

Defining Species: A Multi-Level Approach

Tudor M. Baetu

Received: 7 May 2011 / Accepted: 12 December 2011 / Published online: 27 December 2011
© Springer Science+Business Media B.V. 2011

Abstract Different concepts define species at the pattern-level grouping of organisms into discrete clusters, the level of the processes operating within and between populations leading to the formation and maintenance of these clusters, or the level of the inner-organismic genetic and molecular mechanisms that contribute to species cohesion or promote speciation. I argue that, unlike single-level approaches, a multi-level framework takes into account the complex sequences of cause-effect reinforcements leading to the formation and maintenance of various patterns, and allows for revisions and refinements of pattern-based characterizations in light of the gradual elucidation of the causes and mechanisms contributing to pattern formation and maintenance.

Keywords Species concepts · Speciation · Levels · Mechanisms · Pattern · Process

1 Introduction

Most biologists agree that discrete clusters exist among organisms. They disagree however, about which clusters constitute species, and what are their defining characteristics. Many of these disagreements, especially those touching on the defining characteristics of groups of organisms we may want to call species, stem from divergences in research interests (Harrison 1998; Lidén and Oxelman 1989; Reydon 2005). Evolutionary biologists are interested in the causes leading to the formation and maintenance of distinct clusters, and tend to define species in terms of processes like gene flow, as well as the factors that promote or hinder these processes, such as geographic and ecological barriers that promote speciation

T. M. Baetu (✉)

Department of Philosophy, University of Maryland, 1125D Skinner Building, College Park,
MD 20742, USA
e-mail: tbaetu@hotmail.com

[e.g., (Mayr 1963)]. In contrast, systematists are more interested in the final effects than in the causes leading to the formation and maintenance of species, namely the formation of diagnosable clusters or exclusive groups of common ancestry [e.g., (Cracraft 1983; Shaw 1998)]. More recently, geneticists and molecular biologists became interested in elucidating the molecular basis of speciation. With this novel research interest came yet another family of species concepts, placing an emphasis on the biological mechanisms present in each organism (as opposed to extrinsic barriers of a geographical or ecological sort) that make possible or hinder processes like gene flow [e.g., (Coyne and Orr 1998)].

In light of the above considerations, I identify three levels at which groups of organisms we may want to call species are defined: the pattern-level grouping of organisms into discrete clusters; the underlying process-level of gene flow and other processes operating within/between populations, and ultimately contributing to the formation and maintenance of discrete clusters; or the inner-organismic mechanism-level of the genetic and molecular mechanisms that contribute to species cohesion or promote speciation, most notably by allowing or hindering gene exchange between organisms.¹

It is generally accepted that different species concepts are relevant depending on the level associated with the primary research interest, the organisms under investigation, the techniques used, and the available data (Ereshefsky 1992; Kitcher 1984). This pragmatically-driven pluralism is viewed as both describing a state of affairs in the biological sciences, and as having the positive effect of promoting a variety of investigative projects and approaches.

In this paper, I defend an integrative alternative to a strictly pragmatic pluralism aiming to capture the fact that the formation and maintenance of species is a multi-level phenomenon involving many possible sequences of cause-effect reinforcements affecting different levels. I argue that a multi-level framework is required in order to track the diverse sequences of events leading to the formation and maintenance of patterns, as well as to revise and refine pattern-based characterizations in light of the gradual elucidation of the causes and mechanisms contributing to pattern formation and maintenance.

The paper is organized as follows. I begin by introducing process-, pattern-, and mechanism-level species concepts (Sects. 2.1–3), followed by a discussion of recent attempts to integrate available concepts by assigning them different operational roles, or by ordering them along a ‘natural history of a species’ (Sects. 3.1–2). In Sect. 3.3, I elaborate my own version of an integrative, multi-level framework for a species concept, and I illustrate how such a framework makes it possible to justify relevance decisions that would have otherwise been arbitrary, as well as to disentangle complex patterns associated with rapidly evolving organisms. I end my paper with a summary of my findings (Sect. 4).

¹ There are other ways of classifying species concepts, such as operational versus theoretical (Mayden 1999), prospective versus retrospective (O’Hara 1993), history-based versus character-based (Baum and Donoghue 1995), mechanistic versus historical (Luckow 1995), pattern- versus process-based (Cracraft 1983; Harrison 1998), class versus entity concepts (Reydon 2005).

2 Species Concepts

2.1 Process-Level Species Concepts

The best known concept is the biological species concept (BSC), according to which species are “groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups” (Mayr 1963, 19). What is explicitly emphasized by this concept is a selective capacity to exchange genetic material (gene flow), which, for the most part, connects only the members of the species-group. Species are defined relative to the process that maintains the cohesion of the species-group and separates it from other such groups (Fig. 2, panel b). This very same process also provides at least a partial explanation of why organisms tend to group into diagnosable clusters (i.e., because of shared inherited traits), and of how speciation occurs (i.e., via reproductive isolation).

While gene flow is an important process leading to the formation and stabilization of diagnosable clusters, other evolutionary processes may also play an important role. According to the ecological species concept (EcSC), a species is “a lineage (or a closely related set of lineages), which occupies an adaptive zone minimally different from that of any other lineage in its range and which evolves separately from all lineages outside its range” (Van Valen 1976, 233).² This concept differs from BSC because it does not require the absence of gene flow between species. Rather, recognizable clusters form and are maintained because different organisms populate different ecological niches, each requiring a distinct set of adaptations. Such clusters evolve independently of each other because of the distinct selective forces associated with each niche. Just like BSC, EcSC defines species relative to the process that maintains the cohesion of the species-group and separates it from other such groups. However, while the end result is the same, namely pattern-level clusters are formed, the process leading to the formation and maintenance of species is different.

Finally, the evolutionary species concept (EvSC) can be viewed as a general version of process-level concepts, stating that a species is a “single lineage of ancestor–descendant populations which maintains its identity from other such lineages and which has its own evolutionary tendencies and historical fate” (Wiley 1978, 18). Unlike BSC and EcSC, EvSC does not commit to any particular evolutionary processes that may lead to the formation and maintenance of distinct clusters. According to Mayden (1999), EvSC provides a non-operational, but ‘highly consilient’ theoretical framework capable of accommodating a variety of operational concepts and criteria.

Relative to the level of a process such as gene flow or divergent selection, one can also consider the patterns forming as a result of the process (diagnosable phenotypic/genetic clusters, phylogenetic patterns), or the underlying level

² “An adaptive zone is some part of the resource space together with whatever predation and parasitism occurs on the group considered. It is a part of the environment, as distinct from the way of life of a taxon that may occupy it, and exists independently of any inhabitants it might have” (Van Valen 1976, 234). Only “minimally different” adaptive zones are taken into account in order to ensure that higher taxa do not count as ecological species.

specifying the genetic and molecular basis of the biological mechanisms contributing to or interfering with the functioning of the process. Before discussing the species concepts associated with these other levels, it is worth emphasizing the advantages of process-level concepts. Defining species exclusively in terms of pattern is problematic because recognizable patterns also occur within species (e.g., intra-population polymorphisms, races) and higher rank taxa (Ayala and Escalante 1996; Berlocher 1999). As a result, pattern-level concepts must specify thresholds of statistical relevance for distinguishing two clusters as two distinct species (Coyne and Orr 2004, 449), or specify which and how many gene trees are to be used, as well as the extent to which these trees have to agree in order to reconstruct phylogeny (Avice and Wollenberg 1997). In relation to the mechanism level, it has been argued that, in as much as the same causal process may be underpinned by different mechanisms in different organisms, the process-level provides the best level of generality while avoiding the problems afflicting pattern-based concepts (Ereshefsky 2010).

2.2 Pattern-Level Species Concepts

Instead of targeting causal processes, one can alternatively consider the discrete phenotypic/genotypic clusters and phylogenetic patterns forming as a result of such processes (Fig. 2, panel a). It is generally acknowledged that, while species may be defined relative to the causal processes that maintain their identity, they are seldom diagnosed this way: “All species delimitations (even those based on isolation or cohesion concepts) ultimately depend on inferences from patterns of variation and character state distributions” (Harrison 1998, 21). Diagnosis criteria are deemed operationally valuable because they pick up correlates or surface markers of underlying causal connections between organisms (Futuyma 1998, 463).

In practice, the most commonly used operational test is recognizability of organisms in sympatry (i.e., living in the same location): if hybridization occurs between two populations characterized by distinct morphological/genetic characters, the distinctive differences between the two populations would disappear (Coyne and Orr 2004; de Queiroz 2005; Mallet 1995; Mayr 1996; Schilthuizen 2000). In more general terms, diagnosis involves an assessment of overall genetic or phenotypic similarity (the observed patterns), in sympatry or otherwise, without any further claims about gene flow or other causal connections between organisms (the processes underlying these patterns).

The genotypic cluster species concept (GCSC) equates diagnosis in sympatry with a definition of species: “A species is a morphologically or genetically distinguishable group of individuals that has few or no intermediates when in contact with other such clusters” (Mallet 1995, 296). This concept doesn’t pick the causal connections between the members of a species (the process); instead, it targets the surface level of the effects resulting from these causal connections, namely genotypic/phenotypic clusters (the pattern). This preference of pattern over process doesn’t stem from a disregard for the deeper causal connections responsible for the formation and maintenance of observable clusters, but from an attempt to take into account the fact that there is no universal causal process that connects all

and only the members of a species. Mallet (1995) observes that several processes may play a role in maintaining the distinctiveness of groups in sympatry (gene flow, stabilizing selection, and ‘historical inertia’), and that organisms may exchange genetic material in a variety of ways (sexual and asexual, vertical and lateral gene transfer). GCSC is meant to circumvent these complications by remaining independent of assumptions about the nature of the processes leading to the formation and maintenance of diagnosable clusters.

Diagnosis criteria and GCSC diagnose/define species based on geographical distributions of character states at a given time, and, in this sense, are phenetic in nature [they diagnose/define species clusters based on overall similarity (Sneath and Sokal 1973)]. In contrast, phylogenetic concepts define species in terms of phylogenetic patterns, that is, the grouping of living organisms in distinct lineages constituting the various branches and sub-branches of the tree of life. The phylogenetic species concept (PSC) defines a species as “an irreducible (basal) cluster of organisms that is diagnosably distinct from other such clusters, and within which there is a parental pattern of ancestry and descent” (Cracraft 1989, 34), where diagnosis is based on one or more synapomorphic characters that identify a monophyletic group of individuals.

The genealogical species concept (GSC) further demands that species must be groups of organisms whose members are more closely related to each other than they are to any organisms outside the group (Baum and Shaw 1995; Shaw 1998). It has been argued that this concept too favors the pattern-level in the sense that it targets the grouping of organisms (as exclusive groups) rather than the processes/mechanisms leading to the formation and maintenance of such groups (Harrison 1998; Luckow 1995).

Pattern-based concepts and criteria are justified on at least three accounts. Advocates of pattern-level concepts argue that there is no single causal connection at work in all organisms, such that the most general species concept can only be formulated at the higher level of the clusters that form as a result of a variety of distinct causal connections (Mallet 1995); and that, at least in some cases, biologists are interested in the final effects, rather than the causes leading to the formation/maintenance of species and associated phenotypic/genotypic clusters (Cracraft 1983).

In addition, the pattern-level grouping of organisms in discrete clusters provides a description of a phenomenon³ which process- and mechanism-level concepts aim to explain by pointing to the causes and mechanisms contributing to its generation. It is generally acknowledged that process/mechanistic species concepts have an explanatory component. For example, Luckow (1995, 590) argues that such concepts “begin with a theory of how speciation (evolution) works”, while Coyne and Orr (1998; 2004, ch. 1) argue that process- and mechanism-level concepts aim to answer the question “Why do species exist?” If these concepts have an

³ The description of the phenomenon varies from GCSC to PSC/GSC. The patterning captured by PSC/GSC has a temporal component (i.e., phylogeny); in contrast, the patterning captured by GCSC and diagnostic criteria associated with BSC focuses on the spatial/geographical aspect (i.e., diagnosable clusters in sympatry). However, the two are related aspects of the same phenomenon, namely the formation of spatial-temporal discrete clusters among organisms (Avice and Wollenberg 1997).

explanatory component, then there must be an explanandum. That this explanandum is the grouping of organisms in discrete clusters is clearly indicated by the fact that the accuracy and adequacy of species concepts is measured in terms of their ability to match the groupings of organisms picked by operational diagnosis criteria. This supports the claim that species concepts should preserve a certain degree of autonomy vis-à-vis explanation just in case currently proposed explanations turn out to be incomplete or false.

2.3 Mechanism-Based Concepts

Famously, Dobzhansky (1935, 353) proposed a version of BSC stating that “a species is a group of individuals fully fertile inter se, but barred from interbreeding with other similar groups by its physiological properties (producing either incompatibility of parents, or sterility of hybrids, or both).” These ‘physiological properties’ are explicitly framed as intrinsic properties of organisms that determine whether certain processes (interbreeding, gene flow) are possible. For example, Coyne and Orr (2004, 29) define isolating barriers as “those biological features of organisms that impede the exchange of genes with members of other populations. These barriers are usually, but not invariably, based on genetic differences between populations.” Similar points are made by Futuyma (1998, 450) and Harrison (1998, 23) by drawing a contrast between intrinsic barriers/mechanisms and extrinsic causes that contribute to speciation.

Reference to the mechanisms that allow or restrict gene flow between organisms is not the unique feature of BSC. The recognition species concept (RSC) defines species as “the most inclusive population of individual biparental organisms which share a common fertilization system [specific mate recognition system]” (Paterson 1985, 15). RSC is very similar to BSC with the exception that it emphasizes the mechanisms that promote gene flow rather than the barriers obstructing it. Paterson defines a ‘common fertilization system’ as a set of features (behaviors, morphological traits, pheromones) by means of which organisms recognize each other as potential mates; whether an organism possesses these features is a property of that organism. Any genetic mutations or gene combinations that result in a modification of these features constitute a form of reproductive isolation (loss of the ability to interbreed and exchange genes).

A more generalized version of BSC/RSC is the cohesion species concept (CSC), according to which a species is “the most inclusive population of individuals having the potential for phenotypic cohesion through intrinsic cohesion mechanisms [genetic and/or demographic exchangeability]” (Templeton 1989, 12). One of the most important differences between the GCSC and the CSC is that the latter includes an explanation of why individuals within clusters remain genetically/phenotypically similar. Just like RSC, CSC differs from BSC in the sense that it emphasizes the biological features of organisms that allow (rather than impede) the gene exchange. CSC is also meant to be more generally applicable than BSC and RSC because it aims to accommodate instances of gene exchange that are not associated with sexual reproduction (Templeton 1989, 13). For example, it has been argued that the mechanisms of lateral genetic transfer in viruses and bacteria act as

‘cohesion mechanisms’ analogous to the mechanisms of sexual reproduction in higher organisms (Griffiths 1999).

Dobzhansky (1937) proposed an exhaustive list of intrinsic reproductive mechanisms/barriers and devised a schematic genetic model of reproductive isolation according to which functional divergence between two interacting loci in different lineages yields incompatible interactions in their hybrids. A growing body of evidence demonstrates that post- and pre-zygotic isolation have indeed a genetic component (Coyne 1992; Noor and Feder 2006; Orr et al. 2004; Wu 1994). Putative mechanisms involving transposons (Masly et al. 2006), transmission ratio disorders (Phadnis and Orr 2009), single genes (Bradshaw and Schemske 2003), and chromosomal rearrangements (Navarro and Barton 2003) have also been sketched out. These genetic/molecular mechanistic determinants constitute manipulable variables that allow researchers and evolutionary processes (e.g., drift via spontaneous mutation) to ‘take control’ over the causal connections between the members of a species.

One of the most impressive successes of this approach is the partial elucidation of the genetic basis and molecular mechanisms underlying pre-zygotic isolation in *Drosophila* (Fig. 2, panel c). In *Drosophila* species, mutations affecting the number of Thr-Gly repeats of the *period* (*per*) gene product contribute to changes in the rhythmic interval in the ‘courtship song’ (wing vibrations) of male flies. Genetically engineered male *D. melanogaster* flies expressing *D. simulans* *per* gene display the ‘song’ rhythm typical of *D. simulans*, and fail to attract females of their own species (Yu et al. 1987).

From a theoretical point of view, mechanism-level concepts are supported by ‘the point of no return’ argument: distinct taxa cannot coexist in sympatry without the eventual evolution of mechanisms regulating gene exchange. According to Mayr (1969), speciation is initiated when a sub-population becomes geographically isolated (allopatric speciation); alternatively, in the case of ecological speciation, adaptation to a specific ecological niche may provide circumstances similar to geographical isolation (Kirkpatrick and Ravigné 2002). However, such barriers are external circumstances that may change in the future, such that previously distinct species collapse back together if there is nothing intrinsic to organisms that prohibits inter-species hybridization. What renders speciation irreversible at an evolutionary scale, that is, at a scale that includes many changes in external circumstances, are intrinsic mechanisms regulating gene flow (Coyne and Orr 2004, 38).⁴

3 Integrative Approaches to Species Concepts

The general tendency in the recent literature is to accommodate a multitude of species concepts. As pointed out earlier, it is generally agreed that different species concepts are best suited depending on the primary research interest, the organisms

⁴ Hudson and Coyne (2002, 1564) further argue that mechanism considerations are also relevant to genealogical species as well, on the grounds that the coalescence of genes alone does not guarantee that groups remain exclusive when isolated populations are once again able to exchange genes.

under investigation, the techniques used, and the available data. Ultimately, however, the goal is not merely to assert the pragmatic value of pluralistic practices, but to integrate these practices by showing how different species concepts fit together. A strictly pragmatically-driven pluralism is compatible with the possibility that there might be no communication and no need for communication between branches of biology using different species concepts because these branches study completely different objects. In contrast, an integrative approach builds on the premise that, even though different species concepts are custom-tailored to serve the interests of different investigative projects, all these projects investigate different aspects of the same phenomenon.

3.1 Species as Separately Evolving Lineage Segments

In a series of recent articles, de Queiroz (1998, 2007) argues that different species concepts, as well as the plethora of characterizations and criteria encompassed under each concept, are not as incompatible as they may seem at first sight. De Queiroz (2007), 879 claims that most species concepts “agree in treating existence as a separately evolving metapopulation lineage as the primary defining property of the species category”, where the term ‘metapopulation’ refers to an inclusive population composed of connected subpopulations (to be contrasted with demes and local populations in general) and the term ‘lineage’ refers to an ancestor–descendant series (a population as it extends in time). The disagreement, argues de Queiroz, is not as much about what species are, but about the particular “properties acquired by lineages during the course of divergence (e.g., intrinsic reproductive isolation, diagnosability, monophyly) as secondary defining properties (or secondary species criteria).” In short, the claim is that all species concepts define species relative to speciation events along lineages, and in this sense agree that species are lineage segments delimited by speciation (branching) events, but disagree on the criteria on the basis of which speciation is judged to be complete and the lineage is deemed to be separately evolving.⁵

De Queiroz aims to provide a general theoretical framework for conceptualizing species, although it is not entirely clear that he succeeds in achieving this aim. Not all species concepts incorporate or rely on the assumption that species are lineage segments; at least some influential concepts (e.g., BSC, GCSC) define species in terms of synchronic populations rather than lineages (Reydon 2005). Nevertheless, one interesting suggestion made by de Queiroz is that different criteria for speciation (the secondary species criteria) are realized at different points in time and therefore don’t contradict each other (Fig. 1, left side). Also, since not all speciation events fully satisfy all the speciation criteria, not all species concepts apply in all

⁵ Mayden makes a similar argument in reference to EvSC: “the non-operational Evolutionary Species Concept [...] should serve as the theoretical concept appropriate for the category Species. As operational concepts, the remaining ideas have been incompatible with one another in their ability to encompass species diversity because each has restrictive criteria as to what qualifies as a species. However, the operational concepts can complement one another and do serve a vital role under the Evolutionary Species Concept as fundamental tools necessary for discovering diversity compatible with the primary theoretical concept” (1999, 95).

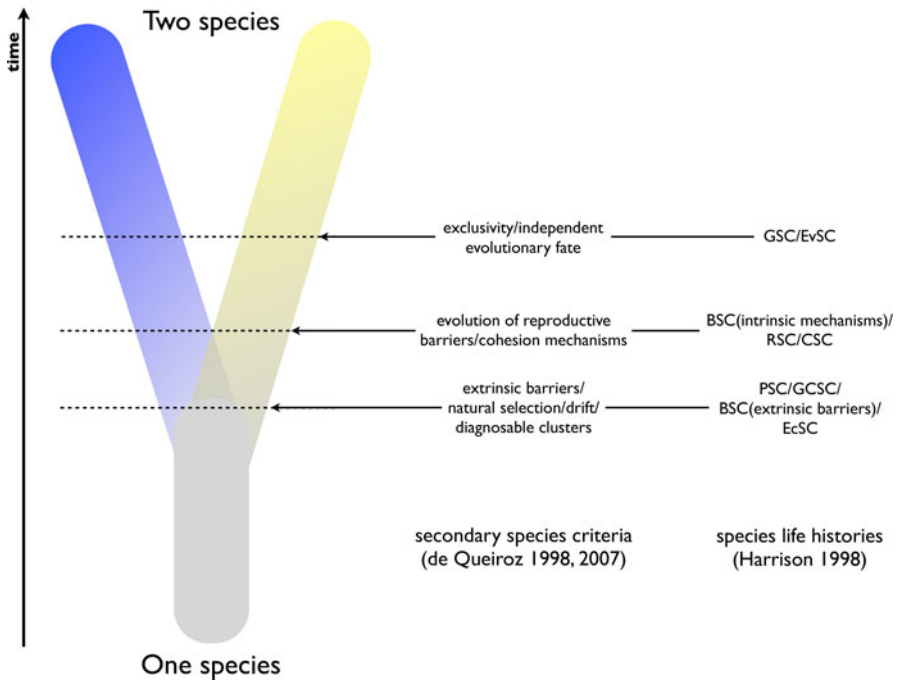


Fig. 1 The temporal integration of species concepts

situations. This suggests that different species concepts latch onto characteristics species acquire at different points in time, and that many disagreements about species concepts stem from the fact that not all species go through the same stages, or in the same order.

3.2 Harrison's "Life History of Species" Approach

Harrison (1998, 20) articulates more explicitly the temporal integration approach suggested by de Queiroz: "each species [...] has a distinctive life history, which includes a series of stages that correspond to some of the named species concepts." For instance, in many cases, the formation of diagnosable clusters precedes the evolution of intrinsic reproductive barriers and/or new cohesion/common fertilization mechanisms, themselves preceding genealogical exclusivity and evolutionary independence (Fig. 1, right side). It is worth noting that de Queiroz aims at a conceptual unity that can be viewed as a form of soft monism. In contrast, Harrison's 'life history' approach does not require that the concepts themselves share core elements (e.g., the notion of lineage), and in this sense it is unmistakably pluralistic: different concepts may be built around very different core notions; what 'keeps them together' is their common reference to an external object of inquiry, namely the temporal progression of a lineage as it diverges from other lineages.

One of the most interesting corollaries of Harrison's approach is that species concepts refer to specific time points in the 'life history of a species', while leaving in-between transition periods not covered by any of the available species concepts. The events that take place during these transition periods provide the key for understanding how different concepts relate to each other. This notion that different species concepts relate to each other not as much in terms of core conceptual commonality, but rather in terms of causal connections between the stages characterized by these concepts is illustrated by Harrison's discussion of the conditions required for attaining exclusivity.

According to the GSC, a species is a "basal, exclusive group of organisms, whose members are all more closely related to each other than they are to any organisms outside the group [the group is 'exclusive'], and that contains no exclusive group within it [the group is 'basal']" (Shaw 1998, 48). Genetically speaking, a "group of organisms is exclusive if their loci coalesce more recently within the group than between any member of the group and any organisms outside the group" (Baum and Shaw 1995, 296). In practice, two groups are recognized as distinct genealogical species when they are reciprocally monophyletic at a majority or a consensus of gene genealogies. Coalescence theory predicts that selection and drift will typically cause most alleles within a lineage to descend from a single ancestral copy occurring within that lineage (Avice and Ball 1990; Hudson and Coyne 2002). Harrison notes that coalescence models make several assumptions about process-level contributions to speciation, most notably about the interruption of gene flow and the nature of selective pressures at work prior to the formation of exclusive groups. For example, "time to reciprocal monophyly [...] depends on the nature of natural selection (assumed to be absent [...]). Divergent selection (different alleles favored in the two daughter populations) will speed the approach to exclusivity. In contrast, balancing selection will slow it down (and can, in some cases like human and chimp MHC, prevent populations that have diverged for a very long time from becoming exclusive groups). Finally, all of the above models invoke complete interruption of gene flow" (Harrison 1998, 28). The upshot of this argument is that genealogical species cannot exist prior and in the absence of biological or ecological species because exclusivity is unlikely to be attained in the absence of reproductive isolation and/or divergent selection (Coyne and Orr 2004, 469–70; de Queiroz and Donoghue 1988; Hudson and Coyne 2002).⁶

3.3 A Multi-Level Framework for Species Concepts

As discussed in Sects. 2.1–3, each concept tends to favor a particular level (pattern, process or mechanism). For example, species can be defined in terms of diagnostic characters specific to each cluster, as dictated by PSC or GCSC (Fig. 2, panel a); in terms of their ability to interbreed within each group or in terms of restricted gene flow between the two groups, as proposed by process versions of BSC (panel b); or

⁶ Wiley (1978, 20) acknowledges a similar dependence vis-à-vis EvSC: "Separate evolutionary lineages (species) must be reproductively isolated from one another to the extent that this is required for maintaining their separate identities, tendencies, and historical fates."

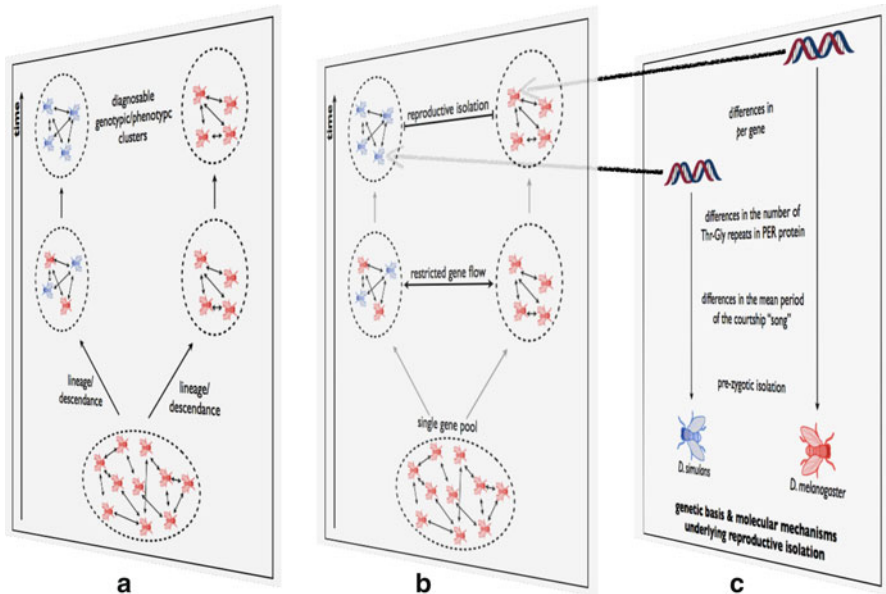


Fig. 2 The pattern, process and mechanism levels. **a** level of the observed phenomenon (*diagnosable phenotypic clusters and/or lineages*). **b** underlying casual networks responsible for generating or maintaining the observed phenomenon (*gene flow*). **c** underlying biological mechanism enabling organism to participate to a casual network (*intrinsic reproductive barriers*)

again in terms of species recognition mechanisms and genetic changes that promote barriers/loss of recognition, as proposed by RSC and mechanism versions of BSC (panel c).

One of the points raised by de Queiroz and Harrison is that there is no unique sequence of events leading to the formation and maintenance of species (i.e., each species has its own ‘history’). This is due, at least in part, to the fact that different speciation events can be initiated at different levels. Another reason is that, within any given level, speciation may be achieved by means of one or more distinct processes or mechanisms. Finally, the initiation of speciation at one level requires a reinforcement at other levels in order to complete the speciation event and maintain the identity of the newly formed species.

In the paradigmatic case of allopatric speciation (Mayr 1963), the interruption of gene flow due to extrinsic geographical barriers leads, by drift or natural selection, to the formation of diagnosable clusters, as well as the eventual evolution of intrinsic reproductive barriers. It is assumed here that what we witness at the pattern and mechanism levels are the effects of something initiated at the process-level. Based on this observation, one may be tempted to argue that the process-level is privileged because it captures the main, or at least the triggering cause of speciation. However, this would be a mistake. For one thing, as discussed previously, without reproductive isolation, allopatric speciation may be a temporary event, the effects of which can be reversed once geographical barriers are removed; for instance, without the evolution of intrinsic reproductive barriers, ring species could not form. Second,

a growing body of evidence demonstrates that not all speciation events follow this path. For example, in ecological (sympatric) speciation, the interruption of gene flow is not the cause, but the effect. In this case, speciation is triggered by divergent selection of ordinary phenotypic traits, has an effect on the mechanisms contributing to species recognition, and ends up by affecting the gene-flow process (Schluter 2009).

It is also possible that speciation may proceed bottom-up, starting with a mechanism-level genetic change resulting in an alteration of mating preferences or mate availability, in turn altering accessibility to a given gene pool. The change in mating preference/availability may itself be adaptive, or it may grant access to adaptive genes, thus leading to the evolution of a new species. For example, the pink-flowered *Mimulus lewisii* is pollinated by bumblebees, while its red-flowered species *M. cardinalis*, is pollinated by hummingbirds; the difference in color is mainly due to differences at the YUP locus, which controls the presence of yellow carotenoid pigments in the petals. When the YUP allele from one species is substituted into the sister species, a major switch in pollinator is observed, suggesting that pre-zygotic barriers can rapidly evolve as a result of single mutations (Bradshaw and Schemske 2003).

Given this diversity of triggering causes and sequences of events leading to the formation and maintenance of species, I suggest an alternative integration strategy to the temporal progression approach adopted by Harrison and de Queiroz, namely a multi-level framework capable of accommodating multiple species 'life histories.' De Queiroz argues that all concepts agree that species are lineage segments found in-between speciation (branching) events, but rely on criteria for speciation realized at different points in time; here, the notion of 'lineage segment' provides a common conceptual feature onto which one can graft a variety of speciation criteria. Harrison argues for an alternative, and more flexible solution that does not require that all concepts share a common conceptual feature; according to his account, the framework for accommodating a diversity of concepts is the real object to which all these concepts refer, namely the temporal progression of a lineage. I want to go a step further and argue that in as much as the formation and maintenance of clusters we may want to call species is a multi-level phenomenon relying on diverse sequences of cause-effect reinforcements of events affecting different levels, the very nature of this object of investigation requires a multi-level framework capable of accommodating a diverse set of species 'life histories.'

A multi-level framework approach is not meant to replace the plethora of species concepts in use today, but rather to integrate the various levels at which species are defined and highlight the causal connections between these levels. These connections are relevant to the formation and maintenance of species, yet are not captured by any of the available concepts, which, as discussed previously, tend to concentrate on a single level and to disregard prior causes and subsequent effects at other levels. A multi-level framework approach both supports and is supported by pragmatic pluralism. In as much as the inter-level dependencies associated with species formation/maintenance constitute a legitimate domain of investigation, a multi-level framework suitable for the investigation of these dependencies is on equal footing with the pattern-, process-, and mechanism-level concepts already

in place. At the same time, a multi-level framework can better justify pluralistic practices (i.e., justify them beyond the fact that they describe what seems to be a status quo successful scientific practice) by showing that different concepts target different cause-effect levels, and that they complement each other by shedding light onto different aspects of species formation/maintenance. In turn, this promotes the integration and transfer of knowledge between fields traditionally favoring different concepts.

According to a multi-level framework approach, the pattern-level grouping of organisms into diagnosable clusters and/or phylogenetic patterns is causally dependent on a multitude of possible processes, where at least some processes may be underpinned by a multitude of alternate molecular mechanisms regulating gene flow (Fig. 2). This does not entail, however, (i) that speciation/species maintenance can be reduced to strictly molecular explanations, or (ii) that species are best defined in terms of pattern because our understanding of the causes and mechanisms underlying speciation/maintenance of species can change in light of new discoveries, while the phenomenon to be explained, namely the pattern, remains always the same.

Claim (i) is rejected because it fails to take into account the multitude of causal pathways leading to the formation and maintenance of species. As discussed above, there is no unique level that captures all the causes, and not even the main or the triggering cause leading to the formation and maintenance of diagnosable clusters. Furthermore, what happens at the process-level of a population affects what happens at the mechanism-level within individual organisms, and vice versa. This dynamic interplay and reciprocal cause-effect reinforcement plays an important role in the formation and maintenance of species: gene flow and intrinsic reproduction isolation/cohesion mechanisms contribute to the formation and maintenance of diagnosable clusters; intrinsic mechanisms underlying reproductive isolation reinforce restricted gene flow; restricted gene flow contributes to the evolution of reproductive barriers by drift; divergent selection contributes to the reinforcement of restricted gene flow by selecting traits contributing to reproductive isolation/cohesion. Thus, what we witness at the pattern-level may be the result of multiple and/or alternate causal pathways, each underpinned by its distinct set of mechanisms.⁷

Claim (ii), according to which species can be defined independently of our understanding of the causes/mechanisms of speciation, highlights the relevance of pattern-level concepts as descriptions of the phenomena to be explained. However, it is not uncommon that the phenomenon under investigation is more and more precisely characterized as our understanding of its underlying cause(s) and mechanism(s) progresses.⁸ It is good scientific practice to delineate as clearly as

⁷ Claims that the process-level is best suited to define species [e.g., (Ereshefsky 2010)] can be ruled out for the same reasons.

⁸ Examples include the Hertz-Thomson characterization of cathode rays discussed by Achinstein (2001), modern nosological classifications whereby symptomatic descriptions encompass several distinct diseases defined in terms of etiology (main cause) and/or pathology (mechanism of the disease) [e.g., diabetes mellitus (Alberti and Zimmet 1998)], the splitting of the memory phenomenon into more and more specific kinds of memory (Roediger et al. 2002).

possible the phenomenon under investigation from other phenomena, accidental correlations, and experimental artifacts. Such efforts often involve several re-characterizations of the phenomenon under investigation in light of empirically supported findings about its causes and mechanisms.

In the case of species, strictly pattern-based characterizations can be problematic because recognizable patterns also occur within species (e.g., intra-population polymorphisms, races) and higher rank taxa (Ayala and Escalante 1996; Berlocher 1999), meaning that further criteria are needed in order to distinguish relevant from irrelevant patterns (Avice and Wollenberg 1997; Coyne and Orr 2004, 449). A multi-level framework makes it possible to justify relevance decisions that would have otherwise been arbitrary, as well as to disentangle complex, transient patterns associated with rapidly evolving organisms.

The first scenario, concerning decisions about which patterns are relevant, is illustrated by biological cases where, depending on the stringency of statistical relevance thresholds and the genes/traits taken into consideration, pattern-level concepts yield incongruent groupings. For example, in the case of genealogical species (GSC) discussed by Harrison, MHC gene trees support alternate groupings according to which humans and chimps (Ayala and Escalante 1996), or mice and rats (Figueroa et al. 1988) are one and the same species. The decision to acknowledge certain patterns (i.e., the consensus trees) while ignoring other patterns (alternative trees, such as MHC gene trees) is justified in light of the causal processes/mechanisms known or suspected to affect the organisms under investigation. In this particular case, it has been proposed that hybridization constantly injects new variants in a population, while selection favors the preservation of several alleles (balanced selection) (Tajima 1983), such that exclusivity at MHC loci is never attained (Hey and Kliman 1993). A partial understanding of what goes on at the process and mechanism levels (e.g., the adaptive value of MHC diversity and the presence of molecular mechanisms generating diversity) explains observed discrepancies in gene genealogies and provides threshold criteria for genealogical speciation [e.g., determine which and how many loci should be assessed and in what percentage they must concord in order to infer exclusivity (Avice and Ball 1990; Hudson and Coyne 2002)].

The second scenario, involving complex phenetic and phylogenetic patterns, constitutes a frequent issue in microbiology. For example, some HIV (Human Immunodeficiency Virus) strains are phylogenetically closer to SIV (Simian Immunodeficiency Virus) strains than other HIV strains, and vice versa. Most famously, HIV-1 is closer to SIVcpz than HIV-2 (Gao et al. 1998), while HIV-2 is closer to SIVsm than HIV-1 (Hirsch et al. 1989). This is due to the fact that, in all probability, the transmission of the retrovirus from non-human primates to humans occurred very recently and on several separate occasions. One reason for which HIV is nevertheless clearly distinguished from SIV is that, given the primary mode of transmission of these viruses, once the virus passed from non-human primates to humans, it is effectively ensured an independent evolutionary future likely to favor more and more pronounced pattern-level divergences between HIV and SIV. Furthermore, the probability of multiple infections is much higher for viral strains infecting the same host species, meaning that HIV strains are much more likely to

swap genes among themselves than with SIV strains, and therefore maintain their pattern-level relatedness. Thus, even though a concept such as BSC does not apply directly to viruses and other microbial parasites, which are almost exclusively classified in terms of phenetic clusters subsequently revised and refined in light of phylogenetic analysis, this is not to say that species differences among hosts do not have a say in determining whether the parasites infecting these hosts are likely to have an independent evolutionary history ultimately leading to changes observable at the pattern-level.

4 Conclusion

In this paper, I classify different species concepts according to the level at which they define groups of organisms we may want to call species. These are: the pattern-level grouping of organisms into discrete clusters; the level of the processes leading to the formation and maintenance of these clusters; or the mechanism-level of the genetic and molecular basis of cohesion/barrier mechanisms that contribute to species cohesion or promote speciation. The formation and maintenance of species is a multi-level phenomenon, typically requiring cycles of diverse reciprocal cause-effect reinforcements of events affecting different levels. I argue that a multi-level framework is required in order to track the complex sequences of events leading to the formation and maintenance of various patterns, as well as to allow for revisions and refinements of pattern-based characterizations in light of the gradual elucidation of the causes and mechanisms contributing to pattern formation and maintenance.

Acknowledgments I would like to thank Lindley Darden, Erika Milam, Eric Sidel, Pamela Henson, Joan Straumanis, the DC History and Philosophy of Biology group, as well as three anonymous reviewers and the editor for helpful comments and discussion. This work is supported by 'Fonds de la recherche sur la société et la culture', Québec, Canada [grant number 127231].

References

- Achinstein P (2001) *The book of evidence*. Oxford University Press, Oxford
- Alberti K, Zimmet P (1998) Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 15(7):539–553
- Avise J, Ball R (1990) Principles of genealogical concordance species in concepts and biological taxonomy. *Oxf Surv Evolut Biol* 7:45–67
- Avise J, Wollenberg K (1997) Phylogenetics and the origin of species. *Proc Natl Acad Sci* 94:7748–7755
- Ayala F, Escalante A (1996) The evolution of human populations: a molecular perspective. *Mol Phylogenet Evol* 5:188–201
- Baum D, Donoghue M (1995) Choosing among alternative “phylogenetic” species concepts. *Syst Bot* 20(4):560–573
- Baum DA, Shaw KL (1995) Genealogical perspectives on the species problem. In: Hoch PC, Stevenson AG (eds) *Experimental and molecular approaches to plant bsystematics*. Missouri Botanical Garden, St. Louis, pp 289–303
- Berlocher S (1999) Host race or species? Allozyme characterization of the ‘flowering dogwood fly’, a member of the *rhagoletis pomonella* complex. *Heredity* 83:652–662
- Bradshaw H, Schemske D (2003) Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* 426:176–178

- Coyne J (1992) Genetics and speciation. *Nature* 355:511–515
- Coyne J, Orr H (1998) The evolutionary genetics of speciation. *Philos Trans Royal Soc B* 353:287–305
- Coyne J, Orr H (2004) Speciation. Sinauer, Sunderland
- Cracraft J (1983) Species concepts and speciation analysis. *Curr Ornithol* 1:159–187
- Cracraft J (1989) Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. In: Otte D, Endler JA (eds) Speciation and its consequences. Sinauer, Sunderland
- de Queiroz K (1998) The general lineage concept of species, species criteria, and the process of speciation. In: Howard DJ, Berlocher SH (eds) Endless forms: species and speciation. Oxford University Press, New York, pp 57–75
- de Queiroz K (2005) Ernst Mayr and the modern concept of species. *Proc Natl Acad Sci* 102(1): 6600–6607
- de Queiroz K (2007) Species concepts and species delimitation. *Syst Biol* 56(6):879–886
- de Queiroz K, Donoghue M (1988) Phylogenetic systematics and the species problem. *Cladistics* 4:317–338
- Dobzhansky T (1935) A critique of the species concept in biology. *Philos Sci* 2:344–355
- Dobzhansky T (1937) Genetics and the origin of species. Columbia University Press, New York
- Ereshefsky M (1992) Eliminative pluralism. *Philos Sci* 59:671–690
- Ereshefsky M (2010) What's wrong with the new biological essentialism. *Philos Sci* 77:674–685
- Figueroa F, Günter E, Klein J (1988) MHC polymorphism pre-dating speciation. *Nature* 335:265–267
- Futuyma DJ (1998) Evolutionary biology, 3rd edn. Sinauer Associates, Sunderland
- Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, Sharp PM, Hahn BH (1998) Origin of HIV-1 in the chimpanzee pan troglodytes troglodytes. *Nature* 397:436–441
- Griffiths P (1999) Squaring the circle: natural kinds with historical essences. In: Wilson R (ed) Species: New Interdisciplinary Studies. MIT Press, Cambridge
- Harrison R (1998) Linking evolutionary pattern and process: the relevance of species concepts for the study of speciation. In: Howard D, Berlocher S (eds) Endless forms: species and speciation. New York, Oxford University Press
- Hey J, Kliman R (1993) Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. *Mol Biol Evol* 10(4):804–822
- Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR (1989) An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature* 339:389–392
- Hudson H, Coyne J (2002) Mathematical consequences of the genealogical species concept. *Evolution* 56(8):1557–1565
- Kirkpatrick M, Ravigné V (2002) Speciation by natural and sexual selection: models and experiments. *Am Nat* 159:S22–S35
- Kitcher P (1984) Species. *Philos Sci* 51:308–333
- Lidén M, Oxelman B (1989) Species—pattern or process? *Taxon* 38:228–232
- Luckow M (1995) Species concepts: assumptions, methods, and applications. *Syst Bot* 20:589–605
- Mallet J (1995) A species definition for the modern synthesis. *Trends Ecol Evol* 10:294–299
- Masly JP, Jones CD, Noor MAF, Locke J, Orr H (2006) Gene transposition as a cause of hybrid sterility in *Drosophila*. *Science* 313(5792):1448–1450
- Mayden RL (1999) Consilience and a hierarchy of species concepts: advances toward closure on the species puzzle. *J Nematol* 31(2):95–116
- Mayr E (1963) Animal species and evolution. Belknap Press, Cambridge
- Mayr E (1969) Principles of systematic zoology. McGraw Hill, New York
- Mayr E (1996) What is a species and what is not? *Philos Sci* 63:262–277
- Navarro A, Barton N (2003) Chromosomal speciation and molecular divergence—accelerated evolution in rearranged chromosomes. *Science* 300:321–324
- Noor M, Feder J (2006) Speciation genetics: evolving approaches. *Nat Rev Genet* 7(11):851–861
- O'Hara RJ (1993) Systematic Generalization, Historical Fate and the Species Problem. *Syst Biol* 42:231–246
- Orr H, Masly JP, Presgraves DC (2004) Speciation genes. *Curr Opin Genet Dev* 14(6):675–679
- Paterson H (1985) The recognition concept of species. In: Vrba E (ed) Species and speciation. Transvaal Museum, Pretoria
- Phadnis N, Orr H (2009) A single gene causes both male sterility and segregation distortion in *drosophila* hybrids. *Science* 323(5912):376–379

- Reydon T (2005) On the nature of the species problem and the four meanings of 'species'. *Stud Hist Philos Biol Biomed Sci* 36:135–158
- Roediger H, Marsh E, Lee S (2002) Kinds of memory. In: Pashler H, Yantis S, Medin D, Gallistel R, Wixted J (eds) *Stevens' handbook of experimental psychology*. Wiley, New York
- Schilthuizen M (2000) Dualism and conflicts in understanding speciation. *BioEssays* 22:1134–1141
- Schluter D (2009) Evidence for ecological speciation and its alternative. *Science* 323:737–741
- Shaw KL (1998) Species and the diversity of natural groups. In: Howard DJ, Berlocher SJ (eds) *Endless forms: species and speciation*. Oxford University Press, Oxford, pp 44–56
- Sneath PHA, Sokal RR (1973) *Numerical taxonomy: the principles and practice of numerical classification*. W. H. Freeman, San Francisco
- Tajima F (1983) Evolutionary relationship of dna sequences in finite populations. *Genetics* 105:437–460
- Templeton A (1989) The meaning of species and speciation: a genetic perspective. In: Otte D, Endler J (eds) *Speciation and its consequences*. Sinauer, Sunderland, pp 3–27
- Van Valen L (1976) Ecological species, multispecies, and oaks. *Taxon* 25:233–239
- Wiley E (1978) The evolutionary species concept reconsidered. *Syst Zool* 27:17–26
- Wu C (1994) Genetics of postmating reproductive isolation in animals. *Annu Rev Genet* 28:283–308
- Yu Q, Colot HV, Kyriacou CP, Hall JC, Rosbash M (1987) Behaviour modification by in vitro mutagenesis of a variable region within the period gene of *Drosophila*. *Nature* 326:765–769