

The evolution of evolutionary engines

Marcelo O. Magnasco

August 8, 2001

Center for Studies in Physics and Biology,
Rockefeller University, 1230 York Avenue, New York
magnasco@rockefeller.edu

Abstract

Abstract: Evolution is based on a complex and intertwined interplay between generation of new variants and their selection by competition in the environment. Much research has established that the generation of evolutionary diversity is not a passive process that just "occurs" to organisms: rather, a fluid and active set of strategies is in place so that living beings control the ways in which diversity is generated. We describe some of the consequences for our current views of evolutionary theory and survey how the understanding of the field is changing.

These notes attempt to introduce some current views of the mechanisms through which living beings evolve. This attempt is both incomplete and idiosyncratic. An introduction to elementary biological concepts was given first in the Geilo lectures and is obviously required; I obviate it here since there are excellent textbooks, and since I used lots of colorful diagrams taken from them anyway. The reader unfamiliar with concepts such as DNA transcription, regulation of gene expression, etc, may wish to consult the original sources behind my introduction: Alberts et al's *Molecular Biology of the Cell* [1], Darnell et al's *Molecular Cell Biology* [2], and T. A. Brown's *Genomes* [3]. An excellent (though mostly nonmathematical) introduction to evolutionary theory is given in Ridley's *Evolution* [4]. I shall henceforth assume general background knowledge.

1 Evolution and mutations.

"the mutation rate of living beings is itself encoded in their DNA, and so it is both inheritable and mutable"

David Thaler, May 1994, 3am. (the meaning and consequences of this statement are the subject of these notes)

The classical view of selection is that it is a two-step process, in which alternate generation of diversity (G.O.D., previously known as "descent with modification") and

selection. See Figure 1. Variants are generated, by imperfect copying of the genome or by damage in transit. Since the genome is a blueprint, changes to the blueprint cause changes in the actual organisms, which affect their ability to function in their environment. That this is some sort of optimization recipe is self-evident and I shall not explore it much further; the question is whether this is what we actually see happening in the organisms around us. These ideas have been widely used as computational methods to obtain optimal solutions to difficult problems in a technique called Genetic Algorithms, with which the condensed matter community got acquainted mostly by way of the study of glasses.

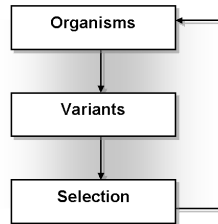


Figure 1: The classical view of selection

Already Darwin stated quite forcefully that evolution through generation of variants and natural selection would be an enormously slow process, very gradual, requiring many generations. The variants are very, very similar to their parents. Gradualism has been an article of faith within the mainstream of evolutionary theory for the past century; attempts against it are derided with such names as "hopeful monsters theories" and other such. Even someone like Steven Jay Gould felt the need to defend his punctuated equilibrium theory [5] by claiming that the rapid changes the theory describes still takes many generations. It is instructive to notice that, in contrast, laboratory mutagenesis experiments routinely come up with rather gross deviants following mutations: since the blueprint in the genome is not a passive description of the structure but rather encodes a developmental program to be executed, mutation of an instruction somewhere can result in a rather large alteration of the final structure. A famous example from years ago was a mutant fly in which a perfectly formed leg was coming off its eye socket—somewhere the subroutine call "insert eye here" was mutated to "insert leg". To see examples of variants, see for instance Lindquist's paper [6].

However, Darwin's gradualism actually was a point against him for a very long time. In Darwin's time, the prevailing view of "genetics" was one of "blending inheritance": since children look a bit like mom and a bit like dad, the notion was that sexual mating "blended" the traits together. Now, a bit of thought reveals that given blending inheritance in a population that stays overall constant in size, an advantageous trait which randomly appears through generation of diversity will be exponentially dissolved into the population and there will be no net evolution. Thus Darwin's theory fell into disrepute for some time, until at the turn of the century Mendel's work was rediscovered after half a century of oblivion. Mendel's experiments had shown that traits are inherited in full quanta: either you inherit the trait or you don't; since you get

two copies of everything, some traits are only displayed when you have inherited two copies (recessive) and for some one suffices (dominant); thus, *displaying* the trait is slightly different from actually *carrying* it; Mendel had shown that the combinatorics for displaying traits showed that some were controlled by a single dichotomic variable, while others responded to more complex combinations. Mendelian inheritance now matched Darwinian theory perfectly: the advantageous trait might be lost, by chance, but also by chance it could propagate undiluted; and it stood, by virtue of being advantageous, a better chance at propagating. In this way, it could in due time take over, with undiluted strength, the entire population. (Homework: a classic result states that a trait which is advantageous by p percent over the wild-type has a probability of p of being fixed in the entire population after sufficient time. Can you see why?)

This prompted serious mathematical developments; today we call this golden era of mating Darwin, Mendel and mathematics the neo-Darwinism, or *the evolutionary synthesis* [7]. Its founding fathers, H.B.S. Haldane, Sewall Wright and others put forth the following framework. We consider the population to be a distribution over all possible phenotypes. The fitness of an organism is the number of its offspring which reach reproductive maturity; divided by two in the case of sexual species. (This assumes a constant generation span). For each phenotype, there shall be an "expected fitness": while the actual fitness of any organism will be a random variable, it has an expected value based on the advantages and disadvantages of the particular phenotype under consideration. Thus, there is a function called "expected fitness" whose range is the set of all phenotypes; it's usually called the "fitness landscape"[8, 9]. Thus, descent with modification is an operator which smears the population distribution in the space of phenotypes, while natural selection both re-normalized the distribution, and skews it towards those phenotypes having a higher fitness. When a simple case is considered, the math becomes alarmingly familiar [10].

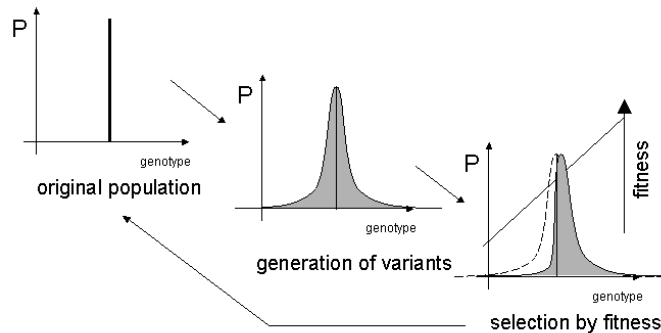


Figure 2: selection as a random walk on a landscape.
Descent with modification is quite akin to a diffusion operator, broadening the

amount of variants around the mode of the population, while natural selection introduces a drift term pushing the mode to higher fitnesses. It is not a coincidence that this theory was being developed at the same time as the theory of random walks was providing a foundation for chemical kinetics. The Haldane/Wright view can be mapped into a random walker undergoing thermal motion in an (inverted) potential.

2 Fitness landscapes are rugged

It has been argued that fitness landscapes are far from smooth, they are “rugged”, meaning they have fractal-like structure at all scales [11]. I have yet to find a prettier or more insightful description of what a “rugged fitness landscape” is than Theodosius Dobzhansky’s 1951 analogy to the large-scale structure of mountain ranges:

Thus, the ecological niche occupied by the species "lion" is relatively much closer to those occupied by tiger, puma and leopard than to those occupied by wolf, coyote and jackal. The feline adaptive peaks form a group different from the group of canine peaks. But the feline, canine, ursine [...] peaks form together the adaptive "range" of the carnivores, which is separated by deep adaptive valleys from the "ranges" of rodents, bats, ungulates [...] In turn, these ranges [form] the adaptive system of mammals, which [differs from those of] birds, reptiles, etc.

T. Dobzhansky, *Genetics and the origin of species* [12]

Darwin made a very clear point in *Origin of Species* that the struggle for survival starts at home: most organisms are sufficiently well adapted to the “elements” and thus their struggle is not a competition against them, but rather against other organisms—first and foremost the nearby members of its own species, then the members of species it directly interacts with, such as their prey or predators, and only then the “elements”, such as inclement weather. A striking example of this is the emperor penguin, the only large vertebrate to brave the winter in the south pole. As spring arrives, the females leave the colony to search for food, while the male stays (still fasting) taking care of the egg. The female then must fight to grab as much squid as she can—a fight against both the squid and the other females. Thus, a correct description of the environment of an organism is not one of a fixed, given and immutable landscape, but rather a landscape dynamically generated by interactions with other living beings. It was a long while after Haldane until it became clear how to do this.

The proper way to describe such dynamically changing interactions is in the framework of game theory. Game theory was invented by von Neumann and Morgenstern [13] to describe the actions of individual agents in economics. One of the most famous game-theoretic constructs is the prisoner’s dilemma, a game in which depending upon the parameters the optimal game strategy results in cooperation between two competing agents. Another version, the iterated prisoner’s dilemma, in which many rounds are played; the competing agents thus have a chance to choose many times whether to compete or defect. It has as one optimal strategy tit-for-tat, in which each player decides to cooperate or defect by copying what the previous player has done before. Game theory was introduced into evolutionary models by John Maynard-Smith, precisely to explain

the evolution of cooperativity. Contrary to the common misconception that Darwinian evolution depicts savage competitions, Nature is replete with examples of cooperative strategies, from the evolution of two sexes [14], to scores of examples of symbiosis [15]. It had been a deep mystery to the early practitioners of the neo-darwinian synthesis how to actually treat such stuff formally. For an introduction, see J.M-S's *Games, sex and evolution* [16]. A game-theoretic fixed point (a la Nash) is called an "evolutionarily stable strategy". Recently people have argued that self-organization of an ecosystem results in this ruggedness [17].

Most laboratory experiments, however, deal with fixed challenges, a shortcoming that only now is becoming evident. (See last section). In this context one may model the experiments in terms of a fixed and immutable landscape.

3 Summary of background

Let us review the background we have covered up to now:

- organisms encode themselves (or better stated, encode algorithms to generate themselves) in their genome. This is accomplished through nested tiers of information-bearing and enzymatically active polymers.
- genome gets translated, somehow, into what the organism actually is (in functional terms) called the phenotype. This process is extremely complex, ill-understood, obscure.
- classically evolution was conceived as an iteration of two distinct noninteracting steps,
 - descent with modification, also known as generation of diversity ("GOD"), in which variants are generated without regard as to their ultimate utility,
 - and natural selection in which they are filtered regardless of how they were made.
- changes to DNA causes mutations. Mutations manifest themselves as either nothing happening (most of the time [18]) or alterations that range from the subtle to the quite dramatic. There exist single letter changes that are lethal (though they are few)
- fitness, namely the number of offspring who reach reproductive maturity, is the outcome of game-theoretic competition. Nevertheless, within a large population, any given individual does not substantially change by itself the environment, and hence has a probabilistically defined "expected fitness". Confounding the two has caused much misery. (I have perpetrated such confusion many times)
- thus we can draw or infer a fitness landscape. Fitness landscapes are rugged, meaning their local minima have lots of structure.
- generation of diversity acts as a diffusion operator in genotype space

- selection skews population distributions according to fitness and acts as a drift
- continued rounds of selection and generation-of-diversity act as a diffusion-drift system, I.e., as a brownian walker in the fitness landscape.

4 Mutation rates

Mutation rates are measured in terms of mistakes per letter copied per generation of copying.

organism	mutation rate	comments
viral	10^{-2} to 10^{-4}	both mutators and antimutators
in vitro (taq)	10^{-2} to 10^{-5}	stringency depends on buffer[19, 20]
<i>E. coli</i>	10^{-4} to 10^{-6}	variations of 1000 between strains[21, 22, 23]
<i>Drosophila</i>	10^{-8}	variations of 10 between strains[24]
Mammals	$< 10^{-10}$	(except for cancerous hypermutators [25, 26, 27])

Please notice that these numbers are quite impressive for the density of information we're discussing: information is stored as letters in a polymer, in aqueous solution, at room temperature. The raw rate of errors in man-made storage media (before various algorithms like parity correction take care of them) is not above these: modem transmission has error rates comparable to viral replication, diskettes and tapes have error rates similar to *E. coli* rates, and until quite recently hard drives (whose surfaces are sequestered away from the environment in hermetic seals) did not exceed *Drosophila* error rates.

It has been noted that the decreasing error rate for more complex organisms exactly matches the increasing genome length, in the case of animals[28, 29]. If instead of measuring the mutation rate in errors per letter copied per cell division, one measures it in total letter errors in the genome (i.e., per letter times genome length), per generation (i.e., times the number of cell divisions in the germ line), one obtains a number between $\frac{1}{2}$ and $\frac{1}{4}$ for all organisms from *E. coli* through mammals. Plants do not fit in this scheme since they can have huge genomes—we do not really know why, but the common lily has a genome 20 times larger than the human one. The implication is clearly that evolutionary behaviour tunes the mutation rate to an appropriate level. However, these studies only considered the average mutation rate of a population (or better stated, the *typical* mutation rate of the population) and not individual variation.

5 Mutagenesis experiments.

A strain of *E. coli* can be made such that the lacZ gene has a lethal single point mutation; i.e., this particular strain cannot metabolize lactose. Only a single letter needs to be changed for the bacterium to be able to eat lactose. Putting such a strain in lactose-rich medium leads to reversion to lac competence within 100 generations or so in colonies of 10⁸ bacteria. In such experiments, it is not unusual to recover *E. coli*

strains which are mutators, i.e., they have mutation rates much larger than the wild type [21, 22, 23, 30].

Usually, this happens because they have defective error-correcting genes. DNA is precious and fragile. It gets attacked frequently by physical and chemical means. Many mechanisms have evolved to repair and ensure faithful copying [31, 1]. During replication, the accuracy of DNA polymerase is enhanced by a double-checking mechanism known as kinetic proofreading [32, 33]. After replication, some number of errors in copying still have been made. Is there any chance of correcting them? Yes! In order to distinguish the original and the copy strands, cells methylate their DNA. Thus, when mismatches are detected (bulges in the double helix) the unmethylated strand can be removed and you get a second go at copying. After faithful reproduction has been ascertained, the daughter strands are methylated and the pathway is turned off. Quite important in bacteria is the methyl-directed pathway, involving mutS and associated proteins [31, 34, 35]. Other pathways involve checking for oxidative damage, checking for strand breakage, maintaining the purity of the nucleotide pool, etc. In higher organisms, in addition, cells are asked to commit suicide (apoptosis) upon finding irreparable damage to the cell (which could result in cancer: p53).

Repair of oxidative damage example: "Spontaneous oxidation of guanine residues in DNA generates 8-oxoguanine (oxoG). By mispairing with adenine during replication, oxoG gives rise to a G≡C → T=A transversion, a frequent somatic mutation in human cancers. The dedicated repair pathway for oxoG centres on 8-oxoguanine DNA glycosylase (hOGG1), an enzyme that recognizes oxoG≍C base pairs, catalysing expulsion of the oxoG and cleavage of the DNA backbone." [36].

6 The *no-multiple-mutations* fallacy

[Fred Hoyle] claims that the origin of new major groups is impossible without such intervention [extraterrestrial DNA] because, "What mutations cannot do is to find improvements which demand the simultaneous change of several base pairs". *Evolutionary biologists would agree that a change requiring a number of base changes, each of which is without value until all are present, cannot occur by natural selection.* They have therefore concluded that the origin of major groups has been a stepwise process, with each genetic change being an advantage on its own [...]. If there is no stepwise path up the mountain, natural selection won't climb it. Much thought has been given to the nature of the intermediate steps.

John Maynard-Smith, reviewing Fred Hoyle's book *Mathematics of Evolution*, in *Nature* **403**, 594 - 595 (2000). (Italics my emphasis)

Under what circumstances is the joint probability of two mutations happening simultaneously the product of individual probabilities that any of them will happen?

$$P(A \wedge B) = P(A)P(B)$$

if and only if:

- the mutation process is not physiologically triggered
- the mutation process is not directed
- the mutation rates are constant in time
- the mutation rates are constant along the genome

Any other situation results in $P(A \wedge B) \neq P(A)P(B)$ i.e., mutations will appear to be statistically correlated. As an example, let us consider two subpopulations: a "wild type" subpopulation with low mutation rate, and a "mutator" subpopulation with higher mutator rate. Let's say 99% of the population has 10^{-6} mutation rate, while 1% has 10^{-3} . What is the probability of seeing a mutant? a double mutant? a triple mutant?

- single mutants: $0.99 * (10^{-6}) + 0.01 * (10^{-3}) = 1.1 * 10^{-5}$
- double mutants: $0.99 * (10^{-6})^2 + 0.01 * (10^{-3})^2 = 10^{-8} \neq (1.1 * 10^{-5})^2$
- triple mutants: $0.99 * (10^{-6})^3 + 0.01 * (10^{-3})^3 = 10^{-11} \neq (1.1 * 10^{-5})^3$

So, what Thaler meant when he said "the mutation rate of living beings is itself encoded in their DNA, and so it is both inheritable and mutable" was that insults to DNA, both physical and chemical, happen all the time. The damage they cause is not yet a mutation: it's called a premutagenic lesion. It will become a mutation if and only if the enzymes of DNA metabolism do not correct it. Therefore the mutation rate of an organism is a direct function of its error-correcting enzymes, which are encoded in DNA like any other protein. If these enzymes are changed, the mutation rate will be changed heritably.

7 A more complete view of evolution

So in a more complete view of evolution, the environment enters directly in three ways [30]: The environment is "perceived" by the organism, which used this perception to modify their own physiology, as in operon induction, and their genetic metabolism, as in the SOS pathways. Finally, the organism modifies the environmental interaction with the genome as in metabolic activation or detoxification. Selection does not act only upon those inheritable traits leading to a phenotype that will directly intervene in the struggle for survival: all of the genes in charge of generating genetic variability are subject to inheritance and selection. Thus the environment interacts with genetics in various ways:

- it is the proximate agent of selection,
- it directly impinges on the DNA via such agents as radiation and chemical mutagens, and
- it interacts with DNA via the genes of DNA metabolism.

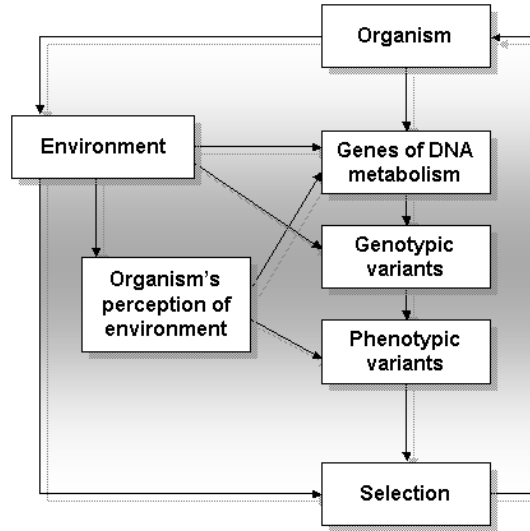


Figure 3. A more complete view of selection. From [30]

All of the conditions for statistical independence in the previous section are known to be violated. Stress responses trigger high mutation rates physiologically (e.g. in the SOS pathway). Mutation rates can also transiently be increased in hypermutable states [37]. There are various mutation rate and recombination hotspots in the genome [45]. Various mutation processes can appear to be directed, because they function through the various mechanisms which we use to perform genetic engineering [38, 39, 40].

8 Allegorically

Given the allegory “evolution” \approx “brownian walk on a landscape”, then mutation rate \approx temperature. Based on this rather slim rationale, we modeled a brownian walker capable of modifying its own temperature:

$$\begin{aligned}\dot{x} &= -\partial_x V + g(y)\xi_1(t) \\ \dot{y} &= -\beta y + g(y)\xi_2(t)\end{aligned}$$

where and we use the Ito convention. Anomalous features are evident the moment we jot down the Fokker-Planck equation for this system and plug in the usual double-well potential to check out escape rate times [41]. First, the equations are structurally incapable of detailed balance: currents only cancel if either F or T are constant. Second, given a double well (or double peak in F), the system is free to choose at what temperature T_c to cross the well. A saddle-node calculation quickly shows $T_c \approx Q/\log Q$ for

sufficiently large Q and hence from the Kramers estimate $\tau_K \approx e^{Q/T_c} \approx Q^\alpha$ for some unspecified α . Thus, as an evolutionary challenge becomes steeper and steeper, it will be surmounted by mutator strains with higher and higher mutation rates—but it still shall not be insurmountable in the “exponentially small probability” sense that classic evolutionary theory (as incarnated in the Maynard-Smith quote above) would like us to think. *By having a mutation rate that is another inheritable property subject to variability and selection, evolution acquires a means to change its own pace dynamically. This ability to change the pace results in exponentially enhanced ability to surmount difficult challenges.*

This result was substantially strengthened by a serious, nonallegorical calculation showing the coevolution of a wild-type and mutator strain in the face of an evolutionary challenge [42]. Mutator phenotypes provide alternative, faster pathways through which adaptive genes can be generated. This is true even though the mutator phenotype is extremely rare. Double-challenge mutagenesis experiments in bacteria have shown a surprising result: the time for a genotype to arise which undoes two simultaneous, independently lethal mutations is not much bigger than the sum of the individual times. Thus adaptive two-point mutations have arisen (in the lab at least) within reasonable (ie algebraic, not geometric) timescales.

Further to this exploration, one can consider not only time variability of the mutation rate, but a variability along the genome. Contingency loci [43] are sites in the genome (loci) where different genes are encoded in such a way that a mutation can transform one into another (through frameshift, for instance), and these genes are used for sporadically encountered situations (contingency) Example (Moxon [44, 45]): *H influenzae* has a number of iterated sequences lying around. Polymerases tend to stutter on them and generate frameshifts. Most (14) iterated sequences lying within coding regions encode hypermutable virulence factors, all surface proteins involved in contact with the host such as pili. Time-independent set of several different mutation rates: maps to a multiple-temperature problem.

9 Towards "natural" selection

"I think the stability of the genome is a lab artifact, a direct consequence of culturing bacteria under constant conditions"

F. Taddei, private communication (1997)

Experiments in evolutionary theory can only be done in experimental organisms having a fast reproductive time so that many generations can be observed within experimentally reasonable timescales. Bacteria are an obvious choice, though the applicability of the results may be circumscribed—multicellular sexual organisms evolve in ways that emphasize other mechanisms. For studies of evolutionary phenomena intrinsic to these "higher" taxa, fast reproducing insects such as fruit flies have been the choice. The game-theoretic aspects of sexual selection, for instance, have been recently studied by allowing for instance only males to evolve: every generation of flies, all females in the experimental population are destroyed and new females are introduced from a constant, "thermal bath" of females kept under constant conditions. This results in the

males accumulating changes in their sexual strategies which the females cannot adapt to. The result was astounding: within a few generations the male flies had acquired awfully aggressive mutations, such as toxic sperm,[46] barbed harpoons in their sexual apparatus, etc. By not allowing the females to evolve counterdefenses, this experiment showed the true intensity of the evolutionary arms race that underlies sexual species [47, 48].

Experiments in the bacterial realm have long been performed under extremely artificial conditions. Sometimes this is really what is experimentally sought: for instance, Sid Brenner proposed culturing bacteria in heavy water, where their metabolism slows down enormously, so that they would be forced to shed as many nonessential genes as possible (to accelerate the DNA replication time)—this might result in a "minimal" organism having the fewest number of genes necessary to sustain life in a nonthreatening environment. But many times what is being desired by culturing the bacteria is not to "evolve" them under constant conditions, but simply to "keep" them. Bacteriology labs keep strains by culturing them. Alas, as you culture them, they necessarily evolve, adapting to their constant environment. For instance, cancerous cell lines are kept in labs so that experiments can be done on them—to test chemotherapy agents, for instance. After a while, these cell lines become so much the standard gold test of anything related to cancer that any study of a new chemo agent must include a test of efficiency on them. Some immortalized cancerous cell lines date back to the fifties—they have been continuously growing in petri dishes ever since, long outliving the patient they killed. According to some researchers, they don't look like human cells anymore: they've adapted to growth on lab media, they can stick to the plastic of the petri dishes. (Sandy Simon, private communication).

The lungs of cystic fibrosis (CF) patients are chronically infected for years by one or a few lineages of Pseudomonas aeruginosa. These bacterial populations adapt to the highly compartmentalized and anatomically deteriorating lung environment of CF patients, as well as to the challenges of the immune defenses and antibiotic therapy. These selective conditions are precisely those that recent theoretical studies predict for the evolution of mechanisms that augment the rate of variation. Determination of spontaneous mutation rates in 128 P. aeruginosa isolates from 30 CF patients revealed that 36% of the patients were colonized by a hypermutable (mutator) strain that persisted for years in most patients. Mutator strains were not found in 75 non-CF patients acutely infected with P. aeruginosa.

Oliver et al [49]

The steady realization that cultured organisms look nothing like wild-type organisms (poodles vs. wolves) has prompted a growing revolt against the use of constant culture conditions [50]. Many cancer researchers now think that only fresh clinical isolates are appropriate for serious study, and that immortalized cell lines are not really "cancerous" anymore—malignant tumors are characterized by the ferocity with which they misappropriate the resources of their host, and the culturing process likely is taming them, destroying the very property that one wishes to study and understand. The same is true of bacteria. Most bacteria under active study now are pathogens. (Curiously,

the vast majority of bacteria are non-pathogenic. About two tenths of a percent of the bacterial species that coexist with humans cause illnesses—many other bacteria coexist symbiotically with us, like the intestinal flora which colonizes our guts almost as soon as they are made, on the fourth month of gestation). Many pathogens lose virulence after being cultured for long enough. But in addition to virulence, the mechanisms through which one can adapt to new conditions, the evolutionary engine, may be lost as well. In particular, if one cultures bacteria in periodically changing conditions (as opposed to constant), mutator strains are stabilized. According to F. Taddei, (private communication) clinical isolates of *E. coli* from patients suffering acute infection show that 30% of the bacteria are various mutator strains which stabilize or "evanesce" as soon as they are cultured. This has led him to believe that the wild-type genome is a substantially more fluid object than the genome of the tame cultured strains of *E. coli*. Thus, him and many other specialists [51] now believe that studies of the evolutionary mechanisms better be carried out in a more natural environment;—for example, in the gut of live mice [52]. The natural clinical isolates from patients themselves [49] are an unwelcome but naturally occurring evolutionary laboratory; Oliver and colleagues studied infections by *P. aeruginosa* of the lungs of cystic fibrosis patients and showed that 36% of a test group of patients had been infected by hypermutable strains which persisted for years (see quote above).

10 Conclusion

Living beings possess mechanisms to change, transiently or in heritable ways, the rate and kinds of their own genetic variation. These mechanisms are under tight biological control and interact with the physiology of the organism, and with the information the organism has about its environment. The generation of diversity is not something that "happens" to organisms: living beings have actually taken charge of, and actively control, the kinds and rates of variation on which their fate rests. Attempts to model evolutionary behavior in such a "self-referential" context, by, i.e., explicitly modeling a wild-type and mutator subpopulations, have shown qualitatively distinct behavior from the standard models. There is thus not just the evolution of survival strategies, but also of strategies for generating evolutionary novelty itself: a "gear shift" for evolution and a set of rules to use it.

I'd like to thank David Thaler, John Maynard-Smith, Dawn Field, Miroslav Radman, François Taddei, Charlie Doering, Jim Shapiro, Itzak Rabin, Guillermo Cecchi and Gustavo Stolovitzky, for the teaching me all I know about this subject, and a truly large bunch of other people. My research in this area was supported by the Mathers Foundation and the Sloan Foundation.

Figure 2

References

- [1] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson, *Molecular Biology of the Cell*, 3rd ed, Garland, New York, 1994.

- [2] H. Lodish *et al.*, *Molecular Cell Biology*, W. H. Freeman & Co., 1999.
- [3] T. A. Brown, *Genomes*, Wiley-Liss (1999)
- [4] Mark Ridley, *Evolution*, Second Edition, Blackwell Science, Cambridge, 1996
- [5] Eldredge N., Gould S.J., in *Models in Paleobiology*, Schopf T. J. M. ed., San Francisco: Freeman, Cooper and Co. (1972)
- [6] S. L. Rutherford and S. Lindquist, *HSP90 as a capacitor for morphological evolution*, *Nature*, **396** 336, (1998)
- [7] E. Mayr, *Some thoughts on the history of the evolutionary synthesis*, in *The evolutionary synthesis: perspectives on the unification of biology*, E. Mayr and W. Provine, eds, Harvard University Press, Cambridge, 1980.
- [8] N. Eldredge, *Macro-evolutionary dynamics: species, niches, and adaptive peaks*, McGraw Hill (1989)
- [9] S. Wright, *Proc. Sixth Int. Congr. Genetics* **1** 356 (1932)
- [10] Newman CM, Cohen JE, Kipnis C, *Nature* **315** 400 (1985)
- [11] Kauffman, S. A. *The origins of order: self-organization and selection in evolution*. Oxford: Oxford University Press (1993)
- [12] T. Dobzhansky, *Genetics and the Origin of Species*, 3rd ed., Columbia University Press (1951)
- [13] J. von Neumann and O. Morgenstern, *Theory of games and economic behaviour*, Princeton University Press, Princeton, 1944.
- [14] J. Maynard-Smith, *The evolution of sex*, Cambridge University Press, Cambridge, UK, 1978.
- [15] J. Maynard-Smith and E. Szathmary, *The major transitions in evolution*, W.H. Freeman/Spektrum, Oxford UK and New York, 1995
- [16] J. Maynard-Smith, *Games, sex, and evolution*, Harvester-Wheatsheaf, New York, 1988.
- [17] P. Bak and K. Sneppen, *Punctuated equilibrium and criticality in a simple model of evolution*, *Phys Rev. Lett.* **71** 4083-4086 (1993)
- [18] Kimura, M. *Genet. Res. (Camb.)* **9** 23 (1967), and Maruyama T. and Kimura M. *Evolution* **28** 161 (1974)
- [19] Eckert, K. A. and Kunkel, T. A. *PCR Meth. and Applic.* **1** 17 (1991)
- [20] Kunkel, T. A. *J. Biol. Chem.* **267** 18251 (1992)
- [21] Rayssiguier, C.A., Thaler, D.S. and Radman, M. *Nature* **342** 396 (1989)

- [22] Goodman M. F., Creighton, S., Bloom, L. B. and Petruska, J. *Crit. Revs. Biochem. and Mol. Biol.* **28** 83 (1993)
- [23] Radman, M., Matic I., Halliday J. A. and Taddei F. *Phil. Trans. Royal Soc. London B* **347** 97 (1995)
- [24] Sturtevant, A. H., *Quat. Rev. Biol.* **12** 464 (1937)
- [25] Nowell, P. C., *Science* 194 23 (1976)
- [26] Fishel R. and Kolodner, R. D., *Curr. Opin. Genet. & Devel.* **5** .3 382 (1995)
- [27] Boyer JC, Thomas DC, Maher VM, McCormick JJ, Kunkel TA, *Cancer Res.* **53** .14 3270 (1993)
- [28] Drake, J. W. *Proc. Natl. Acad. Sci. USA* **88** 7160 (1991)
- [29] Drake, J. W. *BioEssays* **14** .2 137 (1992)
- [30] Thaler, D. S. *Science* **264** 224 (1994)
- [31] Kornberg A., Barker T., *DNA Replication* , Chapter 21, W. H. Freeman (1992)
- [32] J. J. Hopfield, *Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity*, *Proc. Natl. Acad. Sci. U.S.A.* **71** 4135-4139 (1974)
- [33] J. Ninio, *Kinetic amplification of enzyme discrimination*, *Biochimie* **57** 587-595 (1975)
- [34] M. Cascalho et al, *Mismatch Repair co-opted by hypermutation*, *Science* **279** 1207 (1998)
- [35] F. Taddei et al, *Correction by MutT Protein of transcriptional errors caused by oxidative damage*, *Science*, **278**, 128 (1997)
- [36] S. D. Bruner et al, *Nature* **403** 859 (2000)
- [37] J. Ninio, *Biochimie* 73, 1517 (1991)
- [38] Shapiro, J.A., *Adaptive mutation: who's really in the garden?* *Science* **268** 373-374 (1995)
- [39] Shapiro, J.A. *Transposable elements as the key to a 21st century view of evolution*, *Genetica* **107** (1-3): 171-179 (1999)
- [40] Shapiro, J.A. *Structured pathways for variation: Evolution as a biological function*, *Ann NY Acad. Sci.* **870** 95-98 (1999)
- [41] M. Magnasco and D. Thaler, *Changing the pace of evolution*, *Phys. Lett. A* **221**(5) 287-292 (1996)

- [42] F. Taddei et al, *Role of mutator alleles in adaptive evolution*, *Nature* **387** 700 (1977)
- [43] D. Field et al, *Contingenci loci, mutator alleles and their interactions*, *Ann. NY Acad Sci* **870** 378-382 (1999)
- [44] E. R. Moxon et al, *Curr. Biol.* **4** 24-33 (1994)
- [45] E. R. Moxon and D. Thaler, *The tinkerer's evolving toolbox*, *Nature* **387** 659 (1997)
- [46] Rice W. R., *Sexually antagonistic male adaptation triggered by experimental arrest of female evolution*, *Nature* **381** 232-234 (1996)
- [47] N. K. Michiels and L. J. Newman, *Sex and violence in hermaphrodites*, *Nature* **391** 647-64 (1998)
- [48] Crudgington, H. S. and Siva-Jothy M. T., *Genital damage, kicking and early death: The battle of the sexes takes a sinister turn in the bean weevil*, *Nature* **407** 855-856 (2000)
- [49] A. Oliver et al, *High Frequency of Hypermutable Pseudomonas aeruginosa in Cystic Fibrosis Lung Infection*, *Science* **288** 1251 (2000) .
- [50] M. Radman, I. Matic, F. Taddei, *Evolution of evolvability*, *Ann. N.Y. Acad. Sci.* **870** 146-155 (1999)
- [51] I. Matic and M. Radman, *Highly variable mutation rates in commensal and pathogenic Escherichia coli*. *Science*, **277**, 1833 (1997), and see the following response by J. LeClerc and T. Cebula
- [52] A. Giraud et al, *Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut*, *Science* **291** 2606 (2001)