

REVIEW ARTICLE



Unpacking the species conundrum: philosophy, practice and a way forward

KARTIK SHANKER^{1,2,3*}, S. P. VIJAYAKUMAR¹ and K. N. GANESHAIAH^{3,4}

¹Centre for Ecological Sciences, Indian Institute of Science, Bengaluru 560 012, India

²A-001, Samvridhhi Gardenia Apartments, 88/3, Byataranapura, Near Sahakara Nagar A Block, Bengaluru 560 092, India

³Royal Enclave, Srirampura, Jakkur Post, Bengaluru 560 064, India

⁴School of Ecology and Conservation, University of Agricultural Sciences, GKVK, Bengaluru 560 065, India

*For correspondence. E-mail: kshanker@gmail.com.

Received 20 January 2017; revised 27 June 2017; accepted 29 June 2017; published online 19 July 2017

Abstract. The history of ecology and evolutionary biology is rife with attempts to define and delimit species. However, there has been confusion between concepts and criteria, which has led to discussion, debate, and conflict, eventually leading to lack of consistency in delimitation. Here, we provide a broad review of species concepts, a clarification of category versus concept, an account of the general lineage concept (GLC), and finally a way forward for species discovery and delimitation. Historically, species were considered as varieties bound together by reproduction. After over 200 years of uncertainty, Mayr attempted to bring coherence to the definition of species through the biological species concept (BSC). This has, however, received much criticism, and the last half century has spawned at least 20 other concepts. A central philosophical problem is that concepts treat species as ‘individuals’ while the criteria for categorization treats them as ‘classes’. While not getting away from this problem entirely, the GLC attempts to provide a framework where lineage divergence is influenced by a number of different factors (and correlated to different traits) which relate to the different species concepts. We also introduce an ‘inclusive’ probabilistic approach for understanding and delimiting species. Finally, we provide a Wallacean (geography related) approach to the Linnaean problem of identifying and delimiting species, particularly for cases of allopatric divergence, and map this to the GLC. Going one step further, we take a morphometric terrain approach to visualizing and understanding differences between lineages. In summary, we argue that while generalized frameworks may work well for concepts of what species are, plurality and ‘inclusive’ probabilistic approaches may work best for delimitation.

Keywords. species concepts; delimitation; allopatry; Wallacean shortfall; morphometric terrain; sister lineages.

A brief history of species

The earliest documented effort at a systematic classification of natural objects in ‘Western science’ is Aristotle’s principle of logical division, where every object (living or nonliving) was classified through a series of binary steps. At every step, each object was defined on the basis of five predicables: species (eidos), genus (genos), differentia (diaphora), property, and accident. For example, a human (species) is an animal (genus) with the ability to reason (differentia). The genus and species referred to general and specific categories the object belonged to rather than the particular taxonomic sense in which we use them today. Many objects reached their final classification in just a few steps, while others needed several steps of subdivision. In his ‘*Historia animalium*’ (History of Animals), Aristotle provided the natural

history of more than 500 species of animals (Young 2007).

Following his extraordinary contributions to natural history, a millennium and a half passed before naturalist-philosophers turned their attention to the problem of classification. In the 17th century, most plants and animals were classified on the basis of their relationship to humans. John Ray (sometimes Wray) and Francis Willughby made what might be considered the first ‘scientific’ attempt at classification, the former working on plants and latter on birds. Ray and Willughby documented hundreds of new species or varieties and classified them for the first time on the basis of the form (and function) of parts. Ray and Willughby’s ‘*The ornithology*’ followed the principle of logical subdivision based on form and structure (beak, foot and overall size). They divided birds into land and water birds; land birds were divided into birds with straight and

curved beaks; and those with curved beaks into prey and fruit eaters, etc. Some final groupings were achieved after three subdivisions (flightless birds) while others needed up to eight subdivisions. Many of their final groupings have good correspondence with some groups today, such as crows, woodpeckers, geese, etc. (see [Young 2007](#) for a detailed history of the discovery of evolution).

Crucially, Ray also recognized that species or types were bound together by reproduction. He wrote: ‘no surer criterion for determining species has occurred to me than the distinguishing features that perpetuate themselves in propagation from seed’ (Ray 1686). Inherent in this is the idea of gene flow that has remained central to notions of species for three and a half centuries. While ‘evolution’ had not been discovered, it was then widely believed that species or kinds were mutable and could arise spontaneously, which explained ‘sports’ or ‘monsters’. Hence, Ray’s contributions strengthened the idea that species were largely constant in time with relatively small variations ([Young 2007](#)).

Linnaeus strengthened the notion of fixity of species by assigning ‘species’ a binomial nomenclature. In addition, he was also interested in the ecological relationships of species. That species were adapted to tropical and temperate climates led him to telescope the Biblical flood and creation, and hypothesise that all species originated on a mountain which would contain the range of environments necessary for different kinds of plants and animals ([Young 2007](#)). The idea of ‘ecological species’ has its roots here.

The third and final class of ‘evolutionary’ species concepts, based on ancestry and descent derives from the work of Charles Darwin. Though Darwin wrote that he ‘was much struck how entirely vague and arbitrary is the distinction between species and varieties’ (1859), his work on the origin of species by natural selection provides a framework for understanding species as lineages. In a letter to Hooker, Darwin wrote (dated August 1857) ([Darwin 1887](#)):

‘I am got [*sic*] extremely interested in tabulating, according to mere size of genera, the species having any varieties marked by Greek letters or otherwise: the result (as far as I have yet gone) seems to me one of the most important arguments I have yet met with, that varieties are only small species—or species only strongly marked varieties. ... It is good to have hair-splitters and lumpers’

Mayr and Dobzhansky grappled with this species question (i.e. what they were rather than how they originated), leading to Mayr’s formulation of the biological species concept (BSC) (1942, 1963). [Dobzhansky \(1950\)](#) viewed species as a ‘reproductive community of sexual and cross-fertilizing individuals which share a common gene pool’. He however remained skeptical and wrote: ‘Of late, the futility of attempts to find a universally valid criterion

for distinguishing species has come to be fairly generally, if reluctantly, recognized’ ([Dobzhansky 1937](#)). Mayr first defined species as ‘a group of actually and potentially interbreeding populations that do not breed with any other ...’ in 1942, but removed the ‘potentially’ in subsequent definitions. [Mayr \(1963\)](#) defined biological species as ‘groups of interbreeding natural populations that are reproductively isolated from other such groups’.

[Mayr and Ashlock \(1991\)](#) recognized that the multitude of ways in which speciation occurs makes the process of species delimitation difficult. They suggested that one could use the degree of difference in ‘good species’ (those that were clearly morphologically separated from their sibling species) to make inferences about the status of cryptic or sibling species (those that exhibited conserved morphological traits despite deeper divergences). They advocated that the morphology of sympatric species could be used as a guide for delineating allopatric taxa. While the BSC maintained pre-eminence as the central definition of species, phenetic or morphological methods were still used in practice.

The latter half of the 20th century received several centuries of accumulated thought and wisdom, but little clarity about what a species was. Philosophical arguments about species ranged from realism (species are real objects that exist independent of our existence or knowledge or description of them) to nominalism (species are human constructs), and whether they are a single or multiple kinds of entities. We propose that biologists and philosophers fall into three main schools of thought: (i) a ‘Darwinian’ view that species are no different from varieties, and we should not expend effort in trying to define them, rather we should try to understand the evolutionary processes that underlie variation (ii) a ‘Dobzhansian’ view that species are a special class of discontinuities, but it is futile to find a single universal concept and one should adopt a pluralistic approach, and (iii) a ‘Mayresque’ doctrine in which there is a universal species concept (it’s not just physics envy) that we should endeavour to derive from the principles of ecology and evolutionary biology (monism).

There have already been numerous reviews of species concepts (see [Wilkins 2009](#); [Zachos 2016](#)). However, what may be lacking is a clear enunciation of why the gap exists between philosophical theorizing about species, and practical attempts to delimit them. In this paper, we first provide a historical narrative, and broadly classify and review species concepts. We then elaborate on the general lineage concept (GLC) and introduce an ‘inclusive’ probabilistic approach to understanding species. We clarify the distinction between category and concept by examining the ontology of species as individuals versus classes, or alternately as fuzzy sets. Finally, using the GLC and inclusive approaches as a backdrop, we provide a way forward for species discovery and delimitation using a ‘Wallacean’ framework and a ‘morphometric terrain’ approach.

A historical account of species concepts

Both Darwin and Wallace had considered the importance of sterility and fertility in defining species. Darwin (1859) realized that neither offered ‘any clear distinction between species and varieties’. There were many cases where ‘good species’ had few sterility barriers, and others where sterility was not associated with speciation. Wallace (1865) had posited a species definition based on interbreeding (similar to BSC), but immediately dismissed it due to the practical difficulty of making all the necessary crosses to test compatibility. He also believed that there was circularity in such a definition; when speciation involved a reduction in the ability to interbreed, this could not then be used as a criterion to define species (c.f. Mallet 1995) as it confused process (cause) and pattern (effect).

Following its formal postulation by Mayr (1942, 1963), the BSC has remained at the centre of discussions around what species are for three quarters of a century, including in recent decades. However, one of the most scathing critiques was itself written nearly 50 years ago; Sokal and Crovello (1970) reviewed the BSC and examined whether it goes beyond phenetic approaches in practice. In the BSC, species are defined by the ability of individuals to breed with each other, and hence reproductive isolation plays a central role. However, reproductive isolation between populations need not be binary, but can vary by degree; in the last two decades, molecular methods have allowed the estimation of the level of migration between populations. Hence, some populations could be completely isolated while others are isolated to greater or lesser extents. Given this, could it be (as Darwin suggested) that variation above and below the ‘species’ level is a continuum? Or, as Sokal and Crovello (1970) ask: ‘Is there a special class of discontinuities that delimits units (biological species) whose definition and description should be attempted because they play a significant role in evolution or in understanding it?’

They point out that, in the BSC, species are (i) defined by distinctiveness (reproductive gaps) rather than by morphological (phenetic) difference, (ii) made up of populations rather than unconnected individuals, and (iii) more unequivocally defined by their relation to non conspecific populations (isolation). Hence, to establish species under the BSC, one has to find (i) some individuals which lack distinctiveness from each other—these are grouped into populations, (ii) a group of populations which interbreeds, and (iii) a group that does not interbreed with other such groups. This begs the key questions of what interbreeding is, and how it can be detected.

Sokal and Crovello (1970) also pointed out that practical taxonomy rarely (if ever) resorted to the criterion of isolation and used mainly phenetic difference (and now in addition, genetic). To answer their question about whether the concept was useful in understanding evolution, they

pointed out that homologies were used to infer evolutionary relationships, that biological populations were the units of such enquiry, and that population genetics was the heuristic tool that was used to determine relationships between these units. In their opinion, the BSC served neither a practical role in describing species, nor a conceptual role in understanding evolution.

Despite this, the idea that reproductive isolation is essential to defining species remained embedded in the numerous other species concepts that were proposed. In his review, Ridley (1993) discussed seven concepts (phenetic, biological, recognition, ecological, cladistic, pluralistic, and evolutionary) and advised using a combination of four concepts (biological, cladistic, ecological, and recognition). King (1993) reviewed eight (morphological, biological, recognition, ecological, cladistic, cohesion, evolutionary, and phylogenetic) and concluded that the BSC was ideal.

We will define some of the key concepts here, though more than 20 have been proposed (Mayden 1997; Wilkins 2006a, b; Zachos 2016). In general, these can be categorized as being based on (i) some form of gene flow and reproductive isolation (biological, recognition, ecological), (ii) lineage divergence (evolutionary, cladistic, phylogenetic) and (iii) morphological or phenotypic divergence. Finally, there are sets of what might be considered attempts to unify the different approaches to defining species.

Reproduction based species concepts

Biological (isolation): These are groups of interbreeding natural populations that are reproductively isolated from other such groups (Mayr 1963).

Recognition: In contrast to isolation which formed the core of the BSC, Paterson (1985) proposed a concept based on mate recognition (critical for both gene flow and isolation) that a species is ‘that most inclusive population of individual, biparental organisms which share a common fertilization system’.

Ecological: Van Valen (1976) proposed the idea of adaptive zones or niches (the sum of environmental factors that conspecifics interact with) as defining the boundaries of species. This is an attractive proposition, as it brings phenotype–environment interactions into what constitutes a species.

Lineage based species concepts

Evolutionary: First defined by Simpson (1961) and further modified by Wiley (1978), the evolutionary species concept (ESC) defines species as lineages (an ancestral-descendent sequence of populations) that are evolving separately from others and with their own unitary evolutionary role and

tendencies. Within the ESC, all organisms belong to some evolutionary species, reproductive isolation is complete, morphological distinction is not necessary and they are monophyletic.

Phylogenetic: The phylogenetic species concept (PSC) takes different forms, among which is the Hennig (1966) concept, grouping together the set of organisms between two speciation events or between a speciation and extinction event; an ancestor becomes extinct when the lineage splits. Some authors use diagnosability as a key criterion (Nelson and Platnick 1981; Eldredge and Cracraft 1980; Cracraft 1983; Nixon and Wheeler 1990), while others focus on monophyly (Rosen 1979; Donoghue 1985; Mishler 1985), i.e. species are monophyletic groups of organisms recognized as lineages on the basis of shared derived characters and ranked as species because of causal factors. Cracraft (1983) defined species as ‘the smallest diagnosable cluster of individual organisms within which there is a parental pattern of ancestry and descent’. Unlike the BSC and ESC, the PSC does not depend on reproductive isolation, examines shared derived characters and looks at the result of evolution rather than the process.

Phenotype based concepts

Morphological: Traditionally the most widely used concept, this is merely a diagnostic account which differentiates species on the basis of morphological characters.

Phenetic: This refers to a cluster of characters (Sokal and Sneath 1963; Michener 1970; Sokal and Crovello 1970) and may be considered as the phenotypic analogue of Mallet’s genotype cluster concept.

Geopolitical: Partially in jest, Karl and Bowen (1999) proposed a geopolitical species concept, where populations could be designated as species for geographical/political reasons. While the latter part of their argument is intentionally facetious and intended as a barb against overzealous conservationists, the former (geography) draws attention to an often underestimated criterion in species delimitation, namely geographical separation (though not along political boundaries).

Unifying concepts

Cohesive: Templeton (1989) proposed a concept to unify biological and evolutionary approaches. He suggested that species can be viewed as ‘evolutionary lineages bounded by cohesion mechanisms that cause reproductive communities, particularly genetic exchange, and ecological interchangeability’. Here, a species is the largest delimited population that functions as an internal mechanism ensuring mutual phenotype cohesion of its members.

Genotypic cluster: Mallet (1995) proposed that species could be considered as genotypic clusters (i.e. with an absence of intervening heterozygotes). He considered this a ‘modern

synthesis’ where he defined species in a Darwinian way as ‘groups of individuals that have few or no intermediates when in contact’; he included polytypic species in the definition and recommended use of knowledge from genetics. Most importantly, within this concept, species are affected by gene flow, not defined by it.

Inclusive: Following on the idea of inclusive fitness (Hamilton 1964), we propose the idea that species could be considered as ‘that inclusive group of individuals that have finite probabilities of contributing to a common gene pool’ (first presented by Shanker 2001). This is similar to Mallet’s (1995) genotype cluster concept, but adds the probability that in certain cases, individuals have a finite probability of contributing to more than one cluster (figure 1). In some instances (hybrid zones), these probabilities could be high even in ecological time, whereas in many cases, the probabilities are likely to be nonzero in evolutionary time but have a low chance of detection by current methods.

Most importantly, the factors affecting gene flow (both prezygotic and postzygotic mechanisms) can be explicitly studied and modelled, from proximate individuals to populations to species level lineages. The role of biogeography, i.e. the combination of topography, medium of dispersal, historical and current climate, ecology and organism traits (dispersal ability) can be explicitly examined to make inferences about the probability of gene flow. This is more closely considered in ‘Understanding species boundaries: a Wallacean solution to a Linnaean problem’ section, where we offer a biogeographic approach to the discovery and delimitation of species.

GLC: De Queiroz (1998) proposed a generalized approach to understanding lineages, integrating biological, evolutionary and other processes as stages in the divergence between lineages. This differs from the cohesive and genotype cluster approaches in that it explicitly identifies different processes as occurring independently at different stages of lineage evolution and divergence. Given the general acceptance of GLC as the most unifying/universal of current theories of species, we provide a more elaborate treatment in the following section.

The general lineage concept

De Queiroz (1998) first provided a generalized scheme for understanding species/lineages, treating them as ‘independently evolving metapopulation lineages’. The term ‘independently evolving lineages’ emphasizes the ideas of ancestor-descendent lineages from phylogenetic concepts and independent history from the evolutionary concept. The use of ‘metapopulations’ drives home the idea of varying degrees of isolation and gene flow between populations over time. de Queiroz further attempts to generalize his theory by suggesting that the criteria adopted by various other concepts appear at different points of time during lineage separation, and not necessarily in the same order.

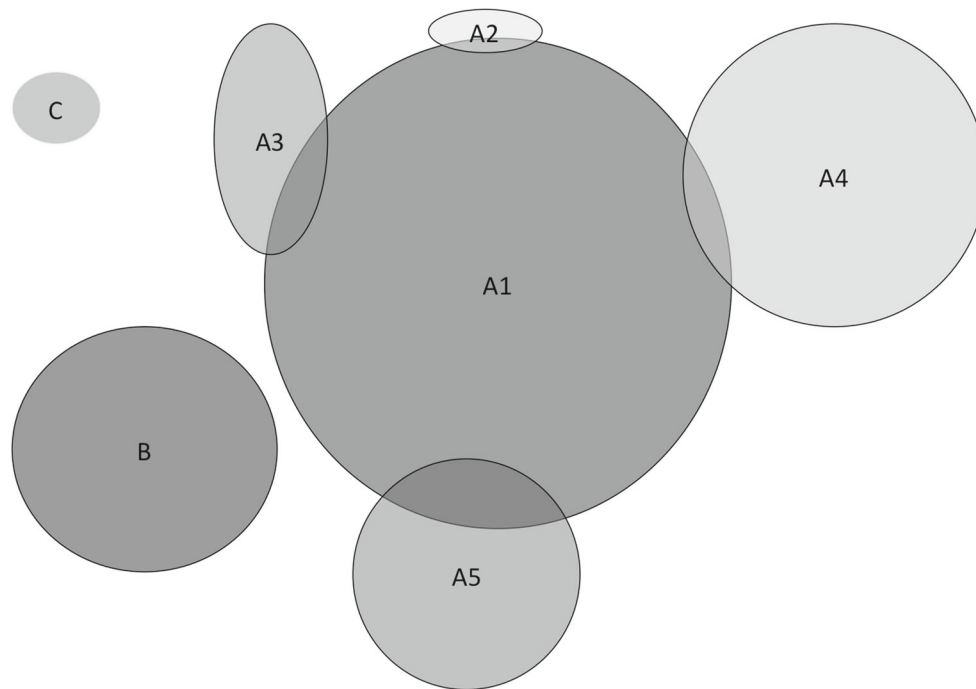


Figure 1. A schematic representation of inclusive species in genetic and/or trait space. Each circle represents a genotype cluster (sensu Mallet 1995). We extend this by suggesting that such clusters have finite but varying probabilities of gene flow with other such clusters. A1 to A5 represent clusters of different sizes (with clear morphological or geographical separation) that have varying degree of gene flow between them (A2 to A4 with A1). B and C are genotype clusters that have gene flow with each other or A. Areas of overlap represent probabilities; this can either be interpreted as the probability of contributing genes to a particular cluster (species as individual), or the probability of assignment of an individual to a particular cluster (species as class).

In this definition, a lineage is an ancestral-descendant sequence of populations. Populations are different from lineages only along a temporal axis. Simpson (1951) points out that a lineage is a population extended through time and, conversely, a population is a lineage at a particular moment in time (i.e., a temporal cross section of a lineage). Here, species are not entire lineages, but a small segment of a lineage (a branch in the line of descent); they can be viewed as lineage segments between branching events (internodes of phylogenetic trees) though there may be disagreement about what constitutes a branching event or node.

De Queiroz (2011) argues that Darwin conceptualized species as branches in the lines of descent (segments of lineages). He also highlights the fact that, contrary to the views held by Mayr (1963) and others like Coyne and Orr (2004), Darwin touched upon the role of geographical isolation in species formation. De Queiroz argues that Darwin perceived species as real, despite his comments on the arbitrary distinction between species and varieties. In his view, 'Darwin's treatment of the species as a rank and his adoption of amount of difference as a ranking criterion led him to adopt the implicit position that all species are separately evolving lineages but not all separately evolving lineages are species. Only those separately evolving lineages that had evolved a certain amount of difference from other such lineages were to be ranked as species; those that had

not evolved the requisite amount of difference were to be ranked as subspecies, varieties, or nothing at all.'

Many biologists have treated 'species category as a rank to be conferred on a lineage' after it has reached a certain level of divergence (de Queiroz 1998). Lineages that had not reached a stage were grouped under varieties by Darwin and as subspecies by others. Considered from the evolutionary species perspective, each ranking criterion represents a 'threshold crossed by separately evolving lineages as they diverge from one another' (de Queiroz 1998). One can debate whether there is any real possibility for reconciliation across concepts, as these thresholds are highly sensitive to the questions that biologists are interested in.

De Queiroz (1998) defines the 'species problem' as our failure to distinguish between the concept of what a species is and the criteria that we use to identify them. He defines species as a segment of independently evolving metapopulation lineages (better visualized as a section between an internode on a phylogenetic tree) (de Queiroz 1998, 2007) and the 20 plus species concepts that are extant in the literature as just different secondary criteria that reflect the diversity of changes or 'events' that occur during speciation (de Queiroz 1998, 2011). The conflict arises because of temporal differences in rate of change of different secondary species criteria (e.g., reproductive isolation, niche differentiation, monophyly) during speciation (de Queiroz

2007). The different properties recognized in species delimitation are emergent characters as a consequence of the evolution of species criteria (de Queiroz 2007).

De Queiroz's concept is appealing particularly because it incorporates evolution into systematics and emphasises underlying processes, and shifts the focus from static views that treat species as taxonomic ranks. In this evolutionary sense, which is apparently an extension of Darwin's initial conceptualization, all identifiable independently evolving lineage segments are species irrespective of whether the lineages have acquired a certain property (de Queiroz 2011). This proposal attempts to unify (hence the name unified species concept, *sensu* de Queiroz 2007) the underlying common thread in conflicting concepts of what species are and paves the way for researchers to focus on species delimitation—determining species boundaries—which hinges on testing for one or more of the criteria that lend support for lineage separation (de Queiroz 2007). According to this proposal, any single criterion can be used in recognizing species, as long as it lends strong evidence for the independently evolving nature of a lineage. In the past, conflict has likely arisen from the different methods involved, and can be easily avoided by focussing on multiple lines of evidence to support species separation (de Queiroz 2007).

De Queiroz (2011) says: 'In sum, there is evidence from the history of biology that what I have called "the unified species concept" has been developing ever since Darwin's revolutionary proposal to equate species with branches in the lines of descent. Indeed, the emerging unified concept appears to be a natural outgrowth, if not a logical consequence, of accepting an evolutionary underpinning for the concept of species.'

Are species classes or individuals?

De Queiroz's distinction between concepts and criteria can be viewed in the light of the ontology of species as 'individuals' and 'classes'. Ghiselin (1974) and Hull (1978) made a critical philosophical distinction that paves the way for understanding some of the crisis in reconciling concepts and criteria. Ghiselin (1974) argues that species should be treated as 'individuals' and not 'classes'. As individuals, species would have proper names, would not have defining properties and their constituent organisms would be parts, not members. He defines species as the 'most extensive units in the natural economy such that reproductive competition occurs among their parts' and compares them to firms in economy. Treating species as 'individuals' implies that a species is a defined body of objects through its history, i.e. a very specific collection of organisms that either are or are not part of its history, bound by an ancestor-descendant relationship. Therefore, regardless of morphological or any other trait, an individual is either part of that historically bound chain of organisms, or is

not! Being bound by reproduction and gene flow, organisms within a 'species' may be similar in morphology, and more likely to breed with each other, etc. but none of this is essential conceptually.

Most biological and evolutionary species concepts treat species as individuals. They are bound together by reproduction and as ancestor-descendent lineages, but reference different slices of time. The GLC also treats species as individuals in the same manner, except that it generalizes the temporal component. As de Queiroz points out, morphological and other trait changes can occur at different points in lineage history in different clades, and different species concepts focus on different slices of time. Thus, e.g., reproductive isolation may precede morphological or ecological divergence in one clade while it may follow them in another. Species concepts focussing on any single aspect are therefore capturing only a different stage of divergence.

Treating species as 'classes' on the other hand requires a definition of the traits of that class and not the objects that belong to it. Once the traits or properties of that class or set are defined, objects may or may not belong to that class. An object can belong to more than one class, with or without an associated probability, as long as it satisfies the properties of that class. Many species concepts (morphological based ones) and all criteria treat species as classes. Let us take the example of a morphological mutant. Assuming that this is still capable of reproduction, but wildly divergent from all members of its species, this would still belong to the same species if we were to treat species as 'individuals'. However, if we were to treat it as a class, it is likely that this organism would not be classified along with the other members of its species, especially if information on ancestry were not available (as it most often is not).

As long as we treat species as classes, an individual could potentially belong to more than one species, but this contrasts wildly with our notion of what species are. This can be put down to the paradoxical situation where our 'notion' relies on 'concept' (=species as individuals) but delimitation derives from categorization of species into classes, which is true even for genetic/phylogenetic concepts, where actual assignment is based on genetic traits.

Kitcher (1984) provides an ontologically neutral stance by treating species as sets of organisms. Thus, some organisms can be spatio-temporally restricted sets (= individuals), while others are spatio-temporally unrestricted sets. Kitcher suggests that proximate causes (similarity in genetics or morphology) correspond to species defined as the spatio-temporally unrestricted sets, while ultimate/evolutionary causes correspond to species as individuals.

Another view is that species are natural kinds (individuals) as homeostatic property clusters (HPC) (Boyd 1999; Millikan 1999; Wilson 1999), i.e. groups of entities which share stable similarities. While HPC does better at treating species as natural kinds than traditional

essentialism (which posits that species have unique properties), it nevertheless suffers from weaknesses in dealing with geographical variation and polymorphism, and more importantly does not treat species as genealogical lineages (Ereshefsky 2007).

A practical solution to this thorny problem may lie in a probabilistic approach (as proposed in the inclusive concept) (figure 1). Individual organisms can be assigned a probability (rather than a binary 1–0) of belonging to a species, either when treated as individuals or classes, i.e. as fuzzy sets. When species are treated as individuals, probabilities can mean that (i) given our lack of knowledge, we can only guess (probability) at which lineage (=individual ancestor-descendant line) the organism belongs to; (ii) organisms are contributing to multiple lines to varying degrees which could be relevant in plants (ploidy), hybrid zones, etc. On the other hand, when species are treated as classes, probabilities can mean that an organism has traits that lie between two trait distribution clusters.

Since all delimitation methods (criteria) must treat species as classes, a fuzzy set approach may suit the species problem best, that some are restricted while others are unrestricted sets. The unrestricted sets will have varying degrees of fuzziness, i.e. biologically, varying probabilities of gene flow with their nearest neighbours, and practically, varying probabilities with which individuals can be assigned to one or another set.

Understanding species boundaries: a Wallacean solution to a Linnaean problem

Many of the world's described and undescribed species are concentrated in the global biodiversity hot spots (Myers *et al.* 2000; Joppa *et al.* 2011). In these regions, the 'Linnaean shortfall', or the problem of undescribed lineages co-occurs with the 'Wallacean shortfall', the lack of distribution data (Whittaker *et al.* 2005) and both form a major impediment to our understanding of biogeographic and evolutionary history. One of the main challenges with reference to the 'Linnaean shortfall' is that of species delimitation. Though delimitation of lineages has been facilitated by molecular systematics (Vences *et al.* 2005), recent discoveries also highlight uncertainties in the estimation of new species. The underlying cause for the uncertainty in these estimates remains unclear but could, at least in part, be attributed to the larger problem of recognizing and delimiting lineages (Wiens and Penkrot 2002). This is usually compounded by an incomplete understanding of divergence in space and time, which is the result of incomplete sampling or the 'Wallacean shortfall'.

Significance of sisters

Sister species are the closest relatives—independently evolving lineages—that share a recent common ancestor.

The identification of sisters is critical to understanding species boundaries. Since sister species are a product of speciation, they are good models for understanding the processes driving speciation which can provide novel insights into species boundaries and species delimitation. Sister lineages that are deeply divergent can be diagnosed as independently evolving with considerable certainty. In these cases, as suggested in the GLC, any single criterion can be used in recognizing species, as long as it provides strong evidence that the lineages are evolving independently. However, most problems in species boundaries arise at the shallow phylogenetic scale of recently diverged sister lineages. Recognizing species with sympatric distributions is generally less problematic than delimiting species which are isolated in space. Conventional species concepts such as the BSC cannot be used in such cases, and other criteria are required to test the separation of such lineages over evolutionary time. Recent conceptual contributions (GLC) provide an avenue to address the species delimitation problem in allopatric lineages. Addressing this problem requires multiple lines of evidence to support species separation (de Queiroz 2007).

Conventionally, most species have been delimited based on morphology. In other cases, characteristics like mating calls have been used successfully in recognizing species, including those that are otherwise morphologically cryptic (Funk *et al.* 2008). The advent of molecular tools has enabled the use of phylogenetic frameworks and genetic (temporal) axes to delineate species (Avice and Johns 1999). While genetic distance has been strongly advocated as a defining variable (Fouquet *et al.* 2007), it has been suggested that lineage delimitations based purely on arbitrary cut-offs in a phylogenetic tree and those based on genetic divergence can be misleading (Linkem *et al.* 2010). Where allopatry is shown to be the dominant mode of speciation, geographic range can be used as an important criterion in recognizing isolated lineages (Mayr and Ashlock 1991). The presence of unsuitable areas between geographically disjunct sister lineages indicates a lower probability of dispersal and supports isolation (Wiens and Graham 2005). Hence, lineages can be explored along multiple axes including genetic relatedness, geographical distribution and phenotypic (morphology/behaviour) traits (figure 2). Here, we use a case study on bush frogs in the Western Ghats of Peninsular India to explore a hierarchical, multicriteria approach for delimiting lineages based on a mtDNA haplotype phylogeny.

Significance of geography in understanding species boundaries: the case of bush frogs

We addressed the problem of species delimitation using bush frogs (Family: Rhacophoridae) in one of the global biodiversity hot spots, the Western Ghats mountains of Peninsular India. Bush frogs refer to three large clades

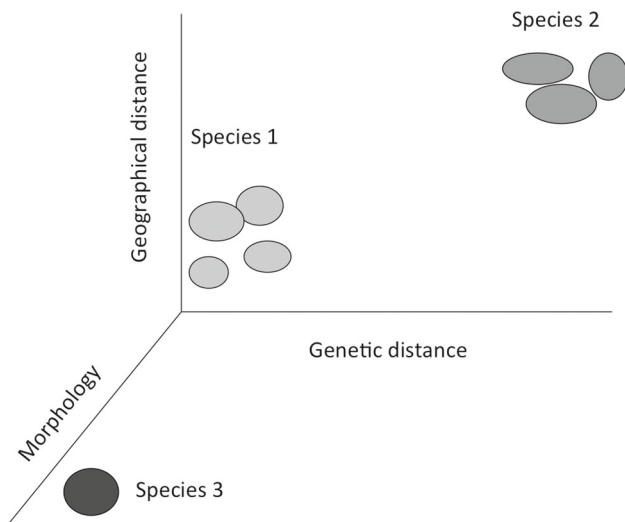


Figure 2. A conceptual representation of differences between species in genetic, geographical and morphological space. Theoretically, the degree to which species differ from each other along different axes can vary from pair to pair.

of frogs: *Raorchestes*, *Pseudophilautus* and *Philautus*. The combined range of the three clades extends from parts of South Asia (Peninsular India, Sri Lanka and northeastern Indian regions) to South-East Asia (Burma to Borneo). We focus on *Raorchestes*, a recently discovered, large vertebrate radiation in Peninsular India, with its centre of diversification in the Western Ghats (Vijayakumar et al. 2014, 2016). The basic approach is as follows:

- (1) The Western Ghats is a highly heterogeneous mountain range with a complex topography and associated climatic gradients that have created multiple opportunities for lineage separation and speciation. We used a sampling design to account for the heterogeneity at the regional scale and hence uncover as many bush frog lineages as possible along the major hill complexes across both elevational and latitudinal gradients. This geographically explicit approach, which incorporated both historical and current barriers, as well as ecological gradients, was critical for the identification of sister lineages, which in turn is critical for understanding lineage divergence.
- (2) The phylogenetic tree (using 16S, ND4, Cytb genes) was used to identify exclusive haplotypes (Wiens and Penkrot 2002) based on individual clusters with strong bootstrap values (>70).
- (3) Genetic distance for three genes was obtained for all haplotype clusters with their sister lineages or the closest branch on the tree obtained from step 2.
- (4) We then adopted two strategies, both yielding similar results:
 - (a) Following Mayr and Ashlock's (1991) suggestion of using the degree of difference between 'good'

species to make inferences about allopatric taxa within a genus, we selected a sympatric sister—lineage pair (from step 1) that exhibited the lowest genetic divergence and yet showed strong morphological divergence and could be distinguished by their advertisement calls in the field.

- (b) Considering that such sympatric sister lineages may not always occur, and the problems associated with genetic cut-offs, one can potentially skip steps 4 and 5 and jump to step 6.
- (5) The genetic distance of the sympatric lineage pair was used as a cut-off to filter haplotype clusters identified in step 2. Lineages that met the cut-off distances in at least two genes (step 3) were retained.
- (6) Potential lineages were further classified into low (1% to $<2\%$), moderate (2% to $<4\%$), high (4% to $<6\%$) and very high ($>6\%$) levels of divergence based on 16S rRNA distance values. A majority of the lineages with moderate to very high divergence levels could be easily distinguished in field. This classification into discrete bins is arbitrary and a continuous distribution can be used.
- (7) We then tested the low divergent sister lineages for overlap in geographical (step 7a) and morphological space (step 7b).
 - (a) Based on geographical range overlap, low divergence lineages were identified as allopatric (with zero overlap), parapatric (with partial overlap on range margins or abutting ranges) and sympatric (overlapping ranges) and uncertain (the data to assign them to a category was inadequate).
 - (b) Shallow divergent sister lineages were also examined for separation in multivariate morphological space using principal component analysis.
- (8) Using species distribution modelling, we tested for isolation of allopatric sister-lineages in geographical space.

The above approach uses a combination of molecular haplotype phylogeny, genetic distance, geographic distribution, morphological and acoustic variables to delimit potential candidate lineages in the Western Ghats bush frog clade. Considering that species are a result of evolutionary and biogeographic processes, the hierarchical approach used here allowed us to objectively determine lineage boundaries and accommodate the resulting uncertainties in understanding the drivers of diversification of these fundamental units of evolution.

Our results showed a significant increase in the number of potential lineages of bush frogs in this region (Vijayakumar et al. 2014). Both described and newly delimited lineages fell under a range of divergence levels (16S) from low to very high, suggesting that previous explorations were not systematic. The results suggest that

the approach was effective in uncovering lineages varying in their levels of divergence and particularly shallow divergent lineages; up to 45% of the potential new lineages belonged to the shallow divergence category. Geographic isolates are the fundamental units of speciation (Mayr 1963) and understanding the nature of range limits among putative lineages provides important insights into species delimitation (Mayr and Ashlock 1991). The incorporation of geography as a variable along with phylogenetic relationship played a significant role in uncovering a large number of missing lineages. Overlaying shallow divergent sister lineages on geography indicated a predominance of allopatric lineages. Notably, all recent allopatric sister lineages exhibited a divergence of around 1% on the 16S gene. The pattern can be tested across other clades in the Western Ghats and for other taxa and regions across the world.

Fouquet *et al.* (2007) advocate 3% divergence on the 16S rRNA gene for Neotropical amphibians. This cut-off is higher than the moderate divergent lineages recognized here in bush frog lineages. Our results suggest that one may risk missing many genetically shallow but morphologically divergent lineages if one were to apply a uniform criterion like genetic distance. Isolation by distance could also have a bearing on the genetic structure of widespread species (Fouquet *et al.* 2007), emphasizing the need for additional data from intervening populations to rule out the effect of distance in these lineages.

The two sympatric lineages that we discovered from a montane region exhibited distinct morphology despite low levels of genetic divergence ($\sim 1\%$ for 16S rRNA). If the genetic divergence of sympatric lineages recovered in this study is used as a yardstick, all the geographical isolates could be considered as independently evolving lineages. To reduce the uncertainty in characterizing these lineages as evolving independently, we further examined the shallow divergent lineages in morphological space. Some allopatric lineages showed separation along multivariate morphological axes while others did not. These results suggest a general disconnect between genetic divergence and morphological divergence among bush frog lineages (figures 3 & 4).

Using multiple criteria, we found that many lineages exhibit divergence across multiple axes; not all lineages are equally divergent across all axes examined (i.e., they occupy the entire conceptual space; figure 2). The simplified visual model presented here shows the expected trend along geographical and morphological axes across three levels of genetic divergence (figure 4). Based on preliminary analysis, we anticipate that most of the allopatric isolates—with varying levels of genetic divergence—show low to moderate variation in morphological space. In contrast, sympatric/parapatric sister pairs—despite low to moderate levels of genetic divergence—exhibit strong separation along morphological and other axes (e.g. acoustics).

Our stepwise approach suggests that a majority of these lineages showed separation along three axes: phylogeny and genetic distance, geography and morphology. Allopatric isolates generally are known to be conserved in morphological space, but our results suggest these lineages may be under different selection regimes. A few of these shallow divergent lineages showed strong signals of selection as exhibited in the high divergence in morphology.

Further, our distribution modelling supports our latitude based range overlap estimates that most species are allopatric. The underlying driver is allopatric speciation due to geographical isolation (Vijayakumar *et al.* 2016). This inference is based on lack of evidence for range expansion and connectivity across the current geographical isolates during the last glacial maximum (Vijayakumar *et al.* 2016). Explicit incorporation of geography in understanding species boundaries brings species delimitation closer to the underlying biogeographic and evolutionary processes that drive lineage diversification.

We mapped the divergence along different axes onto the lineages on the phylogenetic tree (figure 3). This can be viewed as varying degrees of separation along different axes as proposed by de Queiroz (1998). Two points are critical. First, the use of any single criterion would not provide a nuanced understanding of diversification or an adequate delimitation of lineages in the clade. Second, even within this clade, differentiation along the axes (genetics, geography and morphology) does not take place in a consistent temporal sequence (figure 3), and they are uncorrelated with each other.

The ‘morphometric terrain’ approach to species delimitation

While the above approach can provide a method to discover and delimit species incorporating genetics, geography and morphology at a lineage level, it still does not necessarily help in assigning individuals to clusters, especially when there are fuzzy boundaries for any of these axes. Here, we propose a ‘morphometric terrain’ approach to visualize trait distributions in related populations and to demarcate species boundaries. This approach provides a way forward to combine morpho-based taxonomy with evolutionary processes and gene flow.

The principle of classical taxonomy, which uses mostly morphological features, assumes that the individuals of a species exhibit high resemblance among themselves and differ from those of other species. Following Darwin, it has become clearer that natural classification and organization is a result of reproductive isolation among species and of genetic coherence among the individuals within species. Consequently, when the evolutionary based concepts of species were proposed, the basic edifice of classical taxonomy—the morpho-based classification—was not seriously affected as it reflected two important

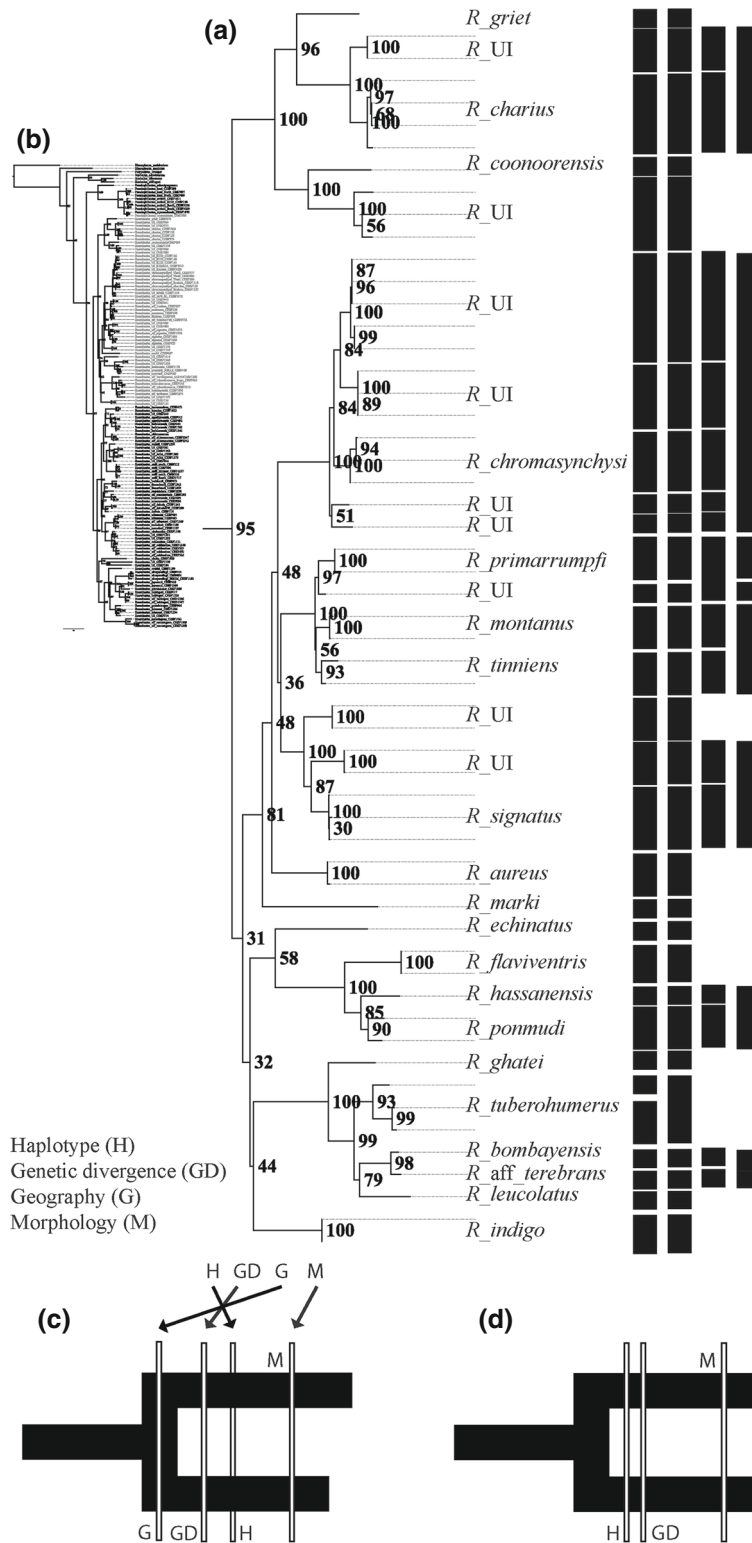


Figure 3. Haplotype phylogenetic tree, based on three mitochondrial genes, of (a) a subclade (north clade) within (b) the genus *Raorchestes* (from Vijayakumar et al. 2016). Values at the nodes represent maximum likelihood based bootstrap estimates. Separation of lineages along the different axis: haplotype clusters, genetic divergence, geography, and morphology have been indicated as black bars. In the case of geography and morphology, separation has been represented for low/shallow divergence levels (1–2% - 16S gene). (c & d) Abstract representations (based on de Queiroz 1998) of the sequence of occurrence of different processes, (c) geographical separation, followed by the genetic divergence and the evolution of haplotype clusters and morphological variation, and (d) genetic divergence and haplotype clusters followed by morphological variation with or without geographical separation. Processes can change in sequence or be absent depending on the underlying processes driving lineage splits.

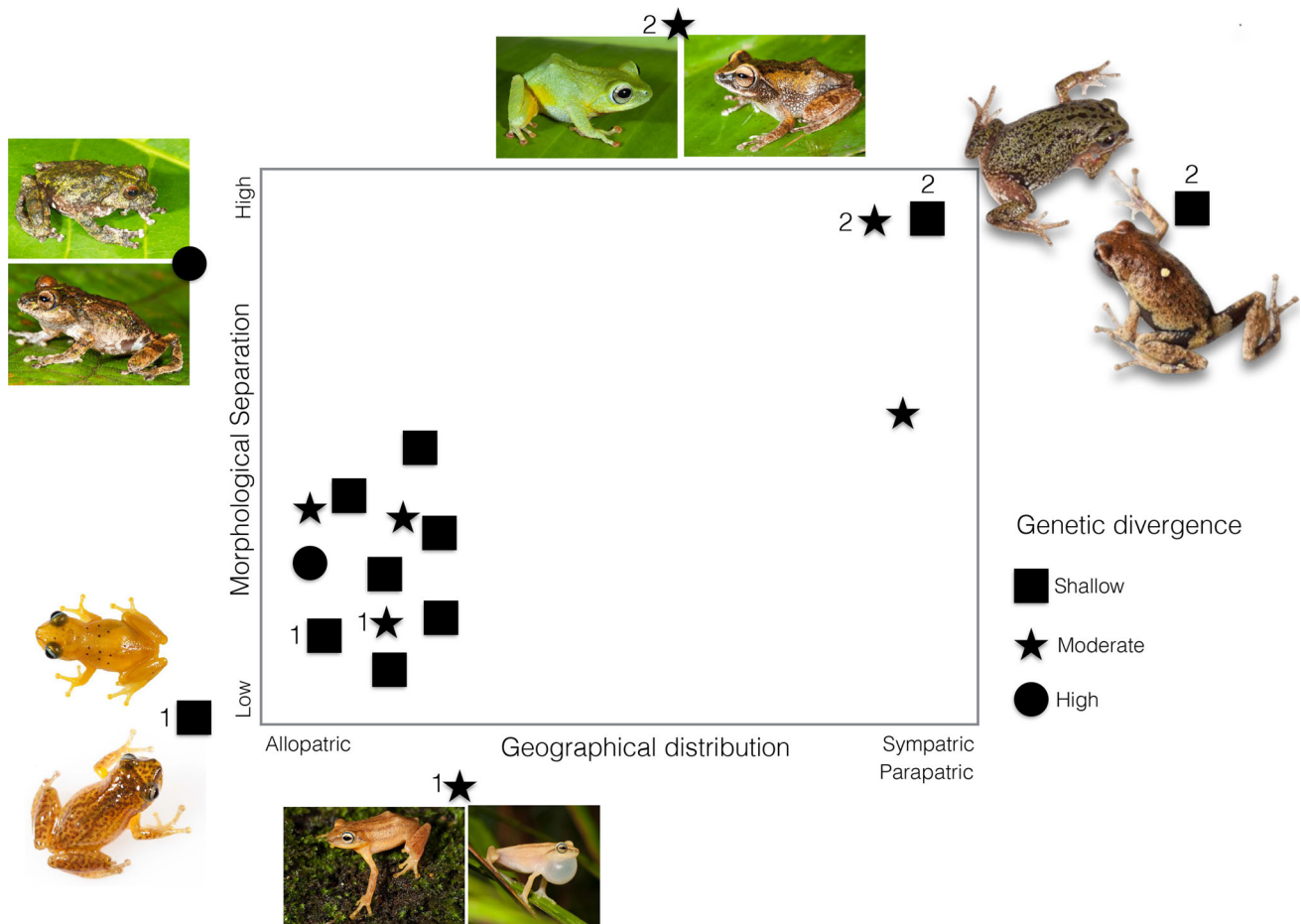


Figure 4. Separation of sister lineages with different levels of genetic divergence in geographical and morphological space. Each point refers to a sister-pair and symbols represent three different levels of genetic divergence. The position of the points in the plot are approximations and trends expected based on preliminary analysis. Representative images of such sister pairs are shown. The numbers 1 and 2 with the symbols refer to the corresponding images of sister-pairs.

pillars: while the individuals of a species resemble each other owing to their genetic ancestry, they differ morphologically from others owing to reproductive isolation.

Thus, though practicing taxonomy overlapped with the expectations from the theory of evolution, the variation deriving from evolutionary processes has not been incorporated into taxonomic classification. Here, we show that (i) a 'morphometric terrain' that reflects the density distribution of phenotypic variants can be developed for a given species based on intraspecific variations, and (ii) the visualization of such a terrain can offer insights for both practicing taxonomy and understanding species concepts. Further, such a morphological terrain can also be used to demarcate boundaries of species, identify species complexes and help resolve conflicts among taxonomists in such situations. We also show that such density terrains of morphological variants (or other traits) can provide an opportunity to define the probability with which an individual can be assigned to a particular species.

The principle

All individuals of a species are not exactly alike owing to intraspecific variations in genetic composition and their interaction with the environment. Thus, though individuals of a species are expected to be similar, they do exhibit morphological, physiological and biochemical variations. In accordance with these variations, the survival ability, reproductive capacity and hence the fitness of individuals also varies within species. The resultant fitness variations are more often continuous than discrete. In a Darwinian world, individuals with the most adaptive features, i.e. those with the highest fitness advantage, reproduce better and hence are likely to be represented in greater numbers in the population. While morphological traits may not directly influence or reflect this fitness variance, one can examine the continually varying morpho-types within each species as a representation of the consequence of ecological and evolutionary processes. If the morpho-types of a species can be located in two dimensional space, then

their population densities can be represented in the third dimension, as an undulating mat or terrain of varying heights. Generally speaking, in such a terrain, the most common morpho-types would appear as the peak or hill in the centre of distribution of different types. In a combined morphometric terrain of several species, the raised peaks of the landscape represent the mode of the species, variations within species appear as the shape of the terrain around the peak, and their limiting boundaries by the spread at base of the hill. The height of the hill and the depth of the valley surrounding it may, to some degree, reflect the extent of stabilizing selection on the most common morpho-types. We propose that such a morphometric terrain can be used as a basis for recognizing (i) the boundaries and variance of morpho-types of populations within species, (ii) the peak or the hill centre as the holotype of the species and, (iii) the species complexes existing in the group.

Developing the morphometric landscape: a case study using dung beetles

The process of developing a morphometric landscape or terrain for taxonomic purposes was attempted by [Chandrasekhara et al. \(1998\)](#). Fifteen species of dung beetles belonging to four genera were studied ([table 1](#)) to estimate interspecific and intraspecific variation. A minimum of 30

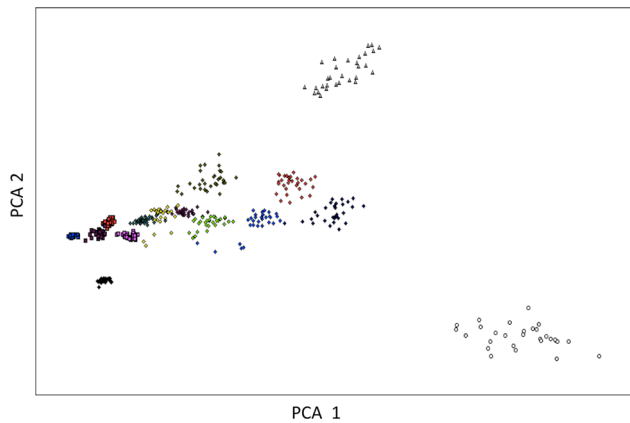


Figure 5. Separation of individuals of 15 species of dung beetles on two PC Axes. The axes explained about 65% of the variability. The colours indicate different taxonomic species. Individuals of each species are coloured and or given a different shape index.

individuals from each species were randomly sampled, and morphometric data on 26 traits of these species were collected ([table 1](#)). The data was then subjected to a Principal Component Analysis (PCA) to obtain the spread of all the individuals on two PC axes, both of which together extracted about 65% of the total variability. This two-dimensional map ([figure 5](#)) was then divided into a grid of 50 rows and 60 columns. Based on the number of individuals in each grid, a three dimensional density map of individuals was plotted on this grid space ([figure 6a](#)). This density map shows the morphometric terrain of the populations studied and their segregation into different groups or ‘species’ as peaks or hills.

These peaks were then compared with known taxonomic species. Two major patterns emerged:

1. [Figure 6b](#) shows that four of the five species of the genus *Cacobius* are distinctly isolated on the morphometric terrain; they show discrete peaks and deep valleys between them suggesting that they are completely reproductively isolated and that stabilizing selection has shaped their distinctness represented by the peaks. Clearly, such species do not offer any problems in morpho-based classification.
2. Species 4 and 5 (at the front of [figure 6a](#)) show a bridge between them as a minor ‘mound’ between their peaks. Obviously individuals comprising this mound are not a part of the two major peaks, nor do they constitute the shoulders of the two peaks. Rather, the mound (and hence the individuals in it) appears to be isolated from the two species in the valley suggesting that these individuals are a result of intraspecific polymorphism, or interbreeding or both. In other words, they are perhaps evolving, and may even emerge as a new taxonomic group; or they may disappear (local extinction); or merge into one of the neighbouring peaks.

Table 1. List of characters studied in dung beetles.

List of characters studied	
1	Maximum width of head
2	Anterior width of thorax
3	Interocular distance
4	Length of prothorax
5	Maximum width of thorax
6	Length of elytra
7	Width of elytra
8	Median length of elytra
9	Height of thorax
10	Number of teeth on tibia
11	Length of basitarsus (foreleg)
12	Length of foretibia
13	Width of foretibia
14	Inter-coxal distance of foreleg
15	Inter-coxal distance of midleg
16	Inter-coxal distance of hindleg
17	Length of basitarsus (midleg)
18	Length of midtibia
19	Length of basitarsus (hindleg)
20	Length of hindtibia
21	Width of hindtibia
22	Median length of last abdominal sternite
23	Length of pygidium
24	Width of pygidium
25	Number of horns on head
26	Number of horns on thorax

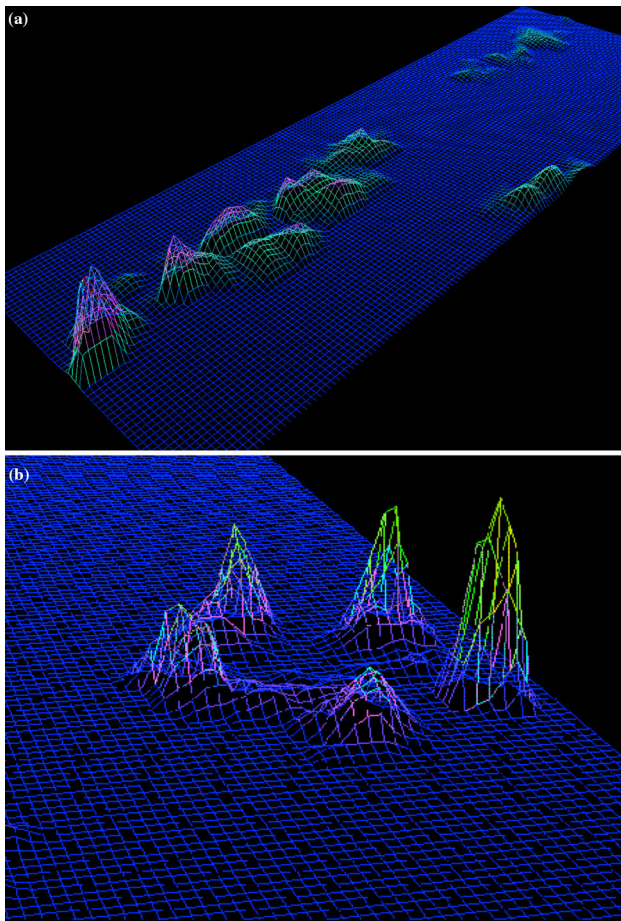


Figure 6. (a) Morphometric terrain developed based on the density of individuals of the 15 species of dung beetles. (b) Expanded view of the left cluster of peaks of the morphometric terrain map. The three peaks (species) on the right and middle are clearly separated by valleys. The rare individuals that constitute the mound in the valley between the left most peak and that in the mid-front may be a consequence of infrequent or chance interbreeding between the species represented by the two peaks. Taxonomists are likely to attribute these individuals to either of the peaks or set up a new peak (species). Either way, these individuals are sure to lead to taxonomic conflicts.

Figure 6a also shows species complexes in the middle and to the right of the morpho-space. While peaks do show up, there are no clear boundaries among the species. Often, there are also shoulders adjoining the peaks that extend to the terrain of other groups or diffuse into the valley. Clearly it is these shoulders that represent the varieties or ecotypes or even subspecies compared to the main peak; perhaps they are shaped by a different habitat or other such factors. These mounds, shoulders and valleys are likely the individuals/populations that pose serious taxonomic problems to practitioners. The morphometric terrain maps thus developed demonstrate that there may always be a certain proportion of individuals that cannot be grouped into recognized taxonomic species (e.g. species 4 and 5 in figure 6b). However, the intrinsic tendency of taxonomists to

leave nothing unclassified usually prompts them to group such individuals into neighbouring species or classify them as new species, both of which may be questionable, leading to taxonomic conflicts.

Probability based classification

In this approach, a species may be viewed as a hill or a peak—a conical or bell shaped peak with shoulders (varieties) and extensions (sub species). Note that all the individuals under the bell belong to the species though the central morph represented by the peak with the highest density located within the area of mode ± 1 SD of population, may be considered to constitute the representative morpho-types of the species. Perhaps this population, the core population, is expected to correspond to the type specimen defined by classical Linnaean taxonomists. However, all other types may also constitute members of the species but with varying levels of certainty depending upon their position within this density domain. It should be noted, however that some species/populations may show polymorphism in traits which could result in bimodal or multimodal distributions for individual species.

In general, though, one may consider that the further away an individual is from the central core, the less is its probability of representing the species. If an individual is found beyond the encompassing boundary of the species (say beyond the mode ± 4 SD), then its affinity with the species becomes uncertain. Such rare types may not even survive in due course. It is not unlikely that some very rare species described by the taxonomists (and not recovered easily by others) could in fact belong to such fluctuating populations. Thus it may be essential to define such rare individuals by the probability with which they represent known species rather than to set up exclusively new species for them. It is possible to develop protocols for defining the probabilities with which rare individuals can be assigned to known species.

From philosophy to practice: a way forward

The holy grail or physics envy: a unified species concept

Ecologists and evolutionary biologists of all hues have grappled with species concepts over the last century. The battle between the search for a unified concept (monism) versus pluralism continues to date. [Mishler and Donoghue \(1982\)](#) made an early case for pluralism, arguing that given the many different ‘species situations’, various species concepts would be more applicable and relevant. Pluralism itself has taken many forms. For example, [Ereshefsky \(1998\)](#) suggests that an organism can belong to two entities, for example, an interbreeding species and a phylogenetic species, even though these may not fully overlap with each other.

Many others advocate a less radical form of pluralism that is purely definitional in nature (Michler and Donoghue 1982). Wilson (1992), while still holding out support for the BSC, believes that multiple concepts will continue to be applied in different contexts. Reydon (2005) has criticized this, suggesting that multiple operational definitions for different contexts/clades does not constitute true pluralism. He suggests that species concepts address four independent ontological concepts, two entity and two class concepts. The two entity concepts include units that are evolving sets of populations (evolverons) or segments of a phylogenetic tree (phylons), roughly a separation of ecological/spatial (former) versus evolutionary/temporal (latter) processes. Correspondingly, the class concepts include organism kinds and evolveron kinds. Reydon (2005) assigns all species concepts to these four categories.

While the GLC (de Queiroz 1998, 2007) may indeed lay fair claim to being the ‘holy grail’ of species concepts—and it does come closer than any other at this time—there is some sleight of hand in this framing. By absorbing the various other concepts as elements/stages in this framework, the GLC in a sense, embraces rather than shuts out plurality. In that sense, the GLC is not strictly a universal species concept, but a framework for understanding multiple evolutionary (and ecological) processes in lineage divergence. In our case study, we demonstrate a practical approach to using GLC in delimitation, mapping the framework onto a phylogenetic tree (figure 3). In addition, we demonstrate that the processes are not correlated (figure 4) and can occur in different sequences and temporal scales (figure 3), even within a small clade.

Given current knowledge of the processes of evolution, particularly speciation, there is no doubt that gene flow (or lack thereof) must remain a key element. However, to avoid the circularity of cause and effect pointed out by Wallace (see Mallet 1995), gene flow must only be considered as process. The advantage here is that probabilities of gene flow can be explicitly tested for each pair or set of closely related populations/lineages (as demonstrated in our case study). In fact, Mallet’s (1995) definition focusses on what might be considered as the central tenet of the GLC, the creation of genotype clusters by various biological processes. While the definition differs from phylogenetic concepts only in that it uses population genetics as a heuristic rather than a phylogenetic tree, this explicitly allows gene flow where a phylogenetic framework assumes to some degree that gene flow does not occur. Dover (1995) points out that this concept still views species as far too static. We therefore recommend a more ‘inclusive’ approach, where genotypic clusters exist, but individual organisms have a finite probability of contributing to more than one such cluster (figure 1), which can be estimated using recent molecular genetic, particularly genomic, tools.

Rings, complexes and hybrids

The ‘inclusive’ approach provides both an understanding as well as a clear visualization (figure 1) of traditionally difficult species problems such as ring species (Irwin et al. 2001) and species complexes. Further, it provides the clearest framework for acknowledging the role of hybridization in the history of lineages.

Hybridization was once considered an anomaly to the extent that hybrids were not considered worthy of conservation, as in the case of North American wolves (Jenks and Wayne 1992; Gittleman and Pimm 1991; but see Wilson et al. 2009; Nowak 1995). Hybridization is now known to be much more widespread in vertebrates than earlier believed and has gained importance in evolutionary biology (Mallet 2005). Recent work has suggested a strong role for hybridization and gene introgression across vertebrate taxa including in the adaptive radiation of fishes (cichlids, Meier et al. 2017) and birds (Darwin’s finches, Grant and Grant 2015). Hybrids between species belonging to different genera (that may have diverged over 30 mya) have been recorded from sea turtles (Karl et al. 1995). Other than wolves, hybridization has been shown to have influenced the evolutionary history of mammals including the European bison (Soubrier et al. 2016). In bears, there is gene flow not just between closely related species like brown and polar bears, but also with the more distant Asiatic black bear (Kumar et al. 2017). Ackermann et al. (2016) suggest that hybridization has been an ‘essential creative force’ in the evolution of modern humans.

We suggest that our ‘inclusive’ framework allows the explicit incorporation of probabilities of gene flow between ‘clusters’ that may be either ‘populations’ or ‘species’, a distinction that rests on other axes.

The ontological problem

For nearly 2000 years, species were considered as natural kinds with essences (Ereshefsky 2016). This essentialism has been abandoned as it is not possible to conceive of any trait (essence) that all members of a species possess and is unique to it i.e., no other species possesses it (Hull 1965). Ghiselin (1974) suggested that species should be treated as individuals rather than classes. There has been considerable debate over this, with further suggestions to treat species as homeostatic property clusters and as fuzzy sets.

Irrespective of the merits of these propositions, what seems to be true is that difference between concepts and criteria (sensu de Queiroz) may have arisen due to the confusion between treating species as individuals/entities and classes. All true concepts, including de Queiroz’s GLC treat species as individuals, though the GLC does so the least when applied to the fuzzy boundaries of lineage divergence

(particularly after), when some individuals may contribute to multiple evolving lineages. Criteria must treat species as a class, where membership is conferred based on the possession of certain traits and characters.

We recommend a probabilistic framework for understanding species when treated as individuals, and for delimiting species when treated as classes, thereby bridging (if not unifying) philosophy and practice. These do not replace the GLC, but provide an approach to further understanding and delimiting within the GLC. In concept, an individual organism has a finite probability of belonging to two species (as defined by a clustering approach). In categorization (typically using traits), the probability refers to the uncertainty of assignment to a particular cluster which can be estimated using the morphometric terrain approach (figure 6) among others.

A practical approach to discovery and delimitation

While arguments about concepts have raged on, there have also been both conceptual and methodological discussions about species delimitation (de Queiroz 2007). The idea of ‘good’ species has become embedded in the discourse over the last few decades, with lineages that are clearly separated along morphological and genetic axes (with clear evidence of reproductive isolation) representing the ideal end of the spectrum. Following Mayr and Ashlock (1991), ‘good’ sympatric species (i.e., those with morphological separation and reproductive isolation) have been used as a yardstick for setting a genetic cut-off for other closely related species (see Baker *et al.* 1995 on kiwis, and Dutton *et al.* 1996 on sea turtles, for early examples). When the degree of difference between ‘good’ species has been used to delimit other species within that clade, as in the above examples, the approach has been relatively useful. However, the idea of cut-offs has also led to more simplistic propositions such as ‘barcoding’ where a standard genetic distance on a particular gene has been recommended for the designation of species.

Indeed, the proposal by de Queiroz (1998) that different processes occur at different stages of lineage divergence provides a evolutionary scaffolding for understanding species, while at the same time incorporating other axes such as morphology, behaviour and geography. This offers a more nuanced understanding of both the process of speciation and species as units. Cryptic species and intraspecific polymorphism provide sufficient evidence that genetic and morphological divergence are not correlated. In fact, the rate of change along all axes of variation are not correlated either between or even within clades, and equally importantly do not necessarily occur in the same sequence (see figure 3). In practice, the challenge lies in using this understanding to discover and delimit species in lesser known clades, especially in the context of biodiversity inventory and conservation.

Here, we provide a geography based approach (borrowing Wallacean ideas) for resolving the species problem of discovery and delimitation (the Linnaean problem). We highlight the need to test for geographical range overlap in sister species pairs, which can either be allopatric, sympatric or parapatric. The degree of separation of populations/species can be quantified along a continuum by modelling geographic separation combined with species traits. This contributes process related insights into species delimitation and allows one to expand on this further to test for divergence given gene flow and other processes. Most critically, this calls for incorporating both geography and traits in the sampling design. Most studies that result in species delimitation (either by design or accident) do not explicitly incorporate one or the other, especially geography.

Among vertebrates, an allopatric mode of speciation is dominant among birds (Phillimore *et al.* 2008) but equivocal in other taxa including mammals (Coyne and Orr 2004; Fitzpatrick and Turelli 2006). In the case of invertebrates, different patterns have been found; for example, in the butterfly genus *Ithomiola*, parapatry takes precedence (Hall 2005). In contrast, a high percentage of allopatry is seen in the gastropod genus *Nerita* (Frey 2010). The importance of allopatry can vary within a clade over time, with an increase of other modes of speciation at deeper nodes (Vijayakumar *et al.* 2016).

In the case of bush frogs, allopatry was found to be the dominant mode among the shallow divergent lineages. The sporadic cases of sympatric sister lineages, in contrast to shallow allopatric sisters, showed high levels of divergence in morphology and other axes rendering their recognition less problematic. Clades that contain sympatric sister lineages that are cryptic can be hard to recognize and delimit presenting some of the most difficult challenges in the discovery and delimitation of species. Among such cases, examining divergence along multiple axes is critical. Our morphometric terrain approach offers a trait based approach to visualizing and understanding these relationships, especially at shallow levels of divergence.

Conclusion

Discovery and delimiting species in a clear and consistent manner is important, not just for studies of ecology, evolution and biogeography, but for conservation as well. The importance of correctly classifying species in a conservation context has been emphasized for a variety of reasons. For example, endangered species may be denied legal protection and resources may be wasted on abundant species. In management, incorrectly diagnosed species may be hybridized with other species and populations that could be used to improve fitness of inbred populations may be overlooked. As a consequence, unrecognized endangered species may become extinct (which

is likely already occurring in lesser known taxa). Hence, the contentious philosophical issue of what species are has practical consequences for the future of biodiversity conservation.

In summary, we believe that our approaches offer a practical and process oriented way forward for species delimitation, and propose that a probabilistic approach helps bridge philosophy and practice. We suggest that concepts which consider species as ‘individuals’, especially a general framework like GLC, provide the best current understanding of what species are. On the other hand, treating species as classes with ‘inclusive’ probabilistic approaches for criteria and categorization offer the best heuristic tools for delimitation.

Acknowledgements

The authors, like most other evolutionary biologists around the world, have been pondering the species question throughout their respective careers. KS started his musings in a not-so-quiet corner of the Wildlife Institute of India, aided by spirited conversations with Jagdish Krishnaswamy, M. S. Chaitra and Meera Anna Oommen, and inspired by a ‘scruffy dog’ that became a symbol of the ‘good’ vs ‘bad’ species conundrum. SPV got embroiled in it when he joined KS’s lab as a graduate student and became obsessed with questions regarding the origin and diversification of bush frogs and other herpetofauna in the Western Ghats. KNG has got into the habit of attempting to solve intractable problems over several decades, abetted by his colleagues R. Uma Shaanker and K. Chandrashekara. We would like to thank all the reviewers and editors for their constructive comments on the manuscript.

References

- Ackermann R. R., Mackay A. and Arnold M. L. 2016 The hybrid origin of “modern” humans. *Evol. Biol.* **43**, 1–11.
- Avise J. C. and Johns G. C. 1999 Proposal for a standardized temporal scheme of biological classification for extant species. *Proc. Natl. Acad. Sci. USA* **96**, 7358–7363.
- Baker A. J., Daugherty C. H., Colbourne R. and McLennan J. L. 1995 Flightless brown kiwis of New Zealand possess extremely subdivided population structure and cryptic species like small mammals. *Proc. Natl. Acad. Sci. USA* **92**, 8254–8258.
- Boyd R. 1999 Homeostasis, species, and higher taxa. In *Species: New Interdisciplinary Essays* (ed. R. Wilson), pp. 141–185. MIT Press, Cambridge, USA.
- Chandrasekhara K., Kumar A. R. V. and Ganeshaiah K. N. 1998 *Morphometric terrain as a technique to demarcate the species limits: a study using dung beetles*. Project report submitted to Department of Science and Technology, GOI New Delhi, India.
- Coyne J. A. and Orr H. A. 2004 *Speciation*. Sinauer Associates, Sunderland, USA.
- Cracraft J. L. 1983 Species concepts and speciation analysis. In *Current Ornithology* (ed. R. F. Johnston), pp. 159–187. Plenum Press, New York, USA.
- Darwin C. 1859 *On the Origin of Species by Means of Natural Selection, Or, the Preservation of Favoured Races in the Struggle for Life*. J. Murray, London.
- Darwin F. 1887 *The life and letters of Charles Darwin, including an autobiographical chapter*. John Murray, London, UK.
- de Queiroz K. 1998 The General Lineage Concept of species, species criteria, and the process of speciation. In *Endless forms: species and speciation* (ed. D. J. Howard and S. H. Berlocher), pp. 57–75. Oxford University Press, Oxford, UK.
- de Queiroz K. 2007 Species concepts and species delimitation. *Syst. Biol.* **56**, 879–886.
- de Queiroz K. 2011 Branches in the lines of descent: Charles Darwin and the evolution of the species concept. *Biol. J. Linn. Soc.* **103**, 19–35.
- Dobzhansky T. 1937 *Genetics and the origin of species*. Columbia University Press, New York, USA.
- Dobzhansky T. 1950 Mendelian populations and their evolution. *Am. Nat.* **84**, 401–418.
- Donoghue M. J. 1985 A critique of the biological species concept and recommendations for a phylogenetic alternative. *Bryologist* **Oct**, 172–181.
- Dover, G. 1995 A species definition: a functional approach. *Trends Evol. Biol.* **10**, 489–490.
- Dutton P. H., Davis S. K., Guerra T. and Owens D. 1996 Molecular phylogeny for marine turtles based on sequences of the ND4-leucine tRNA and control regions of mitochondrial DNA. *Mol. Phylogenet. Evol.* **5**, 511–521.
- Eldredge N. and Cracraft J. L. 1980 *Phylogenetic patterns and evolutionary process*. Columbia University Press, New York, USA.
- Ereshefsky M. 1998 Species pluralism and anti-realism. *Philosophy of Science* **65**, 103–120.
- Ereshefsky M. 2007 Foundational issues concerning taxa and taxon names. *Syst. Biol.* **56**, 295–301.
- Ereshefsky M. 2016 Species. In *The Stanford encyclopedia of philosophy* (summer 2016 edition) (ed. E. N. Zalta) (<https://plato.stanford.edu/archives/sum2016/entries/species/>).
- Fitzpatrick B. M. and Turelli M. 2006 The geography of mammalian speciation: mixed signals from phylogenies and range maps. *Evolution* **60**, 601–615.
- Fouquet A., Vences M., Salducci M. D., Meyer A., Marty C., Blanc M. and Gilles A. 2007 Revealing cryptic diversity using molecular phylogenetics and phylogeography in frogs of the *Scinax ruber* and *Rhinella margaritifera* species groups. *Mol. Phylogenet. Evol.* **43**, 567–582.
- Frey M. A. 2010 The relative importance of geography and ecology in species diversification: evidence from a tropical marine intertidal snail (*Nerita*). *J. Biogeogr.* **37**, 1515–1528.
- Funk W. C., Angulo A., Caldwell J. P., Ryan M. J. and Cannatella D. C. 2008 Comparison of morphology and calls of two cryptic species of *Physalaemus* (Anura: Leiuperidae). *Herpetologica* **64**, 290–304.
- Ghiselin M. T. 1974 A radical solution to the species problem. *Syst. Biol.* **23**, 536–544.
- Gittleman J. L. and Pimm S. L. 1991 Crying wolf in North America. *Nature* **351**, 524–525.
- Grant P. R. and Grant B. R. 2015 Introgressive hybridization and natural selection in Darwin’s finches. *Biol. J. Linn. Soc.* **117**, 812–822.
- Hall J. P. W. 2005 Montane speciation patterns in Ithomiola butterflies (Lepidoptera: Riodinidae): are they consistently moving up in the world? *Proc. R. Soc. London, Ser. B* **272**, 2457–2466.
- Hamilton W. D. 1964 The genetical evolution of social behaviour. II. *J. Theor. Biol.* **7**, 17–52.
- Hennig W. 1966 *Phylogenetic systematics* (ed. D. D. Davis and R. Zangerl) University of Illinois Press, Urbana-Champaign, USA.
- Hull D. 1965 The effect of essentialism on taxonomy: two thousand years of stasis (I). *Brit. J. Philos. Sci.* **15**, 314–326.
- Hull D. 1978 A matter of individuality. *Philos. Sci.* **45**, 335–360.

- Irwin D. E., Irwin J. H. and Price T. D. 2001 Ring species as bridges between microevolution and speciation. *Genetica* **112–113**, 223–243.
- Jenks S. M. and Wayne R. K. 1992 Problems and policy for species threatened by hybridization: the red wolf as a case study. In *Wildlife 2001: populations* (ed. D. R. McCullough and R. H. Barrett), pp. 237–251. Springer Netherlands, New York, USA.
- Joppa L. N., Roberts D. L., Myers N. and Pimm S. L. 2011 Biodiversity hotspots house most undiscovered plant species. *Proc. Natl. Acad. Sci. USA* **108**, 13171–13176.
- Karl S. A. and Bowen B. W. 1999 Evolutionary significant units versus Geopolitical taxonomy: molecular systematics of and endangered sea turtle (Genus *Chelonia*). *Conserv. Biol.* **13**, 990–999.
- Karl S. A., Bowen B. W. and Avise J. C. 1995 Hybridization among the ancient mariners: characterization of marine turtle hybrids with molecular genetic assays. *J. Hered.* **86**, 262–268.
- King M. 1993 *Species evolution: the role of chromosome change*. Cambridge University Press, Cambridge, USA.
- Kitcher P. 1984 Species. *Philos. Sci.* **51**, 308–333.
- Kumar V., Lammers F., Bidon T., Pfenninger M., Kolter L., Nilsson M. A. and Janke A. 2017 The evolutionary history of bears is characterized by gene flow across species. *Sci. Rep.* **7** (doi:10.1038/srep46487).
- Linkem C. W., Hesed K. M., Diesmos A. C. and Brown R. M. 2010 Species boundaries and cryptic lineage diversity in a Philippine forest skink complex (Reptilia; Squamata; Scincidae: Lygosominae). *Mol. Phylogenet. Evol.* **56**, 572–585.
- Mallet J. 2005 Hybridization as an invasion of the genome. *Trends Ecol. Evol.* **20**, 229–237.
- Mallet J. 1995 A species definition for the modern synthesis. *Trends Ecol. Evol.* **10**, 294–299.
- Mayden R. L. 1997 A hierarchy of species concepts: the denouement in the saga of the species problem. In *Species: the units of diversity* (ed. M. F. Claridge, H. A. Dawah and M. R. Wilson), pp. 381–423. Chapman and Hall, New York, USA.
- Mayr E. 1942 *Systematics and the origin of species*. Columbia University Press, New York, USA.
- Mayr E. 1963 *Animal species and evolution*. Harvard University Press, Massachusetts, USA.
- Mayr E. and Ashlock P. D. 1991 *Principles of systematic zoology*, 2nd edition. McGraw-Hill, New York, USA.
- Meier J. I., Marques D. A., Mwaiko S., Wagner C. E., Excoffier L. and Seehausen O. 2017 Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nat. Commun.* **8**, 14363.
- Michener C. D. 1970 Diverse approaches to systematics. *Evol. Biol.* **4**, 1–38.
- Millikan R. 1999 Historical kinds and the “special sciences”. *Philos. Stud.* **95**, 45–65.
- Mishler B. D. and Donoghue M. J. 1982 Species concepts: a case for pluralism. *Syst. Zool.* **31**, 491–503.
- Mishler B. D. 1985 The morphological, developmental, and phylogenetic basis of species concepts in bryophytes. *Bryologist* **Oct**, 207–14.
- Myers N., Mittermeier R. A., Mittermeier C. G., da Fonseca G. A. and Kent J. 2000 Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858.
- Nelson G. J. and Platnick N. I. 1981 *Systematics and biogeography: cladistics and vicariance*. Columbia University Press, New York, USA.
- Nixon K. C. and Wheeler Q. D. 1990 An amplification of the phylogenetic species concept. *Cladistics* **6**, 211–223.
- Nowak R. 1995 Hybridization: the double-edged threat. *Canid News* **3**, 2–6.
- Paterson H. E. H. 1985 The recognition concept of species. In *Species and speciation* (ed. E. S. Vrba), pp. 21–29. Transvaal Museum, Pretoria, South Africa.
- Phillimore A. B., Orme C. D. L., Thomas G. H., Blackburn T. M., Bennett P. M., Gaston K. J. and Owens I. P. F. 2008 Sympatric speciation in birds is rare: Insights from range data and simulations. *Am. Nat.* **171**, 646–657.
- Reydon T. A. C. 2005 On the nature of the species problem and the four meanings of ‘species’. *Stud. Hist. Philos. Biol. Biomed. Sci.* **36**, 135–158.
- Ridley M. 1993 *Evolution*. Blackwell Scientific, Boston, USA.
- Rosen D. E. 1979 Fishes from the uplands and intermontane basins of Guatemala: revisionary studies and comparative biogeography. *Bull. Am. Mus. Nat. Hist.* **162**, 267–376.
- Shanker K. 2001 A review of species concepts: ideas for a new concept and implications for the green-black turtle debate. Twenty first Annual Symposium on Sea Turtle Biology and Conservation. Philadelphia, USA.
- Simpson G. G. 1951 The species concept. *Evolution* **5**, 285–298.
- Simpson G. G. 1961 *Principles of animal taxonomy*. Columbia University Press, New York, USA.
- Sokal R. R. and Sneath P. H. A. 1963 *Principles of numerical taxonomy*. Freeman, San Francisco, USA.
- Sokal R. R. and Crovello T. J. 1970 The biological species concept: a critical evaluation. *Am. Nat.* **104**, 127–153.
- Soubrier J., Gower G., Chen K., Richards S. M., Llamas B., Mitchell K. J. et al. 2016 Early cave art and ancient DNA record the origin of European bison. *Nat. Commun.* **7** (doi:10.1038/ncomms13158).
- Templeton A. R. 1989 The meaning of species and speciation: a genetic perspective. In: **Speciation and its Consequences** (eds D. Otte and J. A. Endler), pp. 3–27. Sinauer, Sunderland, USA.
- Van Valen L. 1976 Ecological species, multispecies, and oaks. *Taxon* **25**, 233–239.
- Vences M., Thomas M., Bonett R. M. and Vieites D. R. 2005 Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philos. Trans. R. Soc. London, Ser. B* **360**, 1859–1868.
- Vijayakumar S. P., Dinesh K. P., Prabhu M. V. and Shanker K. 2014 Lineage delimitation and description of nine new species of bush frogs (Anura: *Raorchestes*, Rhacophoridae) from the Western Ghats Escarpment. *Zootaxa* **3893**, 451–488.
- Vijayakumar S. P., Menezes R. C., Jayarajan A. and Shanker K. 2016 Glaciations, gradients, and geography: multiple drivers of diversification of bush frogs in the Western Ghats Escarpment. *Proc. R. Soc. London, Ser. B* **283**, 20161011.
- Wallace A. R. 1865 On the phenomena of variation and geographical distribution as illustrated by the Papilionidae of the Malayan region. *Trans. Linn. Soc. Lond.* **25**, 1–71.
- Whittaker R. J., Araújo M. B., Paul J., Ladle R. J., Watson J. E. and Willis K. J. 2005 Conservation biogeography: assessment and prospect. *Divers. Distrib.* **11**, 3–23.
- Wiens J. J. and Graham C. H. 2005 Niche conservatism: integrating evolution, ecology, and conservation biology. *Annu. Rev. Ecol. Syst.* **36**, 519–539.
- Wiens J. J. and Penkrot T. A. 2002 Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). *Syst. Biol.* **51**, 69–91.
- Wiley E. O. 1978 The evolutionary species concept reconsidered. *Syst. Zool.* **27**, 17–26.
- Wilkins J. S. 2006a Species, kinds, and evolution. *Rep. Cent. Natl. Sci. Edu.* **26**, 36–45.
- Wilkins J. S. 2006b A list of 26 species ‘concepts’. *Sci. Blogs* (<http://scienceblogs.com/evolvingthoughts/2006/10/01/a-list-of-26-species-concepts/>, accessed 5 January 2017).
- Wilkins J. S. 2009 *Species: a history of the idea*. University of California Press, Berkeley, USA.
- Wilson E. O. 1992 *The diversity of life*. Penguin Books, London.

- Wilson R. 1999 Realism, essence, and kind: resuscitating species essentialism? In *Species: new interdisciplinary studies* (ed. R. Wilson). MIT Press Cambridge, Cambridge, USA.
- Wilson P. J., Grewal S. K., Mallory F. F. and White B. N. 2009 Genetic characterization of hybrid wolves across Ontario. *J. Hered.* **100**, 80–89.
- Young D. 2007 *The discovery of evolution*. Cambridge University Press, Cambridge, UK.
- Zachos F. E. 2016 *Species Concepts in Biology: Historical Development, Theoretical Foundations and Practical Relevance*. Springer.

Corresponding editor: T. N. C. VIDYA