan KO, Alexander AM, Grismer LL, Su Y-C, Grismer JL, Quah ESH, Brown RM (2017) Species delimitation with gene flow: A methodological comparison and population genomics approach to elucidate cryptic species boundaries in Malaysian Torrent Frogs. Molecular Ecology 2017: 5435–5450. https://doi.org/10.1111/mec.14296

Chan-ard T, Makchai S (2011) A new insular species of *Cyrtodactylus* Gray, 1827 (Squamata, Gekkonidae), from the Surin Islands, Phang-nga Province, southern Thailand. The Thailand Natural History Museum Journal 5: 7–15.

Collins RA, Cruickshank RH (2013) The seven deadly sins of DNA barcoding. Molecular Ecology Resources 13: 969–975. https://doi.org/10.1016/j.tree.2014.02.002

Collins RA, Cruickshank RH (2014) Known knowns, known unknowns, unknown unknowns, and unknown knowns in DNA barcoding: A comment on Dowton et al. Systematic Biology 63: 1005–1009. https://doi.org/10.1093/sysbio/syu060

Connette GM, Oswald P, Thura MK, Connette KJL, Grindley ME, Songer M, Zug GR, Mulcahy DG (2017) Rapid forest clearing in a Myanmar proposed national park threatens two newly discovered species of geckos (Gekkonidae: *Cyrtodactylus*). PLOS ONE 12: e0174432. https://doi.org/10.1371/journal.pone.0174432

Costello MJ, May RM, Stork NE (2013) Can we name Earth's species before they go extinct? Science 339: 413–416. https://doi.org/10.1126/science.1230318

Daudin FM (1802) Histoire Naturelle, générale et particulièredes reptiles, ouvrage faisant suite, à l'histoire naturelle, générale et particulière composée par Le clerc de Buffon, et rédigée par Sonnini CS, vol. 3. F. Dufart, Paris.

de Sá RO, Streicher JW, Sekonyela R, Forlani MC, Loader SP, Greenbaum E, Richards S, Haddad CFB (2012) Molecular phylogeny of microhylid frogs (Anura: Microhylidae) with

emphasis on relationships among New World genera. BMC Evolutionary Biology 12: 241. http://www.biomedcentral.com/1471-2148/12/241

Dever JA, Fuiten AM, Konu Ö, Wilkinson JA (2012) Cryptic torrent frogs of Myanmar: An examination of the *Amolops marmoratus* species complex with the resurrection of *Amolops afghanus* and the identification of a new species. Copeia 2012: 57–76. https://dx.doi.org/10.1643/CH-10-180

Diechmann J, Mulcahy DG, Vanthomme H, Tobi E, Wynn AH, Zimkus BM, McDiarmid RW (2017) How many species and under what names? Using DNA barcoding and GenBank data for West Central African amphibian conservation. PLOS ONE 12: e0187283. https://doi.org/10.1371/journal. pone.0187283

Dowling HG, Jenner JV (1988) Snakes of Burma: Checklist of reported species and bibliography. Smithsonian Herpetological Information Service 76: 1–19. https://doi.org/10.5479/si.23317515.76.1

Dubois A (1992) Notes sur la classification des Ranidae (Amphibiens anoures). Bulletin Mensuel de la Société Linnéenne de Lyon 61: 305–352. https://doi.org/10.3406/linly.1992.11011

Duméril AMC, Bibron G (1836) Erpetologie Générale ou Histoire Naturelle Complète des Reptiles. Vol. 3. Librairie Encyclopédique Roret, Paris, 528 pp.

Frost DR (2017) Amphibian species of the world: An online reference. Version 6.0 (Oct. 2017). American Museum of Natural History, New York, USA. http://research.amnh.org/herpetology/amphibia/index.html

Garnett ST, Christidis L (2017) Taxonomy anarchy hampers conservation. Nature 546: 25–27. https://www.nature.com/news/taxonomy-anarchy-hampers-conservation-1.22064 https://doi.org/10.1038/546025a

Geoffroy SH (1809) Mémoire sur les tortues molles. Nouveau Bulletin des Sciences, Société Philomathique de Paris 22: 363–367.

Giller G (2014) Are we any closer to knowing how many species there are on Earth? Scientific American April 8: 2014. https://www.scientificamerican.com/article/are-we-any-closer-to-knowing-how-many-species-there-are-on-earth/

Gmelin JF (1789) Caroli a Linné Systema naturae. Tomus 1 Pars III. Leipzig, 1033–1516.

Gravenhorst JLC (1829) Deliciae Musei Zoologici Vratislaviensis. Fasciculus primus. Chelonios et Batrachia. Leopold Voss, Leipzig.

Gray JE (1831) Description of two new genera of frogs discovered by John Reeves, Esq. in China. Zoological Miscellany, Part 1: 38.

Gray JE (1845) Catalogue of the specimens of lizards in the collection of the British Museum. Trustees of the British Museum/Edward Newman, London, 289 pp.

Gray JE (1853) Descriptions of some undescribed species of reptiles collected by Dr. Joseph Hooker in the Khassia Mountains, East Bengal, and Sikkim Himalaya. Annals and Magazine of Natural History 12: 386–392. https://doi.org/10.1080/03745485709495063

Grismer LL (2011) Amphibians and Reptiles of the Seribuat Archipelago (Peninsular Malaysia): A field guide. Edition Chimaira, Frankfurt am Main, 239 pp.

Grismer LL, Sumontha M, Cota M, Grismer JL, Wood Jr. PL, Pauwels OSG, Kunya K (2010) A revision and redescription of the rock gecko *Cnemaspis siamensis* (Taylor 1925) (Squamata: Gekkonidae) from Peninsular Thailand with descriptions of seven new species. Zootaxa 2576: 1–55. www.mapress.com/zootaxa/2010/f/z02576p055f.pdf

Grismer LL, Wood Jr. PL, Quah ESH, Anuar S, Muin MA, Sumontha M, Ahmad N, Bauer AM, Wangkulangkul S, Grismer JL, Pauwels OSG (2012) A phylogeny and taxonomy of the Thai-Malay Peninsula Bent-toed geckos of the *Cyrtodactylus pulchellus* complex (Squamata: Gekkonidae): combined morphological and molecular analyses with descriptions of seven new species. Zootaxa 3520: 1–55. http://www.mapress.com/j/zt/article/view/14344

Grismer LL, Wood Jr. PL, Aowphol A, Cota M, Murdoch ML, Aguilar C, Grismer MS (2016) Taxonomy, phylogeny, and distribution of *Bronchocela rayaensis* (Squamata: Agamidae) on the Thai-Malay Peninsula. Zootaxa 4092: 414–420. http://www.mapress.com/j/zt/article/view/zootaxa.4092.3.6

Grismer LL, Wood Jr. PL, Aowphol A, Cota M, Grismer MS, Murdoch ML, Aguilar C, Grismer JL (2017) Out of Borneo again and again: biogeography of the Stream Toad genus *Ansonia* Stoliczka (Anura: Bufonidae) and the discovery of the first limestone cave dwelling species. Biological Journal of the Linnean Society 120: 371–395. https://doi.org/10.1111/bij.12886

Grosjean S, Bordoloi S, Chuaynkern Y, Chakravarty P, Ohler A (2015a) When young are more conspicuous than adults: a new ranid species (Anura: Ranidae) revealed by its tadpole. Zootaxa 4058: 471–498. http://dx.doi.org/10.11646/zootaxa.4058.4.2

Grosjean S, Ohler A, Chuaynkern Y, Cruaud C, Hassanin A (2015b) Improving biodiversity assessment of anuran amphibians using DNA barcoding of tadpoles. Case studies from Southeast Asia. Comptes Rendus Biologies 338: 351–361. http://dx.doi.org/10.1016/j.crvi.2015.03.015

Hasan M, Islam MM, Khan MMR, Alam MS, Kurabayashi A, Igawa T, Kuaramoto M, Sumida M (2012a) Cryptic anuran biodiversity in Bangladesh revealed by mitochondrial 16S rRNA gene sequences. Zoological Science 29: 162–172. http://dx.doi.org/10.2108/zsj.29.162

Hasan M, Mitsuru M, Islam MM, Alam MS, Khan MMR, Sumida M (2012b) A new species of *Hoplobatrachus* (Anura, Dicroglossidae) from the coastal belt of Bangladesh. Zootaxa 3312: 45–48. http://www.mapress.com/zootaxa/2012/f/z03312p058f.pdf

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003a) Biological identifications through DNA barcodes. Proceedings of the Royal Society B Biological Sciences 270: 313–21. https://doi.org/10.1098/rspb.2002.2218

Hebert PDN, Ratnasingham S, DeWaard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society B Biological Sciences (Suppl.) 270: S96–S99. https://doi.org/10.1098/rsbl.2003.0025

Honda M, Ota H, Kobayashi M, Nabhitabhata J, Yong HS, Hikida T (1999) Phylogenetic relationships of the flying lizards, genus *Draco* (Reptilia, Agamidae). Zoological Science 16: 535–549. http://dx.doi.org/10.2108/zsj.16.535

Honda M, Ota H, Kobayashi M, Nabhitabhata J, Yong HS, Sengoku S, Hikida T (2000a) Phylogenetic relationships of the family Agamidae (Reptilia: Iguania) inferred from mitochondrial DNA sequences. Zoological Science 17: 527–537. http://dx.doi.org/10.2108/0289-0003(2000)17[527:PROTFA]2.0.CO;2

Honda M, Ota H, Kobayashi M, Nabhitabhata J, Yong HS, Hikida T (2000b) Phylogenetic relationships, character evolution, and biogeography of the subfamily Lygosominae (Reptilia: Scincidae) inferred from mitochondrial DNA sequences. Molecular Phylogenetics and Evolution 15: 452–461. https://doi.org/10.1006/mpev.1999.0767

Howlader MSA, Nair A, Merilä J (2016) A New Species of Frog (Anura: Dicroglossidae) Discovered from the Mega City of Dhaka. PLOS ONE 11: e0149597. https://doi.org/10.1371/journal.pone.0149597

Inger RF, Stuart BL, Iskandar DT (2009) Systematics of a widespread southeast Asian frog, *Rana chalconota* (Amphibia: Anura: Ranidae). Zoological Journal of the Linnean Society 155: 123–147. Https://doi.org/10.1111/j.1096-3642.2008.00440.x

Islam MM, Kurose N, Khan MMR, Nishizawa T, Kuramoto M, Alam MS, Hasan M, Kurniawan N, Nishioka M, Sumida M (2008) Genetic divergence and reproductive isolation in the genus *Fejervarya* (Amphibia: Anura) from Bangladesh inferred from morphological observation, crossing experiments, and molecular analyses. Zoological Science 25: 1084–1105. http://dx.doi.org/10.2108/zsj.25.1084

Kathriner A, O'Shea M, Kaiser H (2014) Re-examination of *Hemidactylus tenkatei* van Lidth de Jeude, 1895: populations from Timor provide insight into the taxonomy of the *H. brookii* Gray, 1845 complex (Squamata: Gekkonidae). Zootaxa 3887: 583–599. http://dx.doi.org/10.11646/Zootaxa.3887.5.5

Koch A (2012) Discovery, Diversity, and Distribution of the Amphibians and Reptiles of Sulawesi and its Offshore Islands. Edition Chimaira, Frankfurt am Main, 374 pp.

Kotaki M, Kurabayashi A, Matsui M, Kuramoto M, Djong TH, Sumida M (2010) Molecular phylogeny of the diversified frogs of genus *Fejervarya* (Anura: Dicroglossidae). Zoological Science 27: 386–395. http://dx.doi.org/10.2108/zsj.27.386.s1

Kuhl H (1820) Beiträge zur Zoologie und vergleichenden Anatomie. Hermannsche Buchhandlung, Frankfurt, 152 pp.

Kuraishi N, Matsui M, Hamidy A, Belabut DM, Ahmad N, Panha S, Sudin A, Yong HS, Jiang J-P, Ota H, Thong HT, Nishikawa K (2012) Phylogenetic and taxonomic relationships of the *Polypedates leucomystax* complex (Amphibia). Zoologica Scripta 42: 54–70. https://doi.org/10.1111/j.1463-6409.2012.00562.x

Lee JL, Myint Kyaw Thura, Mulcahy DG, Zug GR (2015) Three colubrid snakes new to Myanmar. Herpetology Notes 8: 217–220. https://www.biotaxa.org/hn/article/view/7805

Lee JL, Miller AH, Connette GM, Oo K Swe, Zug GR, Mulcahy DG (2018) First record of the Malaysian Bridle Snake, *Dryocalamus subannulatus* (Duméril, Bibron & Duméril, 1854), in Myanmar (Reptilia, Serpentes, Colubridae). Check List 14: 341–345. https://doi.org/10.15560/14.2.341

Lesson RP (1831) Catalogue des Reptiles qui font partie d'une Collection zoologique recueille dans l'Inde continentale ou en Afrique, et apportée en France par M. Lamare-Piquot. Bulletin des Sciences Naturelles et de Géologie, Paris 25: 119–123.

Lidth de Jeude TW van (1895) Reptiles from Timor and the neighbouring islands. Notes from the Leyden Museum 16: 119–127.

Linnaeus C (1758) Systema naturæ per regna tria naturæ, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Tomus I. Editio decima, reformata. Laurentii Salvii, Holmiæ, 824 pp.

Lipscomb D, Platnick N, Wheeler Q (2003) The intellectual content of taxonomy: a comment on DNA taxonomy. Trends in Ecology and Evolution 18: 65–66. http://dx.doi.org/10.1016/S0169-5347(02)00060-5

Macey JR, Larson A, Ananjeva NB, Fang Z, Papenfuss TJ (1997) Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Molecular Biology and Evolution 14: 91–104. https://doi.org/10.1093/oxfordjournals.molbev.a025706

Mallet J, Willmott K (2003) Taxonomy: renaissance or Tower of Babel? Trends in Ecology and Evolution 18: 57–59. https://doi.org/10.1016/S0169-5347(02)00061-7

Matsui M, Khonsue W, Nabhitabhata J (2005a) A new *Ansonia* from the Isthmus of Kra, Thailand (Amphibia, Anura, Bufonidae). Zoological Science 22: 809–814. https://doi.org/10.2108/zsj.22.809

Matsui M, Shimada T, Ota H, Tanaka-Ueno T (2005b) Multiple invasions of the Ryukyu Archipelago by Oriental frogs of the subgenus *Odorrana* with phylogenetic reassessment of the related subgenera of the genus *Rana*. Molecular Phylogenetics and Evolution 37: 733–742. https://doi.org/10.1016/j.ympev.2005.04.030

Matsui M, Nabhitabhata J (2006) A new species of *Amolops* from Thailand (Amphibia, Anura, Ranidae). Zoological Science Tokyo 23: 727–732. https://doi.org/10.2108/zsj.23.727

Matsui M, Shimada T, Liu WZ, Maryati M, Khonsue W, Orlov N (2006) Phylogenetic relationships of Oriental torrent frogs in the genus *Amolops* and its allies (Amphibia, Anura, Ranidae). Molecular Phylogenetics and Evolution 38: 659–666. https://doi.org/10.1016/j.ympev.2005.11.019

Matsui M, Hamidy A, Belabut DM, Ahmad N, Panha S, Sudin A, Khonsue W, Oh H-S, Jiang J-P, Nishikawa K (2011) Systematic relationships of Oriental tiny frogs of the family Microhylidae (Amphibia, Anura) as revealed by mtDNA genealogy. Molecular Phylogenetics and Evolution 61: 167–176. https://doi.org/10.1016/j.ympev.2011.05.015

McMahan C, Zug GR (2007) Burmese *Hemidactylus* (Reptilia, Squamata, Gekkonidae): geographic variation in the morphology of *Hemidactylus bowringii* in Myanmar and Yunnan, China. Proceedings of the California Academy of Sciences 58: 485–509.

Mulcahy DG, Macdonald III KS, Brady SG, Meyer C, Barker K, Coddington JC (2016) Greater than X kb: A quantitative assessment of preservation conditions on genomic DNA quality, and a proposed standard for genome-quality DNA. PeerJ 4: e2528. https://doi.org/10.7717/peerj.2528

Mulcahy DG, Lee JL, Miller AH, Zug GR (2017) Troublesome Trimes: Potential cryptic speciation of the *Trimeresurus* (*Popeia*) *popeiorum* complex (Serpentes: Crotalidae) around the Isthmus of Kra (Myanmar and Thailand). Zootaxa 4347: 301–315. https://doi.org/10.11646/zootaxa.4347.2.6

Nabhitabhata J, Chan-ard T, Chuaynkern Y (2004) Checklist of Amphibians and Reptiles in Thailand. Office of Environmental Policy and Planning, Bangkok, Thailand. [Print date 2000]

Nishikawa K, Matsui M, Yong HS, Ahmad N, Yambun P, Belabut D, Sudin A, Hamidy A, Nikolai O, Ota H, Yoshikawa N, Tominaga A, Shimada T (2012) Molecular phylogeny and biogeography of Asia (Amphibia, Gymnophiona, Ichthyophiidae), with special reference to high cryptic species diversity in Sundaland. Molecular Phylogenetics and Evolution 63: 714–723. https://doi.10.1016/j.ympev.2012.02.017

Ohler A, Wollenberg KC, Grosjean S, Hendrix R, Vences M, Ziegler T, Dubois A (2011) Sorting out Laos: description of new species and additional taxonomic data on megophryid

frogs from northern Indochina (genus *Leptolalax*, Megophryidae, Anura). Zootaxa 3147: 1–83. http://www.mapress.com/zootaxa/2011/2/zt03147p083.pdf

Oliver LA, Prendini E, Kraus F, Raxworthy CJ (2015) Systematics and biogeography of the *Hylarana* frog (Anura: Ranidae) radiation across tropical Australasia, Southeast Asia, and Africa. Molecular Phylogenetics and Evolution 90: 176–192. http://dx.doi.org/10.1016/j.ympev.2015.05.001

Palumbi S, Martin A, Romano S, McMillan WO, Stice L, Grabowski G (1991) The Simple Fool's Guide to PCR, Version 2.0. Privately published, USA.

Pan SL, Dang NX, Wang JS, Zheng YT, Rao DQ, Li JT (2013) Molecular Phylogeny Supports the Validity of *Polypedates impresus* Yang 2008. Asian Herpetological Research 4: 124–133. www.ahr-journal.com/OA/pdfdow.aspx?Sid=20130006 https://doi.org/10.3724/SP.J.1245.2013.00124

Pauwels OSG, Chan-ard T (2006) Reptiles of Kaeng Krachan National Park, Western Thailand. Natural History Bulletin of the Siam Society 54: 89–108.

Pauwels OSG, Laohawat O-A, David P, Bour R, Dangsee P, Puangjit C, Chimsunchart C (2000) Herpetological investigations in Peninsular Thailand, with a list of reptile species and notes on their biology. Dumerilia 4: 123–154.

Pauwels OSG, Laohawat O-A, Naaktae W, Puangjit C, Wisutharom T, Chimsunchart C, David P (2002) Reptile and amphibian diversity in Phang-nga Province, southern Thailand. The Natural History Journal of Chulalongkorn University 2: 25–30.

Pauwels OSG, Chan-ard T, Wanchai P, Kheowyoo B, Bauer AM (2009) Reptile diversity of Phetchaburi province, western Thailand: an update. Hamadryad 34: 9–21.

Peters WCH (1867) Herpetologische Notizen. Monatsberichte der Königlichen Preussische Akademie des Wissenschaften zu Berlin 1867: 13–37.

Puillandres N, Strong EE, Bouchet P, Boisselier MC, Couloux A, Samadi S (2009) Identifying gastropod spawn from DNA barcodes: possible but not yet practicable. Molecular Ecology Resources Sep 9: 1311–21 https://doi.org/10.1111/j.1755-0998.2009.02576.x

Raposa MA, Stopiglia R, Brito GRR, Bockmann FA, Kirwan GM, Gayon J, Dubois A (2017) What really hampers taxonomy and conservation? A riposte to Garnett and Christidis (2017). Zootaxa 4317: 179–184. https://doi.org/10.11646/zootaxa.4317.1.10

Ratnasingham S, Hebert PDN (2013) A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. PLOS ONE 8: e66213. https://doi.org/10.1371/journal.pone.0066213

Reynolds RG, Miemiller ML, Revell LJ (2014) Toward a Tree-of-Life for the boas and pythons: multilocus species-level phylogeny with unprecedented taxon sampling. Molecular Phylogenetics and Evolution 71: 204–213. http://dx.doi.org/10.1016/j.ympev.2013.11.011

Schlegel H (1837) Abbildungen neuer oder unvollständig bekannter Amphibien, nach der Natur oder dem Leben entworfen, herausgegeben und mit einem erläuternden Texte begleitet. Part 1. Arnz & Co, Düsseldorf.

Schneider JG (1799) Historiae Amphibiorum narturalis et literariae. Fasciculus primus, continens Ranas. Calamitas, Bufones, Salamandras et Hydros. Frommanni, Jena, 266. https://doi.org/10.5962/bhl.title.78757

Schneider JG (1801) Historiae Amphibiorum naturalis et literariae. Fasciculus secundus continens Crocodilos, Scincos, Chamaesauras, Boas. Pseudoboas, Elapes, Angues. Amphisbaenas et Caecilias. Frommanni, Jena, 374 pp

Sclater WL (1892) On some specimens of frogs in the Indian Museum, Calcutta with description of several new species. Proceedings of the Zoological Society of London 1892: 341–348.

Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analyses. Canadian Journal of Zoology 69: 82–90. https://doi.org/10.1139/z91-013

Sheridan JA, Bickford D, Su KF-Y (2010) An examination of call and genetic variation in three wide-ranging Southeast Asian anuran species. Raffles Bulletin of Zoology 58: 369–379. http://scholarbank.nus.edu.sg/handle/10635/100064

Sheridan JA, Stuart, BL (2018) Hidden species diversity in *Sylvirana nigrovittata* (Amphibia: Ranidae) highlights the importance of taxonomic revisions in biodiversity conservation. PLOS One 13(3): e0192766. https://doi.org/10.1371/journal.pone.0192766

Smith MA (1940) The amphibians and reptiles obtained by Mr. Ronald Kaulback in Upper Burma. Records of the Indian Museum 42: 465–486.

Stuart BL, Inger RF, Voris HK (2006a) High levels of cryptic species diversity revealed by sympatric lineages of Southeast Asian forest frogs. Biology Letters 2006: 470–474. https://doi.org/10.1098/rsbl.2006.0505

Stuart BL, Chuaynkern T, Chan-ard T, Inger RF (2006b) Three species of frogs and a new tadpole from eastern Thailand. Fieldiana Zoology 1543: 1–10. http://dx.doi.org/10.3158/0015-0754(2006)187[1:TNSOFA]2.0.CO;2

Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues ASL, Fischman DL, Waller RW (2004) Status and trends of amphibian declines and extinctions worldwide. Science 306: 1783–1786. https://doi.org/10.1126/science.1103538

Sumida M, Kotaki M, Islam MM, Djong TH, Igawa T, Kondo Y, Matsui M, De Silva A, Khonsue W, Nishioka M (2007) Evolutionary relationships and reproductive isolating mechanisms in the Rice Frog (*Fejervarya limnocharis*) species complex from Sri Lanka, Thailand, Taiwan and Japan, inferred from mtDNA gene sequences, allozymes, and crossing experiments. Zoological Science 24: 547–562. https://doi.org/10.2108/zsj.24.547

Sumontha M, Pauwels OSG, Kunya K, Limlikhitaksorn C, Ruksue S, Taokratok A, Ansermet M, Chanhome L (2012) A new species of Parachute Gecko (Squamata: Gekkonidae: genus *Ptychozoon*) from Kaeng Krachan National Park, western Thailand. Zootaxa 3513: 68–78. http://www.mapress.com/zootaxa/2012/f/z03513p078f.pdf

Sumontha M, Pauwels OSG, Panitvong N, Kunya K, Grismer LL (2015) A new lowland forest Bent-toed Gecko (Squamata: Gekkonidae: *Cyrtodactylus*) from Ranong Province, peninsular Thailand. Zootaxa 3911: 106–118. https://doi.org/10.11646/zootaxa.3911.1.6.

Suwannapoom C, Yuan Z-Y, Chen J-M, Hou M, Zhao H-P, Want L-W, Nguyen TS, Nguyen TQ, Murphy RW, Sullivan J, McLeod DS, Che J (2016) Taxonomic revision of the Chinese *Limnonectes* (Anura, Dicroglossidae) with the description of a new species from China and Myanmar. Zootaxa 4093: 181–200. ttp://dx.doi.org/10.11646/zootaxa.4093.2.2

Swofford DL (2002) PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP (2003) A plea for DNA taxonomy. Trends in Ecology and Evolution 18: 70–74. https://doi.org/10.1016/S0169-5347(02)00041-1

Theobald W (1868) Catalogue of the reptiles of British Birma, embracing the Provinces of Pegu, Martaban, and Tenasserim; with descriptions of new or little-known species. Journal of the Linnean Society, Zoology 4: 4–67. https://doi.org/10.1111/j.1096-3642.1868.tb02007.x

Tonione MA, Reeder N, Moritz CC (2011) High Genetic Diversity Despite the Potential for Stepping-Stone Colonizations in an Invasive Species of Gecko on Moorea, French Polynesia PLOS ONE 6: e26874. https://doi.org/10.1371/journal.pone.0026874

Uetz P, Freed P, Hošek J (2018) The Reptile Database, http://www.reptile-database.org [Accessed 31 January 2018]

Vogel G, David P, Pauwels OSG, Sumontha M, Norval G, Hendrix R, Vu NT, Ziegler T (2009) A revision of *Lycodon ruhstrati* (Fischer 1886) *auctorum* (Squamata Colubridae), with the description of a new species from Thailand and a new subspecies from the Asian mainland. Tropical Zoology 22: 131–182.

Vogel G, David P (2012) A revision of the species group *Xenochrophis piscator* (Schneider, 1799) (Squamata: Natricidae). Zootaxa 3473: 1–60.

Vogt T (1911) Beitrag zur Amphibien-fauna der Insel Formosa. Sitzungsberichte der Gesellschaft Naturforschender Freunde zu Berlin 1911:179–184.

Weigt LA, Driskell AC, Baldwin CC, Ormos A (2012) DNA Barcoding Fishes. In: Kress WJ, Erickson DL (Eds) DNA Barcodes: Methods and Protocols, Humana Press, New York, 109–126. https://doi.org/10.1007/978-1-61779-591-6_6

Wilson EO (2017) Biodiversity research requires more boots on the ground. Nature Ecology & Evolution 1: 1590–1591. https://doi.org/10.1038/s41559-017-0360-y

Wijayathilaka N, Garg S, Senevirathne G, Karunarathna N, Biju SD, Meegaskumbura M (2016) A new species of *Microhyla* (Anura: Microhylidae) from Sri Lanka: an integrative taxonomic approach. Zootaxa 4066: 331–342. http://doi.org/10.11646/zootaxa.4066.3.9

Wilkinson JA, Sellas AB, Vindum JV (2012) A new species of *Ansonia* (Anura: Bufonidae) from northern Tanintharyi Division, Myanmar. Zootaxa 3163: 54–68. http://www.mapress.com/zootaxa/2012/f/z03163p068f.pdf

Wilkinson M, Presswell B, Sherratt E, Papadopoulou A, Gower DJ (2014) A new species of striped *Ichthyophis* Fitzinger, 1826 (Amphibia: Gymnophiona: Ichthyophiidae) from Myanmar. Zootaxa 3785: 45–58. http://dx.doi.org/10.11646/zootaxa.3785.1.4

Wogan GOU, Vindum JV, Wilkinson JA, Koo MS, Slowinski JB, Win H et al. (2008) New country records and range extensions for Myanmar amphibians and reptiles. Hamadryad 33: 83–96.

Wogan GOU, Stuart BL, Iskandar DT, McGuire JA (2016) Deep genetic structure and ecological divergence in a widespread human commensal toad. Biology Letters 12: 20150807. http://dx.doi.org/10.1098/rsbl.2015.0807

Wood PL, Grismer LL, Aowphol A, Aguilar CA, Cota M, Grismer MS, Murdoch ML, Sites Jr. JW (2017) Three new karst-dwelling *Cnemaspis* Strauch, 1887 (Squamata; Gekkonidae) from Peninsular Thailand and the phylogenetic placement of *C. punctatonuchalis* and *C. vandeventeri*. PeerJ 5: e2884. https://doi.org/10.7717/peerj.2884

Zug GR (2011) Tropical Asian Dry-Forest Amphibians & Reptiles: A Regional Comparison of Ecological Communities. In: McShea WJ, Davies SJ, Bhumpakphan N (Eds) The Ecol-

ogy and Conservation of Seasonally Dry Forest in Asia. Smithsonian Institution Scholarly Press, Washington DC, 275–303.

Zug GR (2015) Morphology and systematics of *Kalophrynus interlineatus-pleurostigma* populations (Anura: Microhylidae: Kalophryninae) and a taxonomy of the genus *Kalophrynus* Tschudi, Asian Sticky Frogs. Proceedings of the California Academy of Sciences 62: 135–190.

Zug GR, Blackburn III JL, Sai Wunna Kyi (2006a) Checkered Keelbacks (*Xenochrophis* – Reptilia, Serpentes, Natricidae) at the Moyingye Wetland Bird Sanctuary, Myanmar. Hamadryad 30: 157–166.

Zug GR, Brown H, Schulte J, Vindum JV (2006b) Systematics of the garden lizards, *Calotes versicolor* group (Reptilia, Squamata, Agamidae), in Myanmar: central dry zone populations. Proceedings of the California Academy of Sciences 57: 35–68.

Zug GR, Mulcahy DG, Vindum JV (2017) Resurrection of *Bronchocela burmana* Blanford, 1878 for the Green Crested Lizard (Squamata, Agamidae) of southern Myanmar. ZooKeys 657: 141. https://doi.org/10.3897/zookeys.657.11600.

Zug GR, Vindum JV, Koo M (2007) Burmese *Hemidactylus* (Reptilia, Squamata, Gekkonidae): taxonomic notes on tropical Asian *Hemidactylus*. Proceedings of the California Academy of Sciences 58: 387–405.

Zug GR, Win H, Thin T, Min TZ, Lhon WZ, Kyaw K (1998) Herpetofauna of the Chatthin Wildlife Sanctuary, north-central Myanmar, with preliminary observations on their natural history. Hamadryad 23: 111–120.

Zuniga JD, Gostel MR, Mulcahy DG, Barker K, Funk VA, Coddington JC (2017) Data Release: DNA Barcodes of Plant Species Collected for the Global Genome Initiative Gardens Project, National Museum of Natural History, Smithsonian Institution. PhytoKeys 88: 119–122. https://doi.org/10.3897/phytokeys.88.14607

Supplementary material 1

12S and 16S neighbor-joining trees

Authors: Mulcahy DG, Lee JL, Miller AH, Chand M, Thura MK, Zug GR

Data type: molecular data

Link: https://doi.org/10.3897/zookeys.757.24553.suppl1