

On the Unit of Selection in Sexual Populations

Richard A. Watson

School of Electronics and Computer Science, University of Southampton, SO17 1BJ. U.K.
raw@ecs.soton.ac.uk

Abstract. Evolution by natural selection is a process of variation and selection acting on replicating units. These units are often assumed to be individuals, but in a sexual population, the largest reliably-replicated unit on which selection can act is a small section of chromosome – hence, the ‘selfish gene’ model. However, the scale of unit at which variation by spontaneous mutation occurs is different from the scale of unit at which variation by recombination occurs. I suggest that the action of recombinative variation and mutational variation together can enable local optimization to occur at two different scales simultaneously. I adapt a recent model illustrating a benefit of sexual recombination to illustrate conditions for two scales of optimization in natural populations, and show that the operation of natural selection in this scenario cannot be understood by considering either scale alone.

1 Nucleotides, genes, and individuals

Although it is often convenient when providing evolutionary explanation to suppose that selection acts on individual organisms, in fact, in sexual populations the combination of alleles represented in an individual’s genotype is not reliably transferred to its offspring [1][2]. In sexual populations the largest unit of genetic material that reproduces with reliable fidelity is a subsection of chromosome small enough to avoid being disrupted by crossover. The size of these units will be determined by the crossover rate¹, with higher rates defining smaller units, but for common purposes it is taken that the relevant unit is about the size of individual genes [1]. Accordingly, since the gene is the largest genetic unit that reproduces reliably, the gene is the largest unit on which natural selection can act [1] – hence, the well-known “selfish gene” framework [2]. These observations support fundamental axioms underlying the way evolution is defined, i.e. the change in frequencies of alleles in a gene pool [3], placing attention on the frequencies of individual alleles, not the frequencies of genotypes. Although the selective unit might seem unambiguous in evolutionary algorithms because evaluation is always applied to individuals and individual fitnesses determine reproduction, in fact the same issues apply if sexual recombination, or crossover, is used. Although *individuals* are selected to reproduce, the genetic material of individuals is broken-up by crossover, so it is only *fragments* of individuals whose frequencies can be affected by selection – hence, attention on “schema” in evolutionary computation theory [4].

¹ The probability of recombining adjacent loci (not the probability that crossover is applied to an individual, as sometimes meant in evolutionary computation). See *C* in Table 1.

Biologically, the genes on which selection acts may consist of thousands of nucleotides. New variants of genes (new alleles) are introduced by spontaneous point mutation affecting one or a small number of individual nucleotides within a gene. Accordingly, the scale of unit that is manipulated by mutation (individual nucleotides) is quite a different scale of unit from that which is manipulated by recombination (alleles of genes). However, since selection acts on whole alleles it is common in population genetics models to abstract away the nucleotide-level details and simply refer to the variant alleles of a gene by unique labels, e.g. A and a . Each allele label represents a different combination of maybe thousands of nucleotides, but if one allele is produced by mutation of the other, then they may differ in only a few nucleotide substitutions. This abstraction of the internal detail, hiding the level of individual nucleotides on which mutation acts, is appropriate for some purposes. Indeed, it sufficed perfectly well for all the population genetic results derived prior to the discovery of the molecular structure of DNA [5]. Historically, an allele is simply defined as a particulate unit of Mendelian inheritance (thus being intimately linked to the action of recombination), and the fact that, on a molecular basis, each consisted of thousands of nucleotides was not known. Following conventional population genetic models, evolutionary computation models [4] used in artificial life rarely make the distinction between genes and nucleotides. Individuals are generally modelled as binary strings where each bit is taken to be synonymous with the allele of a gene at a particular locus and mutation changes the bit between one allele and the other, 0 or 1, standing in for A and a . The problem is that this abstraction is not consistent with a model of a gene containing a thousand nucleotides and 4^{1000} different alleles – consider the probability of a reversion, or “back mutation”, for example, or more generally, the probability of finding a particular allele by random mutation. This inconsistency can cause more than a mere terminological problem, especially in cases where mutation and recombination are applied together.

So, in population genetics and in evolutionary computation, does it matter whether genes are modelled as collections of nucleotides or simply abstracted into particulate alleles? There are many scenarios, lying within common assumptions, where it does not. But there are other scenarios that are biologically plausible where it does matter. The aim of this paper is to discuss the conceptual issues involved and examine the implications of moving outside common assumptions. I use some simulation results as an example just to illustrate some of the salient points. My claim is that existing definitions of the unit of selection and common simplifying assumptions preclude some interesting phenomena, and more specifically, that it is necessary in some cases to recognise more than one level of optimisation to understand the action of evolutionary systems. Cases (like those examined here) that step outside the usual simplifying assumptions are necessary for understanding the operation of evolution in natural populations. They are also necessary for understanding how to use and model evolution in artificial life experiments, and in addressing the complex relationship that these observations have to concepts of the unit of selection and related processes such as Shifting Balance Theory [6].

The following two sections introduce the basic ideas about why two levels of optimisation may be required to find fit genotypes in a fitness landscape, and how this might be provided. Section 4 describes a set of simulation experiments to illustrate the

effect of changing the scales at which optimisation is applied. Specifically, I show that optimisation at any one scale is insufficient to find fit genotypes, and a combination of optimisation scales is required. I illustrate two different ways in which these two scales can be provided, but the main idea is that mutation manipulates the nucleotides and recombination manipulates the alleles of genes. The quite simple idea that mutation and recombination afford local optimisation in different spaces by operating on different units motivates a rethink about simplistic notions of the unit of selection.

2 Background: selective units and local optima

I suggested above that there are evolutionary scenarios that afford optimisation on some unit of genetic material and also on combinations of that unit: specifically, nucleotides and combinations of nucleotides (i.e. alleles of genes). The consequences of this are the same for other scales of units, for example, alleles and combinations of alleles. In fact, discussion about the possibility of selection on combinations of alleles has quite a history in population genetics and we can use this to better understand the implications of having optimisation occur at the lower scales of nucleotides and combinations of nucleotides.

If selection acts on individual units (of whatever scale), then this has important consequences for what evolution by natural selection can and cannot do. If selection acts on individual units, then evolution can only respond to the net fitness effects of individual units. Although a piece of genetic material like an allele may have different fitness effects in different genetic backgrounds, the change in frequency of that unit over several generations cannot be controlled by its fitness in any particular background in a sexual population (as its presence in a given background is not reliably reproduced). Instead, Fisher argues, the change in frequency of an allele will be controlled by its average *fitness excess* [7], roughly, its average fitness effect over the backgrounds in which it occurs. A consequence of selection on the average effect of individual units is that if a particularly fit *combination* of units involves units that are *individually* disfavoured then, even if that combination should happen to arise in some individual, it cannot take hold in the population.

A particular instance of this may occur when a population is stuck at a local fitness peak in a fitness surface. Consider a genotype *A* that is locally optimal and a second genotype *B* that is fitter. All single allele substitutions to *A* are deleterious (by the definition of being locally optimal) so *B* (necessarily) differs from *A* by several allelic substitutions. These allelic differences are thus collectively favourable but individually unfavourable to a genotype at *A*. Since sexual recombination re-assorts alleles into the different genetic backgrounds of other individuals in the population, even if such a combination of alleles were introduced to an individual in a population located at *A*, selection would act on the alleles individually and remove them from the population. Sewall Wright considered the escape from a local fitness peak to a genotype of higher fitness to be the central problem of evolution [8], and devised Shifting Balance Theory, SBT, [6] to explain how population subdivision might enable an evolving population to achieve this. However, some consider the conditions for SBT to not be widely, if at all, available in natural populations [9].

Actually, it would be easy to facilitate selection on combinations of alleles if the crossover rate could be modified. That is, if the crossover rate were very low then alleles would not assort independently and would instead be selected as a unit, and in the limit, the entire genotype of an *asexual* individual replicates as a whole. If an asexual population was centred on *A*, and the combination of alleles required to reach *B* were introduced to an individual in the population, then selection would have no problem in promoting the resultant *B* genotype to take over in the population. But selection on combinations of alleles has its disadvantages as well. In some circumstances it can be favourable to have selection act on the average fitness effect of individual alleles rather than on combinations of alleles. For example, when a sub-optimal combination of alleles arises in a genotype, containing some favourable alleles and some disfavoured alleles, an asexual population is unable to promote the good alleles without also promoting the bad alleles. This is the basis of the classic Fisher/Muller model for the benefit of sexual recombination [7]. In short, in some circumstances it is preferable to have selection act on individual units and in others it is preferable to have selection act on combinations of units. The former cannot select for good combinations of units unless they involve only units which are also individually good, and the latter cannot select for good individual units unless they are in a collectively good combination of units.

Fisher and Wright were concerned with the action of selection on alleles and combinations of alleles, but here we are concerned (for the most part) with nucleotides and combinations of nucleotides (alleles). The same reasoning applies: in some cases selection on individual nucleotides may be beneficial, and in others selection on alleles may be beneficial. In natural populations, it does not make sense to imagine that recombination rates are so high that selection acts on individual nucleotides. But I will argue that mutation effects local optimisation in nucleotide sequence space, and that if, at the same time, recombination is manipulating whole alleles, then this can effect a two-level process of optimisation. Such a scenario is precluded in prior models by common assumptions about the epistasis model, and also by the abstraction of nucleotide combinations into particulate alleles. In the next section I clarify what I mean by local optimisation and the different ways that evolving populations can provide it.

3. Modes of local optimisation

Given a combinatorial search space $\langle u_1, u_2, u_3, \dots, u_N \rangle$ (where each point in the space is a combination of values for each of the units u_1 to u_N) local optimisation involves the movement of a point through this space where successive points are neighbouring or nearby. For example, *hill climbing* optimisation follows trajectories formed by moving to adjacent points that are higher in fitness. An evolving population is often conceived as a hill climbing process. However, there are two qualitatively different means by which this may be provided in evolution by natural selection (one being more common in evolutionary computation and the other being more common in population genetics) and these are often not distinguished properly: I call these *local mutation* and *direct selection*. Fig.1a. illustrates local mutation as provided by, for example, an asexual population with a low rate of spontaneous point mutation on nucleotides. Here, al-

though selection promotes entire genotypes, mutation modifies one or a small number of nucleotides (e.g. third locus) enabling only local movement in the combinatorial space of nucleotides. This maps very naturally to bit-mutation hill climbers used in computational optimisation. Fig.1.b. illustrates local optimisation via direct selection on individual units as provided by, for example, a sexual population that exhibits free recombination between alleles (that is, alleles assort independently during sexual reproduction). In this case, an evolving population can (with some simplifying assumptions about linkage equilibrium [10] and epistasis) be usefully described as a point in allele frequency space [8]. If selection acting on individual units makes small adjustments to the frequencies of these units, then the population has performed local movement in this space. This sense of local optimisation is the normal interpretation of evolution in population genetics models (though not with all the same terminology).

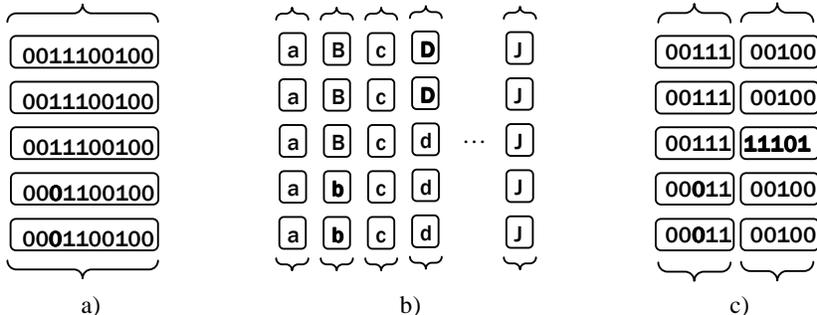


Fig. 1. Rows are individuals in a population; boxes show units of genetic material; braces show selection pools. a) local mutation. b) direct selection. c) see text.

Formally, local mutation and direct selection move in different spaces, genotype sequence space and allele frequency space, respectively. But note that genotype sequence space is coincident with the vertices of allele frequency space (i.e. points where frequencies are either 1 or 0). The more important difference is that the units in the natural interpretation of local mutation (in the post-molecular-genetics era) are nucleotides, whereas the units in the natural interpretation of direct selection are alleles (Mendelian units of inheritance). If the units that the two mechanisms manipulate are not the same scale then this makes a profound difference in the ‘locality’ of the spaces in which this optimisation occurs. Although local optimisation at any scale can become stuck on local optima at that scale, local optimisation at different scales ‘sees’ different fitness landscapes, different fitness gradients, and different local optima.

Unlike Fig. 1b where alleles are abstract, Fig. 1c illustrates alleles that are each composed of many nucleotides. Here if variation introduces new alleles that differ by only one nucleotide (left of Fig. 1c) then we still have local optimisation at the nucleotide scale. In contrast, if variation introduces completely new alleles (right of Fig. 1c), this effects local optimisation at the allele scale but not at the nucleotide scale. A combination of mutation on nucleotides and recombination of whole alleles has the potential to provide optimisation at two different scales simultaneously, and when the local optima of one scale are different from those of the other scale, the interaction of the two can provide optimisation that neither one scale can provide alone.

4. Examining a model landscape

The consequences of optimisation at more than one scale will only be seen if we remove common simplifying assumptions about epistasis that would cause local optima at different scales to be coincident. Previous work [11] described a fitness landscape designed to illustrate a benefit of sexual recombination that is well suited for our purposes here. This work showed that a subdivided sexual population can discover the highest fitness genotypes of this landscape easily, but an asexual population (under the same conditions) cannot. In this paper I use this model to demonstrate that there are two levels of optimisation involved in this effect by explicitly examining the local mutation and direct selection mechanisms at the levels of nucleotides and alleles.

In this model a genotype consists of two genes each containing many nucleotides. The epistasis in the model has strong ‘synergy’ of good mutations within a gene, and also a general field of random epistatic interactions among all mutations. The fitness of a genotype is given by:²

$$f(G)=R_{ij}(2^i+2^j) \quad (1)$$

where i is the mutational (Hamming) distance of one of the genes (the first half of the genotype, $\{g_1, g_2, \dots, g_n\}$) from an ideal allele, and j is the mutational distance from an ideal allele in the other gene (the second half of the genotype $\{g_{n+1}, g_{n+2}, \dots, g_{2n}\}$), and each R_{ij} is a random value drawn uniformly in the range $[0.5, 1]$ (for one instance of this random landscape, each R_{ij} is a constant). The basic form of modularity used here, where genes are constituted by a large number of nucleotide sites that are grouped both functionally (with epistasis) and physically (by location on the chromosome), is also seen in natural systems where the nucleotides of a gene are grouped functionally and physically by virtue of the transcription and translation machinery. Without loss of generality, the maximum fitness allele for each gene can be that where all nucleotides are 1s, and the maximum fitness genotype is that where both genes have their maximum fitness alleles, i.e. the all-1s genotype. This function can be conveniently drawn as a two dimensional fitness landscape where the two axes are the number of 1s in each of the two genes (Fig. 2).

The idea behind this model is that mutation will search locally to find good combinations of nucleotides within each gene, and crossover will make new combinations of alleles to bring these the two good alleles together and hence find fit genotypes. This is a simple idea but it is worth pointing out that in most cases over-simplistic assumptions about epistasis preclude the need for a two-level description of the evolutionary process, even if the two processes are available. That is, in simple landscapes, if selection can find good alleles by finding good nucleotide mutations, then fit genotypes can be found by simply doing more of this, i.e. finding the good nucleotide mutations in all genes. However, in this landscape, it is easy to find the best allele for gene 1 by accumulating beneficial mutations only when gene 2 is not yet well optimised, and vice versa. Once either of the genes becomes well optimised it then becomes difficult to find the best allele for the second gene without disrupting the fitness contributions

² This is modified to keep the maximum fitness value to be 3 (and make it insensitive to n , the number of nucleotides per gene) but the original shape of the function is retained.

of the first. This can be seen by considering the problem of escaping the local optima in the fitness ridges shown at the back of Fig. 2 (left). Specifically, each local optimum shown is a local optimum in nucleotide sequence space, so although it is easy to find a good combination of nucleotides for either gene by selecting on nucleotides, it is not the case that continued selection on individual nucleotides will find the fittest genotypes.

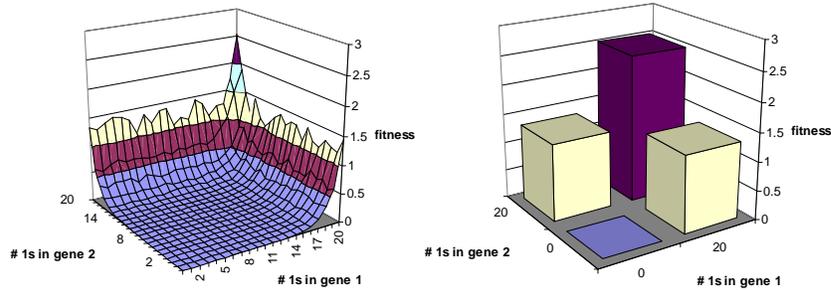


Fig. 2. The fitness landscape defined by Equation 1.

(left) as seen by nucleotide variation, (right) as seen by combinations of good and bad alleles.

Thus local optimisation on nucleotides will easily find a good allele for one of the genes but will then be stuck on a local optimum. In a subdivided population different demes may find individuals on different local peaks, some having optimised gene-1 (but not gene-2) and others having optimised gene-2 (but not gene-1). But that is as far as it goes with optimisation only on nucleotides – to the nearest peak in nucleotide sequence space (Fig. 2. left). Additional selection on the allele scale, (taking the good alleles from different demes), can easily find fit genotypes. There are only two genes, and the best genotypes are simply the union of best alleles from these two genes (Fig. 2 right) – so local optimisation in allele space will easily find good genotypes if good alleles for these two genes are provided. However, if there were no nucleotide-scale optimisation there would be no good alleles on which this allele-scale optimisation could operate. Specifically, the number of possible combinations of nucleotides in a gene is exponential in the number of nucleotides it contains, and if fit alleles of the gene are rare then neither initial standing variation nor spontaneous mutation can be guaranteed to provide these fit alleles. In this example, the best alleles for each gene are unique in a space of 2^n possible alleles for each gene. For large n , without local optimisation in nucleotide space, finding them relies on chance and is infeasible. Local optimisation on alleles alone will only be able to select on the best alleles that are in the initial population or any better alleles that might be provided by random search in the set of alleles for each gene. Thus although optimisation at the scale of alleles is not troubled by local optima in nucleotide sequence space, it also cannot exploit local search in nucleotide space as is required to find fit alleles.

To examine the above reasoning the simulation experiments use the parameters given in Table 1. To maintain population diversity the population was subdivided with a total population of 10,000 individuals, subdivided into 100 demes of 100 individuals each. (These demes are bigger than those used in [11] so as to preclude the need to

use elitism which would explicitly create a means for whole genotypes to reproduce without modification and confound these studies). Migration between demes was such that one individual in each new generation in each deme was a migrant from some other randomly selected deme. Each sub-population independently creates a new generation by fitness proportionate selection (with replacement) [4]. Each individual in each deme is initialized to a random binary string of 100 bits – representing two genes of 50 nucleotides each. In the crossover methods, crossover is applied to all reproduction events. The data recorded are the number of generations until the first occurrence of the fittest genotype, and also the number of generations for the fittest alleles to arise in both genes (numbers in brackets give the standard deviation). 30 independent runs of each parameter set were performed. Runs were terminated at a limit of 2000 generations.

In Table 1, $L=100$ is the length of the genotype in bits. Mutation values are the probability that each bit of the genotype is replaced with a new random bit. ‘NRA’=New Random Allele, i.e. a randomly generated combination of nucleotides for a whole gene (this is equivalent to *genewise* crossover with a random string, and thereby confirms that the success of crossover in experiment 3(b) is not the result of “macromutation” [12]). The crossover value, C , is the probability that a crossover point occurs between adjacent loci. $C=0.5$ is free recombination, or uniform crossover. ‘Genewise’ means that crossover occurs only between the two genes and never within them (given that the coding regions of biological genes are very small compared to the intergenic distances, this scenario is actually the most biologically realistic). Finally, $C=1/L$ produces on average one crossover point per reproduction but, unlike *genewise* crossover, its location is random.

Experiment	Mutation	Crossover	Generations to find:	
			~ best alleles	~ best genotype
1.a)	$1/L$	0	36 (2)	> 2000
b)	0	0.5	9 (1)	> 2000
2.a)	NRA	0	> 2000	> 2000
b)	0	genewise	> 2000	> 2000
3.a)	$1/L$	genewise	36 (2)	50 (4)
b)	0	$1/L$	90 (25)	181 (32)

Table 1: parameters and results of simulation experiments.

Exp. 1 examines optimisation on the scale of nucleotides only, a) provided by local mutation on individual nucleotides (with selection on whole genotypes, i.e. $C=0$), b) provided by direct selection on individual alleles (i.e. $C=0.5$). Exp. 2 examines optimisation on the scale of whole alleles only (50 nucleotides each), a) provided by a ‘local’ mutation model that creates new random alleles (with selection on whole alleles via *genewise* crossover), b) provided by direct selection on standing variation (also via *genewise* crossover). Exp. 3 examines optimisation on the scales of nucleotides and alleles simultaneously, a) uses a combination of mechanisms from 1a and 2b, i.e. local mutation on individual nucleotides and direct selection on alleles. Exp. 3b uses direct selection at two scales provided by selection on variable sized sections of chromosome using a low per locus crossover probability (initial standing variation provides the necessary nucleotides).

Table 1 also gives the results of these simulations. In Exps. 1 and 2 all 30 runs failed to find the best genotype in 2000 generations. Exps. 1 and 3 succeeded in finding the best alleles for both genes in all 30 runs, but only Exp. 3 succeeds in finding the best genotypes in any runs, and in fact finds them in all runs. As hypothesised we see that optimisation on the scale of nucleotides only can find fit alleles but not fit genotypes, optimisation on the scale of alleles only cannot find fit alleles, and optimisation on both nucleotides and alleles is necessary and sufficient to find fit genotypes. Unsurprisingly, Exps. 1 and 2 are no better in other simulations using a panmictic population, failing in all 30 runs, as does Exp. 3b. in 28 of the runs. But interestingly, Exp. 3a. does succeed in 14 of 30 runs with a panmictic population (mean 71 (98)) indicating that the two-scale optimisation is not entirely dependent on population subdivision. Exp. 3b is interesting because it uses recombination to provide optimisation at both scales.³ This appears to be dependent on the subdivision model and further examination is required to ascertain its similarities and contrasts with SBT. Comparing with the simulations performed in [11] which use a recombination rate of $1/L$ (as in Exp. 3b) but also use mutation at a rate of $1/L$ (as in 3a), Exps. 3a and 3b separate out the mechanisms that might be responsible for the result in [11].

5. Conclusions

The simulations show that, in a scenario like the example modelled, two scales of optimisation are in operation and are required to find fit genotypes. What does this tell us about the relevant units of selection? Arguably, in Exp. 3 the unit of selection is still the gene because this is the largest unit that replicates reliably under recombination at these rates. But it would be a mistake to conclude that this is the whole story. Let us consider further what has historically been the main purpose of defining the unit of selection: i.e. to identify the unit that adaptation by natural selection acts in the interest of? A lot of relevant discussion has been focussed on evolution for the good of the group versus evolution for the good of the individual [1], but also on evolution for the individual versus evolution for the gene [2]. Should we add to this list the question of evolution for the good of the gene versus evolution for the good of the nucleotide? Although these are the two scales in question here, ascribing benefit to a nucleotide does not seem conceptually useful to me. Moreover, it is not clear to me that this is the right question to be asking, or that it has a sensible answer. In contrast, I find it relatively unambiguous to state that evolution is performing local optimisation at the nucleotide scale and at the allele scale. Moreover, Exp. 3b shows that this could in principle apply at other scales such as genes and combinations of genes, and that it would be incorrect to insist that any one scale was sufficient to understand the process.

Accordingly, a preoccupation solely with the unit defined by the recombination rate does not necessarily capture all the important scales in evolutionary processes. In particular, sexual recombination and spontaneous mutation may provide different

³ This means that multi-scale optimisation is not necessarily restricted just to nucleotides and alleles, but in principle to alleles and combinations of alleles, although the utility of optimisation at the scale of combinations of alleles would depend on the structure of the genetic map and its correspondence with epistatic interactions [13].

levels of optimisation. This suggests that it is not always appropriate to abstract the combinations of nucleotides within the alleles of genes into indivisible units simply because this is the unit that is particulate under crossover. Simplifying assumptions about epistasis may preclude the necessity for such distinctions, but biologically plausible epistatic structures are complex, and simplistic models may overlook significant structure that makes these distinctions important, as shown in the example model. Similarly, simplifying assumptions about population structure (like panmixia), and recombination models (like uniform crossover), are also self-reinforcing in that they each exclude the phenomena that make the other interesting. Natural populations often lie outside these simplifying assumptions about sex, population structure, and epistasis and require a more sophisticated treatment of the mechanisms involved.

More generally, these observations challenge our understanding of the underlying algorithmic principles of evolution by natural selection – in this example we cannot model the action of evolution as a hill-climbing process that operates at any one scale. Two-scale optimisation is more closely allied to a divide and conquer process of problem decomposition [13]. Such a distinction is implied in the evolutionary computation literature on the building block hypothesis [4][14][15]. Issues of selection on parts and wholes are also important to artificial life in understanding mechanisms that scale-up the processes of evolution [16], and to understand when evolutionary processes can do more than simple hill-climbers can [13].

References

1. Williams, G.C., 1966, *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton University Press, Princeton, N.J.
2. Dawkins, R., 1976, *The Selfish Gene*, Oxford University Press, NY.
3. Hardy, G.H., 1908. Mendelian proportions in a mixed population, *Science*, **28**, 49-50.
4. Holland, J.H., 1975, *Adaptation in Natural and Artificial Systems*, Ann Arbor, MI: The University of Michigan Press.
5. Watson, J.D. & Crick, F.H., 1953, Molecular structure of Nucleic Acids, *Nature* **171**, 737-738.
6. Wright, S., 1977, *Evolution and the Genetics of Populations, Volume 3*, U. of Chicago Press.
7. Fisher, R.A., 1930. *The genetical theory of natural selection*. Dover publisher, Inc., New York.
8. Wright, S., 1931, Evolution in Mendelian populations, *Genetics* **16**: pp. 97-159.
9. Coyne, J.A., Barton, N.H., and Turelli, M., 2000, Is Wright's Shifting Balance Process Important in Evolution, *Evolution* **54**(1), pp.306-317.
10. Ridley, M. 1996. *Evolution*, 2nd ed. Cambridge, MA: Blackwell
11. Watson, R.A., 2004, A Simple Two-Module Problem to Exemplify Building-Block Assembly Under Crossover, in *Parallel Problem Solving from Nature (PPSN VIII)*, Xin Yao, et al (Eds.), LNCS 3242, Springer, pp 161-171.
12. Jones, T., 1995, *Evolutionary Algorithms, Fitness Landscapes and Search*, PhD dissertation, 95-05-048, University of New Mexico, Albuquerque.
13. Watson, R.A., 2006, *Compositional Evolution: The Impact of Sex, Symbiosis, and Modularity on the Gradualist Framework of Evolution*. MIT Press, Cambridge, MA. (in press)
14. Goldberg, D.E., 1989, *Genetic Algorithms in Search, Optimization and Machine Learning*, Reading Massachusetts, Addison-Wesley.
15. Forrest, S., & Mitchell, M., 1993a, Relative Building block fitness and the Building block Hypothesis, in *Foundations of Genetic Algorithms 2*, Whitley, D, ed., Morgan Kaufmann.
16. Maynard Smith, J.M. & Szathmary, E., 1995, *The Major Transitions in Evolution*, Freeman.