
The internodal species concept: a response to ‘The tree, the network, and the species’

JOEL D. VELASCO*

Department of Philosophy, 5185 Helen C. White Hall, University of Wisconsin, Madison, WI 53706, USA

Received 17 January 2007; accepted for publication 9 June 2007

Samadi and Barberousse attempt to provide a precise, formal definition of the species category that is faithful to the internodal species concept (Samadi S, Barberousse A. 2006. The tree, the network, and the species. *Biological Journal of the Linnean Society* **89**: 509–521). Here, it is argued that their study is technically flawed. Most of the necessary corrections to their definitions are provided in order to accurately portray the internodal concept. It is then argued that the internodal concept is fundamentally flawed; it does not solve the ‘classical problems’ that Samadi and Barberousse claim it does. In particular, it does not allow for the possibility of interspecies hybridization. In addition, the proposal is unworkable in practice, and also theoretically problematic because it entails that, in many lineages, speciation events are taking place every few generations. © 2008 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2008, **93**, 865–869.

ADDITIONAL KEYWORDS: speciation – systematics.

INTRODUCTION

Sarah Samadi and Anouk Barberousse’s ‘The tree, the network, and the species’ (Samadi & Barberousse, 2006: 510) attempts to provide a precise, formal definition of the species category that is faithful to the internodal species concept favoured by many authors. That concept is simply the idea that, as Samadi and Barberousse (henceforth ‘S&B’) put it, ‘A species is thus a branch segment of the tree of life delimited either by two nodes or branching points (i.e. by two speciation events) or by a node and the end of a branch (i.e. an extinction event).’

Kornet (1993) gives a formalization that intuitively matches well with the internodal concept, which Kornet, Metz, & Schellinx (1995) claim uniquely identifies the internodal concept. S&B admit that their definition is close to that of Kornet, but they claim it has a major advantage: ‘it strictly derives from, and is motivated by, the theory of evolution’ (Samadi & Barberousse, 2006: 510). I argue to the contrary that S&B have not actually produced a species definition at all,

that the concept they are apparently aiming for has no advantages over Kornet’s, that it is not justified by the theory of evolution, and, despite their claims to the contrary, that it fails as both a practical and a theoretical proposal.

THE S&B SPECIES CONCEPT

Before we can judge the success of S&B’s concept, we need to carefully describe exactly what this concept is. Since they claim that genealogical history is the one and only factor which matters for classification, species membership must be completely fixed by the giant, or global genealogical network (GGN) that connects all organisms via the relationship of reproductive descent. To do so, S&B define a relationship R which they then use in defining species classes. Here is what they say:

R is defined as holding between two organisms a and b in a generation G_i such that aRb , if a and b have common offspring in generation G_{i+1} . The relationship R , which is an equivalence relation (see above), formalizes the concept of community of descent upon which the unicity of the genealogical pattern of the history of life is based... It must be stressed

*Corresponding author. E-mail: jdvelasc@wisc.edu

that the introduction of relationship R is the key to obtaining a definition of species as equivalence classes of organisms. (Samadi & Barberousse, 2006: 512)

Notice that S&B claim that R is an equivalence relation and, furthermore, that this fact is central to their project (it does play an important role in their ‘proof’ that their definition of species yields equivalence classes.) However, R is clearly not transitive. In order for R to be transitive, it would have to be the case that if a and b share an offspring and b and c share an offspring, then a and c must also share an offspring. A moment of thought shows that this violates what we know about many types of sexual organisms with exclusively male and female types; the relation is not transitive and therefore is not an equivalence relation. Since R is not an equivalence relation, F , the familial relation, will not be either. This means that the algorithm that S&B provide for dividing the GGN into equivalence classes fails.

FIXING THE DEFINITIONS

Although the definition that S&B give is not enough to divide all organisms into equivalence classes, it is clear what they intend to do, and thus we can easily reconstruct a definition that works. We can then examine whether we should accept this revised definition.

Instead of using the supposed fact that R is an equivalence relation, we can simply form our classes in a way that forces them to be equivalence classes. One way to do this is to form $[x]$ (the equivalence class that contains x) directly by saying that $[x]$ is the smallest class such that $[x]$ contains x and for all $y \in [x]$, if yRz then $z \in [x]$. For example, let us imagine that aRb and bRc . Now $[a]$ contains a (by definition) and since aRb , $b \in [a]$. Now since $b \in [a]$ and bRc , $c \in [a]$ as well, so $[a]$ contains at least three members. If a , b , or c share offspring with anyone else, they are in $[a]$ as well. This may have been what S&B had in mind all along, although this is not clear from their text.

The next step is to use these classes to get the ‘family’ of x which consists of every organism in $[x]$ and any offspring of any organism in $[x]$. Thus, everyone is in the same family as their parents, any offspring of either of their parents, any parents of these other offspring, etc. Now that we can divide two generations into sets of families, we are ready to form lineages. What follows is S&B’s definition of a lineage with a few necessary corrections placed in bold brackets. In addition to those changes, in order for the definition to be technically correct, we must understand ‘a set of families’ to mean the union of that set. That way, individual organisms, rather than families, can be members.

A family lineage $L(j, n)$ between generations G_j and G_n is a set of families $F(i, i + 1)$ (i varying from j to n). [should vary from j to $n - 1$]. It is a subset of family partitions [should be a set of families] for each pair of consecutive generations between G_j and G_n , fulfilling the following three conditions:

- 1 All the parents P_{j+1} of the set of families $F(j + 1, j + 2)$ are the offspring of the offspring D_j [should be ‘are the offspring of the parents P_j ’] of the set of families $F(j, j + 1)$.
- 2 All the offspring D_j of the set of families $F(j, j + 1)$ leaving offspring at the following generation G_{j+2} [the ‘leaving offspring. . .’ clause is unnecessary] belong to the set of families $F(j + 1, j + 2)$.
- 3 $L(j, n)$ cannot be subdivided into several family lineages as defined by the first two conditions. (Samadi & Barberousse, 2006: 513)

As was true of the R relation, it is fairly clear what S&B want a lineage to be; yet they do not give a correct formal definition. As before, this is easy to fix. The importance of this definition of lineage is that it places each organism in any given timespan (between G_j and G_n) into exactly one lineage $L(j, n)$. Since families are (supposedly) exclusive, lineages (over the same generations) are too.

Here the importance of each organism being placed into a particular generation becomes clear. S&B define a generation as:

the set of all organisms leaving offspring between two close instants of time, with the length a generation lasts depending on the taxa considered. This notion is an idealization, for when generations overlap, adult organisms belong to several generations. Such an idealization is, however, justified by the fact that the mathematical structure of the genealogical network remains unaltered when generations overlap. (Samadi & Barberousse, 2006: 512)

First, notice that this definition places only organisms with offspring into generations and therefore into lineages. Although this problem could be fixed, there is a far more serious problem: organisms from different generations cannot share offspring. A simple example of an incestuous relationship between parent and offspring provides a direct counterexample, but no such incest is required. Any generational overlap will cause families to overlap and thus not form equivalence classes. It is therefore false that the mathematical structure ‘remains unaltered’. It is unclear how any concept of generation could uniquely place every organism in exactly one generation, but this is what is required if we are to create families and lineages in the way that S&B suggest.

For now, let us assume that S&B could solve this problem. The next step is to determine what speciation is. Although S&B repeatedly refer to a speciation event as a permanent lineage divergence, they fail to give a definition for speciation that matches this idea. Incidentally, S&B believe that anagenesis does occur

(Samadi & Barberousse, 2006: 511) although this is clearly inconsistent with any internodal species concept. Attempting to define speciation as lineage branching, here is what S&B say:

Assume that there is a speciation at generation G_i . A new speciation occurs at generation G_j (with $j > i + 1$) in the family lineage $L0(i)$ (spreading between G_i and the end of the GGN) if and only if: (i) there exist n ($n > 2$) family lineages (spreading between G_i and the end of the GGN) noted $L1(j), L2(j), \dots, L_n(j)$, and including $L0(i)$ and (ii) there exists a unique family lineage (spreading between G_{j-1} and the end of the GGN), and including $L0(i)$. (Samadi & Barberousse, 2006: 514)

Besides the strange usage of lineage names (which is inconsistent with the rest of their paper), there are two problems with the definition. First, it entails that the number of family lineages between G_j and the end of the GGN has to be greater than two. Typically, we imagine a lineage splitting into exactly two new lineages; it certainly does in all of the diagrams that S&B give. With the definition they give as is, none of their figures contain any speciation events. Let us assume that this is simply a mistake and that S&B meant that n must be greater than 1.

The other problem is that S&B do not allow speciation events to take place in consecutive generations. Combined with the fact that there must be a unique lineage between G_{j-1} and the end of the GGN, this causes problems. For example, if there is a speciation at G_i splitting into lineages A and B and then a permanent lineage split at G_{i+1} in lineage A , then no matter what else happens in lineage A (including any number of further splittings at any later generations), there could never be a speciation in that lineage since there will never be a unique lineage at any G_{j-1} . As a strange example (but not unrealistic given the way fission occurs), imagine a large chunk of the network that is bifurcating each generation. By this definition, there could not be any speciations occurring although there are constant lineage splits.

Once again, this problem is easily fixed. Simply change the requirements on j and n to $j > i$ and $n > 1$. This will match speciation with permanent lineage splitting. Extinction is defined in the obvious way: a lineage with members at G_i but none at G_{i+1} .

Given all the logical definitions so far, it is surprising that S&B do not actually give a careful logical definition of species. S&B say that a species is simply a lineage between two speciation events or between speciation and extinction. This is not a fully determinate logical definition, but it is easily made into one. Let us say that two organisms, x and y , are members of the same species iff x and y are members of some lineage Z where x is a member of generation G_i , y is a member of generation G_k , and no speciation event has occurred in lineage Z at any j such that $i < j \leq k$.

This puts organisms 'at the border' with new species rather than old. Since speciation events are defined with reference to the full GGN, no species categorizations can be determined by looking only at the generations between i and k .

IMPLICATIONS OF THIS SPECIES CONCEPT

Now that we have finally partitioned the set of organisms in the way that S&B want, can we safely identify these equivalence classes with species? No, we cannot unless we are ready to swallow some very counterintuitive conclusions. S&B argue that their view is superior to alternatives in the literature because it solves three 'classical problems'. I argue that, in each case, S&B's criterion is no better than major alternatives and is clearly worse than some.

UNIPARENTAL REPRODUCTION

In their discussion of uniparental organisms, S&B seem to think that asexual reproduction is a problem only for versions of the biological species concept (the BSC) because the concept of interbreeding does not make sense for uniparentals. However, their own species concept is no better. Since every single asexual lineage is divergent from every other (unless one is just a subset of the other), if any organism has two clonal offspring that go on to reproduce, there are now two divergent lineages and thus a speciation event.

Many authors, such as Templeton (1989) and Van Valen (1976), who criticize versions of the BSC because of how they deal with asexual organisms, would certainly think that S&B do not do any better. The problem of uniparental reproduction is that there are genuinely coherent groups of asexual organisms that should be considered a species. Although S&B's criterion does apply to asexuals, it implies that they are very often in a species with just two or three members and so their proposal does not solve the problem.

TEMPORAL DEPTH

It is unclear what S&B mean by the problem of 'temporal depth.' They say that evolution requires that species be 'temporally extended units of evolution'. Their solution is to require that species be historical entities. Of course it is not at all clear that an historical entity is automatically going to be a 'unit of evolution.' Whatever that means, if S&B's criterion solves this problem, it would seem that any lineage-based view solves it as well. As de Queiroz (1988) points out, this includes just about every reasonable view put forward in the literature. So S&B's concept

has no advantage over competitors. In fact, as we will see, S&B's concept implies that species generally exist for only a few generations, so it is their concept, not others, that actually has a problem with temporal depth.

HYBRIDIZATION

Here, the problem is that many views have difficulty saying what happens during hybridization, which is usually defined as two organisms of different species producing offspring. S&B's 'solution' is to simply deny the existence of the phenomenon. In their view, hybridization is actually impossible because, if two organisms share an offspring, then they must be members of the same species. S&B severely understate their own position on this and do not appear to recognize the consequences of their own view. They say 'This does not mean that hybridization events are evolutionarily unimportant, for they may give rise to new genuine species.' Their caption in figure 5 implies that this occurs when a new lineage is 'reproductively isolated from their parents.' (S&B, 2006: 516). However, their definition of species does not have anything to do with reproductive isolation and, since a new species requires a speciation event (i.e. a lineage splitting), new hybrid species cannot form in the way they describe. Thus, S&B are simply mistaken about the consequences of their own view.

DIAGNOSIS OF THE PROBLEM

Just based on the 'classical problem' of hybridization above, we can see that S&B's concept is not acceptable. Some lineage splits that should be genuine speciation events are not really permanent splits if we are looking at the organismal level. This is a general problem with any internodal concept and is not specific to S&B's theory. Kornet herself acknowledges that this is extremely counterintuitive, but she thinks that no other species concept really does any better. Perhaps she is right if we refuse to allow any vagueness at all in the definition of speciation, but a more promising way to proceed is to acknowledge that this concept is vague and proceed from there.

Although it may be correct to say that a species is a segment of a lineage between speciation events (and many authors have said just that), the difficult part is to say what constitutes a speciation event. We can say that a speciation is 'a kind of lineage branching' but, as Baum & Shaw (1995) have pointed out, it is not at all easy to say what constitutes a 'significant enough' lineage branching, especially since part of being significant is that it will have lasting effects in the future, which cannot be seen at the time. S&B solve this problem by assuming: (1) all definitions such as

that of lineage, speciation, and species are determined at the end of time after all life is gone and (2) all and only permanent divergences count as speciation. Of course (1) causes obvious practical problems (and possibly metaphysical problems – am I a member of any species now? If so, does that mean the future is determined, etc.) but let us focus on problems with clause 2.

With hybridization, we have seen that there is a problem in determining whether a temporary lineage split is significant enough to warrant calling it a speciation event. It does not seem correct to demand that the split must be permanent. Current systematic practice dictates that hybridization is a fairly common event and so any view that implies it is impossible differs wildly from this practice.

This problem appears to be that some significant divergences are not really permanent lineage splits. Although this is certainly a difficult problem, an even more severe problem is that S&B's concept creates far too many groups. In other words, there are many permanent divergences that are not significant. As Kornet (1993) points out, permanent splits are very common, often occurring only a few generations apart. For example, imagine that *a* and *b* have an offspring *c* (in generation G_i) that dies without producing any offspring of its own. If neither *a* nor *b* have any other offspring, then this constitutes an extinction event and the lineage traces back until it reaches organisms that do have descendants in the G_{i+1} generation. Thus, if *a* or *b* had siblings that have grandchildren who would be in generation G_{i+1} , then {*a*, *b*, *c*} constitute an entire species unto themselves. An asexual that has only one offspring, and which then does not reproduce, could be a member of a species of two. These toy examples are extremely common and extend to similar cases so that lineages consisting of only a very small number of generations are going to be the norm rather than the exception. Assuming no future lineage convergences, S&B's own examples of genealogies (figs 2, 4) contain 86 organisms in seven species and 151 organisms in seven species, respectively. Kornet's example, figure 7 (which she has delineated for us), consists of 192 organisms in 14 species. These numbers are simply biologically unrealistic.

It seems clear that if we look at the genealogies of individual organisms, either Kornet's concept or that of S&B is going to have far too many equivalence classes to allow us to identify these classes with species. Kornet acknowledges that these groups cannot be species because they are far too small, but she thinks that they are theoretically important in building larger groups that might then be given the status of a species. Perhaps this provides a way forward for a concept related to the internodal species

concept, but then we are back to the problem of determining which lineage splits are significant enough to be speciation events. Kornet & McAllister (2005) answer this question by using character states as criteria to determine which splits are speciation events. Whether or not these are the right criteria, this view abandons (appropriately) the idea that species are determined completely by the reproductive relationships between organisms which is the stated foundation for S&B's internodal species concept.

REFERENCES

- Baum DA, Shaw KL. 1995.** Genealogical perspectives on the species problem. In: Hoch PC, Stephenson AC, eds. *Experimental and molecular approaches to plant biosystematics*. St Louis, MO: Missouri Botanical Garden, 289–303.
- Kornet DJ. 1993.** Permanent splits as speciation events: a formal reconstruction of the internodal species concept. *Journal of Theoretical Biology* **164**: 407–435.
- Kornet DJ, McAllister JW. 2005.** The composite species concept: a rigorous basis for cladistic practice. In: Reydon TAC, Hemerik L, eds. *Current themes in theoretical biology*. Dordrecht: Springer, 95–127.
- Kornet DJ, Metz JAJ, Schellinx HAJM. 1995.** Internodons as equivalence classes in genealogical networks: building-blocks for a rigorous species concept. *Journal of Mathematical Biology* **34**: 110–122.
- de Queiroz K. 1988.** The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In: Howard DJ, Berlocher SH, eds. *Endless forms: species and speciation*. Oxford: Oxford University Press, 57–75.
- Samadi S, Barberousse A. 2006.** The tree, the network, and the species. *Biological Journal of the Linnean Society* **89**: 509–521.
- Templeton AR. 1989.** The meaning of species and speciation: a genetic perspective. In: Otte D, Endler JA, eds. *Speciation and its consequences*. Sunderland, MA: Sinauer Associates, 3–27.
- Van Valen L. 1976.** Ecological species, multispecies, and oaks. *Taxon* **25**: 233–239.