

Mean Field Theory and the Gene Centered View of Evolution

Yaneer Bar-Yam

New England Complex Systems Institute

277 Broadway Cambridge MA 02139, USA

Abstract

While recent years have seen a movement away from the gene centered view of evolution, it continues to have a strong hold on the conceptual foundations of biology. A formal understanding of the strengths and weakness of this view is lacking. In this article we show that the gene-centered view directly corresponds to a mean-field approximation in the reproduction-selection dynamics. This explains both why the gene centered view is useful and limited in application to evolution. Effective gene fitness result from (time dependent) averages over the current organism pool to obtain a mean-field environment for the gene. Such averaging is justified if mixing by sequel reproduction of the population is rapid compared o trait divergence of sub-populations. When trait-divergence is important, the mean field approximation breaks down. The latter is particularly important over larger time scales in understanding the global properties of evolution, where trait divergence and speciation are essential features to be understood.

A basic formulation of evolution requires reproduction (trait heredity) with variation and selection with competition. At a particular time, there are a number of organisms which differ from each other in traits that affect their ability to survive and reproduce. Differential reproduction over generations leads one organisms offspring to progressively dominate over others and changes the composition of the population of organisms. Variations during reproduction allows offspring to differ from the parent and an ongoing process of change over multiple generations is possible.

One of the difficulties with this conventional view of evolution is that many organisms reproduce sexually and the offspring of an organism are thus often as different form the parent as other organisms that it is assumed to be competing against. Note that this conceptual difficulty does not apply to asexually reproducing organisms. To address this fundamental paradox, the gene-centered view was introduced. In the gene-centered view there are assumed to be indivisible elementary units of the genome (thought of as individual genes) that are preserved from generation to generation. Different versions of the gene (alleles) compete and mutate rather than the organism as a whole. Thus the subject of evolution is the allele, and, in effect, the selection is of alleles rather than organisms.

The simple picture that allelic competition (gene-centered evolution) is the fundamental process of evolution was strongly advocated by some evolutionary biologists, while others maintained more elaborate pictures which, for example, differentiate between vehicles of selection (the organisms) and replicators (the genes). In this article we will review the mathematics of some standard conceptual models of evolution to clarify the relationship between gene-centered and organism-based notions of evolution. We will show that the gene centered view is equivalent to a mean field approach where correlations between the different genes are ignored. Each gene evolves in an effective environment formed within the organism and its environment. This effective environment is an average environment (mean field) within a sexually reproducing population (e.g. species). By showing that the gene-centered view of evolution is a mean field approach, we can recognize why is is useful and we can also recognize when it is invalid—when correlations are relevant.

Correlations between genes arise when the presence of one allele in one place in the genome affects the probability of another allele appearing in another place in the genome. One of the confusing points about the gene-centered theory is that there are two stages in which the dynamic introduction of correlations must be considered: selection and sexual

reproduction (gene mixing). Correlations occur in selection when the probability of survival favors certain combinations of alleles, rather than being determined by a product of terms given by each allele separately. Correlations occur in reproduction when parents are more likely to mate if they have certain combinations of alleles. If correlations only occur in selection and not in reproduction, the mean field approach continues to be at least partially valid. However, if there are correlations in both selection and sexual reproduction then the mean field approach and the gene-centered view break down. Indeed, there are cases for which it is sufficient for there to be very weak correlations in sexual reproduction for the breakdown to occur. For example, populations of organisms are distributed over space and an assumption that reproductive coupling is biased toward organisms that are born closer to each other can self-consistently generate allelic correlations in sexual reproduction by symmetry breaking. Thus, this is particularly relevant to considering trait divergence of sub-populations.

To clarify how standard models of evolution are related to this picture, it must be recognized that the assumptions used to describe the effect of sexual reproduction are as important as the assumptions that are made about selection.

A standard first model of sexual reproduction assumes that recombination of the genes during sexual reproduction results in a complete mixing of the possible alleles not just in each pair of mating organisms but rather throughout the species—the group of organisms that is mating and reproducing. Offspring are assumed to be selected from the ensemble which represents all possible combinations of the genomes from reproducing organisms.

If we further simplify the model by assuming that each gene controls a particular phenomic trait for which selection occurs independent of other gene-related traits, then each gene would evolve independently; a selected allele reproduces itself and its presence within an organism is irrelevant.

Without this further assumption, selection should be considered to operate on the genome of organism. Thus, correlations may be induced in the allele populations in the surviving (reproducing) organisms. Nevertheless, due to the assumption of complete sexual mixing, the correlations disappear in the offspring. From the point of view of a particular allele at a particular gene, the complete mixing means that at all other genes alleles will be present in the same proportion that they appear in the population—there are no allele correlations after reproduction. Nevertheless, because selection operates on the genome, fitness depends

not on individual genes but rather on gene combinations. As the presence of one allele in the population changes in the population due to evolution over generations, the fitness of another allele at a different gene will be affected. However, due to the assumption of complete mixing in sexual reproduction only the average effect (mean field) of one gene on another is relevant. Thus the assumption of complete mixing in sexual reproduction is equivalent to a gene based mean-field approximation.

This qualitative discussion of standard models and their relationship to the mean-field approximation can be shown formally. We write a two-step model for sexual reproduction:

$$\{N(s, t)\} = R[\{N'(s; t - 1)\}] \quad (1)$$

$$\{N'(s; t)\} = D[\{N(s; t)\}] \quad (2)$$

The first equation describes reproduction. The number of offspring $N(s; t)$ having a particular genome s is written as a function of the reproducing organisms $N'(s; t - 1)$ from the previous generation. The second equation describes selection. The reproducing population $N'(s; t)$ is written as a function of the same generation at birth $N(s; t)$. This reflects selection—the differential survival of organisms from birth to reproduction. The brackets on the left indicate that each of these equations actually represents a set of equations for each value of the genome. The brackets within the functions indicate, for example, that each of the offspring populations depends on the entire set of parent populations.

A mean field approximation is performed by assuming that the reproduction step (not necessarily the selection step) depends only on the proportion of alleles and not on their specific combinations in the reproducing population. This proportion can be written as the number of organisms which have a particular allele s_i at gene i divided by the total number of organisms:

$$P'(s_i; t) = \frac{1}{N'_0(t)} \sum_{s_j, j \neq i} N'(s; t) \quad (3)$$

where $s = (s_1, \dots, s_N)$ represents the genome in terms of alleles s_i . The sum is over all alleles of genes j except gene i . $N'_0(t)$ is the total reproducing population at time t . According to our assumption about reproduction, the same offspring would be achieved by a population

with a number of reproducing organisms given by

$$\tilde{N}'(s, t) = N'_0(t) \prod_i P'(s_i; t) \quad (4)$$

since this has same proportions as Eq. 3. The form of this equation indicates that the probability of a particular genome is a product of the probabilities of the individual genes—they are independent. Thus complete reproductive mixing assumes that:

$$R[\{\tilde{N}'(s; t)\}] \approx R[\{N'(s; t)\}] \quad (5)$$

It follows that a complete step including both reproduction and selection can also be written in terms of the allele probabilities in the whole population. The update of an allele probability is:

$$P'(s_i; t) \approx \frac{1}{N'_0(t)} \sum_{s_j, j \neq i} D[R[\{\tilde{N}'(s; t-1)\}]] \quad (6)$$

Given the form of Eq. 4 we could write this as an effective one-step update

$$P'(s_i; t) = \tilde{D}[\{P'(s_i; t-1)\}] \quad (7)$$

which describes the allele population change. Thus the assumption of complete mixing by sexual reproduction allows us to write the evolution of a single allele in this way. However, because Eq. 7 is a function of all the allele populations, the fitness of an allele is coupled to the evolution of other alleles.

Eq. 4 describes the neglect of allele correlations in reproduction consistent with a mean field approximation. It should be apparent that this is only a first approximation. It is valid only when the allelic correlations induced by selection are weak enough to be reversed by the gene mixing during sexual reproduction. In more realistic models correlations between genes affect both reproduction and selection.

We can provide a specific example of breakdown of the mean field approximation using a simple example, which has a conceptual history (discussed below) in the controversy of the gene-centered view. We start by using a simple model for population growth to define a fitness parameter λ . An organism that reproduces at a rate of λ offspring per individual

per generation has a population growth described by an iterative equation:

$$N(t) = \lambda N(t - 1) \tag{8}$$

We obtain a standard model for fitness and selection by taking two equations of the form Eq. 8 for the two populations $N_1(t)$ and $N_2(t)$ with λ_1 and λ_2 respectively, and normalize the population at every step so that the total number of organisms remains fixed at N_0 . We have that

$$\begin{aligned} N_1(t) &= \frac{\lambda_1 N_1(t - 1)}{\lambda_1 N_1(t - 1) + \lambda_2 N_2(t - 1)} N_0 \\ N_2(t) &= \frac{\lambda_2 N_2(t - 1)}{\lambda_1 N_1(t - 1) + \lambda_2 N_2(t - 1)} N_0 \end{aligned} \tag{9}$$

The normalization does not change the relative dynamics of the two populations, thus the faster-growing population will dominate the slower-growing one according to their relative reproduction rates. If we call λ_i the fitness of the i th organism we see that according to this model the organism populations grow at a rate that is determined by the ratio of their fitness to the average fitness of the population.

Consider now sexual reproduction where we have multiple genes. In particular, consider two nonhomologue genes [1] with selection in favor of a particular combination of alleles on genes. Specifically, after selection, when allele A_1 appears in one gene, allele B_1 must appear on the second gene, and when A_{-1} appears on the first gene allele B_{-1} must appear on the second gene. We can write these high fitness organisms with the notation $(1, 1)$ and $(-1, -1)$, and the organisms with lower fitness (for simplicity, $\lambda = 0$) as $(1, -1)$ and $(-1, 1)$. When correlations in reproduction are neglected there are two stable states of the population with all organisms $(1, 1)$ or all organisms $(-1, -1)$. If we start with exactly 50% of each allele, then there is an unstable steady state. In every generation 50% of the organisms reproduce and 50% do not. Any small bias in the proportion of one or the other will cause

[1] Homologue genes are genes on homologue chromosomes that are at the same location and thus serve the same organismal function and allow the same alleles as a result of crossover during sexual reproduction. It is helpful to recall that during sexual reproduction an offspring obtains half of the chromosomes of nuclear DNA from each parent. The chromosomes are paired in function—homologous pairs. Each homologue chromosome of the offspring is formed in a parent by a process (crossover during meiosis) that combines segments of DNA from both parents homologues. The case of considering two homologue genes can also be treated (see reference 4) but does not serve as useful an example for this discussion.

there to be progressively more of one type over the other, and the population will eventually have one set of alleles.

We can solve this example explicitly for the change in population in each generation when correlations in reproduction are neglected. It simplifies matters to realize that the reproducing parents (either $(1, 1)$ or $(-1, -1)$) must contain the same proportion of the correlated alleles (A_1 and B_1) so that:

$$\begin{aligned} P_{1,1}(t) + P_{1,-1}(t) &= P_{1,1}(t) + P_{-1,1}(t) = P_1(t) \\ P_{-1,1}(t) + P_{-1,-1}(t) &= P_{1,-1}(t) + P_{-1,-1}(t) = P_{-1}(t) = (1 - P_1(t)) \end{aligned} \quad (10)$$

The reproduction equations are:

$$\begin{aligned} P_{1,1}(t) &= P_1(t-1)^2 \\ P_{1,-1}(t) &= P_{-1,1}(t) = P_1(t-1)(1 - P_1(t-1)) \\ P_{-1,-1}(t) &= (1 - P_1(t-1))^2 \end{aligned} \quad (11)$$

The proportion of the alleles in the generation t is given by the selected organisms:

$$P_1(t) = \left(P'_{1,1}(t) + P'_{1,-1}(t) \right) \quad (12)$$

Since the less fit organisms $(1, -1)$ and $(-1, 1)$ do not reproduce this is described by:

$$P_1(t) = P'_{1,1}(t) = \frac{1}{P_{1,1}(t) + P_{-1,-1}(t)} P_{1,1}(t) \quad (13)$$

This gives the update equation:

$$P_1(t) = \frac{P_1(t-1)^2}{P_1(t-1)^2 + (1 - P_1(t-1))^2} \quad (14)$$

which has the behavior described above and shown in Fig. 1. This problem is reminiscent of an Ising ferromagnet at very low temperature. Starting from a nearly random state with a slight bias in the number of UP and DOWN spins, the spins align becoming either all UP or all DOWN.

Since we can define the proportion of a gene in generation t and in generation $t + 1$ we

can always write an expression for allele evolution in the form:

$$\begin{aligned}
 P(s_i; t) &= \lambda_{s_i} P(s_i; t - 1) \\
 \sum_{s_i} \lambda_{s_i} &= 1
 \end{aligned}
 \tag{15}$$

so that we have evolution that can be described in terms of gene rather than organism behavior.

The fitness coefficient λ_1 for allele A_1 and B_1 is seen from Eq. 12 to be:

$$\lambda_1(t) = P_1(t)
 \tag{16}$$

with the corresponding $\lambda_{-1} = 1 - \lambda_1$. One difficulty with this equation is in the time dependence of the fitness through its dependence on the changing population. In steady stat, λ values would not change. Of course, in steady state there is no need to describe the dynamics. One could argue that from the perspective of describing the evolution of organisms in terms of fitness values the equation is only useful as a description of the dynamics if the values of λ are slowly varying in time compared to the changes in P .

It is interesting, however, to consider when this picture breaks down more severely due to a breakdown in the assumption of complete reproductive mixing. In this example, if there is a spatial distribution in the organism population with mating correlated by spatial location and fluctuations so that the starting population has more of the alleles represented by 1 in one region and more of the alleles represented by -1 in another region, then patches of organisms that have predominantly (11) or $(-1 - 1)$ will form after several generations. This symmetry breaking, like in a ferromagnet, is the usual breakdown of the mean field approximation. Here, it creates correlations in the genetic makeup of the population. When the correlations become significant then the species has a number of types. The formation of organism types depends on the existence of correlations in reproduction that are, in effect, a partial form of speciation—what is important is whether interbreeding occurs in reality, not whether it is possible.

Thus we see that the most dramatic breakdown of the mean field approximation / gene centered view occurs when multiple organism types form. This is consistent with our understanding of ergodicity breaking, phase transitions and the mean field approximation.

Interdependence at the genetic level is echoed in the population through the development of subpopulations. We should empathize again that this symmetry breaking required both selection and reproduction to be coupled to gene correlations [2].

The simple example we have discussed has an interesting conceptual history. It is analogous to the example of the right-handed and left-handed rowers used by Dawkins [3] and Lewontin [4] to argue for and against the gene-centered view. We will review their arguments and show how they are related to this discussion.

In the rowers analogy, there are two types of rowers, left-handed and right-handed. A boat gains an advantage in speed when it is formed from more same-handed rowers—rowers of different handedness interfere with each other (of course this is not necessarily justified by direct analysis of rowing but that is not the point of the example). The rowers compete in heats of a certain number of boats. At the beginning of a race, rowers are assigned at random to boats and the winners of each heat are replicated, replacing the ones that were defeated. Details like the number of rowers per boat, or the number of boats per race, are not essential to the analysis. The mathematical analysis given above corresponds to having two rowers per boat, but it can be easily generalized.

In this picture we can see that over time, one of the types of rowers will come to dominate the other kind of rower, because starting from a bias (even a random bias) in the number of left or right handed rowers, it will be more likely for the dominant type to have more of its type of rower in a boat. Thus, as pointed out by Dawkins, even though the boats are selected as winners, the rowers reproduce to increase the number of the dominant kind. In the analog of the gene-centered view of evolution, we see that one rower type will dominate the other and the selection of boats has served to select rowers.

The analysis by Lewontin of this situation argued that the claim of Dawkins was misleading because it was impossible to assign meaningful fitnesses to each of the rower types. This argument uses the result of Eq. 16 to suggest that the fitnesses vary with time and are not

[2] We note that if there is a small bias in the fitness of (11) over $(-1 - 1)$ then the formation of the two types will not persist due to competition between them. To enable the distinct types to persist there is need for the existence of multiple resources each resource causing one of the types to be more fit. This is a general feature of evolution models not restricted to the one discussed here.

[3] Dawkins, R. (1989); *The Selfish Gene*, 2d ed. (Oxford University Press: Oxford) p. 86

[4] Lewontin, R. in R. N. Brandon and R. M. Burian, eds. (1984); *Genes, Organisms, Populations: Controversies Over the Units of Selection* (MIT Press, Cambridge)

given by the underlying properties of the rowers and thus are not helpful in understanding the evolutionary process.

This limitation of the mean-field approach however, can be seen to be only part of the story. By introducing small correlations in rower selection, we can create two populations of left-handed and right-handed boats, which correspond to symmetry broken subpopulations. This is the example of trait divergence of organisms which, as discussed above, is important for the larger scale properties of evolution.