

# Optimal Mutation Rate Control under Selection in Hamming Spaces

Elizabeth Aston<sup>1</sup>, Alastair Channon<sup>1</sup>, Roman V. Belavkin<sup>2</sup>, Rok Krašovec<sup>3</sup> and Christopher G. Knight<sup>3</sup>

<sup>1</sup> School of Computing and Mathematics, Keele University, ST5 5BG, UK

<sup>2</sup> School of Engineering and Information Sciences, Middlesex University, London, NW4 4BT, UK

<sup>3</sup> Faculty of Life Sciences, The University of Manchester, M13 9PT, UK

{e.j.aston, a.d.channon}@keele.ac.uk, r.belavkin@mdx.ac.uk, {rok.krasovec, chris.knight}@manchester.ac.uk

## Abstract

We investigate the effect of selection in a meta-genetic algorithm designed to optimize mutation rate control, based on the fitness of sequences relative to a defined optimum, in asexual evolution. Multiple innovations in the algorithm are required to achieve the evolution of optimal mutation rate control under selection. Before implementing selection, results from this improved algorithm clarify the optimal relationship of mutation rate to distance from the optimum as being a double sigmoid for binary sequences. Furthermore, the results clarify how such control functions depend on alphabet size, sequence length and the time horizon over which evolution is assessed. Incorporating selection leads to a distinctive shape of optimal mutation rate control function. This function has a mutation rate less than a third of  $1/\text{length}$  at a Hamming distance of one from the optimum and beyond. This surprising result for a simple, universally monotonic single-peaked fitness landscape highlights the need for further research using models such as this. Future work will therefore explore how this control function may vary, for instance with population size and alternative selection mechanisms common in Artificial Life models.

## Introduction

Evolution is dependent on the balance between mutation, recombination, selection and genetic drift. These processes can act in opposition to one another, for example mutation introduces variation while drift causes a reduction in diversity. In other circumstances they can act together, for instance mutation rate is positively correlated with recombination rate in human single nucleotide polymorphisms (Lercher and Hurst, 2002), in the Human Immunodeficiency Virus (Schlub et al., 2014), and in other retroviruses and RNA viruses (Tromas et al., 2014). Evolutionary processes can also interact in more complex ways, for instance, the type of recombination (crossover of single sequences versus crossover between diploid maternal and paternal sequences) can affect the magnitude of the critical mutation rate, above which sequences with greater robustness to mutation are favoured over individuals with greater fitness (Aston et al., 2013).

Genetic variation is at the centre of this balance. Both recombination and mutation introduce variation, while ge-

netic drift and selection reduce it, effectively favouring certain individuals through random sampling or based on fitness or respectively. Selection and mutation in particular oppose one another in mutation-selection balance (Kimura and Maruyama, 1966; Bull et al., 2005). Selection affects the persistence of sequences depending on their fitness relative to other sequences in the population. The amount of variation maintained reduces as the strength of selection increases (Avila et al., 2014). Mutation rates and their levels relative to theoretically optimal and critical mutation rates, can determine how efficiently adaptation can occur and, to a degree, whether or not adaptations are lost.

These processes also interact with population parameters. For instance, genetic drift has a greater effect, relative to selection, in smaller populations due to random sampling. At a larger scale, Sung et al. (2012) show an inverse relationship between mutation rate and effective population size across the tree of life, including unicellular eukaryotes, eubacteria, and multicellular eukaryotes.

Together, these interactions occur in the context of a fitness landscape (Wright, 1932), in which each organism has a fitness value representing its relative replication rate (Domingo and Wain-Hobson, 2009). Under selection the population tends to move toward ‘fitter’ regions of the fitness landscape to form a quasispecies (Eigen and Schuster, 1979; Nowak, 1992; Bull et al., 2005; Nowak, 2006). In a landscape with a single fitness peak, a quasispecies is able to maintain its position surrounding the top of the peak, in mutation-selection balance, so long as the mutation rate does not exceed the error threshold above which adaptation is lost (Eigen et al., 1988; Nowak and Schuster, 1989; Tannenbaum and Shakhnovich, 2004; Bull et al., 2005; Saakian et al., 2006; Nowak, 2006; Takeuchi and Hogeweg, 2007; Domingo and Wain-Hobson, 2009; Schuster, 2009; Tejero et al., 2011). If the different evolutionary processes vary systematically across such a fitness landscape it has the potential to affect the course of evolution profoundly. Landscape-dependent control of evolutionary parameters has been considered in evolutionary algorithms (Eiben et al., 1999) but is much less understood in biology. For the key process of

mutation, we have started to explore how its rate varies in practice (Krasovec et al., 2014). Here, for the first time, we consider how such rates might be controlled optimally, in simple fitness landscapes, in the presence of selection.

### Mutation rate control in Hamming spaces

Mutation rate control has been studied theoretically and *in silico*. Belavkin et al. (2011) derived the probability of adaptation as a function of mutation rate, for populations of asexual self-replicating organisms with a monotonic fitness landscape, in the absence of selection. The probability of adaptation is dependent on the rate of mutation. This introduces the possibility that organisms may maximize the expected fitness of their offspring through mutation rate control. Belavkin et al. (2011) noted that, in general, analytical solutions are not available for long sequences evolving for more than one generation, and used a meta-genetic algorithm (Meta-GA) to evolve populations of mutation rate control functions. Each function  $\mu(f)$  was used to control mutation rates in another GA, referred to as the Inner-GA. The Inner-GA constitutes the evolutionary model, while the Meta-GA acts to optimise mutation rate control.

**The Inner-GA.** The Inner-GA is a simple generational genetic algorithm in which each genotype is a sequence  $\omega \in H_\alpha^l$ , where  $H_\alpha^l := \{1, \dots, \alpha\}^l$  is the space of sequences of length  $l$  and  $\alpha$  letters and is equipped with the Hamming metric  $d_H(a, b) := |\{i : a_i \neq b_i\}|$ . Inner-GA populations, of 100 individuals each, are initialised by the Meta-GA (see below) and evolved by the Inner-GA with the objective being to maximise a fixed fitness function  $f(\omega)$ . Mutation is according to a per-locus mutation rate control function  $\mu(f)$  specified by an individual genotype from the Meta-GA (below). In previous work (Belavkin et al., 2011) the Inner-GA ran for 500 generations and used no recombination and no selection, allowing for very rapid execution on GPGPU architectures due to thread independence.

**The Meta-GA.** The Meta-GA is a simple generational genetic algorithm that uses tournament selection (a good choice when little is known or assumed about the structure of the landscape) to maximise mean fitness in the final generation of the Inner-GA. Each genotype in the Meta-GA is a mutation rate control function  $\mu(f)$ , which is a sequence of real values  $\mu \in [0, 1]$  representing probabilities of mutation at different fitness values or binned intervals. In the case of Inner-GA fitness being the Hamming distance between sequences in  $H_\alpha^l$ , Meta-GA genotypes have length  $l + 1$  so as to cover the range of fitnesses from 0 to  $l$ . Each run has a population of 100 functions, with each function initialised to  $\mu(f) = 0$  for all  $f$ . In each generation of the Meta-GA, the Inner-GA is run once for each mutation rate control function (so 100 times), with each Inner-GA run evolving a population of 100 sequences.

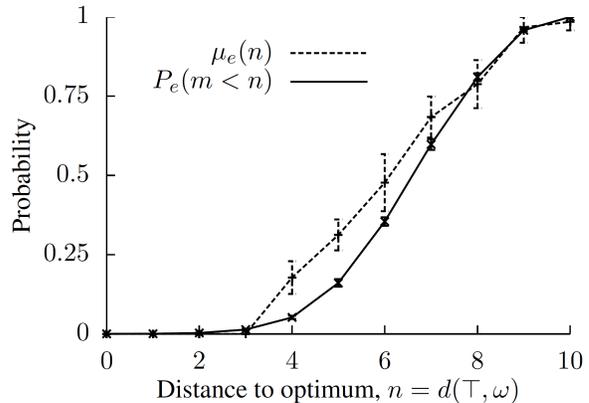


Figure 1: Mutation rate control function  $\mu_e$ , evolved to maximise mean fitness  $f(\omega) = -d_H(T, \omega)$  after 500 generations of evolution without selection,  $\omega \in H_4^{10}$ . Averaged over 20 runs. Error bars show standard deviations. Taken from previous work (Belavkin et al., 2011, figure 3). The CDF  $P_e$  is also shown but not discussed in this paper.

The Meta-GA randomly selects three individuals from the population and replaces the least fit with a mutated single-point crossover of the other two. This process repeats until every individual in the Meta-GA population has been selected, or until fewer than three individuals remain, so producing the next generation. This method allows for excellent parallelisation, as triples can be selected at the start of each generation. The Meta-GA returns the fittest mutation rate control function  $\mu(f)$  in its final generation.

In previous work (Belavkin et al., 2011), the Meta-GA ran for 500000 generations and its mutation procedure added a uniform-random number  $\Delta\mu \in [-.1, .1]$  to one randomly selected value  $\mu$  (mutation rate, bounded to be within  $[0, 1]$ ) in the individual mutation rate control function. In each generation of the Meta-GA, the 100 Inner-GA runs were seeded with identical populations. These were generated randomly, in each Meta-GA generation, so as to have the same number of individuals at each possible fitness value. Figure 1 shows the optimal mutation rate function previously found for short (length 10) base 4 sequences evolving for 500 generations without selection.

### Mutation rate control in biological landscapes

The mutation rate control function evolved by the Meta-GA is dependent on the fitness landscape used in the Inner-GA. If fitness  $f(\omega)$  corresponds to negative Hamming distance to the optimum,  $-d_H(T, \omega)$ , then the optimal mutation rate can be seen to increase with  $d_H(T, \omega)$  (Belavkin et al., 2011). Biologically relevant landscapes are likely to be more complex than this simple case. In more rugged landscapes, fitness does not provide all necessary information about the position of a sequence in the landscape.

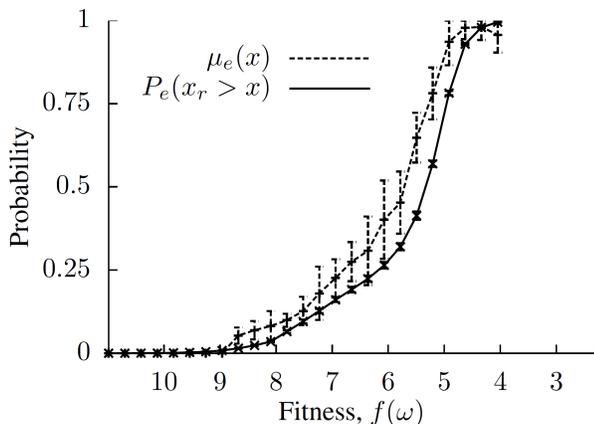


Figure 2: Mutation rate control function  $\mu_e$ , evolved to maximise mean fitness on an aptamer landscape after 500 generations of evolution without selection,  $\omega \in H_4^{10}$ . Averaged over 20 runs. Error bars show standard deviations. Taken from previous work (Belavkin et al., 2011, figure 4).

There is extensive variation in mutation rates among species in nature (Sung et al., 2012). The degree to which biological organisms can control their mutation rate is a more open question, but variation in mutation rates within species and in response to the environment (mutation rate plasticity) could be widespread (MacLean et al., 2013; Krasovec et al., 2014). Such propensity for mutation rate variation in nature raises the question of whether the findings about optimal control of mutation rate in simple landscapes in silico could be extended to biological systems. Belavkin et al. (2011) evolved mutation rate control functions with fitness defined by an aptamer landscape. An aptamer is a nucleic acid that can be selected to bind to a particular target molecule (Knight et al., 2009). This complete DNA-protein affinity landscape was described by Rowe et al. (2010), and represents a rugged landscape with many local optima. Figure 2 shows the average of evolved mutation control functions for the aptamer landscape. This provides evidence that the simpler results for  $f(\omega) = -d_H(\top, \omega)$  have relevance to biology. However, neither figure 1 nor figure 2 was produced using a model that included selection in the Inner-GA. Selection is a key aspect of evolution and introducing it into the model is therefore the next logical step.

### Innovations in the Meta-GA

In order to achieve the results presented in this paper, it was necessary to make significant improvements to the Meta-GA. Most notably, the introduction of selection into the Inner-GA results in very much more rapid evolution to the target sequence, and so in Inner-GA populations spending very little search time far from the optimum. In the selection-based runs reported below, less than 1% of Inner-

GA evaluations were at fitnesses in the lower quartile, and less than 5% were in the lower half. This necessary paucity of evaluations far from target greatly hinders the optimisation of mutation rates in these fitness regions. This is compounded by the requirement for inter-process communication and synchronisation, brought about by selection’s need for fitness comparisons and the copying of genotypes between threads. This significantly reduces performance on GPGPU architectures, although these still provide a pronounced increase in performance when compared to a straightforward CPU implementation. In each run reported below, the Meta-GA was restricted (for practical considerations) to run for just 2000 generations.

Having explored a number of strategies for improving the Meta-GA’s ability to optimise mutation rate control functions, and after searching over a wide range of parameters, the following modifications to the Meta-GA (as described in our previous work and this paper’s introduction) were employed for the experiments reported here.

1. To evaluate a mutation rate control function from the Meta-GA, the Inner-GA is run not once but  $10(l + 1)$  times, i.e. 110 times for  $l = 10$  and 310 times for  $l = 30$ . The fitness of a mutation rate control function is summed over these runs.
2. Rather than seeding Inner-GA populations such that each begins with the same number of individuals at each possible fitness value, the multiplicity of runs introduced above is utilised further in order to improve the optimisation of mutation rates across the full range of fitness values. 10 runs are initiated with all individuals at the target ( $d_H(\top, \omega) = 0$ ), and 10 runs are initiated with their populations randomised such that each individual is at distance 1, 10 runs at distance 2, ... 10 runs at distance  $l$ . Within each generation of the Meta-GA, all mutation rate control functions are evaluated using copies of the same  $10(l + 1)$  seed populations.
3. Uniform-random<sup>1</sup> *delta* mutation is reduced to  $\Delta\mu \in [-.01, .01]$ . This allows for better fine-tuning of mutation rates.
4. 10% of mutations in the Meta-GA are *resets* to a uniform-random value  $\mu \in [1, 1]$ ; the other 90% remain *delta* mutations ( $\Delta\mu$ , see above). This aids in the optimisation of mutation rates at low fitnesses.

These innovations alone lead to interesting new findings about optimal mutation rate control in the absence of selection, as they generate clearer relationships between fitness and optimal mutation rate. Figure 3 demonstrates very

<sup>1</sup>Parallel sets of runs, to those reported below, employed Gaussian mutation over a range of mutation rates, with and without modification 4 (above). The results were consistent with those below but uniform-random mutation resulted in lower errors (variances) within the sets of 20 runs each.

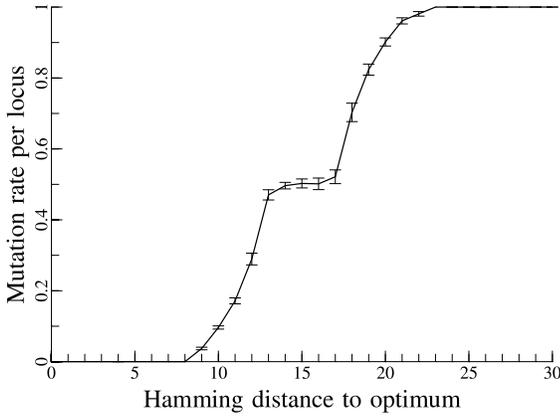


Figure 3: Mutation rate control function  $\mu(f)$ , evolved to maximise mean fitness  $f(\omega) = -d_H(\top, \omega)$  after 100 generations of evolution without selection,  $\omega \in H_2^{30}$ . Averaged over 20 runs. Error bars show standard deviations.

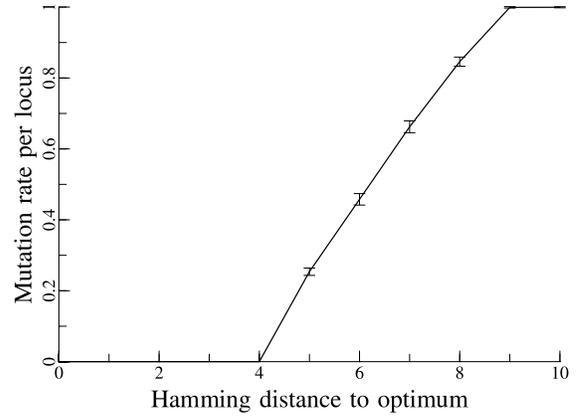


Figure 5: Mutation rate control function  $\mu(f)$ , evolved to maximise mean fitness  $f(\omega) = -d_H(\top, \omega)$  after 100 generations of evolution without selection,  $\omega \in H_4^{10}$ . Averaged over 20 runs. Error bars show standard deviations.

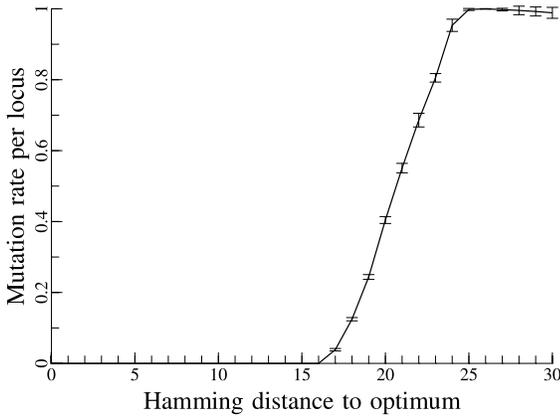


Figure 4: Mutation rate control function  $\mu(f)$ , evolved to maximise mean fitness  $f(\omega) = -d_H(\top, \omega)$  after 100 generations of evolution without selection,  $\omega \in H_4^{30}$ . Averaged over 20 runs. Error bars show standard deviations.

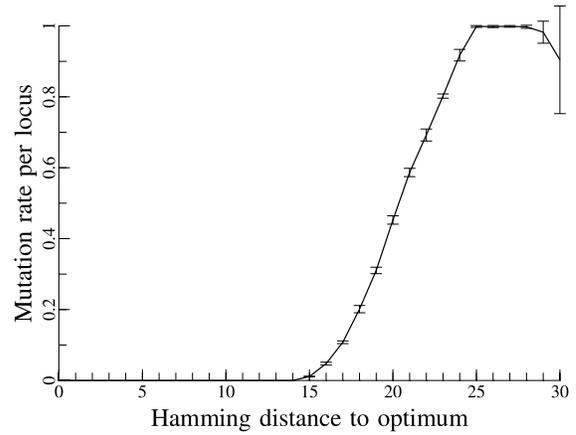


Figure 6: Mutation rate control function  $\mu(f)$ , evolved to maximise mean fitness  $f(\omega) = -d_H(\top, \omega)$  after 500 generations of evolution without selection,  $\omega \in H_4^{30}$ . Averaged over 20 runs. Error bars show standard deviations.

clearly that for a binary sequence, the optimal mutation rate control function is not a simple sigmoid but a double-sigmoid. Our intuition is that this may relate to binary's unique property, among alphabet sizes ( $\alpha$ ), that two mutations in succession, at the same locus, are guaranteed to return the original value. Certainly figure 4 shows that the optimal mutation rate control function for  $\alpha = 4$  (as for DNA) is sigmoidal, not double-sigmoidal.

Figure 5, when compared to figure 4, shows that shortening sequence length from 30 to 10 results in a less sigmoidal, more "z-like" mutation rate control function; likewise comparing figures 7 and 6. Comparing figures 4 and 6 shows that reducing the number of (Inner-GA) generations before eval-

uation from 500 to 100 results in a steeper, more step-like optimal mutation rate control function, which rises above zero further from the optimum and more rapidly than for the less urgent function; likewise comparing figures 7 and 5.

Note also that figure 7 is a match (with reduced error) for the evolved mutation function shown in figure 1, taken from (Belavkin et al., 2011), demonstrating that the innovations noted above have not affected the core result, other than reducing error.

### Mutation rate control under selection

The main advantage of the innovations above are that they enable the evolution of mutation rate control functions un-

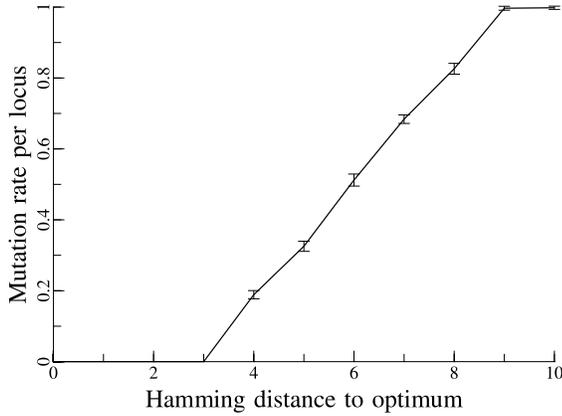


Figure 7: Mutation rate control function  $\mu(f)$ , evolved to maximise mean fitness  $f(\omega) = -d_H(\top, \omega)$  after 500 generations of evolution without selection,  $\omega \in H_4^{10}$ . Averaged over 20 runs. Error bars show standard deviations.

der selection, making this work much more relevant to both biology and general evolutionary algorithms. We introduce selection into the Inner-GA through the following modification: in every generation of the Inner-GA, each genotype is compared to a randomly chosen (per genotype) other genotype and replaced by a copy of that if that is fitter. Mutation is then applied, as for the without-selection case. The population is fully mixed.

Figure 8 shows the evolved mutation rate control function for 30-long sequences of base 4 (such as DNA) when evolving for 100 generations under selection. Values for highly unfit sequences (worse than random, i.e.  $d_H > 22$ ) have not been highly optimised but the majority of the function shows very little variation among independent runs. There is a clear, distinctive shape to the optimal mutation rates moving out from the optimum. At the optimum the optimal mutation rate is of course zero. One mutation away it jumps to a non-zero value and then slowly rises from there. Most intriguingly, the optimal mutation rate one mutation from the origin is significantly lower than the “ $1/L$ ” heuristic of conventional wisdom, despite the maximal simplicity of the fitness landscape.

### The $1/L$ heuristic

It can be seen from figure 8 that the optimal mutation rate found for an Inner-GA that incorporates selection does not follow the  $1/L$  heuristic.  $1/L$  has been suggested as a general value for the per-locus mutation rate in a GA, where  $L$  is sequence length (Mühlenbein, 1992; Bäck and Schütz, 1996; Ochoa et al., 2000; Ochoa, 2002). Mühlenbein (1992) states that  $\mu = 1/L$  is optimal for general unimodal functions. Ochoa (2002) used GAs with bit-strings to test the limitations of the  $1/L$  heuristic. The change in optimal mu-

tation rate with time was studied, along with the interaction between mutation rate, selection pressure, and population size. It was found that a mutation rate of  $1/L$  produces optimal or near optimal results. It was also found that increasing the selection pressure increases the magnitude of optimal mutation rates; and that decreasing population size (without changing selection pressure) can result in a decrease in optimal mutation rate at small population sizes, specifically in the case of the Knapsack Problem when reducing population size from 50 to 10. It was concluded that a rate of  $1/L$  is sub-optimal only when the selection pressure is either extremely weak or extremely strong, or when the population size is very small (Ochoa, 2002). The optimal mutation rate for Hamming distances close to the optimum in figure 8 is close to 1%, where as following the  $1/L$  heuristic it is expected to be  $1/30$ , i.e. approximately 3.3%. The fact that the observed optimal mutation rate is less than a third of the expected rate may relate to the small population size of 100 used in the Inner-GA, although population sizes of this magnitude are frequently used in Artificial Life work. The effect of tournament selection (also frequently used in Artificial Life work) on selective advantage could also be an explanatory factor.

Error threshold is related to optimal mutation rate, and is therefore also dependent on selection pressure. Ochoa et al. (2000) note that selection pressure is an informal term and that “Loosely the selection pressure measures the ratio of maximum to average fitness in the population”. In a fitness landscape with a single peak with fitness  $\sigma > 1$ , and with all other sequences having fitness 1, the error threshold, denoted by Ochoa et al. (2000) as  $p$ , is given as:

$$p = \frac{\ln(\sigma)}{L}$$

$\sigma$  is the selective advantage of the master (optimum) sequence over all other sequences in the population, i.e. the selection pressure.  $L$  is the sequence length. In the simplest instance,  $\sigma$  represents the ratio of the reproduction rate of the master sequence to the average reproduction rate of the rest of the sequences in the population. Ochoa et al. (2000) looked at error thresholds and optimal mutation rates when solving both toy and real-world problems with a genetic algorithm. They concluded that error thresholds and optimal mutation rates are correlated. In addition, selection pressure was shown to have a significant effect on the magnitude of both the error threshold and the optimal mutation rate; the stronger the selection pressure, the greater the error threshold and optimal mutation rate (Ochoa et al., 2000). However, it has been observed that the error threshold has an exponential dependence on population size that is more noticeable in small populations (Nowak and Schuster, 1989; Ochoa and Harvey, 1998; Ochoa et al., 1999; Channon et al., 2011; Aston et al., 2013). The population in the Inner-GA may be considered small enough to be affected by the exponential

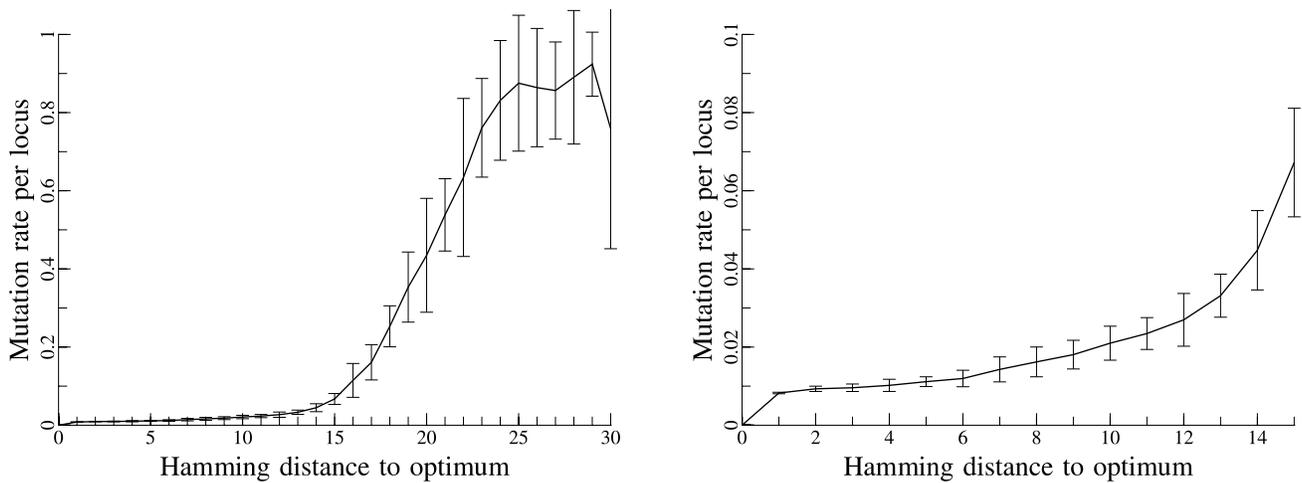


Figure 8: Mutation rate control function  $\mu(f)$ , evolved to maximise mean fitness  $f(\omega) = -d_H(\mathbb{T}, \omega)$  after 100 generations of evolution under selection,  $\omega \in H_4^{30}$ . Averaged over 20 runs. Error bars show standard deviations. Shown in full (left) and in detail near the optimum (right).

model, potentially influencing both the error threshold and optimal mutation rate, and providing a possible explanation for the fact that the observed optimal mutation rate is less than a third of the expected rate following the erroneous  $1/L$  heuristic.

## Conclusions

The innovations in the Meta-GA have provided a demonstration that for binary strings the optimal mutation rate control function is a double sigmoid, and also enabled us to clarify the relationship of optimal mutation rate to both sequence length and number of generations in this scenario.

When applied to the evolution of mutation rate control under selection, we have successfully determined the distinctive shape of the optimal mutation rate control function. This has a mutation rate less than a third of  $1/\text{length}$  at and near (above) a distance of one mutation from the optimum, despite the single-peak everywhere-monotonic simplicity of the fitness landscape. We have noted a number of possible reasons for this difference. This highlights a need for further research using models such as this, that can take population size and selection mechanisms common in Artificial Life models into account.

In previous work (Belavkin et al., 2011) we demonstrated that, in the absence of selection, the optimal mutation rate control function for a DNA-protein affinity fitness landscape resembles that for the monotonic case of fitness being negative Hamming distance. Further work is needed to evaluate whether or not the characteristics of optimal mutation rate control functions under selection generalise to such non-monotonic landscapes.

Future work may also incorporate more of the evolution-

ary pressures, namely recombination and genetic drift. It is expected that optimal mutation rate will counteract those pressures that reduce variation, and be complemented by those pressures that increase variation.

## Acknowledgements

This work was supported by Engineering and Physical Sciences Research Council (EPSRC) grant EP/H031936/1 and Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/L009579/1. The dataset underpinning the results is openly available from Zenodo at <http://doi.org/4pv>.

## References

- Aston, E., Channon, A., Day, C., and Knight, C. G. (2013). Critical mutation rate has an exponential dependence on population size in haploid and diploid populations. *PLoS ONE*, 8(12):e83438.
- Avila, V., Pérez-Figueroa, A., Caballero, A., Hill, W. G., García-Dorado, A., and López-Fanjuan, C. (2014). The action of stabilizing selection, mutation, and drift on epistatic quantitative traits. *Evolution*, 68(7):1974–1987.
- Bäck, T. and Schütz, M. (1996). Intelligent mutation rate control in canonical genetic algorithms. In *Proceedings of the Ninth International Symposium on Foundations of Intelligent Systems*, volume 1079.
- Belavkin, R. V., Channon, A., Aston, E., Aston, J., and Knight, C. G. (2011). Theory and practice of optimal mutation rate control in Hamming spaces of DNA sequences. In *Advances in Artificial Life, ECAL 2011*:

*Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems.*

- Bull, J. J., Meyers, L. A., and Lachmann, M. (2005). Quasispecies made simple. *PLoS Computational Biology*, 1(6):e61.
- Channon, A., Aston, E., Day, C., Belavkin, R. V., and Knight, C. G. (2011). Critical mutation rate has an exponential dependence on population size. In *Advances in Artificial Life, ECAL 2011: Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems*.
- Domingo, E. and Wain-Hobson, S. (2009). The 30th anniversary of quasispecies. *EMBO Reports*, 10:444–448.
- Eiben, A. E., Hinterding, R., and Michalewicz, Z. (1999). Parameter control in evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 3(2):124–141.
- Eigen, M., McCaskill, J., and Schuster, P. (1988). Molecular quasispecies. *Journal of Physical Chemistry*, 92:6881–6891.
- Eigen, M. and Schuster, P. (1979). *The hypercycle*. Springer, New York.
- Kimura, M. and Maruyama, T. (1966). The mutational load with epistatic gene interactions in fitness. *Genetics*, 54:1337–1351.
- Knight, C., Platt, M., Rowe, W., Wedge, D., Khan, F., Day, P., McShea, A., Knowles, J., and Kell, D. (2009). Array-based evolution of DNA aptamers allows modelling of an explicit sequence-fitness landscape. *Nucleic Acids Research*, 37(1):e6.
- Krasovec, R., Belavkin, R. V., Aston, J. A. D., Channon, A., Aston, E., Rash, B. M., Kadirvel, M., Forbes, S., and Knight, C. G. (2014). Mutation rate plasticity in rifampicin resistance depends on escherichia coli cell cell interactions. *Nature Communications*, 5:3742.
- Lercher, M. J. and Hurst, L. D. (2002). Human SNP variability and mutation rate are higher in regions of high recombination. *Trends in Genetics*, 18(7):337–340.
- MacLean, R. C., Torres-Barcelo, C., and Moxon, R. (2013). Evaluating evolutionary models of stress-induced mutagenesis in bacteria. *Nat Rev Genet*, 14(3):221–7.
- Mühlenbein, H. (1992). How genetic algorithms really work: Mutation and hillclimbing. *Parallel Problem Solving from Nature*, 2:15–26.
- Nowak, M. A. (1992). What is a quasispecies? *Trends in Ecology and Evolution*, 7:118–121.
- Nowak, M. A. (2006). *Evolutionary dynamics: Exploring the equations of life*. Harvard University Press.
- Nowak, M. A. and Schuster, P. (1989). Error thresholds of replication in finite populations: Mutation frequencies and the onset of Muller’s ratchet. *Journal of Theoretical Biology*, 137:375–395.
- Ochoa, G. (2002). Setting the mutation rate: Scope and limitations of the 1/l heuristic. In *Proceedings of Genetic and Evolutionary Computation Conference (GECCO-2002)*, pages 315–322.
- Ochoa, G. and Harvey, I. (1998). Recombination and error thresholds in finite populations. In *Foundations of Genetic Algorithms (FOGA-5)*, pages 245–264.
- Ochoa, G., Harvey, I., and Buxton, H. (1999). Error thresholds and their relation to optimal mutation rates. In *Proceedings of the Fifth European Conference on Artificial Life*, pages 54–63.
- Ochoa, G., Harvey, I., and Buxton, H. (2000). Optimal mutation rates and selection pressure in genetic algorithms. In *Proceedings of Genetic and Evolutionary Computation Conference (GECCO-2000)*.
- Rowe, W., Platt, M., Wedge, D. C., Day, P. J., and Kell, D. B. (2010). Analysis of a complete DNA-protein affinity landscape. *Journal of Royal Society Interface*, 7(44):397–408.
- Saakian, D. B., Muñoz, E., Hu, C., and Deem, M. W. (2006). Quasispecies theory for multiple-peak fitness landscapes. *Physical Review E*, 73:041913.
- Schlub, T. E., Grimm, A. J., Smyth, R. P., Cromer, D., Chopra, A., Mallal, S., Venturi, V., Waugh, M. C., Mak, J., and Davenport, M. P. (2014). Fifteen to twenty percent of hiv substitution mutations are associated with recombination. *Journal of Virology*, 88(7):3837–3849.
- Schuster, P. (2009). Genotypes and phenotypes in the evolution of molecules. *European Review*, 17(2):281–319.
- Sung, W., Ackerman, M. S., Miller, S. F., Doak, T. G., and Lynch, M. (2012). Drift-barrier hypothesis and mutation-rate evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 109(45):18488–18492.
- Takeuchi, N. and Hogeweg, P. (2007). Error-threshold exists in fitness landscapes with lethal mutants. *BMC Evolutionary Biology*, 7:15.
- Tannenbaum, E. and Shakhnovich, E. I. (2004). Solution of the quasispecies model for an arbitrary gene network. *Physical Review E*, 70:021903.

- Tejero, H., Marin, A., and Montero, F. (2011). The relationship between error catastrophe, survival of the flattest, and natural selection. *BMC Evolutionary Biology*, 11:2.
- Tromas, N., Zwart, M. P., Poulain, M., and Elena, S. F. (2014). Estimation of the in vivo recombination rate for a plant RNA virus. *Journal of General Virology*, 95:724–732.
- Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In *Proceedings of the Sixth International Congress on Genetics*, pages 355–366.