

# Laws of Adaptation

A course on biological evolution in eight lectures  
by Carlo Matessi

## Lecture 3

Invading mutants, or the “long-term dynamics” point of view

Part I – time scales of evolution

Wednesday October 4, 16:00-17:00

## Natural selection on multiple genes

- ❖ Genotypes are defined by a pair of *gametes* (string of genes), received from the mother and from the father respectively
- ❖ Set of possible gametes:  $\mathcal{G} = \{g_1, g_2, \dots, g_n\}$ ; set of possible genotypes:  $\mathcal{G} \times \mathcal{G}$
- ❖  $X_{ij}(t)$ ,  $Y_{ij}(t)$  = proportion of  $g_i g_j$  among adults, newborn of generation  $t$
- ❖ Individuals produce gametes by a rearrangement (*recombination*) of parental gametes

$$R(i|jk) = \text{Probability that genotype } g_j g_k \text{ produces gamete } g_i, \quad R(i|jk) = R(i|kj) \quad \forall i, \quad \sum_i R(i|jk) = 1 \quad \forall (jk)$$

- ❖  $p_i(t)$  = frequency of  $g_i$  among gametes produced by adults of generation  $t$

$$p_i(t) = \sum_{jk} X_{jk}(t) R(i|jk)$$

- ❖ In the general recurrence equation let:  $f_{klrs}(ij) = R(i|kl) R(j|rs)$

$$Y_{ij}(t+1) = \sum_{kl} \sum_{rs} X_{kl}(t) X_{rs}(t) R(i|kl) R(j|rs) = p_i(t) p_j(t)$$

- ❖ Gametic frequencies are sufficient state variables:

$$p_i(t+1) = \frac{\sum_{jk} w_{jk} p_j(t) p_k(t) R(i|jk)}{\bar{w}(t)}, \quad \bar{w}(t) = \sum_{jk} w_{jk} p_j(t) p_k(t)$$

$$Y_\alpha(t+1) = \frac{\sum_{\beta} \sum_{\gamma} X_\beta(t) X_\gamma(t) f_{\beta\gamma}(\alpha)}{\sum_{\delta} \sum_{\beta} \sum_{\gamma} X_\beta(t) X_\gamma(t) f_{\beta\gamma}(\delta)}$$

$$X_\alpha(t+1) = \frac{w_\alpha Y_\alpha(t+1)}{\sum_{\beta} w_\beta Y_\beta(t+1)}$$

## The case of two genes with two alleles

❖ Genes and alleles: gene A  $\equiv \{A_1, A_2\}$  ; gene B  $\equiv \{B_1, B_2\}$

❖ Gametes:

label	1	2	3	4
gamete	$A_1B_1$	$A_1B_2$	$A_2B_1$	$A_2B_2$

❖ Recombination:

$\frac{A_iB_k}{A_jB_l} \rightarrow$	gamete	prob.	type
	$A_iB_k$	$\frac{1-r}{2}$	parental
	$A_iB_l$	$\frac{r}{2}$	recombinant
	$A_jB_k$	$\frac{r}{2}$	recombinant
	$A_jB_l$	$\frac{1-r}{2}$	parental

❖ Recurrence equations

$$p_i(t+1) = \frac{p_i(t)\bar{w}_i(t) + r w_{14} \eta_i [p_2(t)p_3(t) - p_1(t)p_4(t)]}{\bar{w}(t)} \quad , \quad \eta_1, \eta_4 = 1 \text{ and } \eta_2, \eta_3 = -1$$

$$\bar{w}_i(t) = \sum_j w_{ij} p_j(t) \quad , \quad \bar{w}(t) = \sum_{ij} w_{ij} p_i(t) p_j(t) \quad , \quad w_{14} = w_{23}$$

$$p_2(t)p_3(t) - p_1(t)p_4(t) = \text{linkage disequilibrium}$$

## Failure of fitness maximization

$$p_i(t+1) = \frac{p_i(t)\bar{w}_i(t) + r w_{14} \eta_i [p_2(t)p_3(t) - p_1(t)p_4(t)]}{\bar{w}(t)}$$

where :  $\eta_1, \eta_4 = -1$  and  $\eta_2, \eta_3 = 1$

- ❖ Any fully polymorphic equilibrium satisfies:

$$\bar{w}_i - \bar{w} = r w_{14} \frac{\eta_i}{\hat{p}_i} (\hat{p}_1 \hat{p}_4 - \hat{p}_2 \hat{p}_3)$$

hence it cannot correspond to a maximum of the mean fitness, unless linkage equilibrium prevails, which generally is not the case

❖ Moreover, fitness does not necessarily increase throughout generations, as shown in this example:

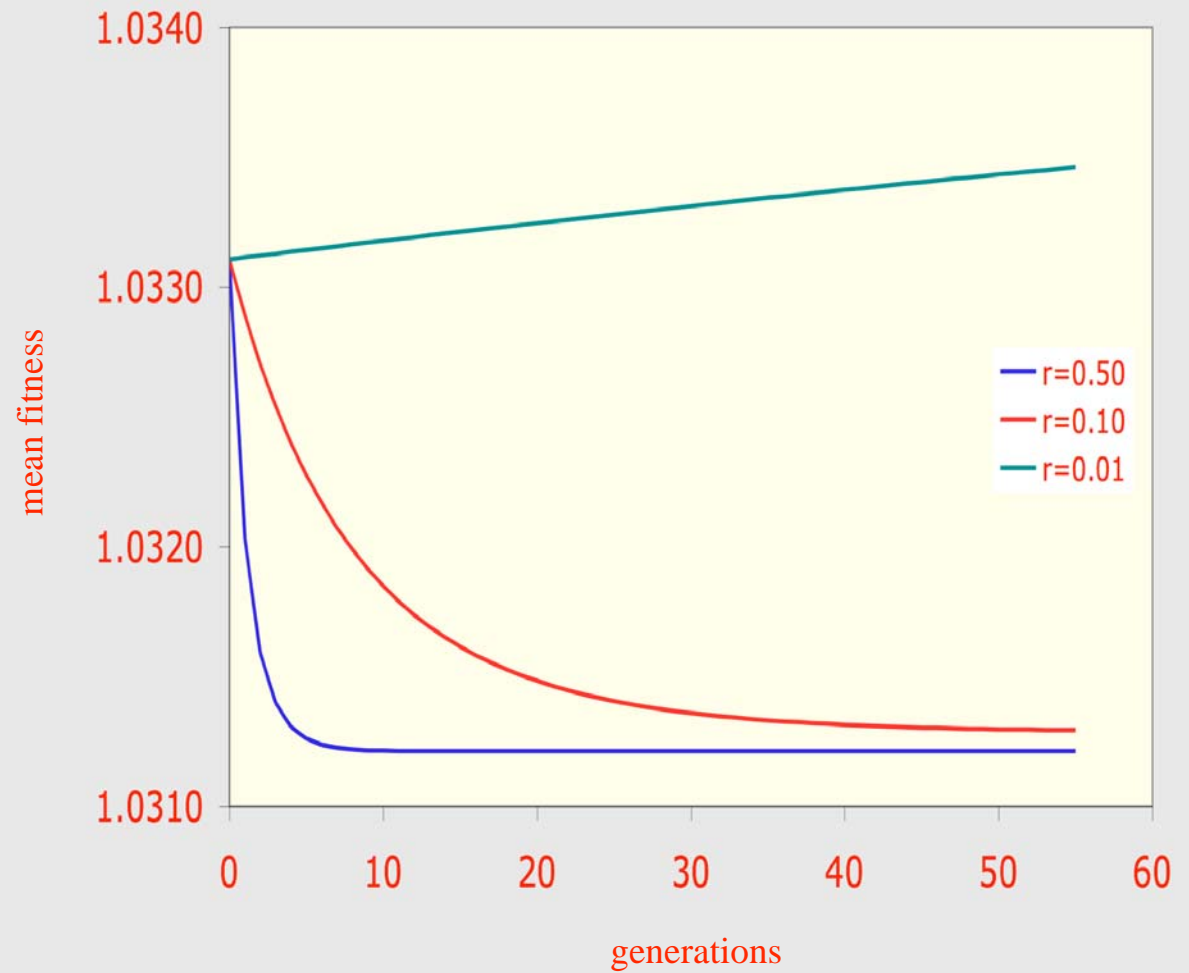
Fitness matrix

	B <sub>1</sub> B <sub>1</sub>	B <sub>1</sub> B <sub>2</sub>	B <sub>2</sub> B <sub>2</sub>
A <sub>1</sub> A <sub>1</sub>	1.000	1.024	1.021
A <sub>1</sub> A <sub>2</sub>	1.025	1.066	1.026
A <sub>2</sub> A <sub>2</sub>	1.018	1.019	1.007

initial gametic frequencies

A <sub>1</sub> B <sub>1</sub>	A <sub>1</sub> B <sub>2</sub>	A <sub>2</sub> B <sub>1</sub>	A <sub>2</sub> B <sub>2</sub>
0.168	0.362	0.292	0.178

example taken from Ewens (1979), p. 58



## Limitations of the optimization approach

- ❖ The postulate on which the approach is based, that fitness is maximized in the short term dynamics driven by natural selection, is not generally valid
- ❖ And, quite apart from the maximization issue, in certain cases the very choice of the appropriate fitness measure is questionable and has led to discussions and divergences of opinions. Examples of the biological complications where this uncertainty has arisen are: age structure; density dependence; chaotic population dynamics; kinship among interacting individuals
- ❖ Thus, if the limits of the “short term dynamics” approach are consequences of excessive attention to details of dynamics, the limits of the optimization approach are due to the opposite extreme of neglecting altogether this aspect of the problem
- ❖ A careful consideration of the different time scales of evolution, and the resulting proposal to formulate the evolutionary process on a time scale larger than that which is characteristic of short term dynamics, has opened the possibility to overcome most of the difficulties intrinsic to the two classical approaches discussed so far
- ❖ Shifting the analysis to a larger time scale permits to: (i) temperate considerably in many cases the difficulties due to complex dynamics; (ii) eliminate altogether the burden of *guessing* the “right” form of the fitness measure; (iii) recover in a new clothing rigorous principles of optimality

# Time scales of evolution

## Natural selection

- ❖ The time unit of natural selection processes is generation length
- ❖ At any given moment, selection operates on a restricted set of genetic variants, much inferior to the total variation that is potentially possible for that organism
- ❖ Due to differences in some demographic parameters, resulting in different reproduction rates, selection produces changes of the relative abundances of the various genotypes in the population
- ❖ The process generally terminates with the attainment of an equilibrium composition of the population, which often is achieved, or almost so, in a moderate number of generations (~50–500)
- ❖ The total change in the physical traits of the species (*phenotype*) effected by selection in this time span is generally minute

## Mutation

- ❖ Fresh genetic variation is introduced in the population by mutation. It is only through this force that a sustained walk in the phenotypic space can be performed
- ❖ Mutation is a rare phenomenon: typical rates being of the order of  $10^{-6}$  per gene, per individual, per generation. Thus, for example, if a trait is controlled by 10 genes, in a population of 10,000 individuals, after 100 generations only about 10 new genes affecting that trait would have appeared (carried initially by as many individuals); but almost all of these would be harmful and eliminated by natural selection, and the very few that are not so, have a chance of being swamped by pure random events
- ❖ Therefore, it is reasonable to assume that significant mutations appear in a population so rarely that any previous selection episode has already achieved a state of equilibrium
- ❖ Whenever a successful mutation appears in the population, carried initially by a very small minority of individuals, a new selection episode is set in motion till a new equilibrium, that incorporates somehow the genetic novelty, is attained
  
- ❖ In this perspective, it is clear that there are two distinct time scales. In the short time scale of demographic processes, with a time unit of generation length, the driving force is natural selection which filters mutations and moves the population to a new (transient) equilibrium which establishes the prevalence of favourable mutants
- ❖ In the long time scale the driving force is mutation which introduces new genetic variants. The time unit of this scale corresponds to the time span between successive appearances of favourable mutation, generally at least as long as the time necessary to attain a transient equilibrium of the short term selection process



# General properties of long-term evolution

- ❖ Long-term evolution is the evolutionary process as *perceived* in the long time scale
- ❖ A succession of transient population states that, in the short time scale would correspond to states of *equilibrium* of natural selection dynamics
- ❖ Transition from one state to the next is caused by the occurrence of a successful mutation that *invades* (is established in) the population
- ❖ It follows that this process has an intrinsically stochastic nature: both the interval of time between transitions and the size of change in the evolving traits are random variables. The direction of change (i.e., which mutations can invade and which don't) and the population state attained eventually (new frequencies of genotypes) instead are not random, being determined by natural selection

## Long term equilibria (LTE)

- ❖ A population state (which necessarily is an equilibrium of the short term process) that cannot be invaded by *any* mutation, because *all possible mutants* of the evolving trait (or traits) are eliminated by natural selection while still rare in the population
- ❖ Biologically, a long term equilibrium has the same meaning as an evolutionarily stable strategy, but mathematically the two notions are very different:
  - For ESS* non-invasibility is synonymous with superiority of fitness with respect to rare alternatives, on the base of a preassigned fitness measure without reference to any underlying dynamics
  - For LTE* non-invasibility must be verified directly from an explicitly formulated short term dynamics, so that it does not depend on any optimality principle or assumption of maximization of a preassigned fitness function