

CHANCE AND MACROEVOLUTION*

ROBERTA L. MILLSTEIN†‡

Department of Philosophy

California State University, Hayward

*Received July 2000; revised August 2000.

†I would like to thank John Beatty, Ron Giere, and Ken Waters for their helpful comments on earlier versions of this paper. I also thank the anonymous referees for their constructive remarks. This work was supported by a Faculty Support Grant from California State University, Hayward.

‡Send reprint requests to the author, Department of Philosophy, California State University, Hayward, 25800 Carlos Bee Boulevard, Hayward, CA 94542.

Abstract

When philosophers of physics explore the nature of chance, they usually look to quantum mechanics. When philosophers of biology explore the nature of chance, they usually look to microevolutionary phenomena such as mutation or random drift. What has been largely overlooked is the role of chance in macroevolution. The stochastic models of paleobiology employ conceptions of chance that are similar to those at the microevolutionary level, yet different from the conceptions of chance often associated with quantum mechanics and Laplacean determinism.

1. Introduction. Whereas microevolution concerns changes at or below the species level, macroevolution concerns changes at or above the species level. Paleontologists are often interested in the macroevolutionary processes that have led to the formation and extinction of species and higher categories as exhibited in our fossil record (part of the sub-discipline of paleontology known as paleobiology). In 1972, a group of paleontologists held two “informal meetings” at the Marine Biological Laboratory in Woods Hole, Massachusetts. The outgrowth of these and other subsequent meetings was a series of papers suggesting that stochastic processes might play a large role in the evolution of species and higher orders. The main tool of the researchers (who came to be known as the “Woods Hole Group”¹) was a computer program known as the MBL program, which used a Monte Carlo approach to model the macroevolutionary processes of speciation and extinction. Comparing the results of the model to fossil data, the researchers called into question the ubiquity of so-called “deterministic” explanations for macroevolutionary phenomena, arguing that stochastic processes could account for the same data.

I say “so-called” because when paleontologists refer to “deterministic” explanations, they are referring to a particular kind of explanation rather than to an ontological claim about the nature of the universe. They do not mean “deterministic” in the Laplacean sense employed by most philosophers (where a deterministic universe is one in which any given state uniquely determines a future state). Similarly, proponents of stochastic models do not deny

Laplacean determinism. Rather, both deterministic and stochastic macroevolutionary explanations are consistent with either Laplacean determinism or indeterminism. Thus, the stochasticity of the models does not depend on the stochasticity of the underlying phenomena; the underlying phenomena could be ontologically random or Laplacean deterministic (Raup et al. 1973, Gould et al. 1977). The macroevolutionary senses of “deterministic” and “stochastic” will be discussed further below.

The stochasticity of macroevolutionary processes has gone largely unexamined by philosophers of biology;² analyses of the role of chance in evolution focus primarily on random drift and mutation, quintessentially microevolutionary phenomena. Yet the research of the Woods Hole group is striking in what it reveals about concept of chance and the nature of explanation in evolutionary theory. Furthermore, there are many interesting parallels between chance at the microevolutionary level (i.e., random drift) and chance at the macroevolutionary level. In this essay, I will explore these issues.

The structure of the essay is as follows. First, I describe the deterministic explanations that have been traditionally been given in paleobiology, and contrast these explanations with the stochastic model of the MBL program. Second, I explore the question of what makes the “stochastic” model “stochastic” by discussing the various conceptions of chance implicit in the model. Third, I discuss the apparent explanatory pluralism that is suggested by the addition of stochastic models to the more traditional deterministic explanations. I then explain how it is possible for the stochastic model to account for phenomena that have been traditionally explained deterministically, through a discussion of the nature of a Markov

process. Finally, I summarize the similarities between the random drift and stochastic macroevolutionary processes, and suggest some broader conclusions about chance within evolutionary theory.

2. Background: Deterministic Explanations vs. Stochastic Models.

2.1. Deterministic Explanations in Paleobiology. Before we examine the stochastic macroevolutionary models of the MBL program, it will be helpful to have an understanding of more traditional, deterministic explanations.³ Peterson (1983) provides an example of this kind of explanation. Peterson argues that four genera of iguanid lizards (*Anolis*, *Chamaeleolis*, *Chamaelinorops*, and *Phenacosaurus*, referred to as anoline lizards or anoles) have been more successful than other iguanid lizard genera due to their possession of a subdigital adhesive pad:

The pad complex appears to be a key innovation (*sensu* Liem, 1973) in the successful radiation of the group. Compared to animals that must rely on their claws and toe position for grip, anoles appear to have greater range of usable locomotor substrates. They can grip or adhere to a perch of almost any diameter (or radius of curvature), spatial orientation, or surface texture (e.g., surfaces as smooth as glass or as rough as bark). It may become possible to state the relationship between the pad complex and physical aspects of the habitat more precisely in the future, but the adhesive pad seems to be among the morphological adaptations that permit anoles to adopt highly acrobatic

locomotor behavior and to exploit the spatial heterogeneity of arboreal niches (Peterson 1983, 245).

In this explanation, the phenomenon being explained is the successful radiation of the anoline lizards, i.e., the greater number of anoline lizard species as compared to other species of iguanid lizards. (In macroevolutionary contexts, a “radiation” is a marked increase in the number of sub-groups within a group; in this case, the radiation in question is the increase in the number of species in the four anoline lizard genera). According to Peterson, there are more anoline lizard species than other iguanid species because the possession of a subdigital adhesive pad provided the anoline lizards with a superior grip, giving them greater agility and enabling them to more fully explore their habitat as compared to lizards which must rely on claws and toe position. This is referred to as a “deterministic” explanation because it invokes a specific cause (the presence of the sub-digital adhesive pad) as an explanation for the macroevolutionary phenomenon in question (the greater radiation of anoline lizards). Such an explanation is “taxonbounded” in that it does not treat all taxa alike; rather, it claims that the physical characteristics of some taxa (the anoline lizard genera) caused them to have a higher rate of branching than other taxa (non-anoline iguanid lizard genera).

Taxonbounded explanations are one kind of deterministic explanation. Another kind of deterministic explanation is a “timebounded” explanation (Gould et al. 1977). Timebounded explanations do not treat all times alike. For example, one could give a timebounded explanation for why mammals were a very small group up until the Tertiary period; prior to

that time, they were kept small by a successful competitor: the dinosaurs. Once the dinosaurs became extinct, times were “good” for mammals, and the number of genera expanded (Gould et al. 1977, 37). Because this explanation presupposes that some times are “better” for a group than others, it is considered timebounded. Other timebounded explanations claim that taxa have specifiable life histories – for example, that older taxa are diverse early in their history, whereas younger taxa in the same group are diverse late in their history (Raup 1985, 42). Since these explanations invoke specific causes (in these cases, specific time periods) as the explanation for the phenomenon in question, they too are considered deterministic explanations.⁴

2.2. The MBL Program: A Stochastic Model for Macroevolution. In contrast to deterministic explanations, stochastic models of macroevolution are neither taxonbounded nor timebounded; that is, they “treat all times and taxa alike” (Gould et al. 1977, 25).⁵ In the discussion that follows, I will show how the concepts of untimeboundedness and untaxonboundedness are embodied by the model used in the MBL program. Although elements of this model have been criticized since it was first proposed, and alternative models have been proffered, the MBL model (as described in Raup et al. 1973, Gould et al. 1977, and Raup 1977) continues to be cited in textbooks (see, e.g., Futuyma 1986, Ridley 1993) and recent papers (see, e.g., Uhen 1996, Eble 1999). My focus will therefore be on the model used by the Woods Hole Group in order to: 1) describe the “classic” stochastic model from which other models are derived, and 2) to delineate the basic elements of stochastic

models (the untimebounded, unaxonbounded, and Markovian qualities) without getting bogged down in the details of the differences between different models. With these goals in mind, in the description of the MBL program that follows, I will discuss the details of the program only insofar as they pertain to the discussion of stochasticity. For further details of the program, see Raup et al. 1973 and Gould et al. 1977.

The MBL program attempts to model the evolutionary process by producing simulated phylogenetic trees. (A phylogenetic tree is a tree that attempts to depict the genealogical history of a group of organisms). It begins with a single lineage, from which a “life history” is generated. At each time interval, and for each current lineage, one of three things occurs: 1) The lineage becomes extinct, 2) The lineage splits into two daughter lineages, or 3) The lineage persists to the next interval without branching.⁶ Which of these three fates occurs is determined by a computer-generated random number, based on predetermined probabilities. There are two main versions of the MBL program. In one version, the “freely floating” model, the probabilities of branching and extinction are equal (probability of branching = probability of extinction = 0.1) and unchanging throughout the simulation (Gould et al. 1977). In the other version, the damped-equilibrium model, there is an optimum number of co-existing lineages for any interval (the same optimum number applies to all intervals), and the probabilities of branching and extinction are adjusted to obtain this optimum value.⁷ This adjustment occurs by raising the probability of branching when the number of lineages is below the optimum number, and by raising the probability of extinction when the number of lineages is above the optimum number. Thus, in the initial phase of the program, where the

number of lineages is few, the probability of branching will exceed the probability of extinction, causing an “explosion” of lineages (Gould et al. 1977).

Both of these stochastic models can be said to be “untaxonbounded” and “untimebounded” to some extent. Both are untaxonbounded because the same probabilities of branching and extinction apply to all current, endpoint lineages in the simulation. Thus, all taxa are treated alike. This contrasts with the taxonbounded, deterministic explanation given by Peterson, where the rate of extinction was less for some taxa (the anoline lizards) than it was for other taxa (non-anoline lizards), and the rate of branching was greater for some taxa (the anoline lizards) than it was for other taxa (the non-anoline lizards).

The freely floating model is completely untimebounded; the probabilities of branching and extinction remain the same throughout all the intervals in the simulation.⁸ The damped-equilibrium model is untimebounded in the sense that the optimum number of lineages applies to all time intervals. However, it is not completely untimebounded because not all time intervals are alike; as previously mentioned, in the initial phase of a simulation the probability of branching exceeds the probability of extinction. However, for the rest of the simulation, the model is untimebounded; although the probabilities of branching and extinction change, they do not do so with respect to any particular phase of the simulation. Thus, all times after the initial phase are treated alike. This contrasts with the timebounded, deterministic explanation given above, where prior to the Tertiary period, mammals were not very successful (high extinction rate, low speciation rate), but once dinosaurs became extinct, times were “better” for mammals (low extinction rate, high speciation rate).

Thus, explanations that are both untimebounded and untaxonbounded are referred to as stochastic; explanations that are either timebounded or taxonbounded are referred to as deterministic. As mentioned above, this sense of determinism is distinct from Laplacean determinism; deterministic explanations might also be called “determinate cause” explanations, since they refer to specific causes related to the properties of particular lineages or time periods.

3. The Stochasticity of Stochastic Models. So, if stochastic models are not a denial of Laplacean determinism, in what sense can stochastic accounts be said to be stochastic? Eble’s recent discussion of the various conceptions of chance used in microevolutionary and macroevolutionary contexts provides a useful starting point. After an analysis of these conceptions, I will show how they apply to the stochastic macroevolutionary models of the MBL program.

3.1. Eble’s Conceptions of Chance. Eble (1999) characterizes five (not entirely distinct) conceptions of chance that are potentially relevant to evolutionary theory. First, chance sometimes means “uncaused,” implying a denial of Laplacean determinism. This is the conception frequently associated with quantum mechanics and nonequilibrium thermodynamics. Second, chance sometimes means “not by design or deliberate plan,” a conception associated with Paley and natural theology. According to Eble, neither of these two conceptions is particularly relevant to present-day evolutionary theory.⁹ Of greater

relevance, according to Eble, are two conceptions which taken together elaborate what he calls the “statistical meaning” of chance:

(1) An event occurs at random because it is unpredictable, due to our ignorance of causes. This ‘ignorance’ interpretation (dating back to Laplace) is probably the most frequently used. It is usually associated with the assumption of probabilistic behavior and indiscriminate sampling (viewed as a property of independent events in nature, not of experimental design...). (2) An event may be said to be random when it is the result of the confluence of independent causal chains, as in traffic accidents (this “coincidence” interpretation stems from Aristotle)... In effect, these two versions of chance imply and complement each other, and both stand as renditions of the statistical meaning of chance (1999, 76-77).

Eble argues that the fifth conception of chance, the “evolutionary meaning” of chance, is also highly relevant to present-day evolutionary theory. The evolutionary meaning of chance is when “*events are independent of an organism’s need and of the directionality provided by natural selection in the process of adaptation*” (Eble 1999, 77; italics in original). For example, the claim that mutations are due to chance usually implies this meaning – chance mutations are mutations that can be beneficial, harmful, or neutral. They are not the result of an organism’s need.

Without taking a position on Eble's claim that it is primarily the statistical and evolutionary conceptions of chance that are important in present-day evolutionary theory, I will follow Eble in focusing my attention on these conceptions. Eble's conceptual scheme provides a useful starting point for a discussion of the meaning of chance in evolutionary theory, but it is in need of further clarification. In particular, there are three conceptions of chance implicit in Eble's "statistical meaning," not two. In the first of the two statistical meanings of chance in the quote above, Eble conflates "chance as ignorance" with "chance as indiscriminate sampling." However, a simple example will show that these two meanings are not identical. Consider the canonical example of indiscriminate sampling, the blindfolded sampling of colored balls from an urn. Suppose the balls are identical except for the fact that some are red and some are green. One person, fully blindfolded, picks balls out of the urn, while another person watches. We can even suppose that our observer is Laplace's demon himself, an observer who is aware of the picker's mental states and any other factors that might be relevant to the picking of balls from the urn.¹⁰

According to such an observer, the picking of balls would not be chance as ignorance, since he is fully aware of all the relevant causes. However, would the demon say that he had observed indiscriminate sampling? In one sense, perhaps not, because the demon would not claim that each ball was equally likely to be picked – given the mental states of the picker and the physical distribution of the balls, some balls get picked whereas other balls do not. However, in another, important sense, the demon would claim that he *had* observed indiscriminate sampling – namely, sampling that was indiscriminate with respect to color. In

other words, whatever else may be causally relevant to which balls were picked, the *color of the balls* is not causally relevant to which ball is picked. Thus, although the demon may not lay claim to indiscriminate sampling *simpliciter*, he would concede to indiscriminate sampling with respect to color, in spite of the fact that he is not ignorant of the relevant causal factors. The demon would acknowledge that the picking of the balls involved “chance as indiscriminate sampling,” where the sampling occurs with respect to a given property, although it does not involve “chance as ignorance.”

This example serves to show that “chance as indiscriminate sampling” does not imply “chance as ignorance.” However, “chance as ignorance” does not imply “chance as indiscriminate sampling,” either. For example, in discussing the meaning of “chance variations,” Darwin stated that by “chance” he meant only to acknowledge “our ignorance of the cause of each particular variation” ([1859] 1964, 131). There was no indiscriminate sampling involved – nor was there necessarily any “chance as coincidence” – yet Darwin invokes “chance as ignorance.” I would thus suggest that “chance as ignorance” is a notion of chance that has applications outside of the statistical realm, and is not properly part of the statistical meaning of chance.

But what of the two remaining “statistical meanings” of chance – “chance as indiscriminate sampling” and “chance as coincidence” – do they imply one another? Consider the example Eble mentions, a traffic accident – “the result of the confluence of independent causal chains” (1999, 76). In other words, the causal chain of events that leads one car to be at a particular place and time is independent from the causal chain of events that

leads a second car to be at the same place and time, leading to the accident. It is hard to see how this could be construed as indiscriminate sampling, *unless* one considers a “population” of cars involved in traffic accidents – “sampling” implies that there is a population in a way that “coincidence” does not. If a population of cars is examined, one might discover that car color is causally irrelevant to traffic accidents, so that there would be indiscriminate sampling with respect to color. On the other hand, one might discover – as American car insurance companies claim – that red cars are more likely to be involved in traffic accidents than other car colors. (We can characterize such an event as *discriminate* sampling, where color *is* causally relevant to traffic accidents.) Of course, there may be indiscriminate sampling with regard to some *other* property besides color, but there is no guarantee that a population of coincidental events will involve indiscriminate sampling. Thus, “chance as indiscriminate sampling” and “chance as coincidence” are not inextricably connected.

What of Eble’s “evolutionary meaning” of chance – not caused by an organism’s need – how does that relate to the statistical meanings? Eble claims that these meanings are also distinct, although they are often used simultaneously in evolutionary explanations. In the discussion that follows, I will argue that Eble (1999) is correct in that these meanings are sometimes used simultaneously, through an examination of the stochastic macroevolutionary models of the MBL program. I will also compare and contrast this usage to the conceptions of chance involved in random drift, an area that Eble (1999) argues uses both the statistical and evolutionary meanings of chance. However, given the independence of the three “statistical meanings” I will refer to “chance as indiscriminate sampling,” and “chance as

coincidence” separately. These meanings may or may not occur in concert with the Laplacean notion of “chance as ignorance,” and are thus distinct from it.

3.2. Eble’s Chances in Stochastic Macroevolutionary Models. As we saw above, stochastic models are characterized by explanations that are untimebounded and untaxonbounded. These characteristics imply that differences between taxa are causally irrelevant to differences in rates of branching and extinction within the taxa (this is the untaxonbounded element) and that different time periods are causally irrelevant to differences in rates of branching and extinction (this is the untimebounded element). In other words, stochastic models involve “chance as indiscriminate sampling” – sampling that is indiscriminate with respect to differences in time intervals and taxa. For example, physical differences between genera do not confer different rates of speciation and extinction upon different species of the genus (as would be the case in a deterministic, taxonbounded explanation, such as the one given for anoline lizards above); rather, origination and extinction occur randomly among the different *species* of the taxa.

The analogy between random drift and the stochastic macroevolutionary process implicit in the MBL program is striking; however, they are not identical processes. In random drift, physical differences between *organisms* are causally irrelevant to differences in reproductive success; that is, there is indiscriminate sampling with respect to physical differences between organisms. An untaxonbounded macroevolutionary explanation, on the other hand, claims that physical differences between *taxa* are causally irrelevant to differences in rates of

branching and extinction within the taxa (differences between the properties of the higher taxonomic group are not causally relevant to the branching and extinction of the taxa of the lower category).¹¹ That is, there is indiscriminate sampling with respect to physical differences between taxa. Both processes involve “chance as indiscriminate sampling,” but the sampling takes place at different levels; random drift occurs at the microevolutionary level, whereas the stochastic model we have been discussing pertains to the macroevolutionary level. Also, there does not seem to be an analog of “untimeboundedness” in the random drift case. The issue of whether some times are “good” or “bad” for organisms does not seem to arise in the microevolutionary case as it does in the macroevolutionary. Nonetheless, stochastic models and random drift share a further similarity: both incorporate “evolutionary chance.” That is, since neither samples discriminately with respect to physical differences (as would be the case with natural selection), both processes are directionless with respect to adaptation. They may lead in an adaptive direction, but they may equally well lead in a neutral or even harmful direction.

So, stochastic models of evolution are stochastic in the sense that they reflect “chance as indiscriminate sampling” and “evolutionary chance.” However, they also express “chance as coincidence.” In denying that macroevolutionary events are determined by specific time intervals or membership in a taxonomic group, proponents of the MBL program suggest that evolutionary events such as multiple extinctions are likely to be the result of many “unrelated or unconnected” causes (Slowinski and Guyer 1989, 908), not a single cause as would be proposed under a deterministic explanation. For example, Raup and Marshall ask: “Did the

order Creodonta die out because of the ‘creodontness’ of its genera (meaning small cranial volume and presumably low intelligence), *or* did the order die out just because all its genera went extinct through independent and widely varying causes?” (1980, 10). Thus, the extinction of *each genus* of an order might have a single cause, but the causes would be unrelated (“independent” events), so that there would be no single cause of the extinction of the order *as a whole*. If the causes of origination or extinction are independent, then it is solely by “coincidence,” or chance, that the higher level macroevolutionary event occurs.

The random drift case is somewhat more complicated. Random drift, unlike the stochastic models, encompasses situations where there seems to be one determinate cause. For example, a severe drought might decimate a population of artichokes, killing artichokes without regard to any physical differences between them. This would be random drift because it is a case where the physical differences between organisms are causally irrelevant to differences in reproductive success. Because there seems to be one determinate cause, the drought, as opposed to a stochastic macroevolutionary process where there are many causes, one might be inclined to suppose that random drift does not involve “chance as coincidence.” However, that is not the case. If we examine why some artichokes live and some die, we will probably find (for example) that some happen to be in areas where the soil was richer or retains more moisture. That these artichokes survive where others do not is a coincidence – a confluence of independent causal chains. Therefore, even when many causes are not apparent, random drift invokes “chance as coincidence,” providing for yet another similarity between it and the stochastic macroevolutionary processes of the MBL program.

4. The Stochastic Model: An Alternative to Traditional Deterministic Explanations. In the previous section, we explored the sense in which the *processes* modeled by the MBL program can be said to be chance processes. In this section, we will focus on the *outcomes* of these processes and the nature of the explanations they provide. Deterministic explanations have traditionally been provided for seemingly non-random phenomena such as patterns of taxonomic diversity (for example, extreme increases or decreases in the number of taxa), the simultaneous increase or decrease of taxonomic diversity in a number of unrelated taxa, and morphological trends (directional evolutionary changes in physical traits that take place over a lengthy time period). Proponents of stochastic accounts claim that their models can explain phenomena that were previously thought to require deterministic explanations (such as the radiation in the anoline lizard genera discussed above).

4.1. Theoretical Pluralism in Paleobiology. However, it should be noted that while stochastic models provide an alternative to deterministic explanations, the two kinds of explanation are not mutually exclusive explanatory paradigms. That is, it may be the case that some macroevolutionary phenomena are best explained stochastically, whereas other macroevolutionary phenomena are best explained deterministically. This is similar to the situation with natural selection and random drift; although natural selection and random drift can sometimes explain the same phenomena, it is not the case that in accepting a random drift explanation one must reject all natural selection explanations (and vice versa). Beatty,

in discussing the debate between neutralists (proponents of random drift) and selectionists (proponents of natural selection) suggests that there are both specific and general issues at stake in the debate:

The more specific issues concern whether or not a change in frequency of a specific set of genetic alternatives can be ascribed to their selective neutrality and the consequent random drift of their frequencies. Complimentarily, one might expect that the more general issues concern whether all or no evolutionary changes – or whether all of a certain kind or none of a certain kind of evolutionary change – are due to random drift alone. It is important to recognize, however, that even the most general issues surrounding this version of the importance of random drift do not boil down to questions of all or none, but to questions of more or less (1984, 199).

That is, the debate between proponents of random drift and proponents of natural selection is sometimes over whether a specific instance – a specific change in a population, for example – is due to natural selection or random drift. Alternatively, such debates might discuss whether the specific evolutionary change was primarily due to random drift or primarily due to natural selection. At other times, the debate between neutralists and selectionists is over whether, in general, most evolutionary changes (or, evolutionary changes of a particular kind) have been due primarily to natural selection, or primarily to random drift.

With regard to debates between proponents of deterministic explanations and proponents of stochastic explanations, the situation is strikingly similar.¹² Sometimes, disputants disagree over the proper explanation for a specific phenomenon. For example, Flessa and Imbrie (1973) and Flessa and Levinton (1977) argue that common diversity patterns across taxa during the Phanerozoic call for a deterministic explanation, whereas Smith (1977) maintains that these patterns can be accounted for by a stochastic model. At other times, the debate concerns the more general issues. As the Woods Hole Group asks:

How different, then, is the real world from the stochastic system? How, in other words, is the real world “taxonbound” and “timebound” – i.e., in need of specific, causal explanations involving uniqueness of time and taxon at various stages of earth history? The answer would seem to be “not very” (Gould et al. 1977, 32).

However, immediately following the claim that the “real world” can be largely explained stochastically, the Woods Hole Group proceeds to outline three instances in which deterministic explanations are required. Similarly, one group of deterministic proponents admits: “There are, of course, stochastic elements in the genesis and development of all clades, including those that we designate as [the consequence of some condition or event]” (Stanley et al. 1981, 117). However, these same proponents are careful to circumscribe the extent of those stochastic elements: “Only within small taxa, such as genera comprising a small number of species, do chance factors have more than a small probability of prevailing”

(Stanley et al. 1981, 125). Thus, proponents on both sides assert that their form of explanation should predominate, while carefully circumscribing a small area where the other form of explanation is appropriate. Proponents of stochastic models do not assume that all macroevolutionary phenomena should be explained stochastically; the same is true of proponents of deterministic explanations. Beatty argues that “disputants [in evolutionary biology] defend the importance of their favorite modes of evolution without ruling the others entirely out of the question” (1984, 207); the disputants in this debate do just that. This suggests that with regard to macroevolutionary explanations, paleontologists are “theoretical pluralists” in the sense that Beatty elaborates: “A proponent of theoretical pluralism with respect to a particular domain believes that the domain is essentially heterogeneous, in the sense that a plurality of theories or mechanisms is required to account for it, *different items in the domain requiring explanations in terms of different theories or mechanisms*” (1995, 65; italics in original).¹³ They are also engaging in the kind of “relative significance” debate that Beatty (1995) describes.

4.2. “*Non-randomness*” *Out of Randomness*. Thus, it is often the case that proponents of stochastic models seek not merely to “fill in” where deterministic explanations seem inappropriate, but rather to replace specific deterministic explanations (without eliminating deterministic explanations entirely). Yet it seems puzzling that stochastic models would be able to account for the same sorts of phenomena that deterministic explanations can account for; intuitively, one would expect very different sorts of results from a stochastic model than

from a deterministic explanation. For example, you might expect that a stochastic process would show fluctuations in taxonomic diversity over time, but you would not necessarily expect it to show a marked increase in the number of taxa. However, the particular kind of system embodied by the stochastic models – the Markov process – can often result in seemingly non-random results that we would not intuitively expect from a stochastic process. In order to understand the nature of a Markov process, we will first examine the simpler case of random drift.

A Markov process is a process where the conditional probability of a future event given all the previous events equals the conditional probability of that future event given only the present event. Thus, in determining the probability of a future event, only the present outcome is relevant; the particular history that led to the present outcome is irrelevant. Random drift is a particular kind of Markov process, called a random walk. The easiest way to understand a random walk is through an ideal coin-tossing game. Let heads represent an increase of one unit, tails a decrease of one unit, and the probability of heads or tails be equal. (This probability is known as the *transition probability*, the probability of changing from the current state to a subsequent state). We can then describe a sequence of “winnings” s_1, s_2, \dots, s_n , (starting at $s_0=0$) where each s_i is the total amount “won” (positive or negative) after the i th trial (Feller 1968). This is a Markov process because it does not matter what sequence of events has led to the present outcome; for example, the sequence *heads, heads, heads, tails* is equivalent to the sequence *heads, tails, heads, heads*; in both cases, our

winnings have the value 2, and in both cases, our possible values of our winnings after the next toss are either 1 (if we get tails) or 3 (if we get heads).

Note that in our ideal coin tossing game, for each trial i , s_i depends on the previous outcome s_{i-1} and the transition probability, which is 0.5 for an increase of 1 (heads) and 0.5 for a decrease of 1 (tails). So, for example, if our winnings are currently 7 (which would imply that we had turned up heads 7 more times than we had turned up tails), there are two possibilities for our next toss of the coin: 6 (if the coin turns up tails) and 8 (if the coin turns up heads). However, since heads and tails are equally likely, it is just as likely that we will have winnings of 8 as it is that we will have winnings of 6; there is no tendency to “equalize out” the number of heads and tails. The outcome of the previous toss thus constrains the outcomes of future tosses. Each *toss* is independent, but the *outcome* of each toss, since it is added to the outcome of the previous toss, is not independent. Thus it can happen that outcomes can “walk” in one direction or another; the values of our coin tosses will probably fluctuate somewhat, but the value may continue to increase overall (although presumably not indefinitely).

In fact, our winnings may not return to zero as quickly as one might (intuitively) think. Our intuition suggests that, in a long series of tosses, our winnings would fluctuate from negative to positive values frequently, so that in any series of tosses, it is likely that our winnings recently had a value of zero. However, as Feller has demonstrated, in any sequence of $2n$ trials, where the $2k$ th trial is the last trial where the number of heads and tails is equal, it is just as likely that $k = i$ as $k = n - i$.¹⁴ So, for example, in sequence of 100 coin tosses, it

is just as likely that the last point our winnings equalized was at the 98th toss as it is that the last point our winnings equalized was at the 2nd toss. This somewhat non-intuitive result implies that “it is quite likely that in a long coin-tossing game one of the players remains practically the whole time on the winning side, the other on the losing side” (Feller 1968, 81). For example, for a large number of tosses, the probability that our winnings are positive for 96.7 percent of the time is 0.2; the probability that our winnings are positive for 99.4 percent of the time is 0.1 (Feller 1968, 82).¹⁵

Random drift is analogous to the ideal coin tossing game in the following way. Recall that random drift is a process in which physical differences are causally irrelevant to differences in reproductive success. So, if random drift occurs in the absence of selection, in any given generation it is equally likely that one genotype will be as successful as another genotype (just as it is equally likely that we would get heads or tails). Furthermore, if we consider each generation as a trial and the number of alleles as the outcome of that trial, the number of alleles in each generation depends only on the number of alleles of the previous generation and the various transition probabilities T_{ij} , (where i is the present number of A alleles, j is the number of A alleles in the next generation, and T_{ij} is the probability of going from state i to state j). The transition probability is defined by the binomial distribution:

$$T_{ij} = [(2N)! / (2N - j)! j!] (i/2N)^j (1 - i/2N)^{2N - j} \quad (\text{Roughgarden 1996, 65-6})$$

To get a general idea of the properties of this distribution, let us look at a transition probability matrix for a population of size 3 ($2N = 6$ genes) containing the allele A (Table 1). Note that for each state i at t , the most probable state at $t + 1$ is i . (That is, the most likely result is that the number of alleles will not change from one generation to the next). However, other transitions occur with a relatively high probability. In particular, note that when allele A is predominant in the population, it is still fairly likely that it will increase in frequency in the next generation. For example, in state 4 at time t , the probability of changing to state 5 at time $t + 1$ is 0.263, whereas the probability of changing to state 3 at time $t + 1$ is 0.219. Similarly, when allele A does *not* predominate in the population, it is still relatively likely that it will decrease in the generation. For example, in state 2 at time t , the probability of going to state 1 at time $t + 1$ is 0.263, whereas the probability of going to state 3 at time $t + 1$ is only 0.219. (States 0 and 6 are the cases where no alleles are A 's or all alleles are A 's, respectively, and thus no change is possible unless there is a mutation).

Thus, as with the ideal coin tossing game, there is no tendency to return to the equalization point, which for random drift is the point at which all alleles are represented equally (in a population where there are 2 alleles at a locus, this is the point at which there are N copies of each allele, or half of the total of $2N$ alleles). When an allele predominates in the population, it may continue to increase in the population (or it may decrease). In addition, since any future generation will be formed from the previous generation, the frequencies of the present generation will tend to constrain the values of the future generation. Since frequency values, unlike coin toss values, can change by variable amounts

at each time t , future values are not as constrained by the most recent values in the random drift case as in the coin tossing case. However, the smaller the changes in frequency from generation to generation, the more future values will be constrained by the value of the previous generation. Larger frequency changes in a population mean that frequency values may return towards the equalization point in just one generation, but they also mean that the frequency may diverge even farther from the equalization point. Thus, as with the ideal coin tossing game, it is possible for values to “walk” in one direction or the other. Furthermore, since random drift is a random walk process like the coin tossing game, the chances of one allele predominating for a long period of time are reasonably high.

In this way, random drift, a Markov process, is able to explain apparent directionality. Again, the situation for stochastic models of macroevolution is very similar to the microevolutionary case. Recall that a Markov process is a process in which only the present state is relevant and past states are irrelevant. The stochastic models we have been discussing incorporate a kind of Markov process called a branching process (Feller 1968). The models can be seen as embodying a branching process because at each point in time, each “particle” (in this case, a taxon) has a certain probability of terminating (going extinct), branching into two “particles” (producing two new taxa), or continuing into the next time interval (persistence). Because each future outcome (two new taxa, extinction of taxon, or persistence of taxon) is dependent on the present taxa without being dependent on any of the outcomes prior to the present taxa, the branching process of the stochastic model is Markovian. Similar to the random walk process discussed above, branching processes also

have the property that future outcomes are partly dependent on the present taxa and partly dependent on the transition probability to the future outcome. In the “freely floating” model described above, the transition probability from the present state to a branching event is 0.1, the transition probability from the present state to an extinction state is also 0.1, and the transition probability from the present state to a “persistence state” (i.e., the taxon simply persists in time without branching or terminating) is 0.8. In the damped-equilibrium model, transition probabilities change over the course of a simulation. In the beginning of the simulation, the probability of branching is greater than the probability of termination, but once the optimum diversity is reached, the probabilities of branching and termination remain relatively constant and equal.

Despite the fact that the basic model used in the MBL program is a branching process, the program also uses a random walk model by superimposing a random walk on top of the branching process. This is done by counting the number of coexisting lineages at each time interval, where lineages are produced by a branching process. Gould et al. (1977) use this method in the “freely floating” model described above. As with the random drift case, the random walk produced by this model differs from the random walk of an ideal coin tossing game in that the number of coexisting lineages is not limited to incrementing or decrementing by one unit. For example, if there are four lineages at time t , and at time $t + 1$ three of them branch while one goes extinct, the number of coexisting lineages increases from four to six.

Above, we saw how, in a random walk, the dependency of a future state on the present state has a tendency to constrain the future state; values “walk” from the present state to the future state. Furthermore, there is no general tendency to return to an equalization point; if the values in the system have been increasing up to the present state, they are no more likely to decrease at the next state than they are to increase. This is the situation with the “freely floating” model. For example, if the number of taxa at time $t - 2$ is 10, the number of taxa at $t - 1$ is 12, and the number of taxa at t is 15, the number of taxa at $t + 1$ will be an increment or a decrement from the value of 15. The fact that the two previous time intervals represented an increase in diversity value does not make it any more likely that there will be a decrease at $t + 1$. The damped-equilibrium model behaves similarly, except that there *is* some tendency to return to the optimum value that the program has set. Note, however, that since we are dealing with probabilities, even when the diversity exceeds the optimum value, it may continue to increase in the next time period. It is just less *likely* to do so (and the farther away from the optimum value a simulation goes, the less likely such a departure becomes, by the design of the program). The characteristics of these two models imply that the values of these simulations may increase (or decrease) for long periods of time without returning to the equalization point. (Of course, this is less true for the damped-equilibrium model than it is for the “freely floating” model).

Consequently, when the simulations are run, the patterns of diversity exhibit considerable variation. Some show what you would intuitively expect from a stochastic model: fluctuations in diversity, with no marked increases or decreases. Others, however, exhibit

patterns where there are marked increases in diversity, or marked decreases in diversity. In Figure 1, diagrams 2 and 6 are what one might intuitively expect from a stochastic process: small amounts of changes in diversity over time. Notice, however, the sudden increase in diversity in diagram 9, similar to that of diagram 12, and the long periods of stability in both diagram 7 and diagram 19.

We are now in a position to answer the question raised in the first paragraph of this section, namely, how is it possible that a stochastic model can account for the same phenomena that were previously thought to require deterministic explanations? The Woods Hole Group charges that deterministic explanations are sometimes proposed merely because the phenomena being explained appear to be nonrandom (Gould et al. 1977). However, stochastic Markov processes can also (with varying degrees of probability, depending on the situation) produce these same sorts of results. Thus, if seemingly non-random patterns are the *only* evidence, stochastic models can often provide explanations of the same phenomena. As D. C. Fisher notes, “An important lesson of Markovian models is that some degree of apparent order can arise by chance ... As such, directional trends do not necessarily require explanations framed in terms of single causes acting throughout the duration of the trend” (1986, 106). It follows that there is an underdetermination of patterns; you cannot determine whether a stochastic or a deterministic explanation is appropriate just from looking at the pattern. Because the MBL program models a Markov process, we can expect it to produce a great variety of patterns; within the history of a lineage, it may undergo periods of increasing or decreasing diversity. With Markov processes, the production of directionality is not

unusual, but rather quite common. Markov processes are processes that may be constrained by the immediate past; this leads them to produce orderly results regularly.

Thus, the empirical question remains as to how, in any given instance, one is to decide whether a deterministic or stochastic model provides the better explanation. Stochastic models have been proposed as explanations for a variety of macroevolutionary phenomena. There have been studies of phylogenetic patterns of diversity (Raup et al. 1973; Gould et al. 1977; Smith 1977); morphology (Raup and Gould 1974); rates of evolution (Schopf et al. 1975; Schopf 1979; Raup and Marshall 1980; Bookstein 1987); and faunal extinctions (Schopf 1974, Simberloff 1974). Each of these studies discusses phenomena to which stochastic models can potentially be applied. My purpose here is simply to show how stochastic models can provide an alternative to deterministic explanations.

5. Conclusion. There are many parallels between macroevolutionary stochastic models (and the “opposing” deterministic style of explanation) and random drift (and the “opposing” process of natural selection). The scare quotes here indicate that in both the macroevolutionary and the microevolutionary cases, the two forms of explanation are not really opposing; they may 1) simultaneously explain different aspects of the same phenomena, 2) be applied to different phenomena by different people, or even 3) be applied to different phenomena by the same people. Thus, there is theoretical pluralism. Nonetheless, there is a sense in which the stochastic models have (historically) provided a

challenge to more traditional, Darwinian, selectionist explanations, and this aspect cannot be overlooked.

In addition, the stochastic processes at the microevolutionary level and the macroevolutionary level can both be represented as a random walk, implying that seemingly nonrandom patterns can be produced, even though the processes themselves are random. Given this similarity, it does not seem very farfetched to suggest that the parallels between the microevolutionary process of random drift and the stochastic macroevolutionary models are deliberate, although the Woods Hole Group pointed more towards ecology than population genetics as their inspiration.

The MBL program models stochastic processes that incorporate conceptions of chance distinct from both “chance as ignorance” and “chance as uncaused”: “chance as indiscriminate sampling,” “chance as coincidence,” and “evolutionary chance.” These conceptions of chance play a role at the microevolutionary level as well, in the form of random drift. Thus, when evolutionary biologists invoke “chance,” it is more than an admission of ignorance (they may or may not be ignorant of the relevant causes in a particular case); it is an expression of a particular kind of process. Moreover, this expression is independent of quantum mechanical conceptions of chance. The examination of chance in evolution thus provides reason to rethink some of our traditional ideas about chance. It also raises the intriguing possibility that the evolution of life is stochastic on a grand scale – that major events in the biological history of our earth are due to chance

REFERENCES

- Beatty, John (1984), "Chance and Natural Selection", *Philosophy of Science* 51: 183-211.
- (1995), "The Evolutionary Contingency Thesis", in Gereon Wolters and James Lennox (eds.), *Concepts, Theories, and Rationality in the Biological Sciences*. Pittsburgh: University of Pittsburgh Press, 45-81.
- Bookstein, Fred L. (1987), "Random Walk and the Existence of Evolutionary Rates", *Paleobiology* 13: 446-464.
- Brandon, Robert and Scott Carson (1996), "The Indeterministic Character of Evolutionary Theory: No 'No Hidden Variables' Proof But No Room for Determinism Either", *Philosophy of Science* 63: 315-337.
- Darwin, Charles ([1859] 1964), *On the Origin of Species: A Facsimile of the First Edition*. Cambridge, Massachusetts: Harvard University Press.
- Eble, Gunther J. (1999), "On the Dual Nature of Chance in Evolutionary Biology and Paleobiology", *Paleobiology* 25: 75-87.
- Feller, William (1968), *An Introduction to Probability Theory and Its Applications*. New York: John Wiley & Sons.
- Fisher, Daniel C. (1986), "Progress in Organismal Design", in David M. Raup and David Jablonski (eds.), *Patterns and Processes in the History of Life*. Berlin: Springer-Verlag, 99-117.

- Flessa, Karl W. and John Imbrie (1973), "Evolutionary Pulsations: Evidence from Phanerozoic Diversity Patterns", in Donald H. Tarling and S. K. Runcorn (eds.), *Implications of Continental Drift to the Earth Sciences*. London: Academic Press, 247-285.
- Flessa, Karl W. and Jeffrey S. Levinton (1975), "Phanerozoic Diversity Patterns: Tests for Randomness", *Journal of Geology* 83: 239-248.
- Futuyma, Douglas J. (1986), *Evolutionary Biology*. Sunderland: Sinauer Associates.
- Gould, Stephen Jay, David M. Raup, J. John Sepkoski, Thomas J. M. Schopf and Daniel S. Simberloff (1977), "The Shape of Evolution: A Comparison of Real and Random Clades", *Paleobiology* 3: 23-40.
- Grantham, Todd A. (1999), "Explanatory Pluralism in Paleobiology", *Philosophy of Science (PSA 1998 Proceedings)* 66 (Supplement): S223-S236.
- Peterson, Jane A. (1983), "The Evolution of the Subdigital Pad in *Anolis*. I. Comparisons Among the Anoline Genera", in Anders G. J. Rhodin and Kenneth Miyata (eds.), *Advances in Herpetology and Evolutionary Biology: Essays in Honor of Ernest E. Williams*. Cambridge, MA: Museum of Comparative Zoology, Harvard University, 245-283.
- Raup, David M. (1977), "Stochastic Models in Evolutionary Paleontology", in Anthony Hallam (ed.), *Patterns of Evolution As Illustrated by the Fossil Record*. Amsterdam: Elsevier, 59-78.
- (1985), "Mathematical Models of Cladeogenesis", *Paleobiology* 11: 42-52.

- Raup, David M. and Stephen Jay Gould (1974), "Stochastic Simulation and the Evolution of Morphology – Towards a Nomothetic Paleontology", *Systematic Zoology* 23: 305-322.
- Raup, David M., Stephen Jay Gould, Thomas J. M. Schopf and Daniel S. Simberloff (1973), "Stochastic Models of Phylogeny and the Evolution of Diversity", *Journal of Geology* 81: 525-542.
- Raup, David M. and Larry G. Marshall (1980), "Variation Between Groups in Evolutionary Rates: A Statistical Test of Significance", *Paleobiology* 6: 9-23.
- Ridley, Mark (1993), *Evolution*. Boston: Blackwell Scientific Publications.
- Roughgarden, Jonathan (1996), *Theory of Population Genetics and Evolutionary Ecology: An Introduction*. Upper Saddle River, NJ: Prentice Hall.
- Schopf, Thomas J. M. (1974), "Permo-Triassic Extinctions: Relation to Sea-Floor Spreading.", *Journal of Geology* 82: 129-143.
- (1979), "Evolving Paleontological Views on Determinism and Stochastic Approaches", *Paleobiology* 5: 337-352.
- Schopf, Thomas J. M., David M. Raup, Stephen Jay Gould and Daniel S. Simberloff (1975), "Genomic Versus Morphology Rates of Evolution: Influence of Morphologic Complexity", *Paleobiology* 1: 63-70.
- Simberloff, Daniel S. (1974), "Permo-Triassic Extinctions: Effects of Area on Biotic Equilibrium", *Journal of Geology* 82: 267-274.

- Slowinski, Joseph B. and Craig Guyer (1989), "Testing the Stochasticity of Patterns of Organismal Diversity: An Improved Null Model", *The American Naturalist* 134: 907-921.
- Smith, Charles A. F. (1977), "Diversity Associations as Stochastic Variables", *Paleobiology* 3: 41-48.
- Stanley, Steven M., Philip W. Signor III, Scott Lidgard and Alan F. Karr (1981), "Natural Clades Differ From "Random" Clades: Simulations and Analyses", *Paleobiology* 7: 115-127.
- Uhen, Mark D. (1996), "An Evaluation of Clade-Shape Statistics Using Simulations and Extinct Families of Mammals", *Paleobiology* 22: 8-22.

FOOTNOTES

¹ The first paper produced by this group was written in 1973 by David Raup, Stephen Jay Gould, Thomas Schopf, and Daniel Simberloff. Subsequent papers include Raup and Gould (1974); Schopf, Raup, Gould, and Simberloff (1975); Gould, Raup, J. John Sepkoski, Jr., Schopf, and Simberloff (1977); Raup (1977); and Schopf (1979).

² For a notable exception, see Grantham (1999).

³ As I mentioned above, the use of the term “deterministic” here does not imply Laplacean determinism. Throughout this essay, unless I specify Laplacean determinism, my usage of the term “deterministic” should be taken to imply the macroevolutionary, paleontological meaning that I elaborate in this section.

⁴ It might seem strange to speak of a time period as a cause, particularly in the case of the dinosaur example given. In speaking in this way, the Woods Hole group seems to be thinking in terms of abstract models – that is, a model where there are distinct periods of time where the probability of extinction is different from the probability of branching (timebounded) versus a model where there are no such distinct periods (untimebounded).

⁵ It should be noted that I am contrasting a kind of *explanation* that has commonly appeared in the literature (deterministic explanation) with a stochastic *model* that can be used for evolutionary explanations. In other words, there is no specific “deterministic model,” although such models have been developed (see Raup 1985 for discussion).

⁶ The model assumes that speciation only occurs when there is branching, a restriction that would be questioned by those who would argue that speciation can occur through the accumulation of gradual changes over time (anagenesis).

⁷ It might seem odd that one of the stochastic models is an equilibrium model. The reason for the use of an equilibrium model is to see if the successful use of equilibrium models in ecology and population biology can be replicated in paleontology (Raup et al. 1973). Furthermore, such a model may be more biologically realistic: “The maintenance of an equilibrium diversity in the present work implies that an adaptive zone or a geographic area becomes saturated with taxa and remains in a dynamic equilibrium determined by the opposing forces of branching (speciation) and extinction” (Raup et al. 1973, 529).

⁸ In the MBL model, the probabilities of branching and extinction are equal as well as constant. Other stochastic models allow the probabilities of branching and extinction to differ from one another (Raup 1985).

⁹ Brandon and Carson (1996) take a different position concerning the quantum mechanical conception of chance.

¹⁰ This supposition suggests that the example is Laplacean deterministic in that the given state uniquely determines a future state, a position that I will grant for the sake of argument.

¹¹ Note that whereas the “success” of organisms is measured by the number of offspring they produce, the “success” of a taxon is measured by the number of taxa at the lower level.

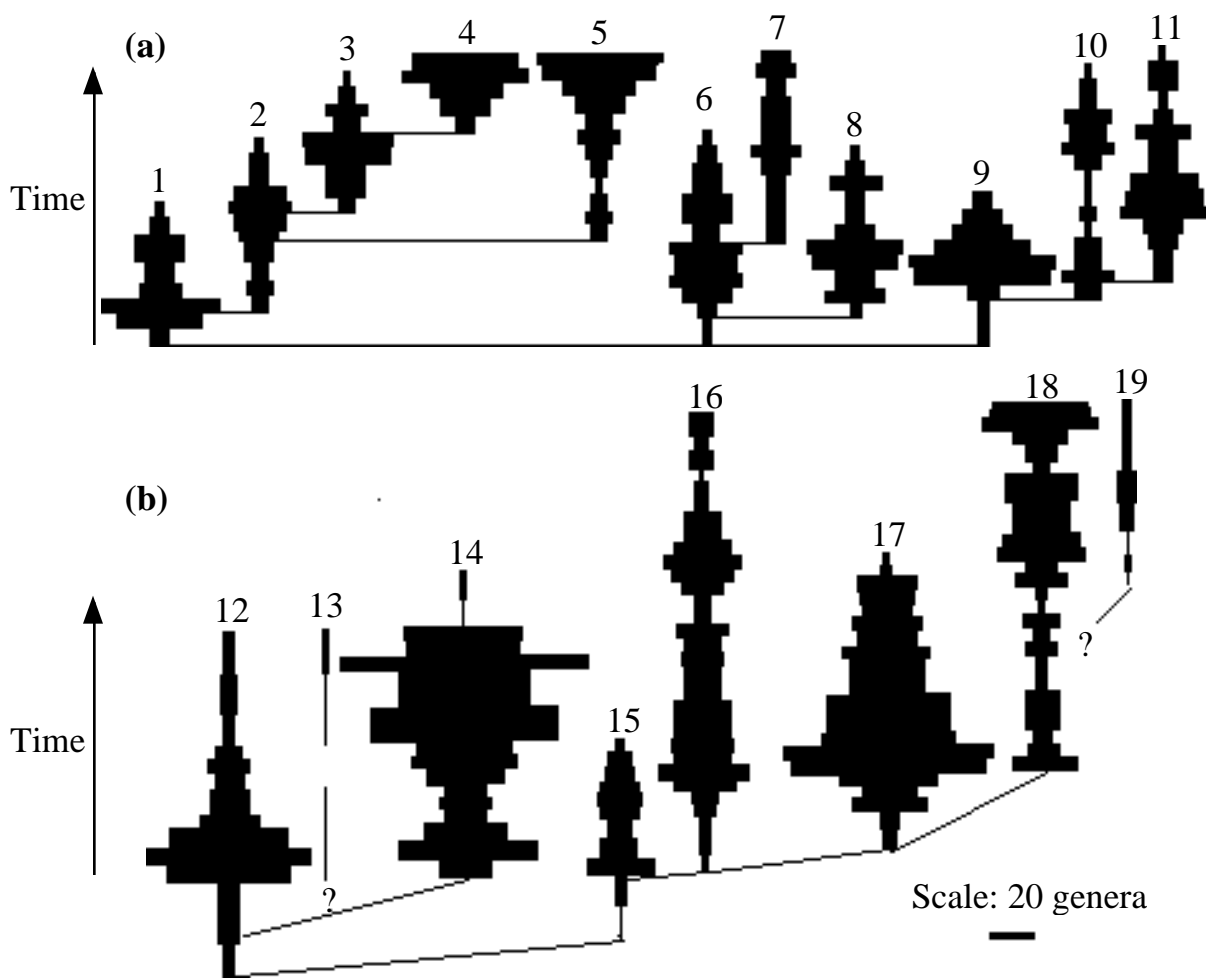
For example, one genus is more successful than another genus if it contains a greater number of species (implying that the species of the genus have had either a greater speciation rate, a lower extinction rate, or both).

¹² The *situation* is similar, but the point of disagreement is not the same. Raup and Gould, for example, explain that they are “quite consciously avoiding the subject of evolutionary mechanisms”; they insist that their model is consistent with selection in randomly fluctuating environments as well as the fixation of mutations by random drift (1974, 309). Thus, because selection *may* underlie the stochastic model, the debate between neutralists and selectionists does not fall along the same lines as the debate between proponents of deterministic explanations and proponents of stochastic models.

¹³ Grantham argues for the stronger thesis of “explanatory pluralism” in paleobiology, in which “both strategies provide correct explanations for the same event” (1999, S225).

¹⁴ We use a value of $2k$ since, of course, our winnings can only be equal when there has been an even number of tosses.

¹⁵ Likewise, the probability that our winnings are *negative* for 96.7 percent of the time is 0.2; the probability that our winnings are negative for 99.4 percent of the time is 0.1.



FIGURE/TABLE LEGENDS

Figure 1: Comparison of stochastically generated diversity patterns with real diversity patterns. Numbers above diagrams are for ease of identification only. The width of a diagram at a given “point” in time represents the number of taxa at that time (more taxa = greater diversity). (a) Diversity diagrams for one run of the MBL program with probabilities of branching and extinction set to 0.1. (b) Diversity diagrams for genera within orders of Brachiopods (From Gould et al. 1977).

Table 1: Transition probability matrix for a population of size $N = 3$ ($2N$ genes). Numbers generated by computer program and rounded to the third decimal place.