Nonlinearity **9** (1996) 607–629. Printed in the UK

Can nonlinear epigenetic interactions obscure causal relations between genotype and phenotype?

Andreas Wagner

Yale University, Department of Biology, Center for Computational Ecology, OML 327, PO Box 208104, New Haven, CT 06520-8104, USA

Received 14 July 1995
Recommended by H Levine

Abstract. Does natural selection always penetrate the web of epigenetic interactions that produced the phenotypic character it acts upon and thereby affect individual genes? Or, alternatively, are there instances where character evolution is partly decoupled from DNA level events, i.e. are there characters that are 'irreducible', 'emergent', 'collective' features of the system that produced them? These are two persistent questions in evolutionary biology, and for good reasons: to answer them means to determine what level of biological organization is relevant to our understanding of evolution. Conceptual problems associated with this notion of collective behaviour are not the only reason for the elusiveness of these questions: they have to be answered by experiment, but available experimental methods are insufficient to resolve them. Since the nature of the problem is such that adequate experimental methods are not yet available, preliminary studies on models motivated by experimental evidence are called for. Such a study is presented here. It is based on a biochemically motivated model for sets of genes encoding for transcriptional regulators which mutually regulate each other's expression. The ubiquity of these 'gene networks' in higher eukaryotes and their central role in development make their evolution an important subject of study. On the basis of several statistical criteria, it is argued that selection acting on the level of individual genes within a network is most likely a minor factor in network evolution. Reasons are given for why these results are likely to be robust to alterations in model architecture and why they may thus be expected to apply to real gene networks. Possible implications of the results to evolutionary theory and molecular studies of development are discussed.

AMS classification scheme numbers: 92A12, 92A10

1. Introduction

Do selective forces that act on a phenotypic character in general affect individual genes, or are there 'irreducible', 'emergent', 'collective' features of the epigenetic system, arising through complex interaction patterns between many gene products? In other words, can the patterns of causation in the production of a phenotypic character be so complex that the underlying variation in individual genes cannot be linked to character variation? 'Gene selectionists', such as Dawkins (1982) and Williams (1966, 1992) argue in favour of reducibility. Other researchers (Lewontin 1974, Wimsatt 1980) argue against it. The discussion is ideologically charged, since an ancient philosophical problem is at stake. The above questions are part of a more general debate on the appropriate 'levels' or 'units' of biological organization on which natural selection acts. Most of this debate has revolved around a qualitative analysis of different levels of biological organization and their role in the evolutionary process (Brandon 1982, Dawkins 1976, 1980, Hull 1980, Lewontin 1970,

0951-7715/96/030607+23$19.50

Lloyd 1988, Williams 1966, 1992). Much of the terminology endemic to this debate, such as the notion of a 'unit of selection' itself, is used in slightly different ways by different authors (cf Buss 1987, Wagner 1990). Its usage will therefore be avoided here as much as possible. In this paper, a different, quantitative approach is pursued. A model system is used in which a phenotypic character is expressed via the action of a small number of genes that interact in a nonlinear way. To detect causal relations between the phenotypic level and the genic level in this model, operational criteria in the form of statistical tests are applied. Statistical approaches to a problem of this kind are not unprecedented; they have been used in sociology (Boyd and Iversen 1982) as well as in biology (Goodnight *et al* 1992) for the related problem of group selection versus individual selection.

The current emphasis on molecular studies of evolution and development indicates the implicit assumption that much of organismal evolution and development can be understood by studying molecules. While the molecular perspective has caused a quantum leap in our understanding of these processes, it seems surprising that it should always be possible to partition causally the immense number of interacting factors involved in the expression of the phenotype of an organism. It seems that, aside from ideological reasons, two major methodological limitations can be held responsible for this perceptional bias. Current methods in molecular biology and genetics seldom allow for rapid, highly accurate and simultaneous measurements of many biochemical and genetic parameters. Such data are required to establish testable models that provide the minimal level of complexity required to demonstrate the existence of collective behaviour. Since not even the necessary data are available, the problem cannot be addressed adequately. Thus, one might want to resort to predictions of existing models in mathematical evolutionary biology. A rich literature on complex phenomena in physics (e.g. Haken 1977, Stein 1990) strongly suggests that their occurrence is associated with nonlinear interactions among a system's parts. Although such nonlinear behaviour is probably the rule rather than the exception in the interaction of gene products (Wright 1968), most existing models are based on the unrealistic assumption of linear (additive) gene interactions and will therefore not be useful in this context. Again, there are good reasons for this assumption. The mathematical intricacies of models involving nonlinear (epistatic) gene interactions are formidable and rarely permit analytical approaches. Moreover, one particular complication is characteristic for nonlinear models: while there is only one kind of linear interaction, there are uncountably many kinds of epistatic interactions. Crucial features of a system may change with subtle alterations in the type of interaction involved. Robustness of results to small changes in a model and its parameters becomes a critical issue.

The above statements regarding (i) the absence of adequate data and (ii) the peculiar nature of nonlinear systems impose several restrictions on any attempt to answer the questions considered here on a more than purely qualitative level. First, any such attempt at present is likely to involve a mathematical model of a biological process rather than an experimental system. This model should stay as close to experimental evidence as possible, without incorporating too many immeasurable parameters. Importantly, the conclusions drawn from the model should be robust to changes in model parameters and, if possible, to changes in model architecture. Second, the results will most likely depend on the particular system under consideration. The immense variety of qualitatively different nonlinear systems in the physical sciences (e.g. Haken 1977, Stein 1990), together with the absence of a 'taxonomy' for these systems suggests that no universal body of theory is likely to cover the range of possible phenomena. Case studies will be necessary to explore this range. Due to the methodological limitations discussed above, most available

evidence argues in favour of a 'reductionist' perspective. This perspective might sometimes be inadequate, as indicated by the results presented below.

The conceptual framework underlying this study was made explicit by Lewontin (1970), who listed a set of conditions that are necessary for any evolutionary process to occur. First, there has to be phenotypic variation in a population. Second, different phenotypes have to have different fitness and, third, there has to be heritability in fitness. The notion of a phenotype is used in a non-standard meaning here, designating only the relevant level of organization for a particular evolutionary phenomenon. A phenotype may be a biochemical property of an RNA molecule in a hypothetical RNA world, a morphological trait of an organism, or some suitably defined property of a biological species. In this paper, any feature of an organism to which one can assign fitness in a biologically meaningful way is considered a phenotypic character. Character variation then implies variation in fitness, and if this variation is heritable, evolution becomes possible. The expression of the character will involve one or more genes. If more than one gene is involved, a question arises about the role that individual genes have in the expression of the character. A closely related question is how variation in fitness can be traced to variation on the level of individual genes. This is the question that will be addressed here. On the basis of several statistical criteria and for the model system under consideration, it will be argued that variation in fitness cannot be decomposed into independent contributions to this variation by individual genes. One important criterion used here was introduced by Wimsatt (1980) and is discussed in detail by Lloyd (1988). It is based on a notion of non-additivity—a property already known to be associated with collective behaviour in physical systems. According to this criterion, a character is a higher order 'unit of selection', if additive variance at the fitness level cannot be decomposed into a sum of contributions to this variance by the genic level.

Basic pattern formation processes during animal development involve evolutionarily highly conserved proteins that regulate gene expression on the transcriptional level. Empirical evidence suggests that these transcriptional regulators frequently interact in a network-like fashion to establish patterns of gene expression which, in turn, determine basic *Bauplan* features of the organism. Their central role in development make such gene networks and their constituent genes an important and popular subject of research in evolutionary and developmental biology (Ingham 1988, McGinnis *et al* 1990, McGinnis and Krumlauf 1992, Olson 1990, Rosenfeld 1991, Kappen *et al* 1989). A mathematical model for their evolution, based on biochemical evidence, is used here to address the issues raised above (see also Wagner 1994, 1996). The gene expression pattern produced by such a gene network is viewed as a phenotypic character. The fitness associated with this phenotype is defined by the extent of deviation of this expression pattern from a hypothetical, optimal pattern. Statistical relations (i) between variation in fitness and units of genetic variation as well as (ii) among units of genetic variation are explored. The analysis suggests that there may be systems of only few genes, whose (nonlinear) interaction patterns are such that it is effectively impossible to detect meaningful associations between variation on the phenotypic level and variation on the genic level, even in the absence of noise and with large data sets (i.e. high statistical significance). These findings are robust to changes in all model parameters. The notions of 'emergence' and 'collective' behaviour will hereafter be used in this rather technical sense. To decide whether the behaviour of such a system is 'irreducible', because it cannot be decomposed into the behaviour of individual parts, or whether it is 'in principle' reducible, because interactions among its parts are governed by deterministic laws, is beyond the scope of this contribution. Possible implications of the results to the theory of evolution and to molecular studies of development are discussed.

2. The model

Only a fraction of the genes encoding transcriptional regulators are likely to be expressed in any given cell and during any given ontogenetic stage of an organism. Moreover, expression patterns of these genes may vary from cell to cell and from stage to stage. The model to be used here refers to the expression pattern of transcription factor genes only in one developmental stage and only in one cell or a body region that shares an expression pattern, e.g. a set of nuclei in a part of a *Drosophila* blastoderm expressing a specific subset of segmentation genes (Ingham 1988). A set of N such genes, denoted as $(G_1, \ldots, G_N)$, whose products mutually regulate each other's expression on the transcriptional level, will be referred to as a 'network'. Due to cross-coupling between regulatory pathways (e.g. Schüle and Evans 1991), the number of regulatory proteins (transcriptional regulators and others) involved in most intracellular regulatory circuits is probably large. However, those circuits presumably relying to a large extent on transcriptional regulation may be quite small, involving 10–100 or fewer genes, according to the available circumstantial evidence (Ingham 1988, He *et al* 1989, McGinnis and Krumlauf 1992).

Regulation of transcription from RNA Polymerase II promoters in eukaryotes is a process in which DNA interacts with multi-protein complexes (Johnson and McKnight 1989, Mermelstein 1989). Many of the protein–protein and protein–DNA interactions involved are poorly understood, and a considerable number of genes important for the process are probably not even cloned and characterized (Weinzierl *et al* 1993). Based on the available empirical data, it seems therefore unlikely that a good qualitative model—let alone a quantitative theory— of transcriptional regulation will be available in the near future. For these reasons and in order to arrive at an analytically and computationally tractable formalism, a number of simplifying assumptions will be used in the model presented here. It is assumed (i) that expression of the genes in the network is regulated exclusively on the transcriptional level, (ii) that each gene of the network produces one and only one species of an active transcriptional regulator and (iii) that enhancer elements mediating one regulator's effect on expression of the target gene act independently from enhancer elements for other regulators of the same gene.

In the model used here and motivated in a more formal way by Wagner (1994), a gene network is represented by a dynamical system whose state variables correspond to expression states of the network's genes. They are denoted as

$$\vec{S}(t) := (S_1(t), \ldots, S_N(t)) \tag{1}$$

where $S_i(t)$ is the expression state of the ith gene at some time $t \geqslant 0$ during a developmental process in which the network acts. For reasons of computational simplicity, it is assumed that $S_i(t)$ can only assume two values, namely $(+1)$ and (-1), corresponding to a situation in which the gene G_i is expressed or not expressed, respectively, at time t. The gene expression state $\vec{S}(0)$ at time $t = 0$ is called the initial expression state. It can be conceptualized as being imposed onto the network by the products of one or more 'upstream' genes that are not themselves part of the network. Such products might be extracellular signalling molecules, such as growth factors or differentiation signals, but also transcriptional regulators, e.g. a retinoic acid receptor acting on homeobox genes in a developing vertebrate limb. The boundary of a network is therefore somewhat arbitrary: a gene is defined as 'upstream' by virtue of the fact that it regulates the expression of network genes, but is not regulated by these genes.

Starting from the initial gene expression pattern, $\vec{S}(0)$, cross- and auto-regulatory interactions among network genes cause the expression state to change. These changes

are modelled by the set of difference equations

$$S_i(t+\tau) = \sigma\left[\sum_{j=1}^{N} w_{ij} S_j(t)\right] = \sigma[h_i(t)]\,. \tag{2}$$

Here, the expression state of gene G_i at time $t+\tau$, $S_i(t+\tau)$, is a function of a weighted sum, $h_i(t)$, of the expression state of all network genes at time t. $h_i(t)$ represents the sum of the regulatory effects of all network genes on gene G_i. $\sigma(x)$ is the sign function ($\sigma(x) = -1$ for $x < 0$, $\sigma(x) = +1$ for $x > 0$ and $\sigma(0) = 0$), and τ is a time constant characteristic for the process under consideration. Its value will depend on biochemical parameters such as the rate of transcription, or the time necessary to export mRNA into the cytoplasm for translation. The real constants w_{ij} represent the 'strength' of regulatory interaction of the product of G_j with G_i, i.e. the degree of transcriptional activation ($w_{ij} > 0$) or repression ($w_{ij} < 0$) that the transcriptional regulator produced by gene G_j has on gene G_i. Such regulatory interactions are known to be mediated by regulatory (enhancer) DNA sequences. In biological terms, individual w_{ij}'s can be thought of as some compound measure of the binding constant and the transcriptional activation (repression) ability of the factor produced by G_j at the enhancer element that mediates its interaction with G_i. Alternatively, and in line with the structure of (2), w_{ij} can be thought of as a measure of the influence that the product of G_j has on G_i relative to other gene products. In this sense, it is the relative size of the w_{ij}'s that is relevant to the dynamics of (2). The ith row of w, $w_i := \{w_{ij} | 1 \leqslant j \leqslant N\}$ corresponds to the entire enhancer of gene G_i with all regulatory DNA elements that affect the expression of G_i. The 'connectivity matrix' $w = (w_{ij})$ corresponds to all DNA elements relevant to regulatory interactions among network genes. Any non-zero diagonal element, $w_{ii} \neq 0$, corresponds to autoregulation of G_i by its own gene product (e.g. Regulski *et al* 1990; Sucov *et al* 1990). Some (or most) entries of w may be zero. The fewer non-zero entries w has, the fewer regulatory interactions exist among network genes. An important model parameter is therefore the fraction of entries different from zero, denoted by c ($c \in (0, 1)$), which will be called the 'connectivity density' of the network. The discrete-time dynamical system (2) can also be viewed as the limiting case of a system of differential equations, in which concentrations of gene products, rather than binary (on–off) gene expression states change (Wagner 1994), but computational limitations prohibited the use of such a system here. It should be noted that the structure of (2) is similar to 'spin glass' (Binder and Young 1986) or neural net (Amit 1989, Hopfield and Tank 1986) type models of gene networks first introduced by Kauffman (1969, 1993). However, (2) is conceptually different from Kauffman's models in that a specific type of gene interaction, namely transcriptional regulation, is considered. Recently, models conceptually similar to (2) have been successfully used to describe and predict regulatory gene interactions in early *Drosophila* embryogenesis (Mjolsness *et al* 1991, Reinitz *et al* 1995, Reinitz and Sharp 1995).

The dynamics of (2) will lead to the attainment of an asymptotic (equilibrium) gene expression state, which may be a fixed point of (2) or a limit cycle. The genes expressed in the asymptotic state will affect the expression of genes outside ('downstream') of the network. Possible downstream genes may include structural genes or genes encoding proteins involved in signal transduction processes, but also transcriptional regulators that do not themselves regulate the expression of genes within the network (Budd and Jackson 1991). Many experimental genetic studies (e.g. McGinnis and Krumlauf 1992) suggest that deviations in a network's expression pattern from the wild type pattern causes developmental perturbations that often lead to deleterious effects on the adult phenotype. In

a biological population, such individuals would be eliminated by natural selection. These observations motivate the assumption that an optimal asymptotic gene expression state, denoted as $\vec{S}^{opt}(\infty)$, exists for networks acting in a developmental process. For reasons of computational tractability, it is assumed that this state is a fixed point of (2). If a network attains an asymptotic state $\vec{S}(\infty)$ that is different from this optimal state, developmental perturbations will result and the fitness of the respective adult organism will be reduced. Deviations from an optimal state may occur for a variety of reasons, one of them being mutations in genes within or outside of the network. The pattern of regulatory interactions within a gene network represents its most interesting 'organizational' properties, and thus mutations in regulatory DNA regions, represented by changes in the matrix w, will be the focus of this study.

The evolutionary scenario envisaged here involves a gene network acting in an ontogenetic process in each of the members of a population of organisms. It is assumed that both the initial gene expression pattern, $\vec{S}(0)$, and the optimal asymptotic gene expression pattern, $\vec{S}^{opt}(\infty)$, are the same for all organisms in that population. The organisms are subject to mutations of regulatory DNA regions, recombination among network genes, genetic drift, and selection on the attainment of $\vec{S}^{opt}(\infty)$. This requires a notion of an individual's fitness, which is modelled in the following way. A measure, $D(w)$, for the distance between a network's asymptotic state, $\vec{S}(\infty)$, to the optimal state, $\vec{S}^{opt}(\infty)$, is given by

$$D(w) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} d_h[\vec{S}(t), \vec{S}^{opt}(\infty)] \tag{3}$$

where

$$d_h[\vec{S}(t), \vec{S}^{opt}(\infty)] = \frac{1}{2} - \frac{1}{2N} \sum_{i=1}^{N} S_i(t) S_i^{opt}(\infty) \tag{4}$$

is the Hamming distance (e.g. Amit 1989), which counts the number of expression states of individual genes that are different in the two states, and normalizes it to the interval $(0, 1)$. Note that (3) reduces to $D(w) = d_h[\vec{S}(\infty), \vec{S}^{opt}(\infty)]$, if the state $\vec{S}(\infty)$ is a fixed point of (2). Based on $D(w)$, the fitness of an individual is then given via a Gaussian fitness function as

$$f(w) := \exp\left[-\frac{D(w)^2}{s}\right]. \tag{5}$$

The notations $D(w)$ and $f(w)$ indicate that mostly the dependence of fitness on the pattern of regulatory interactions, represented by w, will be of interest here. The parameter s ($s > 0$) represents the strength of selection, small values of s implying strong selection against deviations from the optimal state. Transformations of $f(w)$ will be used for some statistical tests.

Individual genes belonging with a network are assumed to be unlinked (subject to free recombination), with the following rationale: with tight linkage between genes, the whole network acts as a 'supergene'. In the scenario considered here, the only phenotypic differences between parents and offspring would then be those caused by mutations, which implies high heritabilities. Heritabilities different from zero are much less intuitive if a large amount of genetic information is shuffled in every generation. Thus, free recombination between genes *a priori* provides a more stringent test for the heritability criterion that has to be met by an evolvable system. The fact that a haploid model was chosen simplifies the formalism considerably, and is unlikely to affect conclusions qualitatively, as will be discussed below.

3. Numerical methods

3.1. Population initialization

Network sizes used in the simulations ranged from $N = 5$ through $N = 10$. Two pseudorandom arrays in $\{-1, 1\}^N$, corresponding to $\vec{S}(0)$ and $\vec{S}^{opt}(\infty)$ were chosen stochastically and independently with $\boldsymbol{P}(S_i = 1) = p$ for each entry. Then, an ordered set ('population'), $W := \{w^{(1)}, \ldots, w^{(P)}\}$, of $P = 500$ (except where noted otherwise below) $N \times N$ pseudorandom matrices $w^{(k)} = (w_{ij}^{(k)})$ of regulatory interaction strengths was generated. In these matrices, cN^2 entries (cN per row) were assigned a value of zero, where c denotes the 'connectivity density' of a network, as introduced above. In order to apply some of the statistical tests used here appropriately, it was necessary that matrix entries identical to zero occur at the same position in each of the matrices. In other words, if $w_{ij}^{(k)} \neq 0$, $w_{ij}^{(l)} \neq 0$ held for all $l \neq k$. The $(1 - c)N$ non-zero matrix entries in each row were chosen randomly and independently from those in other rows. Individual non-zero entries were distributed identically and independently within and among matrices, following a continuous probability density $p(x)$. Gaussian ($p(x) = 1/(\sqrt{2\pi}\sigma)\exp[-x^2/(2\sigma^2)]$) or 'reflected Gamma' densities ($p(x) = [2\Gamma(a)]^{-1}\mathrm{e}^{-|x|}|x|^{a-1}$, $x \in \Re \setminus \{0\}$, $a > 0$) were used, since they cover a wide range of qualitatively different shapes of symmetric distributions. This allows one to assess how robust results are to changes in the distribution of regulatory interactions.

3.2. Simulated evolution

A population thus initialized was subject to a process of simulated evolution, in the course of which various statistics (detailed below) were evaluated. The following five steps represent one generation of simulated evolution for a population of networks. The steps were carried out in the order given here. The number of iterations of the five steps (i.e. the number of generations) depended on the actual statistic measured during the process as well as on the individual population (see below), but did not exceed 5000 generations.

(i) *Recombination.* In pairs of consecutive matrices in the population, starting with $w^{(1)}$ and $w^{(2)}$, rows were swapped with probability 0.5, corresponding to free recombination between genes and tight linkage within an enhancer. Note that randomness in the order of matrices is implied by the selection algorithm (step four).

(ii) *Mutation.* One per cent of all non-zero connectivities were mutated stochastically in the following way. For each connectivity in each of the matrices in the population, starting with $w_{11}^{(1)}$, a pseudorandom number r with uniform distribution on $(0, 1)$ was generated. If its value was greater than 0.01, the respective connectivity was not changed. If it was less than 0.01, it was checked whether that connectivity was different from zero, and if so, the connectivity was replaced by a pseudorandom number following the same distribution $p(x)$ as that used in the initialization of the population W. This approach is essentially the house-of-cards approximation that is frequently used in population genetic modelling (see, e.g. Zeng and Cockerham 1993, Turelli 1985).

(iii) *Fitness evaluation.* The dynamics (2) was evaluated numerically for each of the networks in the population W. If a given network attained a fixed point $\vec{S}(\infty)$, its fitness was evaluated according to (5). If a network did not reach a fixed point, it was subjected to 100 additional time steps of (2), which is a number of steps an order of magnitude greater than the average period of a limit cycle. In the course of this process, the average Hamming distance to $\vec{S}^{opt}(\infty)$ was calculated using (3), and from it, the fitness (5).

(iv) *Selection.* After fitness had been assigned to all networks in the population, it was normalized such that the maximum fitness in the population was equal to one. Then, a network was chosen at random and a pseudorandom number, r, with uniform distribution on $(0, 1)$ was generated. If r was less than the fitness of the network, the network 'survived'. This process was repeated until a new population of the same size as the old population had been generated, i.e. sampling of networks was carried out with replacement. Networks with high fitness spread through the population because of their greater probability of survival.

(v) *Adjustment of selection strength.* Selection of matrices that arrive at some $\vec{S}^{opt}(\infty)$ is a difficult optimization task in a high dimensional space. In order to facilitate the generation of populations with high mean fitness, a technique related to simulated annealing (Kirkpatrick *et al* 1983) was used, in which selection strength is increased during the course of the simulation, depending on a population's mean fitness. In generation one, the selection strength was set to $s = 1$. In each subsequent generation, it was checked whether the mean fitness of the population exceeded a value of 0.9, and if so, the selection parameter s was decreased according to $s' = s\mathrm{e}^{-\lambda}$, where $\lambda = 0.01$.

The process of simulated evolution carried out here has two different stages, an initial stage, in which mean fitness increases, and a second stage, in which the population is in a (quasi-)equilibrium with respect to mean fitness. For the gene networks under consideration here, the latter situation is probably the biologically relevant one. However, some of the tests utilized below require a sufficient amount of intrapopulational fitness variation. Thus, the earlier phase, during which mean fitness evolves is important here for technical reasons, because of the higher levels of variation it provides.

3.3. Statistics

During the process of simulated evolution, various statistics were taken from the population after fitness evaluation (step 3 above). Results obtained from these statistics were qualitatively insensitive to variations in model parameters. They will therefore be discussed only for $N = 10$, $p = 0.5$, Gaussian distribution of connectivities with $\sigma = 0.1$, $c = 1$ as well as, for comparison, $c = 0.2$. Events in the underlying probability space are sufficiently complex, e.g. the periodic shuffling of network parts by recombination, that accurate analytical estimates for the statistics given here can most probably not be obtained. In this section, only some general remarks are made. A more detailed discussion of individual statistical tests is given in the next section along with the results.

Most statistics used here are based on three common measures of association among random variates, namely Pearson product-moment correlation coefficients, linear regression coefficients and Kendall's coefficient of rank correlation ('Kendall's τ'; Sokal and Rohlf 1981). The random variates under consideration were network connectivities or fitness. To be precise, measures closely related to fitness were used in these calculations. For reasons of computational simplicity, the distance measure (3) itself was used in *all* calculations involving Kendall's τ. To make distributions approximate more closely normal distributions, $-\log[D(w) + 5/100]$ was used in *all* estimates of Pearson correlation and regression coefficients. For all three measures of association, a significance test (Sokal and Rohlf 1981) was carried out for the null hypothesis that the respective measure is different from zero. The significance test for Pearson's correlation coefficient is based on a tabulated t-value for 120 degrees of freedom. Since populations used were larger than this value, the test is slightly too conservative.

Type II errors represent a potential problem in this system because of the limited analytical insight into its behaviour. It is difficult to find a rational, testable alternative

hypothesis that could serve to evaluate the power of a test (e.g. Sokal and Rohlf 1981, p 150). To address this problem, statistics were taken at two different significance levels, namely $\boldsymbol{P} = 0.01$ and $\boldsymbol{P} = 0.001$. The qualitative results discussed below turned out to be independent of this change in significance level and they will therefore be discussed in detail only for $\boldsymbol{P} = 0.01$. The very large sample sizes (at least 500 for each test) lend additional confidence to the results.

There is variation among simulations in the time needed to reach fitness equilibrium, and there are therefore also different time courses of mean fitness evolution. This implies that averaging results over a large number of independent simulations, although desirable, is not feasible here, because simulations cannot be compared on a generation-by-generation basis. Therefore, results from representative simulation runs are given below. No simulation runs occurred in which results were qualitatively different from those reported.

4. Results

This section addresses two issues. First, is non-zero heritability in fitness (and therefore evolution) possible in this model? Second, can variation in fitness be traced to variation patterns of individual genes?

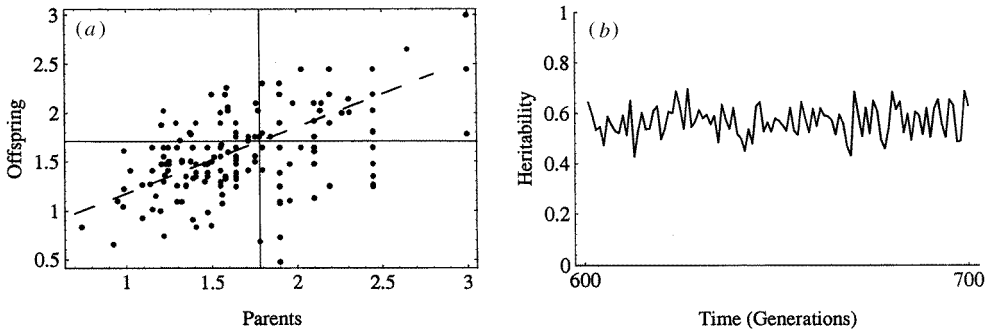

Figure 1. Fitness is heritable. (*a*) Plot of mean parental versus mean offspring fitness ($-\log[D(w) + 0.05]$) during mean fitness evolution. Heritability estimated from this plot was $h^2 \approx 0.68$ ($\boldsymbol{P} < 0.01$). (*b*) Time series of heritability estimates obtained as in (*a*) for a network with connectivity density $c = 1$ from generation 600 through generation 700 after the start of the simulation. All values obtained were significantly different from zero ($\boldsymbol{P} < 0.01$). $\overline{h^2} = 0.56$ for the time window shown. See text for details. Other parameters: $N = 10$, $p = 0.5$, $P = 500$, $p(x)$ Gaussian with $\sigma = 0.1$.

4.1. Fitness is heritable

Fitness of two 'offspring' networks obtained from two 'parents' via recombination and mutation was calculated for each pair of parents in a population. Figure 1(*a*) shows a plot of mean parental fitness versus mean offspring fitness for all parent–offspring pairs in one generation during the phase of mean fitness increase. The discreteness of the fitness measure is reflected in the occurrence of data points that are horizontally or vertically aligned. Heritability, i.e. the slope of the regression line through the set of data points is approximately equal to 0.68 ($\boldsymbol{P} < 0.01$). Figure 1(*b*) shows a time series of heritabilities

from generation 600 through 700 after initialization, corresponding to the phase of mean fitness evolution. This time series was calculated in the same way as the individual value obtained from figure 1(*a*). All values shown were significantly ($\boldsymbol{P} < 0.01$) different from zero. (The fact that successful selection on increasing mean fitness is possible in and by itself implies non-zero heritability.) Although the exact values of heritabilities are irrelevant, it should be mentioned that they are high compared to experimentally measured heritabilities for some quantitative characters (Falconer 1981). The reason is most likely the absence of environmental variance in the model. Non-zero parent–offspring regressions were also observed in populations with sparsely connected networks, and in populations that were in mean fitness equilibrium (results not shown).

4.2. Individual enhancer sites contain little information about macroscopic network properties

The observation that the distributions or even the strengths of individual regulatory interactions remain the same in a population with high and stable mean fitness would provide a first indication that macroscopic network features can be traced to the action of individual enhancer elements. Those sites, if they exist, would be prime candidates for elements important to a network's function. Such an observation could indicate the result of a 'symmetry breaking' process among the enhancer sites, during which some sites in the network became more important than others.

The following simulation approach addresses this issue. A population in fitness equilibrium was 'duplicated' and the duplicate as well as the original population were subject to the same regime of selection on $\vec{S}(\infty) = \vec{S}^{opt}(\infty)$, except that they evolved independently from each other after duplication. Each non-zero entry w_{ij} of the connectivity matrix w may assume a variety of values in each of the two populations. For each non-zero entry, a Kolmogorov–Smirnov test on the identity of its distribution in the two populations (Sokal and Rohlf 1981, p 440) was carried out to find out whether the two sets of values represent samples from the same underlying probability distribution. Sample sizes used for the tests were equal to population size, $P = 1000$. This large population size was used to minimize the effects of genetic drift, which tend to homogenize the distributions for reasons unrelated to selection. The fraction $\xi(t)$ of tests for which the null hypothesis of identity in distribution was rejected (on some significance level $\boldsymbol{P}$) was recorded in each generation t. A value $\xi(t) = 1$ means that none of the entries have the same distribution, whereas $\xi(t) = 0$ means that all entries have the same distribution. It is not trivial to arrive at a significance measure for a compound quantity such as $\xi(t)$, given $\boldsymbol{P}$, since high and low order correlations between distributions of different entries in the same population may exist. To address this issue, tests were carried out using $\boldsymbol{P}$-values ranging from $\boldsymbol{P} = 0.01$ to $\boldsymbol{P} = 10^{-6}$. Results obtained were qualitatively identical for all these values and are therefore shown only for $\boldsymbol{P} = 0.01$.

The time series $\{\xi(t)\}$ is plotted for a population of densely connected networks ($c = 1$) and a population of sparsely connected networks ($c = 0.2$) in figures 2(*a*) and 2(*b*), respectively. In both cases, $\{\xi(t)\}$ converges in a small number of generations to $\xi(t) = 1$. A decrease in $\boldsymbol{P}$ leads to a delay in the number of generations elapsed before $\xi(t) = 1$ is first reached. However, regardless of $\boldsymbol{P}$, $\xi(t) = 1$ is first reached before 100 generations have elapsed. This implies that after a small number of generations, *all* distributions of connectivities have diverged, even for two populations in which almost all networks have highest fitness. Therefore, two sets of networks generating the same gene expression pattern may have little in common in terms of individual regulatory interactions, or even in terms of the distributions of interaction strengths. This fact is even more remarkable for a network

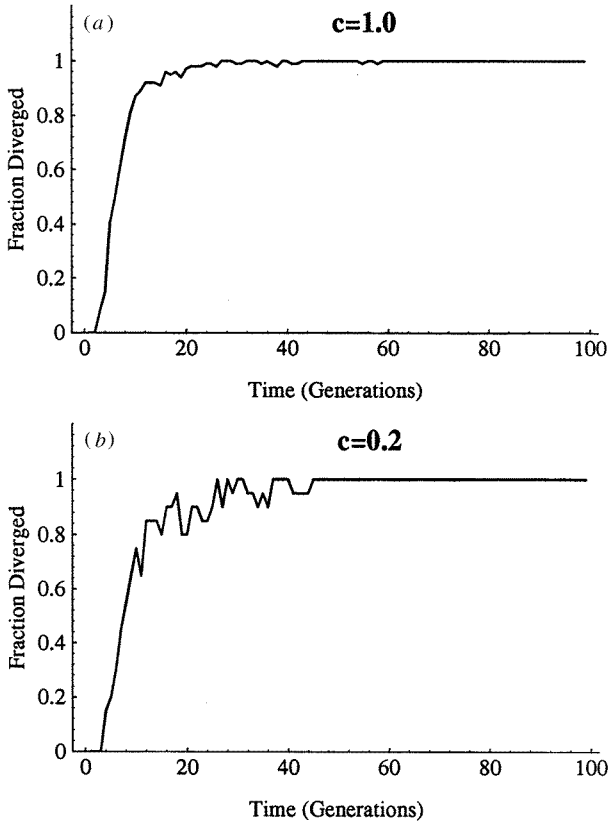

Figure 2. Connectivity distributions in populations evolving in parallel diverge rapidly. A 'duplicate' of a population of networks in mean fitness equilibrium and under strong selection was generated. After duplication, the duplicate and the original population were independently subjected to the same regime of strong selection. The fraction, $\xi(t)$, of marginal distributions of corresponding individual connectivities that differed in the two populations was plotted for the first 100 generations after the duplication event. (*a*) $c = 1$; (b) $c = 0.2$. The plot in (*b*) appears more coarse grained than the plot in (a), because $\xi(t)$ can only assume $cN^2 = 20$ different values, as opposed to $cN^2 = 100$ different values in (*a*). Other parameters: $p = 0.5$, $P = 1000$, $p(x)$ Gaussian with $\sigma = 0.1$.

in which few regulatory interaction strengths are different from zero, such as that shown in figure 2(*b*). In this case, a fewer overall number of enhancer sites accounts for network dynamics, and one might suppose that some connectivities play more important roles than others, which is not the case.

The test used here only provides a necessary criterion and a first suggestion towards the possibility that it may be difficult to pinpoint variation in fitness to variation in individual enhancer elements. Its main statement is that there are many possible genotypes producing the same 'phenotype' (gene expression pattern), a property that this system shares with simpler, additive polygenic systems. Thus, it has to be complemented by other criteria.

4.3. Few and varying binding sites show correlation to fitness

If there exist simple causal relations between the strengths of individual regulatory interactions and the gene expression pattern generated by a network, the strengths of these

interactions should be correlated with fitness. The case of a linear model, although not comparable to the network model considered here, is instructive in this context. In such a model, the genotypic value, G, would be represented as the sum of n underlying genetic variables, X_i, and the correlation between individual such variables and G would be $1/\sqrt{n}$. With a sample size as large as that used here, sample correlation coefficients would likely be significantly different from zero, even for small $\boldsymbol{P}$. In the nonlinear network model studied here, a comparable result does not hold.

Two measures of correlations were calculated in populations during the phase of mean fitness evolution. First, Pearson correlation coefficients between fitness and individual entries of w were calculated for all entries of w different from zero. Note that a calculation of these correlations is not expedient in mean fitness equilibrium, since in that case many of the individuals in the population may have fitness one. Only the fraction $\zeta_p(t)$ of those regulatory interaction strengths that showed a non-zero Pearson correlation coefficient was recorded in each generation t. A parallel, control simulation showed that there exist background correlations unrelated to selection that have to be accounted for. In this simulation, a population was used that had evolved for the same number of generations as the above population, but in the absence of selection (using $s = 100$ in (5)), i.e. it was only under the influence of mutation and genetic drift. For the evaluation of correlation measures in the control population, the fitness of each network was calculated using $s = 1$, i.e. as if the population had been under selection. The time series for the control population is denoted by $\zeta_p^c(t)$.

Analogously, Kendall's rank correlation coefficient τ was calculated between connectivities and fitness, resulting in two time series of the fraction of non-zero τ's, $\zeta_\tau(t)$ and $\zeta_\tau^c(t)$, for a population under selection and a control population, respectively. This measure of association should give more accurate results if fitness and network connectivities are not normally distributed. In many simulations, both $\zeta_p(t)$ and $\zeta_p^c(t)$ were consistently greater than their counterparts $\zeta_\tau(t)$ and $\zeta_\tau^c(t)$, respectively. This was interpreted to be caused by false non-zero Pearson coefficients due to deviations from normal distributions. Aside from this discrepancy, results were qualitatively identical for the two measures of association, and only $\zeta_\tau(t)$ as well as $\zeta_\tau^c(t)$ will therefore be discussed in detail.

Figures 3(*a*) and 3(*b*) show $\zeta_\tau(t)$ and $\zeta_\tau^c(t)$ for a densely connected ($c = 1$) network. The amount of background correlation can be seen from figure 3(*b*). Importantly, selection itself does not generate a large number of significant correlations. If one uses $\overline{\zeta_\tau(t)} - \overline{\zeta_\tau^c(t)} = 0.097$ as a crude measure of the average number of correlations related to selection (where bars indicate temporal averages over the time windows shown in figures 3(*a*) and 3(*b*)), one finds that less than ten such significant correlations exist on average. Figures 3(*c*) and 3(*d*), showing the same quantities but for a sparsely connected network with $c = 0.2$, represent an even more extreme situation. Here, $\overline{\zeta_\tau(t)} - \overline{\zeta_\tau^c(t)} = 0.18$, corresponding to less than four significant correlations. How likely is it that this low a number of significantly correlated connectivities, e.g. four or less in the case of the sparsely connected network, could account for all variation in fitness? Further doubt in the relevance of these correlations is raised by considering the large amount of fluctuation in all time series shown, as well as the large amount of background correlation, especially for the sparsely connected case (figure 3(*d*)). A more detailed analysis of the correlation patterns will further support these doubts.

If two snapshots of a population at time t_1 and time t_2 are taken, will the patterns of correlation between enhancer sites and fitness be similar and, more specifically, will the same enhancer elements show correlations to fitness? Such enhancer elements, if they exist, might be important to network function. The fluctuations observed in $\zeta_\tau(t)$ (figures 3(*a*)

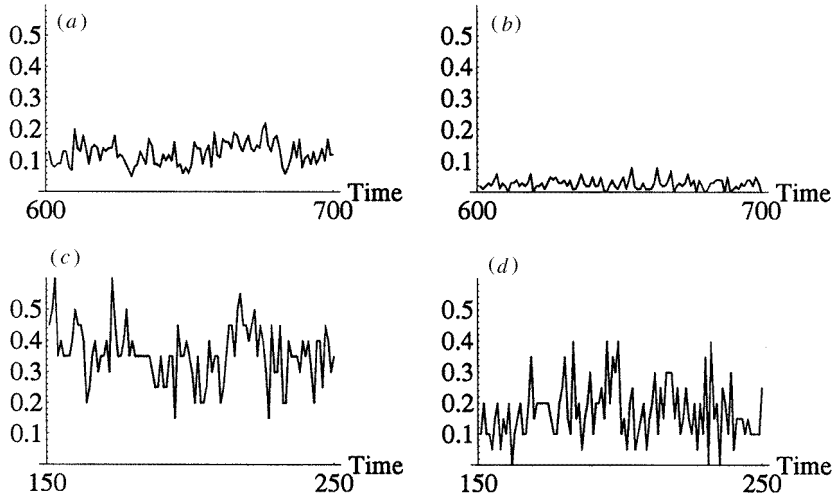

Figure 3. Few connectivities are correlated to fitness. Fraction, $\zeta_\tau(t)$, of non-zero Kendall's τ correlation coefficients between connectivities and fitness, plotted against the number of generations, t, after start of the simulation. (a) Densely connected networks ($c = 1$), population under strong selection, mean fraction, $\overline{\zeta_\tau} = 0.124$, where the bar indicates a temporal average over the time window shown. (b) $c = 1$, no selection, $\overline{\zeta_\tau^c} \approx 0.027$. ($c$) Sparsely connected network ($c = 0.2$), population under strong selection, $\overline{\zeta_\tau} \approx 0.357$. ($d$) $c = 0.2$, no selection, $\overline{\zeta_\tau^c} \approx 0.173$. ($b$) and ($d$) were included in order to account for correlations not caused by selection. Both types of population were in the same stage in terms of the evolution of their mean fitness. The differences in the time windows shown are due to differences in the time course of mean fitness evolution in the two network types. Other parameters used: $N = 10$, $p = 0.5$, $P = 500$, $p(x)$ Gaussian with $\sigma = 0.1$. Note the large fluctuations in (a) and (c), as well as the considerable amount of background correlations for $c = 0.2$, as shown in (d).

and 3(c)) already argue against such a conservation. In order to answer this question more rigorously, two different approaches were pursued.

(i) *Autocorrelation of Kendall's τ correlation matrices based on Euclidean distance measure.* A time series of $N \times N$ matrices $\{Y(t)\} = \{(Y_{ij})_{i,j=1,\ldots,N}(t)\}$ of Kendall's coefficient τ, $Y_{ij}(t)$, between w_{ij} and fitness at time t was evaluated. The question that has to be answered is: how rapidly does the similarity between these matrices decay over time? A fast decay implies that correlation patterns between individual enhancer elements and fitness are not conserved. To answer this question, an autocorrelation function of matrix similarity was designed. The distance measure d between matrices is defined as

$$d[Y(t_1), Y(t_2)] := \frac{1}{2cN^2} \sum_{i,j=1}^{N} |Y_{ij}(t_1) - Y_{ij}(t_2)|,$$

where the normalization factor $1/(2cN^2)$ takes into account that there are cN^2 matrix elements different from zero and that individual correlation coefficients can assume values in $(-1, +1)$, thus restricting d to the interval $(0, 1)$. Based on d, the following autocorrelation function of the time series of matrices is defined.

$$\rho_Y(\delta) := 1 - \frac{\sqrt{\pi}}{(T-\delta)\overline{\sigma}_\tau} \sum_{t=1}^{T-\delta} d[Y(t), Y(t+\delta)]$$

For sample sizes $n > 40$, Kendall's τ follows approximately a normal distribution ($N(0, \overline{\sigma}_\tau^2)$, Sokal and Rohlf 1981, p 607). It is straightforward to show that the expectation of d, under the assumption of identically and independently distributed correlation coefficients, is equal to $\overline{\sigma}_\tau/\sqrt{\pi}$. Thus, the normalization factors in the definition of ρ_Y assure that $\rho_Y(\delta) = 0$ for all $\delta > 0$, if correlation coefficients within and across matrices are stochastically independent, i.e. if correlation patterns in one generation yield no information about correlation patterns in the following generation. Note that $\rho_Y(0) = 1$ and $\rho_Y(\delta) \leqslant 1$, for all $\delta > 0$. In the simulations, $\overline{\sigma}_\tau$ was estimated as the average $(1/T)\sum_{t=0}^{T} \overline{\sigma}_\tau(t)$, where $\overline{\sigma}_\tau(t)$ is the standard deviation of Kendall's τ values in matrix $Y(t)$.

(ii) *Autocorrelation of matrices of significant Kendall's τ coefficients.* A time series of $N \times N$ matrices $\{Z_\tau(t)\} = \{(Z_{\tau,i,j})_{i,j=1,\ldots,N}(t)\}$ was constructed as follows:

$$Z_{\tau,i,j}(t) := \begin{cases} 1 & w_{ij} \text{ is correlated with fitness at time } t \\ 0 & w_{ij} \text{ is not correlated with fitness at time } t \end{cases} \quad \forall i, j, t\,.$$

Entries of each (binary) matrix in this series indicate which regulatory interaction strengths are correlated to fitness. The question is, then, to what extent is the structure of these matrices conserved over time? Given two matrices, $Z(t)$ and $Z(t+\delta)$, the hypothesis that the zero-one patterns in these matrices are stochastically independent was tested by a 2×2 χ^2 test for independence (Sokal and Rohlf 1981), grouping the numbers of zeros and ones into a 2×2 table according to their co-occurrence in the matrices. The quantity

$$\chi(t, \delta) := \begin{cases} 1 & Z(t),\ Z(t+\delta) \text{ are not stochastically independent} \\ 0 & Z(t),\ Z(t+\delta) \text{ are stochastically independent} \end{cases} \quad \forall t, \delta$$

was used to define the autocorrelation function ρ_Z,

$$\rho_Z(\delta) := \frac{1}{(T-\delta)} \sum_{t=1}^{T-\delta} \chi(t, \delta), \tag{6}$$

with $t \in \{1, \ldots, T\}$. Note that $\rho_Z(0) = 1$ and $\rho_Z(\delta) \leqslant 1$, for all $\delta > 0$. The faster $\rho_Z(\delta)$ decreases as δ increases, the less conserved are the patterns of zeros and ones that indicate which enhancer sites show correlation to fitness.

Results from these different autocorrelation functions are similar, in that both autocorrelation functions decay rapidly. They will therefore only be discussed for ρ_Z. Figures 4(*a*) and 4(*b*) show the autocorrelation function $\rho_Z(\delta)$ for a densely connected ($c = 1$) network and a sparsely connected ($c = 0.2$) network. The significance level for both the χ^2-test and the significance test for Kendall's τ was $\boldsymbol{P} = 0.01$. The length of the time series used is $T = 300$ generations, with a maximum lag of $\delta = 50$ generations. The time window used to calculate ρ_Z was comparable to that shown in figure 3, i.e. the population was in the phase of mean fitness evolution. To check the performance of the test, a time series of length $T = 300$ of $N \times N$ pseudorandom matrices with entries zero and one was generated, in which entries were stochastically independent within and across matrices. The probability of each entry being one was equal to 0.5. The sample autocorrelation function of this control series was evaluated (see figure 4(*c*)) and the results were as close to $\rho(\delta) = 0\ \forall \delta > 0$ as could be expected from the given significance level. It is obvious from figures 4(*a*) and 4(*b*) that there is a rapid decay of 'memory' in the time series, regardless of the density of connections in the network. This strongly suggests that connectivities correlated to fitness are unlikely to be the same at times t and $t + \delta$, even if

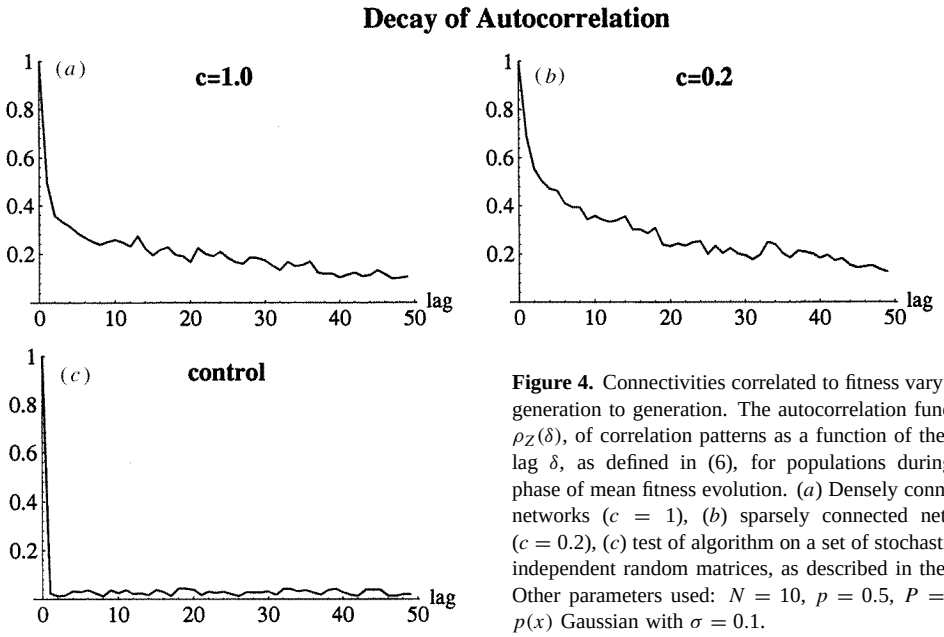

Figure 4. Connectivities correlated to fitness vary from generation to generation. The autocorrelation function, $\rho_Z(\delta)$, of correlation patterns as a function of the time lag δ, as defined in (6), for populations during the phase of mean fitness evolution. (*a*) Densely connected networks ($c = 1$), (*b*) sparsely connected network ($c = 0.2$), (*c*) test of algorithm on a set of stochastically independent random matrices, as described in the text. Other parameters used: $N = 10$, $p = 0.5$, $P = 500$, $p(x)$ Gaussian with $\sigma = 0.1$.

δ is small. Qualitatively identical results were obtained for a time series similar to Z_τ, but involving Pearson correlation coefficients instead of Kendall's τ.

In sum, those regulatory interactions that show non-zero correlations to fitness are few, their number fluctuates and they tend to show correlations only for a very short period of time. For these reasons, pairwise correlation measures do not indicate the existence of meaningful associations between individual enhancer elements and fitness. If there are only few and temporally variable correlations between individual connectivities and fitness, could there be more complex association patterns on the level of entire enhancers? Such associations could be viewed as supporting a gene selectionist viewpoint, since the enhancer (i.e. all binding sites influencing the expression of a transcription unit) most closely corresponds to the classical notion of a gene. One possible measure consists in a multiple linear regression coefficient between all binding sites on an enhancer and fitness, or, similarly, the corresponding measure of multiple correlation. Technically, the existence of some non-zero pairwise Pearson correlation coefficients is required for the calculation of these quantities (Sokal and Rohlf 1981, p 618). Given that these coefficients are few and rapidly changing, it is not likely that such quantities can even be calculated in all generations and, therefore, that a meaningful higher order linear association pattern can be found.

4.4. Variance in fitness cannot be linearly decomposed into variance components on the genic level

A predictor for the selection response of a quantitative character Z, such as fitness in this model, is the additive variance $\sigma_A^2(Z)$ of Z, which can only be predicted analytically if the genetic architecture is sufficiently simple. Here, it will be estimated as (Falconer 1981)

$$\sigma_A^2(Z) = \sigma_T^2(Z)h^2(Z) \tag{7}$$

where $\sigma_T^2(Z)$ is the total variance in fitness, which is identical to the genetic variance in this model. $h^2(Z)$ is the heritability of Z. The observation that $\sigma_A^2(Z)$ cannot be accounted

for by a linear superimposition of variance contributions of the genic level, would imply collective behaviour (Wimsatt 1980, Lloyd 1988).

Assume that a linear decomposition of $\sigma_T^2(Z)$ is possible:

$$\sigma_T^2(Z) = \mathrm{Cov}(Z, Z) = \mathrm{Cov}(Z, X_1 + \cdots + X_2) = \mathrm{Cov}(Z, X_1) + \cdots + \mathrm{Cov}(Z, X_n) \tag{8}$$

which relies on the stochastical independence of individual variables. The fraction of additive variance explained by such a linear decomposition of genetic variance is then given by

$$\frac{\mathrm{Cov}(Z, X_1) + \ldots + \mathrm{Cov}(Z, X_n)}{\sigma_T^2(Z)h^2(Z)} \tag{9}$$

using (8). $\sigma_T^2(Z)$ can then be evaluated directly as the fitness variance in the population, $h^2(Z)$ is calculated by parent–offspring regression, and the covariances $\mathrm{Cov}(Z, X_i)$ are evaluated between enhancer elements and fitness. Their estimates are based on the calculation of a Pearson correlation coefficient between connectivities and fitness. Two different approaches were pursued in evaluating (9). In the first approach, only those covariances were included in (9) that were significantly different from zero at $\boldsymbol{P} < 0.01$. The second approach used all cN^2 covariances. Results from these two approaches were qualitatively identical, and will therefore only be reported for the latter case. A time series of (9) was calculated during mean fitness evolution. Results are displayed for a densely and a sparsely connected network in figures 5(*a*) and 5(*b*), respectively. Obviously, the quantity shown fluctuates around zero and it is consistently smaller than one. This finding indicates that independent contributions of the genic level cannot account for variation in fitness.

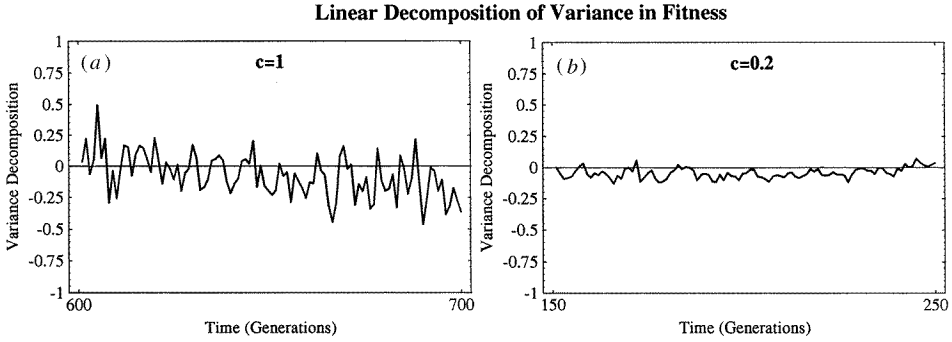

Figure 5. Superimposition of variance contributions by the genic level cannot account for variation in fitness. The quantity defined in (9) was plotted against the number of generations, t, elapsed after the start of the simulation. (*a*) Densely connected networks ($c = 1$), mean of time series is approximately -0.07, (*b*) sparsely connected network ($c = 0.2$), mean of time series is approximately -0.04. The differences between (*a*) and (*b*) in absolute generations for the time windows shown are due to differences in the speed with which the two types of network evolve. Other parameters: $N = 10$, $p = 0.5$, $P = 500$, $p(x)$ Gaussian with $\sigma = 0.1$. See text for details.

A conceivable alternative approach would have consisted in decomposing the total variance, i.e. evaluating a time series for $[\mathrm{Cov}(Z, X_1) + \ldots + \mathrm{Cov}(Z, X_n)]/\sigma_T^2(Z)$. This approach would have yielded nearly identical results, since the latter formula differs from (9) only in the factor $h^2(Z)$, and $h^2(Z)$ does not undergo large temporal fluctuations (figure 1(*b*)).

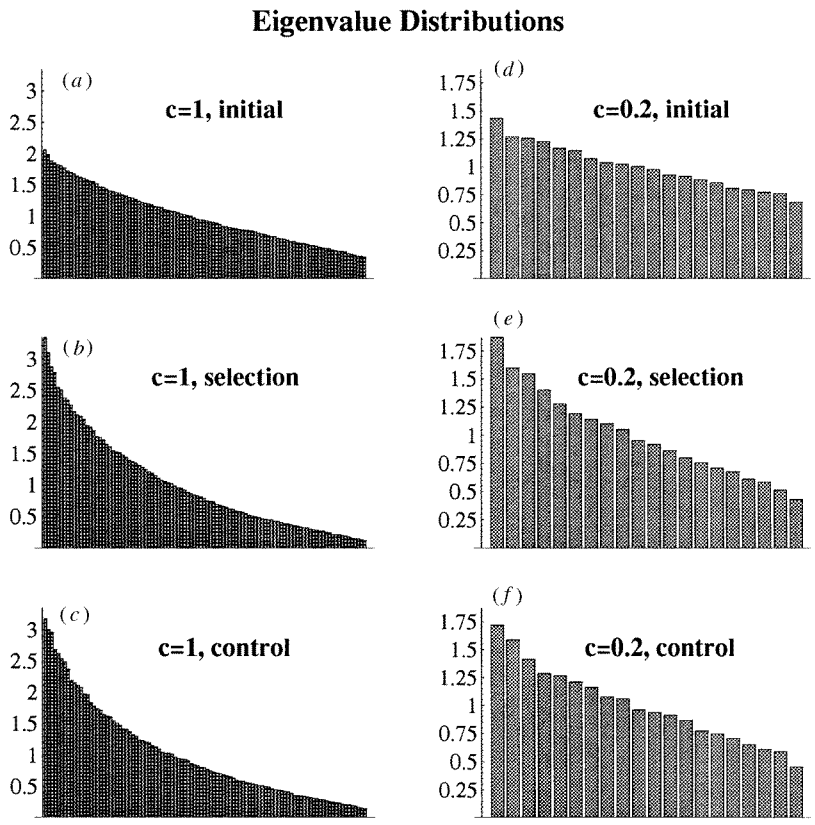

Figure 6. Low order linkage disequilibrium does not characterize optimal networks. The eigenvalue distribution of the $cN^2 \times cN^2$ sample covariance matrix of all connectivities in a population is shown. Eigenvalues are sorted by decreasing magnitude. (*a*) Densely connected networks ($c = 1$), stochastically initialized population in generation zero. (*b*) $c = 1$, population close to mean fitness one, i.e. under strong selection in mean fitness equilibrium. (*c*) $c = 1$, a population that had evolved the same number of generations as the population in (*b*), but in the absence of selection ($s = 100$). (*d*) through (*f*) are identical to (*a*) through (*c*), respectively, except that $c = 0.2$. Note that there is a smaller number of eigenvalues in these cases, since there is a smaller number of non-zero connectivities. Note also the only minor differences between (*b*) and (*c*), as well as between (*e*) and (*f*). Other parameters: $N = 10$, $p = 0.5$, $P = 500$, $p(x)$ Gaussian with $\sigma = 0.1$.

4.5. Low order linkage disequilibrium does not characterize optimal networks

In a population of networks that all arrive at a state $\vec{S}^{opt}(\infty)$ given $\vec{S}(0)$, patterns of associations between connectivities must exist. These associations along with the distributions of individual connectivities define the set of networks. The results presented so far indicate that there is no simple predictor of fitness that is based on individual connectivities, and on a linear superposition of individual connectivities. Given that, a question arises regarding the level of organization on which relevant patterns of association among connectivities can be detected, i.e. will there be associations between pairs, triples, quadruples etc of connectivities? This problem is potentially non-trivial, since there is a combinatorial explosion of the number of possible correlations between connectivities as the order of correlation is increased, i.e. there are $\binom{N}{k}$ k-tuples of connectivities and each

k-tuple might show correlation.

For technical reasons, empirical and theoretical population genetic studies restrict themselves mostly to patterns of associations between two units of genetic variation (pairwise linkage disequilibrium). For the network model considered here, the relevant information is unlikely to be contained on this low a level of organization, and it must therefore be represented by higher order correlations. To demonstrate this, the set of connectivity matrices in a population was conceptualized as a sample of the distribution of a multivariate random variable of dimension cN^2. The $cN^2 \times cN^2$ symmetric matrix of all pairwise sample Pearson correlation coefficients between connectivities was evaluated and its eigenvalue distribution was calculated. This was carried out for two different populations: first, in a population in mean fitness equilibrium with a mean fitness close to one; second, in a control population that was subject to mutation and genetic drift only. This control population had evolved independently from the population under selection, but for the same number of generations. Figure 6(a) shows the eigenvalue distribution for a population of densely ($c = 1$) connected networks at time zero, i.e. it is the eigenvalue distribution of the sample correlation matrix obtained from 500 random matrices with independently, normally distributed entries. Figures 6(b) and 6(c) show the eigenvalue distribution of populations of densely connected ($c = 1$) networks under strong selection and under mutation–drift balance, respectively. Figures 6(d), 6(e) and 6(f) are completely analogous to figures 6(a), 6(b) and 6(c), respectively, except that the networks here were sparsely connected ($c = 0.2$). No matter whether networks are sparsely or densely connected, there is little difference in the eigenvalue distribution between selected population and control population (compare figures 6(b) and 6(c), as well as figures 6(e) and 6(f)), especially with regard to the smallest and the largest eigenvalue observed in the distributions. This suggests that selection does not play an important role in the generation of second order correlations between connectivities. In the case of sparsely connected networks, this observation moreover suggests that the correlations and, thus, the information that they represent, are 'distributed' over more than one gene.

5. Discussion

In the nonlinear system used here as a dynamical model of a network of transcriptional regulators, fitness is a heritable quantitative character (figure 1). The observation that selection for an optimal gene expression pattern is possible here is a non-trivial result in and by itself. Without this prerequisite, the following, principal question of this contribution would be meaningless. Can causal relations of individual genes to fitness be discerned? In other words, to which extent is high fitness associated with particular genes or units of genetic variation? First, a conceptual remark is necessary: here, the elementary units of genetic variation are not transcription units, but the binding sites of transcription regulators on the enhancers of genes encoding for other transcriptional regulators. A whole 'gene', in terms of the concepts used here, comprises an entire enhancer with several binding sites (a row of the connectivity matrix) as well as its associated transcription unit. It will be argued that whatever the desired elementary level of variation, enhancer sites or genes, variation in fitness is not determined on that level.

That individual units of genetic variation may not be related to fitness variation in any obvious way is already suggested by the rapid divergence of the marginal distributions of *all* individual regulatory DNA elements in two populations evolving in parallel (figure 2). This result demonstrates that (i) a large number of different regulatory interaction patterns

may generate the same gene expression pattern, and (ii) the differences in these matrices are such that a variety of different values is admissible for any given enhancer site (matrix entry). Although it has to be followed by more sophisticated approaches, this is a hint towards collective behaviour: there are many degrees of freedom in the system and the strength of any individual regulatory interaction (enhancer element) is not critical to a network's function. The information necessary to characterize networks may be distributed over groups of connectivities or groups of genes.

In order to investigate which enhancer elements (regulatory interactions) in a gene network are important to the establishment of a specific gene expression pattern, one has to look for associations of the strength of individual regulatory interactions with fitness. Such measures of association were obtained here. The main observations (figures 3 and 4) are that (i) only a small fraction of all enhancer elements is associated with fitness at any given point in time, (ii) the number of such elements strongly fluctuates over time, and (iii) those elements that display association are unlikely to remain the same over time. These results cast into doubt the utility of pairwise association patterns for a delineation of the causal relations in this system. Large samples of highly accurate raw data are used in the statistical tests here. Even in the unlikely event that an experimental system could generate data at this level of resolution, one would not be able to identify any meaningful pairwise association patterns. The next higher level of complexity would consist in multiple linear regression coefficients or the corresponding measure of multiple correlation, e.g. between all binding sites of an enhancer and fitness. As already discussed, the results obtained for pairwise associations, especially their sparsity and their variability over time, suggest that such patterns do not exist. This means that also on the level of a gene (an entire enhancer) no meaningful pattern of association between DNA level variation and variation in fitness will be detectable. Would it be possible to find some nonlinear pattern of association between individual units of genetic variation and fitness? The model used here implies that fitness is some nonlinear function of the connectivity matrix. Therefore, there exists some measure of nonlinear association—represented by this function itself—between connectivities and fitness. The likely absence of a closed form for solutions of (2) (e.g. Binder and Young 1986) makes it very unlikely that simple nonlinear dependencies of few enhancer elements and fitness can be found. Therefore, one can probably not do better than measure linear association patterns in this system (as one would in experimental systems of comparable complexity). In fact, one can view the question for nonlinear association patterns as a paraphrase of the problem: which types of nonlinearity make multivariate nonlinear functions exhibit collective properties?

A further case against the possibility of decomposing variation in fitness into independent contributions of individual genetic factors is made by Wimsatt's criterion (Wimsatt 1980, Lloyd 1988). As shown in figure 5, variation in fitness is not accounted for by superimposing contributions of the genic level, implying that interaction effects of different enhancer elements, or of groups of elements are involved in its generation.

Robustness is an issue that must be addressed in models of nonlinear gene interactions. Are the results obtained robust to changes in the model? Could a real gene network be expected to behave in similar ways? Alterations in essentially all the parameters entering the model were tested, although only a small subset of the results was shown: qualitatively different distribution types for connectivities were used, and connectivity densities as well as network sizes were varied. All statistical tests, where applicable, yielded results that support the statements made above. Moreover, an alteration of the model towards a more realistic, albeit much more complicated one would probably support these statements further. The crucial simplifying assumption to make the model formally tractable was an assumption

of linearity: the effects of proteins binding to an enhancer *additively* contribute to the activation state of a gene. However, cooperation or antagonism between transcriptional regulators, although difficult to model, is likely to occur in nature (e.g. Han *et al* 1989, Lamb and McKnight 1989). It represents a deviation from this assumption of linearity, but because nonlinearity is at the root of the phenomena observed here, this deviation would most likely underscore the existence of these phenomena. The same line of reasoning holds for other possible extensions, such as diploid models or models that take the evolution of the transcription units themselves into account.

It has been claimed that genomes act as cohesive wholes in evolution and that linkage disequilibrium, i.e. patterns of association between units of genetic variation, is the primarily responsible factor. What is its role in the model analysed here? There must be patterns of association among connectivities in a population of networks (organisms) with the same gene expression pattern. Since the model used here is deterministic, these associations along with the distributions of individual connectivities completely specify such a set of networks, given $\vec{S}(0)$ and $\vec{S}(\infty)$. Linkage disequilibrium is understood as any pattern of association between two or more enhancer elements. Although this is not the traditional usage of the term, which applies to transcription units and frequently implies only pairwise correlations, it is the usage appropriate for this model. However, given the large number of variables, it is not feasible to measure high order correlations between connectivities. The results demonstrate that second order (pairwise) correlations are not very informative: a population under strong selection and a control population under mutation–drift balance have very similar eigenvalue distributions of the correlation matrix of all connectivities (see figure 6). This strongly suggests that only correlations higher than those of second order are useful in characterizing the networks. Rarely would it be possible to measure such correlations experimentally in any given system. Thus, this is a case where experiments would not have detected any amount of (relevant) linkage disequilibrium, although it must exist. It would therefore be even less likely to find linkage patterns between entire genes (enhancers). One further instructive point can be made by considering results from networks with a low number of regulatory interactions. A network with $N = 10$ genes and connection density $c = 0.2$ has binding sites for two different transcriptional regulators on the enhancer of each gene. Thus, no correlations higher than second order are needed to describe the organization of elements within any given enhancer. If contributions of individual genes to a network's function could be singled out, one might hope to detect patterns of correlations among enhancer elements within genes. The eigenvalue analysis for $c = 0.2$ does not support this possibility, since no conspicuous pairwise correlations occur. However, the cautionary remark made above regarding possible *nonlinear* enhancer-fitness associations applies also to associations among regulatory DNA elements.

What are the consequences of these findings for experimental studies of gene networks whose properties resemble those of the model used here? The study of eukaryotic enhancers is an important topic in and by itself, but caution may be necessary in the prediction of organizational network features by studying individual genes. Some of the results reported above may seem quite intuitive to the mathematically oriented biologist, given the nonlinear nature of the network model, but they may have important implications on the amount of information that can be obtained by currently common experimental techniques. Experimental alteration of individual enhancer elements (e.g. via mutagenesis or deletion) is a common strategy used to understand the role of enhancers in gene regulation, and to understand regulation in gene networks. As the above analysis shows, even in the unlikely event that one could arrive at a complete understanding of the composition and function of all the enhancers in a network, it might sometimes not be possible to describe how their

interaction makes a gene expression pattern emerge. An accumulation of genetic data on phenotypic effects of mutations in regulatory regions may not be sufficient to elucidate how a wild-type expression pattern is generated: it may be impossible to infer a simple law from studies of individual genes, if the necessary information is 'distributed' over a network's genes. Given current methods, it is hard to conceive of an experimental approach to this problem. A similar issue arises for studies of network evolution. If many degrees of freedom are involved in the evolution of gene networks, it will be difficult to draw conclusions about network organization and, thus, network function by comparing the structure of enhancers between different taxa. Even if the biochemical function and the expression pattern of a set of genes is highly conserved in evolution, the interactions that produce these expression patterns might be very different. The added complication caused by sometimes obscure relations between the structure (sequence) and function (regulatory specificity) of enhancers elements (e.g. Gehring *et al* 1994) is likely to further complicate comparative studies.

A conceptual distinction is important with regard to what has been said so far. There are instances where the expression of a gene in a body region of an embryo determines the developmental fate of that body region. Examples of such 'homeotic' genes are abundant in arthropods. They include genes encoding transcriptional regulators (Ingham 1988). The existence of such genes might be taken as evidence against the results presented here: the fact that mutations of an individual gene can be responsible for homeotic transformations of a whole body part might indicate that only that gene would be 'causal' to the formation of the part. This argument neglects, however, that it is most likely to be complex crossregulatory and autoregulatory interactions between a number of genes that lead to the *establishment* of a stable expression pattern of every one of these genes. Whether the information necessary for regional determination is distributed over a set of genes or concentrated in an individual gene may depend on the system studied. However, it must be distinguished from the question of how the gene expression pattern is established.

The possible existence of phenotypic characters, on whatever level of organization, whose variation cannot be traced to genic variation, would raise conceptual problems. What if a set of genes has two different functions during different times in the life cycle: how could pleiotropy be defined if effects of individual genes cannot be discerned in either context? Would it be likely that an organizational hierarchy exists in the epigenetic system, each level of which would be an emergent feature of the next lower level, or would there merely be overlapping sets of genes, within each of which a task would be distributed over the whole set? What would happen at the intersection of such sets? Finally, if there were characters whose variation was most efficiently described on a high level of epigenetic organization, whereas in others a description in terms of the genic level was more efficient, how would selective forces have to be partitioned between these different levels?

Complex, nonlinear processes with multiple components are abundant in biological systems: networks of enzymatic reactions, cross-talking signal transduction pathways with complicated feedback mechanisms, and intercellular communication processes in development are but three examples. Whether such systems cannot be described efficiently on the level of their individual parts can only be decided by experiment. One has to be aware, though, that collective properties of the system analysed here may represent the rule rather than the exception.

Acknowledgments

I would like to thank M D Laubichler and G P Wagner, who first raised my interest in the issues addressed here. They, as well as M J Blanco, L W Buss, B Misof, C Pazmandi, J Stites

and R Vaisnys contributed significantly to this work by providing constructive criticism and invaluable suggestions on earlier drafts of the manuscript. This paper is communication Number 11 from the Center for Computational Ecology at the Yale Institute for Biospheric Studies.

References

Amit D J 1989 *Modelling Brain Function.* (Cambridge: Cambridge University Press)

Binder K and Young A P 1986 Spin glasses: experimental facts, theoretical concepts and open questions *Rev. Mod. Phys.* **58** 801–976

Boyd L H and Iversen G 1982 *Contextual Analysis* (Belmont, CA: Wadsworth)

Brandon R 1982 *The Levels of Selection.* PSA 1982, volume 1, 315–22 (East Lansing, MI: Philosophy of Science Association)

Buss L W 1987 *The Evolution of Individuality* (Princeton, NJ: Princeton University Press)

Chisaka O and Capecchi M R 1991 Regionally restricted developmental defects result from targeted disruption of the mouse homeobox gene hox-1.5 *Nature* **350** 473–9

Dawkins R 1976 *The Selfish Gene* (New York: Oxford University Press)

—— 1978 Replicator selection and the extended phenotype *Z. Tierpsychol.* **47** 61–76

—— 1982 *The Extended Phenotype* (San Francisco: Freeman)

Falconer D 1981 *Introduction to Quantitative Genetics* (London: Longman)

Gehring W J, Affolter M and Bürglin T 1994a Homeodomain proteins *Ann. Rev. Biochem.* **63** 487–526

Goodnight C J, Schwartz J M and Stevens L 1992 Contextual analysis of models of group selection, soft selection, hard selection and the evolution of altruism *Am. Nat.* **140** 743–61

Haken H 1977 *Synergetics: an introduction* (New York: Springer)

Han K, Levine M S and Manley J L 1989 Synergistic activation and repression of transcription by Drosophila homeobox proteins *Cell* **56** 573–83

He X, Treacy M N, Simmons D M, Ingraham H A, Swanson L W and Rosenfeld M G 1989 Expression of a large family of POU-domain regulatory genes in mammalian brain development *Nature* **340** 35–42

Hull D L 1980 Individuality and selection *Ann. Rev. Ecol. Syst.* **11** 311–32

Ingham P W 1988 The molecular genetics of embryonic pattern formation in Drosophila *Nature* **335** 25–32

Johnson P F and McKnight S L 1989 Eukaryotic transcriptional regulatory proteins *Ann. Rev. Biochem.* **58** 799–839

Kappen C, Schughart K and Ruddle F H 1989 Two steps in the evolution of Antennapedia-class vertebrate homeobox genes *Proc. Natl. Acad. Sci. USA* **86** 5459–63

Kauffman S A 1969 Metabolic stability and epigenesis in randomly connected nets *J. Theor. Biol.* **22** 437

Kauffman S A 1993 *The Origins of Order* (New York: Oxford University Press)

Kirkpatrick S, Gelatt C D Jr and Vecchi M P 1983 Optimization by simulated annealing *Science* **220** 671–80

Lamb P and McKnight S L 1991 Diversity and specificity in transcriptional regulation: the benefits of heterotypic dimerization *Trends in Genetics* **16** 417–22

Lewontin R C 1970 The units of selection *Ann. Rev. Ecol. Syst.* **1** 1–18

—— 1974 *The Genetic Basis of Evolutionary Change* (New York: Columbia University Press)

Lloyd E A 1988 *The Structure and Confirmation of Evolutionary Theory* (Westport, CT: Greenwood Press)

Lufkin T, Dierich A, LeMeur M, Mark M and Chambon P 1991 Disruption of the Hox-1.6 homeobox gene results in defects in a region corresponding to its rostral domain of expression *Cell* **66** 1105–19

McGinnis W and Krumlauf R 1992 Homeobox genes and axial patterning *Cell* **68** 283–302

McGinnis, N, Kuziora M A and McGinnis W 1990 Human hox-4.2 and Drosophila deformed encode similar regulatory specificities in Drosophila embryos and larvae *Cell* **63** 969–76

Mermelstein F H, Flores O and Reinberg D 1989 Initiation of transcription by RNA Polymerase II *Biochim. Biophys. Acta* **1009** 1–10

Mjolsness E, Sharp D H and Reinitz J 1991 A connectionist model of development *J. Theor. Biol.* **152** 429–53

Olson E 1990 MyoD family: A paradigm for development *Gene. Dev.* **4** 1454–61

Regulski M, Dessain S, McGinnis N and McGinnis W 1991 High-affinity binding sites for the Deformed protein are required for the function of an autoregulatory enhancer of the Deformed gene *Genes Dev.* **5** 278–86

Reinitz J and Sharp D 1995 Mechanism of formation of eve stripes *Mech. Dev.* **49** 133–58

Reinitz J, Mjolsness E and Sharp D 1995 Cooperative control of positional information in Drosophila by bicoid and maternal hunchback *J. Exp. Zool.* **271** 47–56

Rosenfeld M G 1991 POU-domain transcription factors: pou-er-ful developmental regulators *Genes. Dev.* **5** 897–907

Schüle R and Evans R E 1991 Cross-coupling of signal transduction pathways: zinc finger meets leucine zipper *Trends in Genetics* **7** 377–81

Sokal R R and Rohlf F J 1981 *Biometry* (New York: Freeman)

Stein D L (ed) 1989 *Lectures in the sciences of complexity* (Redwood City, CA: Addison-Wesley)

Sucov H M, Murakami K K and Evans R M 1990 Characterization of an autoregulated response element in the mouse retinoic receptor type β gene *Proc. Natl. Acad. Sci. USA* **87** 5392–6

Turelli M 1985 Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits *Genetics* **111** 165–95

Wagner A 1994 Evolution of gene networks by gene duplications: a mathematical model and its implications on genome organization *Proc. Natl. Acad. Sci. USA* **91** 4387–91

—— 1996 Does evolutionary plasticity evolve? *Evolution* at press

Wagner G P 1990 The domestication of replicators: a neo-Darwinian commentary on the concept of replicator selection *Evolutionary Trends in Plants* **4** 71–3

Weinzierl R O J, Dynlacht B D and Tjian R 1993 Largest subunit of Drosophila transcription factor IID directs assembly of a complex containing TBP and a coactivator *Nature* **362** 511–7

Williams G C 1966 *Adaptation and Natural Selection* (Princeton, NJ: Princeton University Press)

—— 1992 *Natural Selection: Domains, Levels and Challenges* (New York: Oxford University Press)

Wimsatt W C 1980 *Units of Selection and the Structure of the Multi-level Genome* PSA 1980, vol 2, pp 122–83 (East Lansing, MI: Philosophy of Science Association)

Wright S 1968 *Evolution and the Genetics of Populations* vol 1 (Chicago, IL: University of Chicago Press)

Zeng Z-B and Cockerham C C 1993 Mutation models and quantitative genetic variation *Genetics* **133** 729–36