*Syst. Biol.* 63(4):518–533, 2014

DOI:10.1093/sysbio/syu017
Advance Access publication March 13, 2014

# Global Biodiversity Assessment and Hyper-Cryptic Species Complexes: More Than One Species of Elephant in the Room?

MARK ADAMS[1,2,*], TARMO A. RAADIK[3,4], CHRISTOPHER P. BURRIDGE[5], AND ARTHUR GEORGES[4]

[1]*Evolutionary Biology Unit, South Australian Museum, North Terrace, SA 5000, Australia;*
[2]*School of Earth and Environmental Sciences, University of Adelaide, Adelaide, SA 5005, Australia;*
[3]*Aquatic Ecology Section, Arthur Rylah Institute for Environmental Research, Department of Environment and Primary Industries, 123 Brown Street, Heidelberg, VIC 3084, Australia;*
[4]*Institute for Applied Ecology and Collaborative Research Network for Murray-Darling Futures, University of Canberra, ACT 2601, Australia and*
[5]*School of Biological Sciences, University of Tasmania, Private Bag 55, Hobart, TAS 7001, Australia*
[*]*Correspondence to be sent to: Evolutionary Biology Unit, South Australian Museum, North Terrace, SA 5000, Australia; E-mail: mark.adams@samuseum.sa.gov.au*

*Received 31 July 2013; reviews returned 6 November 2013; accepted 6 March 2014*
*Associate Editor: Adrian Paterson*

*Abstract*.—Several recent estimates of global biodiversity have concluded that the total number of species on Earth lies near the lower end of the wide range touted in previous decades. However, none of these recent estimates formally explore the real "elephant in the room", namely, what proportion of species are taxonomically invisible to conventional assessments, and thus, as undiagnosed cryptic species, remain uncountable until revealed by multi-gene molecular assessments. Here we explore the significance and extent of so-called "hyper-cryptic" species complexes, using the Australian freshwater fish *Galaxias olidus* as a proxy for any organism whose taxonomy ought to be largely finalized when compared to those in little-studied or morphologically undifferentiated groups. Our comprehensive allozyme (838 fish for 54 putative loci), mtDNA (557 fish for 605 bp of cyt*b*), and morphological (1963–3389 vouchers for 17–58 characters) assessment of this species across its broad geographic range revealed a 1500% increase in species-level biodiversity, and suggested that additional taxa may remain undiscovered. Importantly, while all 15 candidate species were morphologically diagnosable *a posteriori* from one another, single-gene DNA barcoding proved largely unsuccessful as an *a priori* method for species identification. These results lead us to draw two strong inferences of relevance to estimates of global biodiversity. First, hyper-cryptic complexes are likely to be common in many organismal groups. Second, no assessment of species numbers can be considered "best practice" in the molecular age unless it explicitly includes estimates of the extent of cryptic and hyper-cryptic biodiversity. [Galaxiidae; global estimates; hyper-diverse; mountain galaxias; species counts; species richness.]

Whenever biologists contemplate the big questions in systematics, they usually include one of the most fundamental of all, namely "How many species are there on Earth?". Many have provided their best-guess estimates, with some targeting all life forms (Chapman 2009), others restricting themselves to eukaryotes (Scheffers et al. 2012; Costello et al. 2013), protists (Pawlowski et al. 2012) or prokaryotes (Curtis et al. 2002), and still others focusing on single prominent groups (e.g., arthropods, Basset et al. 2012; fungi, O'Brien et al. 2005; spiders, Platnick 1999) and biomes (marine vs. terrestrial, Mora et al. 2011; Appeltans et al. 2012; Costello et al. 2012; benthos, Grassle and Maciolek 1992). These estimates typically employ various combinations of empiricism and theory, and reflect a range of approaches that build from what is known (or thought to be known) via various multiplying factors (May 1988; Scheffers et al. 2012). Given this great array of methodologies, target scenarios, and assumptions, it comes as no surprise that estimates of global species counts for eukaryotes have varied >50-fold, from a low of around 2 M to a high of over 100 M (Costello et al. 2012). Consequently, there has been much contention, particularly over some of the more extreme estimates (e.g., Erwin's methodology based on rainforest insects; see e.g., Erwin 1982; Erwin 1991; Gaston 1991; Stork 1999; ØDegaard 2000; Hamilton et al. 2010).

Despite this diversity in focus group and methods, two common themes often lie at the core of most "how many species" (HMS) estimates. First, key reference groups that are comparatively well-known taxonomically (e.g., birds, butterflies, all vertebrates, all plants in Britain etc.; May 1988; Stork 1999; Dolphin and Quicke 2001) are nominated to serve as the foundation for extrapolating outwards to all other groups, regions, or biomes under scrutiny. Second, various "rate of taxonomic discovery" plots are used to predict when and where plateaus are likely to occur for the X (time) and Y (number of species) axes. Importantly, implicit in these themes are two underlying assumptions, namely (1) our existing taxonomic knowledge of the key reference groups is a reasonably accurate reflection of the real situation and thus not itself a major unknown in the HMS process; and (2) the ease of species discovery only reflects extrinsic factors (e.g., the number of taxonomists, funding levels for taxonomy) rather than itself decreasing over time, as species counts approach plateaus.

Given their central importance to HMS estimation, these assumptions deserve to be critically examined across a range of organismal groups and biomes. Here we explore one phenomenon that arguably challenges the legitimacy of both assumptions, namely that of the "hyper-cryptic" species complex. These are defined herein as any taxon currently regarded as a

single "species" or any related group of taxonomically confused "species", that in reality consists of a "large" number of valid but undiagnosed species. For convenience, we have arbitrarily defined "large" to reflect any increase in species-level biodiversity including and beyond a 4-fold increase (i.e., one "species" becomes four or more valid species; four named "species" in a complex together comprise 16 or more valid species etc.). More importantly, our definition specifically excludes the phenomenon of "taxonomic inflation", whereby extra species are recognized under a rigid application of the phylogenetic species concept (Isaac et al. 2004). Instead, we have used a more conservative, total-evidence approach, based on reproductive isolation and/or diagnosability at multiple, independent, nuclear genes (including of course genes that influence morphology, Frankham et al. 2012; Hammer et al. 2013). Moreover, while this approach is best characterized as the "unified species concept" (de Queiroz 2007), all of the candidate taxa identified herein are valid species under most other "primary" species concepts (Hammer et al. 2013), including the morphological (all taxa), biological (all sympatric combinations), and evolutionary (all taxa) concepts.

Some might contend that hyper-cryptic species are just a more extreme case of the now well-established phenomenon of cryptic species (morphologically similar species that remain undiagnosed, *sensu* Bickford et al. 2007; often confused with the term "sibling species", which we and others (Bickford et al. 2007) prefer to employ for near-relatives, regardless of comparative morphology). We will argue, drawing on the results of a detailed molecular and morphological investigation on a widespread "species" of Australian freshwater fish, that novel considerations emerge when the number of undiagnosed species within a species complex goes beyond a small number (herein a 15-fold increase). Moreover, we believe that focusing on the prevalence and extent of cryptic species is necessary in any case, because most assessments of planetary biodiversity either do not address the issue (May 1988; Erwin 1991; Stork 1993; Mora et al. 2011; Costello et al. 2012) or underplay its significance (Stork 1999; Costello et al. 2013).

Several recent, high-profile HMS assessments have concluded that planetary eukaryote species counts are more likely to be at the low end of the 2-100 M range (Chapman 2009; Mora et al. 2011; Appeltans et al. 2012; Costello et al. 2012; Costello et al. 2013). However, none explicitly acknowledge that "cryptic biodiversity" (*sensu* Beheregaray and Caccone 2007) is likely to prove a major impediment to the rate of taxonomic discovery, nor do they explore future scenarios in which next-generation DNA sequencing might greatly facilitate the taxonomic splitting of single "morphospecies" into their component evolutionary species (Emerson et al. 2011). In stark contrast, many studies have pointed to the prevalence of cryptic species (Bickford et al. 2007; Scheffers et al. 2012), and their broad distribution across most organismal groups and biomes (Pfenninger and Schwenk 2007; Pawlowski et al. 2012). Further reflecting this trend, the yearly number of published studies that refer to cryptic or sibling species is maintaining the exponential increase noted by Bickford et al. (2007) since the late 1980's (Supplementary Fig. S1), fuelled by current and expected advances in next-generation DNA sequencing technologies (Lemmon and Lemmon 2013) and in the phylogenetic analysis of multi-gene sequences (Song et al. 2012). Thus we suggest the time has come for all future HMS assessments to fully acknowledge the elephant in the room (an animal that, befittingly, has its own complex taxonomic history of between two and nine species, Laursen and Bekoff 1978; Shoshani and Eisenberg 1982).

Despite their disparity in scope, a general consensus is evident across HMS studies that there ought to be far less taxonomic uncertainty (and therefore fewer hyper-cryptic complexes) in organismal groups that are (1) multi-cellular; (2) well-studied taxonomically; (3) of large body size, with many assessable morphological features; (4) located in temperate rather than tropical regions; (5) common and readily accessible to collectors; and (6) found in affluent nations with an established taxonomic infrastructure and workforce. Under these criteria our target species, the mountain galaxias, *Galaxias olidus* Günther, 1866 (Teleosti: Galaxiidae), clearly merits a comparatively low probability of being hyper-cryptic (Fig. 1).

Mountain galaxias are found throughout the temperate southeastern portion of mainland Australia (Fig. 2), where they are widely distributed throughout numerous river basins, spread across 4 of the country's 12 drainage divisions (AWRC 1976). A small to moderate-sized fish (average size 60-80 mm, maximum 140 mm; Allen et al. 2002), *G. olidus* displays a wide range of morphotypic forms (Supplementary Fig. S2) and occupies a variety of habitats, from large lowland rivers to small, high-altitude streams (altitudinal range 50–>2000 m). As befitting a prominent inhabitant of one of the most accessible and studied freshwater ecosystems in Australia, mountain galaxias have been subjected to a number of taxonomic revisions since 1866 (reviewed by Raadik 2011). These have variously resulted in the recognition of between one and six species. Importantly however, the most recent revision, coauthored in 1981 by a world authority on galaxiid taxonomy, concluded that all mountain galaxias belonged to one morphologically variable species (McDowall and Frankenberg 1981).

The revision of McDowall and Frankenberg (1981) has been largely followed in all subsequent relevant publications (Allen 1989; Allen et al. 2002; Hoese et al. 2006; Humphries and Walker 2013), with one interesting exception. Many populations of mountain galaxias have suffered extirpation or considerable decline since European settlement, reflecting its occurrence throughout the most intensively populated and freshwater-regulated region of the country (Tilzey 1976; Raadik 2011). While this process has afflicted many regions and morphotypic forms, it has been particularly severe on one of the most distinctive forms, the barred galaxias (originally described as

| Group attribute | Greater ← Potential for cryptic biodiversity → Lesser | | |
|---|---|---|---|
| Kingdom | prokaryote/protists → | 'basal' eukaryotes → | 'derived' eukaryotes |
| Morphological complexity | simple → | moderate → | complex |
| Biome | tropical → | temperate → | extreme |
| Taxonomic effort | little → | moderate → | intensive |
| Molecular framework | none → | basic → | detailed |
| Ease of access to specimens | very difficult → | not easy → | easy |
| Infrastructure of host nation(s) | undeveloped → | developing → | developed |

FIGURE 1. Summary of major factors influencing the extent to which a species or species group is likely to be hyper-cryptic or harbor cryptic biodiversity. Our study species, *G. olidus*, would receive a relatively low ranking for most factors (shown underlined), compared to a large proportion of the planet's other "species".

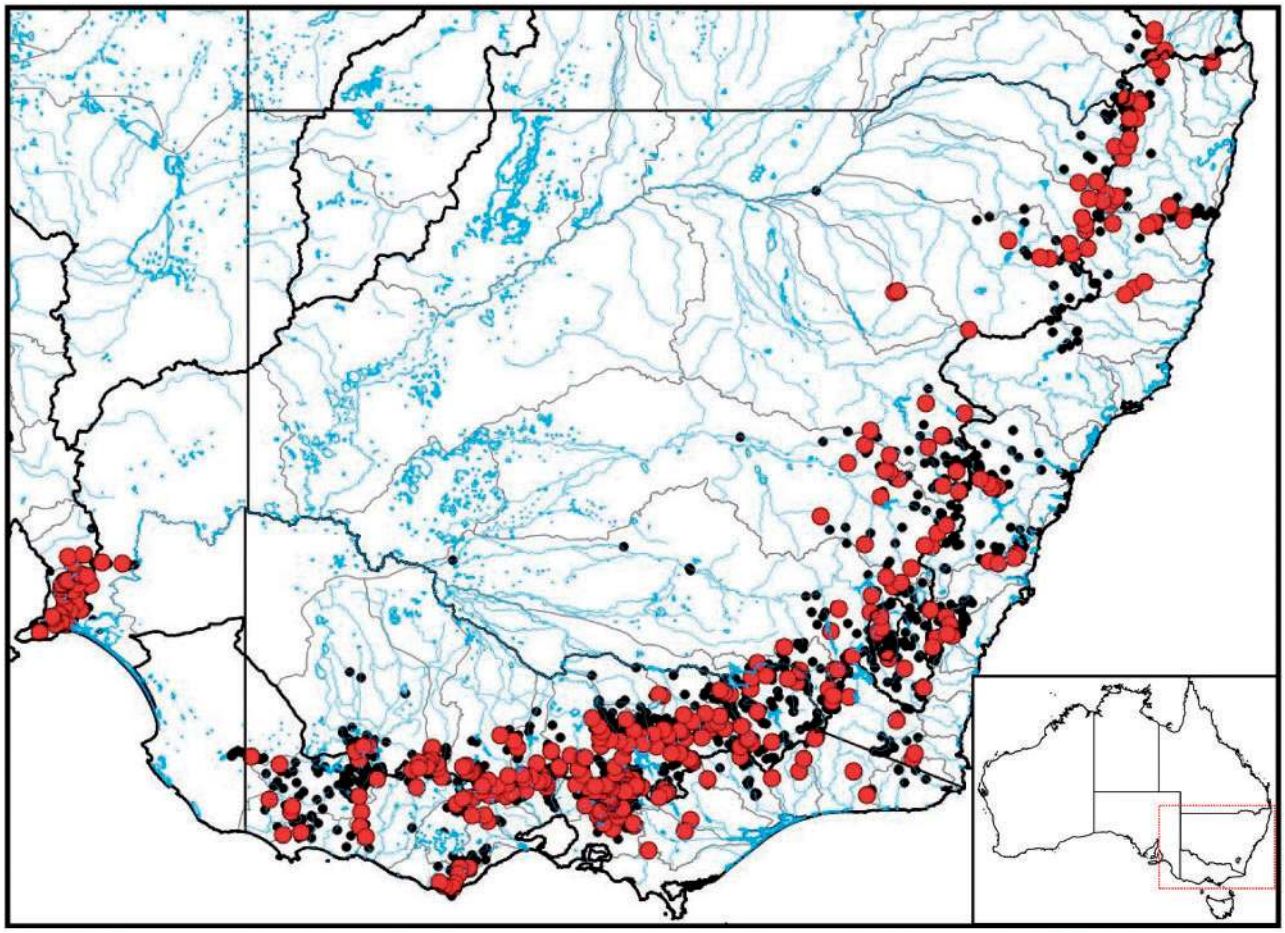

FIGURE 2. Map indicating the broad geographic distribution of *G. olidus sensu* McDowall and Frankenberg (1981) and the location of all sites surveyed in the allozyme study. *Legend*: large circles = sites included in the allozyme study; small dots = additional sites for which either morphological vouchers or ethanol-only tissues were examined.

*G. fuscus* Mack 1936), leading most workers to follow the precautionary principle and recognize it as a valid species (Raadik et al. 1996; Allen et al. 2002; Humphries and Walker 2013).

This study has three broad aims. First, it introduces the concept of hyper-cryptic complexes and explores the concept in our target species *G. olidus*. Second, it hypothesizes that hyper-cryptic complexes are common across many organismal groups and likely to be extremely common in those groups that make up the base of the pyramid of planetary biodiversity. Finally, it advocates that future HMS assessments take into account the reality of hyper-cryptic complexes in particular and cryptic biodiversity in general.

TABLE 1. Summary of taxon codes and sample sizes per data set for each of the 15 candidate taxa identified within the *G. olidus* complex by the allozyme analyses and thereafter diagnosed morphologically

| Taxon code | Informal name | Allozymes | mtDNA | Morphometrics | Meristics | Supplementary morphology |
|---|---|---|---|---|---|---|
| AR | "arte" | 7 (1) | 4 (1) | 15 (1) | 33 (2) | 24 |
| BA | "bass" | 131 (54) | 71 (44) | 297 (38) | 401 (39) | 141–171 |
| DA | "dargo" | 3 (1) | 7 (1) | 10 (1) | 41 (1) | 21–48 |
| FU | "fuscus" | 34 (15) | 36 (17) | 74 (14) | 188 (23) | 114–136 |
| GE | "genoa" | 19 (3) | 19 (5) | 31 (3) | 123 (5) | 28–65 |
| JI | "jibolaro" | 6 (1) | 5 (1) | 6 (1) | 9 (1) | 6–9 |
| KO | "kosciusko" | 14 (2) | 14 (2) | 26 (2) | 44 (2) | 22–30 |
| OL | "olidus s.s. " | 393 (178) | 239 (176) | 1074 (142) | 1535 (146) | 464–500 |
| OR | "oliros" | 164 (80) | 93 (73) | 303 (55) | 512 (70) | 89–236 |
| RF | "riffle" | 26 (9) | 30 (12) | 66 (10) | 394 (43) | 89–103 |
| RI | "rintoul" | 5 (1) | 7 (1) | 11 (1) | 17 (1) | 11–17 |
| RO | "rodger" | 7 (1) | 5 (1) | 15 (1) | 28 (1) | 23–26 |
| SH | "shaw" | 5 (1) | 7 (1) | 5 (1) | 28 (2) | 5–23 |
| ST | "stoney" | 8 (1) | 8 (1) | 15 (1) | 21 (1) | 15–19 |
| TA | "tantangara" | 6 (1) | 6 (1) | 15 (1) | 15 (1) | 15 |
| | Total | 828[a] (349[a]) | 557 (337) | 1963 (272) | 3389 (338) | 1067–1422 |

Note: The number of individuals and sites (in parentheses) are shown for all major data sets, whereas only the range of sample sizes examined is presented for the 12 supplementary morphological characters.
[a] An additional 10 individuals of likely hybrid ancestry (involving 1 additional site) were also genotyped.

## METHODS

### *Sampling*

Tissues and morphological vouchers were collected over a 5-year period (2001-2005) by comprehensively surveying 1187 sites across all rivers in southeastern Australia known or likely to harbor mountain galaxias (Fig. 2, Supplementary Table S1). All retained specimens ($n = 9740$) were ethically euthanized using a solution of clove oil in water (Adams et al. 2011). Representatives of every morphotypic form at a site (usually only one, sometimes two, rarely three) were snap-frozen whole in liquid nitrogen for future genetic analysis, and the remainder fixed in formalin for subsequent morphological analysis. This contemporary sampling provided the majority of specimens used in the molecular appraisals and a significant proportion of the vouchers used in the detailed morphological assessments. Additional tissues were obtained from the South Australian Museum's frozen tissue collection, and a large number of older museum vouchers were examined to confirm that morphological diagnoses based on contemporary vouchers could also be employed retrospectively. Sample sizes for the allozyme, mtDNA, and morphological components of this study are shown in Table 1, as are the taxon codes employed for the 15 candidate species. Maps were generated using the open source GIS software Quantum GIS, version 1.8.0 (http://qgis.org/).

### *Allozyme Analyses*

A range of molecular data sets can now be employed to provide the type of nuclear genetic markers required to assist with species delineation (Hammer et al. 2013). Of these, the five most widely used are allozymes, microsatellites, nuclear DNA sequences (nDNA), SNPs (single nucleotide polymorphisms), and AFLPs (amplified fragmented length polymorphisms). In this study we ruled out all but one data-type as either being too variable or hypervariable for wholesale species delineation (microsatellites and SNPs; Frankham et al. 2010; Hammer et al. 2013), not comprising simple, codominant markers (nDNA sequences, AFLPs), or too expensive (nDNA, SNPs). As a consequence, we chose allozyme markers as the primary means of identifying taxa. Allozyme analysis has a long and successful history of delineating species (Avise 1975; Richardson et al. 1986; Hammer et al. 2007). In the hands of experienced practitioners, the technique can rapidly provide a genetic assessment across a large number ($> 50$) of independent genetic markers, each displaying codominant and biparentally inherited alleles. Moreover, allozyme markers can readily detect instances of hybridization and/or introgression (both common phenomena among congeneric fishes; Vespoor and Hammar 1991) and thereafter identify the likely parental taxa (Bertozzi et al. 2000; Adams et al. 2011).

Allozyme electrophoresis of muscle homogenates was undertaken on cellulose acetate gels following the methods of Adams et al. (2011). The following enzymes or non-enzymatic proteins were successfully surveyed: ACON, ACP, ADA, ADH, AK, ALD, AP, CA, CK, ENOL, FDPASE, FUM, GAPD, GDA, GLO, GOT, GP, G6PD, GPI, GSR, IDH, LDH, MDH, ME, MPI, NDPK, NP, PEPA, PEPB, PEPD, PGAM, 6PGD, PGM, PK, SOD, SORDH, TPI, and UGPP. Details of enzyme/locus abbreviations, enzyme commission numbers, electrophoretic conditions, and stain recipes are presented elsewhere (Richardson et al. 1986; Hammer et al. 2007).

Genetic markers are best suited for delineating candidate species when individuals are used as the unit of analysis (Horner and Adams 2007). However, efforts to use the Bayesian clustering programs STRUCTURE, GENELAND, and STRUCTURAMA to define candidate taxa using the multi-locus data set for individuals were unsuccessful, for two predictable reasons. First, our sample sizes per site (mean = 2.4; range = 1–8) are too small to provide the statistical power needed to define populations from first principles. Second, these programs identify populations based on their constituent individuals conforming to Hardy–Weinberg expectations across all loci, an attribute rarely displayed by any taxonomic units beyond the level of population, let alone those occurring in highly fragmented freshwater habitats (Hammer et al. 2013).

Instead we used the multivariate clustering procedure of Principal Coordinates Analysis (PCoA; Gower 1966) to identify cohesive genetic groups among individuals, independent of other *a priori* considerations of locality, mtDNA haplotype, morphology, or ecology. PCoA is most powerful when employed in a stepwise fashion that is, starting with all individuals and thereafter progressively restricting follow-up analyses to those subsets of individuals clearly defined by diagnostic differences at multiple allozyme loci from one another in previous PCoAs. This approach is especially necessary when there are more than a handful of genetically similar taxa, because well-defined genetic clusters often overlie one another in the first few dimensions of the initial PCoA, despite being readily distinguishable in deeper dimensions (Georges and Adams 1992; Horner and Adams 2007). Importantly, follow-up PCoAs are undertaken on every primary and secondary genetic group thus diagnosed, until there remains no evidence of further diagnosable taxa using the nominated criteria for inferring species-level differences. To further confirm the integrity of all putative taxa, stepwise PCoAs were also undertaken on the individuals occurring in each of three regions, namely "northern", "central", and "western" (defined in Supplementary Table S1).

Our two genetic criteria for delineating candidate species in this study were (1) major differences ($\Delta p > 50\%$) in allele frequency at $n \geq 5$ allozyme loci for PCoA-defined groups in full or partial sympatry, or (2) fixed or "near-fixed" differences (the latter allowing a cumulative tolerance of 10% for any shared alleles per locus, Adams et al. (2013)) at multiple allozyme loci for PCoA-defined groups in allopatry. Given the decreased analytical power inherent in working with smaller sample sizes, we set the minimum number of fixed/near-fixed differences required to two for a combined $n \geq 20$, and to three for a combined $n < 20$. Ultimately, most candidate species were diagnosable from one another at many more than these minimum numbers of loci. The operational details of stepwise PCoA are presented in Hammer et al. (2007).

Once all individuals had either been assigned to candidate species ($n = 828$) or flagged as having a putative hybrid ancestry ($n = 10$), the overall genetic affinities among "pure" species were depicted in a NJ network, constructed from a pairwise matrix of unbiased Nei's D values (procedures detailed in Adams et al. 2013).

### *Morphological Analyses*

Given the sheer number of sites surveyed and individuals collected, it was not practicable for the primary field collector (Raadik) to sample the same fish for both genetic and morphological analysis. Consequently, all morphological vouchers were assigned to an allozymically defined candidate taxon on the basis of (1) being collected at the same site and time; and (2) possessing the same external morphotype as the individuals used for the genetic determination of each taxon. While clearly not the ideal scenario, the only likely bias introduced by this protocol would be to underestimate the true number of species, as could occur if a rare and sympatric novel species was not selected in the allozyme analyses plus remained morphologically cryptic after taxonomic reappraisal. Data for individuals of suspected hybrid ancestry (i.e., all vouchers from populations identified allozymically as containing putative hybrids) were excluded from all analyses.

A total of 29 morphometric, 17 meristic, and 12 "supplementary" morphological characters were used to reappraise the morphological affinities of the candidate species identified by allozyme profiling. Summaries of all characters are provided in Supplementary Text 1 and full details of all characters are presented in Raadik (2011). A range of univariate analyses were initially undertaken on each individual data set to identify outliers and to determine which characters were likely to be taxonomically informative for pairwise combinations of candidate species. Thereafter, the multivariate procedure of Principal Components Analysis (PCA; Pearson 1901), using the covariance matrix, was used to assess morphological diagnosability separately for the meristic and morphometric data sets. PCAs were conducted both on the entire data set and on all pairwise combinations of candidate species.

Where pairwise PCA failed to unequivocally discriminate two candidate species (i.e., taxon ellipses overlapped in the first three dimensions), the data were subjected to pairwise Discriminant Functions Analysis (DFA; Fisher 1936). Here, taxa were considered morphologically distinctive and diagnosable when DFA correctly classified 80% or greater of individuals to the appropriate taxon. In the three instances where classification success was $< 80\%$ (always due to one or both taxa being geographically widespread, and hence exhibiting considerable within-taxon variation), a further DFA was conducted on a reduced "regional" data set, which consisted only of those individuals from the catchment(s) most proximate to one another. All statistical analyses were undertaken using SPSS 15.0.0 and R, version 2.8.1 (R Development Core Team 2009). More detailed summaries of the univariate and multivariate procedures are presented

in Supplementary Text 2 and full details are contained in Raadik (2011).

### *MtDNA Analyses*

A large mtDNA barcoding component was included in this study to (1) compensate for the inability of the allozyme data set to provide detailed phylogenetic and phylogeographic insights among and within candidate taxa; (2) provide an independent assessment of the prevalence of introgression among taxa; and (3) objectively assess the efficacy of mtDNA barcoding for delimiting taxa from first principles.

Total genomic DNA was extracted using a chelex-based procedure (Walsh et al. 1991). Following an initial trial of three mitochondrial genes (ATPase, CR, and cyt*b*) on a panel of 25 ingroup and outgroup tissues, cyt*b* was chosen as the most reliable of these equally informative mtDNA markers. The cyt*b* region was amplified in a 25 μL polymerase chain reaction (PCR) containing 20-100 ng template DNA, 1x PCR buffer, 2.5 mM MgCl2, 0.2 mM each dNTP, 0.2 μM of primers CytB Glu 5′-GAAAAACCACCGTTGTTATTCA-3′ and CytB Thr 5′-CGACTTCCGGATTACAAGACT-3′ (Waters et al. 2001), and 0.5 units Taq polymerase (Bioline Red Taq). The PCR cycling conditions were as follows: 95°C for 3min, followed by 32 cycles of 95°C for 30s, 55°C for 30s, and 72°C for 90s, and a final extension of 72°C for 5min. All reactions were run on an Eppendorf Mastercycler 5333 version 2.30 thermal cycler. PCR products were purified using a standard polyethylene glycol procedure (Sambrook and Russell 2001). Purified DNA was dried, resuspended in 15 μL TLE buffer (1 mM Tris, 10 mM EDTA), and sent to Macrogen (Seoul, Korea), for sequencing on an ABI automated sequencer.

Sequence data were edited and aligned using CLUSTAL X (Thompson et al. 1997) in the program Geneious Pro 3.8.2 (Biomatters Ltd). The final data set comprised 557 sequences for 605 bp of cyt*b*, plus single sequences for each of six outgroup species from southeastern Australia, namely *G. parvus* (the likely sister lineage to *G.* "olidus"; Burridge et al. 2012), *G. brevipinnis*, *G. fontanus*, *G. maculatus*, *G. robustus*, and *G. truttaceus*. Following the conclusions of Srivathsan and Meier (2011) that *p*-distances are more appropriate than K2P distances for DNA barcoding studies, we used MEGA 5.10 (Tamura et al. 2011) to generate a NJ tree (bootstrapped 1000 times) for all unique haplotypes. To ensure that this barcoding analysis was not concealing useful phylogenetic information, we also undertook a full maximum-likelihood phylogenetic analysis, following the methodologies used by Unmack et al. (2012).

## Results

### *Allozyme Delineation of Candidate Species*

The final allozyme data set comprised 838 individuals, collected from 350 sites and genotyped at 54 putative allozyme loci. Overall, more than 40 PCoAs, each requiring an associated assessment of allozyme diagnosability for the groups evident therein, were undertaken on the composite and pure groups or subgroups diagnosed by the stepwise PCoA procedure. Here we present a small but representative subset of these PCoAs (Fig. 3; Figs. S3–S5) plus a summary table of the final candidate species identified by these analyses and the pairwise number of diagnostic allozyme loci among them (Table 2).

An initial PCoA on all individuals (labeled for convenience according to their final taxon assignment) revealed three distinctive genetic groups (Fig. 3) ultimately shown to correspond to one "pure" taxon (taxon OR; no evidence of diagnosable subgroups in the follow-up PCoA), one composite group (*Grp1*; follow-up PCoA presented in Supplementary Fig. S3), and a group of three individuals whose genotypes and geographic location were consistent with a recent hybrid ancestry between taxon OR and taxon OL. These three putative "hybrids" and a further seven individuals that displayed evidence of hybrid ancestry under all three criteria (i.e., intermediate PCoA position, genotypic consistency for key diagnostic loci, a geographic overlap or adjacency of putative parental taxa) were excluded from subsequent analyses. Two follow-up PCoAs are also presented, one demonstrating the diagnosability of the two taxa FU and RF that comprised composite group *Grp2* and the second presenting a regional assessment of taxon KO versus taxon OL and a single admixed site (Supplementary Figs. S4 and S5, respectively).

A total of 15 diagnosable taxa were defined through the application of stepwise PCoA to our allozyme data set (allele frequency data presented in Supplementary Table S2). As shown in Table 2, each taxon was diagnosable from all others at a minimum of two allozyme loci (mean = 9.2; range 2-17), with most diagnosable at multiple loci (95% of values ≥4; 90% of values ≥6). An unrooted NJ network depicting the genetic relationships among these 15 taxa (Fig. 4) revealed little evidence of any correlation between the genetic affinities of allopatric taxa and their geographic proximity to one another (Fig. 5).

### *Morphological Assessments of the Taxonomic Validity of Candidate Species*

Two comprehensive data sets were generated and used as the primary means for assessing the taxonomic distinctiveness of the 15 taxa identified by allozyme analysis. As expected given its current taxonomic status as a single species, neither of the two "all taxa" PCAs on the morphometric (1963 individuals for 27 characters; Fig. 6) or meristic (3389 individuals for 17 characters; PCA not shown) data sets revealed any distinctive clusters that might indicate the presence of multiple species. However, a completely different perspective is obtainable when pairwise comparisons are undertaken

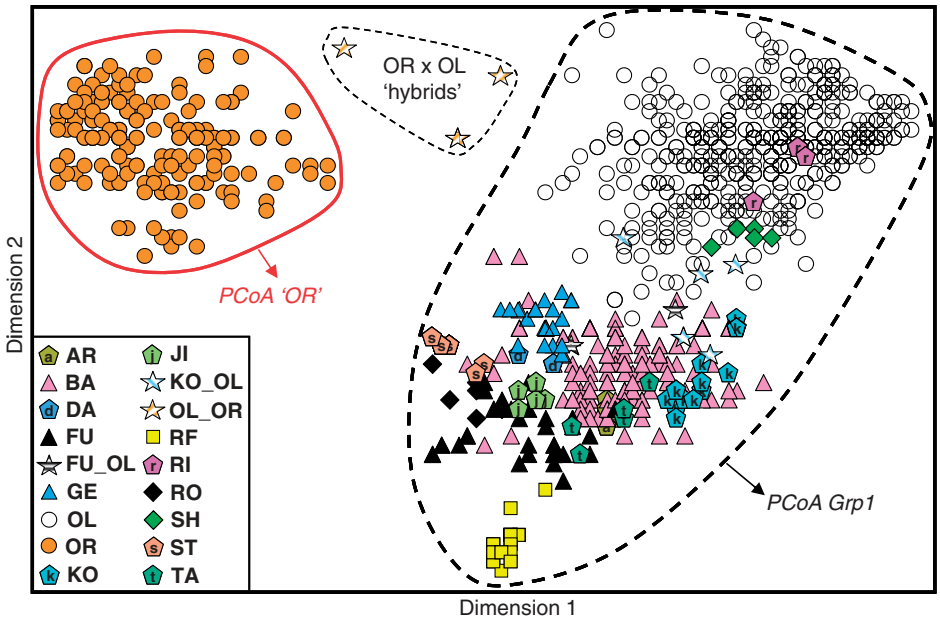


FIGURE 3. PCoA of all 838 individuals genotyped in the allozyme study. The relative PCoA scores have been plotted for the first (*X*-axis) and second (*Y*-axis) dimensions, which individually explained 19% and 11%, respectively, of the total multivariate variation present in 837 dimensions. Individuals are labeled according to their ultimate taxon identification, using the symbols indicated. The number of points is <838 because many individuals had identical PCoA scores in the first two dimensions. *Envelope legend*: thick dashed line = group subsequently found to include multiple taxa in follow-up PCoAs; thick solid line = no evidence of multiple taxa in a follow-up PCoA; thin dashed line = PCoA position and allozyme profiles consistent with a hybrid ancestry between the nominated taxa.

among the individual taxa identified by the allozyme analyses (see example in Supplementary Fig. S6).

The results of pairwise comparisons for our two primary data sets and for the secondary data set comprising 12 additional morphological features are summarized in Table 2. All 105 pairwise multivariate analyses (PCA alone if taxa were fully defined, otherwise DFA) of the morphometric data proved capable of diagnosing the 15 taxa from one another under the criteria employed. Moreover, all but three pairwise multivariate analyses of the meristic data were also indicative of these taxa being morphologically diagnosable, with even these three "failures" being only marginally below our DFA threshold of 80% (Table 2). After combining all allozyme and morphological "characters" (conservatively counting the morphometric and meristic analyses as one "super-character" each), the mean number of characters found to differentiate these 15 taxa was 13.1 (range 4-23; Supplementary Table S3). Most importantly, all taxa were diagnosable by a suite of individual morphological characters (Raadik 2011), ensuring that all merit recognition as valid biological (in sympatry) or evolutionary (in allopatry) species. Our current understanding of the geographic distribution of these 15 candidate species is depicted in Figure 5.

### *MtDNA Barcoding*

While our cyt*b* barcoding data set encompassed fewer individuals than were represented in the allozyme data set ($n = 557$ vs. $n = 838$), the two molecular analyses were similar in terms of geographic coverage (337 vs. 350 sites, with substantial overlap; Supplementary Table S1). A total of 271 unique haplotypes were identified for the ingroup, reflecting variation at 206 (161 parsimony informative) of the 605 bp sequenced. All ingroup and outgroup haplotype sequences have been deposited in GenBank (accession numbers KJ511493–KJ511763 for ingroup haplotypes 1–271 and KJ511764–KJ511769 for the six outgroup haplotypes), and the sequence alignment has been deposited in both Dryad (http://dx.doi.org/10.5061/dryad.tk043) and TreeBase (http://purl.org/phylo/treebase/phylows/study/TB2:S15464). Pairwise *p*-distances among ingroup haplotypes ranged from 0.002 to 0.076 (mean = 0.031), compared to the minimum ingroup–outgroup distance of 0.141 (range 0.141–0.196), and the minimum distance among outgroup species of 0.081 between *G. maculatus* and *G. rostratus*.

A NJ tree among these 271 haplotypes, rooted using the six outgroup species, is presented in Figure 7. Only a handful of well-supported clades were evident, with most defining terminal or shallow nodes. Importantly, none of the putative taxa defined by any of the various barcoding procedures used in the literature (van Velzen et al. 2012) corresponded with species diagnosed by multiple allozyme and morphological characters. This lack of concordance was particularly evident in the six most widespread species (BA, FU, GE, OL, OR, and RF; Fig. 5), most of which harbored haplotypes spread throughout the shallower portion of the tree, and all of which shared haplotypes with other sympatric or proximately distributed species (Fig. 7).

TABLE 2. Summary of all pairwise evaluations of taxon diagnosability for the four categories of taxonomic character employed on the *G. olidus* complex

| Taxon | AR | BA | DA | FU | GE | JI | KO | OL | OR | RF | RI | RO | SH | ST | TA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AR | — | $S^{2}$, $D^{96\%}$ | $S^{1}$, $D^{95\%}$ | $S^{2}$, $D^{90\%}$ | $S^{6}$, $D^{98\%}$ | $S^{1}$, $D^{100\%}$ | $S^{1}$, $D^{92\%}$ | $S^{2}$, $D^{92\%}$ | $S^{2}$, $D^{99\%}$ | $S^{4}$, $D^{91\%}$ | $S^{1}$, $D^{96\%}$ | $S^{1}$, $D^{91\%}$ | $S^{4}$, P | $S^{1}$, $D^{96\%}$ | $S^{1}$, $D^{100\%}$ |
| BA | $A^{7}$, $D^{87\%}$ | — | $Dr^{88\%}$ | $S^{1}$, $D^{84\%}$ | $S^{6}$, $D^{91\%}$ | $S^{1}$, $D^{88\%}$ | $D^{98\%}$ | $Dr^{83\%}$ | $S^{3}$, $D^{86\%}$ | $S^{2}$, $Dr^{86\%}$ | $Dr^{88\%}$ | $S^{1}$, $D^{96\%}$ | $Dr^{90\%}$ | $S^{1}$, $D^{90\%}$ | $S^{1}$, $D^{94\%}$ |
| DA | $A^{8}$, P | $A^{7}$, $D^{94\%}$ | — | $S^{1}$, [a]$D^{77\%}$ | $S^{7}$, $D^{95\%}$ | $D^{94\%}$ | $D^{98\%}$ | $Dr^{86\%}$ | $S^{3}$, $D^{95\%}$ | $S^{2}$, $D^{85\%}$ | $D^{81\%}$ | $D^{96\%}$ | $D^{81\%}$ | $D^{91\%}$ | $S^{1}$, $D^{98\%}$ |
| FU | $A^{9}$, P | $A^{5}$, $D^{98\%}$ | $A^{8}$, P | — | $S^{8}$, $D^{96\%}$ | $S^{1}$, $D^{90\%}$ | $S^{1}$, $D^{96\%}$ | [a]$D^{75\%}$ | $S^{2}$, $D^{93\%}$ | $S^{3}$ $D^{93\%}$ | $S^{1}$, $D^{89\%}$ | $S^{1}$, $D^{94\%}$ | $S^{1}$, $D^{83\%}$ | $S^{1}$, $D^{81\%}$ | $S^{2}$, $D^{95\%}$ |
| GE | $A^{12}$, P | $A^{9}$, $D^{99\%}$ | $A^{10}$, P | $A^{11}$, P | — | $S^{6}$, $D^{96\%}$ | $S^{6}$, P | $S^{6}$, $D^{90\%}$ | $S^{7}$, $D^{92\%}$ | $S^{9}$, $D^{88\%}$ | $S^{6}$, $D^{92\%}$ | $S^{7}$, P | $S^{6}$, $D^{96\%}$ | $S^{7}$, P | $S^{5}$, P |
| JI | $A^{9}$, P | $A^{8}$, $D^{88\%}$ | $A^{10}$, P | $A^{8}$, P | $A^{13}$, P | — | $S^{1}$, P | $D^{81\%}$ | $S^{2}$, $D^{89\%}$ | $S^{2}$, $D^{86\%}$ | $D^{89\%}$ | $S^{1}$, P | $D^{80\%}$ | $S^{1}$, $D^{97\%}$ | $S^{1}$, P |
| KO | $A^{4}$, $D^{98\%}$ | $A^{2}$, $D^{99\%}$ | $A^{6}$, P | $A^{8}$, $D^{100\%}$ | $A^{9}$, P | $A^{7}$, P | — | $S^{1}$, $D^{93\%}$ | $S^{3}$, $D^{96\%}$ | $S^{2}$, $D^{100\%}$ | P | $D^{91\%}$ | P | $D^{96\%}$ | $S^{1}$, P |
| OL | $A^{8}$, $D^{84\%}$ | $A^{2}$, $D^{84\%}$ | $A^{7}$, $D^{92\%}$ | $A^{5}$, $D^{99\%}$ | $A^{9}$, $D^{100\%}$ | $A^{7}$, $D^{88\%}$ | $A^{2}$, $D^{96\%}$ | — | $S^{2}$, $D^{90\%}$ | $S^{2}$, $Dr^{92\%}$ | $Dr^{88\%}$ | $S^{1}$, $D^{91\%}$ | [a]$D^{78\%}$ | $S^{1}$, $D^{83\%}$ | $S^{1}$, $D^{91\%}$ |
| OR | $A^{10}$, $D^{96\%}$ | $A^{2}$, $D^{95\%}$ | $A^{8}$, $D^{96\%}$ | $A^{7}$, P | $A^{11}$, $D^{100\%}$ | $A^{9}$, $D^{95\%}$ | $A^{7}$, $D^{100\%}$ | $A^{2}$, $D^{98\%}$ | — | $S^{5}$, $D^{93\%}$ | $S^{3}$, $D^{96\%}$ | $S^{3}$, $D^{99\%}$ | $S^{3}$, $D^{95\%}$ | $S^{3}$, $D^{95\%}$ | $S^{2}$, $D^{98\%}$ |
| RF | $A^{9}$, P | $A^{7}$, $D^{100\%}$ | $A^{9}$, P | $A^{6}$, P | $A^{11}$, P | $A^{10}$, P | $A^{7}$, P | $A^{8}$, $D^{98\%}$ | $A^{9}$, P | — | $S^{2}$, $D^{90\%}$ | $S^{2}$, $D^{100\%}$ | $S^{2}$, $D^{85\%}$ | $S^{2}$, $D^{97\%}$ | $S^{3}$, $D^{98\%}$ |
| RI | $A^{13}$, P | $A^{8}$, $D^{92\%}$ | $A^{9}$, P | $A^{12}$, $D^{99\%}$ | $A^{14}$, P | $A^{15}$, P | $A^{9}$, P | $A^{6}$, $D^{94\%}$ | $A^{11}$, $D^{98\%}$ | $A^{13}$, P | — | $S^{1}$, $D^{95\%}$ | $S^{1}$, $D^{89\%}$ | $S^{1}$, $D^{95\%}$ | $S^{1}$, $D^{91\%}$ |
| RO | $A^{7}$, P | $A^{10}$, $D^{87\%}$ | $A^{12}$, P | $A^{13}$, P | $A^{16}$, P | $A^{13}$, P | $A^{7}$, P | $A^{12}$, $D^{92\%}$ | $A^{12}$, $D^{98\%}$ | $A^{13}$, P | $A^{17}$, P | — | $S^{1}$, P | $S^{1}$, $D^{96\%}$ | $S^{1}$, $D^{100\%}$ |
| SH | $A^{12}$, $D^{100\%}$ | $A^{6}$, $D^{86\%}$ | $A^{9}$, P | $A^{10}$, P | $A^{11}$, P | $A^{14}$, P | $A^{8}$, P | $A^{4}$, $Dr^{95\%}$ | $A^{8}$, $D^{96\%}$ | $A^{13}$, P | $A^{10}$, P | $A^{16}$, P | — | $S^{1}$, $D^{97\%}$ | $S^{1}$, $D^{100\%}$ |
| ST | $A^{12}$, P | $A^{9}$, $D^{83\%}$ | $A^{6}$, P | $A^{10}$, P | $A^{14}$, P | $A^{13}$, P | $A^{9}$, P | $A^{9}$, $Dr^{87\%}$ | $A^{9}$, $D^{94\%}$ | $A^{12}$, P | $A^{14}$, $D^{92\%}$ | $A^{14}$, P | $A^{10}$, $D^{90\%}$ | — | $S^{1}$, $D^{100\%}$ |
| TA | $A^{8}$, P | $A^{6}$, $D^{95\%}$ | $A^{8}$, P | $A^{8}$, $D^{97\%}$ | $A^{10}$, P | $A^{10}$, P | $A^{5}$, P | $A^{5}$, $Dr^{86\%}$ | $A^{10}$, $D^{97\%}$ | $A^{7}$, P | $A^{11}$, P | $A^{11}$, P | $A^{12}$, $D^{90\%}$ | $A^{12}$, P | — |

The lower triangle presents the number of fixed allozyme differences (shown as $A^{n}$), followed by a summary outcome of the multivariate analyses of the morphometric data (P = no taxon overlap in scatterplot based on the first two PCA dimensions; $D^{x\%}$ = percentage of individuals correctly allocated to taxon in a full DFA of all individuals; $Dr^{x\%}$ = percentage of individuals correctly allocated to taxon in a DFA of all individuals found within the same or abutting catchments. The upper triangle presents the number of diagnostic features evident among the 12 "supplementary morphological characters" (shown as $S^{n}$), followed by a summary outcome of the multivariate analyses of the meristic data (format as for the morphological data above).

[a]Percentage discrimination below threshold adopted for this study.

Not surprisingly, given the low levels of sequence divergence amongst haplotypes, the maximum-likelihood tree (not presented, but available from TreeBase; http://purl.org/phylo/treebase/phylows/study/TB2:S15464) was also not able to delineate species from first principles, nor did it provide any strong phylogenetic support for any node not also supported in the barcoding analysis.

Despite their failure to identify candidate species from first principles, both DNA barcoding and maximum-likelihood phylogenetic analysis nevertheless did provide some hindsight support for most of the nine candidate species represented by only a single site or river basin (AR, DA, KO, JI, RI, RO, SH, ST, and TA; Fig. 5). With one exception (KO), species-restricted haplotypes clustered together (albeit often without bootstrap support) to the exclusion of haplotypes found in other species, and most of these lineages were relatively basal in the overall tree. In addition, both barcoding and the maximum-likelihood tree also identified two quite distinctive mtDNA lineages in sympatry at several sites in the Hunter catchment (average *p*-distance between-lineages = 0.053 vs. average within-lineage values of 0.005 and 0.013; haplotypes identified in Fig. 7), for a selection of ethanol-only tissues (i.e., no companion allozyme or morphological data). This raises the possibility that yet another, geographically restricted taxon may be present in this catchment (location shown by the asterisk in Fig. 5) in addition to the widespread species OL (as found in other catchments for the narrow-range endemics DA, JI, KO, and TA).

## DISCUSSION

Several recent and prominent HMS assessments have concluded that a considerable proportion of the planet's species has already been "discovered" and named (often

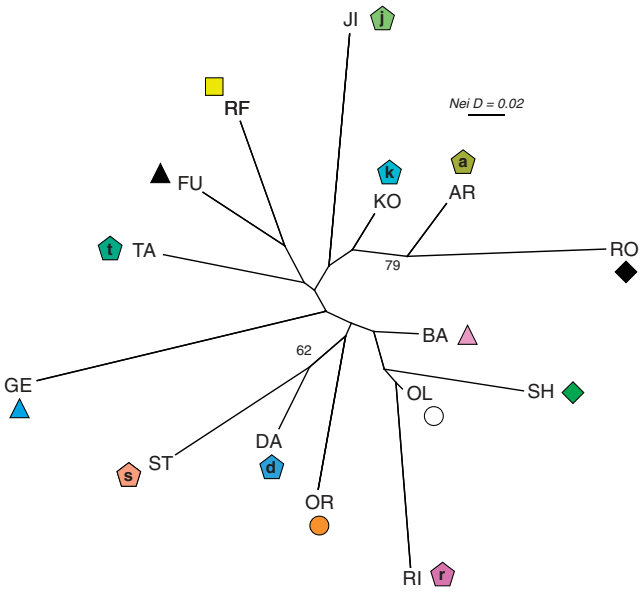


FIGURE 4. Neighbor Joining network, based on pairwise Nei's D-values for the allozyme data, among 15 diagnosable taxa in the mountain galaxias complex. Bootstrap values above 60% are shown.

multiple times; see later discussion on the issue of synonymy). This perspective is perhaps best illustrated by quoting some of these authors e.g., " …. recent analyses indicate that we may already have named one to two-thirds of all species." (Costello et al. 2013); "If the current trend [of species discovery] continues, most [marine] species will be discovered this century." (Appeltans et al. 2012); "We predict that there may be 1.8–2.0 million species on Earth, of which about 0.3 million are marine, significantly less than some previous estimates." (Costello et al. 2012). Importantly, these optimistic assessments of the extent of our taxonomic knowledge either completely ignore the issue of cryptic species or briefly present it as a minor encumbrance. However, when the real species counts for taxonomically well-studied "species" (based on comparisons across all eukaryotes, as per Fig. 1) in groups where our state of taxonomic knowledge is considered by most to be near-complete (e.g., an estimated ~78% of vertebrates and ~79% of fish have been described, Chapman 2009; Scheffers et al. 2012) can be off by at least 1500%, as is the case for mountain galaxias, there is a clear need to formally consider whether such hyper-cryptic species are a rare phenomenon or the tip of a taxonomic iceberg.

Our intensive molecular and morphological revision of *G. olidus* s.l. revealed it to be a hyper-cryptic complex of at least 15 genetically and morphologically diagnosable species, all valid under multiple species concepts. Moreover, the high proportion of narrow-range endemics discovered (9 of 15), suggests two obvious caveats concerning the levels of cryptic biodiversity detected. First, it is not unlikely that further intensive sampling in other peripheral catchments or in remote streams and headwaters of already-sampled catchments would turn up additional taxa (or would have in the past, before the introduction of salmonids, which are highly efficient predators of non-migratory, upland galaxiids; McDowall 2006). Indeed, mtDNA barcoding suggests this may be the case in the Hunter River (Figs. 5 and 7). Second, a less intensive field survey of ~150 sites (~43% of the number sampled herein but still a relatively intensive study) would likely have sampled <50% of these new species.

Both of the above caveats are important practical reminders of another key but under-addressed principle in the HMS debate, namely that a considerable proportion of species are rare, both in the wild (Scheffers et al. 2012) and as represented in museum collections (Lim et al. 2012). Therefore, the relationship between collecting effort and number of new species sampled must be highly non-linear, and indeed has been shown to be so (Lim et al. 2012). This simple reality alone is enough to demonstrate that HMS assessments ought not to rely heavily on the "estimating the asymptote" approach that has been at the core of many recent assessments. Moreover, our hypothesis that hyper-cryptic complexes are common throughout most organismal groups adds further weight to this conclusion. This is because the "degree of difficulty" in discovering new species is likely to increase as species counts approach an asymptote within any group. Once all the "easy to diagnose" morphospecies have been discovered, a higher proportion of the remaining undiscovered species will be, by definition, cryptic to *a priori* morphological diagnosis. Thus, diagnosing these species will require considerable extra effort to (1) comprehensively collect tissues and companion vouchers; (2) generate companion data sets of multi-locus genotypes; and (3) where additional candidate species are present, assess whether they are genuinely cryptic or only apparently so (as in the *G. olidus s.l.* complex).

Many HMS assessments have rightly discussed the problem of synonymy that is, where two or more taxonomic names have been applied to a single species. Typical estimates for the extent of synonymy are >20% across all species (Costello et al. 2012) and up to 70% in some organismal groups (Scheffers et al. 2012). While we have no quarrel with such assessments nor with the rationale *per se*, we believe that future estimates should also consider the flip side of synonymy, as revealed by our data for *G. olidus* s.l. Here, eight taxonomic names were synonymized under the one species by an expert taxonomic assessment (McDowall and Frankenberg 1981), an appropriate decision when only morphological data were available, but inappropriate with the benefit of genetic and morphological hindsight. The clear message is that just because "species" have been or should be synonymized based on morphological criteria alone, does not guarantee that any such species will not prove to be a cryptic or hyper-cryptic complex in the future. Indeed, we suggest that such "species" may even be a sensible starting point to search for cryptic biodiversity in many "higher" organisms (e.g., Welton et al. 2013).

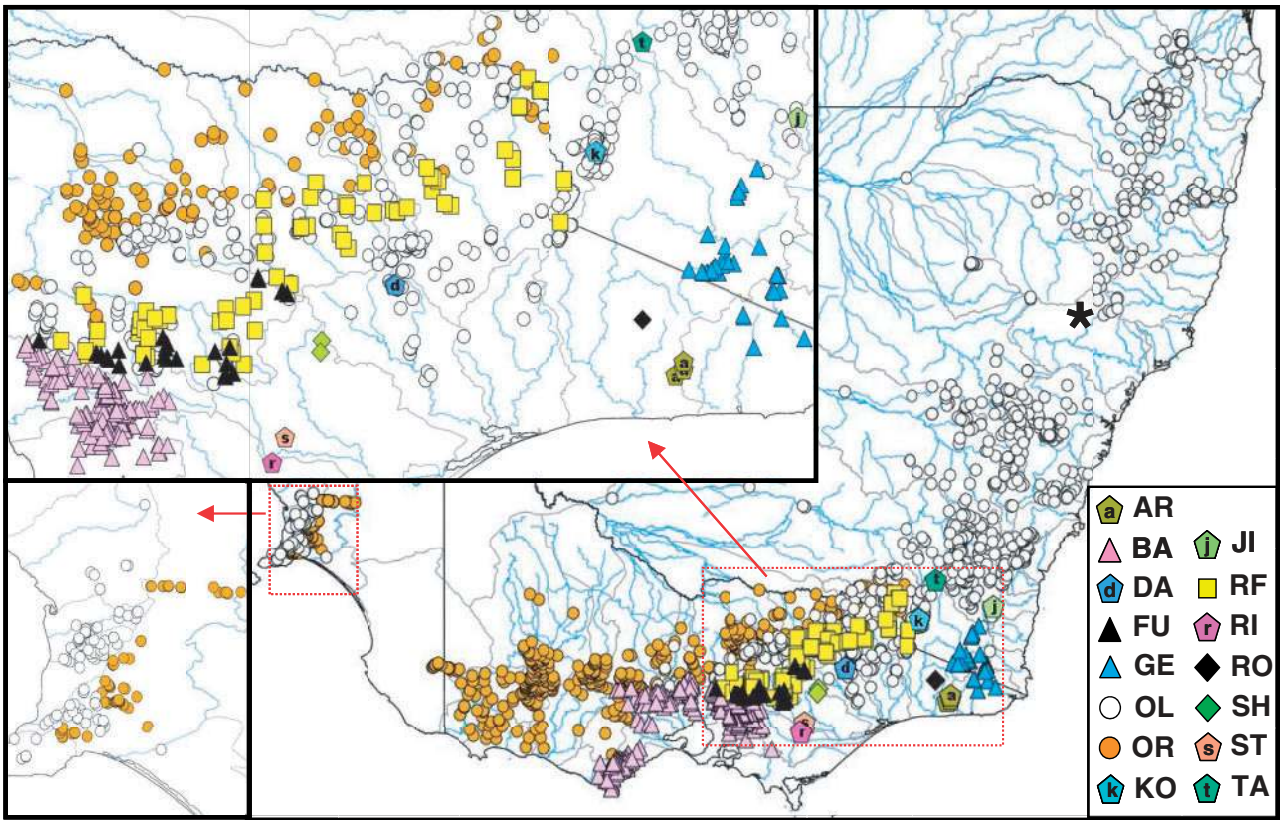


FIGURE 5. Map showing the geographic distribution of the 15 candidate species ultimately identified in this study. As shown, taxon codes and symbols follow Table 1 and Figure 3, respectively. The Hunter River catchment is identified by an asterisk.

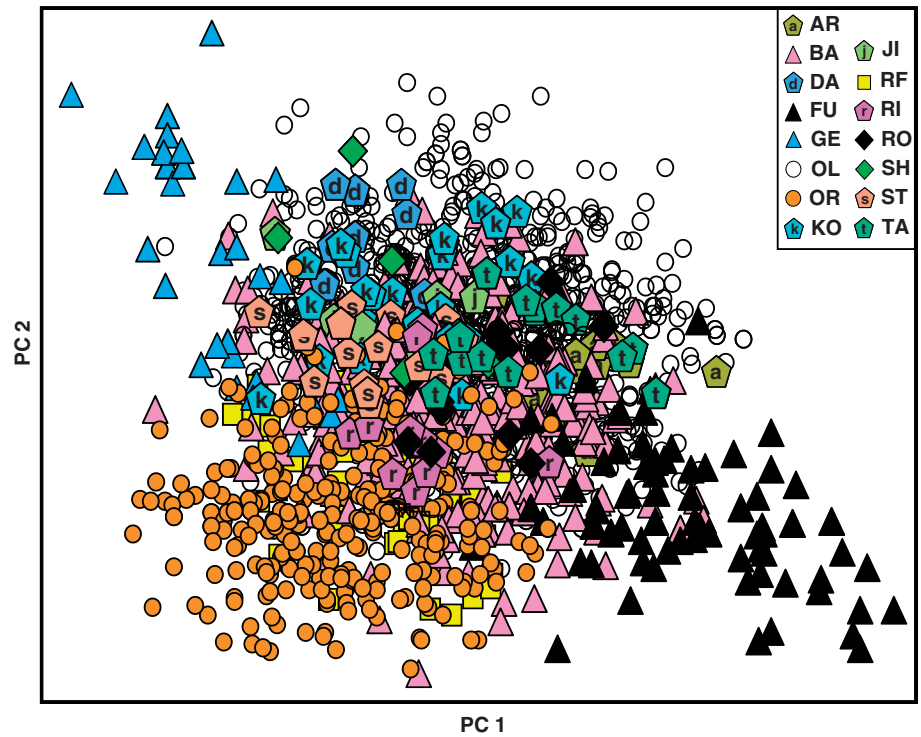


FIGURE 6. Scatterplot of factor scores of first (PC1) and second (PC2) principal components of an initial PCA of 1963 individuals from all 15 allozyme-defined candidate taxa, based on the morphometric data. PC1 and PC2 explained 29.1% and 17.0%, respectively, of the total variance. Taxon symbols follow Figure 3.

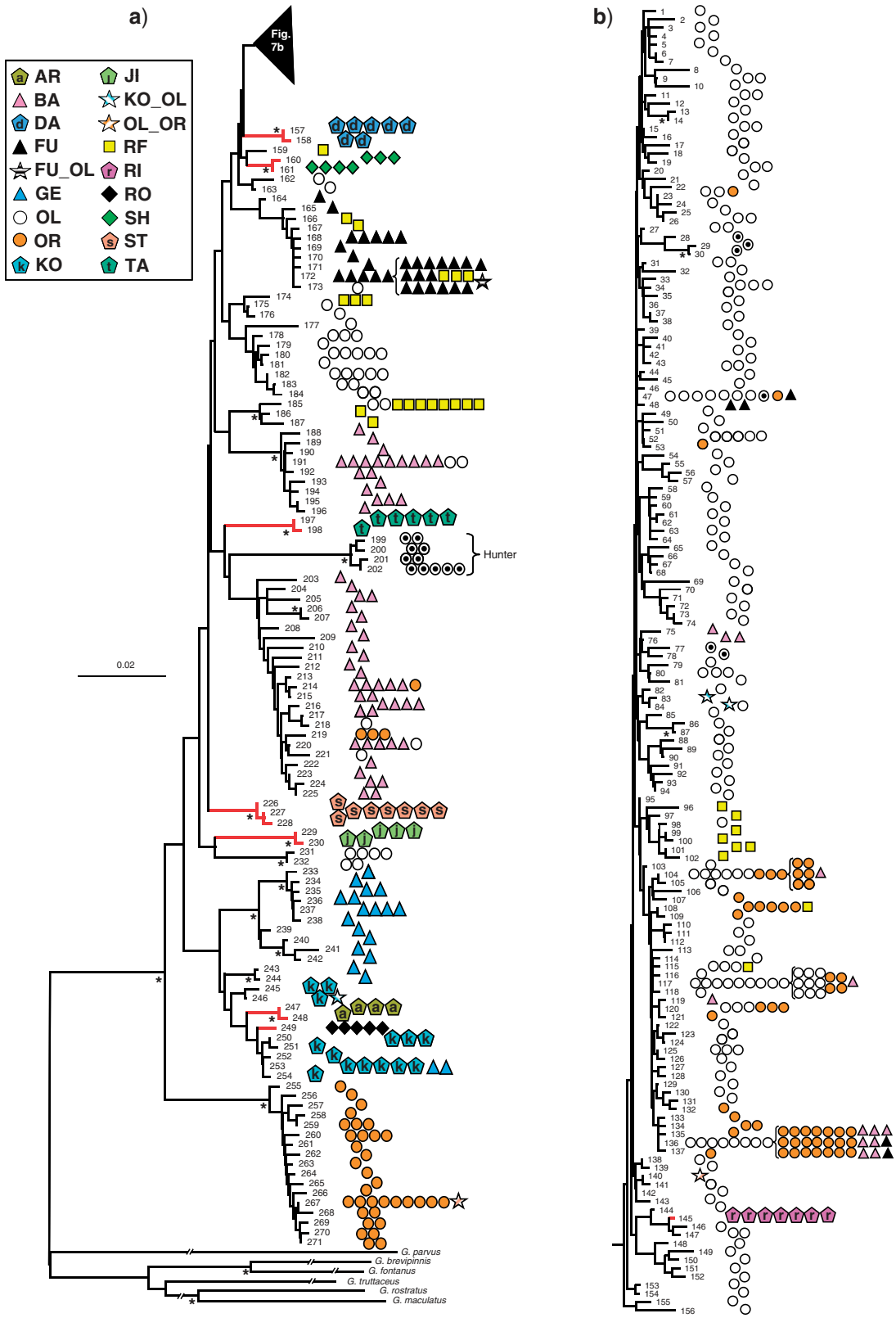


FIGURE 7. Neighbor Joining tree, based on *p*-distances and rooted using six outgroup species of *Galaxias*, for 271 cyt*b* haplotypes for the *G. olidus* s.l. complex. Haplotypes are numbered sequentially 1–271, and each is also labeled by symbols indicating the number of individuals displaying that haplotype and their known (for individuals in the allozyme study) or inferred (based on morphology) taxon assignment. Nodes supported by bootstrap values above 90% are asterisked. Lineages that only contain exemplars of a single candidate species are drawn using thicker lines. The symbol for the OL haplotypes from the Hunter River have a small solid circle at their center.

Our detailed revision of mountain galaxias also highlights some of the reasons why documenting cryptic biodiversity has major implications beyond the purview of global biodiversity assessment. Prior to this study, mountain galaxias were thought by most researchers to comprise one narrowly distributed, endangered, and alien trout-impacted species (*G. fuscus*; herein species FU) and one very widespread (and therefore presumably vagile) and resilient (i.e., able to subsist alongside trout) species (all other populations; Allen et al. 2002). Instead, the true pattern is almost the exact opposite. The majority of the "new" species occur as narrow-range endemics, and nine of these (AR, DA, JI, KO, RI, RO, SH, ST, and TA) are considered more vulnerable than *G. fuscus* (Raadik 2011; DSE 2013), as each occurs only at 1–2 locations and all are trout-impacted, to the point where wildlife managers are now actively attempting to lessen the threat of extinction (Raadik 2013). Furthermore, only four species (BA, FU, OL, and OR) show any genetic evidence of contemporary dispersal across catchments (Fig. 7; allozyme analyses not shown), and only OL and OR appear to fit the ecological profile previously attributed to the widespread "species" (Raadik 2011). Examples of how perspectives change following the realization that a presumed widespread species is actually a hyper-cryptic complex are known in many applied disciplines, such as conservation biology (Murphy et al. 2013), environmental monitoring (Sharley et al. 2004), public health (White et al. 2011), pest management (Gauthier 2010), and applied ecology (Georgieva et al. 2013). Many of these changes in perspective also have obvious conservation implications for any new, narrow-range species previously considered part of a hyper-cryptic complex, beginning with the obvious need to assess its conservation status (Murphy et al. 2013).

An important question to ask here is how common are hyper-cryptic species? Clearly, we cannot know until molecular frameworks involving multi-locus nDNA markers are available for a wide range of groups. However, even a superficial glance at the literature reveals definitive or likely instances of hyper-cryptic species across almost all organismal groups e.g., mammals (Aplin et al. 2011; Nicolas et al. 2012), reptiles (Oliver et al. 2009; Brown et al. 2012), amphibians (Fouquet et al. 2007; Funk et al. 2012), other fish (Niemiller et al. 2011; Cooke et al. 2012), including other galaxiids (Wishart et al. 2006), echinoderms (O'Loughlin et al. 2011), arthropods (Alquezar et al. 2010; Porco et al. 2012; Lee et al. 2013; Miller et al. 2013), molluscs (Vrijenhoek 2009; Puillandre et al. 2012), onychophorans (Ruhberg and Daniels 2013), helminths (Baverstock et al. 1985; Poulin 2011), rotifers (Leasi et al. 2013), fungi (O'Brien et al. 2005; Smith et al. 2011; Hawksworth 2012), plants (Grundt et al. 2006; Jolles and Wolfe 2012; Medina et al. 2012), photosynthetic protists (Fraser et al. 2013), other eukaryotic protists ((Adams et al., 1989); (Xu et al., 2012); (Tarcz et al., 2013)), and bacteria (Koeppel et al. 2008; Clermont et al. 2011; Serkebaeva et al. 2013). Such ubiquity across groups further argues the need for future HMS assessments to consider this phenomenon as part of formally estimating levels of overall cryptic biodiversity. Most importantly, this need exists independently of whatever threshold is used to define a complex as hyper-cryptic.

While all candidate species in the mountain galaxias complex are obviously closely related species that diverged relatively recently in evolutionary time (Fig. 7), it does not necessarily follow that most hyper-cryptic complexes will follow this pattern. As just one of many vertebrate examples, the hyper-cryptic Australian gecko *Crenadactylus ocellatus* has recently been shown to comprise at least 12 candidate species (with additions likely in the biodiverse north-western region), many of which diverged in the mid- to late-Miocene (Oliver et al. 2010; Oliver et al. 2012). This alternate scenario for the evolutionary relatedness of hyper-cryptic congeners is likely to be even more common among the "less specialized" eukaryotes and among prokaryotes. For example, preliminary molecular studies of *Trichoplax adhaerens*, the only described species in the supposedly monotypic phylum Placozoa, have concluded it may comprise as many as ~200 morphologically indistinguishable species, some displaying levels of genetic divergence comparable to those found between different families within other basal metazoan phyla (Eitel et al. 2013).

It is becoming increasingly evident that single-gene barcoding cannot serve as a stand-alone panacea for taxonomic discovery (Dupuis et al. 2012; Taylor and Harris 2012), let alone define candidate species in hyper-cryptic complexes involving closely related species plus sporadic hybridization and introgression (as appears to have occurred in the *G. olidus* complex; Figs. 3 and 7). Thus all future detailed exploration of such complexes will require multi-locus nuclear genetic data, and increasingly those generated by various next-generation DNA technologies (Harrison and Kidner 2011; Lemmon and Lemmon 2013). As inferred by our study, such data sets may need to encompass dozens or even hundreds of independent genetic markers to accurately diagnose all sampled species. Here we surveyed >100 potential taxonomic characters in *G. olidus* s.l. to find a sufficient number of concordant diagnostic characters to define species, equating to success rates of ~4–25% per character, depending on the species pair being diagnosed. Fortunately, and unlike the more "traditional" data sets presented herein, such next-generation DNA data sets will likely be able to both diagnose species and robustly assess their true phylogenetic relationships (McCormack et al. 2013).

### *A Different Perspective on "Where are all the new species?"*

Of course there are many uncataloged species that have never been collected, perhaps because they live deep in the benthos, high atop a rainforest tree, or (perhaps just as likely) survive as a rare, range-restricted species in a more conventional habitat (Stork et al.

2008). However, the stark and unappealing reality is that a good proportion of them are either in museum collections awaiting companion molecular data or have yet to be vouchered because they have been assigned by default to existing, already-cataloged species. Both such scenarios were realities for *G. olidus* s.l. prior to this study and both remain a possibility for this complex into the future. Given that discovering "new" species under either scenario requires further targeted collecting of specimens and companion tissues for molecular analysis, it is imperative that field researchers continue to collect and lodge such material with an appropriate museum, even for groups thought to be blessed with taxonomic stability.

## Conclusions

In addition to reinforcing our contention that all future HMS assessments should explicitly factor in estimates of cryptic and hyper-cryptic biodiversity, we believe the present study offers up six important lessons for the overall debate. First, even comparatively well-studied species in relatively charismatic groups can be hyper-cryptic complexes. Second, cryptic biodiversity is theoretically possible for any species that has been defined *solely* on morphological criteria. Third, nuclear genetic data sets, reflecting multiple "goldilocks" genes (i.e., not too variable, not too conservative) and codominant alleles, are required to truly assess whether a species is likely to harbor cryptic taxa. Fourth, while single-gene barcoding may assist in this process, it can never substitute for it. Fifth, additional genetic markers are required when a hyper-cryptic complex involves sibling species (i.e., sister species, *sensu* Bickford et al. 2007), particularly if there has been some hybridization and introgression. Finally, the ultimate and crucial taxonomic outcome of diagnosing candidate species (or concluding that they cannot be diagnosed morphologically) requires considerable intensive effort in two key traditional and labor-intensive components of taxonomic endeavor that is, targeted collecting of vouchers, and detailed morphological analyses; in other words, *there are no short-cuts to reach the final taxonomic outcome*. While a daunting prospect to some, we hope that others will be inspired by the magnitude and intellectual challenge of the task ahead.

## Supplementary Material

Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.tk043.

## Funding

This work was supported by the University of Canberra; the Co-operative Research Centre for Freshwater Ecology (Program C220); and the South Australian Museum.

## Acknowledgments

Numerous people assisted with one or more aspects of this study, including field collection and photography of specimens, acquisition of genetic data, data analysis, and assistance with preparing the manuscript. We thank Erika Alacs, Lachlan Farrington, Anika Gardner, Martin Gomon, Michael Hammer, Peta Hill, Marion Hoen, Rudi Kuiter, Emily Moriarty Lemmon, Ken Sharpe, Peter Unmack, and Rachel Walsh. We also thank the editor, associate editor, and various referees for their helpful comments. All fish were collected under permits issued by the relevant State authorities, and the research was conducted in accordance with animal ethics regulations and approvals.

## References

Adams M., Andrews R.H., Robinson B., Christy P., Baverstock P.R., Dobson P.J., Blackler S.J. 1989. A genetic approach to species criteria in the amoeba genus *Naegleria* using allozyme electrophoresis. Int. J. Parasitol. 19:823–834.

Adams M., Page T.J., Hurwood D.A., Hughes J.M. 2013. A molecular assessment of species boundaries and phylogenetic affinities in *Mogurnda* (Eleotridae): a case study of cryptic biodiversity in the Australian freshwater fishes. Mar. Freshw. Res. 64:920–931.

Adams M., Wedderburn S.D., Unmack P.J., Hammer M.P., Johnson J.B. 2011. Use of congeneric assessment to reveal the linked genetic histories of two threatened fishes in the Murray-Darling Basin, Australia. Conserv. Biol. 25:767–776.

Allen G.R. 1989. Freshwater fishes of Australia. Neptune City (NJ): T. F. H. Publications.

Allen G.R., Midgeley S.H., Allen M. 2002. Field guide to the freshwater fishes of Australia. Perth: Western Australian Museum.

Alquezar D.E., Hemmerter S., Cooper R.D., Beebe N.W. 2010. Incomplete concerted evolution and reproductive isolation at the rDNA locus uncovers nine cryptic species within *Anopheles longirostris* from Papua New Guinea. BMC Evol. Biol. 10:392.

Aplin K.P., Suzuki H., Chinen A.A., Chesser R.T., ten Have J., Donnellan S.C., Austin J., Frost A., Gonzalez J.P., Herbreteau V., Catzeflis F., Soubrier J., Fang Y.-P., Robins J., Matisoo-Smith E., Bastos A.D.S., Maryanto I., Sinaga M.H., Denys C., Van Den Bussche R.A., Conroy C., Rowe K., Cooper A. 2011. Multiple geographic origins of commensalism and complex dispersal history of black rats. PLoS ONE 6:e26357.

Appeltans W., Ahyong S.T., Anderson G., Angel M.V., Artois T., Bailly N., Bamber R., Barber A., Bartsch I., Berta A., Błażewicz-Paszkowycz M., Bock P., Boxshall G., Boyko C.B., Brandão S.N., Bray R.A., Bruce N.L., Cairns S.D., Chan T.-Y., Cheng L., Collins A.G., Cribb T., Curini-Galletti M., Dahdouh-Guebas F., Davie P.J.F., Dawson M.N., De Clerck O., Decock W., De Grave S., de Voogd N.J., Domning D.P., Emig C.C., Erséus C., Eschmeyer W., Fauchald K., Fautin D.G., Feist S.W., Fransen C.H.J.M., Furuya H., Garcia-Alvarez O., Gerken S., Gibson D., Gittenberger A., Gofas S., Gómez-Daglio L., Gordon D.P., Guiry M.D., Hernandez F., Hoeksema B.W., Hopcroft R.R., Jaume D., Kirk P., Koedam N., Koenemann S., Kolb J.B., Kristensen R.M., Kroh A., Lambert G., Lazarus D.B., Lemaitre R., Longshaw M., Lowry J., Macpherson E., Madin L.P., Mah C., Mapstone G., McLaughlin P.A., Mees J., Meland K., Messing C.G., Mills C.E., Molodtsova T.N., Mooi R., Neuhaus B., Ng Peter K.L., Nielsen C., Norenburg J., Opresko D.M., Osawa M., Paulay G., Perrin W., Pilger John F., Poore G.C.B., Pugh P., Read G.B., Reimer J.D., Rius M., Rocha R.M., Saiz-Salinas J.I., Scarabino V., Schierwater B., Schmidt-Rhaesa A., Schnabel K.E., Schotte M., Schuchert P., Schwabe E., Segers H., Self-Sullivan C., Shenkar N., Siegel V., Sterrer W., Stöhr S., Swalla B., Tasker M.L., Thuesen Erik V., Timm T., Todaro M.A., Turon X., Tyler S., Uetz P., van der Land J., Vanhoorne B., van Ofwegen L.P., van Soest R.W.M., Vanaverbeke J., Walker-Smith G.,

Walter T.C., Warren A., Williams G.C., Wilson S.P., Costello M.J. 2012. The magnitude of global marine species diversity. Curr. Biol. 22:2189–2202.
Avise J.C. 1975. Systematic value of electrophoretic data. Syst. Zool. 23:465–481.
AWRC. 1976. Review of Australia's water resources. Canberra: Australian Water Resources Council.
Basset Y., Cizek L., Cuenoud P., Didham R.K., Guilhaumon F., Missa O., Novotny V., Odegaard F., Roslin T., Schmidl J., Tishechkin A.K., Winchester N.N., Roubik D.W., Aberlenc H.-P., Bail J., Barrios H., Bridle J.R., Castano-Meneses G., Corbara B., Curletti G., da Rocha W.D., de Bakker D., Delabie J.H.C., Dejean A., Fagan L.L., Floren A., Kitching R.L., Medianero E., Miller S.E., de Oliveira E.G., Orivel J., Pollet M., Rapp M., Ribeiro S.P., Roisin Y., Schmidt J.B., Sorensen L., Leponce M. 2012. Arthropod diversity in a tropical forest. Science 338:1481–1484.
Baverstock P.R., Adams M., Beveridge I. 1985. Biochemical differentiation in bile duct cestodes and their marsupial hosts. Mol. Biol. Evol. 2:321–337.
Beheregaray L.B., Caccone A. 2007. Cryptic biodiversity in a changing world. J. Biol. 6:9.
Bertozzi T., Adams M., Walker K.F. 2000. Species boundaries in carp gudgeons (Eleotrididae: *Hypseleotris*) from the River Murray, South Australia: evidence for multiple species and extensive hybridization. Mar. Freshw. Res. 51:805–815.
Bickford D., Lohman D.J., Sodhi N.S., Ng P.K.L., Meier R., Winker K., Ingram K.K., Das I. 2007. Cryptic species as a window on diversity and conservation. Trends Ecol. Evol. 22:148–155.
Brown R.M., Siler C.D., Grismer L.L., Das I., McGuire J.A. 2012. Phylogeny and cryptic diversification in Southeast Asian flying geckos. Molec. Phylogenet. Evol. 65:351–361.
Burridge C.P., McDowall R.M., Craw D., Wilson M.V.H., Waters J.M. 2012. Marine dispersal as a pre-requisite for Gondwanan vicariance among elements of the galaxiid fish fauna. J. Biogeogr. 39:306–321.
Chapman A.D. 2009. Numbers of living species in Australia and the world. 2nd ed. Canberra: Australian Biological Resources Study.
Clermont O., Gordon D.M., Brisse S., Walk S.T., Denamur E. 2011. Characterization of the cryptic *Escherichia* lineages: rapid identification and prevalence. Environ. Microbiol. 13:2468–2477.
Cooke G.M., Chao N.L., Beheregaray L.B. 2012. Five cryptic species in the Amazonian catfish *Centromochlus existimatus* identified based on biogeographic predictions and genetic data. PLoS ONE 7:e48800.
Costello M.J., May R.M., Stork N.E. 2013. Can we name Earth's species before they go extinct? Science 339:413–416.
Costello M.J., Wilson S., Houlding B. 2012. Predicting total global species richness using rates of species description and estimates of taxonomic effort. Syst. Biol. 61:871–883.
Curtis T.P., Sloan W.T., Scannell J.W. 2002. Estimating prokaryotic diversity and its limits. Proc. Natl Acad. Sci. USA 99:10494–10499.
de Queiroz K. 2007. Species concepts and species delimitation. Syst. Biol. 56:879–886.
Dolphin K., Quicke D.L.J. 2001. Estimating the global species richness of an incompletely described taxon: an example using parasitoid wasps (Hymenoptera: Braconidae). Biol. J. Linn. Soc. 73:279–286.
DSE. 2013. Advisory list of threatened vertebrate fauna in Victoria. Melbourne: Department of Sustainability and Environment.
Dupuis J.R., Roe A.D., Sperling F.A.H. 2012. Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. Mol. Ecol. 21:4422–4436.
Eitel M., Osigus H.-J., DeSalle R., Schierwater B. 2013. Global diversity of the Placozoa. PLoS ONE 8:e57131.
Emerson B.C., Cicconardi F., Fanciulli P.P., Shaw P.J.A. 2011. Phylogeny, phylogeography, phylobetadiversity and the molecular analysis of biological communities. Philos. Trans. R. Soc. Lond. B Biol. Sci. 366:2391–2402.
Erwin T.L. 1982. Tropical forests: their richness in Coleoptera and other arthropod species. Coleopt. Bull. 36:74–75.
Erwin T.L. 1991. How many species are there?: revisited. Conserv. Biol. 5:330–333.
Fisher R.A. 1936. The use of multiple measurements in taxonomic problems. Ann. Eugen. 7:179–188.
Fouquet A., Vences M., Salducci M.D., Meyer A., Marty C., Blanc M., Gilles A. 2007. Revealing cryptic diversity using molecular phylogenetics and phylogeography in frogs of the *Scinax ruber* and *Rhinella margaritifera* species groups. Molec. Phylogenet. Evol. 43:567–582.
Frankham R., Ballou J.D., Briscoe D.A. 2010. Introduction to conservation genetics. 2nd ed. Cambridge, UK: Cambridge University Press.
Frankham R., Ballou J.D., Dudash M.R., Eldridge M.D.B., Fenster C.B., Lacy R.C., Mendelson J.R. 3rd, Porton I.J., Ralls K., Ryder O.A. 2012. Implications of different species concepts for conserving biodiversity. Biol. Conserv. 153:25–31.
Fraser C.I., Zuccarello G.C., Spencer H.G., Salvatore L.C., Garcia G.R., Waters J.M. 2013. Genetic affinities between trans-oceanic populations of non-buoyant macroalgae in the high latitudes of the southern hemisphere. PLoS ONE 8: e69138.
Funk W.C., Caminer M., Ron S.R. 2012. High levels of cryptic species diversity uncovered in Amazonian frogs. Proc. R. Soc. Lond. B Biol. Sci. 279:1806–1814.
Gaston K.J. 1991. The magnitude of global insect species richness. Conserv. Biol. 5:283–296.
Gauthier N. 2010. Multiple cryptic genetic units in *Hypothenemus hampei* (Coleoptera: Scolytinae): evidence from microsatellite and mitochondrial DNA sequence data. Biol. J. Linn. Soc. 101:113–129.
Georges A., Adams M. 1992. A phylogeny for Australian chelid turtles based on allozyme electrophoresis. Aust. J. Zool. 40:453–476.
Georgieva S., Soldanova M., Perez-del-Olmo A., Dangel D.R., Sitko J., Sures B., Kostadinova A. 2013. Molecular prospecting for European *Diplostomum* (Digenea: Diplostomidae) reveals cryptic diversity. Int. J. Parasitol. 43:57–72.
Gower J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53:325–338.
Grassle J.F., Maciolek N.J. 1992. Deep-sea species richness: regional and local diversity estimates from quantitative bottom samples. Am. Nat. 139:313–341.
Grundt H.H., Kjolner S., Borgen L., Rieseberg L.H., Brochmann C. 2006. High biological species diversity in the arctic flora. Proc. Natl Acad. Sci. USA 103:972–975.
Hammer M.P., Adams M., Hughes J.M. 2013. Evolutionary processes and biodiversity. In: Walker K., Humphreys P., editors. Ecology of Australian freshwater fishes. Collingwood: CSIRO Publishing. p. 49–79.
Hammer M.P., Adams M., Unmack P.J., Walker K.F. 2007. A rethink on *Retropinna*: conservation implications of new taxa and significant genetic substructure in Australian smelts (Pisces: Retropinnidae). Mar. Freshw. Res. 58:327–341.
Hamilton A.J., Basset Y., Benke K.K., Grimbacher P.S., Miller S.E., Novotny V., Samuelson G.A., Stork N.E., Weiblen G.D., Yen J.D.L. 2010. Quantifying uncertainty in estimation of tropical arthropod species richness. Am. Nat. 176:90–95.
Harrison N., Kidner C.A. 2011. Next-generation sequencing and systematics: what can a billion base pairs of DNA sequence data do for you? Taxon 60:1552–1566.
Hawksworth D.L. 2012. Global species numbers of fungi: are tropical studies and molecular approaches contributing to a more robust estimate? Biodivers. Conserv. 21:2425–2433.
Hoese D.F., Bray D.J., Paxton J.R., Allen G.R. 2006. Fishes. In: Beesley P.L., Wells A., editors. Zoological catalogue of Australia. Vol. 35. Australia: ABRS and CSIRO Publishing. Part 1, p. xxiv 1–670.
Horner P., Adams M. 2007. A molecular systematic assessment of species boundaries in Australian *Cryptoblepharus* (Reptilia: Squamata: Scincidae) – a case study for the combined use of allozymes and morphology to explore cryptic biodiversity. Beagle, Suppl. 3:1–19.
Humphries P., Walker K.F., editors. 2013. Ecology of Australian freshwater fishes. Collingwood: CSIRO Publishing.
Isaac N.J.B., Mallet J., Mace G.M. 2004. Taxonomic inflation: its influence on macroecology and conservation. Trends Ecol. Evol. 19:464–469.
Jolles D.D., Wolfe A.D. 2012. Genetic differentiation and crypsis among members of the myco-heterotrophic *Pyrola picta* species complex (Pyroleae: Monotropoideae: Ericaceae). Syst. Bot. 37:468–477.
Koeppel A., Perry E.B., Sikorski J., Krizanc D., Warner A., Ward D.M., Rooney A.P., Brambilla E., Connor N., Ratcliff R.M., Nevo E., Cohan F.M. 2008. Identifying the fundamental units of bacterial

diversity: a paradigm shift to incorporate ecology into bacterial systematics. Proc. Natl Acad. Sci. USA 105:2504–2509.
Laursen L., Bekoff M. 1978. *Loxodonta africana*. Mamm. Species 92:1–8.
Leasi F., Tang C.Q., De Smet W.H., Fontaneto D. 2013. Cryptic diversity with wide salinity tolerance in the putative euryhaline *Testudinella clypeata* (Rotifera, Monogononta). Zool. J. Linn. Soc. 168:17–28.
Lee W., Park J., Lee G.-S., Lee S., Akimoto S. 2013. Taxonomic status of the *Bemisia tabaci* complex (Hemiptera: Aleyrodidae) and reassessment of the number of its constituent species. PLoS ONE 8:e63817.
Lemmon E.M., Lemmon A.R. 2013. High-throughput genomic data in systematics and phylogenetics. Annu. Rev. Ecol. Evol. Syst. 44:99–121.
Lim G.S., Balke M., Meier R. 2012. Determining species boundaries in a world full of rarity: singletons, species delimitation methods. Syst. Biol. 61:165–169.
May R.M. 1988. How many species are there on Earth? Science 241:1441–1449.
McCormack J.E., Hird S.M., Zellmer A.J., Carstens B.C., Brumfield R.T. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Molec. Phylogenet. Evol. 66:526–538.
McDowall R.M. 2006. Crying wolf, crying foul, or crying shame: alien salmonids and a biodiversity crisis in the southern cool-temperate galaxioid fishes? Rev. Fish Biol. Fish. 16:233–422.
McDowall R.M., Frankenberg R.S. 1981. The galaxiid fishes of Australia. Rec. Aust. Mus. 33:443–605.
Medina R., Lara F., Goffinet B., Garilleti R., Mazimpaka V. 2012. Integrative taxonomy successfully resolves the pseudo-cryptic complex of the disjunct epiphytic moss *Orthotrichum consimile* s.l. (Orthotrichaceae). Taxon 6:1180–1198.
Miller A.D., Skoracka A., Navia D., de Mendonca R.S., Szydło W., Schultz M.B., Michael Smith C., Truol G., Hoffmann A.A. 2013. Phylogenetic analyses reveal extensive cryptic speciation and host specialization in an economically important mite taxon. Molec. Phylogenet. Evol. 66:928–940.
Mora C., Tittensor D.P., Adl S., Simpson A.G.B., Worm B. 2011. How many species are there on Earth and in the ocean? PLoS Biol. 9:e1001127.
Murphy N.P., Adams M., Guzik M.T., Austin A.D. 2013. Extraordinary micro-endemism in Australian desert spring amphipods. Molec. Phylogenet. Evol. 66:645–653.
Nicolas V., Herbreteau V., Couloux A., Keovichit K., Douangboupha B., Hugot J.-P. 2012. A remarkable case of micro-endemism in *Laonastes aenigmamus* (Diatomyidae, Rodentia) revealed by nuclear and mitochondrial DNA sequence data. PLoS ONE 7:e48145.
Niemiller M.L., Near T.J., Fitzpatrick B.M. 2011. Delimiting species using multilocus data: diagnosing cryptic diversity in the southern cavefish, *Typhlichthys subterraneus* (Teleostei: Amblyopsidae). Evolution 66:846–866.
O'Brien H.E., Parrent J.L., Jackson J.A., Moncalvo J.-M., Vilgalys R. 2005. Fungal community analysis by large-scale sequencing of environmental samples. Appl. Environ. Microbiol. 71:5544–5550.
O'Loughlin P.M., Paulay G., Davey N., Michonneau F. 2011. The Antarctic region as a marine biodiversity hotspot for echinoderms: diversity and diversification of sea cucumbers. Deep-Sea Res. Part II – Topical Stud. Oceanogr. 58:264–275.
ØDegaard F. 2000. How many species of arthropods? Erwin's estimate revised. Biol. J. Linn. Soc. 71:583–597.
Oliver P.M., Adams M., Doughty P. 2010. Molecular evidence for ten species and Oligo-Miocene vicariance within a nominal Australian gecko species (*Crenadactylus ocellatus*, Diplodactylidae). BMC Evol. Biol. 10:386.
Oliver P.M., Adams M., Lee M.S.Y., Hutchinson M.N., Doughty P. 2009. Cryptic diversity in vertebrates: molecular data double estimates of species diversity in a radiation of Australian lizards (*Diplodactylus*, Gekkota). Proc. R. Soc. Lond. B Biol. Sci. 276:2001–2007.
Oliver P.M., Doughty P., Palmer R. 2012. Hidden biodiversity in rare northern Australian vertebrates: the case of the clawless geckos (*Crenadactylus*, Diplodactylidae) of the Kimberley. Wildl. Res. 39:429–435.
Pawlowski J., Audic S., Adl S., Bass D., Belbahri L., Berney C., Bowser S.S., Cepicka I., Decelle J., Dunthorn M., Fiore-Donno A.M., Gile G.H., Holzmann M., Jahn R., Jirků M., Keeling P.J., Kostka M., Kudryavtsev A., Lara E., Lukeš J., Mann D.G., Mitchell E.A.D., Nitsche F., Romeralo M., Saunders G.W., Simpson A.G.B., Smirnov A.V., Spouge J.L., Stern R.F., Stoeck T., Zimmermann J., Schindel D., de Vargas C. 2012. CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. PLoS Biol. 10:e1001419.
Pearson K. 1901. On lines and planes of closest fit to systems of points in space. Philos. Mag. 2:559–572.
Pfenninger M., Schwenk K. 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. BMC Evol. Biol. 7:121.
Platnick N.I. 1999. Dimensions of biodiversity: targeting megadiverse groups. In: Cracraft J., Grifo F., editors. The living planet in crisis: biodiversity science and policy. New York: Columbia University Press. p. 33–52.
Porco D., Bedos A., Greenslade P., Janion C., Skarżyński D., Stevens M.I., Jansen van Vuuren B., Deharveng L. 2012. Challenging species delimitation in Collembola: cryptic diversity among common springtails unveiled by DNA barcoding. Invertebr. Syst. 26:470–477.
Poulin R. 2011. Uneven distribution of cryptic diversity among higher taxa of parasitic worms. Biol. Lett. 7:241–244.
Puillandre N., Modica M.V., Zhang Y., Sirovich L., Boisselier M.C., Cruaud C., Holford M., Samadi S. 2012. Large-scale species delimitation method for hyperdiverse groups. Molec. Ecol. 21:2671–2691.
R Development Core Team. 2009. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
Raadik T.A. 2011. Systematic revision of the mountain galaxias, *Galaxias olidus* Günther, 1866 species complex (Teleostei: Galaxiidae) in eastern Australia. [Ph.D. thesis], Canberra: University of Canberra. Available from: URL http:www.canberra.edu.au/researchrepository/items/eee1b111-513d-f98a-1363-5b4071b784cb/1; last accessed: 10/6/2013.
Raadik T.A. 2013. Fact file: conserving the threatened Shaw Galaxias. Critical conservation actions in the Alpine National Park. Melbourne: Parks Victoria.
Raadik T.A., Saddlier S.R., Koehn J.D. 1996. Threatened fishes of the world: *Galaxias fuscus* Mack, 1936 (Galaxiidae). Environ. Biol. Fishes 47:108.
Richardson B.J., Baverstock P.R., Adams M. 1986. Allozyme electrophoresis: a handbook for animal systematics and population studies. Sydney: Academic Press.
Ruhberg H., Daniels S.R. 2013. Morphological assessment supports the recognition of four novel species in the widely distributed velvet worm *Peripatopsis moseleyi* sensu lato (Onychophora: Peripatopsidae). Invertebr. Syst. 27:131–145.
Sambrook J., Russell D.W. 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
Scheffers B.R., Joppa L.N., Pimm S.L., Laurance W.F. 2012. What we know and don't know about Earth's missing biodiversity. Trends Ecol. Evol. 27:501–510.
Serkebaeva Y.M., Kim Y., Liesack W., Dedysh S.N. 2013. Pyrosequencing-based assessment of the *Bacteria* diversity in surface and subsurface peat layers of a northern wetland, with focus on poorly studied phyla and candidate divisions. PLoS ONE 8:e63994.
Sharley D.J., Pettigrove V., Parsons Y.M. 2004. Molecular identification of *Chironomus* spp. (Diptera) for biomonitoring of aquatic ecosystems. Aust. J. Entomol. 43:359–365.
Shoshani J., Eisenberg J.F. 1982. *Elephas maximus*. Mamm. Species 182:1–8.
Smith M.E, Henkel T.W., Aime M.C., Fremier A.K., Vilgalys R. 2011. Ectomycorrhizal fungal diversity and community structure on three co-occurring leguminous canopy tree species in a Neotropical rainforest. New Phytol. 192:699–712.
Song S., Liu L., Edwards S.V., Wu S. 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proc. Natl Acad. Sci. USA 109:14942–14947.
Srivathsan A., Meier R. 2011. On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. Cladistics 27:1–5.

Stork N.E. 1993. How many species are there? Biodivers. Conserv. 2:215–232.
Stork N.E. 1999. The magnitude of global biodiversity. In: Cracraft J., Grifo F., editors. The living planet in crisis: biodiversity science and policy. New York: Columbia University Press. p. 3–32.
Stork N.E., Grimbacher P.S., Storey R., Oberprieler R.G., Reid C., Slipinski S.A. 2008. What determines whether a species of insect is described? Evidence from a study of tropical forest beetles. Insect Conserv. Diver. 1:114–119.
Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28:2731–2739.
Tarcz S., Przyboś E., Surmacz M. 2013. An assessment of haplotype variation in ribosomal and mitochondrial DNA fragments suggests incomplete lineage sorting in some species of the *Paramecium aurelia* complex (Ciliophora, Protozoa). Molec. Phylogenet. Evol. 67:255–265.
Taylor H.R., Harris W.E. 2012. An emergent science on the brink of irrelevance: a review of the past 8 years of DNA barcoding. Mol. Ecol. Resour. 12:377–388.
Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F., Higgins D.G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24:4876–4882.
Tilzey R.D.J. 1976. Observations on interactions between indigenous Galaxiidae and introduced Salmonidae in the Lake Eucumbene catchment, New South Wales. Aust. J. Mar. Freshw. Res. 27:551–564.
Unmack P.J., Bagley J.C., Adams M., Hammer M.P., Johnson J.B. 2012. Molecular phylogeny and phylogeography of the Australian freshwater fish genus *Galaxiella*, with an emphasis on Dwarf Galaxias (*G. pusilla*). PLoS ONE, 7:e38433.
van Velzen R., Weitschek E., Felici G., Bakker F.T. 2012. DNA barcoding of recently diverged species: relative performance of matching methods. PLoS ONE 7:e30490.
Vespoor E., Hammar J. 1991. Introgressive hybridization in fishes: the biochemical evidence. J. Fish Biol. 39 (Suppl. A): 309–334.
Vrijenhoek R.C. 2009. Cryptic species, phenotypic plasticity, and complex life history: assessing deep-sea fauna diversity with molecular markers. Deep-Sea Res. Part II - Topical Stud. Oceanogr. 56:1713–1723.
Walsh P.S., Metzger D.A., Higuchi R. 1991. Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 10:506–513.
Waters J., Craw D., Young J., Wallis G. 2001. Genes meet geology: fish phylogeographic pattern reflects ancient, rather than modern drainage connections. Evolution 55:1884–1851.
Welton L.J., Siler C.D., Oaks J.R., Diesmos A.C., Brown R.M. 2013. Multilocus phylogeny and Bayesian estimates of species boundaries reveal hidden evolutionary relationships and cryptic diversity in Southeast Asian monitor lizards. Molec. Ecol. 22:3495–3510.
White B.J., Collins F.H., Besansky N.J. 2011. Evolution of *Anopheles gambiae* in relation to humans and malaria. Annu. Rev. Ecol. Evol. Syst. 42:111–132.
Wishart M., Hughes J., Stewart B., Impson D. 2006. Extreme levels of intra-specific divergence among Cape Peninsula populations of the Cape galaxias, *Galaxias zebratus* Castelnau 1861, reveals a possible species complex. Afr. J. Aquat. Sci. 31:99–106.
Xu F.F., Jerlstrom-Hultqvist J., Andersson J.O. 2012. Genome-wide analyses of recombination suggest that *Giardia intestinalis* assemblages represent different species. Mol. Biol. Evol. 29:2895–2898.