

# Phenotypic plasticity can potentiate rapid evolutionary change

Narayan Behera<sup>a,1</sup>, Vidyanand Nanjundiah<sup>a,b,\*</sup>

<sup>a</sup>Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560004, India

<sup>b</sup>Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560012, India

Received 9 June 2003; accepted 19 August 2003

## Abstract

Using a computational model of string-like haploid genotypes, we verify the conjecture (J. Theor. Biol. 188 (1997) 153) that phenotypic plasticity can speed up evolution. The corresponding real-life situation was realized by Waddington in experiments carried out on the fruit fly *Drosophila*. Waddington found that after selecting for an environmentally induced trait over a number of generations, a new, true-breeding phenotype resulted that was absent in the starting population. The phenomenon, termed ‘genetic assimilation’, continues to attract interest because of the rapidity of the effect and because of its seemingly Lamarckian implications. By making use of a genetic algorithm-based approach developed previously, we show that conventional Darwinian selection acting on regulatory genes can account for genetic assimilation. The essential assumption in our model is that a structural gene can be in either of three allelic states. These correspond to its being (a) ‘on’ or ‘off’ constitutively or (b) in a plastic state in which the probability that it is ‘on’ or ‘off’ is influenced by regulatory loci in a dosage-dependent manner.

© 2003 Elsevier Ltd. All rights reserved.

**Keywords:** Canalization; Genetic assimilation; Phenotypic plasticity

## 1. Introduction

The conventional picture of evolution by natural selection is built on two foundations (Fisher, 1958). One, that the raw material of evolution, heritable variation, is generated spontaneously, that is, without any reference to environmental conditions; acquired traits cannot be inherited. Two, that evolution is an inherently slow process: the possibility of rapid evolutionary change is said to be foreclosed by the fact that evolution consists of one well-adapted state going over to another. In a series of laboratory experiments on the fruit fly *Drosophila*, Waddington (1956, 1961) obtained results that seemed to question both these tenets of evolutionary theory. Firstly, he showed that it was possible for major changes in the phenotype to occur in a small number of generations—about 15—; the new phenotypes bred true. Secondly, he managed to select

for novel phenotypes that appeared in a wild-type population following exposure to a stressful environment. In other words, he showed that an acquired trait (that also happened to result in a major change in the phenotype) could eventually become inherited. On top of the rapidity with which it was elicited, the seemingly Lamarckian implications of the outcome continue to draw attention from evolutionary and developmental biologists (see Gilbert, 2000).

A typical experiment, sketched here to illustrate the principle rather than describe details, goes as follows. To begin with, independent lines of brother–sister matings are set up between true-breeding wild-type flies. ‘Wild-type’ means that they (a) possess a set of standard phenotypic traits that is so designated (because it corresponds to the norm in the wild) and that when tested in a standard—ideally natural—environment, (b) transmit the traits reliably to their offspring. The progeny from each single-pair mating are divided into two groups. One group (‘Test’) is subjected to a brief, non-mutagenic environmental shock at a definite stage of development (for example, as pupae). The shock might be exposure to ether or, in some experiments, elevated temperature. The other (‘Control’) group is

\*Corresponding author. Fax: +91-80-3600683.

E-mail address: [vidya@ces.iisc.ernet.in](mailto:vidya@ces.iisc.ernet.in) (V. Nanjundiah).

<sup>1</sup>Present address: Institute of Bioinformatics and Applied Biotechnology, G 05, Tech Park Mall, International Tech Park, Whitefield Road, Bangalore 560 066, India.

allowed to develop undisturbed. Among those individuals that survive the shock and become adults, a few turn out to exhibit grossly altered morphologies; and in cases, their phenotypes resemble those of known mutant genotypes. Mutants of the *bithorax* series, in which part or all of the haltere can develop as wing, are an example: flies with the extreme or Ultrabithorax phenotype have four wings rather than the normal complement of two wings and two halteres. Because the phenocopies tend to be of impaired vitality, the next generation is best generated by sib-selection: by breeding, not from the phenocopies themselves, but from single-pair lines generated from their ‘Control’ brothers and sisters. One set among the offspring of each pair of parents is once again subjected to the shock: sibs of those that display a strong phenocopy are chosen to act as parents for the third generation, and so on. After as few as about 15 generations, flies with the desired phenotype start appearing among ‘Control’ individuals and breed true for the *new* phenotype. What is striking is that neither the parents nor any other ancestors of these individuals had ever experienced the shock. The phenomenon, termed ‘Genetic Assimilation’ by Waddington, has been re-discovered on many occasions (Bateman, 1959; Ho et al., 1983; Gibson and Hogness, 1996).

Evidently, the population has undergone a heritable, therefore genetic, change, but how? A possible explanation was offered by Waddington himself (1961). The crux of the explanation is that the starting population is genetically heterogeneous. Normally, the heterogeneity is masked on account of selection in the past having favoured the development of the phenotype that is best adapted to the natural environment, the ‘wild-type’ phenotype. The masking of genetic heterogeneity implies that development is normally strongly buffered, or, as Waddington put it, canalized. Canalization works via well-adapted combinations of regulatory genes, the adaptation being to the range of environments usually encountered. When an environmental shock is applied, this range is exceeded. Gene expression gets disrupted and canalization fails. The outcome—in so far as development takes place at all—is disruptive and there is an increase in phenotypic variance. However, the appearance of many phenotypes permits selection to be carried out from one generation to the next. In the beginning, the phenotype that is selected for must be elicited anew in each generation with the help of the same environmental shock as at the beginning. But the genotypes which give rise to this phenotype are not the same in every generation. New combinations of regulatory genes arise at every meiosis; some of these combinations lead to a stronger response to the shock than others, or have a lower threshold for responding in the desired fashion, or both. Therefore, while selecting for a particular phenotype, one is also—albeit indir-

ectly—selecting for those genotypes that favour the phenotype. In other words one is selecting for canalization of the *new* phenotype. If the response to the environmental shock is of an all-or-none character, a certain combination of regulatory genes could even lead to a constitutive crossing of the threshold for response. The result would be a ‘response’ without the presence of the stimulus: the assimilated phenotype would breed true.

On the face of it, Waddington’s proposal is plausible. Should it work, it would seem to offer important clues in thinking about phenotypic modifications of the sort that have characterized major evolutionary transitions. In spite of this, it has never been examined within the framework of an explicitly genetic model. The lines of a possible approach are implicit in Hinton and Nowlan’s (1987) pioneering investigation of how learning might guide evolution. However, to the best of our knowledge, the approach was not followed through explicitly. We (Behera and Nanjundiah, 1995, 1996, 1997) have investigated a series of computational models that clarify the strengths and weaknesses of the Hinton–Nowlan approach. In doing so, we have shown that phenotypic plasticity, combined with gene regulation, can promote rapid evolution. We proceed to demonstrate that a version of these models can also explain genetic assimilation. The essence of the explanation is that the phenotype is capable of being tuned, as it were, via the action of regulatory genes that influence the inherently plastic state of activity of structural genes; the strength of this influence—the strength of regulation—is sensitive to the environment.

## 2. Model

The model is based on a computational algorithm for studying the role of phenotypic plasticity in evolution (Hinton and Nowlan, 1987; Behera and Nanjundiah, 1995, 1996, 1997). Genotypes are haploid and consist of randomly generated arrays of two strings of genetic loci. One is a ‘structural string’ (which determines the phenotype) and the other is a ‘regulatory string’ (which determines the functioning of the structural string). Each structural locus has possible alleles 1, 0 and X (frequencies  $p_{1s}$ ,  $p_{0s}$ ,  $p_{xs}$ ) which can be thought of as giving rise to ‘on’, ‘off’ and ‘either on or off’ states of the gene, respectively. Regulatory genes (possible alleles 1, 0; frequencies  $p_{1r}$ ,  $p_{0r}$ ) specify the probability that an X at a structural locus functions as a 1 or a 0; thus they influence the phenotype only indirectly (the two sets of 1’s and 0’s are not the same of course). To begin with, whatever their genotype, all individuals possess the same wild-type phenotype. The desired, or target, phenotype is defined by a string of 1’s at the structural loci.

In environment A, individuals of the organism of interest express the wild-type phenotype. In every generation some of them are made to develop in environment B; the rest remain in A. The nature of B is such that it makes them prone to express the altered, desired phenotype. The sibs of those that come closest to expressing the altered phenotype after exposure to B are selected for mating and constitute the parents of the next generation of individuals. The aim is to see whether over a number of generations the altered phenotype is expressed in those members that are retained in A—that is, is expressed by those belonging to a lineage that has never been exposed to B. In order to compare our model with Waddington's experiments, A can be thought of as permitting development under conditions similar to that in the wild and B, as corresponding to the condition of a brief ether shock at a sensitive stage. The desired phenotype can be thought of as the full ultrabithorax phenotype.

A detailed description of the model follows. The parameter values given below are the ones used to generate the results displayed in Figs. 2a–c. These values have been chosen in order to make the point that Waddington's observations can be mimicked successfully. As will be pointed out in the Discussion, the output of the model is quite robust with respect to variations in the parameters. A genotype is represented by two chromosomes idealized as strings of length  $N_s$  ( $s$ =structural) and  $N_r$  ( $r$ =regulatory), respectively ( $N_s = 14$  and  $N_r = 12$ ). Ninety genotypes are specified at the start by making use of a random number generator to decide whether, given certain a priori probabilities, an allele at a particular locus is a 0, a 1 or (only in the case of a structural locus) an X. The starting probabilities are denoted  $p_{0s}$ ,  $p_{1s}$  and  $p_{xs}$  for structural loci and  $p_{0r}$  and  $p_{1r}$  for regulatory loci, respectively. At later times, the same symbols stand for allele frequencies: the fraction of an individual's structural loci that are 0, the fraction that are 1, and so on. The 1 allele at a structural locus can be imagined to mean that the locus is constitutively active or 'on', and a 0 to mean that the locus is constitutively inactive or 'off'. An X on the other hand means that the status of the locus is facultative; it can be either 'on' or 'off'. The choice is made independently and on a probabilistic basis, locus-by-locus. Each X allele tosses a coin, so to speak, and a set of independent coin tosses makes each X in a structural gene set function in the manner of a 1 or a 0. We refer to this as an  $X \rightarrow 1$  or  $X \rightarrow 0$  'switch', it being understood that it is not the gene that is changing but merely its functional status within that generation. The total number of 'coin-tossing' attempts permitted is  $2^m$ , where  $m = N_s p_{xs}$ .

The central point of our model is that the probability of a  $X \rightarrow 1$  switch depends on the fraction of 1's in the regulatory loci. The essential difference between the two

environments lies in these probabilities, the probability, that a X behaves like a 1 is significantly smaller in environment A than in B. The probabilities are taken to be  $p_{1r}^3$  in environment A and  $p_{1r}$  in environment B. Once the functional status of all the X's in a genotype is determined in this manner for all the individuals in a generation, it becomes possible to assign a fitness to every genotype, or rather to the phenotype manifested by every genotype. Fitness is computed as the weighted sum of two components (Behera and Nanjundiah, 1995) which take into account both the energetic cost of attaining the final phenotype and the degree to which the final phenotype matches the desired phenotype. The desired phenotype is taken to be the one corresponding to a genotype made up of a string of 1's alone. This phenotype can be present at the start if the starting genotype is a string of 1's. If not, it can be thought of as a target to be attained by a random search (via 'coin-tossing', as explained below). The first component of fitness reflects the extent of overlap between the actual phenotype and the desired phenotype. It is defined as

$$W_D = \frac{1}{N_s} \sum_{i=1}^{N_s} d_i, \quad (1)$$

where  $i$  stands for the  $i$ -th genetic locus and  $d_i = 1$  if the  $i$ -th locus has a 1 allele and  $d_i = 0$  otherwise. The other component of fitness highlights the plastic aspect of the phenotype. It depends on the fraction of the total number of 'coin-tossing' trials left unused and, in effect, favours genotypes that require only a few coin-tossing trials over those that require many. This component of fitness is defined as

$$W_P = n/(2^{N_{PX}}), \quad (2)$$

where  $n$  is the number of trials that remain after the genotype has attained the target and  $N_{PX}$  is the expected number of X alleles in a genotype.  $n$  is zero if all the  $2^{N_{PX}}$  trials are gone through (whether or not the target has been attained). One might say that  $W_P$  represents the physiological cost of coin-tossing. The total fitness function  $W_T$  is calculated as the average of  $W_P$  and  $W_D$  as

$$W_T = \frac{1}{2}(W_P + W_D). \quad (3)$$

We have previously explored the consequences of weighting the relative contributions of  $W_P$  and  $W_D$  differently (Behera and Nanjundiah, 1995). Fig. 1 represents the mating scheme in the form of a flow-chart. Mating takes place between 45 pairs of individuals chosen at random from the 90 starting haploid genotypes. The haploid genotypes that constitute the next generation are produced via meiosis. Following a mating between a pair of haploid parental genotypes, five independent meiotic events lead to the appearance of 20 haploid progeny genotypes. Meiosis is accompanied by recombination involving independent crossovers

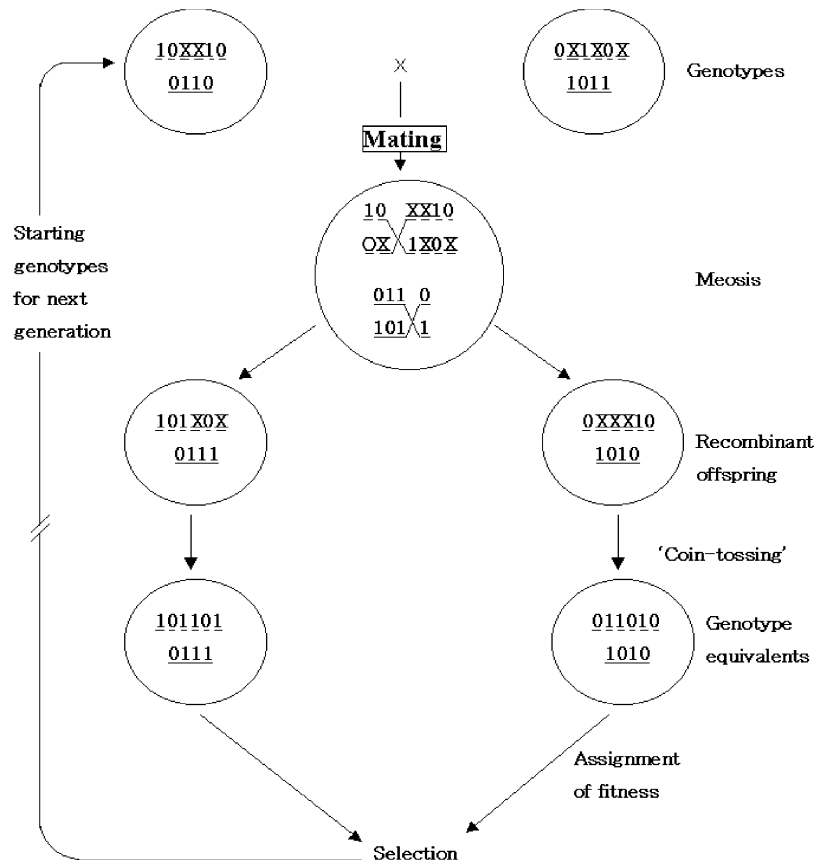


Fig. 1. A sketch of the mating scheme (based on Behera and Nanjundiah, 1997). A genotype is represented as two haploid chromosomes made up of one string of structural (---) loci and one string of regulatory loci (—). Meiosis is accompanied by a single obligatory crossover event in each chromosome.

in which the crossover points are chosen at random; one crossover invariably takes place within the structural loci and the other one within the regulatory loci. Out of the possible four haploid recombinant offspring that result from a single meiosis, we pick just two in the following manner. The genotype of the first offspring is chosen by copying all alleles from the (arbitrarily designated) first parent starting from the left up to either crossover point, and from the second parent beyond the crossover point. The genotype of the second offspring consists of all alleles from the second parent to the left of either crossover point and from the first parent beyond the crossover point. The two parental genotypes are also included, so that even though the haploid chromosome number is two, all four meiotic products are represented equally in the offspring. Since five independent crossover events occur, a single mating pair gives rise to 20 progeny and 900 progeny result from 45 distinct single-pair matings.

Each set of 20 progeny of a single sib-pair is further partitioned into halves of 10 each. One set of 10 is designated C and the remaining set T; they correspond to the ‘Control’ and ‘Test’ groups, respectively (see Introduction). After subjecting the members of each test group to the stressful environment (i.e., B) and assigning

a phenotype to each genotype as explained above, the fitness, averaged over all members, is computed for each group belonging to the T class. The group with the highest average fitness—again, as measured in the stressful environment B—is noted and the next generation is raised by carrying out random matings between the 10 members of the C group corresponding to that T group, which means between the siblings of the T group with the highest average fitness. The number of possible matings is  $^{10}C_2$  or 45; each mating yields two recombinant genotypes as already explained. The 90 progeny become the starting genotypes for the next cycle and define the population whose progress we follow. Thus a cycle consists of monitoring members of the T group following their exposure to environment B followed by mating between members of the selected C group and monitoring their fitness in environment A. No member of a C group is ever exposed to environment B. The question is, does the fitness of the members of successive C groups, as determined by their phenotype in the normal environment A, keep increasing over the course of generations, and if so, in what manner? In other words, does genetic assimilation take place?

We monitor the efficacy of genetic assimilation by evaluating two quantities. One,  $W$ , is the mean

population fitness as measured in environment A ( $W_A$ ) or in environment B ( $W_B$ ). The other ( $H$ ) is the fraction of the population that expresses the desired phenotype, either after exposure to environment B ( $H_B$ ) or without any such exposure ( $H_A$ ).  $H$  is defined as follows. It is the fraction of the population which consists of individuals who (i) have only 1's or X's in their structural loci and (ii) following coin tossing, acquire the phenotype corresponding to that of the genotype consisting of a string of 1's. Thus  $H_A$  is a measure the efficacy of genetic assimilation. The difference between  $W$  and  $H$  should be noted. In terms of Waddington's experiment,  $H_A$  is the fraction of the population that expresses the Ultrabithorax phenotype without any prior exposure to ether and  $H_B$  is the fraction that does so after ether exposure.

### 3. Results

As is only to be expected given the small population size, eventually fixation is reached at all loci. The approach to fixation is best described in terms of the allele frequencies at structural ( $p_{0s}$ ,  $p_{1s}$ ,  $p_{xs}$ ) and regulatory ( $p_{0r}$ ,  $p_{1r}$ ) loci. In general,  $p_{0s}$  decreases steadily,  $p_{xs}$  remains more or less steady while  $p_{1s}$  increases with generation number (data not shown). Thus, selection acts against the 0 allele and in favour of the 1 allele in the structural loci, all as expected. But, what is at first sight unexpected,  $p_{1r}$ , the frequency of the 1 allele in the regulatory loci, increases although there is no direct selection acting on it:  $p_{1r}$  goes up by 'hitch-hiking' on the structural genes (Behera and Nanjundiah, 1997).

Both fitness estimates,  $W_A$  and  $W_B$ , increase over the course of generations (Fig. 2a), and so do  $H_A$  and  $H_B$  (Fig. 2b), the relative fractions of the population that display the desired phenotype in environments A and B. The relatively higher starting value of  $W_A$  in Fig. 2a is because the frequency of 0 alleles in the initial population is very low. Correspondingly,  $H_A$  too is low to begin with (Fig. 2b), meaning that hardly any individual exhibits the desired novel phenotype: in the beginning the population is canalized for the wild-type phenotype. Both  $H_A$  and  $H_B$  increase from one generation to the next (albeit with ups and downs), with  $H_B$  always being higher than  $H_A$ . In other words, a larger fraction of the population expresses the desired phenotype in the stressful environment than in the normal environment. This too is expected because of the manner in which regulatory genes affect the phenotype of a  $X \rightarrow 1$  transition at structural loci in the two environments. However, already in the first generation a few individuals exhibit the desired phenotype upon exposure to B whereas none do so in the control population ( $H_A = 0.012$  and  $H_B = 0.076$  initially and

the population size is 90). Genetic assimilation can be observed in a small fraction of the population by the third generation (Fig. 2b). By generation 15, assimilation is essentially complete, with close to 90% of the control population exhibiting the altered phenotype in the normal environment. One might say that the flies breed true for the 'assimilated' phenotype. For the value of  $H$  at fixation, i.e. the efficacy of genetic assimilation, to be non-zero, the initial value of  $p_{xs}$ —in effect, the degree of plasticity that the genome can exhibit—has to be beyond a threshold (Fig. 2c). Similarly, we find that a minimum value of  $p_{1r}$  is needed at the start for assimilation to be effective, this value is 0.2 under the conditions in Fig. 2c.

### 4. Discussion

We begin by drawing attention to a basic difference between the present model and that of Hinton and Nowlan (1987). What Hinton and Nowlan postulated was a zone of increased fitness around the optimal genotype and a relatively flat fitness function elsewhere. That enabled genotypes that were 'almost there', so to speak, to persist until favourable mutations drove them even closer to attaining peak fitness. Hinton and Nowlan (1987) required a very large number of generations in order for an appreciable degree of genetic change to occur in the population (Behera and Nanjundiah, 1995). A reason for this could be that the model lacked regulatory genes that could sense the environment and so speed up phenotypic change. In fact, because of the all-or-nothing dependence of fitness on genotype, what Hinton and Nowlan actually put forward was a model for the Baldwin effect—a phenomenon rather different from genetic assimilation (Nanjundiah, 2003). In the Baldwin effect, a facultative phenotypic response to a specific environment can become constitutive if the right mutational changes take place (Simpson, 1953). The Baldwin effect does not require that the population be genetically heterogeneous whereas genetic assimilation does.

The choice of parameters made by us and the robustness of the outcome deserve comment. Our reason for the specific choice of  $p_{1r}$  and  $p_{1r}^3$ , respectively, for the  $X \rightarrow 1$  transition probability in the two environments was because (a) this resulted in no genetic assimilation without exposure to the test environment and (b) following exposure and sib selection, the entire population exhibits genetic assimilation after about 14 generations, i.e. as rapidly as in Waddington's experiments. If we take  $p_{1r}^2$  instead of  $p_{1r}^3$ , full assimilation occurs at about the sixth generation, i.e. too rapidly, while with  $p_{1r}^4$  it takes more than 25 generations, i.e. too slowly, besides which the fraction of the population with the assimilated phenotype is very low. Very likely other



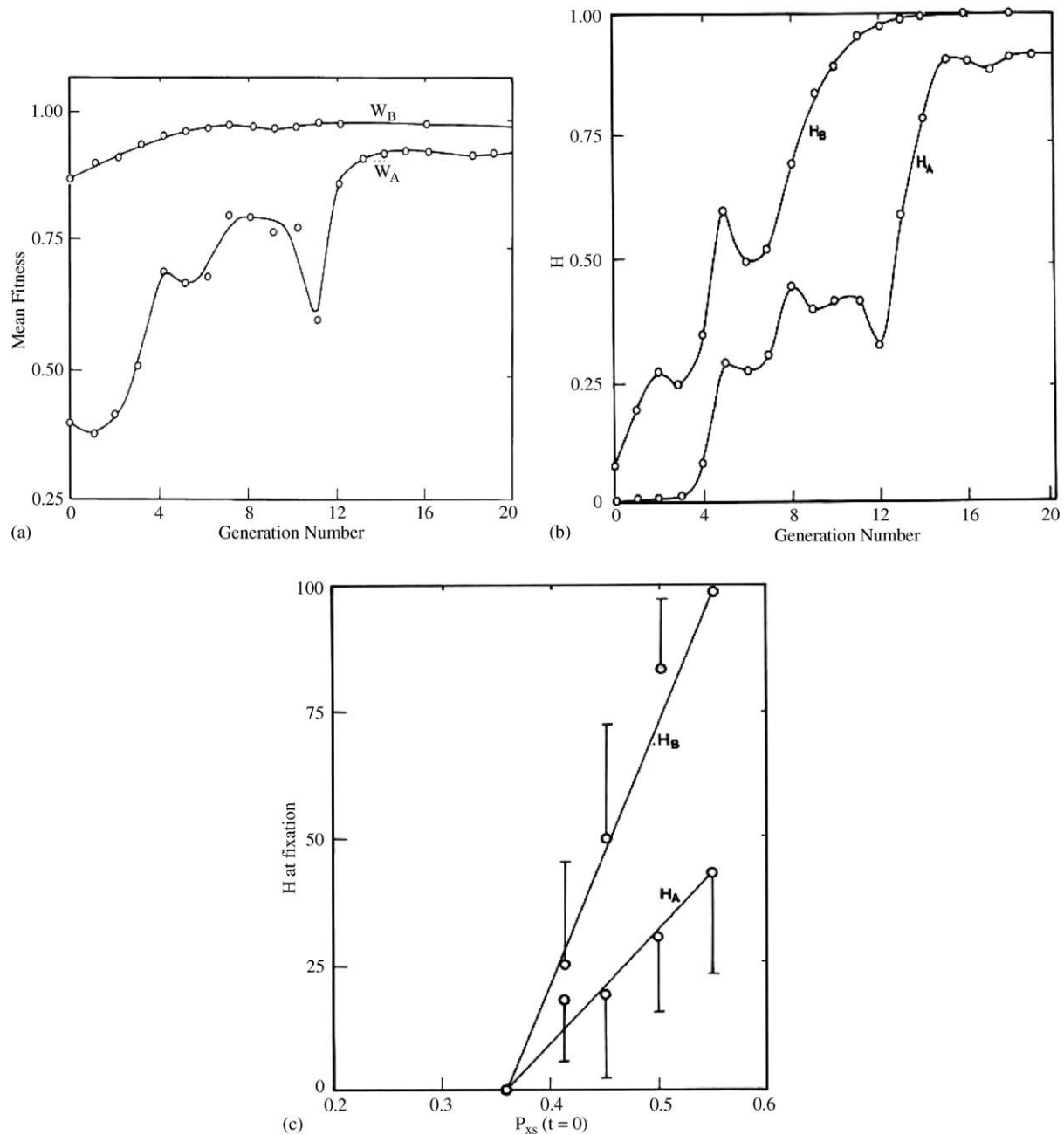


Fig. 2. (a) Changes in mean population fitness ( $W$ ) as a function of time measured in numbers of generations; output from a representative simulation. Initial conditions:  $N_s = 14$ ,  $N_r = 12$ ,  $p_{1r}(t=0) = 0.5$ ,  $p_{0s}(t=0) = 0.08$  and  $p_{xs}(t=0) = 0.43$ . (b) The fraction of the total population that exhibits the assimilated phenotype ( $H$ ) as a function of time (representative simulation); initial conditions as before. (c)  $H$  as a function of the initial degree of plasticity in the genotype ( $p_{xs}(t=0)$ ). Means of 12 simulations with standard deviations. Initial conditions:  $N_s = 22$ ,  $N_r = 12$ ,  $p_{1r}(t=0) = 0.5$ ,  $p_{0s}(t=0) = 0.1$ . In all three figures,  $W$  and  $H$  prior to fixation are values that pertain to the sib-group that has the highest fitness. The suffixes A and B refer to the control and test environments, respectively.

combinations of probabilities may lead to the same outcome; our purpose was merely to show that one combination would. The mating scheme used by us includes both parental and recombinant genotypes in the offspring generation. The aim is to introduce a degree of realism in that this is what one would expect in a real-life situation: generations would overlap. Note that this leads to some degree of slurring over the

difference between the test and control groups and means that in our model genetic assimilation occurs more slowly than it would otherwise. Genetic algorithm models used for DNA sequence alignment have used both parental and recombinant genotypes to define the post-mating generation. (Notredame et al., 1997).

It takes approximately the same number of generations to achieve the end result in our model as in the

Table 1  
Dependence of  $G$ , the number of generations required for fixation, on structural and regulatory loci

$N_s$	$H_B$	$H_A$	$W_B$	$W_A$	$G$
(a) $N_r = 12, p_{1r}(t = 0) = 0.5; p_{xs} = 0.41; p_{0s} = 0.09$					
14	100 (0)	87.15 (14.64)	0.99 (0)	0.93 (0.06)	8.08 (2.82)
22	83.30 (37.25)	33.63 (42.13)	0.98 (0.01)	0.63 (0.28)	8.38 (1.84)
26	66.67 (47.14)	24.20 (36.86)	0.95 (0.01)	0.62 (0.26)	8.52 (2.53)
$N_r$	$H_B$	$H_A$	$G$		
(b) $N_s = 14, p_{1r}(t = 0) = 0.5; p_{xs} = 0.41; p_{0s} = 0.09$					
1	100 (0)	100 (0)	8.33 (4.44)		
7	99.98 (0.06)	56.57 (37.65)	8.16 (2.67)		
50	81.87 (36.73)	25.61 (37.33)	8.22 (1.58)		
100	54.98 (46.63)	0.78 (1.51)	8.42 (3.45)		

A and B refer to the control and test environments, respectively.  $H_A, H_B$ , percentage of assimilation;  $W_A, W_B$ , mean population fitness. Other symbols as in the text. The values shown are good to two places of decimals and represent means with standard deviations in parentheses.

experiments (Waddington, 1956,1961; Bateman, 1959; Ho et al., 1983; Gibson and Hogness, 1996). This merely reflects the fact that parameter values have been chosen for this purpose, among them the numbers of structural and regulatory loci ( $N_s = 14$  and  $N_r = 12$  in Figs. 2a–c).  $N_s$  and  $N_r$  should be thought of as representing the effective size of the genome—the number of genes capable of affecting the trait in question—not the number of genes in the genome as a whole. The total population size in our model, 900, is of the same order as that in the most recent experiments (1500 in the experiments of Gibson and Hogness, 1996). If  $N_s$  or  $N_r$  is made much smaller, fixation occurs too rapidly for significant assimilation to take place. In general, the number of generations required for fixation to be reached is weakly dependent on  $N_s$  or  $N_r$  but the effectiveness of assimilation is, as measured either by the final fitnesses,  $W_A, W_B$ , or by the fraction of the population that attains the desired phenotype,  $H_A, H_B$  (Table 1a,b). While the mean fitness at fixation and the fraction of the population that exhibits genetic assimilation both decrease as a function of  $N_s$  or  $N_r$ , the number of generations required for fixation,  $G$ , increases with them (Table 2). Further,  $G$  is positively correlated with the initial fraction of plastic alleles in the structural loci,  $p_{xs}(t = 0)$ , and negatively correlated with initial value of  $p_{1r}$ ;  $p_{xs}$  itself increases as evolution proceeds (Behera and Nanjundiah, 1997).

We have assumed for computational convenience that the regulatory genes work in *trans*. The regulation of gene expression by *trans*-acting factors is an ubiquitous phenomenon (Watson et al., 2003); the factors are generally DNA binding proteins that bind to *cis*-acting sequences to control gene expression. The influence of *cis*-regulatory DNA sequences on genetic assimilation is an important question that remains to be explored. Our model is able to account for the broad features of

Table 2

Correlation between the outcome—as reflected in the extent of genetic assimilation,  $H$ , the fitness,  $W$ , or the number of generations required for fixation,  $G$ , in the control (A) and test (B) environments—and the number of structural ( $N_s$ ) and regulatory ( $N_r$ ) loci in the genome

Variables	Mean correlation coefficient	s.d.	95% confidence limits
$H_B$ vs. $N_s$	−0.222	0.037	−0.243, −0.201
$H_A$ vs. $N_s$	−0.264	0.031	−0.281, −0.247
$H_B$ vs. $N_r$	−0.651	0.020	−0.662, −0.640
$H_A$ vs. $N_r$	−0.818	0.015	−0.826, −0.810
$W_B$ vs. $N_s$	−0.185	0.040	−0.564, −0.604
$W_A$ vs. $N_s$	−0.168	0.042	−0.192, −0.144
$W_B$ vs. $N_r$	−0.384	0.031	−0.401, −0.367
$W_A$ vs. $N_r$	−0.846	0.018	−0.856, −0.836
$G$ vs. $N_s$	0.214	0.041	0.191, 0.237
$G$ vs. $N_r$	0.121	0.028	0.105, 0.137

genetic assimilation: after a small number of generations of purely phenotypic selection, an environmentally-induced phenotype can be expressed constitutively in the absence of the inducing environment. The requirements for this are that (a) a genotype should be able to give rise to more than one phenotype, whether or not in the same environment (in our model, this is achieved by the existence of ‘plastic’ or X alleles in the structural loci), (b) the expression of structural genes should be modifiable by suitable combinations of regulatory genes (achieved by starting with a non-zero value of  $p_{1r}$ ) and (c) genetic variation must exist in the initial population (achieved by choosing initial genotypes with the help of a random number generator). (b) and (c) need no further comment; and evidence is beginning to build up in favour of (a) (Jablonka and Lamb, 1995; Bissell et al., 1999; John and Surani, 1999; Newman and Müller, 2000; Nanjundiah, 2003). West-Eberhard (1986) has advanced a scheme in which the same genotype can give rise to distinct but equally well-adapted phenotypes in

different individuals belonging to the same population and, over generations, the alternative phenotypes would get sharpened (perhaps because they were adapted to different sub-niches). Eventually, circumstances that favour just one of the alternatives could result in what looks like rapid evolutionary change, even rapid speciation.

Requirements (b) and (c) permit the variation to remain cryptic because the normal, wild-type phenotype is strongly canalized in environment A. Different structural gene combinations result in the same phenotype because of the canalizing action of regulatory genes. Canalization breaks down when an environmental shock is applied during development. The breakdown is manifested in the fact that the probability of a  $X \rightarrow 1$  transition is  $p_{1r}$  in environment B as against  $p_{1r}^3$  in environment A. In this new environment, B, the preferred phenotype corresponds to a genotype with a 1 allele in each structural locus. Selection in favour of the new phenotype indirectly implies selection in favour of novel regulatory gene combinations that now cause the new phenotype to be canalized (also, therefore, to be assimilated). Because the fitness function chosen by us varies in a more or less continuous fashion with the genotype, selection in favour of the new phenotype also occurs quasi-continuously. However, the appearance of the fully assimilated phenotype is discontinuous because it depends on the appearance of a genotype made up of 1's alone (or of an equivalent string of 1's and X's). Plasticity, meaning the existence of structural locus alleles that can function either as 1's or as 0's, is essential for assimilation to occur in our model.

What are the implications of these simulations for understanding how evolution might work? It has been pointed out that in comparison with the wild type, the assimilated phenotype in any of Waddington's experiments—say the Ultrabithorax phenotype—is poorly adapted to the natural environment. Therefore, the morphological transformation involved could in no sense represent an evolutionary advantage under natural conditions. Nevertheless, the point remains that a drastic change has occurred. This leaves open the question whether environmental stress could lead to the appearance of a novel but potentially *adaptive* phenotype. A 'hopeful monster' would be generated, not as the result of a genetic mutation, but epigenetically (Newman and Müller, 2000). If the probability is not too low, our model provides a means for the rapid transition of the population from the old phenotype to the new.

## Acknowledgements

The research presented here was partially supported by a grant from the Department of Biotechnology.

## References

- Bateman, K.G., 1959. The genetic assimilation of the *dumpy* phenocopy. *J. Genet.* 56, 341–351.
- Behera, N., Nanjundiah, V., 1995. An investigation into the role of phenotypic plasticity in evolution. *J. Theor. Biol.* 172, 225–234.
- Behera, N., Nanjundiah, V., 1996. The consequences of phenotypic plasticity in cyclically varying environments: a genetic algorithm study. *J. Theor. Biol.* 178, 135–144.
- Behera, N., Nanjundiah, V., 1997. Trans-gene regulation in adaptive evolution: a genetic algorithm model. *J. Theor. Biol.* 188, 153–162.
- Bissell, M.J., Weaver, V.M., Lelièvre, S.A., Wang, F., Petersen, O.W., Schmeichel, K.L., 1999. Tissue structure, nuclear organization and gene expression in normal and malignant breast. *Cancer Res. (Suppl.)* 59, S1757–S1764.
- Fisher, R.A., 1958. *The Genetical Theory of Natural Selection*, 2nd Edition. Dover, New York.
- Gibson, G., Hogness, D.S., 1996. Effect of polymorphism in the *Drosophila* regulatory gene *ultrabithorax* of homeotic stability. *Science* 271, 200–203.
- Gilbert, S.F., 2000. *Developmental Biology*, 6th Edition. Sinauer Associates, Inc., Sunderland, USA.
- Hinton, G.E., Nowlan, S.J., 1987. How learning can guide evolution. *Complex Syst.* 1, 497–502.
- Ho, M.W., Tucker, C., Keeley, D., Saunders, P.T., 1983. Effects of successive generations of ether treatment on penetrance and expression of the *bithorax* phenocopy in *Drosophila melanogaster*. *J. Exp. Zool.* 225, 357–368.
- Jablonka, E., Lamb, M.J., 1995. *Epigenetic Inheritance and Evolution*. Oxford University Press, Oxford.
- John, R.M., Surani, M.A., 1999. Agouti germ line gets acquisitive. *Nature Gen.* 23, 254–256.
- Nanjundiah, V., 2003. Phenotypic plasticity and evolution by genetic assimilation. In: Newman, S.A., Müller, G. (Eds.), *Origins of Organismal Form*. MIT Press, Cambridge.
- Newman, S.A., Müller, G.B., 2000. Epigenetic mechanisms of character organization. In: Wagner, G.P. (Ed.), *The Character Concept in Evolutionary Biology*. Academic Press, San Diego.
- Notredame, C., O'Brien, E.A., Higgins, D.G., 1997. RAGA: RNA sequence alignment by genetic algorithm. *Nucleic Acid Res.* 25 (22), 4570–4580.
- Simpson, G.G., 1953. The Baldwin Effect. *Evolution* 7, 110–117.
- Waddington, C.H., 1956. Genetic assimilation of bithorax phenotype. *Evolution* 10, 1–13.
- Waddington, C.H., 1961. Genetic Assimilation. *Adv. Gen.* 10, 257–293.
- Watson, J.D., Hopkins, N.H., Roberts, J.W., Steitz, J.A., Weiner, A.M., 2003. *Molecular Biology of the Gene*, 5th Edition. Benjamin Cummings, USA.
- West-Eberhard, M.J., 1986. Alternative adaptations, speciation and phylogeny. *Proc. Natl Acad. Sci. USA* 83, 1388–1392.