

See How She Runs: Towards Visualising Artificial Red Queen Evolution

James A. R. Marshall¹ and Simon Tokumine²

¹Centre for Behavioural Biology, Department of Computer Science, University of Bristol, BS8 1UB, UK.

²Department of Computing, Imperial College London, SW7 2AZ, UK.

marshall@cs.bris.ac.uk

Abstract

Since Darwin originally proposed that one of the primary drives in evolution towards increased complexity might be biotic competition, the search for examples of this phenomenon has been a major objective for students of evolution. The Red Queen hypothesis proposes that this competition is never-ending; as an organism or species improves in some manner relative to its competitors, these competitors are also improving. This results in an evolutionary “arms race” in which neither side maintains its advantage for long. Examples of the Red Queen effect from the natural world can be found, but it is perhaps in artificial evolutionary systems that there is the greatest scope for gaining a deeper understanding of the nature of evolutionary change under such circumstances, whether gradualistic or punctuational for example. Here we present a model of co-adaptation and introduce a visualisation method for tracking changes in the fitness landscape and the population’s location on it. This enables the Red Queen’s endless run within the model to be visualised.

Introduction

It was Darwin that first proposed, with his metaphor of the wedge, that competition between species could drive evolution to create ever more complex adaptations [5]. Since then naturalists and biologists have found many examples of this in the real world, such as brood parasitism [8] to name but one. More recently this idea has been refined and dubbed the “Red Queen” effect [17], or “evolutionary arms race” [6], as adaptations of a species or organism provide it with no net benefit because its competitors are continually keeping up with adaptations of their own. Like the character in Lewis Carroll’s “Through the Looking Glass”, or nuclear-armed superpowers, species and organisms have to run continually just to maintain their position.

It is perhaps in artificial life models of evolution that the greatest potential for gaining deeper understanding of the Red Queen effect lies. As well as confirming that the effect can occur in principle, it is of interest to consider the different evolutionary modes with which it might operate. For example, does Red Queen evolution typically proceed in a gradualistic manner, as traditional Darwinism would expect, or in punctuated bursts of evolutionary activity between intervening periods of stasis [7]? Similarly, can populations be observed to descend fitness gradients, as expected by Wright’s “shifting balance theory” [18]?

Earlier approaches to visualising artificial evolutionary or co-evolutionary systems have included measurement of

co-evolutionary progress in populations [3], observation of the dynamics of co-evolution in ecosystems [10], and visualising fitness landscapes as sampled by populations [1, 9]. In this paper we take an alternative approach and visualise the Red Queen in an artificial life model using Wright’s [18] concept of the fitness landscape, animating this and the population’s position on it as both change.

The Model

The model underlying the visualisation¹ introduced in this paper is that presented by Marshall & Rowe [12, 13, 14] and only the relevant details are very briefly discussed here. The model was designed to study the evolution of cooperation. Interaction within the model takes the form of an iterated 2 x 2 game, typically the Iterated Prisoner’s Dilemma (IPD). Individuals have both a memory of their interactions with other individuals and a strategy that specifies the current action in the game from two choices, A or B, based on the action choices of both participants on their last interaction together. The model is an evolutionary one, with selection pressure applied on payoffs from interactions through the imposition of a per time step living cost, so individuals die and are replaced by the offspring of the individuals in the population that score higher in their interactions. Individuals are also terminated with a random death probability at each time step, to prevent successful individuals becoming “immortal”. Reproduction is sexual and uses crossover and point mutation operators to combine two parents’ chromosomes into an offspring chromosome. The strategy chromosome encoding follows Mar & St Denis [11], but with a binary alphabet, to specify a two-dimensional strategy. Thus five loci are used (one for each potential interaction history, one for the first interaction), each with two alleles (A or B).

Visualisation Method

The inspiration for our visualisation method is Wright’s [18] fitness landscape. Wright posed the question “if the entire field of possible gene combinations be graded with respect to adaptive value under a particular set of

¹ The source code for the model and visualisation tool is downloadable from <ftp://ftp.swarm.org/pub/swarm/apps/objc/contrib/EPD-2.1.2-2.1.1.tar.gz>

conditions, what would be its nature?" (*ibid*). Wright presented a few examples of visualising simple genospaces, from two to five paired allelomorphs (see figure 1), and noted the high dimensionality of even these simple examples, with another dimension also necessary to represent adaptive value or fitness.

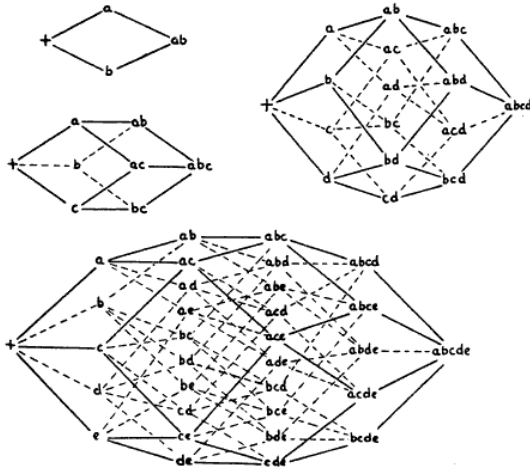


Figure 1. The combinations of from 2 to 5 paired allelomorphs. Reproduced from Wright [18].

To convey his concept of the fitness landscape, Wright thus presented a hypothetical example in which a high-dimensional genospace was compressed into a two-dimensional plane, and drew contour lines to delineate areas of different fitness (see figure 2), much like a cartographer uses contour lines to indicate altitude on a map.

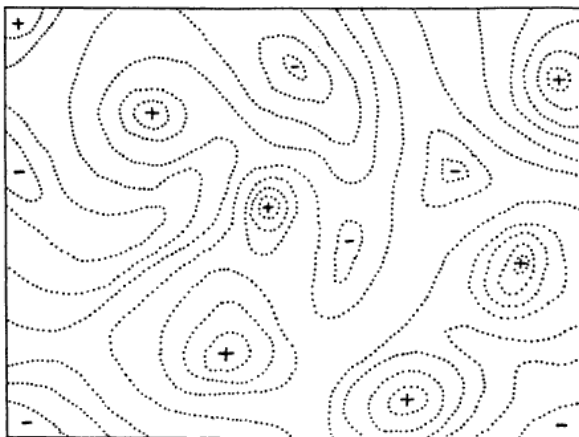


Figure 2. Diagrammatic representation of the field of gene combinations in two dimensions instead of many thousands. Dotted lines represent contours with respect to adaptiveness. Reproduced from Wright [18].

Genospace Compression

The technique presented in this paper therefore aims to

compress the high-dimensional genospace of the Marshall & Rowe model into a representation of sufficiently low dimensionality to present meaningfully on paper or a computer screen (i.e. in two dimensions). A key requirement of this compression, however, is that it maintains the relationships of points in the genospace and points in the compressed representation, i.e. points which are genetically close in genospace (limited genetic difference, or Hamming distance), should be plotted close to each other in the representation. A means of representing the adaptive value of each point in the compressed genospace is also required.

Although the genetic representation used in the Marshall & Rowe model has 5 loci per chromosome and is therefore five-dimensional, as there are only 2 alleles per locus there are only 32 different possible genotypes. The Search Space Matrix technique, described by Collins [4], provides an elegant means of compressing such a genospace into two dimensions, while minimising the Hamming distance between neighbouring points in the compressed representation. The result of applying the Search Space Matrix technique to the genospace of the Marshall & Rowe model is presented below in figure 3.

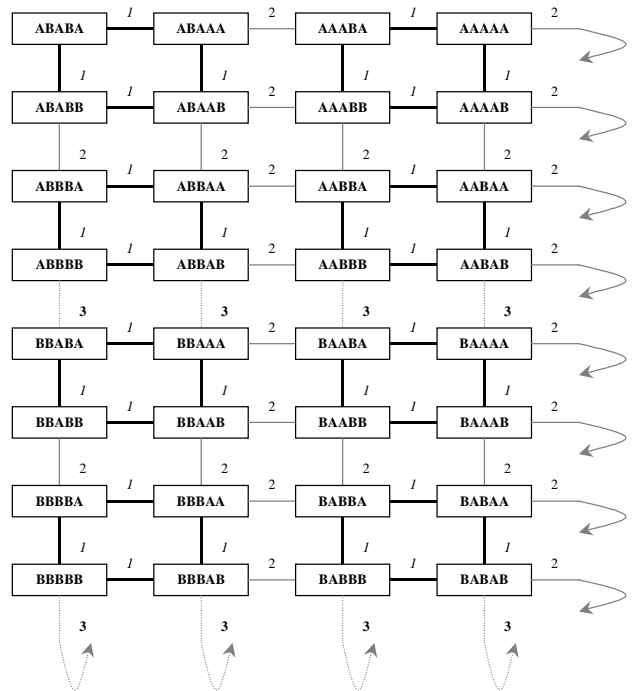


Figure 3. Search Space Matrix representation of genospace in the Marshall & Rowe model. Cells are points in genospace, lines connect neighbouring points and indicate Hamming distance between them. As the Search Space Matrix is toroidal, Hamming distances at the bottom and right edges indicate genetic distance to the opposite edges of the matrix.

Fitness Evaluation

In order to produce a fitness landscape, it is not enough to simply represent the genospace; the adaptive value of each point in that genospace must also be represented. With a two-dimensional representation such as that produced by the Search Space Matrix technique the obvious way to represent fitness is using the third dimension, namely height. However a means of determining the adaptive value of a given genotype must be specified.

In most traditional applications of evolutionary algorithms the objective is to optimise a static problem, therefore the fitness landscape is fixed and the fitness of any individual solution can be evaluated in isolation. However in co-evolutionary or co-adaptive systems such as ours, fitness is a function of the co-evolving species, or of the population as a whole. Here we specify how we calculate fitness in the Marshall & Rowe model, using expected utility from game theory.

The approach to calculating fitness is simple. Each point in the model's genospace represents a unique strategy for an iterated 2 x 2 game. At each model timestep, the frequencies of each of the n strategies are logged from the model in the row matrix

$$[\mathbf{F}] = (f_1, f_2, \dots, f_n). \quad (1)$$

The value of a strategy against another, $V(I|J)$, represents the score of strategy I when interacting with strategy J, e.g. for the IPD $V(\text{ALL D}|\text{TFT}) = T + wP/(1 - w)$ where ALL D is the strategy of unconditional defection, TFT is the well known strategy Tit-for-Tat, T is the temptation payoff in the Prisoner's Dilemma, P is the punishment payoff in the Prisoner's Dilemma, and w is the "shadow of the future" or probability of continued interaction from one iteration to the next in the game [2]. The payoff of any one strategy S_i in the model against all the n other strategies, given a known shadow of the future w , is thus specified by the matrix

$$\mathbf{O} = \begin{pmatrix} V(S_1/S_1) & V(S_2/S_1) & \dots & V(S_n/S_1) \\ V(S_1/S_2) & \ddots & & \\ \vdots & & & \\ V(S_1/S_n) & & \dots & V(S_n/S_n) \end{pmatrix}. \quad (2)$$

Then the expected utility of each strategy in a certain population is given by the row matrix

$$[\mathbf{E}] = [\mathbf{F}] \times \mathbf{O}. \quad (3)$$

It is easy to normalise the values in this matrix using the maximum and minimum values in \mathbf{O} , in order to simplify the calculation of the appropriate height values in the visualisation. The value of w used to calculate \mathbf{O} is set to reflect the average interaction length typically observed in the model with the parameters under investigation. Note that this calculation of expected utility assumes a panmictic population, because the frequencies of strategies as experienced by each individual strategy are identical, whereas the Marshall & Rowe model implements a spatial

population. However the model can be parameterised to be equivalent to a panmictic population, and that is the approach taken here.

Population Distribution

Having determined a means of representing the genospace, and the adaptive value of each point within it, the last remaining component of our visualisation technique is the visualisation of the areas of genospace that the population occupies. In our three-dimensional visualisation the x and y-axes are used to represent the genospace, and the z-axis is used to represent the adaptive value of points in the genospace. The obvious means to represent the population's distribution in the genospace is therefore colouring of the landscape representation, so each point in the genospace is assigned a value from the row matrix $[\mathbf{F}]$ representing the frequency of that genotype in the population. This value is translated to a colour from a colour range and used to colour the landscape in the visualisation.

Results

Figure 4 below presents a sample visualisation from a simulation run of the Marshall & Rowe model using the Apology game².

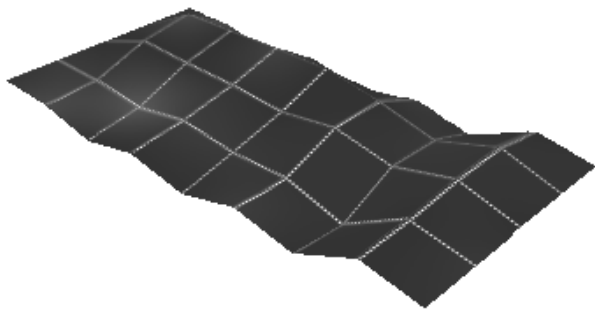
Apology, the payoff matrix for which is shown below, was used in this instance as it generates a more interesting fitness landscape than the Prisoner's Dilemma and consequently more interesting evolutionary dynamics.

	A	B
A	1,1	5,3
B	3,5	0,0

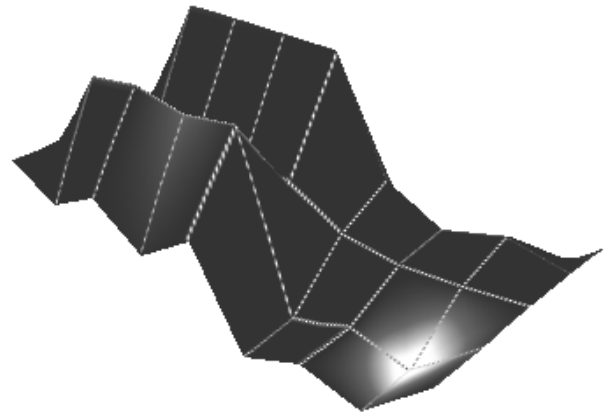
Table 1. Apology payoff matrix (payoffs to row player shown first).

The main difference with the Prisoner's Dilemma is that Apology reserves the highest payoffs for players that make different stage-game choices, while players co-ordinating on one outcome do less well. An evolving population of Apology players should therefore be expected to be especially susceptible to changes in the frequency of different strategies, and hence provide good conditions for observation of the Red Queen effect.

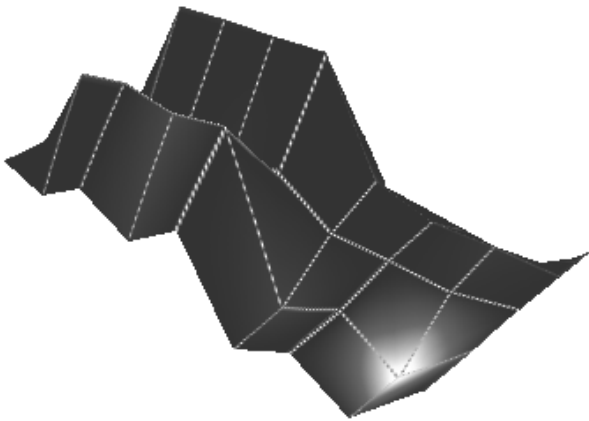
² The model was run with the following parameters: horizontal environment size in cells = 1, vertical environment size in cells = 1, initial individual energy = 15, energy living cost = 3, crossover rate = 0.1, mutation rate = 0.01, maximum population size = 50, initial population size = 25, death probability $d = 0.05$, interaction length per encounter = 3. Apology payoffs were as follows: AA = 1, AB = 5, BA = 3, BB = 0. For the visualisation the "shadow of the future" was set to $w \approx 0.7$.



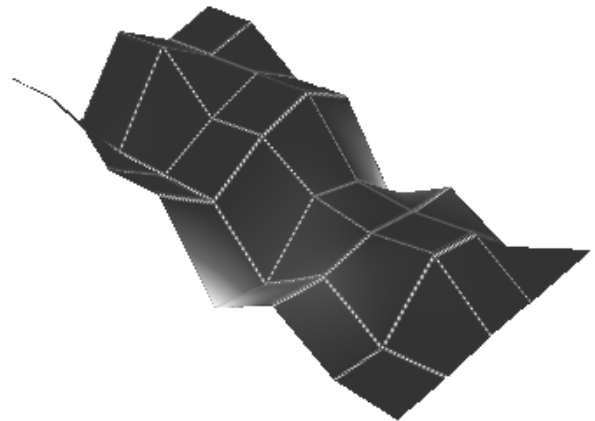
(a) $t = 80$



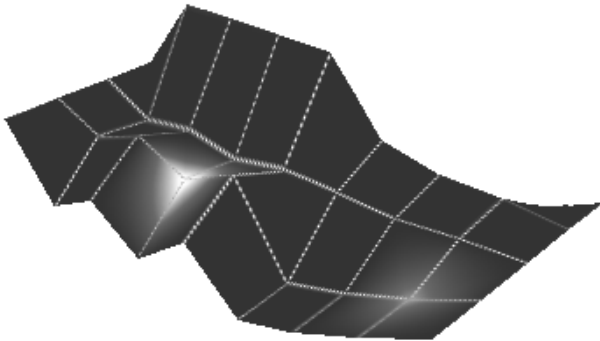
(d) $t = 465$



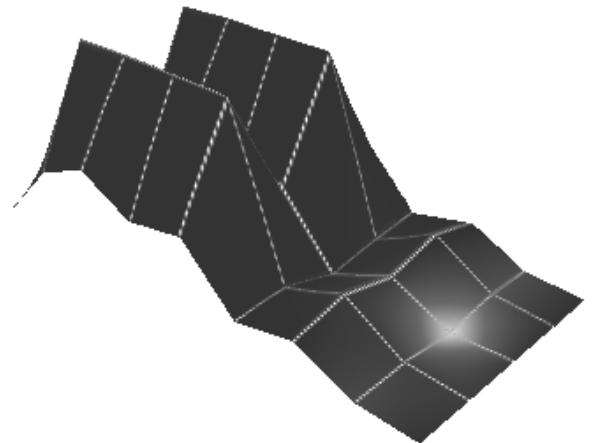
(b) $t = 345$



(e) $t = 1040$



(c) $t = 415$



(f) $t = 1210$

Figure 4. Fitness landscape visualisations for one simulation. Peaks in the landscape indicate high-fitness genotypes. Gridlines are used to indicate the relief of the landscape. Colour represents the frequency distribution of the population in the genospace (white = 1, black = 0). The origin (top left) of the SSM is the point BBBB.



The sequence presented in figure 4 shows the population in the model moving through genospace over a period of simulated time, and the associated changes in the fitness of different points in that genospace. Following several hundred timesteps of what might best be described as genetic drift over a largely uniform but constantly changing fitness landscape (figure 4a), a putative instance of the Red Queen Effect is seen in figures 4b to 4d. In figure 4b the population mostly occupies a fitness valley, yet two obvious fitness peaks exist. Attempting to occupy one of these fitness peaks (4c), that peak becomes depressed by the change in the population's constituency while the fitness valley the population began at is simultaneously elevated. The population abandons the depressed fitness peak for its newly elevated fitness valley, only to achieve, by this shift, a reversion to the former shape of the landscape (4d). This dynamic tension between occupying the fitness peak and remaining in the fitness valley continues for several hundred timesteps, followed by more drift, at the conclusion of which the landscape adopts an entirely new form on which the population finds itself concentrated once again in a fitness valley (4e). In the course of trying to adapt to this new landscape, the population restores the fitness landscape to a state very similar to that first seen in (4b), on which it finds itself in a very similar position to that it previously occupied (4f).

Discussion

While the evolutionary dynamics illustrated by the sample results presented above are interesting, they pose more questions than they answer. Even a more thorough investigation of these results alone, such as explaining the evolutionary dynamics from a game theoretic perspective, is beyond the scope of this paper, concerned as it is with visualisation. However, these results do shed a little light on some questions that have exercised biologists for the past decades or even centuries. For example, the observation of what might be characterised as genetic drift in the model addresses an issue that Wright and even Darwin himself were concerned with, the exclusivity or otherwise of natural selection as a mechanism of evolution. In particular the comparatively uniform fitness landscape of figure 4a could represent a temporary change of the landscape into a "neutral network" [15] through which the population is able to percolate, although with non-adaptive consequences in this case (as the population becomes stuck in a fitness valley in figure 4b). Similarly the question of the mode of evolution, whether gradual or punctuational, finds some answers in the model through the observation of prolonged periods of apparent stasis followed by rapid shifts in the population's genetic constitution. Of course the observation of any such phenomena within a model does not settle these questions, and detailed investigation is required armed additionally with suitable metrics, such as those developed by Cliff & Miller [3]. However a visualisation approach similar to that described here provides an intuitive entry-point for the problem and

perhaps a fruitful one.

We believe that the visualisation technique we have presented here differs from earlier approaches by combining the representation of the fitness landscape with that of the population's position on it. Earlier approaches have visualised only the fitness landscape (by population sampling) [1], or have separately visualised the population's fitness distribution and the population's trajectory through the genospace [9]. We believe that the preliminary results presented here demonstrate both the practicality and the utility of combining both these elements into a single representation, as Wright himself proposed doing [18]. We hope that such visualisation techniques may be usefully applied to other co-evolutionary systems, artificial immune systems for example, and that others will find it interesting to explore further the ideas presented in this paper.

For application to other co-evolutionary systems, several points will have to be taken into account. Most obviously, the target model presented here has a sufficiently limited genospace that it is practical to exhaustively evaluate the adaptive value of every point in it given a certain population distribution. In more complex evolutionary systems this will not be the case, and measures will have to be taken to sample the fitness landscape and build an approximate representation of it (this is indeed the approach taken in [1] and [9]). Also, the genospace compression technique used here, Search Space Matrix, has some limitations, particularly the existence of neighbouring points of high genetic difference (from figure 3 it can be seen that there are two "Hamming cliffs" where neighbouring points are different at three alleles). One approach would be to have geometric space between points in the visualisation related to Hamming distance in the genospace representation. Alternatively, it would be helpful to investigate the application of different compression techniques such as Sammon mapping [16]. In the future we hope to investigate these issues through the visualisation of progressively more complex co-evolutionary systems.

Acknowledgements

Thanks to Tim Kovacs and members of the Ant Lab at the University of Bristol for helpful comments and suggestions.

References

1. Anastasoff, S. J. 2000. A 'Fitness Landscaping' Comparison of Evolved Robot Control Systems. In Proceedings of the Seventh International Conference on Artificial Life, 273-281. Cambridge, Mass.: MIT Press.
2. Axelrod, R. 1984. The Evolution of Cooperation. New York: Basic Books.
3. Cliff, D. and Miller G. F. 1995. Tracking the Red Queen: Measurement of Adaptive Progress in Co-Evolutionary Simulations. In Advances in Artificial Life:

- Proceedings of the Third European Conference on Artificial Life (ECAL95). Lecture Notes in Artificial Intelligence 929, 200-218. Berlin: Springer Verlag.
4. Collins, T. D. 1997. Using Software Visualisation Technology to Help Evolutionary Algorithm Users Validate Their Solutions. In Proceedings of the Seventh International Conference on Genetic Algorithms (ICGA'97), 307-314. San Francisco, Calif.: Morgan Kaufmann.
 5. Darwin, C. R. 1859. On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life. London: John Murray.
 6. Dawkins, R. and Krebs, J. R. 1979. Arms Races Between and Within Species. Proceedings of the Royal Society of London B 205: 489-511.
 7. Eldredge, N. and Gould, S. J. 1972. Punctuated Equilibria: an Alternative to Phyletic Gradualism. In Models in Paleobiology, 82-115. San Francisco, Calif.: Freeman.
 8. Hamilton, W. J. and Orians, G. H. 1965. Evolution of Brood Parasitism in Altricial Birds. Condor 67: 361-382.
 9. Harvey, I. and Thompson, A. 1997. Through the Labyrinth Evolution Finds a Way: A Silicon Ridge. In Proceedings of the First International Conference on Evolvable Systems. Lecture Notes in Computer Science 1259, 406-422. Berlin: Springer Verlag.
 10. Kauffman, S. A. and Johnsen, S. 1991. Coevolution to the edge of chaos: coupled fitness landscapes, poised states, and coevolutionary avalanches. Journal of Theoretical Biology 149: 467.
 11. Mar, G. and St Denis, P. 1994. Chaos in Cooperation: Continuous-Valued Prisoner's Dilemmas in Infinite-Valued Logic. International Journal of Bifurcation and Chaos 4(4): 943-958.
 12. Marshall, J. A. R. and Rowe, J. E. 2000. Investigating the Mechanisms Underlying Cooperation in Viscous Population Multi-Agent Systems. In Proceedings of the Seventh International Conference on Artificial Life, 348-352. Cambridge, Mass.: MIT Press.
 13. Marshall, J. A. R. and Rowe, J. E. 2003a. Kin Selection May Inhibit the Evolution of Reciprocation. Journal of Theoretical Biology 222(3): 331-335.
 14. Marshall, J. A. R. and Rowe, J. E. 2003b. Viscous Populations and Their Support for Reciprocal Cooperation. Artificial Life 9(3): 327-334.
 15. Maynard Smith, J. 1970. Natural selection and the concept of a protein space. Nature 225: 563-564.
 16. Sammon, J. W. 1969. A Nonlinear Mapping for Data Structure Analysis. IEEE Transactions on Computers C 18(5): 401-408.
 17. van Valen, L. 1973. A New Evolutionary Law. Evolutionary Theory 1: 1-30.
 18. Wright, S. 1932. The Roles of Mutation, Inbreeding, Crossbreeding and Selection in Evolution. In Proceedings of the Sixth International Congress on Genetics, 356-366. Ithaca, New York: Brooklyn Botanic Garden.