

## Accepted Manuscript

Diversity loss is predicted to increase extinction risk of specialist animals by constraining their ability to expand niche

Daniel T. Reed , Colin R. Tosh

PII: S0022-5193(19)30214-0  
DOI: <https://doi.org/10.1016/j.jtbi.2019.05.016>  
Reference: YJTBI 9917



To appear in: *Journal of Theoretical Biology*

Received date: 3 July 2018  
Revised date: 15 March 2019  
Accepted date: 23 May 2019

Please cite this article as: Daniel T. Reed , Colin R. Tosh , Diversity loss is predicted to increase extinction risk of specialist animals by constraining their ability to expand niche, *Journal of Theoretical Biology* (2019), doi: <https://doi.org/10.1016/j.jtbi.2019.05.016>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Highlights

- We model animal cognition using neural networks foraging on artificial resources.
- Networks specialise in diverse and resource-abundant historical environments.
- Low diversity environments constrain the ability of specialists to expand range.
- Monoculture farming could have an increased impact on specialist extinctions.
- Negative effects from pesticides and GM defences will further compound this issue.

ACCEPTED MANUSCRIPT

**Diversity loss is predicted to increase extinction risk of specialist animals by  
constraining their ability to expand niche**

Dr Daniel T. Reed and Dr Colin R. Tosh\*

School of Biology, Newcastle University, Ridley Building 2, Newcastle upon Tyne NE1  
7RU, UK

\*Corresponding author: Dan Reed

ndr45@newcastle.ac.uk

Keywords: niche, specialisation, animal behaviour, evolution, dead-end, extinction

**Abstract**

Specialist animals are at a greater risk of extinction in the face of environmental change than generalist ones. The inability of some specialist taxa to expand host range through evolution may exacerbate or cause their high extinction risk. Here we use connectionism (a framework for modelling animal behaviour) to predict the environmental and physiological factors that predispose some specialist taxa to an 'evolutionary dead-end'. Neural networks are evolved to become resource-specialised in a resource-abundant and resource-diverse 'historical' environment while losing 'genes' that should restrict their ability to expand their host range. Networks are subsequently challenged to escape their dead-end by expanding host range in a 'contemporary' environment that may have depleted resource abundance and diversity (as many human impacted environments do). Loss of diversity in available resources universally constrains the ability of networks to expand host range and this effect is very robust to network conformation. Environmental resource abundance is more variable in its effect. Networks are generally robust to loss of genetic diversity during the evolution of specialisation except at very high rates of loss. By omitting historical specialisation, we show that the effect of resource diversity on host range expansion is not a universal network property but something that is often specific to specialist organisms. Historical specialisation also slightly reduces the robustness of networks in the contemporary environment to loss of genetic diversity during the specialisation process. Fundamentally, simulations predict that loss of local resource diversity will further increase the vulnerability of specialists to extinction by containing their ability to expand host range in the face of environmental change.

## Introduction

In attempting to minimise the impact of humans on the natural environment, ecologists identify the traits that predispose species to extinction so that these species can be prioritised for conservation (McKinney, 1997; Purvis *et al.*, 2000; Kotiaho *et al.*, 2005; Lee and Jetz, 2011; Urban, 2015). One such trait is niche specialisation and specialist species in many taxa are at higher risk of extinction than generalists (Harcourt *et al.*, 2002; Colles *et al.*, 2009; Clavel *et al.*, 2011; Gallagher *et al.*, 2015). As organisms may evolve adaptively in response to environmental change (Gienapp *et al.*, 2008; Harmon *et al.*, 2009; Hoffmann and Sgrò, 2011), one mechanism that could contribute to or cause this relationship is the tendency of some specialist species to become stuck in ‘evolutionary dead ends’ from which they are unable to escape, even if niche expansion is selectively advantageous. It was thought that most specialist organisms are stuck in such dead ends but this view is now discredited (Simpson, 1955; Moran, 1988; Forister *et al.*, 2012; Vamosi *et al.*, 2014; Day *et al.*, 2016). At best there may be a slight overall bias towards the evolution of specialisation (vs generalisation), but in some taxa this bias is stronger (Forister *et al.*, 2012; Vamosi *et al.*, 2014; Day *et al.*, 2016). With this research field at the juncture just described it is important to determine the environmental and organismal traits that predispose certain taxa to irreversible specialisation while others escape this fate, and that is the purpose of this article.

We take a modelling approach, using connectionism (Flusberg and McClelland, 2014) to simulate the evolution of behaviour, because behavioural evolution is a major part of any niche contraction or expansion event. Using simple neural network models to model the evolution of animal behaviour (Bain *et al.*, 2007; Lewis *et al.*, 2010; Enquist and Ghirlanda, 2013; Ferrauto *et al.*, 2013; Flusberg and McClelland, 2014) in the natural environment is

attractive for a number of reasons. Neural networks are a core component of many artificially intelligent computational systems and are designed to mimic real animal behaviour. They embody some biologically relevant components such as communicating artificial neurons. Their behaviour can be evolved easily using processes analogous to natural selection. They are also less demanding of processor time than the more complex models used in computational neuroscience.

We consider how three factors at the heart of much ecological research over the last 50 years; resource diversity, resource abundance, and relative host/non-host value (the nature of adaptive trade-off (Fry, 1996) which is based on physiological adaptation to resources), impact the ability of animals to expand resource range following prolonged resource specialisation. It is thought that specialist organisms become stuck in evolutionary dead ends through prolonged loss of genetic variation, due either to intense selection for specialisation or drift in genes not under selection (Futuyma and Moreno, 1988; Jaenike, 1990; Forister *et al.*, 2012; Vamosi *et al.*, 2014). Using a traditional evolutionary genetic framework, the dynamics of gene loss and behavioural constraint following specialisation is fairly trivial, but in reality genes impact components of the complex systems (neural networks) that generate behaviour, and the dynamics of such systems are more complex. Here we embody behavioural evolution in a genetic algorithm where ‘genes’ determine ‘synaptic’ properties of a neural network producing an input (environment) – output (behaviour) mapping (Figure 1). This system has been used successfully to model the evolutionary ecology of animal niche specialisation (Tosh *et al.*, 2009).

Networks are evolved to become behaviourally specialised in an ‘historical’ environment that is both resource diverse and resource abundant. We then assume a transition to an environment where resources may be less diverse and abundant (for example a human-impacted environment). This altered environment we refer to as the ‘contemporary’ environment. Simultaneously, we fix different proportions of the genes determining network function to represent the loss of genetic variation that accompanies niche specialisation. In this contemporary environment we assume that niche expansion is selectively advantageous and we examine the ability of networks to evolve generalist behaviour and so escape their specialised dead-end (Figure 1 Part C). Simulations predict that loss of resource diversity within the geographic range of specialist species will further increase their vulnerability to extinction by constraining their ability to expand host range in the face of environmental change. Specialisation also lowers the robustness of networks in the contemporary environment to the loss of genetic diversity that is thought to occur during the specialisation process in the historical environment.

## **Methods**

### **Scenarios Considered**

Broadly we imagined a scenario where an organism evolves resource specialisation in an ‘historical’ environment that is both resource diverse and abundant, all the while losing ‘genes’ that could potentially help the network to later expand resource range. Subsequently, the environment of this organism changes such that it may be less diverse and abundant. We now assume that it is advantageous for the organism to expand host range through evolution and analyse its ability to do so. From a biological perspective an individual organism would simply be any network in a simulation. As each simulation used 50 networks created together

then these could be seen as individuals of the same species with different populations of 50 organisms/networks being drawn for each simulation.

Essentially here we take a connectionist approach in this paper. Connectionism is a method for modelling animal behaviour that assumes processing information in a parallel and distributed manner is fundamental to the types of behaviour that emerge from neural network systems (Flusberg and McClelland, 2014). For this reason, connectionists feel justified in using grossly simplified and generic representations of cognitive processing just as many fundamental aspects of population biology can be captured using grossly simplified models of population dynamics (May, 2004). Non-generic aspects of our model (different network outputs depending on where appropriate objects are on the retina) are inspired by the coding of visuospatial processing in higher animals (Deco and Rolls, 2004).

Networks were evolved over 500 generations to preferentially select only one of the 40 resources (Figure 1B) in an environment that was resource abundant and resource diverse. An environment of high resource abundance had 5-8 sections of the network input surface occupied by resources selected at random from the 40 in Figure 1B, in each of the 250 projections. An environment of low resource abundance (shown in Figure 1A) had 1-4 sections occupied with resources. In a high diversity environment, resources for placement into network input sections were selected at random from all 40 resources. In a low diversity environment only 20 resources were used for selection. Further specifics can be found in Figure 1B. Improvement in network performance during this stage, when trained as a specialist (choosing a single host) is shown in the Supporting Information (Figure S1).

At the end of the 500 generation specialisation period a percentage of the network weights were fixed and no longer allowed to mutate, representing different levels of loss of genetic diversity that may occur during the specialisation process. Fixing weights was done in steps of 10% weight fixation from 0% through to 100%. To simplify coding we randomly sampled one set of weights from our genetic algorithm's population of 50 and inserted these into the other 49 networks. These weights were subsequently invariant while all other weights could continue to recombine and mutate as described above. This procedure required that variation in the weights of each network in the genetic algorithm's population of 50 was slight and that weights of a single network largely represented all networks in the population. To validate the process, we undertook a pilot study in which we trained 50 populations of 50 networks to specialisation for 500 generations then selected 30% of their 600 weights and compared these to the other weights at that position in their population. Overall mean variance and mean standard deviation of the final weights after the 500 generations were 0.0374 and 0.0506 respectively ( $n = 450000$ ). The mean value of weights across all control networks was 0.5320.

After the 500 generation specialisation period each network then underwent a further 500 generations of evolution in which they were selected to generalise and preferentially choose all of 10 resources (Figure 1B) from the larger pool. This could occur in environments of either high or low abundance or diversity as defined above. The value of non-hosts (either -1 or 1) was kept the same across all 1000 generations of each simulation. Figure 1C is a summary of scenario combinations considered. It was necessary to standardise how well networks under the various scenarios evolved to generalist behaviour at the end of the 500 generation in the 'contemporary' environment because randomly behaving networks vary in fitness in environments of different diversity, abundance and value of non-hosts. A randomly

behaving network in an environment of high abundance and a slight reward for non-host selection will be fitter than the same network in a low abundance environment with a punishment for selecting a non-host. We used the metric ‘progress towards perfect generalist behaviour’. For each simulation run under the different scenarios, we firstly calculated fitness when that network was behaving perfectly and maximising its fitness. We then divided the actual fitness observed at the end of 500 generations in the contemporary environment by this maximum fitness. This gave a metric between 1 and 0 where 1 represents the case where a network has successfully evolved completely optimal generalist behaviour, and 0 when the network has been unable to increase its fitness at all (see vertical axis of Figure 2).

### The Network

The network described here (Tosh *et al.*, 2009) and corresponding to Figure 2 B(i) and (ii) is only one of four network types used. Modifications to this network are described later in the Materials and Methods section. We created 2-layer artificial neural networks with 200 input units and 3 output units. Our networks were fully connected; each input was connected to each of the output units. Output units were binary stochastic elements with  $p_i$ , the probability of firing of the  $i$ th unit, defined by

$$P_i = g \left( \sum_{j=1}^M w_{ij} x_j \right)$$

where  $g(x)$  is the binary sigmoid function  $g(x) = 1/(1+\exp(-x))$ , the  $j$ th input layer unit provides input  $x_j$  to the  $i$ th unit via the connection  $w_{ij}$ , and  $M$  is the number of inputs to the unit.

For the initialising weight arrays each weight was given a value between 1 and -1 for each of the 600 connections. Choosing initial weight values in this general region is a standard method for initialisation (Gallant, 1993; Hinton and Salakhutdinov, 2006; Enquist and Ghirlanda, 2013). Other methods for initialising weights mainly concentrate on reducing the training time rather than final network performance (Psaltis *et al.*, 1988; Nguyen and Widrow, 1992; Yam *et al.*, 1997; Yam and Chow, 2000).

For the network inputs, 40 resources were created using 5x5 arrays of 1s and 0s. (Figure 1, B). Four types of resource were created in groups of ten: bilateral symmetry, radial symmetry, asymmetric and random conformation. The number of pixels in each resource (regardless of its form: asymmetric, symmetric etc.) was rounded from a random sample of the normal distribution with mean of 11.8 and variance 3.1. For more information, see (Tosh *et al.*, 2009).

The input layer for our networks was split into 8, 5x5 sections with each section having a resource projected on to it (host, non-host or empty space) (Figure 1, A). Networks could output one of eight binary output codes, each corresponding to one of the 5x5 sections of the input layer. Starting from the top left section of Figure 1 A and working clockwise (1, 1, 1) = section 1, (1, 1, 0) = 2, (1, 0, 0) = 3, (0, 0, 0) = 4, (0, 0, 1) = 5, (0, 1, 1) = 6, (1, 0, 1) = 7, (0, 1, 0) = 8. If the network output (1,1,1), for example, section 1 of the input layer was examined to determine if a resource was projected within this section. If so, scores were applied which determined ultimate fitness of a network within the genetic algorithm. The basic input–output mapping is not unlike certain systems of visuospatial processing in higher animals (Rolls and

Deco, 2002); however, the model is stripped of all but the most fundamental elements of neural processing (namely parallel distributed processing).

If the network selected an input that was empty, we considered this to be analogous to ‘continuing search’ which had a negligible impact on fitness and was awarded 0 points. Networks were forced to make a single choice and ‘commit’ to a resource as biologically this is what we would expect to see in search behaviours, especially in specialist plant feeding insects. They search till they make a decision and then alight. This is not necessarily a correct choice so they then may move on to make a new choice. The important aspect is that the choice is made each time which is the important feature here. If the network chose a host resource from the input projection it was awarded 5 points. This host resource value was fixed at 5 across all simulations, only the number of potential hosts changed. Networks training as specialists had only a single resource designated as a host but when trained as a generalist a network had 10 possible host resources. The input set used as hosts was the asymmetric set of possible shapes. This choice was arbitrary and previous work (Tosh *et al.*, 2009) has shown that results are robust to the input set used. If a network selected a resource designated as a non-host, the network was either punished with a score of -1 or rewarded with a score of 1. These non-host values were maintained across all 1000 generations of a simulation, so a network that specialised for 500 generations with a -1 non-host value would then generalise for 500 generations with this same value. It may seem counterintuitive to readers to reward a mistake but many organisms do not always suffer from utilising non-hosts resources and can sometimes do quite well if forced to utilise them (Fry, 1996). We considered it intuitively reasonable to allocate a lower absolute score to selection of a non-host than that allocated to selection of a host as most organisms have avoidance mechanisms

to minimise the negative impact of mistakes. The resources designated ‘host’ and ‘non-host’ are shown in Figure 1B.

Two hundred and fifty ‘projections’, where a number of resources, selected randomly from the pool of 40 in Figure 2B, stimulate the input layer (Figure 1A), were input into the network and scores summed across projections to determine overall fitness of a network. The host was fixed as one specific resource of the pool of 40 available, however, as they were chosen randomly, not all projections contained a ‘host’.

### **The Genetic Algorithm**

Fifty networks were created representing the ‘population’ of the genetic algorithm. The top five scoring networks were chosen for mating and paired at random, each mating resulting in recombination with a probability of 0.6. During recombination a position in the 600 element weight array was selected at random and halves swapped between networks. Subsequently, each weight was mutated with a probability of 0.1 with the amount to be added or subtracted to the present value sampled at random from a normal distribution with mean 0 and variance 3. Simulations were repeated 10 times to account for stochastic variation in starting parameters and other stochastic elements and a further 10 times varying which asymmetric resources (Figure 2B) networks specialised or generalised to (see below). Overall fitness was a grand mean calculated first across stochastic repeats and then across shape variant repeats.

### **Generality of Key Effects**

So far we have described only one neural network conformation. We wished to know whether key effects observed were robust to network conformation. We repeated all simulations adding an extra ‘hidden’ layer of 100 stochastic binary neurons into the network (Figure 2C(i)(ii)). Weight recombination and mutation of weights between layers 2 and 3 occurred as described above for weights in the 2-layer network. Addition of this extra complexity rendered simulations more processor-intensive so we only ran simulations at weight fixation rates of 0, 30, 70 and 100%. We returned to the original 2-layer network and removed the stochastic binary component from neurons such that they now output the raw value from the transfer function (Figure 2D(i)(ii)). Lastly, using the original 2-layer network, we shuffled the output encoding (Figure 2E(i)(ii)). Now  $(1, 1, 1) = \text{section 6}$ ,  $(1, 1, 0) = 4$ ,  $(1, 0, 0) = 7$ ,  $(0, 0, 0) = 1$ ,  $(0, 0, 1) = 2$ ,  $(0, 1, 1) = 5$ ,  $(1, 0, 1) = 8$ ,  $(0, 1, 0) = 3$ .

## Results

High resource diversity in the contemporary environment always has a positive impact on the ability of networks to expand resource range. The impact of resource diversity is always greatest in contemporary environments of low resource abundance (Figure 2 B(i)(ii) to E(i)(ii)). In absolute terms, the impact of resource abundance in contemporary environments varies with network conformation. When we state that loss of diversity constrains expansion of host range we are comparing networks in environments with the same resource abundance. Therefore, we would expect that when abundance is at 50% and diversity is at 100% (Figure 2: red lines) networks will perform better than when abundance is 50% and diversity is 50% (Figure 2: green lines). Likewise, when abundance is at 100% and diversity is at 100% (Figure 2: orange lines) networks will outperform those with 100% abundance and 50% diversity (Figure 2: blue lines). When networks are punished for selecting a non-host resource during evolution, ability to

generalise in high diversity contemporary environments either lies within the range of the same measure in low diversity environments or is below the range in low diversity environments (Figure 2 B(i) to E(i)). When networks receive a small reward for selecting a non-host during evolution, specialists are always better able to evolve generalist behaviour in contemporary environments of low resource abundance (Figure 2 B(ii) to E(ii)).

Omitting evolution of specialisation in the historical environment has little impact on evolution in the contemporary environment when networks are rewarded slightly during evolution for selecting a non-host (compare Figure 2 A(ii) to B(ii)). This is not true of simulations where networks are punished for selecting a non-host. Here resource diversity effects in the contemporary environment are little impacted when resource abundance is low but when resource abundance is high these diversity effects are reversed and only facilitate evolution of generalist behaviour in the contemporary environments when networks have undergone historical specialisation (compare Figure 2 A(i) to B(i) - E(i)). Thus in at least a subset of our simulations, the benefits of high resource diversity in the contemporary environment are only realised in organisms that are specialised.

Generally, networks are robust to loss of genetic diversity during the evolution of specialisation. Negative impacts on 'recovery' within the contemporary environment increase with gene fixation rate and major impacts only begin to be realised after around 70% gene fixation (Figure 2 B(i)(ii) to E(i)(ii)). Omitting historical specialisation does however increase the robustness of networks to loss of genetic diversity (compare Figure 2 A(i) and B(i), A(ii) and B(ii)), thus historical specialisation has a slight negative impact on robustness of networks in the contemporary environment to historical genetic diversity loss.

## Discussion

The central finding of these simulations is that low resource diversity within the ‘contemporary’ environment constrains the ability of specialist networks to expand resource range through evolution, presumably increasing the susceptibility of specialists to extinction. Our demonstration that these impacts of diversity are specific to specialists under some conditions (specifically when contemporary environments are resource abundant and organisms are punished for selecting non-hosts by mistake) adds extra significance to our predictions: if effects are not specific to specialists they are unlikely to contribute to the observed increased susceptibility of specialist organisms *per se* to extinction (Harcourt *et al.*, 2002; Colles *et al.*, 2009; Clavel *et al.*, 2011; Gallagher *et al.*, 2015).

Our results would suggest that many human impacted environments such as intensively farmed crops, open pasture or logged areas will have a greater impact on specialists than was initially thought due to the constraints these low resource diversity areas will have on specialists’ ability to evolve to make use of new potential hosts. Previous work (Dunn, 2005; Clavel *et al.*, 2011) has illustrated that specialist insects in particular are more susceptible to extinction and that this can have a knock-on effect in other species that rely on these specialists. Mass homogenisation of their principal resources, plants, through intensive agriculture, logging etc. has been shown to be a major factor in driving this (Fonseca, 2009). Additionally, this will be compounded by the finding that a non-host which carries a negative penalty will further impact this ability. Use of pesticides and GM crop defences will impose this negative penalty on specialist plant feeding insects for example. These resource

abundant, low diversity areas are now common globally and as such could be helping drive the rapid extinction of specialist organisms.

Ours is not the first study to indicate that biodiversity loss can predispose remaining species to local or global extinction. Mechanisms that mediate this effect include: loss of keystone species, increased exposure to competition, and loss of resources that co-existing species depend upon (Borrvall *et al.*, 2000; Koh *et al.*, 2004; Eklöf and Ebenman, 2006; Maynard *et al.*, 2017). Other studies have also shown that organisms may struggle to evolve required niche characteristics in response to dramatic contemporary rates of environmental change (Quintero and Wiens, 2013). Our study is distinctive in linking the effect of resource diversity on extinction risk specifically to specialists, invoking a cognitive mechanism, and implicating restriction of niche width expansion following diversity loss as a key mechanism. How might specialisation, cognition, and niche expansion unite following diversity loss to predispose species to extinction? We suspect the mechanism is fairly simple. Generally, when organisms switch or expand resource range through evolution they do so to resources that are similar in key respects to those from which they came. For example, plant eating insects, which tend to be specialised, have a tendency to evolve onto plants that are closely related and chemically similar to their ancestral host (Becerra, 1997; Pearse and Hipp, 2009). High environmental resource diversity may simply by chance increase the probability of specialists encountering resource traits that are stimulating and therefore facilitate the move onto new resources. Generalists must presumably be sensitive to more resource traits and may therefore benefit less from exposure to this diverse range of traits.

Our study generates the testable prediction that specialist organisms in resource diversity-depleted contemporary environments should be at particular high risk of extinction. This prediction should be amenable to testing using pre-existing comparative methods and data sets that have been used to establish the relationship between specialisation and extinction to date (Harcourt *et al.*, 2002; Colles *et al.*, 2009; Clavel *et al.*, 2011; Gallagher *et al.*, 2015). The more specific prediction that this effect is mediated through cognition may require additional experiments. Artificial selection experiments with specialist species of *Drosophila* could be undertaken, analysing the ability of flies to expand resource range in lab environments of high and low odour diversity, for example.

It is no surprise that we found networks are robust to functional constraint ('gene' loss in our simulations) as neural networks are well known for this property (Haykin, 2004) as are other systems that employ parallel distributed processing such as the brain. People and animals that lose the use of one sense, for example, often compensate by increasing the sensitivity of other senses (Chapman *et al.*, 2010; Merabet and Pascual-Leone, 2010) and people born with only rudiments of normal brain structure may lead relatively normal lives (Fry, 1996) (note that these examples involve structural plasticity as well as redistribution of activity within existing networks). Nevertheless, we also showed that specialisation in the historical environment lowers the robustness of networks in the contemporary environment to loss of diversity in the weights that determine its behaviour. It should be borne in mind that specialisation in our simulations occurred across only 500 generations. In real organisms, where evolution of specialisation may occur over many more generations, this effect of specialisation on network robustness may be more profound and could restrict the ability of specialists to adapt to changing contemporary environments.

**Acknowledgements**

The work presented in this paper was supported by a grant from the Natural Environment Research Council (NERC).

ACCEPTED MANUSCRIPT

## References

- Bain, R.S., Rashed, A., Cowper, V.J., Gilbert, F.S. and Sherratt, T.N. (2007) 'The key mimetic features of hoverflies through avian eyes', *Proceedings of the Royal Society of London B: Biological Sciences*, 274(1621), pp. 1949-1954.
- Becerra, J.X. (1997) 'Insects on plants: macroevolutionary chemical trends in host use', *Science*, 276(5310), pp. 253-256.
- Borrvall, C., Ebenman, B., Jonsson, T. and Jonsson, T. (2000) 'Biodiversity lessens the risk of cascading extinction in model food webs', *Ecology Letters*, 3(2), pp. 131-136.
- Chapman, B.B., Morrell, L.J., Tosh, C.R. and Krause, J. (2010) 'Behavioural consequences of sensory plasticity in guppies', *Proceedings of the Royal Society of London B: Biological Sciences*, 277(1686), pp. 1395-1401.
- Clavel, J., Julliard, R. and Devictor, V. (2011) 'Worldwide decline of specialist species: toward a global functional homogenization?', *Frontiers in Ecology and the Environment*, 9(4), pp. 222-228.
- Colles, A., Liow, L.H. and Prinzing, A. (2009) 'Are specialists at risk under environmental change? Neocological, paleoecological and phylogenetic approaches', *Ecology Letters*, 12(8), pp. 849-863.
- Day, E., Hua, X. and Bromham, L. (2016) 'Is specialization an evolutionary dead end? Testing for differences in speciation, extinction and trait transition rates across diverse phylogenies of specialists and generalists', *Journal of evolutionary biology*, 29(6), pp. 1257-1267.
- Deco, G. and Rolls, E. (2004) *Computational neuroscience of vision*. New York, New York, USA: Oxford University Press.
- Dunn, R.R. (2005) 'Modern insect extinctions, the neglected majority', *Conservation biology*, 19(4), pp. 1030-1036.
- Eklöf, A. and Ebenman, B. (2006) 'Species loss and secondary extinctions in simple and complex model communities', *Journal of animal ecology*, 75(1), pp. 239-246.
- Enquist, M. and Ghirlanda, S. (2013) *Neural networks and animal behavior*. Princeton University Press.
- Ferrauto, T., Parisi, D., Di Stefano, G. and Baldassarre, G. (2013) 'Different genetic algorithms and the evolution of specialization: A study with groups of simulated neural robots', *Artificial life*, 19(2), pp. 221-253.
- Flusberg, S.J. and McClelland, J.L. (2014) *Connectionism and the emergence of mind*. New York, NY, USA: Oxford University Press.
- Fonseca, C.R. (2009) 'The silent mass extinction of insect herbivores in biodiversity hotspots', *Conservation Biology*, 23(6), pp. 1507-1515.
- Forister, M., Dyer, L.A., Singer, M., Stireman, J.O. and Lill, J. (2012) 'Revisiting the evolution of ecological specialization, with emphasis on insect–plant interactions', *Ecology*, 93(5), pp. 981-991.
- Fry, J.D. (1996) 'The evolution of host specialization: are trade-offs overrated?', *The American Naturalist*, 148, pp. S84-S107.
- Futuyma, D.J. and Moreno, G. (1988) 'The evolution of ecological specialization', *Annual Review of Ecology and Systematics*, 19(1), pp. 207-233.
- Gallagher, A.J., Hammerschlag, N., Cooke, S.J., Costa, D.P. and Irschick, D.J. (2015) 'Evolutionary theory as a tool for predicting extinction risk', *Trends in Ecology & Evolution*, 30(2), pp. 61-65.
- Gallant, S.I. (1993) *Neural network learning and expert systems*. MIT press.
- Gienapp, P., Teplitsky, C., Alho, J., Mills, J. and Merilä, J. (2008) 'Climate change and evolution: disentangling environmental and genetic responses', *Molecular ecology*, 17(1), pp. 167-178.
- Harcourt, A.H., Coppeto, S. and Parks, S. (2002) 'Rarity, specialization and extinction in primates', *Journal of Biogeography*, 29(4), pp. 445-456.
- Harmon, J.P., Moran, N.A. and Ives, A.R. (2009) 'Species response to environmental change: impacts of food web interactions and evolution', *Science*, 323(5919), pp. 1347-1350.

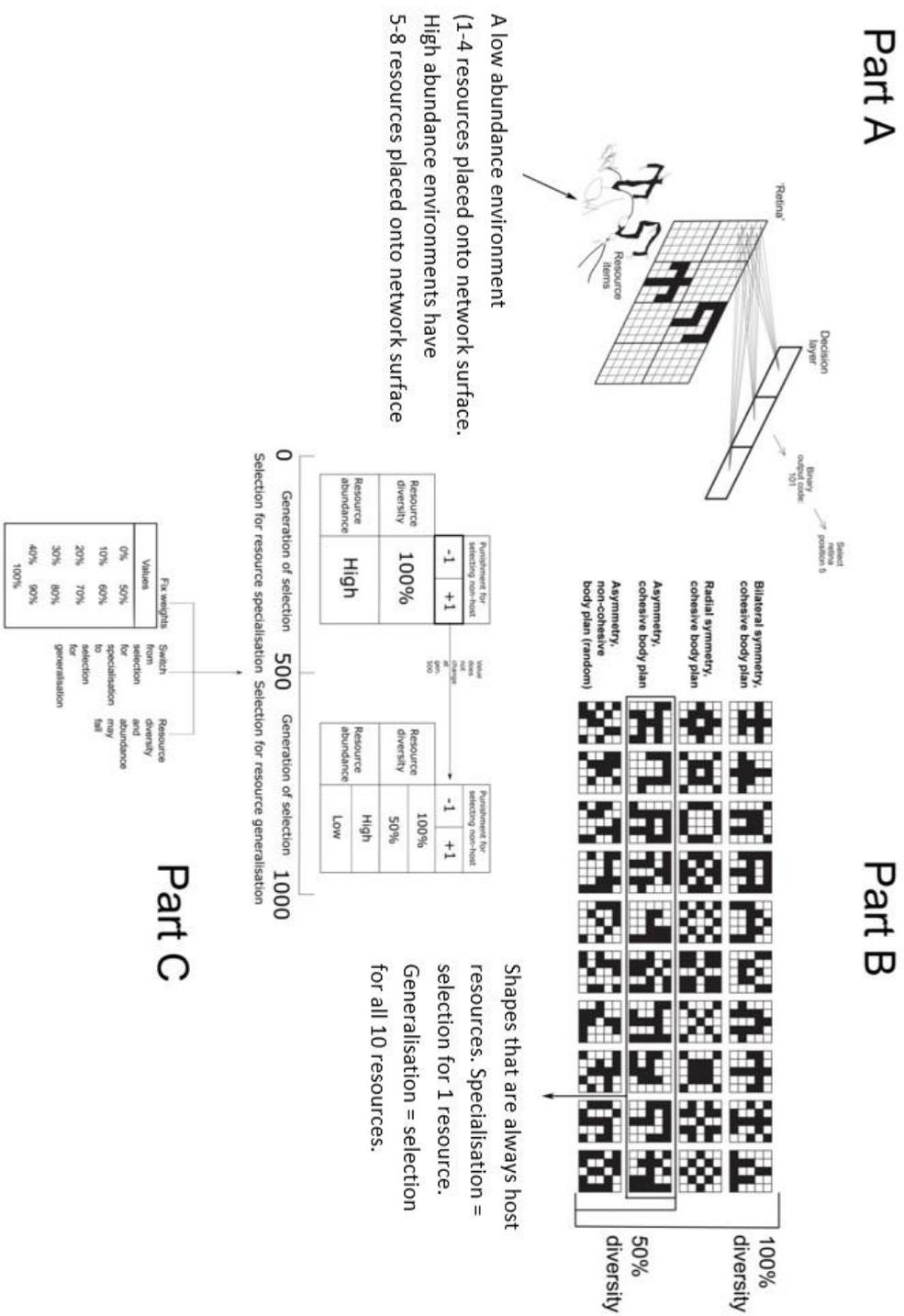
- Haykin, S. (2004) 'Neural networks: a comprehensive foundation', *Neural networks*, 2(2004), p. 41.
- Hinton, G.E. and Salakhutdinov, R.R. (2006) 'Reducing the dimensionality of data with neural networks', *science*, 313(5786), pp. 504-507.
- Hoffmann, A.A. and Sgrò, C.M. (2011) 'Climate change and evolutionary adaptation', *Nature*, 470(7335), p. 479.
- Jaenike, J. (1990) 'Host specialization in phytophagous insects', *Annual Review of Ecology and Systematics*, 21(1), pp. 243-273.
- Koh, L.P., Dunn, R.R., Sodhi, N.S., Colwell, R.K., Proctor, H.C. and Smith, V.S. (2004) 'Species coextinctions and the biodiversity crisis', *science*, 305(5690), pp. 1632-1634.
- Kotiaho, J.S., Kaitala, V., Komonen, A. and Päävinen, J. (2005) 'Predicting the risk of extinction from shared ecological characteristics', *Proceedings of the National Academy of Sciences of the United States of America*, 102(6), pp. 1963-1967.
- Lee, T.M. and Jetz, W. (2011) 'Unravelling the structure of species extinction risk for predictive conservation science', *Proceedings of the Royal Society of London B: Biological Sciences*, 278(1710), pp. 1329-1338.
- Lewis, H., Tosh, C., O'KEEFE, S., Shuker, D., West, S. and Mayhew, P. (2010) 'Constraints on adaptation: explaining deviation from optimal sex ratio using artificial neural networks', *Journal of evolutionary biology*, 23(8), pp. 1708-1719.
- May, R.M. (2004) 'Simple mathematical models with very complicated dynamics', in *The Theory of Chaotic Attractors*. Springer, pp. 85-93.
- Maynard, D.S., Bradford, M.A., Lindner, D.L., van Diepen, L.T., Frey, S.D., Glaeser, J.A. and Crowther, T.W. (2017) 'Diversity begets diversity in competition for space', *Nature ecology & evolution*, 1(6), p. 0156.
- McKinney, M.L. (1997) 'Extinction vulnerability and selectivity: combining ecological and paleontological views', *Annual Review of Ecology and Systematics*, 28(1), pp. 495-516.
- Merabet, L.B. and Pascual-Leone, A. (2010) 'Neural reorganization following sensory loss: the opportunity of change', *Nature Reviews Neuroscience*, 11(1), p. 44.
- Moran, N.A. (1988) 'The evolution of host-plant alternation in aphids: evidence for specialization as a dead end', *The American Naturalist*, 132(5), pp. 681-706.
- Nguyen, D. and Widrow, B. (1992) *Proceedings of IEEE International Joint Conference on Neural Networks*.
- Pearse, I.S. and Hipp, A.L. (2009) 'Phylogenetic and trait similarity to a native species predict herbivory on non-native oaks', *Proceedings of the National Academy of Sciences*, 106(43), pp. 18097-18102.
- Psaltis, D., Sideris, A. and Yamamura, A.A. (1988) 'A multilayered neural network controller', *IEEE control systems magazine*, 8(2), pp. 17-21.
- Purvis, A., Gittleman, J.L., Cowlshaw, G. and Mace, G.M. (2000) 'Predicting extinction risk in declining species', *Proceedings of the Royal Society of London B: Biological Sciences*, 267(1456), pp. 1947-1952.
- Quintero, I. and Wiens, J.J. (2013) 'Rates of projected climate change dramatically exceed past rates of climatic niche evolution among vertebrate species', *Ecology letters*, 16(8), pp. 1095-1103.
- Rolls, E.T. and Deco, G. (2002) *Computational neuroscience of vision*. Oxford university press.
- Simpson, G.G. (1955) *Major features of evolution*. Columbia University Press: New York.
- Tosh, C.R., Krause, J. and Ruxton, G.D. (2009) 'Theoretical predictions strongly support decision accuracy as a major driver of ecological specialization', *Proceedings of the National Academy of Sciences*, 106(14), pp. 5698-5702.
- Urban, M.C. (2015) 'Accelerating extinction risk from climate change', *Science*, 348(6234), pp. 571-573.
- Vamosi, J.C., Armbruster, W.S. and Renner, S.S. (2014) 'Evolutionary ecology of specialization: insights from phylogenetic analysis'. The Royal Society.

Yam, J.Y. and Chow, T.W. (2000) 'A weight initialization method for improving training speed in feedforward neural network', *Neurocomputing*, 30(1-4), pp. 219-232.

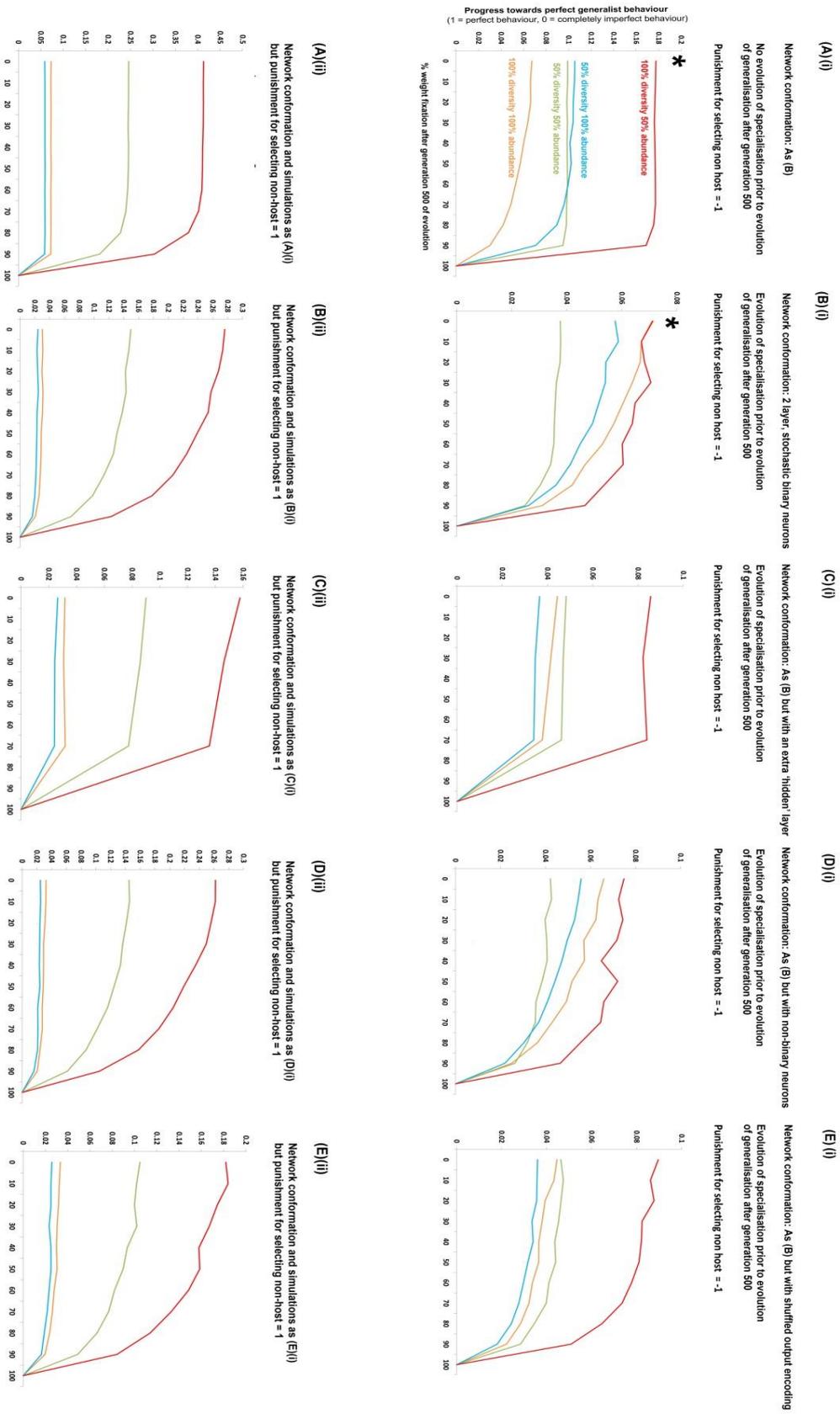
Yam, Y., Chow, T.W. and Leung, C.-T. (1997) 'A new method in determining initial weights of feedforward neural networks for training enhancement', *Neurocomputing*, 16(1), pp. 23-32.

ACCEPTED MANUSCRIPT

Figure 1



**Figure 1.** A summary of the model and simulations. The network model is adapted from (Tosh *et al.*, 2009). **Part A.** Shows one of 250 ‘projections’ that are processed by the neural network. Resources are placed onto the input surface and can appear at any of 8 distinct positions on the surface. This information is processed by stochastic binary neurons (in the first instance; neuron and other properties are modified in a sensitivity analysis) to output 3-element binary codes which determine which of the eight sections, and by extension resource, is selected by the network for ‘consumption’. Consumption has fitness consequences for the network and determines which networks reproduce, mutate, and pass to the next generation. Empty spaces, areas of the input surface containing only zeros, can also be selected. Resource-abundant environments produce 5-8 projections on the input surface of the network and less abundant environments produce 1-4 projections on the input surface. **Part B.** Asymmetric resources are designated ‘host’ resources and all others are non-host resources. Selection of networks for specialisation involves giving a high reward to only one asymmetric resource and selection for generalisation involves giving a high reward to all ten asymmetric resources. Low diversity environments contain all asymmetric resources and asymmetric, non-cohesive resources. High diversity environments contain all asymmetric resources and all of the remaining 30 resources. **Part C.** Networks evolve to specialise in an historical environment of high abundance and diversity. After 500 generations, a proportion of weights are fixed, the environment may become less diverse and/or abundant, and networks are selected to generalise. Reward for selecting a host resource remains the same at +5 but punishment/reward for selecting a non-host is set at either 1 or -1 depending on the simulation.



**Figure 2.** The main predictions of simulations. Networks are evolved to specialise on a small number of resources in an historical environment that is resource-diverse and resource-abundant. After 500 generations of evolution, suddenly environments may become less resource-abundant and diverse (as many contemporary, human-dominated, environments have become) and we now assume that it is selectively advantageous for networks to generalise. At this point different proportions of network ‘genes’ (weights) are fixed to represent the loss of genetic diversity that is thought to occur during prolonged specialisation in nature. Y-axes of plots measure how well networks proceed towards perfect generalist behaviour over a further 500 generations of evolution. X-axes measure the percentage of ‘genes’ fixed in the networks. (i) plots represent simulations in which networks are punished slightly within the genetic algorithm for selection of an ‘incorrect’ resource. (ii) plots represent simulations in which networks are rewarded slightly within the genetic algorithm for selecting an ‘incorrect’ resource. Plots B-E represent identical simulations with networks conformed differently, to determine how well key effects generalise. (A) plots represent simulations where networks did not initially evolve specialisation so comparison of (A) with B-E plots shows the impacts of specialisation in the historical environment on key effects within the contemporary environment. Generally, resource diversity facilitates evolution of generalisation in the contemporary environment but the opposite is true under some conditions when networks do not specialise in the historical environment. This effect of resource diversity in the contemporary environment is, therefore, specific to specialists, at least under some conditions. Networks are generally robust to loss of ‘genes’ and impacts of this factor with respect to treatment are static, but historical specialisation does lower robustness in the contemporary environment to loss of genetic diversity. At 100% weight fixation the networks are unable to evolve and this is reflected in the 0 value. The completely fixed networks could not make any progress toward generalism. \* Note that Y-axes use different scales to ease visualisation.