All in the same boat: A "situated" model of emergent immune response

Tom Hebbron, Jason Noble and Seth Bullock

University of Southampton, SO17 1BJ, UK

Abstract. Immune systems provide a unique window on the evolution of individuality. Existing models of immune systems fail to consider them as situated within a biochemical context. We present a model that uses an NK landscape as an underlying metabolic substrate, represents organisms as having both internal and external structure, and provides a basis for studying the coevolution of pathogens and host immune responses. Early results from the model are discussed; we show that interaction between organisms drives a population to optima distinct from those found when adapting against an abiotic background.

1 Immune systems

From our viewpoint as highly adapted biological agents ourselves, it may seem obvious that there is a boundary between any one individual, its environment, and other individuals in the population. However, if we look at the earliest forms of life, and at the major transitions in evolution, we find that the boundaries of what we would call an individual are not given *a priori*, but are themselves subject to evolutionary change. Furthermore, any form of life, from the simplest autocatalytic set through to a complex multi-cellular animal, represents a local concentration of resources in the environment that could be used by other neighbouring organisms. As the interests of the selfish genes in any opportunistic neighbour are rarely compatible with one's own, a wide variety of mechanisms have arisen to counter such exploitation [1]. The most elaborate of these mechanisms, and a characteristic marker of individuality, is the sophisticated defence of the self/non-self divide that we see in the adaptive immune systems of jawed vertebrates. In everyday life, the adaptive immune system is what makes vaccination possible, and explains why there are many diseases like chickenpox that you are unlikely to catch twice. Preceding adaptive immunity, innate immune mechanisms are a universal feature of even unicellular life [1]. The genome reflects the evolutionary history of pathogenic exposure and successful responses. The continuing selective advantage of these germline-encoded responses depends on the conservation of molecular features of pathogens; features crucial to their metabolism or membrane and difficult to mutate to non-recognised alternatives.

We believe that the study of the immune system is important because it provides a unique window on (and is perhaps even synonymous with) the evolution of individuality itself. Research in immunology is revealing more of the

specific mechanisms by which modern immune systems operate, but there is a gap in our understanding of the evolutionary history and adaptive landscape of immune responses. Artificial life provides the modelling tools to fill that gap.

Clearly, exploitation by pathogens creates selection pressure for a defensive response. But some organisms get by with much simpler immune responses than others. By simulating very simple ecologies we hope to discover the necessary and sufficient conditions for the emergence of an immune response, and in further work, more sophisticated adaptive immune systems. Evolutionary simulation also allows us to probe the consequences an immune system has on the subsequent adaptive landscape for hosts and pathogens.

Artificial immune systems (AIS) research is a growing field. In the same vein as earlier work on genetic algorithms and neural networks, researchers have noted that immune systems, and in particular the adaptive immune system, appear to have computationally useful properties (see [2] for a critical review). Computer security is the most popular application area: detecting computer viruses or spam emails for example.

Taking evolved solutions from nature and applying them to human problems is increasingly prevalent in many branches of engineering. However, when borrowing ideas from biology in this way, there can be pitfalls in both under- and over-reaching. For example, an AIS performing binary classification (e.g., "safe" / "dangerous") is not doing what a real immune system does. The algorithm may do well at the practical task we set for it, but it is not embedded in biochemistry in the way a real immune system is. There is a superficial similarity but we would be wrong to draw conclusions about immunobiology from the AIS. Conversely, building a high-definition model of immune system components might be a valuable scientific goal, but as engineering it would be a misplaced effort if the complex contextual constraints of the real immune system were not present in the target problem. It is important to look closely at the differences between the practical problems tackled by AIS researchers and the biological problem faced by real immune systems, as a false assumption of congruence will lead to either models that are too simple to be good science, or tools that are too byzantine to be good engineering.

Immunology to date has been primarily a medical science, concerned with questions of mechanism and ontogeny [3]. In other words, how does the immune system work, how might we fix it when it's broken, and what is its normal course of development? AIS research draws heavily from this work on proximate mechanism. The question of function however, is often left implicit — immune systems are *for* protecting against pathogens. This question warrants further exploration: for there to be pressure for an immune system in the first place requires certain ecological dynamics to hold: pathogenic behaviour must exist, and defences against it must be evolutionarily accessible. Once an immune system emerges it makes fundamental changes to the future evolutionary pathways available to a lineage; some adjacent possible phenotypes will be incompatible with the defenses now encoded in the genome. Simulation modelling is the logical tool for this approach: "For the past century immunology has, with great

success, been occupied with analyzing the immune system into its molecular building-blocks. The field is now ripe for synthesis." [4, p. 30].

Why should we be interested in modelling the function or adaptive value of immune responses? High-resolution, predictive models of biological mechanisms are certainly useful, for instance in developing therapeutics; but such models are opaque. We can see how they work, but not why they work. Such models are also open to misinterpretation: for instance, low iron levels or a fever may not be a symptom of infection but an immune response, creating an unfavourable environment for a pathogen. Without knowing the function of this mechanism, there is the danger that therapeutics are employed to lower the temperature or restore iron levels, working against the biological mechanisms [5]. This need for functional explanations to understand a mechanism in context is also an example of the "no free lunch" theorem [6]: in short, you cannot evaluate an algorithm for solving a problem without some knowledge of the problem and how it arose, and there are no universal solutions — the success of an algorithm on one problem does not necessarily translate to another. (Compare the notion in [7] that evolution produces *nichiversal* and not universal solutions.) This is significant for AIS practitioners, who are trying to isolate the design principles of a system particularly tightly embedded in its biological context.

Parasitic and subsequent coevolutionary behaviour has been observed in artificial life systems before, notably in Tierra [8], which was not explicitly designed to exhibit the phenomenon. Parasitism in Tierra is an interesting example for us because it shows that a fixed and quite brittle substrate defined by Tierra's virtual machine instruction set gave rise to parasitism and the beginnings of an immune response, suggesting that a more general model should find regions of biological interest without much difficulty.

To investigate functional questions about the evolution of the immune system, we must consider the major transitions in its history. We see these as, first, the isolation of metabolism from random perturbations through some form of membrane. Second, once proto-cellular individuals are present, predation will naturally follow as a strategy; i.e., exploiting other organisms as local concentrations of resources and the corresponding coevolution of defensive counter strategies. Third, the accumulation in the genome of these defences constitutes the emergence of innate immunity. (A fourth stage, not represented in the current model, is the emergence of somatic mutations, i.e., adaptive immunity.)

2   Constructing a model of immune function

A protective membrane will be assumed in our model. Other ALife researchers have modelled the low-level physicochemical interactions required for the emergence of membranes and proto-cells [9], but in our case starting from an artificial chemistry would make the observation of higher-level strategic events difficult. We abstract the underlying physicochemistry to a set of "metabolic components" and their interactions. An organism is a genetically specified phenotype built from these components, which interact to provide energy payoffs.

The degree of complexity of the interactions between components is a key parameter of the model; to tune this complexity we borrowed from Kauffman's NK model of tunably rugged fitness landscapes [10]. An N by N interaction matrix defines the energy gain or loss $[-1.0, 1.0]$ of all possible interactions between N metabolic components. Concentration is not modelled, so a metabolic 'state' in the model is represented simply by a bitstring of length N, i.e., a list specifying which components are present. The values in the interaction matrix specify the energy payoff for the column component against the background of the row component (see Figure 1). For example, the upper-right corner of the example matrix indicates that component 4 against a background of component 1 would result in an energy payoff of 0.6. Note that the matrix is not symmetrical, however. A function *payoff(a,b)* gives the energy payoff for state a against a background of state b. This is the sum of all the relevant component-by-component payoffs.

The values in the interaction matrix determine the ruggedness of the fitness landscape that a population of organisms has to search. For example, a simple "Mt Fuji" single-optima landscape can be defined by having only a single row j populated with non-zero energy payoff values. We expect that interesting regions of parameter space will be those with a slightly rugged landscape, not too simple, but not random. Interaction matrices for this type of landscape are generated using an NK-esque algorithm: each row of the matrix is filled with K values drawn from a normal distribution with mean 0.
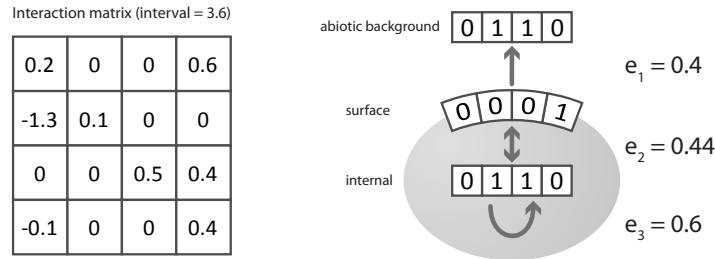

Interaction matrix (interval = 3.6)

0.2	0	0	0.6
-1.3	0.1	0	0
0	0	0.5	0.4
-0.1	0	0	0.4

abiotic background 0 1 1 0

surface 0 0 0 1

internal 0 1 1 0

$e_1 = 0.4$

$e_2 = 0.44$

$e_3 = 0.6$

Fig. 1: An example interaction matrix (left) and the calculation of an organism's net energy gain (right). For details see text.

The evolution of organisms in our model does not take place in a vacuum: a constant abiotic background solution is universally available. Made of the same components as organisms, it is represented by a state bitstring randomly chosen at initialisation, normally with K components present. Organisms themselves are represented by two state bitstrings (each of length N): the first encodes components present in the organism's internal structure, and the second encodes those present on the organism's external surface. This compartmentalisation protects the internal metabolism from outside perturbation and allows organisms to make use of highly reactive components in a controlled environment, whilst exposing selected components well adapted to the rewards and threats of the external environment. Complete decoupling is prevented by requiring that the two sets of components complement one another in order to transfer energy across the

membrane, into the organism for use towards reproduction. The energy chain is represented in the model by three values:

e_1: energy gained or lost in relation to the current external environment: a combination of the abiotic nutrient background bitstring and the surface bitstring of any currently interacting organisms. $e_1 = payoff(surface, context)$.

e_2: a measure of transmission efficiency across the membrane. Calculated as $payoff(surface, internal) + payoff(internal, surface)$, normalised using the interval of all matrix values to $[0.0, 1.0]$.

e_3: a measure of the efficiency of the organism's internal metabolism, calculated as $payoff(internal, internal) + payoff(internal, internal)$. This evaluates the internal components in the context of themselves.

These three factors are combined to give a total change in energy per timestep of $e_1 \times e_2 \times e_3$, except that if e_1 is negative, the change in energy is equal to e_1 alone, i.e., the organism suffers a loss in energy that is not modulated by its internal metabolism. If an organism's energy level falls below zero, it dies.

Organisms are asexual and reproduce when their stored energy reaches a threshold, which is dependent on the number of components present in their phenotype. In other words, it costs more for larger organisms to reproduce. If we define $onecount(string)$ as the number of ones in a bitstring, then the value of the reproduction threshold is $N + (10 \times (onecount(internal) + onecount(surface)))$. At reproduction, the current energy level is split equally between mother and daughter. Mutation of the daughter organism occurs through random flipping of bits in the surface and internal structure strings. A population is initialised by generating a random genotype and initial energy value $[0, N]$ for each organism. During simulation, at each time step the population is shuffled into a random order, this ordering is treated as a ring, and each organism interacts with its left and right neighbours. Energy gains and losses are updated, any offspring are added to the population, dead individuals are removed, and, if necessary, the population is pruned to a predefined carrying capacity by random reaping.

Given a rich interaction matrix and organisms that can interact with each other, the fitness implications of different phenotypes will be complex. In the runs described below, we wanted to limit this complexity by constraining organisms to interact initially only with the nutrient background. This reflects an early ecological stage in which a low-density population experiences no intra-specific competition for resources. In the experiments we will describe here, only a single population is used. However, the model is designed for extension to multiple populations to implement, for example, explicit coevolution between a population of hosts and a population of pathogens.

We anticipate that the primary problem for the evolving organisms in this model will be adaptation to the interaction matrix. This will involve finding a surface structure that works well with the nutrient background, and an internal structure that performs two roles: complementing the surface to pass energy across the membrane, and providing a good internal metabolism by functioning

efficiently in the context of itself. Adapting to the interaction matrix is complicated by the fact that the organisms can interact with each other. A surface structure that works well with the nutrient background may be vulnerable to exploitation by other organisms. Therefore we also expect to see the beginnings of a coevolutionary arms race; exploiting other organisms may give higher payoffs than interacting with the background alone, but when this starts to happen it also provides selection pressure for defensive surface structures, i.e., the beginnings of an immune response.

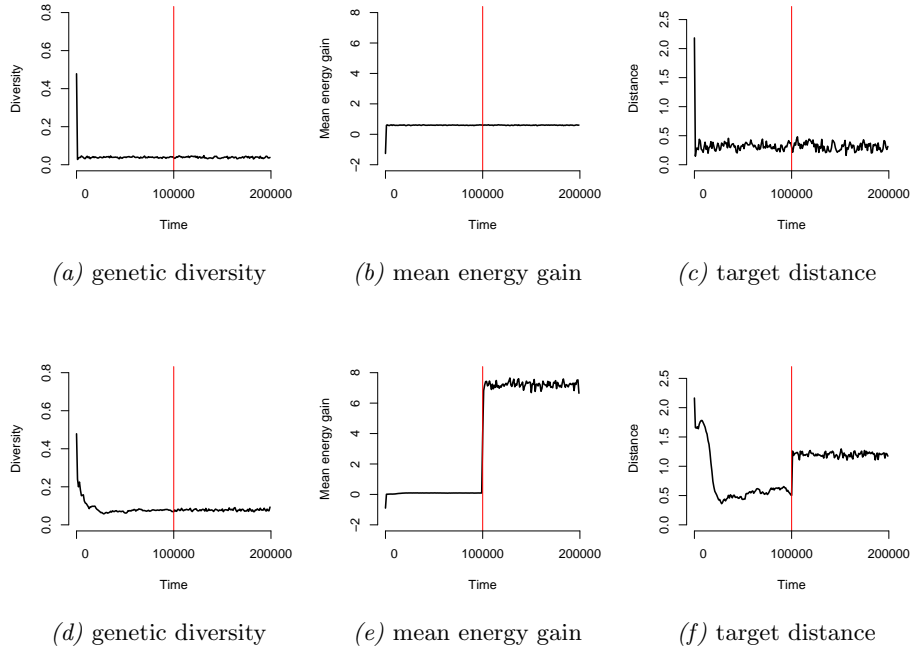## 3   Initial results and discussion

We used a population size of 1000, a mutation rate of 0.02 per locus, and a substrate with $N = 10$ components. The simulation was run for 100,000 timesteps allowing only interaction with the nutrient background, then for a further 100,000 timesteps with inter-organism interactions enabled. Multiple runs were performed, but only a single representative run is described as our aim here is to convey the qualitative dynamics of the model.

We collected data every 200 timesteps on the currently modal genotype, the genetic diversity of the population (measured as the mean disagreement with the modal genotype), the distance between the population centroid and the optimum reached at the end of the initial 100,000 timesteps, and the mean energy gain per timestep across the population.

To confirm that the model was functioning as intended, we devised a "Mt. Fuji" interaction matrix whereby only a single row j had non-zero values, with a minimum of 2 positive values. We paired this with a nutrient background where only component j was present. This meant that there was a global optimum genotype with component j present either on the surface or internally (to enable e_2 to be non-zero and energy to be transferred across the membrane) and any components with a positive interaction with j present both internally and on the surface. Figure 2 (top row a-c) shows the results for this condition.

Evolution given the Mt. Fuji interaction matrix confirmed our expectations. Many organisms in the initial population died off immediately as they had negative net energy gain. The few viable individuals rapidly converged on the anticipated global optimum. Enabling inter-organism interactions at $t = 100,000$ had no effect, but this was also expected as the nature of the payoff matrix meant that there were no benefits available from interaction that were not already available against the background.

Having established that the simulation was producing intelligible results, we constructed a richer interaction matrix to serve as an initial exploration of the space of possible substrates. As described in section 2, an NK-like interaction matrix was constructed with $K = 2$ and payoffs drawn from a normal distribution with a mean of zero and a standard deviation of 0.5. The corresponding nutrient background had two randomly selected components present. In this case, the landscape was more rugged so we expected greater difficulty for the population in finding an initial optimum. We also expected that once inter-organism inter-

(a) genetic diversity   (b) mean energy gain   (c) target distance

(d) genetic diversity   (e) mean energy gain   (f) target distance

Fig. 2: Population genetic diversity, mean per-timestep energy gain, and distance of population centroid from a target optimum over time. Graphs a-c show the Mt. Fuji case; graphs d-f show the rich matrix case. Vertical line at time 100,000 indicates the switching on of interaction.

actions were enabled, there would be a significant change in mean energy gain as many more interaction possibilities became available.

Results for this "rich matrix" condition are shown in Figure 2 (bottom row d-f). Compared to the "Mt. Fuji" condition, genetic diversity remains higher for longer, as the initial no-interaction population searches the rugged landscape and discovers multiple local optima. As the individuals on the fittest of these optima begin to dominate the population, diversity falls: Figure 2f confirms that there is significant movement in the population centroid over this period. Figure 2e shows that mean energy gain, on the other hand, increases only gradually. Once interaction is enabled at $t = 100,000$, new opportunities to exploit the surfaces of other organisms for energy gain become available. These translate as new and higher optima. Genetic diversity is not appreciably affected, but mean energy gain sees a large increase, and Figure 2f shows that the population rapidly converge on the new optima.

At $t = 100,000$ the modal genotype was `0011101100-0011111100` (internal-surface). At the end of the run ($t = 200,000$), and after the effects of interaction, the modal genotype was `0011101100-0011101000`. Note that the internal structure of the modal organism has not changed, but two surface components have been lost. If we examine the $e_1 \times e_2 \times e_3$ energy process for this final modal organism, we find that although it is well adapted to the interactive context, it in fact represents a phenotype that would not have been viable in the earlier

part of the run, i.e., its energy input in the context of the nutrient background, e_1, would have been zero. Note also that the surface and internal components are almost identical, differing only in one position. This makes sense: a state that works well with itself, as an internal structure, gives you a high e_3 value. If this structure is also your surface, then similarly e_2 will be high. Finally, by the same logic, if this string is also the modal surface structure, high values of e_1 will be gained on all sides when organisms interact with each other.

Thus it seems we have discovered a commensal evolutionarily stable state for a single population, rather than having demonstrated the kind of coevolutionary arms race we ultimately would like to see in the model. However, this is not entirely surprising, as coevolution is extremely difficult to engineer in a single population context. In the short term, the obvious next step for the model will be to achieve this by using multiple weakly interacting populations to avoid premature convergence. It may be necessary to explicitly define hosts and pathogens rather than wait for them to evolve; e.g., a host population with a low mutation rate and a minimum number of components per organism (hosts must be large) and a pathogen population with the opposite properties (fast-evolving and small). Beyond that, further development of the model framework will allow us to consider multi-cellular organisms, danger signals [11], and ultimately adaptive immune systems.

References

1. Danilova, N.: The evolution of immune mechanisms. J. Exp. Zool. Part B **306B**(6) (2006) 496–520
2. Garrett, S.M.: How Do We Evaluate Artificial Immune Systems? Evol. Comput. **13**(2) (June 2005) 145–177
3. Murphy, K.: Janeway's Immunobiology. Garland Science (2007)
4. Cohen, I.R.: Tending Adam's Garden. Academic Press (2004)
5. Nesse, R.M., Williams, G.C.: Why We Get Sick. Vintage (1996)
6. Wolpert, D.H., Macready, W.G.: No free lunch theorems for search. IEEE. T. Evolut. Comput. **1**(1) (1997) 67–82
7. Bullock, S.: The fallacy of general purpose bio-inspired computing. In Rocha, L.M., et al., eds.: Tenth Intl. Conf. on Artificial Life, MIT (2006) 540–545
8. Ray, T.S., Xu, C.: Measures of evolvability in Tierra. Artif. Life and Robotics **5**(4) (2001) 211–214
9. McMullin, B.: Thirty years of computational autopoiesis: a review. Artif. Life **10**(3) (2004) 277–295
10. Kauffman, S.A.: The Origins of Order. OUP (1993)
11. Matzinger, P.: The danger model: a renewed sense of self. Science **296**(5566) (2002) 301–305