[LA: 9,097 words; 6 notes; acknowledgements; references; 1 box]

[Received 14 December, 2006; revised May 11, 2007]

Drift: A Historical and Conceptual Overview

Anya Plutynski

Department of Philosophy

University of Utah

plutynski@philosophy.utah.edu

Abstract

There are several different ways in which chance affects evolutionary change. That all of these processes are called "random genetic drift" is in part a due to common elements across these different processes, but is also a product of historical borrowing of models and language across different levels of organization in the biological hierarchy. A history of the concept of drift will reveal the variety of contexts in which drift has played an explanatory role in biology, and will shed light on some of the philosophical controversy surrounding whether drift is a cause of evolutionary change.

Keywords

causation, manipulation, multi-level theory, neutral theory of molecular evolution, random genetic drift, supervenience

The concept of “random drift,” not unlike the concept of “gene” or “fitness,” has become a “concept in tension” (to paraphrase Falk 2000) that has triggered a substantial debate among philosophers of biology. In the last ten years philosophers have debated whether random genetic drift[1] is properly understood as a process or an outcome (Millstein, 2002, 2005; Brandon, 2005), whether the force metaphor is appropriate to describing either selection or drift (Matthen and Ariew 2002; Shapiro and Sober forthcoming), and finally, whether and in what sense drift is a cause (Matthen and Ariew 2002; Walsh et al. 2002; Riesman and Forber 2005; Shapiro and Sober forthcoming). One reading of this exchange is that it is all a tempest in a teapot; drift is as well-defined a concept as any in mathematical evolutionary biology. However, the extent of the philosophical debate suggests that there are larger issues at stake. Arguably, this debate reflects more deep-seated and long-standing issues in evolutionary biology (and metaphysics). Going back to Mayr’s (1959) critique of “beanbag genetics,” there has been a concern among evolutionary biologists about the scope and limitations of classical population genetics, the relationship between the formal and the empirical in biology, and, more broadly, how to integrate understanding of pattern and process at different levels in the biological hierarchy—from molecular evolution to speciation. Drift has become the latest focus around which philosophers of biology debate these issues. Further, the debate is connected to a larger debate in general philosophy of science about levels and modes of causation, or, what may count as a cause.

One element missing from much of the recent philosophical discussion is a historical dimension. Drift, like “gene” and “fitness”

is a concept that has evolved over the past 75 years; understanding how and why may help resolve some of these debates or, at the very least, identify the sources of confusion.

Discussion of the drift concept among philosophers has (in large part) proceeded on the basis of a few key examples and textbooks' definitions of drift. Yet, when one consults textbooks in evolutionary genetics, a variety of not entirely consistent definitions of drift are available. Drift refers to "chance fluctuations of allele frequency" which occur "particularly in small populations, *as a result of* random sampling among gametes"(Hartl 1988: 16) Or, drift is "a *dispersive force* that removes genetic variation from populations," which might be due to either "variation in the number of offspring between individuals, and, if the species is diploid and sexual, from Mendel's law of segregation"(Gillespie 1998: 19). Finally, drift is "the process of change in gene frequency due solely to chance effects"(Grauer and Li 2000: 48). These three texts describe drift as both the "result" of random sampling, the force removing genetic variation, and a process of change due to chance. In other words, drift is spoken of interchangeably as effect and cause, pattern and process. Different accounts of the chief *effect* of drift are offered, from "fluctuations of allele frequency" to "the removal of genetic variation." Drift is also attributed to different causes – "segregation" as well as "chance variation in the number of offspring between individuals." In other words, drift is due to different kinds of mechanisms at work at different time scales and levels of organization, from meiotic segregation and recombination, to chance elements affecting fertilization, to random environmental factors such

as natural catastrophes that affect whether some individuals but not others survive and reproduce.

As the above texts illustrate, drift has become a sort of "kludge" concept in biology— collectively referring to a variety of different ways in which chance broadly understood affects evolving populations over time. Likewise, over the history of the use of this concept, the term "drift" has been used to refer to a variety of outcomes and a variety of causes. As Beatty's has argued, "Random drift is a heterogeneous category of causes and effects… the collective phenomena are very different. Moreover, there are phenomena sometimes included in the category of random drift that have nothing to do with random sampling"(Beatty 1992: 273) With such a diversity of referents and applications of the concept, perhaps it is not surprising that philosophers have so heatedly contested the nature of drift.

In addition to the fact that drift has historically been used to identify a variety of causes and effects, there is a rather large gap between the classical "Wright-Fisher" model of drift and the variety of ways in which chance events effect populations over time. The binomial random sampling model (see box insert at end of text) treats drift as the random sampling of alleles (with replacement) from one generation to the next; where the outcome is a form of "sampling error." As a result of this sampling process, some alleles are fixed and others are lost. In smaller populations, this process is accelerated; or, the time to fixation of an allele will be shorter. The mathematical model of drift thus permits predictions of outcomes

in systematic ways. Taking smaller sample sizes increases the time to fixation or loss of alleles by chance alone.

This basic model was developed in the 1920s and 30s, when very little was known about the mechanisms of inheritance or causes of variability at the molecular level. The drawing of alleles as akin to the drawing of balls from an urn was treated as an appropriate analogy for the process of “sampling” of alleles in a population via genetic recombination, despite disanalogies between this process and the actual processes involved in random assortment and “sampling” at different levels of the biological hierarchy. Nonetheless, this model, and the label of “random genetic drift,” became a placeholder for what, at different levels of organization, are different “engines” and outcomes of chance. Though there are similarities in broad terms across these different contexts (extent of drift is contingent on population sizes), the specific biological factors at work in the “drifting” of nucleotides, cistrons, alleles, whole gene complexes, chromosomes, or for that matter, individuals and groups, are very different. These different contexts require subtle corrections to the original models. For instance, intracistronic recombination rates (rates of recombination between sites within the gene) are far smaller than those between genes, suggesting that the classical model requires subtle modification in this case (Ewens, 2004).

There is no doubt that the classical binomial sampling model has been enormously successful as a tool for understanding the dynamics of evolution at the level of populations. Nonetheless, biology has changed so much since 1930, that historians and philosophers of biology have raised legitimate concerns as to whether the original

models of the dynamics of evolution are to some extent outmoded (Provine 2001), and, whether the ways of speaking of drift as a "dynamic force" are problematic (Matthen and Ariew 2002). In what follows, I will review the history of the tradition of modeling evolution that lead to philosophical controversies about drift as cause and effect, and use this historical analysis as a means of, if not resolving the philosophical debate to everyone's satisfaction, at the least, clarifying issues at stake.

Brief Historical Overview: Origins of Drift <A>

Predecessors to the drift concept are found in Darwin's discussion of variations "neither useful nor injurious" as polymorphic traits "not effected by natural selection" and therefore either left as "fluctuating elements" or "fixed" (Darwin 1859: 46, 81). Darwin did not discuss how or why these variations would become lost or fixed. Fleeming Jenkin, however, understood well the loss of variation due to isolation of small (particularly island) populations. He treated the loss of variation via chance as a challenge to Darwin's theory of natural selection in the review of the *Origin*. That is, however fit a particular individual mutant, or individual, if isolated on a small island population, the advantageous variants will be "swamped" by less fit types (Jenkin illustrated this with a "burrowing hare" and another, (offensively racist) example.) Moreover, Gulick's discussion of the role of "indiscriminate destruction" of some members of a species in differentiation of Hawaiian land snails (Gulick 1889: 209)

was enormously influential for subsequent work on drift, as well as models of speciation. Gulick stressed natural catastrophes as sources of indiscriminate elimination of members of a group. In other words, his focus was on “sampling” at the level of whole subpopulations. However, the earliest formal treatments of random genetic drift focused largely on chance events in sexual reproduction (meiosis and “sampling of alleles”) (Hagedoorn and Hagedoorn 1921), which led to the models of Wright and Fisher.

Fisher (1922) models the effects of drift in the context of discussion of what he calls the Hagedoorn (1921) thesis: namely, “that random survival is a more important factor in limiting the variability of species than preferential survival.” Fisher viewed the Hagedoorn’s argument as a threat to Darwinian evolution by natural selection, and argued that the Hagedoorn effect will not be substantial: “the decay in the variance of a species breeding at random without selection and without mutation, is almost inconceivably slow” (Fisher 1922: 323).

Fisher acknowledges that randomness is crucial in the early stages of a mutation’s appearance in a population, or “While it is rare, its survival will be at the mercy of chance, even if it is well fitted to survive” (Fisher 1922: 326). Likewise, in small populations, the probability of survival of a single mutant will be small. However, Fisher showed how the eventual survival of a favorable new mutation is almost certain if a mutation is recurrent (Fisher 1922: 340). Or, Fisher demonstrates that the chance elimination of alleles is counterbalanced by recurrent mutation, given a high enough rate. This argument was central to Fisher’s demonstration that mutation and selection could, over time, yield substantial change in populations.

The 1922 paper may be viewed as Fisher's attempt to vindicate selection against the threat of random loss of variation. In other words, Fisher views himself as defending Darwin against the very same criticisms that Jenkin raised in his review of the *Origin*. The difference is that Fisher was far more familiar with statistics, and was able to draw an analogy with another discipline that treated statistical changes in large populations of entities – statistical mechanics. Using this analogy, Fisher pictures the variation in a population as a normal distribution; mutation constantly supplies new variation, and chance eliminates new variants from a population. So, if the variation in some trait is pictured as a distribution about the mean, mutation will continually "expand" the bell-shaped curve, whereas chance will "shrink" the curve. In this way, and drawing upon the analogy he drew between evolution and statistical mechanics, Fisher was led to think of selection, mutation, migration, and drift as "forces." Selection coefficients describe the relative advantage of some gene, or if a particular gene's chance of survival is greater or less than one; selection may thus shift or skew the curve in one or another direction. The gene will thus either increase in frequency in a population over time, or decrease, or become eliminated by chance. Selection measures its "velocity" of increase, relative to other factors, such as population size, and rate of mutation. Fisher's object in the 1922 and subsequent papers was to consider the relative role of each of these factors in the dynamics of evolution in populations.

Fisher refers the dynamics of changes of gene frequency as akin to a physical-mechanical process, determined by forces, such as

mutation and selection, and chance elimination: “the effects of selection in modifying gene frequencies are… exhibited… by changes in position with velocities that are uniform and proportional only to the intensity of selection”(Fisher 1930: 79). In contrast, Wright (at least in the early stages of his theoretical work) preferred to speak of the “lability” and “plasticity” of a population – a population was rather like a clay that would be more or less “plastic,” given the extent of inbreeding, outbreeding, etc. A population was thus capable of being molded or “shaped” by breeders, or nature, as the case may be. Fisher’s way of modeling the dynamics of populations, while he used different metaphors of “velocity,” arguably influenced Wright’s thinking concerning the optimal “balance” of factors at work in any population.

There was a very important difference between the two, however. Whereas Fisher viewed the effect of random loss of variation as a threat to selective explanations, and thus (in his view) to “Darwinism” itself, Wright came to view chance constructively or as a causal factor in evolution. This view of drift, as one factor increasing the “lability” of populations, eventually became one of several factors assisting in the eventual “movement” of populations up neighboring adaptive peaks, on Wright’s infamous adaptive landscape. However, Wright did not at first refer to drift per se as a distinct cause of evolution. In most of Wright’s early papers on the role of chance in evolution, “drift” was spoken of as an *effect not a cause*.

In the 1920’s and early 30’s, Wright primarily spoke of drift as an “effect” of chance extinction, subdivision of populations, isolation, and inbreeding; genes would “drift” about equilibrium, and

"drifting" of frequencies of genes was due to other causes; but *drift per se was not spoken of as a cause*. Rather, Wright spoke of inbreeding or "isolation" as the "cause" of drift.

What is inbreeding? Wright developed a formula for calculating the coefficient of inbreeding, or F. F represents the probability that two homologous alleles are identical by descent, or derived from a common ancestor. Inbreeding occurs when mates are more closely related than if chosen at random; in smaller populations, the chance that randomly chosen mates may be relatives is higher than in larger populations. Thus, in smaller populations, inbreeding is higher, and so too, in smaller populations, speed of fixation of genes due to nondeterministic factors is higher. Wright's early reflections on inbreeding led him to consider that population size might play a significant role in evolution.

In sum, Wright, at least in early discussions of the role of chance in evolution, spoke of drift as not a cause, but a consequence of inbreeding. For instance, one of Wright's early disputes with Fisher over the evolution of dominance concerned whether "isolation effect" or, the effects of inbreeding in a population of small size, might compromise the role of selection in the evolution of dominance. At this stage in his debates with Fisher, Wright uses drift as a *verb* describing the behavior of alleles. Rather than view *drift per se as a cause of evolution*, Wright speaks of "isolation effect" (measured by 1/2n) as one of three "factors which controls the fate of the gene." Genetic factors "drift to fixation" as a *consequence* of isolation. Drift, then, is referred to as a consequence, not a cause of evolutionary change. Wright's emphasis on isolation and his dispute

with Fisher was the first round in a lasting debate over the relative significance of drift and selection in evolution. But, at this stage, in the early 1930s, it was not "drift" versus "selection" that was at issue; rather, it was the various causes of chance elimination: isolation, inbreeding, or patterns of mating that were the "causes" of "fluctuating" allele frequencies.

In his classic 1931 paper, when enumerating the factors contributing to genetic homogeneity, Wright *does not list drift*; instead, he speaks of drift as a *consequence* of restricting population size. Alleles may "drift" to fixation, but at this stage, the "cause" of the process Wright had in mind was inbreeding. In sum, at this stage, "drift," referred the random elimination of genes in a population, due to inbreeding effects. The major cause of fixation not due to fitness differences at this time was, in Wright's view, not "drift" per se, but isolation and inbreeding: "the factor of isolation is of utmost importance in evolution" (1929: 279). Isolation and inbreeding due to patterns of mating within small relatively isolated groups was described as *the cause of loss of genetic heterogeneity*.

Drift Extrapolated <A>

What led biologists to speak of drift as an independent causal factor in evolution, and not a consequence of isolation or "inbreeding effect"? There were three main factors. First, the force of the metaphor of selection and drift as "forces" moving populations around the adaptive landscape became an effective way to popularize the

otherwise rather complex and abstruse science of classical population genetics. Second, during the synthesis, “random genetic drift” became identified with the “Sewall Wright effect,” and appeal to this effect for explanations of non-adaptive differences within and between populations became enormously common (Provine 1986: 405). Further, with the polarizing of Wright and Fisher over the relative significance of isolation and chance versus selection, drift became the name for the competing “force” of chance in evolution. In sum, with the popularization of Wright’s “adaptive landscape” model for representing evolutionary change over time, and its invocation in explanations of a variety of otherwise puzzling phenomena, drift became identified as one of several causal mechanisms of evolution. Non-adaptive differences within and between populations and species — from blood groups to patterns on snail shells to patterns of speciation were all explained by drift (see Beatty 1987, 1992).

For instance, Dobzhansky appealed to the “Sewall Wright effect” to explain polymorphisms in man such as blood groups:

polymorphisms in man (e.g. blood group heterozygosis or homozygosis) may… as far as one is able to judge at present, be explained by random fluctuations in gene frequencies in effectively small populations. Such random variations of gene frequencies are referred to as the genetic drift, or Sewall Wright effect. (Dobzhansky 1957: 156) <BLOCK QUOTE>

Dobzhansky elaborated, appealing to Wright’s classical “U” and bell shaped graphs to represent the relative significance of selection and population size in the retention or loss of alleles:

The smaller the effective population size, the greater are random variations in gene frequencies, and the less effective become weak selection pressures. In small populations, alleles favored by selection may be lost and the les favored ones may reach fixation. In very large populations, even very small selective advantages and disadvantages will eventually be effective; but a more rigorous selection must be applied to overcome the genetic drift in small populations. (1957: 161) <BLOCK QUOTE>

Here Dobzhansky contrasts selection and drift, not explicitly as distinct “forces” so much as competing explanations for the retention or loss of an allele. In smaller populations, alleles are more likely to be either fixed or lost due to random factors; in significantly larger populations, the effects of even small selection coefficients will eventually lead to fixation of alleles. The fact that population size “constrains” the power of selection, lends itself to the idea that selection and drift are “competing forces” of evolution that may be decomposed. In other words, the relative significance of each is contingent upon the other.

Dobzhansky goes on to note an important “biological highly significant corollary” of the above, namely, that “a species, broken up into isolated colonies, may differentiate as a result of the restriction of population sizes” (p. 162). After discussion of a few examples, e.g., Hawaiian land snails and non-adaptive differentiation in Drosophila, he writes: “restriction of the genetically effective size of natural populations is in all probability an important agent of differentiation of species into local groups possessing different genotypes” (p. 176).

In other words, Dobzhansky and other authors of the synthesis appropriated the “Sewall Wright effect” to explain everything from nonadaptive differences within species to speciation itself (Provine 1986; Beatty 1987). Huxley also appealed to the “Sewall Wright effect” as an explanation of not only non-adaptive differences between populations, but also similar differences between species, and even whole taxonomic groups. In Huxley’s words, the “Sewall Wright effect” “at one stroke explains many facts which puzzled earlier selectionists,” for instance, the “greater degree of divergence shown by islands,” and other “recent taxonomic discoveries”(Huxley 1942: 199-200, 260).

Despite the fact that Wright originally referred to drift as a consequence, and not a cause of evolutionary change, “random genetic drift” came to replace the “Sewall Wright effect” as the sum of a multi-stage and multi-factor roles of isolation and inbreeding and their effects on the distribution of alleles. Moreover, with the popularization of his adaptive landscape model, the role of drift began to be used to explain the fate of populations and whole species.

For instance, Ernst Mayr’s (1942) “founder effect”—speciation via the isolation of small subpopulations, followed by “drift” to new adaptive peaks—is arguably an extension of the “Sewall Wright” effect up the hierarchy to species differences. Though Mayr denies any debt to Wright on this count, the process uses the same mechanism of change as Wright championed in his classic (1931) paper (Mayr 1963). This constituted a shift of speaking of drift within populations to drift as a cause of change up the hierarchy to species.

Wright (1949, 1955) confused the issue further by referring to randomly fluctuating selection coefficients as due to drift (discussed in Beatty 1992). Wright here expanded the concept of drift even further; explaining any apparently "random" behavior, even if driven by selection.

With the "hardening" of the synthesis in the 1940s (Gould 1983), however, key examples of drift, such as blood groups, were found to have selective explanations. Nonetheless, the multistage and distributed processes of sampling in populations was christened as a single "force," "drift" and the dialectic of the relative significance of drift versus selection was very much in play and continues into the 21st Century. This was solidified by the polarization of Wright's and Fisher's views, in fallouts over the relative roles of chance and selection in explanation of some classic examples of apparent non-adaptive differences between species.[1]

Stepping Back <A>

The concept of random genetic drift has been appealed to in a variety of biological contexts to explain a number of observations over the

[1] Not only did drift get appropriated "up" the hierarchy as playing a significant role in speciation and extinction, but also, it was applied "down" to the molecular level. In 1968, Motoo Kimura and King and Jukes independently proposed the neutral theory of molecular evolution. This is the view that the majority of evolutionary changes at the molecular level are caused by random drift of selectively neutral or nearly neutral alleles. According to Kimura, changes in amino acids are fixed by "drift." The time to fixation of neutral alleles is determined by population size. In this sense, "drift" at the molecular level is analogous to founder and inbreeding effects; the relative impact of drift in both is dependent upon population size, even if the "mechanism" of sampling is likely quite different.

past 75 years. Drift is invoked to explain “fluctuating” polymorphisms, or apparently selectively neutral variation both within and between populations or species, loss of variation in populations, “randomly fluctuating” selection coefficients, and, finally apparently neutral variation at the molecular level. What all these appeals have in common is the notion that random changes in frequencies of genes (or, base pairs, whole gametes, and even populations) are attributable to something like “sampling”—where sampling is just indiscriminate fixation or elimination of genes, gametes, etc., independent of their selective advantage or disadvantage.

In the initial models of drift, drift per se was not referred to as a cause, rather it was a consequence of “inbreeding effect.” Drift, like Darwin’s “fluctuation” was used as a metaphor to refer to the behavior of alleles under various forms of “random sampling.” During and after the early synthesis, the “Sewall Wright effect” was repeatedly invoked to explain patterns of apparently non-adaptive differences between species. Eventually, the term “random genetic drift” was deployed to generically describe the role of chance in populations over time at every level of organization.

What does this historical survey reveal about drift? There are four key conclusions one may draw:

(1) The classical Wright-Fisher model of drift as a binomial sampling process is an artifact of an age in which the mechanics of inheritance were poorly understood, and thus models the variety of ways in which chance affects evolutionary change in an idealized, indirect fashion. The shared feature of all cases of drift is that population size is relevant to predicting the effects of chance over

time. The varieties of modes of “sampling” at different levels of organization are more or less “black boxed” by this model. (2)

(2) That drift is referred to as both effect and cause is not new. Drift was originally a metaphor for the behavior of alleles in isolated, inbred populations. Only relatively recently has drift taken its place as one of several dynamic “causes” of evolutionary change.

(3) That Fisher modeled population genetic theory on statistical mechanics, and further, that Wright adopted the adaptive landscape metaphor for evolutionary change, helped solidify the view that drift, like selection, mutation, and migration, was a “force” of evolutionary change.

(4) Testing claims about the relative significance of drift have been difficult from the beginning; as illustrated by the debates over blood groups, as well as the contentious issues that so divided Fisher and Wright.

This historical overview may assist in assessing a heated philosophical exchange of late over whether and in what sense drift is a cause of evolutionary change. In part, the debate has hinged upon whether it is more appropriate to describe drift as effect or a cause of evolutionary change over time. As we have seen, this question has a historical precedent. For the concept of drift has itself evolved from a metaphor for describing the behavior of alleles due to chance in populations to a distinct “cause” or “force” of evolutionary change.

However, the philosophical debate has extended beyond the matter of whether it is more appropriate to speak of drift as outcome or process. Indeed, there are a family of overlapping issues at stake

here that have, unfortunately, become conflated. First, there is what I will call the metaphysical question: is drift (and, for that matter, natural selection) an epiphenomenal, or a genuinely supervienient, "population level" causal process? More generally, are there any population level "processes" operating in evolution? Second, there is what I will call the epistemological question. Namely, how are claims about drift (or selection) empirically substantiated? If many assessments of the relative significance of drift v. selection are empirically underdetermined; i.e., we see variation about a mean, or departure from expectation, and merely label it "drift," then what supports claims to the effect that there is a distinct "process" is at work here that yields this outcome? Why isn't "drift" simply a "trashbin" that we put unknowns in – a signal that we lack information about causal process, rather than a claim about an indiscriminate sampling process that we know to be operating? It seems that these two distinct issues have been conflated in the debate about drift as a cause of evolutionary change.

The exchange can be traced back to Sober's (1984) influential text on the nature of selection. Sober describes evolutionary theory as a "theory of forces": "In evolutionary theory, the forces of mutation, migration, selection and drift constitute causes that propel a population through a sequence of gene frequencies. To identify the causes of the current state… requires describing which evolutionary forces impinged" (1984: 141). Shapiro and Sober later call this the "conventional view" of evolution:

The conventional view, which WALM [Walsh, Ariew, Lewens and Matthen] oppose, is that natural selection, along with drift, mutation, migration, and mating pattern, are possible causes of evolution. These causes impinge on a population, sometimes changing its state while at other times causing the population to remain in the same state.[3] The causes of evolution behave in some ways like Newtonian forces. If two forces promote the evolution of a trait, it will increase in frequency at a faster rate than if just one of them were in place. And a population can be at equilibrium because opposing forces cancel each other… for example, selection is pushing it to increase in frequency while mutation pressure is pushing it to decline. WALM call this the dynamic view. (Shapiro and Sober forthcoming) <BLOCK QUOTE>

Shapiro and Sober argue that the view that genetic drift is a force or cause (just as are selection, migration, mutation) makes sense insofar as the reduction of population size (which they identify with drift) is one factor associated with changes frequencies of alleles in a population over time (p. 38). In other words, drift "removes" genetic variability in populations. Drift and mutation (or, for that matter, selection and migration) are thus spoken of as "forces" with opposing tendencies; drift removes variation, mutation restores it. These metaphors are surely problematic in some ways; yet, they do capture an important feature of evolutionary change at the level of populations. Isolation and inbreeding reduce genetic variation; and, small subpopulations are likely to contain less variation than large interbreeding populations – e.g., the species as a whole. In this sense, drift, or "chance" sampling, where this is understood to be any chance process involved in reducing the numbers of interbreeding individuals in some study population, is a "force" of change in the sense of (on average) reducing genetic variation.

This is referred to (appropriately) as a “consensus” view, insofar as many biologists today describe drift as if it were a “force” in reducing genetic variation over time due to something like the process of sampling error. (As we have seen, however, this was by no means the uniform sense of drift appealed to historically.) Like selection or mutation, drift can be spoken of metaphorically as opposed by competing “forces” – a population will be in equilibrium when these forces “cancel” one another. So, very small selective differences will not be terribly effective in smaller populations, due to the “effects” of drift (see also, Stephens, 2004).

In reply to Sober, Shapiro, and Stephens, Walsh (2002), Matthen and Ariew (2002) and Walsh et al. (2002) have argued that evolutionary theory is not a theory of forces. The disputants identify this view as a “dynamical” view of evolutionary theory, and contrast it with their preferred “statistical” view. At the center of the dispute is whether individual deaths, births, and interactions of organisms with their environments can or should be supplemented with population level explanations, or whether the latter are simply statistical descriptions with no genuinely causal import.

Matthen and Ariew (2002) claim that there are two concepts of fitness at work in evolutionary biology; one, “vernacular” or “ecological” fitness—fitness as relative adaptation or better “design solution,” versus fitness as a “predictive measure,” which is a statistical measure of the “expected relative rate of increase of some gene (or, trait), in future generations.” The former is not part of evolutionary theory at all, since there are no actual laws describing how fitness differences in this sense are caused. The latter is simply

a statistical measure, and thus, in their view, not a causal parameter at all.

Walsh (2002) similarly argues that natural selection is a mere "shadow" of genuine causal processes; "there is no need to invoke a *distinct* force [of natural selection] operating over populations," when, at the level of individual organisms, there already are the many causes of individual births and deaths (p. 139; italics in original; see also p. 150). Likewise, Walsh et al. (2002) defend the "statistical" over the "dynamical" interpretations of evolutionary theory: "Selection and drift are not forces acting on populations; they are statistical properties of an assemblage of "trial" events: births, deaths, and reproduction. The only genuine forces going on in evolution are at the level of individuals… and none of these can be identified with either selection or drift " (p. 453). Walsh, Lewens, and Ariew (2002) say that drift occurs when and only when a population exhibits a trait frequency that deviates from the frequency that would be predicted if selection alone acted. I.e., the "drifting" or "fluctuations" in gene frequency are simply departures from expectation, where what was expected was based on what WALM call "predictive" fitness. Similarly, but more recently, Pigliucci and Kaplan (2006) write:

Drift is not a process at all; the best sense that can be made of the concept appeals not to some property that particular populations share, but rather to the relative frequency of the kinds of changes that the populations have experienced. In a case in which changes observed are close to the changes statistically expected… we might say that the outcome reflects our expectations from predictive selection;

in a case in which the changes are more distant from the mean, we might say that the outcome does not reflect those expectations—that is, we might choose to call it an example of drift. But that does not imply that any kind of process took place in the latter population that did not take place in the former. (Pigliucci and Kaplan 2006: 28) <BLOCK QUOTE>

There seem to be two separate issues at play in the above passage. First, Pigliucci and Kaplan claim that drift is not a process. This first claim more or less echoes the earlier discussion of WALM, to the effect that "formal" selection and drift are not distinct causal processes. Second, they claim that what we call drift is often only a departure from expectation. This latter claim seems to be more of an empirical critique of the practice of biology than a metaphysical claim about the nature of drift. Arguably, these two issues may be pulled apart; whether or not there are population level causal processes in evolution, and whether biologists provide sufficient evidence in their assessments that drift or selection is operating are two separate questions, which may well have different answers. In other words, one diagnosis of the debate is that there has been a conflation of these two issues.

First, however, there are several respects in which critics of the "consensus" view are exactly correct. Formal models of selection and drift are statistical measures of change within populations over time. They treat average survival or reproductive success of trait groups as causal, and, to the extent that results depart from expectations, the outcome is often explained by "drift." Drift is often identified with the variety of unknown chance factors affecting

survival and reproductive success. Indeed, Sober might well agree with this as a description of much of biological practice.

However, it does not follow, at least without further argument, that drift and selection are not causes, or merely epiphenomenal. Both are genuine, if "supervenient" causal processes, according to Sober and Shapiro. Their argument for this claim depends upon a key assumption: "*investigating whether X causes Y involves figuring out whether wiggling X while holding fixed whatever common causes there may be of X and Y will be associated with a change in Y. It is not relevant to ask what will happen if one wiggles X while holding fixed the micro-supervenience base of X.*" This assumption draws upon Woodward's (2003) manipulationist model of causation; according to Woodward:

> A necessary and sufficient condition for X to be a direct cause of Y with respect to variable set V is that there be a possible intervention on X that will change Y or the probability distribution of Y when one holds fixed at some value all other variables in V.

For Woodward, there are very precise conditions that must be met in order for it to be the case that I is an intervention on X with respect to Y. An intervention on X must change the value of Y; it must be a sort of "switch" on Y, holding all other variables constant. Moreover, X must be a well-defined variable, such that it is clear what changing the value of this variable consists in. In sum, causal claims are essentially claims about how manipulation of one variable, or change in the value of that variable, is capable of changing the value of a second variable.

Shapiro and Sober argue that selection and drift are supervenient causes that meet these conditions. The “supervenience base” of selection, according to Shapiro and Sober, is the individual births and deaths of organisms in a population. Likewise, the “supervenience base” of drift is effective population size. One cannot “hold fixed” effective population size, and change the effects of drift. Claims to the effect that natural selection is not a cause because not distinct from the individual causal interactions that make up its supervenience base are thus not unlike claims to the effect that beliefs, desires, etc., are not causes, because not distinct from neurological states of the brain. Shapiro and Sober summarize:

We reply that while it is true that natural selection is not distinct from its supervenience base in a given token selection process, this is not a reason to deny that selection is a cause. In the same way, we regard the temperature, pressure, and volume of the gas in a container as causes even though they supervene on the states of the molecules making up the gas. Walsh demands that selection contribute something to evolution beyond the contributions made by the causal processes the impinge on individual organisms... Of course selection cannot do this, but that is no argument against its causal efficacy. To assess whether X causes Y, you shouldn’t try to hold fixed the micro-supervenience base of X while wiggling X. <BLOCK QUOTE>

In reply to WALM concerning drift, Shapiro and Sober argue that they confuse outcome with process:

Walsh, Lewens, and Ariew (2002) say that drift occurs when and only when a population exhibits a trait frequency that deviates from the

frequency that would be predicted if selection alone acted... For WALM, drift is a possible outcome; it is a product, not a process.[4]

Drift is part of the evolutionary process whenever population size is finite, just as selection is part of that process whenever there is variation in trait fitness. We see no harm in viewing these two "parts" of the evolutionary process as processes unto themselves... Selection and drift are distinct processes whose magnitudes are represented by distinct population parameters (fitnesses on the one hand, effective population size on the other). Changes in each of these parameters will be associated with changes in the probabilities of different outcomes. If you intervene on fitness values while holding fixed population size, this will be associated with a change in the probability of different trait frequencies in the next generation. And the same is true if you intervene on population size and hold fixed the fitnesses. Selection and drift are causes because they are difference-makers. Fitness values and population size are emphatically not like the barometer in its relationship to the weather. <BLOCK QUOTE>

In sum, on average, reducing population sizes increases the time to fixation of rare alleles. This is all it means to assert that drift is a cause, in Shapiro and Sober's view. Random factors play a role in fixing alleles in smaller populations more quickly than such factors will in larger populations, in the same way in which doing a shorter run of coin flips will be more likely to result in exclusively heads or tails. This is true for any ensemble of populations; smaller populations will show more deviations from expectation, due to "sampling error," where deviations from expectation may be due to anything from lightning striking to genetic recombination.

Forber and Reisman (2005) draw upon Dobzhansky's and Pavlovsky's drift "experiment" to illustrate this argument. Dobzhansky and

Pavlovsky illustrate how the variance of gene frequencies in an ensemble of populations may be increased or decreased as a result of choosing different sample sizes of populations. Insofar as manipulating genetic variance is possible, via selecting larger or smaller populations, drift is a cause of evolutionary change in Woodward's sense. Shapiro and Sober thus argue that drift "supervenes" over changes in Ne, effective population size. Drift is neither epiphenomenal nor causally inert, they say, because you cannot "wiggle" effective population size without "wiggling" drift as well.

It seems that even Pigliucci and Kaplan grant this claim about ensembles of populations in their discussion. They write, "In the formal sense, natural selection can be explanatory at the level of mean changes in frequencies of heritable features in populations — that is, at the level of ensembles of populations… Of course, if there are differences in predictive fitness, there must be discriminate processes at the individual level. However, the particular differences in predictive fitness we find at the formal level do not necessarily reflect those discriminate processes in any straightforward way."(p. 32). In other words, they seem to grant that formal models capture some explanatory relationship; where they stop short is in admitting that this is a causal relationship. And this seems largely to have to do with the *empirical grounds* for claims to the effect that selection is operating, "the particular differences in predictive fitness we find at the formal level do not necessarily reflect those discriminate processes in any straightforward way."

In other words, where it seems that the two camps divide is first, whether there are such things as population level supervenient

causal processes, and second, whether or not biologists are empirically justified in their judgments by and large. That is, on the one hand, all parties seem to agree that population genetics is an adequate "formal" account of the patterns of statistical distributions over time. Where they disagree concerns how models of selection treat property distributions as causes of outcome property distributions. In the case of selection, these properties yield fitness differences; i.e., properties such as fast speed, antibiotic resistance affect survival and reproduction. Thus, at the level of distributions of alleles over time, fitness coefficients associated with the possession of this or that trait predict outcomes. If we can grant that such indirect representations, or the averaging over of the variety of actual causal interactions that "fitness" represents is a kind of causal model, then we can happily grant that selection is a cause. However, not everyone can agree that such indirect representations are genuinely causal, or that there is a distinct "process" of selection. And at least in part, this has to do with the fact that claims about such processes are often underdetermined by the evidence.

Thus, there is at the core of this divide not only a question about the metaphysics of causation, but also, about the matter of evidence for claims about population level patterns and processes. With regard to the metaphysical question, this is entirely open; there remains serious debate in the philosophical literature concerning whether causes are one kind of thing or many, and, whether what may causally interact with what includes all and only token events, or, may also include types of events (Hausman 2005; Godfrey Smith, forthcoming). On the one hand, Sober and Shapiro claim that drift is a

cause, in the same way that taking a small sample is a cause of run of heads for a series of flips of a fair coin. In WALM's view, however, it seems all and only token events are causes; individual births and deaths, or physical interactions between individual organisms and their environments are causal; claims about the 'supervenience' of selection are perhaps merely redundant. One worry about adopting this restrictive view, however, is that we would have to give up making causal claims about all population level processes. Yet, this would mean that taking smaller samples is a not cause of skewed distributions; and, this is to rule out many explanations in statistics, biological and social sciences. It is fairly common, for instance, to argue that small sample sizes are a cause of misleading experimental results; indeed, this claim plays a central role in many of Pigliucci and Kaplan's arguments (2006). On WALM's criteria, gambling does not cause one to lose money; rather, a particular hand of cards or throw of the die is responsible. This seems to unduly rule out a large class of causal explanations that depend upon supervenience relationships.

In sum, what critics of the "consensus view" seem to be arguing is that not simply that "predictive fitness" or drift are not causal, but more fundamentally, that assessments of the relative role of selection and drift are poorly substantiated by evidence. That is, (1) in actual practice, biologists simply assume that the model holds true of the system of interest, and (2) more often than not, drift is applied to whatever departure from expectation we find in the data, with or without sufficient evidence that chance factors rather than some unknown other deterministic causes are at play.

With respect to this empirical claim, they have my sympathies; for, biologists have arguably been inconsistent and not always careful about the meaning and application of drift. And, assessing whether selection has been acting in any population does require more than knowledge of statistical distributions over time; we require further information about ecology, etc., and not simply population level correlations. However, Shapiro and Sober would (likely) be happy to grant this latter, epistemological point. It may well be the case that many claims about the relative significance of selection and drift are unsubstantiated; but this is not sufficient to show that drift and selection are epiphenomenal.

Conclusions <A>

Walsh, Ariew, Lewens, Matthen, Pigliucci, and Kaplan are correct that biologists have reified variables in their models as causal variables, when what these variables measure are rates of survivorship or fecundity, summations over a host of micro-level causal interactions. If you like, the “real” causal events in this context are interactions between discrete individual organisms and their environments. Selection coefficients are a statistical summation over many “actual” causal interactions between organisms and their environment. Kaplan and Pigliucci, on the other hand, seem to be arguing not only that selection is “epiphenomenal,” but also, that a more complete evolutionary explanation would require investigation into the ecological details. In other words, their point is not strictly a

metaphysical one concerning whether selection and drift are properly understood as causal processes, but an epistemological one, concerning whether or not sufficient evidence is mustered for claims about the relative significance of selection v. drift. These two points are quite different, and can and should be kept separate.

On the other hand, Shapiro, Sober, Reisman, and Forber are correct (assuming that Woodward's model of causation is unproblematic, and treating drift as supervenient over effective population size is regarded as requiring no special pleading), that drift is a cause of evolutionary change. In sum, if causal claims are essentially claims about how manipulation of one variable, or change in the value of that variable, on average, is capable of changing the value of a second variable, then in this sense, at least, drift is a cause.

However, the debate concerning the generality of Woodward's model of causation is not over (Cartwright, forthcoming; Hausman, 2005). Whether causal influence requires physical influence, whether causation requires relationships between all and only tokens or also types, not to mention how we are to assess problems of causal intermediaries, all requires further exploration. There seems to be room for a plurality of views of causation; with Cartwright (forthcoming) I would argue that we should not rule out a variety of types of causal explanation.

Critics of the "consensus view" believe treating drift as a supervenient causal process is problematic. How? One potential difficulty is that it is difficult to "locate" drift both historically and conceptually. What does "drift" supervene over, and is drift properly understood as the cause or effect of the distributed random

processes that collectively reduce genetic heterogeneity? Is drift reduction in population size, gamete sampling, meiosis and recombination, the random processes effecting survivorship and fecundity, or all of the above? Finite population size is perhaps more appropriately described as a *condition on the possibility* of sampling error than a cause of reduced heterozygosity. Surely, decreasing population size increases random sampling, but it is also true that the causes and consequences of sampling are distributed and different at different levels of organization. It is thus difficult to "locate" drift at any one stage of this process extended over time.

There is good reason to be troubled here; for historically, drift has referred to any and all of the above – both the variety of conditions on, and outcomes of, isolation, inbreeding, and chance in evolution. Drift is often biologists' catch-all term for any and all changes in gene frequency due to stochastic, non-deterministic events.[2] At different levels of organization, and at different time scales, however, there are very different "engines" or sources of stochasticity. Gamete sampling during sexual reproduction is a different "cause" or "engine" of drift from the random ways in which survivorship and fecundity might be affected. Thus, while it is true that the same basic models of drift are used, at least as a first pass, to model drift at different levels of organization and at

[2] Of course, not all authors agree that there are indeterministic events in evolution. Some (E.g. Rosenberg, 1994; Rosenberg, Horan and Graves, 1999) have argued that to the extent that drift, if it is a process at all, must be a deterministic one, and, deny that evolutionary theory's reliance on statistical explanations is due to reasons that are not purely epistemological. For the purposes of this article, this larger set of issues will be set aside here.

different time scales, arguably, the “real” causal processes are different at these different levels of organization and time scale.

If one may grant that highly idealized, statistical models legitimately represent causes, then population level, supervenient causation via selection and drift are unproblematic. However, if you hold that all and only ecological interactions between organisms and their environments are causal, then treating birth or death rates, or population size to be causes, is at best, not intuitively appealing. Sober’s reply suggests that evolutionary genetic explanations are distinctive; the models, like folk theories of mind, perhaps supervene over much more complicated biological details, and are good short-hand for population level dependency relations. Insofar as our mathematical models eliminate some apparent contingency, or demonstrate how and why this observation depends upon some more basic, better-understood mathematical regularity, it is explanatory. In this way, models of drift explain in virtue of the law of large numbers; smaller sample sizes yield different distributions. Yet, this model of explanation is itself contentious; what Walsh, Matthen, Ariew, and Lewens seem to be contesting is not simply whether and how drift or selection are causes, but whether explanations in the biological sciences are sufficient if based on idealized models using statistical parameters. They seem to be suggesting that we need to understand more about the molecular biology, ecology, development, etc., of the organism in order to satisfactorily explain evolution.

Were we to consult the authors of theoretical population genetics, perhaps Wright would be more sympathetic with WALM, and Fisher would be more sympathetic with the “consensus” view defended by

Shapiro and Sober. Wright knew well that his models were idealizations, or indirect representations of complicated biological interactions, making many strictly speaking false assumptions. Surely, Wright would contend, understanding more about ecology, development, etc., is crucial to both explanation and testing claims about adaptive evolution. Yet, in the Fisherian vein, the models are useful (even if only idealized) tools for understanding population level dynamics.

However, there is, surely, a place for both top-down and bottom-up explanatory strategies in biology. That is, we need both the Wrightian and the Fisherian perspectives on evolutionary dynamics over time. The key is to find a middle ground between the two which integrates our best understanding not only of population level dynamics, but also, of the messy biological details.

Notes

1. For ease of exposition, “drift” will be used interchangeably with “random genetic drift” from this point forward. Thanks to a reviewer for noting the significance of the distinction.
2. King and Jukes (1969) published an article defending the same view in *Science,* with the radical title, “Non-Darwinian Evolution.”
3. The state of a population is characterized by specifying the frequencies of different traits in it.
4. Brandon and Carson (1996: 324-325) view drift in the same way.

References

Beatty J (1992) Random drift. In: Keywords in Evolutionary Biology (Keller EK, Lloyd EA, eds), 273-281. Cambridge, MA: Harvard University Press.

Beatty J (1987) Dobzhansky and drift: Facts, values, and chance in evolutionary biology. In: The Probabilistic Revolution (Krüger L, Gigerenzer G, Morgan MS, eds), Vol. 2, 271-311. Cambridge, MA: MIT Press.

Brandon RN (2005) The difference between selection and drift: A reply to Millstein. Biology and Philosophy 20: 153-170.

Brandon R. and Carson S. 1996. The indeterministic character of evolutionary theory: no 'No Hidden Variables' proof but no room for determinism either. Philosophy of Science 63: 315-337

Cartwright, N. (forthcoming) "Causation: One Word Many Things"

Crow JF (1954) Breeding structure of populations. II. Effective population number. In: Statistics and Mathematics in Biology (Kempthorn O, Bancroft TA, Gowen JW, Lush JL, eds), 543-556. Ames, IA: Iowa State College Press.

Darwin C (1859) The Origin of Species. Facsimile Edition (Mayr E, ed). Cambridge, MA: Harvard University Press.

Dobzhansky T (1937) Genetics and the Origin of Species. New York: Columbia University Press. 2nd ed, 1941; 3rd ed, 1951.

Ewens WJ (2004) Mathematical Population Genetics. 2nd rev. ed. New York: Springer.

Falk R# (2000) The gene: A concept in tension. In: The Concept of the Gene in Development and Evolution: Historical and Epistemological Perspectives (Beurton P, Falk R, Rheinberger H-J, eds), 317-348. Cambridge: Cambridge University Press.

Fisher RA (1922) On the dominance ratio. Proceedings of the Royal Society of Edinburgh 42: 321-341.

Fisher RA (1930) The Genetical Theory of Natural Selection. Oxford: Clarendon press. 2nd ed., 1958. New York: Dover.

Gillespie JH (1998) Population Genetics: A Concise Guide. Baltimore, MD: Johns Hopkins University Press.

Gillespie JH (1991) The Causes of Molecular Evolution. Princeton, NJ: Princeton University Press.

Godfrey-Smith, P. (forthcoming) "Causal Pluralism." Oxford Handbook of Causation, eds. Helen Beebee, Chris Hitchcock, and Peter Menzies. Oxford: Oxford University Press.

Gould SJ (1983) The hardening of the Modern Synthesis. In: Dimensions of Darwinism (Grene M, ed), 71-93. Cambridge: Cambridge University Press

Grauer D, Li WH (2000) Fundamentals of Molecular Evolution. Sunderland, MA: Sinauer.

Gulick JT (1889) Intensive segregation, or divergence through independent transformation. Journal of the Linnaean Society of London 23: 312-380. Reprinted in: Evolution, Racial and Habitudinal (Gulick JT, ed), #-#. Washington, DC: Carnegie Institution of Washington, 1905.

@Hagedoorn AL, and A. C. (1921) The Relative Value of the Processes Causing Evolution. The Hague: Nijhoff.

Hausman D (2005) Causal relata: Tokens, types or variables. Erkenntnis 62: 31-51.

Hartl DL (1988) A Primer of Population Genetics. 2nd ed. Sunderland, MA: Sinauer.

Huxley JS (1942) Evolution: The Modern Synthesis. London: Allen and Unwin.

Kimura M (####)Discussion Forum: Molecular Evolution," TIBS, July.

King JL and Jukes, TH 1969. "Non-Darwinian Evolution." Science. 164: 788-798.

Matthen M, Ariew A (2002) Two concepts of fitness. Journal of Philosophy 94: 55-83.

Mayr E (1942) Systematics and the Origin of Species. New York: Columbia University Press.

Mayr E (1959) Where are we? Cold Spring Harbor Symposium in Quantitative Biology 24: 1-14.

Mayr E (1963) Animal Species and Evolution. Cambridge, MA: Harvard University Press.

Millstein RL (2002) Are random drift and selection conceptually distinct? Biology and Philosophy 17: 33-53.

Millstein RL (2005) Selection vs. drift: A response to Brandon's reply. Biology and Philosophy 20: 171-175.

Provine W (2001) The Origins of Theoretical Population Genetics. 2nd ed. Chicago: University Of Chicago Press.

Reisman K, Forber P (2005) Manipulation and the causes of evolution. Philosophy of Science 72: #-#.

Rosenberg, Alexander (1994), Instrumental Biology or the Disunity of Science. Chicago: University of Chicago Press.

Rosenberg, Horan and Graves (Graves, L., Horan, B.L. et al.: 1999, 'Is Indeterminism the Source of the Statistical Character of Evolutionary Theory?' Philosophy of Science 66, 140—157.

Shapiro L, Sober E (forthcoming) Epiphenomenalism: The do's and don't's. In: Studies in Causality: Historical and Contemporary (Wolters G, Machamer P, eds), #-#. Pittsburgh: University of Pittsburgh Press.

@Skipper R (in press) Stochastic evolutionary dynamics: Drift and draft. Proceedings of the Philosophy of Science Association, 2004.

Stephens C (2004) Selection, drift, and the "forces" of evolution. Philosophy of Science 71: 550-570.

Walsh, D. 2000. Chasing shadows: Natural selection and adaptation. *Studies in the History and Philosophy of Biological and Biomedical Sciences,* 31, 135-153.

Walsh D, Lewens T, Ariew A (2002) The trials of life: Natural selection and random drift. Philosophy of Science 69: 452-473.

Woodward J (2003) Making Things Happen: A Theory of Causal Explanation. Oxford: Oxford University Press.

Wright S (1929) The evolution of dominance: Comment on Dr. Fisher's reply. American Naturalist 63: 556-561.

Wright S [1931] (1986) Evolution in Mendelian populations. Genetics 16: 97-159.

Wright S (1949) "Adaptation and Selection." In Genetics, Paleontology and Evolution, ed. J. L. Jepson, G. G. Simpson, and E. Mayr, 365-389. Princeton: Princeton University Press.

Wright S (1955) "Classification of the Factors of Evolution." Cold Spring Harbor Symposium in Quantitative Biology. 20: 16-24.

[Box]

The Classical Wright-Fisher Model of Drift

The simplest model of drift is often called the Wright-Fisher model, due to its origins in the work of both Wright (1931) and R. A. Fisher (1922). This model treats the production of offspring as the drawing of alleles (where, there are two alleles at a gene “locus”) at random with replacement from parents, or alternatively drawing gametes from an infinite pool to which each parent has contributed equally. The binomial sampling process represents both what was standardly assumed to be both (1) a random process of meiotic segregation and fertilization, and (2) the role of chance the numbers of (successful) gametes produced by parents (e.g., the chance factors while make it the case that some individuals have, say, 10 and some 0 offspring (lightning striking, etc.)). Together, in the Wright-Fisher model, these produce a binomial progeny distribution. In the simplest case, generations are discrete and there are only two alleles at a locus. So, for instance, consider two individuals in a parent generation, one of which is a heterozygote Aa and another a homozygote, AA. Given the assumption that the alleles at this locus are passed on via Mendelian independent assortment, these two individuals can have offspring of one of two sorts, either AA, or Aa, expected in equal numbers. By chance alone, they may have an equal number of AA and Aa offspring, or, alternatively, offspring that are all AA, or offspring that are all Aa. Summed over the population as a whole, the change in distribution of genotypes due to this sampling process is called

drift. In other words, the “cause” of a drift in this sense is simply redistribution due to Mendelian segregation and random fertilization, or, accidents of “sampling” of alleles.

This model makes a number of assumptions; the organisms are diploid, they reproduce sexually, generations are non-overlapping, there is random mating within sample subpopulations, no migration, no mutation, and no selection. All of these assumptions are violated, in at least some contexts. In other words, the Wright-Fisher model, although the pedagogical and theoretical starting point for almost all work on drift today, is hardly ever realized by actual populations. Genes of the same chromosome are linked, so they are not inherited independently. This means that when more than one locus is considered, the theory must be extended to allow for this correlation. There are other complications. For example, the progeny distribution is usually not binomial; there are often separate sexes with unequal numbers; the population may not be mating at random. One way of dealing with these complications is to use Wright's concept of "effective population number."

There are a variety of ways of defining effective population number: the most common are the inbreeding, and the variance effective population size. The inbreeding effective population number is the size of the ideal population that would produce the same probability of identity by descent among selected individuals as exists in the actual population. The variance effective population size is the size of a population with the same dispersion of allele frequencies under drift, or allele frequency variance as the population being studied.

The effective number defined in this way is called the "variance effective number" (see Wright 1931 and Crow 1954: 543-556).

<eof>