Different species problems and their resolution

Kevin de Queiroz

Summary
At least three different issues are commonly referred to by the term “the species problem”: one concerns the necessary properties of species, a second the processes responsible for the existence of species, and a third methods for inferring species limits. Solutions have recently been proposed to the first two problems, which are conceptual in nature (the third is methodological). The first equates species with metapopulation lineages and proposes that existence as a separately evolving metapopulation lineage be considered the only necessary property of species. The second views the species category as a cluster concept and proposes that no single process or set of processes be considered necessary for the existence of species. Although these two solutions have been portrayed as being in conflict, they are, in fact, highly compatible. Moreover, the proposals in question clarify the problem concerning methods for inferring the limits of species, which has for a long time been confused with the problem concerning the necessary properties of species. Together these proposals provide the opportunity for biology to move beyond debates about the definition of the species category and focus on estimating the boundaries and numbers of species as well as studying the diverse processes involved in their origin and persistence. BioEssays 27:1263–1269, 2005. © 2005 Wiley Periodicals, Inc.

The species problem(s)
Uncertainty or disagreement concerning the nature of species, and thus also how to recognize species in practice, is one of the oldest problems in biology. The problem is manifested in the existence of scores of alternative definitions for the term “species,” no fewer than 24 of which have been designated as distinct species concepts. Recently, however, two general solutions have been proposed to this long-standing problem. The first proposes to treat existence as a separately evolving metapopulation lineage as the only necessary property of species; the second proposes that the species category cannot be defined using one or more necessary and sufficient properties and therefore should be treated as a cluster concept. On the surface, these proposals appear to be at odds with one another (as the author of the second proposal has interpreted them); however, I intend to show that they are highly compatible. I argue that the apparent discrepancies stem from using a single term, “the species problem”, to designate at least three distinct problems. One of these problems is solved by the first proposal; another is solved by the second. These proposals concerning the concept of species also clarify a third species problem concerning methods for inferring the boundaries and numbers of species.

Species as metapopulation lineages
The numerous alternative definitions of the term “species” can be classified into several groups that are at least partially incompatible (Table 1). The reason for this incompatibility is that the definitions are based, in part, on different defining (necessary) properties. For example, some definitions require intrinsic reproductive isolation while others require shared specific mate recognition or fertilization systems; phenetic differences, ecological distinctiveness, fixed character state differences, monophyly, the exclusive coalescence of alleles and heterozygote deficits to name some of the widely adopted properties. These properties, which represent thresholds crossed by diverging lineages, commonly arise at different times during the process of speciation. As a consequence, species definitions based on different properties commonly lead to the recognition of different numbers and boundaries of species taxa.

Nevertheless, all modern species concepts—and the definitions that attempt to describe them—are based on a single more general concept of species. All of them either explicitly or implicitly equate species with separately evolving (segments of) metapopulation lineages, where a metapopulation is an inclusive population made up of a set of connected subpopulations, and a lineage (at the population level) is a population extended through time or an ancestral-descendant series of time-limited (instantaneous) populations. Despite this general conceptual uniformity, the various classes of incompatible species definitions (Table 1) differ in treating different properties acquired by metapopulation lineages during their divergence as necessary properties of species—that is, as properties that a metapopulation lineage must possess to be considered a species.

To reconcile incompatible alternative species definitions, I have proposed that only that property shared by all
Table 1. Different classes of contemporary species definitions and the distinctive properties upon which they are based

<table>
<thead>
<tr>
<th>Class of species definition</th>
<th>Property upon which it is based</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological Isolation</td>
<td>Intrinsic reproductive isolation (absence of interbreeding between organisms of different species based on intrinsic properties, as opposed to geographic barriers)</td>
<td>Mayr 1942; Dobzhansky 1970</td>
</tr>
<tr>
<td>Recognition</td>
<td>Shared specific mate recognition or fertilization system (mechanisms by which organisms of the same species, or their gametes, recognize one another for mating and fertilization)</td>
<td>Paterson 1985</td>
</tr>
<tr>
<td>Ecological</td>
<td>Same niche or adaptive zone (all components of the environment with which the organisms interact)</td>
<td>Van Valen 1976</td>
</tr>
<tr>
<td>Phylogenetic</td>
<td>Heterogeneous (see below)</td>
<td>Rosen 1979; Donoghue 1985</td>
</tr>
<tr>
<td>Monophyly</td>
<td>Monophyly (consisting of an ancestor and all of its descendants; commonly inferred from possession of shared derived character states)</td>
<td>Baum and Shaw 1995</td>
</tr>
<tr>
<td>Genealogical</td>
<td>Exclusive coalescence of alleles (all alleles of a given gene are descended from a common ancestral allele not shared with those of other species)</td>
<td>Cracraft 1983; Nixon and Wheeler 1990</td>
</tr>
<tr>
<td>Diagnosable</td>
<td>Form a diagnosable group (qualitative difference)</td>
<td>Grismer 1999, 2001</td>
</tr>
<tr>
<td>Evolutionary</td>
<td>Heterogeneous (see below)</td>
<td>Wiens 2004</td>
</tr>
<tr>
<td>Some interpretations</td>
<td>Form a diagnosable group (qualitative difference)</td>
<td>Michener 1970; Sokal and Crovello 1970; Sneth and Sokal 1973</td>
</tr>
<tr>
<td>Other interpretations</td>
<td>Separation of lineages (intrinsic or extrinsic)</td>
<td></td>
</tr>
<tr>
<td>Phenetic</td>
<td>Form a phenetic cluster (quantitative difference)</td>
<td></td>
</tr>
<tr>
<td>Genotypic cluster</td>
<td>Form a genotypic cluster (inferred from deficits of genetic intermediates, e.g., heterozygotes)</td>
<td>Mallet 1995</td>
</tr>
</tbody>
</table>


contemporary species concepts and definitions be considered a necessary property of species. Thus, species are defined simply as separately evolving (segments of) metapopulation lineages (see also Ref. 27). All of the other properties that have previously been considered additional necessary properties of species (Table 1) are to be reinterpreted as contingent rather than necessary properties of species—that is, as properties that a species may or may not acquire during the course of its existence. Nonetheless, these contingent properties retain their importance in two major ways. First, they serve as lines of evidence for inferring the boundaries and numbers of species. Second, they serve as defining properties of subcategories of the more general species category (e.g. reproductively isolated species, ecologically differentiated species, monophyletic species, diagnosable species, etc.) and thus for identifying those species that are most relevant to addressing particular biological questions. This proposal thus reconciles alternative incompatible species concepts and definitions without denying the importance of the properties that underlie their differences. The result is a general and unified concept of species.

The species as a family resemblance concept
In addition to disagreement about the necessary properties of species, there is also disagreement concerning the processes responsible for the existence of species. Most of the
early definitions that equated species with metapopulation lineages—that is, those proposed during the era of the Modern Evolutionary Synthesis[9,29]—emphasized sexual reproduction and gene flow as the processes responsible for uniting organisms to form populations and thus also species.30,31 Sometimes to the extent that asexual organisms were considered not to form species.30,32,33 Nonetheless, other authors have called attention to the apparent maintenance of separation between metapopulation lineages (as evidenced by distinctive ecologies and morphologies) that appear to exchange genes more or less freely.14,27,34 They have also noted discontinuities as sharp or sharper between asexual species as between sexual ones.35 They are characteristic of species but which the metapopulation lineage proposal supposedly overlooks, one of those that Pigliucci mentioned explicitly, being composed of organisms, and being subjected to a variety of evolutionary forces. Thus, according to Pigliucci (p. 598), being a metapopulation level lineage is not sufficient for being considered a species.

To acknowledge the potential importance of diverse processes in unifying organisms to form species, Massimo Pigliucci proposed that the species category is best viewed as a family resemblance concept. Pigliucci's proposal, which revived an idea put forward almost 40 years earlier by David Hull, is based on Ludwig Wittgenstein's idea of family resemblance groups, which have also been called polythetic groups and cluster concepts.13,40–42

The concept of family resemblance was proposed by Wittgenstein to address the problem that certain words seem impossible to define in terms of a single property or a finite set of properties that are considered both necessary and sufficient to adequately specify the meaning of the word. For example, the diverse activities for which we use the word "game", including track and field, baseball, synchronized swimming, chess and poker, do not seem to share any finite set of properties that also set them apart from activities that we do not call games (though many of them share one or more relevant properties). Similarly, Pigliucci's proposal is that the term "species" (and thus the species category) also cannot be adequately specified using a standard definition — that is, a finite set of necessary and sufficient properties. Instead, it can only be adequately specified by a definition in which no single property or set of properties is necessary and any one of several different sets is sufficient.58 The species category thus forms a cluster, the members of which are linked by partially overlapping sets of shared properties. This proposal obviates the need for identifying a single process (e.g. outcrossing sexual reproduction) as being responsible for the existence of species in diverse types of organisms. In other words, it acknowledges the possibility that a different process or processes may be responsible for uniting organisms to form species in different groups, whether defined functionally (e.g. sexual versus asexual) or taxonomically (e.g. birds versus bacteria).

Supposed discrepancies
In addition to proposing his cluster concept solution to the species problem, Pigliucci suggested that the metapopulation lineage proposal is flawed in two ways. First, it is based on identifying a necessary and sufficient property of species (existence as a separate and evolving metapopulation lineage), which his proposal rejects. Second, it is supposed to be too broad to be useful. On the one hand, it does not stipulate how to distinguish species from metapopulation lineages that have not diverged enough to be considered species. On the other hand, it does not address several additional characteristics that Pigliucci considered necessary (but not sufficient) properties of species, such as being composed of organisms and being subjected to a variety of evolutionary forces. Thus, according to Pigliucci (p. 598), being a metapopulation level lineage is not sufficient for being considered a species.

Contrary to Pigliucci's second criticism, being a metapopulation lineage is sufficient for being considered a species. An important corollary of the metapopulation lineage proposal is that all separately evolving metapopulation lineages are species.4,5,25 In other words, to be considered species, separately evolving metapopulation lineages need not possess any of the other properties that have traditionally been considered necessary for membership in the species category (Table 1), such as intrinsic reproductive isolation, ecological distinctiveness, diagnosability, or monophyly. Being a separately evolving segment of a metapopulation lineage is the only necessary property of species. Pigliucci's second criticism therefore rests on the explicit acceptance of the very proposition that the metapopulation lineage proposal rejects: the proposition that being a separately evolving segment of a metapopulation level lineage is not sufficient for being considered a species—in other words, that a separately evolving metapopulation lineage does not become a species until it acquires some additional property (Table 1).

In this context, it should be evident why the metapopulation lineage proposal does not stipulate how to distinguish species from other separately evolving metapopulation lineages that have not diverged enough to be considered species: it rejects the very notion that those other metapopulation lineages are not to be considered species. Regarding other properties that are characteristic of species but which the metapopulation lineage proposal supposedly overlooks, one of those that Pigliucci mentioned explicitly, being composed of organisms, is implicit in the idea of a metapopulation lineage (populations are composed of organisms and therefore so are population lineages). The other property, being subjected to a variety of evolutionary forces, does not seem to be a logically necessary
property of species, though it may be a physically necessary one (see Ref. 43 for the distinction between logical and physical necessity). Thus, Pigliucci’s second criticism is based on a misunderstanding of the metapopulation lineage proposal.

Pigliucci’s other criticism is that the metapopulation lineage proposal is based on identifying a necessary and sufficient property of species. This idea is supposedly problematical because his proposal rejects the very idea that the species category can be satisfactorily defined in terms of necessary and sufficient properties. Rejection of this idea goes hand in hand with the interpretation of the species category as a family resemblance or cluster concept, which avoids a definition taking the form of a set of properties the members of which are considered separately necessarily and jointly sufficient. In effect, the two proposals reject different fundamental assumptions. The metapopulation lineage proposal rejects the fundamental assumption that a separately evolving metapopulation lineage is not to be considered a species until it has acquired some additional property (see Table 1). In contrast, the cluster concept proposal rejects the fundamental assumption that the definition of the species category is to be stated in terms of necessary and sufficient properties.

There are two different ways that Pigliucci’s cluster concept proposal can be interpreted, one of which is highly compatible with the metapopulation lineage proposal. On the one hand, it can be interpreted as treating the various properties that cause the incompatibilities between alternative definitions of the species category as the basis of a cluster concept definition (i.e. the ones that represent thresholds crossed during lineage divergence, which the metapopulation lineage proposal treats as contingent properties of species). If so, then Pigliucci’s proposal perpetuates the problem of incompatible species definitions in that it does not prevent, or even discourage, different authors from adopting different properties as defining properties of the species category (Table 1). Although no one of those properties would be considered a necessary property of all species, each could still be considered a member of one of several different sets of sufficient properties (which might differ only with respect to the properties in question). Rather than resolving the incompatibilities among alternative definitions of the species category, this interpretation encourages different authors to adopt incompatible definitions, thus perpetuating the current disagreements. Fortunately, Pigliucci does not adopt the interpretation of the properties responsible for incompatible species definitions (Table 1) as those involved in the cluster concept.

Instead, Pigliucci views a different set of properties as the basis of the proposed cluster concept. The properties that he explicitly identified as forming the cluster concept are not properties such as phonetic distinguishability, intrinsic reproductive isolation, monophyly, occupation of a distinct niche, diagnosability, or any of the other properties listed in Table 1. More specifically, they are not thresholds crossed by diverging lineages that are used to decide when a lineage deserves to be recognized as a species. Instead, they are more general properties, most of which are possessed by all metapopulation lineages or their component organisms. The ones that Pigliucci explicitly identified are genetic similarity, reproductive isolation (though in the context of my argument this property might be better described as interbreeding), phylogenetic relationships, ecological role and morphological similarity.

As in the case of the alternative interpretation, the interpretation of these general properties as the basis of the cluster concept does not solve the problem of incompatible alternative definitions of the species category. Because this interpretation does not involve the properties responsible for the conflicts among those alternative definitions, it cannot resolve the conflicts among them. Put another way, in contrast to the case of the concept of games discussed by Wittgenstein, previous authors have not found it impossible or even difficult to formulate a definition of the species category that applies to all of the entities that they consider species. Instead, each author has formulated what he or she considers a perfectly adequate definition of the species category. The problem is that different authors prefer different (and partially incompatible) definitions.

Different species problems

Although the cluster concept proposal does not resolve the problem of the existence of alternative and incompatible definitions of the species category, it solves a different problem concerning the nature of species. As noted above (see The Species as a Family Resemblance Concept), a number of different phenomena, including interbreeding, ecologically mediated natural selection, developmental constraints and common ancestry, have been proposed as important for the existence of species (i.e., for uniting organisms to form species). Not coincidentally, the phenomena proposed as important for the existence of species are closely related to the properties proposed by Pigliucci as the basis of the species as a cluster concept. Thus, genetic similarity (as proposed by Pigliucci) results from interbreeding, common ancestry, and natural selection. Reproductive isolation (if interpreted as either intrinsic or extrinsic) is the complement of interbreeding. Phylogenetic relationships are more or less equivalent to common ancestry. Ecological role is the basis of ecologically mediated natural selection. And morphological similarity results from all of the phenomena (interbreeding, natural selection, developmental constraints, and common ancestry).

In addition, Pigliucci stated that, of the species concepts that he considered, Templeton’s cohesion species concept comes the closest to the idea of family resemblance. The view that several different phenomena are responsible for the existence of species has been developed most thoroughly by Templeton, who discussed the limits of the spread of
genetic variants across a continuum of reproductive modes and identified several phenomena as important for defining those limits. Those phenomena, which he termed “cohesion mechanisms”, include interbreeding, natural selection, ecological similarity, developmental constraints and common ancestry—the same ones identified in the previous paragraph as being closely related to the properties upon which Pigliucci’s family resemblance concept is based. In short, Pigliucci’s proposal is based on properties related to the fact that different processes may be responsible for the existence of species in different groups of organisms, and it solves the problem of how to define the concept of species given this situation by allowing different properties or sets of properties to provide the definition in different cases.

In point of fact, the metapopulation lineage and cluster concept proposals solve different species problems. The metapopulation lineage proposal solves the problem of alternative and partially incompatible definitions of the species category. It does so by removing from those rival definitions the elements responsible for their incompatibilities. The result is a general and unified concept of species as separately evolving (segments of) metapopulation lineages. In contrast, the cluster concept proposal solves a different species problem concerning the phenomena responsible for the existence of metapopulation lineages themselves, which appear to differ among lineages. It does so by allowing different phenomena or sets of phenomena to be responsible (or most important) in different cases.

These two proposals are highly compatible. According to the metapopulation lineage proposal, the species category is best defined by reference to a single necessary and sufficient property—existence as a separately evolving metapopulation lineage. Nonetheless, in agreement with the cluster concept proposal, the idea of a metapopulation lineage may itself be best interpreted as a family resemblance or cluster concept. In this context, any supposed incompatibilities between the two proposals result from confusing two different issues under the single term “the species problem”—that is, disagreements about thresholds that must be crossed by metapopulation lineages for those lineages to be considered species versus disagreements about the phenomena that are most important for the existence of species as metapopulation lineages.

**Species taxa as hypotheses**

In addition to these disagreements, there is at least one additional issue commonly referred to as “the species problem”. This third issue concerns how to recognize species in practice—that is, how to infer the boundaries and thus also the numbers of species. Given that the concept of species influences the criteria and methods that are considered relevant for recognizing species in practice, this third species problem has for a long time been inextricably intertwined with the first (the problem concerning the necessary and sufficient properties of species). Under the unified species concept resulting from the treatment of properties acquired by diverging metapopulation lineages as contingent rather than necessary properties of species, the two problems are clearly separated (though still related). Under this view, the concept of species (and the definition of the species category) is seen not to depend on any of the properties in question (Table 1), and the issue of how to recognize species in practice is seen as a distinct question concerning how to determine whether different organisms or local populations represent separately evolving metapopulation lineages, including the question of how the properties in Table 1 bear on this question.

Part of the reason for this clear separation is that the two questions address fundamentally different concerns. The first species problem is conceptual in that it concerns the basic idea of species; it is related to the question “What are species?” or, more specifically, “What properties must a metapopulation lineage possess to be considered a species?”. The second species problem is also conceptual but is concerned with a different question, namely “What phenomena are responsible for the existence of species (as metapopulation lineages)?”

In contrast, the third species problem is methodological rather than conceptual in that it concerns the criteria and methods for determining the boundaries and numbers of species from empirical data; it is related to the question “How do we recognize species in practice?” Nonetheless, it is legitimately termed a “species problem” in the same sense that we talk about a “phylogeny problem” concerning how to infer or estimate evolutionary relationships.

As such, the third species problem (like the phylogeny problem) has many solutions (reviewed recently in Refs. 44,45). Those solutions involve diverse methods (e.g. correlations between geographic and genetic distances, frequency distributions of genetic distances, ordination, phenetic clustering, phylogenetic analysis, methods for estimating gene flow, spatial analysis, coalescent methods) as well as diverse types of data (e.g. morphological, geographical, behavioral, physiological, ecological, genetic). Moreover, many of the methods are based on lines of evidence that correspond to the properties that were previously treated as necessary properties of species (Table 1) and caused the incompatibilities among alternative species definitions. When these properties are treated as lines of evidence rather than necessary properties of species, the previous existence of incompatible alternative species definitions is seen to have resulted in a large part from confusing conceptual and methodological issues—that is, from confusing the concept of species with the criteria and methods for recognizing species in practice.\(^5,27\)

Implicit in this third species problem (how to infer the limits of species) is the idea that species taxa are hypotheses (e.g. Ref. 46 and references therein). That is, any currently recognized species taxon (i.e. any taxonomic group assigned to the species category) is a set of organisms or populations...
The species problem has plagued biology for years. At least these and many other reasons, there will always be some that there are no definitive answers. Moreover, separation of the boundaries and numbers of species. As a consequence, any disagreements about the interpretation of the results, rather than disagreements about the definition of the species category.

Despite the existence of diverse and increasingly sophisticated methods for inferring the boundaries and numbers of species, ambiguous cases are inevitable. For one thing, hypotheses about species are like all scientific hypotheses in that there are no definitive answers. Moreover, separation between metapopulation lineages can be partial as well as dependent on the time scale over which the case is studied (populations that are separate on a scale of years may be connected on a scale of hundreds or thousands of years). For these and many other reasons, there will always be some cases in which the boundaries and numbers of species are difficult to assess and therefore likely to be disputed. These ambiguities, however, either are of an empirical nature or reflect differences in scale; they do not imply any major unresolved problem concerning the concept (fundamental nature) of species or the definition of the species category.

Conclusions
The species problem has plagued biology for years. At least part of the problem has been that several different issues have been confused under this single term. There is reason to believe, however, that biology is now ready to move beyond the major conceptual problems regarding species. Proposals have recently been put forward that have the potential to resolve both of the long-standing controversies concerning the fundamental nature of species. The proposal to consider existence as a separately evolving metapopulation lineage the only necessary property of species eliminates the incompatibilities among alternative species definitions. By emphasizing the property shared by all contemporary views on species without denying the importance of any of the properties that underlie their differences, it results in a unified concept of species. Similarly, the proposal to consider the species (metapopulation lineage) a family resemblance concept eliminates the need to identify any single process as exclusively responsible for the existence of species. By embracing the idea that different processes may be responsible for the unification of organisms into populations and population lineages in different cases, it results in a concept that can be applied to a diversity of biological situations, including the entire range of the reproductive continuum.

By resolving these long-standing problems concerning the basic nature of species, these proposals shift emphasis away from the definition of the species category to a third species problem concerning the methods for inferring the limits and numbers of species. Fortunately, this problem already has a number of good solutions, though new solutions and refinements of old ones are continually being proposed. It seems that biology is finally in a position to shift its attention away from what once seemed an endless debate about the definition of the species category and focus instead on estimating the boundaries and numbers of species and studying the diverse processes involved in their origin and maintenance.

Acknowledgments
I would like to thank Richard Harrison, Jody Hey, Massimo Pigliucci, David Wake, Adam Wilkins, and an anonymous referee for discussions, correspondence, and comments related to the development of ideas contained in this essay.

References
Problems and paradigms


