1

Introduction

ALEXANDRE K. MONRO

Introduction

This Systematics Association Special Volume is the result of a symposium titled, 'Cryptic Taxa – Artefact of Classification or Evolutionary Phenomena?' held on 17 June as part of the association's 10th Biennial Meeting in 2019. The symposium comprised five presentations, Torsten Struck, Paul Williams, Matt Lavin, Mark Wilkinson and Jim Labisko. For the purposes of this volume, we also invited contributions by Cene Fišer and Klemen Koselj, Alexander Martynov and Tatiana Korshunova, Simon Mayo, Richard Bateman, Marta Álvarez-Presas, and Pablo Muñoz, with the aim of providing a broader perspective on the subject, not only with respect to theory and practice but also with respect to the organisms that they work on.

My motivation for organising the meeting was scepticism. Scepticism that stemmed from the feeling that what was being observed were species whose evolutionary history had resulted in strong genetic partitioning, and that in the absence of a universally accepted species concept, it was arbitrary to designate the rank of species based on DNA alone for organisms where other sources of observations are available. Second, as a taxonomist tasked with producing field guides, identification keys, and identifying biological collections in herbaria, I did not welcome the prospect of taxa that are impossible to identify without access to a DNA laboratory and funds, the latter being difficult to access even in world-leading biological collections institutions.

The first time that I encountered the concept of cryptic species was in the mid-1990s, at which time they were novel and controversial. In the last decade, however, with the wide availability of DNA sequence observations and improved techniques for extracting DNA from biological collections, the description of cryptic species is becoming commonplace (Figure 1.1), including of hitherto well-known species: For example, the Tapanuli Orangutan (*Pongo tapanuliensis,* Nater et al. 2017), the Baltic flounder (*Platichthys solemdali,* Momigliano et al. 2018), or the Kabomani Tapir (*Tapirus kabomani,* Cozzuol et al.
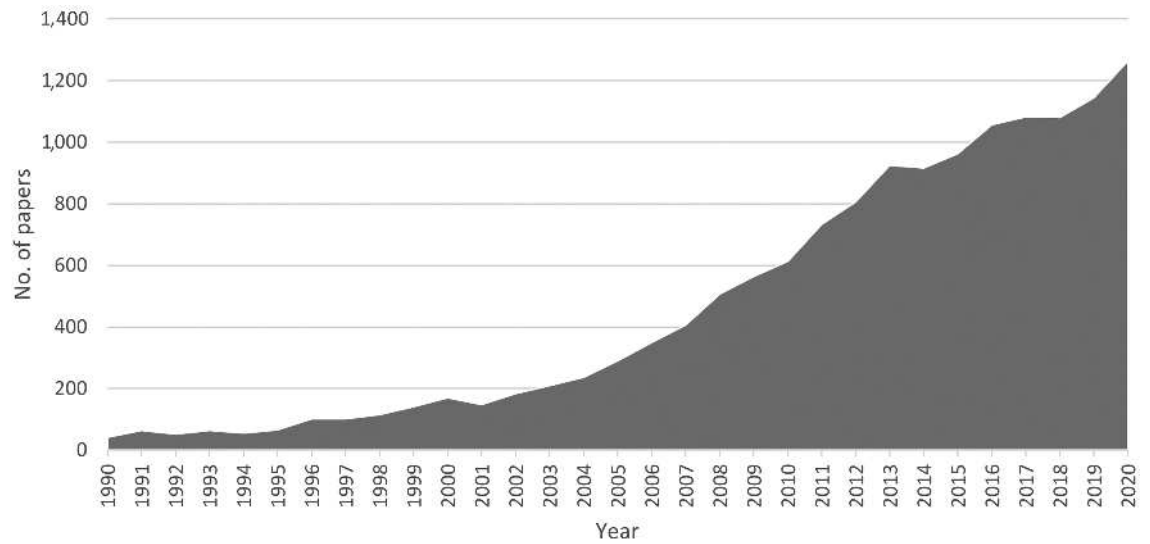
Figure 1.1 Frequency of papers with 'cryptic' and 'species' in the title (1990–2020).
Source: Scopus search 'TITLE-ABS-KEY (cryptic AND & AND species)', 2021 (undertaken 14 June 2021).

2013); but also amongst many less well-known groups of organisms, for example, sponges (Xavier et al. 2010), marine interstitial ghost-worms (Cerca et al. 2020), copepods (Fišer et al. 2015), roundworms (Armenteros et al. 2014), flatworms (Álvarez-Presas et al. 2015; Leria et al. 2020), malaria parasites (alveolates, Bensch et al. 2004), sea-slugs (Korshunova et al. 2020), rotifers (Gabaldón et al. 2015; Mills et al. 2017), cod icefish (Dornburg et al. 2016), lizards (Leavitt et al. 2007), ferns (Bauret et al. 2017), mosses (McDaniel and Shaw 2003), fungi (Muggia et al. 2015), and bumblebees (Williams et al. 2016).

Preparing for the symposium, I began to realise that the notion of – and process of discovery for – cryptic species touches the heart of several major debates in biology, including 'what are species?', 'how should we recognise them?', the notion of punctuated equilibria, and that of morphological stasis in the fossil record. In addition, in the midst of a biodiversity crisis (Koh et al. 2004) the phenomenon of cryptic species indicates that there may be a greater diversity of evolutionary lineages in need of conservation than has been suggested by morphology alone (Funk et al. 2012; Chapters 8–11), implying the need for a more nuanced approach to species conservation (Carroll et al. 2014).

Rather than simply being a distracting artefact of new sequencing technologies, phylogenetic techniques, and opportunism, any consideration of the notion of cryptic species exposes a fundamental and well-documented weakness of contemporary systematic biology: that we do not yet have the conceptual framework or the quality and breadth of observations to be able to say what a species is, and, as a result, to assert crypsis in relation to one. Striving to resolve some of the debates around cryptic species might not only provide the tools and framework to answer some major questions in biology but also make taxon delimitation and the documentation of diversity a more rigorous and useful scientific undertaking.

This book is organised to present overviews of cryptic (sibling) species in the context of species delimitation and the taxonomic method (Chapters 2–4), followed by reviews of cryptic species concepts and their value to evolutionary biology (Chapters 5–8) and then some case examples from diverse groups of organisms (Chapters 3, 5, 9–11).

1.1 Were There Cryptic Species before Darwin?: Cryptic Species and the Concepts of Species

Cryptic species are, logically, by-products of the application of a species concept to group a set of individuals into units referred to as 'species'. The notion of species as units of diversity predates classical Candollean, Linnean, and Aristotelian attempts to classify diversity and can be found in all human cultures (Atran 1998; Berlin 1973, 1992; Berlin et al. 1974; Bulmer et al. 1968; Coyne 2004; Diamond 1966; Ludwig 2017; Majnep and Bulmer 1977; Mayr 1963; Slater 2015). It also likely occurs, at a biological level at least, amongst non-humans (Poelstra et al. 2014; Robinson et al. 2015).

The term 'species' originated in the fourteenth century (Online Etymological Dictionary 1993). It denotes, 'appearance, form, kind' (Oxford University Press 1993) and as such is congruent with morphological species concepts. The notion of species as real entities ('natural things') that existed in nature rather than defined by humans dates back at least

to Locke (Locke 1689, see also Mayo (Chapter 2)). The notion that species are the product of an evolutionary process is most closely associated with Darwin, who emphasised such a relationship in the title of his epoch-defining work, *The Origin of Species* (Darwin 1859). Since Darwin first linked the phenomenon of species to that of evolution, most systematic biologists equate species with separately evolving lineages – equivalent to branches of the 'Tree of Life' (e.g. De Queiroz 2007; Padial and De la Riva 2021), with the logical consequence that the basis and process of species delimitation centres on assigning individuals to a phylogenetic lineage (see Chapter 2 for context). Freudenstein et al. (2016) and Chapter 8 argue, however, that lineage divergence alone is not sufficient to delimit species. Templeton (1989), under his 'Cohesion Species Concept', rather than focussing on isolation or divergence, applies explicitly evolutionary criteria to define species as, 'the most inclusive group of organisms having the potential for genetic and/or demographic exchangeability' (Templeton 1989: 181).

For practical and academic reasons (see Chapter 2), there are now probably as many ways to assign an individual to a species as there are taxonomists doing so, a situation referred to as the 'species problem' (Mayden 1997). Compounding this, even if there was agreement on the criteria for delimiting species, we would rarely have the resources to do so confidently, either with respect to the number of populations sampled or with respect to the observations made from each. The reality is that the incidental evidence (Padial and De la Riva 2021) or operational criteria (De Queiroz 2007) used in the delimitation of the vast majority of species comprise just two classes of observations, morphological and/or molecular (DNA sequence), from a very small sample of individuals (Chapters 3, 4, and 9). Morphological evidence formed the basis of species delimitation for all groups of organisms for over 250 years, albeit mostly from the very small, arbitrary, and biased sample of characters preserved in biological collections and mostly interpreted outside of any explicit hypothesis of homology or species. For the last 25 years, the sequence of nucleotides in DNA has provided an independent class of observations for which increasingly robust statistical analyses have been developed, incorporating complex mathematical inferences for evolutionary phenomena (e.g. coalescent, Bayesian, substitution models), to delimit putative evolutionary lineages. DNA sequence observation, however, also suffers from very small sample sizes, both with respect to the proportion of populations sampled and, to a lesser and varying extent, to the proportion of the genome sampled. As implied by allopatry, geography, sometimes in association with ecological niche, is also a source of observations for the delimitation of species and subspecies (Darwin 1859; Jordan 1905; Rensch 1938) across many species concepts and frequently underpins the decision to delimit new taxa. With the exception of the rank of subspecies, however, geographical observations are rarely applied formally for the purposes of species delimitation but are generally held to be confirming factors (Davis and Heywood 1963) or 'soft characters' (see Chapter 5).

Morphological, DNA sequence, and geographical observations are just three out of an increasing number that could be available for taxon delimitation. For example, developmental (ontogenetic), physiological, transcriptomic, proteomic, behavioural, ecological niche, ecological network, immunological, biochemical, and holobiome could all provide

observation useful in testing hypotheses of species, but they have been largely ignored for the process of species delimitation (Chapters 3 and 4). Both Bateman (Chapter 3) and Martynov and Korshunova (Chapter 4) suggest that no species would likely be considered cryptic were there adequate sampling of populations and morphology together with the inclusion of additional sources of observations, such as ontogenetic (Chapter 4), chemical, electrical, magnetic, sensory, ecological (Chapters 7 and 8), or were morphological characters to be observed and evaluated adequately (Chapter 3) from an effective sample of populations (Chapters 3 and 4).

Within the context of crypsis, discordance between DNA sequence and morphological estimates of divergence can result in two phenomena, (1) that of cryptic species, whereby DNA sequence observations suggest lineage divergence equivalent to a distinct species but morphological observations do not, or (2) of polymorphic species, where morphologically distinct species are suggested by DNA sequence observations to represent a single lineage (Chapter 8; Dexter et al. 2010). This latter group of species has been the focus of far less research.

It is the lack of congruence between the, arguably superficial (Chapter 3) sampling of morphological and DNA sequence observations and their use as incidental evidence that has fuelled a renewed interest in and description of cryptic species. Basically, DNA sequence observations are suggesting greater or lesser lineage divergence than morphological observations and where greater, then this is being used to propose morphologically cryptic lineages at the rank of species. This lack of congruence could be attributed to the identification of early diverging lineages, equivalent to De Queiroz's 'gray zone' of speciation (De Queiroz 2007: Fig. 1; Chapter 7). Struck and Cerca (Chapter 6) and Muñoz et al. (Chapter 5) suggest, however, that this is not the usual case, with crypsis being identified in lineages up to 140 million years old (Chapter 6). In order to prevent early-diverging lineages from being designated as cryptic species, Struck (Struck et al. 2018a, b; Chapter 6) proposes that one should explicitly show that the species are morphologically more similar to each other than would be expected given the time that has passed since their last common ancestor. This is something that is possible to establish, with some degree of error, using DNA sequence observations and/or fossils.

Given the limitations of sample size and bias in the taxonomic process, it could be argued that the current state of knowledge on species can best be described as superficial or tokenistic, as suggested by Bateman (Chapter 3). As a result, we do not have the necessary observations to formulate or apply universal species concepts. More useful is a less mechanistic definition such as Templeton's Cohesion Species Concept (Templeton 1989), or the flexibility to delimit species which conspecific genetic samples resolve as paraphyletic (Freudenstein et al. 2016; Muñoz-Rodríguez et al. 2019; Pennington and Lavin 2016; Chapter 8) and a more rigorous circumscription (see Chapter 3), applied with the recognition of the limitations of our sampling and methods, may, therefore, be more useful for the purposes of exploring evolution, but also for the establishment of a stable classification of life on earth. Within such a context, cryptic species could be viewed as nodes (Chapters 3, 5–8) for which there is evidence of lineage divergence but not of morphological change.

1.2 Cryptic Species, Morphological Stasis: Artefacts of Taxonomic Method

The fact that the definition of cryptic species is problematic does not mean that the phenomena it highlights are not important. In fact, it may be that the use of the term 'cryptic species' within a jungle of problematic species concepts has prevented a key phenomenon, morphological stasis, from getting the research focus it should have. With hindsight, the great debate over the tempo of evolution (punctuated equilibria, gradualism) occurred prematurely, prior to the 'molecular revolution' that has enabled the pairing of the palaeontological perspective of morphology with evaluations of lineage divergence.

There are several reasons for the discordance between DNA sequences and morphological observations. Some of these can be considered experimental error, whereby prior to the application of DNA sequence observations, the delimitation of a species was based on too few morphological observations. For example, the bumblebee, *Bombus kluanensis* (Williams et al. 2016; Chapter 8), was initially recognised from a coalescent analysis of a small subsample of the mitochondrial genome (COX1, Williams et al. 2019). This triggered a morphological re-evaluation of the biological collections that recovered diagnostic morphological character states. Another example is the Tapanuli Orangutan (*Pongo tapanuliensis*, Nater et al. 2017), for which cranio-mandibular and dental characters (albeit from a single individual) were identified following analyses of whole mitochondrial genomes. In both cases, the species turned out not to be cryptic, as morphological differences were observable. They had just not been detected earlier.

There are, however, many cases where morphological crypsis is confirmed and five evolutionary processes can be proposed to account for these (Chapters 6, 8–10): (I) recent lineage divergence that has not yet resulted in morphological divergence, (II) parallel or (III) convergent morphological evolution, (IV) morphological stasis, or (V) introgression. Evidence for all five has been observed for cryptic species (Chapters 6 and 7). These are all, however, distinct, testable, evolutionary phenomena that are not best served by being combined or obscured under the term 'cryptic species' (Chapter 6). Of these phenomena, recent lineage divergence and convergence have been the subject of substantial research effort by evolutionary and population biologists. Parallelism, effectively representing convergence within closely related lineages, has also been the subject of some research from speleo- (Gross 2016; Khalik et al. 2020; Powers et al. 2020) and hydrothermal vent biologists (Yuan et al. 2020).

Morphological stasis, however, remains relatively little studied outside palaeontology (Gingerich 2019), despite being a major feature of the paleontological record (Gould 2002; Stanley 1979) and presumably of evolution. In addition, the explanations for stasis have been controversial (Davis et al. 2014), with both genetic-developmental constraints and stabilising selection being invoked (Charlesworth and Lande 1982; Davis et al. 2014; Estes and Arnold 2007; Raff 1996; Smith 1981). Understanding the causes and implications of morphological stasis in evolution could therefore provide a productive research focus for which cryptic species would be key study organisms/scenarios. It is for this reason that the definition, terminology, and methods used in the recognition of such taxa are important.

Beyond morphological stasis, cryptic species are probably best referred to using terms that highlight the evolutionary phenomena more clearly, such as 'convergence', 'parallelism', or a term explicitly indicative of lineage divergence, such as 'ochlospecies' (Chapter 8). Struck's proposal – that the term 'cryptic species' should be restricted to morphological stasis, defined as lineages morphologically more similar to each other than one would be expected given time since lineage divergence – is useful and pragmatic as highlighted earlier in the chapter.

Delimiting species solely on lineage divergence is pragmatic where there is an abundance of DNA sequence observations and a paucity of other observations. It does, however, invoke operational criteria as definitional concepts, a practice subject to substantial epistemological criticism (De Queiroz 2007). Perhaps more importantly, incongruence between morphological and lineage divergence, whilst identifying important evolutionary phenomena, should not automatically be translated into taxonomic actions. Rather, the identification of incongruence between DNA sequence and morphological observations should be the starting point for hypothesis-testing and the generation of observations from additional sources. For example, the use of geographical, ontogenetic, physiological, behavioural, ecological, and chemical observations. Where these observations corroborate the DNA sequence observations then there is sense in recognising the metapopulation or lineage using a taxonomic rank.

It could be argued that adopting such an approach effectively weights morphological observations over DNA sequence ones. This I think is the reality of a taxonomy designed by and for people, and for a multiplicity of uses.

1.3 Taxonomy Is Not Just About Documenting Evolution

Taxonomy concerns the construction of classifications in general. Here we refer to that long-term enterprise undertaken by biological scientists that results in a classification and identification system founded on species taxa, named according to international codes of nomenclature. This framework, which is intended to encompass all organic life, serves the needs of a range of audiences and applications. Biological classification takes place within a constrained resource, both with respect to observations but also to the number of people delivering it and the narrow window of time in which it is taking place. Scientific practitioners are heavily influenced by evolutionary theory and for many this demands an assumed link between species as units of both diversity and the evolutionary process, but for other users, classification serves as a tool and surrogate for predicting properties (traits) related to usefulness, for measuring biological diversity, for predicting and mitigating the impacts of human activities, and for developing and testing theories about the history of life on earth (evolutionary biology, biogeography) and a shared understanding of the living world (aesthetic).

Evolutionary relationships have provided a robust framework for doing so and they are largely reflected in the classification of life. The aim of taxonomy is not, however, to reveal the footprint of evolution, but rather to use evolutionary relationships to provide a robust

and stable classification and so make the universe of biodiversity accessible to all. A classification needs to meet the requirements for a well-documented and wide range of uses and users. It is the diversity of these, from local farmers, pharmacists, amateur naturalists, archaeologists, anthropologists, ecologists, environmental scientists, physicists to systematic and evolutionary biologists, that places an emphasis on morphometrics, broad predictiveness, and ease of diagnosis and makes taxonomy a fundamentally prag-matic undertaking. It is because of this need for accessibility and ease of diagnosis, the fact that the foundations of post-Linnean, 'Candollean' (or 'natural', see Chapter 2) taxonomy were built on morphological observations, and the importance of integrating fossils, that morphology remains key to species recognition. Whilst DNA sequence observations enable the assigning of individuals to phylogenetic lineages and so provide a major tool for species identification and delimitation, they are limited in their accessibility to a relatively small and wealthy group of academics, and commercial and government agencies. They also still rely on referencing a nomenclatural system wedded to (and so do not function outside of) morphology-based classifications.

1.4 How Best to Document Cryptic Species/Morphological Stasis in Nomenclature

To summarise, the term cryptic species is problematic on two levels. (1) It assumes a lineage-dominated view of species definition and delimitation, which results from a decision to apply operational criteria as definitional concepts. (2) Depending on the definition used (e.g. Struck et al. 2018b; a; Chapter 6), it conflates and obscures noteworthy evolutionary phenomena, the most important of which is probably morphological stasis. Those phenomena are not best served by the *status quo*, whereby 'cryptic' species are described under new binomials that offer no indication of the sister cryptic species(s) and which cannot be diagnosed without the infrastructure and resources to generate DNA sequence observations.

For multicellular and many unicellular organisms, trinomials may be a more useful vehicle for naming cryptic taxa as they flag the relationship between cryptic 'sister' species. In the case of the International Code of Nomenclature for algae, fungi, and plants (Turland et al. 2018), the International Code of Nomenclature of Prokaryotes (Anon 2019), and the International Code of Zoological Nomenclature (Ride et al. 1999) there is the availability of subspecies as a rank, which would fit well within a heuristic framework and cohesive species concept. Within lineage-focussed concepts this may be problematic as there may be a perceived implication of incomplete lineage divergence. In the case of viruses, while the International Code of Virus Classification and Nomenclature does not provide a framework for subspecific ranks, which are devolved to specialist groups (International Committee on Taxonomy of Viruses 2005), there are groups of organisms for which nomenclatural codes do not permit trinomials.

Other notations are possible. Hybrid plant species (nothotaxa) are indicated by placing a multiplication sign before the species epithet (ICBN H.3A.1. 2018). For example,

Verbascum × *schiedeanum* W. D. J. Koch indicates that the taxon is a hybrid. It is conceivable that a similar notation could be used to indicate cryptic status, although the use of the letter 'c' would need to be spaced in a way to avoid orthographic confusion. This would also require the concerted modification of nomenclatural codes, which in turn would require broad consensus, expressed by the votes of the systematics communities.

In the absence of a universal species concept, and given the heuristic nature of species delimitation and recognition, the use of subspecific rank is probably the best way to document morphological stasis.

The symposium and the papers that emerged from it and are presented here show clearly how the topic of 'species' remains central to biodiversity sciences and the subject of wide-ranging and lively debate. In almost every paper there is a call for change, either of direction or for the inclusion of new developments, and their focus ranges from abandoning species altogether (Chapter 4) to highlighting the fact that there is still no accessible reference system for the 300 years-worth of accumulated knowledge of species' delimitation (Chapter 2): our representation of the biological universe is still a chaotic torso. Other authors highlight the need for international cooperation as the only meaningful basis for generating such a representation – a collective effort that requires long-term institutional investment – and that the methodology of monograph production requires a favourable institutional (and political) framework.

Taxonomists need to remember that species, as well as being the products of an evolutionary process, are also conventions on which this language and scientific facts are built (Fleck 1935). The issue of species as units of biological diversity, therefore, goes well beyond the relatively simple problem of scientific definition (Lherminier 2015), because at its root what is involved in the notion of species is a key part of our mental language that we all need for understanding our living world. Now, more than ever, this is a language in which everybody has a stake, as we experience the mass extinction of biodiversity and the loss of the ecosystem services that are, at least in part, derived from them (Haines-Young and Potschin 2010).

Acknowledgements

References

Álvarez-Presas, M., Amaral, S. V., Carbayo, F., Leal-Zanchet, A. M., and Riutort, M. (2015) Focus on the details: Morphological evidence supports new cryptic land flatworm (Platyhelminthes) species revealed with molecules. *Organisms Diversity & Evolution* 15: 379–403. https://doi.org/10.1007/s13127-014-0197-z

Anon (2019) International Code of Nomenclature of Prokaryotes. *International Journal of Systematic and Evolutionary Microbiology* 69: S1–S111. https://doi.org/10.1099/ijsem.0.000778

Armenteros, M., Ruiz-Abierno, A., and Decraemer, W. (2014) Taxonomy of Stilbonematinae (Nematoda: Desmodoridae): Description of two new and three known species and phylogenetic relationships within the family. *Zoological Journal of the Linnean Society* 171: 1–21. https://doi.org/10.1111/zoj.12126

Atran, S. (1998) Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences* 21: 547–569. https://doi.org/10.1017/S0140525X98001277

Bauret, L., Gaudeul, M., Sundue, M. A. A. et al. (2017) Madagascar sheds new light on the molecular systematics and biogeography of grammitid ferns: New unexpected lineages and numerous long-distance dispersal events. *Molecular Phylogenetics and Evolution* 111: 1–17. https://doi.org/10.1016/j.ympev.2017.03.005

Bensch, S., Péarez-Tris, J., Waldenströum, J., and Hellgren, O. (2004) Linkage between nuclear and mitochondrial DNA sequences in avian malaria parasites: Multiple cases of cryptic speciation? *Evolution* 58: 1617–1621. https://doi.org/10.1111/j.0014-3820.2004.tb01742.x

Berlin, B. (1973) Folk systematics in relation to biological classification and nomenclature. *Annual Reviews* 4: 259–271.
  (1992) *Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies*. Princeton University Press, Princeton, NJ.

Berlin, B., Breedlove, D. E., and Raven, P. H. (1974) *Principles of Tzeltal Plant Classification*. Academic Press, New York and London.

Bulmer, R. N. H., Menzies, J. I., and Parker, F. (1968) Kalam classification of birds and reptiles. *Journal of the Polynesian Society* 84: 267–308.

Carroll, S. P., Jørgensen, S. P., Kinnison, M. T. et al. (2014) Applying evolutionary biology to address global challenges. *Science* 346: 313–323. https://doi.org/10.1126/science.1245993

Cerca, J. Meyer, C., Purschke, G., and Struck, T. H. (2020) Delimitation of cryptic species drastically reduces the geographical ranges of marine interstitial ghost-worms (Stygocapitella; Annelida, Sedentaria). *Molecular Phylogenetics and Evolution* 143: 106663. https://doi.org/10.1016/j.ympev.2019.106663

Charlesworth, B. and Lande, R. (1982) Morphological stasis and developmental constraint: No problem for Neo-Darwinism. *Nature* 296: 610–610. https://doi.org/10.1038/296610a0

Coyne, J. A. and Orr, O. H. (2004) *Speciation*. Sinauer Associates, Sunderland, MA.

Cozzuol, M. A., Clozato, C. L., Holanda, E. C. et al. (2013) A new species of tapir from the Amazon. *Journal of Mammalogy* 94: 1331–1345. https://doi.org/10.1644/12-MAMM-A-169.1

Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London.

Davis, C. C., Schaefer, H., Xi, Z. et al. (2014) Long-term morphological stasis maintained by a plant-pollinator mutualism. *Proceedings of the National Academy of Sciences* 111: 5914–5919. https://doi.org/10.1073/pnas.1403157111