Environmental Harshness and Fitness Improving Innovations

Peter Andras[1]

[1]School of Computing and Mathematics, Keele University
p.andras@keele.ac.uk

Abstract

Fitness improving innovations occur in populations of organisms as genetic changes (mutations) that allow better fit with the environmental niche of the organisms. Similarly, fitness improving innovations may occur in the context of human communities as well in terms of socio-economic innovations (e.g. new ways of organizing the military, new products or services) that lead to more efficient use of available resources. Here we explore the link between such innovations and the harshness of the environment, where the populations live. Environmental harshness characterizes the availability of population growth supporting resources in the environment. Our analysis shows that if the harshness of the environment varies smoothly with the distance, the expected extent of fitness improving innovations and of the resource utilization efficiency of populations depends in a combined linear and harmonic manner on the harshness of the environment at the location of origin of the populations. We explore the implications of this result for particular cases of both biological and social environments (e.g. gene drives, business innovation).

Introduction

At any time there is usually a diverse community of organisms at any geographical location (Gould, 2002). These organisms typically belong to more then one, and often many, distinct populations of organisms. The same happen also in terms of humans belonging to culturally distinct populations and companies and organizations belonging to distinct sectors of the economy (Diamond, 1997; Arthur, 2009). Some of these populations are more successful than others. The successful populations spread widely and become dominant, in terms of the number of individuals, among the co-existing populations for a considerable period of time over some extensive geographical range (Andras, 2015; Diamond, 1997; Gould, 2002; Nielsen et al, 2017; Turchin, 2006).

The spreading of populations generally can be described as a reaction-diffusion process (Volpert and Petrovskii, 2009). Generally, individuals belonging a particular population migrate to geographically neighboring areas and establish the presence of their population there. The meaning of neighboring areas depends on the context, e.g. in the case of animal or plant populations it is a close distance sufficiently hospitable location, but in the context of globally mobile companies may mean almost any location around the Earth with sufficient support services present.

Populations of living organisms change by accumulating mutations in their DNA that lead to the emergence of features or behaviors at cellular or organismal scale, which give a selective advantage to the individuals harboring these mutations (Maynard Smith and Szathmary, 2000). This process leads to the emergence of new, genetically and reproductively separated, populations of organisms (Yom-Tov and Geffen, 2011). Similar processes happen at the level of socio-economic organization of human communities as well, leading to the emergence of new, culturally distinct, populations and new, technologically distinct, areas of economic activity filled by a corresponding new population of human organizations (Arthur, 2009; Fukuyama, 2011; Turchin, 2006). The actual process of how these changes accumulate to generate the innovations leading to the emergence of new populations is not yet fully known (Maynard Smith and Szathmary, 2000).

Examples of innovations that define new populations include the development of foot adaptations in animals (Holowka and Lieberman, 2018) or the development of drought resistance in plants (Kooyers, 2015). In the case of human populations an example is the development of settled agriculture (Diamond and Bellwood, 2003) or in the case of companies the development of new products or technologies (Arthur, 2009) such as provision of social media services or the development of diesel engines.

A number of environmental factors influence the spreading of populations (Andras, 2015; Barraclough, Vogler and Harvey, 1998; Fukuyama, 2011). Such factors include the presence of geographical barriers (e.g. mountains, sea, rivers), variation in the harshness of the environment (e.g. cold/warm, arid/wet areas, presence/absence of disease spreading vectors), the extent of competition among co-existing populations (e.g. increased number of populations in overlapping environmental niches may lead to more competition among populations), or the speed of growth of co-existing populations (e.g. plentiful resources may facilitate fast growth of all populations). These environmental factors modulate the spreading of populations by altering the diffusion process that drives the population spreading (Andras, 2015).

Previously it has been reported that the average extent of innovations emerging in populations increases with the harshness of the environment where the new population originates (Andras, 2015). In terms of interpretation it has been suggested that the likely reason is that in harsher environments fitness innovations of larger extent are needed to result in sufficient growth of the population to spread to other areas and in such environments larger innovations have amplified effect on the growth success of the populations. On the other hand, in

less harsh areas, even small innovations may provide sufficient advantage for a comparative growth benefit that may lead to the emergence of a successful population (Andras, 2015). However, the reported findings have also shown an unexplained wavy nature of the variation of the mean extent of innovation with the harshness of the environment superimposed on the above noted linear relationship.

Here we explore and explain the nature of this wavy relationship between the harshness of the environment and the average extent of fitness innovations. We also show that the resource utilization efficiency of successful populations also follows a similar, but phase sifted, wavy relationship with the harshness of the location of origin, similarly superimposed on a linear relationship. These results have implications for a number of research and technological applications area that relate to the emergence of innovations in populations, such as the use of gene drive technologies, emergence of antibiotic resistance, and policies for support of business innovation.

The rest of the paper is organized as follows. First we review briefly the relevant background. Next we describe the core equations and theoretical derivations about the distribution of fitness improving innovations and the impact of environmental factors on this. This is followed by a brief summary of the simulation environment that we used. Then we describe our results and the interpretation and discussion of these. Finally, the paper is closed by the conclusions section.

Background

The spreading of populations of organisms and culturally different human populations is usually modeled using reaction-diffusion equations (Volpert and Petrovskii, 2009). This has the following general form

$$\partial q(x,t) / \partial t = \nabla(D(q,x) \cdot \nabla q(x,t)) + F(q,x,t) \qquad (1)$$

where $q(x,t)$ is the size or quantity (i.e. normalized number of individuals) of a given population at location x and time t, $D(q,x)$ is the diffusivity of the environment at location x and for population size q, and $F(q,x,t)$ is the reaction terms, which expresses the local dynamics of the population given its size $q(x,t)$ at spatial position x and time t. This last term includes the impact of death and birth of individuals on the population size and also the impact of other factors, such as the competition with other populations, the availability of resources at the location and so on. The impact of the harshness of the environment on populations can be incorporated into this term within the reaction-diffusion model of population spreading. The diffusivity element of the model characterizes the ease / difficulty of spreading from one location to a neighboring location. This element would include for example the presence of natural barriers, such as mountains or rivers in the environment (Andras, 2015).

While the reaction-diffusion model of population spreading is very useful for conceptual and formal analysis of the spreading process it has its limitations. These stem mainly from the limits of formal analysis for cases with non-simple forms of the reaction term and also for cases with complex structure for the diffusivity element. Building models that aim to capture realistic conditions and constraints often lead to

models with such non-simple reaction terms and diffusivity elements (Andras, 2015; Cheng et al, 2014). In such cases agent based models can be used very effectively (Montenegro et al, 2016). In these models a discrete version of equation (1) is used in combination with the modeling of populations by a collection of agents that make a few algorithmic and possibly stochastic decisions (e.g. reproduction, movement, division, etc.). Such agent-based models of population spreading can also include the process of accumulation of innovations (e.g. simulating the impact of genetic mutations or of cultural-technical changes).

The emergence of genetic changes that lead to cellular or organismal level fitness improving innovations has been researched for many years (Maynard Smith and Szathmary, 2000). Although there are still many aspects of this process that are unclear, there is important progress in some respects. The best understood aspects of such processes relate to the emergence of antibacterial resistance in bacteria (Blair et al, 2015) and the evolution of viruses (Vijaykrishna, Mukerji, and Smith, 2015). Recent works also show how the replication of mammalian genes leads to the emergence of fitness improving genetic innovations (Carelli et al, 2016) and the emergence and fitness improving usefulness of mutations in plants have been also analyzed using observations of natural spreading of relatively recently introduced invasive plant species (Exposito-Alonso et al, 2018). In the context of emergence cultural-technical changes that lead to fitness improving innovations among companies or human populations there is also considerable related research (Arthur, 2009; Fukuyama, 2014). Genomic studies of human population spreading can indicate the origins of innovations and the directions of spreading, although do not convey information about how the innovations came about (Nielsen et al, 2017). Both biological evolution and simulation-based research show that environmental barriers, which provide relative isolative protection to evolution of species, increase the frequency of fitness improving innovations, while the lack of such barriers reduces this frequency (Mazancourt et al, 2008; Millien, 2006; Yom-Tov and Geffen, 2011; Andras, 2015).

Previous works on modeling the spreading of human populations and animal species have used agent-based models to analyze the location of origin and the spreading pattern of populations (Andras, 2015; Montenegro et al, 2016). Andras (2015) used agent-based simulation to show that the average resource utilization efficiency and average time persistence of successful populations reduces linearly with the harshness of the location of origin of these populations. Similarly the paper also shows that the average extent of fitness improving innovations of successful populations increases linearly with the harshness of the location of origin. However, in all these cases the paper shows that superimposed on the linear relationship there is also a periodic, wavy relationship with the harshness of the location of origin. This aspect of the relationship has not been fully explained in Andras (2015).

Fitness Improving Innovations and Environmental Factors

We assume that the populations exist in an environment characterized at each location by the environmental harshness,

h. This effectively sets the level of resources available at that location, higher *h* implying more scarcity of resources. Note that this definition of environmental harshness is very generic and it may incorporate effects of the physical environment, such as aridity, temperature or availability of shelter, and also effects of the live environment, such as predation risk.

Populations are characterized by their resource utilization efficiency, *r*, which determines to what extent they can use the resources available at a given location. The population growth rate depends on the resource utilization efficiency of the population, on the available resources and also on the general growth support of the available resources and the competition between different co-existing populations for the resources. A relatively general assumption is that the population growth rate, *ρ*, is proportional with the general growth support, *γ*, the extent of the competitiveness of the population, *η*, a gradually saturating function of the resource utilization efficiency of the population (e.g. *ln(1+r)*) and is inversely proportional with the harshness of the environment, *h*. The growth rate should also take account the death rate of the population, *θ*, which is assumed to be the same for all competing populations, but is altered locally additively proportionally with the general growth support, *γ*, and inversely proportionally with the harshness of the environment, *h*. The corresponding equation is:

$$\rho = \gamma \cdot \eta \cdot ln(1+r)) / h - \theta + \gamma / h \quad (2)$$

A fitness innovation is defined as a change in the resource utilization efficiency of a part of a pre-existing population, *Δr*, due to mutations to the genes of the organism or changes to the cultural-technical processes of the human population (Maynard Smith and Szathmary, 2000; Turchin, 2006). A fitness improving innovation is such a fitness innovation that improves the resource utilization efficiency of the population, i.e. when *Δr > 0*. For example, this may mean the emergence of enzymes that allow the animals to digest their food more efficiently, or the emergence of molecular changes that allow the build-up of larger and stronger muscle mass, or the emergence of cultural changes that allow better agricultural practices that make possible the feeding of a larger population without requiring extension of the available agricultural land area (Diamond and Bellwood, 2003; Fukuyama, 2014; Maynard Smith and Szathmary, 2000; Turchin, 2006).

The general growth support of the environment, *γ*, is a general environmental factor that modulates the effectiveness of the impact of the environmental harshness on the populations. For example, large scale climate effects may alter in such way the impact of environmental harshness, or the general accessibility of a technology for all human populations may change the effectiveness of the environmental harshness.

The competitiveness of a population depends on the relative resource utilization efficiency of the population in comparison with other co-existing populations. A conceptually simple approach to defining such competitiveness is to consider a function that saturates both for high and low values, such as the sigmoidal function of an expression that quantifies the relative strength of the resource utilization of the population. For example, we may define the extent of competitiveness as:

$$\eta = 1/(1+exp(-\alpha \cdot r / (\Sigma_k r_{pop-k})) \quad (3)$$

where *α* is an environmental factor that characterizes the steepness of the competition in the given environment, and the summation is applied over all resource utilization values of all co-existing populations. We note that a similar approach is used in the context of calculating the competitiveness of plant species (Kattenborn et al, 2017; Hodgson et al, 1999). In our approach the sigmoidal transformation reduces the competitiveness differences in the top and bottom extremes.

Finally, the spreading of the populations will also be influenced by the presence of natural barriers, such as mountain ranges, rivers, sea, which make difficult to cross the barrier (Andras, 2015; Barraclough, Vogler and Harvey, 1998). The presence of natural barriers can be incorporated into the diffusivity of the environment, *D(x)*, where *x* is the location. The other environmental factors noted above, *h*, *α*, *γ*, *θ*, and the resource utilization efficiency, *r*, get incorporated into the reaction term, *F(q,x,t)*, of the population diffusion equation (1), where *q* is the size of the population and *t* is the time.

We aim to determine the mean resource utilization efficiency of populations originating from locations with environmental harshness *h* and also the mean amount of fitness improving innovation for these populations. Given that in natural environments the harshness varies relatively smoothly with the location in the environment, we adopt this assumption. We also assume that the chance of generating any resource utilization innovation is *β*, where *0 < β < 1*, independently of the location and the amount of innovation. This is a simplifying assumption.

First, let us consider the count of populations with resource utilization efficiency *r* at locations with environmental harshness *h* at time *t*, *g_e(r,h,t)*. At any time the difference in the number of populations moving in and those that move out of the area, all with resource utilization efficiency *r*, is the population move induced change of the number of such populations, in areas with harshness *h*. Some of these populations may die out and some new population with resource utilization efficiency *r* may emerge following a mutation in populations with different resource utilization efficiency. The number of newly emerging populations with resource utilization efficiency *r*, depends on the past number of the source populations summed up over all possible *r* values. In terms of equations we can write

$$\beta \cdot \Sigma_{\Delta r} \, g_e(r - \Delta r, h, t - \Delta t) = \beta \cdot \Sigma_{\Delta r} (g_e(r - \Delta r, h, t - \Delta t) - \quad (4)$$
$$g_e(r - \Delta r, h, t) \,) + \beta \cdot \Sigma_{\Delta r} (g_e(r - \Delta r, h, t)$$

Turning the above equation in continuous form as *Δr* and *Δt* become very small, we get the differential equation formulation for the number of emerging new populations with resource utilization efficiency *r*:

$$- \beta \cdot \int_{\Delta R} \partial g_e(r - \Delta r, h, t) / \partial t \, d\Delta r + \quad (5)$$
$$\beta \cdot \int_{\Delta R} g_e(r - \Delta r, h, t) \, d\Delta r$$

where *ΔR* is the value range of *Δr*-s – we note that this formulation allows fractional counts of populations as well. Summing up the noted parts, we get the equation for the number of populations with resource utilization efficiency *r*, including all changes:

302

$$\partial g_e(r,h,t)/\partial t = -a \cdot \partial^2 g_e(r,h,t)/\partial h^2 - \qquad (6)$$
$$\delta \cdot (h/r) \cdot g_e(r,h,t) -$$
$$\beta \cdot \int_{\Delta R} \partial g_e(r-\Delta r,h,t)/\partial t \, d\Delta r +$$
$$\beta \cdot \int_{\Delta R} g_e(r-\Delta r,h,t) \, d\Delta r$$

where $\delta \cdot (h/r)$ is the extinction multiplier, proportional with the harshness of the environment and inversely proportional with the resource utilization efficiency of the populations. Considering $G_e(r,h,t) = \int_{\Delta R} g_e(r-\Delta r,h,t) \, d\Delta r$, we can write

$$\partial^2 G_e(r,h,t)/\partial t \partial \Delta r = -a \cdot \partial^3 G_e(r,h,t)/\partial h^2 \partial \Delta r - \qquad (7)$$
$$\delta \cdot (h/r) \cdot \partial G_e(r,h,t)/\partial \Delta r -$$
$$\beta \cdot \partial G_e(r,h,t)/\partial t +$$
$$\beta \cdot G_e(r,h,t)$$

We define the following function:

$$u_{r,h,t}(r',h',t') = \beta \cdot \partial G_e(r,h,t)/\partial t + \qquad (8)$$
$$\beta \cdot G_e(r,h,t)$$

which is the number of populations with resource utilization efficiency r that newly emerged at a location with harshness h at time t and then spread all around. This definition implies that

$$u_{r,h,t}(h',t') = 0 \qquad (9)$$

for all $h' \neq h$ and $t' \leq t$ and also for $h' = h$ and $t' < t$. Using the above reasoning, we can write the equation

$$\partial u_{r,h,t}(h',t')/\partial t' = -a \cdot \partial^2 u_{r,h,t}(h',t')/\partial h^2 - \qquad (10)$$
$$\delta \cdot (h'/r) \cdot u_{r,h,t}(h,t)$$

given that we are not interested in this case of innovation induced emergence of new populations.

These populations persist for sufficient time, τ, if

$$\int_H u_{r,h,t}(h',t+\tau) dh' > 0 \qquad (11)$$

where H is the full range of environmental harshness values. Considering the populations that persist for long time, some of these will become dominant and successful in the sense they will account for a considerable part of the overall population in the whole area. The practical meaning of sufficiently long persistence, sufficient size and whole area will depend on the context (e.g. for example this may mean a hundred years, 5% of the total population over an area of the size of an average European country, in the case of human populations). Considering that the populations will spread over locations with all h values, over a long time period, and that there will be always other populations with both higher and lower r values around, the ratio of becoming successful should not change with h, r or t, and this fixed ratio is denoted as κ. We note that if the r values can go only up to a certain limit r^* then there will be an effect on the success ratio that depends on the value of r when this is close to r^*. So, the proportion of successful populations with resource utilization efficiency r originating from an area with harshness h at time t is:

$$\kappa \cdot \int_{T(\tau)} \int_H u_{r,h,t}(h',t+\tau') dh' d\tau' / u_{r,h,t}(h,t) \qquad (12)$$

where $T(\tau) = [\tau, T^* - \tau]$ is the considered time domain, T^* being the maximum time considered for the calculations.

Thus the chance for a population with resource utilization efficiency r originating from an area with harshness h to become a successful population is:

$$q(r,h) = \kappa \cdot \int_{T\#} \int_{T(\tau'')} \int_H (u_{r,h,t}(h', \tau'' + \tau') dh' d\tau' / \qquad (13)$$
$$u_{r,h,t}(h,\tau'')) d\tau''$$

where $T\# = (0,T^*)$ is the full considered time domain and $T(\tau'') = [\tau'', T^* - \tau'']$.

Then we can calculate the chance of having a population as a result of a Δr fitness innovation, while it is originating from an area with harshness h at time t, by summing up (integrating) over all r values. This is:

$$v_{time}(\Delta r,h,t) = \int_R q(r+\Delta r,h) \cdot g_e(r,h,t) dr / \qquad (14)$$
$$\int_R g_e(r,h,t) dr$$

where R is the full range of r values. Integrating this over time, gives us the overall chance of having successful populations originating at a location with harshness h and as a result of a fitness innovation Δr:

$$v(\Delta r,h) = \int_{T\#} (\int_R q(r+\Delta r,h) \cdot g_e(r,h,\tau') dr / \qquad (15)$$
$$G_e(r,h,\tau')) d\tau'$$

where $T\#$ is defined as before.

Using these equations we can calculate the mean resource utilization efficiency and the mean fitness innovation for populations originating at locations with harshness h as follows:

$$r_m(h) = \int_R r \cdot q(r,h) \, dr \qquad (16)$$

$$\Delta r_m(h) = \int_{\Delta R} \Delta r \cdot v(\Delta r,h) \, d\Delta r \qquad (17)$$

where ΔR is the full range of Δr values and R is defined as above.

Given the form of equation (7) that defines $G_e(r,h,t)$ we conclude that if the $\partial^2/\partial h^2$ component of equation (7) is non-zero then $G_e(r,h,t)$ will have a dampened harmonic (i.e. dampened sinusoidal) component in it. Considering the dependence of $q(r,h)$ and $v(\Delta r,h)$ on $G_e(r,h,t)$, the mean values of r and Δr will contain similarly a dampened harmonic component in addition to a linear component. The reason behind this is that equations similar to equation (7) can be written for $r_m(h)$ and $\Delta r_m(h)$ as well. This is valid as long as the environmental harshness, h, varies smoothly with the distance between the locations. If this is not the case, this result no longer holds. For example, if h would vary randomly, then the $\partial^2/\partial h^2$ component of equation (7) would zero out and the expectation would be to have no dampened sinusoidal component in these mean values as a function of the environmental harshness. However, such random variation is very unlikely in natural environments.

The above reasoning explains the finding reported in Andras (2015), which found this unexpected harmonic variation in the mean values of r and Δr. Below we explore experimentally further this nature of the dependence of r and Δr on the harshness, h.

Simulation Environment

To explore the variation of the mean resource utilization efficiency, $r_m(h)$, and the mean fitness innovation, $\Delta r_m(h)$ with the harshness of the locations of origin, h, we used a simulated environment where populations spread and evolve by adopting fitness innovations. The simulation environment that we used has been described in Andras (2015), here we provide a brief description of the key features.

The world is implemented as a 2-dimensional grid of *100 × 60* spatial locations, without gluing of the opposing edges together (i.e. the world is not toroidal). The world contains a number of natural barriers (i.e. simulated mountain ridges), which are located randomly and have randomly set crossing difficulty (i.e. height). In all simulations we used 20 natural barriers. The world has a number of harshness hotspots and the harshness of the world locations vary smoothly with the distance from the centre of the hotspot. In some simulations we used 10 harshness hotspots, while in other simulations we used only one harshness hotspot. We adopted two options for the harshness variation with distance. In the first case the harshness depends on the inverse of the Euclidean distance, while in the second case it depends on the inverse of the Manhattan distance:

$$h(x) = 1 / (\varepsilon + ((x_1 - x_{01})^2 + (x_2 - x_{02})^2)^{1/2}) \qquad (18)$$

$$h(x) = 1 / (\varepsilon + (|x_1 - x_{01}| + |x_2 - x_{02}|)) \qquad (19)$$

where x_0 is the centre of the harshness hotspot and $\varepsilon > 0$.

Initially, around 1.5% of the spatial locations host a population. Each population is characterized by their resource utilization efficiency, r, that is set by calculating the value associated with a population specific sequence of bits (0 or 1) b of length L as follows:

$$r = \Sigma_{k=1,L1} \Sigma_{i=1,L2}\, b_{L1\cdot(k-1)+i} \cdot 2^j \qquad (20)$$

where $L1 \cdot L2 = L$ and in our case $L1 = L2 = 10$. Populations may give rise to a new population by experiencing a mutation in their bit string. The likelihood of such mutations is $\beta = 0.00008$ in our simulations. The fitness innovation corresponding to the mutation is the difference of the r values after and prior to the mutation. Following such mutations a part of the populations adopts the new resource utilization efficiency coding bit string. The population is split proportionally with the two r values.

The populations at each location grow and spread. The spreading follows a discretized version of the reaction-diffusion equation (1):

$$Q(x,t+1) = \qquad (21)$$
$$= \Sigma(\mu,\upsilon) \in \Psi(\varphi x + (\tau,\upsilon),-\mu,-\upsilon,t \cdot Q(x+(\mu,\upsilon),t) - \varphi x,\mu,\upsilon,t \cdot$$
$$Q(x,t))$$
$$+ \rho(x) \cdot Q(x,t)$$

where $Q(x,t)$ is the size of the population at spatial position x at time t, $\Psi = \{(-1,0),(1,0),(0,-1),(0,1)\}$, and $\varphi_{x,\tau,\upsilon,t}$ are stochastic diffusivity parameters, and $\rho(x)$ is the growth rate of the population at the spatial position x. The stochastic diffusivity
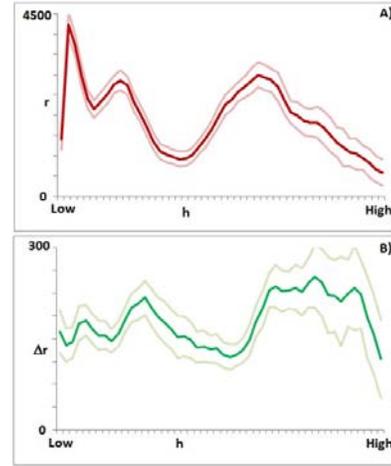

Figure 1. The relationship between environmental harshness, h, and mean resource utilization efficiency, r (A), and mean fitness innovation, Δr (B), in the case of multiple harshness hotspots in the environment (mean value dark line, 95% confidence intervals, light lines).

parameters are set such that

$$\varphi_{x+(\tau,\upsilon),-\tau,-\upsilon,t} = \omega \quad if$$
Q(x,t) is sufficiently large, given the height value at the neighboring location $x+(\mu,\upsilon)$
$$otherwise\ \varphi_{x+(\tau,\upsilon),-\tau,-\upsilon,t} = 0 \qquad (22)$$

where ω is a randomly set value, in the case of the reported simulations $\omega \in (0,0.4)$. If $\varphi_{x+(\tau,\upsilon),-\tau,-\upsilon,t} = 0$ the populations encounters a barrier that cannot be crossed. The growth rate of the population is given by equation (2), with $\theta = 0.005$. The implementation of equation is slightly modified compared to equation (3) by adding constant terms into it:

$$\eta = 0.1/(1 + exp(\alpha / 2 - \alpha \cdot r / (\Sigma_k r_{pop-k})) \qquad (23)$$

The value of α was set to 10 in some simulations and it was varied in the range of 6 to 12 in other simulations. The value of γ in equation (2) was set to 1.4 in some simulations and then it was varied in the range of 0.8 to 1.7 in other simulations.

Each simulation was run for 30,000 time turns. Typically, we started initially with around 90 populations. The populations spread and evolved. The competition for resources drove to extinction some of the populations. The newly emerged populations that innovated successfully spread and became dominant. In the simulations we considered a population dominant if it contained over 0.5% of the total population at any one time turn in the simulated world. We did not set a particular persistence limit, but naturally, for any new population it took some time to become dominant according to the above definition (typically more than 20 time turns).

For each population we considered the harshness of the location of their origin and we calculated for each harshness value, h, the mean value of the resource utilization efficiency, $r_m(h)$, and of the fitness innovation, $\Delta r_m(h)$, considering all populations originating from a location with harshness value h. For each condition that we considered we did 10 different simulations. Each simulation generated in the range of 2,500 –
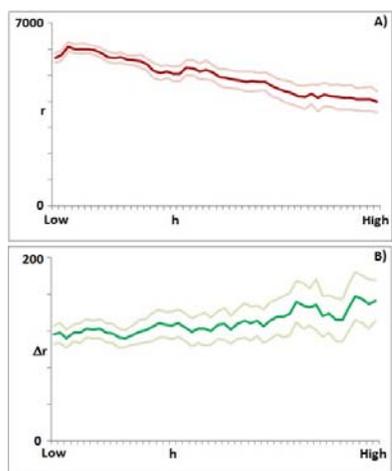
Figure 2. The relationship between environmental harshness, h, and mean resource utilization efficiency, r (A), and mean fitness innovation, Δr (B), in the case of random harshness in the environment (mean value dark line, 95% confidence intervals, light lines).

3,000 successful populations that were taken into consideration. Only the average values, where there was at least 10 different populations originating from locations with a given harshness h, were considered in the analysis and in almost all cases the averages were calculated from more than 20 instances of appropriate populations.

The computer program implementing the simulations described above is available on request from the author.

Results and Discussion

First, we considered the default case with multiple harshness hotspots in the environment and Euclidean distance based determination of the harshness of locations. The results are shown in Figure 1. As expected both relationship displays the harmonic component added onto a linear relationship with a negative slope in the case of the resource utilization efficiency and a positive slope in the case of the fitness innovation. In other words, the mean resource utilization efficiency, r_m, drops with the increase of the harshness of the location of origin and the mean fitness innovation, Δr_m, grows with the increase of the harshness of the location of origin innovation.

Intuitively this means that successful population originating from low harshness environment gain their advantage from being highly efficient in the utilization of resources. On the other hand this result also means that successful populations originating from harsh environments stand out by making high fitness innovations. The harmonic element of the relationships is due to the smooth variation of the environmental harshness with distances between locations, as we explained it earlier. This harmonic element complicates the above outlines picture of the relationships with environmental harshness. This component implies that there will be intermediate ranges of environmental harshness where either the high resource
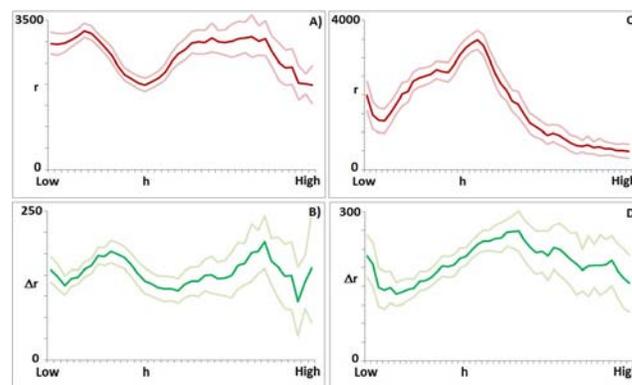

Figure 3. The relationship between environmental harshness, h, and mean resource utilization efficiency, r (A), and mean fitness innovation, Δr (B), in the case of slow variation of harshness with the Euclidean distance with one harshness hotspot; and the mean resource utilization efficiency, r (C), and mean fitness innovation, Δr (D), in the case of fast variation of harshness with the Euclidean distance with one harshness hotspot (mean value dark line, 95% confidence intervals, light lines).

utilization efficiency or high fitness innovation will be more likely than elsewhere for the locally originating new populations.

Next, we considered the un-natural case of random variation environmental harshness. The results are shown in Figure 2. While there is still some very small level of harmonic component in the relationships between h and r and Δr, the relationships are mainly linear as we expected, based on our analysis. The case of random variation of environmental harshness shows the core relationship between harshness and resource utilization efficiency and fitness innovation.

We explored next the case of having a single harshness hotspot with different variation of the harshness with the distance from the centre of the hotspot. We considered equations (18) and (19) for the definition of harshness and we also added a multiplier in the front of them to make the harshness variation fast (high multiplier – 20) or slow (low multiplier – 10). The results are shown in Figures 3 and 4. In these scenarios again we see that the harmonic component is clearly present in addition to the linear variation of r and Δr with h. The data in the figures also shows that the period length of the harmonic component extends as the speed of variation of h increases with the distance, in the case of both distance choices.

This means that faster variation of harshness with the distance comes with slower harmonic variation of r_m and Δr_m with the variation of h. This implies that if harshness changes rapidly with distance, there will be only one or at most a few harshness ranges where high resource utilization efficiency or high fitness innovation may emerge. On the other side, if the harshness changes slowly with the distance, there will be multiple, possibly many, harshness ranges where high resource
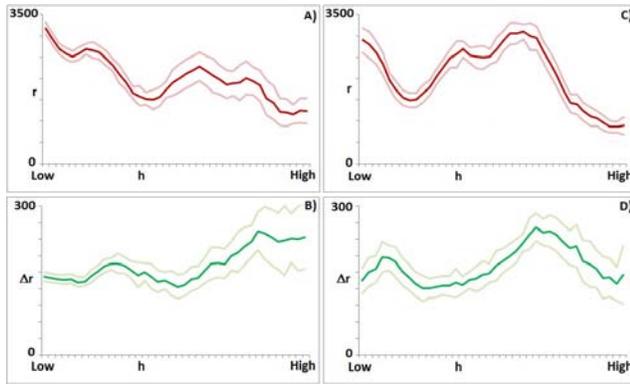
305

Figure 4. The relationship between environmental harshness, *h*, and mean resource utilization efficiency, *r* (A), and mean fitness innovation, *Δr* (B), in the case of slow variation of harshness with the Manhattan distance with one harshness hotspot; and mean resource utilization efficiency, *r* (C), and mean fitness innovation, *Δr* (D), in the case of fast variation of harshness with the Manhattan distance with one harshness hotspot (mean value dark line, 95% confidence intervals, light lines).

utilization efficiency and high fitness innovations may emerge in new populations. A practical prediction derived from this is the expectation to have one or a few ranges of height values in relatively steeply rising mountainous areas, where the most successful new species may arise, while in areas with slow change of height the expectation is to have many ranges of height where highly successful new species may emerge.

We also considered the impact of variation of the competition strength, *α*, and of the general population growth speed parameter, *γ*. The results are shown in Figures 5 and 6. In these respects, we found that stronger competition and faster growth of the populations reduces the period length of the harmonic component. The implication of this is that in environments, which are generally more competitive or support faster growth there will be multiple or many ranges of harshness values where the most innovative and most efficient new populations are likely to emerge. On the other side in environments that stunt the population growth or suppress competition between populations there will be only one or a few ranges of harshness values where the most innovative and most efficient populations may emerge.

We note that in all smooth harshness variation cases the variation of the mean resource utilization efficiency and mean fitness innovation with the harshness follow similar, but phase-shifted harmonic relationships. The conceptual explanation of this phase shift, is the difference between equations (13) and (15), which define *q(r,h)* and *v(Δr,h)*, which are the likelihoods of having a newly emerging population with resource utilization efficiency *r* or fitness innovation *Δr* emerging at a location with harshness *h* (i.e. *v(Δr,h)* is defined using an integral of *q(r,h)*).

Our theoretical analysis provides the explanation for the observed combined harmonic and linear variation of the resource utilization efficiency mean value and fitness innovation mean value with the harshness of the environment at the location of the origin. The simulation experiments confirm the expectations based on the theoretical explanation
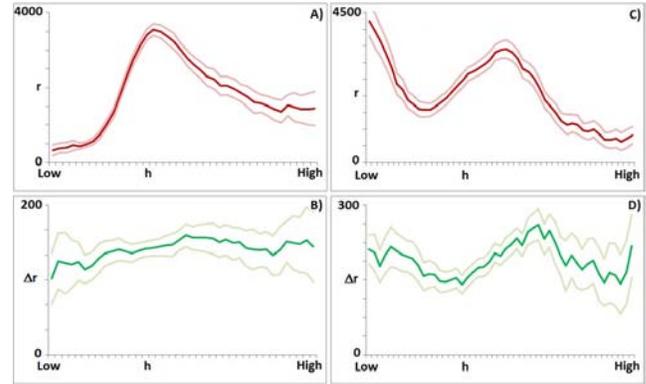

Figure 5. The relationship between environmental harshness, *h*, and mean resource utilization efficiency, *r* (A, C), and mean fitness innovation, *Δr* (B, D), in the case of low (A, B) and high (C, D) level competition (*α*) between populations, using the Euclidean distance and multiple harshness hotspot (mean value dark line, 95% confidence intervals, light lines).

and also offer a heuristic estimation of the impact of the key environmental parameters (i.e. speed of variation of harshness with distance, competition strength, population growth speed) on the nature of the relationship between the environmental harshness and the and the mean resource utilization efficiency and mean fitness innovation (i.e. which change in the parameters.

Our result is interesting because it implies that due to the harmonic component of the studies relationships there will be areas with particular ranges of environmental harshness where the population with the largest fitness innovations and also with the most efficient resource utilization are much more likely to emerge than in other areas with relatively similar features. In other words, the likelihood of emergence of such new species in natural environments does not vary linearly with the harshness of the environment, but follows the combined linear and harmonic relationship.

In a natural context, this implies that for example, genetic mutations that can prevent the effectiveness of a gene drive (Unckless, Clark and Messer, 2017) intervention against mosquitoes, may emerge much more likely in environment locations with certain levels of harshness, and are much less likely to emerge elsewhere. Being able to determine where such locations are likely to be, would help to maintain the effectiveness of the gene drive intervention. Similarly, this kind of analysis may help identify locations where antimicrobial resistance of bacteria (Blair et al, 2015) is more likely to emerge and possibly help to alter the environmental antimicrobial resistance of bacteria (Blair et al, 2015) is more likely to emerge and possibly help to alter the environmental features such that the antimicrobial resistance development happens in areas where it can be identified early and fought effectively.

In a socio-economic setting, this has implications for example for the likelihood of emergence of firms with disruptive innovation and those with high efficiency. Our results imply that in the context of a steeply varying regulatory and economic environment, there will be only a few ranges of regulatory and economic settings that will be conducive for the emergence of highly innovative or efficient firms. Similarly,
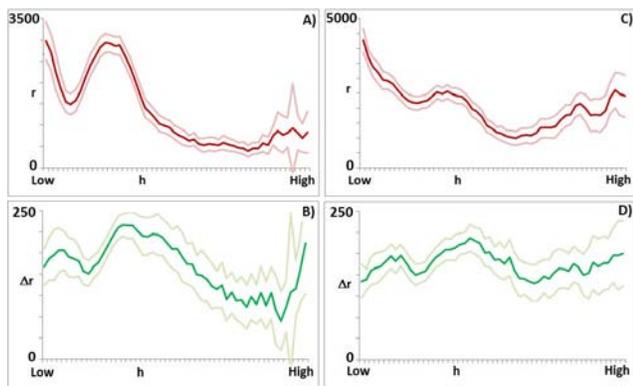
Figure 6. The relationship between environmental harshness, h, and mean resource utilization efficiency, r (A, C), and mean fitness innovation, Δr (B, D), in the case of low (A, B) and high (C, D) general growth support (γ) for populations, using the Euclidean distance and multiple harshness hotspot (mean value dark line, 95% confidence intervals, light lines).

less competition and more barriers of growth will reduce the likely ranges of regulatory and economic settings, in which highly innovative or efficient firms may emerge. On the other side, more competition, fewer growth barriers, and more uniform regulatory – economic landscape favor the presence of multiple ranges of settings in which innovative or efficient firms may emerge. Our results also imply that most likely the preferred regulatory – economic settings will not be the same for the emergence of the most innovative and most highly efficient firms.

Conclusions

In this paper we present a formal analysis of the relationship between environmental harshness and the mean resource utilization efficiency and mean fitness innovation of populations originating from locations with a given environmental harshness. We show that these relationships contain a linear and a harmonic component if the environmental harshness changes smoothly with the distance between locations. We explore experimentally the relationships considering a number of simulation environments with different parameters. The experimental analysis confirms the theoretical expectations. The experimental analysis also shows the dependence of the nature of these relationships on the environmental parameters. The experimental data also highlights that there is a phase shift between the harmonic components of the two relationships.

The experimental results show that faster variation of environmental harshness with distance extends the period length of the harmonic component of the relationships. The data also shows that more general growth support for the populations and more competition between the populations makes the period of the harmonic component of the relationships shorter.

The results are interpreted in both biological and social context. Our analysis may help to identify ways of improving the effectiveness of fight against the emergence of antimicrobial resistance in bacteria or of the application of gene drives. The interpretation in socio-economic context points to

the impact of the regulatory – economic landscape on the likelihood of emergence of firms with disruptive innovations and of highly efficient firms.

References

Andras, P. (2015). Environmental Factors and the Emergence of Cultural – Technical Innovations. *Proceedings of the European Conference on Artificial Life 2015*, pp.130-137.

Arthur, W.B. (2009). The Nature of Technology. Penguin Books, London.

Barraclough, T.G., Vogler, A.P. and Harvey, P.H. (1998). Revealing factors that promote speciation. *Philosophical Transactions of the Royal Society of London B*, 353:241-249.

Blair, J.M.A., Webber, M.A., Baylay, A.J., Ogbolu, D.O. and Piddock, L.J.V. (2015). Molecular mechanisms of antibiotic resistance. *Nature Reviews Microbiology*, 13:42-51.

Carelli, F.N., Hayakawa, T., Go, Y., Imai, H., Warnefors, M and Kaessmann, H. (2016). The life history of retrocopies illuminates the evolution of new mammalian genes. *Genome Research*, 26:301-314.

Cheng, H., Yao, N., Huang, Z.-G., Park, J., Do, Y. and Lai, Y.-C. (2014). Mesoscopic interactions and species coexistence in evolutionary game dynamics of cyclic competitions. *Scientific Reports*, 4:7486.

Diamond, J. (1997). *Guns, Germs, and Steel: The Fates of Human Societies*. Norton, New York.

Diamond, J. and Bellwood, P. (2003). Farmers and their languages: The first expansions. *Science*, 300:597-603.

Exposito-Alonso M., et al. (2018). The rate and potential relevance of new mutations in a colonizing plant lineage. *PLoS Genetics*, 14:e1007155.

Fukuyama, F. (2011). *The Origins of Political Order*. Profile Books, London.

Gould, S.J. (2002). *The Structure of Evolutionary Theory*. Harvard University Press.

Hodgson, J.G., Wilson, P.J., Hunt, R., Grime, J.P. and Thompson, K. (1999). Allocating C-S-R plant functional types: a soft approach to a hard problem. *OIKOS*, 85:282-294.

Holowka, N.B. and Lieberman, D.E. (2018). Rethinking the evolution of the human foot: insights from experimental research. *Journal of Experimental Biology*, 221:jeb174425.

Kattenborn, T. et al. (2017). Linking plant strategies and plant traits derived by radiative transfer modeling. *Journal of Vegetation Science*, 28: 717-727.

Kooyers, N.J. (2015). The evolution of drought escape and avoidance in natural herbaceous populations. *Plant Science*, 234:155-162.

Maynard Smith, J. and Szathmary, E. (2000). *The Origins of Life: From the Birth of Life to the Origin of Language.* Oxford University Press.

De Mazancourt, C., Johnson, E. and Barraclough, T.G. (2008). Biodiversity inhibits species' evolutionary response to changing environments. *Ecology Letter*, 11:380-388.

Millien, V. (2006). Morphological evolution is accelerated among island mammals. *PLoS Biology*, 4: e321.

Montenegro, A., Callaghan, R.T. and Fitzpatrick, S.M. (2016) Using seafaring simulations and shortest-hop trajectories to model the prehistoric colonization of Remote Oceania. *PNAS*, 113:12658-12690.

Nielsen, R., Akey, J.M., Jakobsson, M., Pritchard, J.K., Tishkoff, S. and Willerslev, E. (2017) Tracing the peopling of the world through genomics. *Nature*, 541:302-310.

Turchin, P. (2006). *War and Peace and War. The Rise and Fall of Empires*, Plume, New York.

Unckless, R.L., Clark, A.G. and Messer, P.W. (2017). Evolution of resistance against CRISP/Cas9 gene drive. *Genetics*, 205:827-841.

Vijaykrishna D., Mukerji R. and Smith G.J.D. (2015) RNA Virus Reassortment: An Evolutionary Mechanism for Host Jumps and Immune Evasion. *PLoS Pathogens*, 11:e1004902.

Volpert and V, Petrovskii, S (2009). Reaction-diffusion waves in biology. *Physics of Life Reviews*, 6:267-310.

Yom-Tov, Y., and Geffen, E. (2011). Recent spatial and temporal changes in body size of terrestrial vertebrates: probable causes and pitfalls. *Biological Reviews*, 86: 531-541.