



## EvoEvo Deliverable 3.2

Paulien Hogeweg, Guillaume Beslon

► **To cite this version:**

Paulien Hogeweg, Guillaume Beslon. EvoEvo Deliverable 3.2: Evolution of robustness: mechanisms and consequences. [Research Report] INRIA Grenoble - Rhône-Alpes. 2016. <hal-01577155>

**HAL Id: hal-01577155**

**<https://hal.archives-ouvertes.fr/hal-01577155>**

Submitted on 24 Aug 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



***EvoEvo Deliverable 3.2***

***Evolution of robustness; Mechanisms and consequences***

Due date: M28  
 Person in charge: Paulien Hogeweg  
 Partner in charge: UU  
 Workpackage: WP3 (In silico experimental study of EvoEvo)  
 Deliverable description: Evolution of robustness; Mechanisms and consequences: A report describing how robustness is indirectly selected in the model and its consequences on evolution.

Revisions:

Revision no.	Revision description	Date	Person in charge
1.0	First version	03/09/16	P. Hogeweg (UU)
1.1	Corrections and validation	07/10/16	G. Beslon (INRIA)
1.2	Corrections in the bibliography	26/10/16	G. Beslon (INRIA)



## Table of Contents

<b>1. INTRODUCTION</b>	<b>3</b>
<b>2. SIMPLE EVOLUTIONARY SYSTEMS: FIXED FITNESS LANDSCAPE</b>	<b>3</b>
<b>2.1. MUTATIONAL ROBUSTNESS: BASIC CONCEPTS AND MEASUREMENTS</b>	<b>4</b>
2.1.1. ROBUSTNESS AS FLAT VS. STEEP PEAKS IN THE FITNESS LANDSCAPE.	4
2.1.2. ROBUSTNESS AS FRACTION OF NEUTRAL MUTANTS ( $\lambda$ ) IN THE CLOSE NEIGHBORHOOD.	4
2.1.3. ROBUSTNESS AS THE FITNESS PROFILE OF MUTANTS.	5
<b>2.2. EVOLUTION TOWARD (WHICH TYPE OF) ROBUSTNESS?</b>	<b>5</b>
<b>2.3. EVOLUTIONARY CONSEQUENCES OF EVOLVED MUTATIONAL ROBUSTNESS: EVolvABILITY!</b>	<b>6</b>
<b>3. MUTATIONAL ROBUSTNESS IN GENERALIZED EVOLUTIONARY SYSTEMS: DIFFERENTIAL EFFECT OF MUTATIONAL OPERATORS</b>	<b>6</b>
<b>3.1. INTRODUCTION</b>	<b>6</b>
<b>3.2. EVOLUTIONARY STRATEGIES TO COPE WITH HIGH MUTATION RATES: SHAPING THE MUTATIONAL NEIGHBORHOOD</b>	<b>7</b>
3.2.1. EVOLUTION GENOME STRUCTURE IN MUTATOR STRAINS	7
3.2.2. FUNCTIONAL DIFFERENTIATION OF MUTATIONAL NEIGHBORHOOD	9
3.2.3. EVolvable MUTATIONS	9
<b>4. EVOLUTIONARY STRATEGIES TO COPE WITH VARIABLE ENVIRONMENTS</b>	<b>10</b>
<b>4.1. INTRODUCTION</b>	<b>10</b>
<b>4.2. EXTERNALLY IMPOSED ENVIRONMENTAL VARIABILITY</b>	<b>11</b>
4.2.1. REGULATION AND/OR EVolvABILITY	11
4.2.2. POPULATION VARIABILITY AND REGULATION IN PERIODICAL REFRESHED ENVIRONMENTS	12
<b>4.3. POPULATION INDUCED SPATIAL-TEMPORAL VARIATION</b>	<b>13</b>
<b>5. EVOLUTIONARY ROBUSTNESS</b>	<b>14</b>
<b>5.1. INTRODUCTION</b>	<b>14</b>
<b>5.2. SPATIAL SELF-ORGANIZATION AND MULTILEVEL SELECTION</b>	<b>14</b>
5.2.1. MUTATION RATE AND THE EMERGENCE AND PERSISTENCE OF DIVERSE ECOSYSTEMS	15
5.2.2. MULTILEVEL SELECTION AND THE EVOLUTION OF DNA IN THE RNA WORLD	15
5.2.3. SPATIAL SELF-ORGANIZATION AND 'OPEN ENDED' EVOLUTION IN STRINGMOL	16
<b>6. CONCLUSIONS</b>	<b>16</b>
<b>7. REFERENCES</b>	<b>18</b>



## 1. Introduction

---

A certain degree of robustness is essential for persistence of any (living) system. Robustness is however far from a uniquely defined concept or property. Robustness relates to many different levels and many different timescales, as well as relative to many different disturbances. Most relevant in biological context and the EvoEvo project<sup>1</sup> are robustness at the individual level, the population level, the lineage level and the ecosystem level. Moreover robustness gets a special flavour in evolving system, as these, per definition should be changeable and changing. Indeed studying mechanisms and consequences of robustness in computational experiments cannot be seen without considering variability and evolvability as well. In fact unravelling the intricate relation to these three central concepts of the EvoEvo project is the important outcome of our experiments and discussions. Therefore, although the focus of this report is robustness, we will stress its relations to evolvability and variability.

In this report the focus will be on robustness relative to mutations. We will first discuss different concepts and measurements of robustness developed with simple evolving systems in mind, i.e. systems for which an explicit fitness criterion is defined, which evolve in a constant environment, and are only subjected to point mutations. We analyse when and how robustness does and does not evolve as well as the evolutionary consequences of evolved/evolving robustness at different levels. We next discuss the challenges of observing and analysing robustness in complex evolving systems, and the novel insights obtained on mutational robustness at different levels: individuals, populations and ecosystems. Finally we introduce a novel concept of robustness: evolutionary robustness, which operates on a longer time scales than the classical concepts of mutational robustness.

## 2. Simple Evolutionary systems: fixed fitness landscape

---

Ever since Sewall Wright (1932) introduced it, the metaphor fitness landscape has played a central role in evolutionary thinking. A fitness landscape maps the multidimensional genotype space unto a fitness value. Evolutionary dynamics is then visualized as a trajectory in this landscape. This representation of evolution presupposes a fixed, predefined fitness measure, and a fixed genotype space in which mutational change can be represented as (small) steps, i.e. the mutational neighborhood equals spatial neighborhood in the fitness landscapes. These requirements render it a very restricted view of evolution, as it does not allow genome size evolution, or other mutations than point mutations.

It is useful to split the genotype to fitness mapping by recognizing as intermediate a genotype to phenotype landscape. Unlike fitness, phenotype is not a scalar, and the genotype space is mapped to the similarity of some structural feature due to genotypic change through mutation. Phenotypic features often considered are for example RNA or Protein structure, gene expression, or

---

<sup>1</sup> The applicability and relevance for living technologies of the concepts discussed in the main part of this report will be discussed in the conclusion.

morphology. Clearly different levels of phenotype can be considered as different intermediates. The genotype to phenotype mapping may or may not be considered to be dependent on the environment, and the same is true for the phenotype to fitness mapping. The genotype to phenotype mapping is an evolved property.

In this section we will discuss different notions of robustness within the context of the landscape metaphor, i.e. for evolutionary systems operating within these constraints. Although this is a very restricted model of evolution (as discussed in D3.1) it gives a useful baseline to unravel different concepts of robustness.

In the next section we will examine to what extent these notions are applicable and useful in more general evolutionary systems, and the challenges to study robustness in the general systems as well as the mechanisms and consequences of the robustness which evolves to cope with high mutation rates and environmental variability.

## **2.1. Mutational Robustness: Basic Concepts and Measurements**

Mutational robustness refers to the (the degree of) preservation of fitness under mutation. Within the fitness landscape metaphor robustness is expressed as the (local) shape of the fitness landscape, and the positioning of the population on the landscape. This, however, can be expressed in different ways.

### **2.1.1. Robustness as flat vs. steep peaks in the fitness landscape.**

This is quantified as the average fitness of mutants. Either:

- Relative to the a reference genotype. The reference genotype can be the most common genotype, the fittest genotype, the common ancestor genotype of a later population. This is common in in silico studies, as pioneered in the replicator model of Eigen et al. (1979,1989).
- Relative to a reference population, e.g. the wild-type population relative to the population after mutagenesis. In this case the measurement may or may not include a selection filter of the mutant population. This type of quantification is the most suitable one for in vitro/in vivo evolution.

A steep fitness peak means that an individual is non-robust to mutations as its mutated offspring loses much fitness. However at the same time at the population level a steep fitness landscape can convey high robustness because mutants are quickly purged from the population because of the large fitness difference, leading to a homogeneous population (Krakauer & Plotkin, 2002; Elena et al., 2007; Colizzi and Hogeweg, 2014).

### **2.1.2. Robustness as fraction of neutral mutants ( $\lambda$ ) in the close neighborhood.**

This measure was first used in in silico the studies of RNA phenotype landscapes. A high  $\lambda$  implies a high individual robustness: despite mutations offspring retains fitness. Which connects genotypes with the same fitness (Huynen et al., 1996). Indeed many 'network' properties of the neutral network, like connectivity profiles, are relevant for the fate of an evolutionary system. Such neutral networks can percolate through genotype space, and may be intertwined with other neutral

network. These more global properties can only be measured by extensive protocols (or indeed by evolution itself!). For the discussion about robustness, the most salient point is that large percolating neutral networks lead to large variability in the population.

### 2.1.3. Robustness as the fitness profile of mutants.

A more complete picture is given by studying the fitness profile of mutants, i.e. the frequency diagram of the relative fitness of all mutants relative to a reference genotype (selected as above). In such a profile both “steep” and “flat” properties can co-occur, and are indeed observed experimentally. Observed frequency profiles in yeast (Wloch et al., 2001) and viruses (Sanjuan et al., 2004) show a 'U shape', i.e. mutants tend to be either neutral or strongly detrimental, whereas there are only few slightly deleterious mutants. In other words a combination of steepness and high  $\lambda$  and therewith robustness at the individual and the population level.

*As discussed below our de novo evolutionary simulations show that the evolution of such a U shape is a generic property of evolutionary systems that can be 'tuned' to match evolutionary demands.*

## 2.2. Evolution toward (which type of) robustness?

The degree and the type of Robustness of extant organisms is a product of long-term evolution. Within the landscape metaphor it can be seen as the movement of the population to a region of the landscape where certain robustness properties hold.

It has been recognized for a long time in experimental evolutionary studies that a population recently evolutionary adapted to a novel environment have a 'higher mutational load', i.e. contain more non-fit mutants (Scharloo, 1991). Average fitness of the mutant population was used in these studies.

Classical computational experiments on the effect on robustness of long-term evolution have led to the following seemingly contradictory conclusions:

- If population size and mutation rate is large enough long term evolution leads to a flatter part of the fitness landscape (Huynen and Hogeweg, 1994) and to a higher  $\lambda$  (van Nimwegen et al., 1999.), i.e.  $\lambda$  converges to the largest eigenvalue of the connection matrix of the neutral network iff ( $\lambda N > 1$ ). In other words long-term evolution leads to individual robustness, and large population variability.
- On the other hand Krakauer and Plotkin (2002) as well as Elena et al., (2007) showed that under high mutation rates a large enough population will go to a steep part of the landscape, i.e. to low individual robustness, and low population variability (high population robustness).

These contradictory results can be explained by the different fitness landscapes considered. In the first case landscapes considered were the RNA phenotype landscapes, of RNA sequence to secondary structure mapping. These landscapes contain large neutral networks. Thus the analysis was done during long-term neutral evolution, where maximum fitness remains the same. In contrast, in the latter case considered landscapes designed such that the average fitness around a

steep peak and a flat peak were the same. In both cases evolution leads to higher fitness of the population (quasi-species), in the former case due to more neutral mutants, in the latter case because only the larger populations could sustain a large fraction on the high peak, and hence a higher average fitness of the population.

In simple evolutionary systems the evolutionary trajectory can 'choose' a region in the predefined landscape, and therewith depends on the predefined landscape. More generalized evolutionary systems can through evolution shape their mutational neighborhood. Below we will discuss evolutionary strategies to cope with high mutation rates and environmental fluctuation. We show that the evolution of the mutational neighborhood plays a large role, thus underlining evolution of evolution as an important survival strategy.

### **2.3. Evolutionary consequences of evolved mutational robustness: evolvability!**

Intuitively robustness and evolvability seem to contradict each other. However this is not necessarily the case. If we look at robustness at the individual level, defining robustness as a high fraction of neutral mutants (high  $\lambda$ ) this leads to high genomic variability at the population level. High variability leads in turn to high evolvability, in terms of the chance of the occurrence of innovations in the population leading to higher fitness. A high degree of neutrality allows a population to explore genotype space by moving over a neutral network.

When the neutral network percolates genotype space, and when neutral networks of different phenotype are interlaced such exploration will enhance innovation. Indeed 'natural' fitness landscapes are structured in this way. This was first shown in the case of the RNA genotype to phenotype mapping (Huynen, 1996), where also the evolution toward robustness and the concurrence of robustness and evolvability was quantified. Similar results apply for protein landscapes and regulatory landscapes (Wagner 2005,2008; Ferrada & Wagner, 2008, Ciliberty et al., 2007)

On the other hand, when we consider population level robustness generated in a steep landscape, consisting of isolated peaks (i.e. low  $\lambda$ , highly deleterious mutants) such robustness indeed limits evolvability: the population is concentrated around the fitness peak.

The above mentioned U shaped mutational neighborhood, which is observed in yeast and in viruses, and evolves in several of our computational evolutionary experiments combines high neutrality and high selection pressure, does allow for high evolvability. We therefore conclude that *evolvability and robustness are generally compatible in evolutionary systems.*

## **3. Mutational robustness in generalized evolutionary systems: Differential effect of mutational operators**

---

### **3.1. Introduction**

As discussed in D3.1, point mutations are not the only, and probably not even the most important mutational operator in evolving systems. Gene duplication and deletions as well as gross

chromosomal rearrangement play an important role both in vivo and in silico evolutionary systems. This leads to variable genome size.

Strictly speaking, the landscape metaphor breaks down: the variable genome size implies a variable dimensionality of the landscape, and the mutational operators prevent the spatial mapping of genotypes based on mutational proximity. Nevertheless the measurements of mutational robustness listed above are of course feasible. However in setting up these measurements, and interpreting the results, one should be aware of the following complications.

- Typically, we see early genome inflation, followed by streamlining (Cuypers and Hogeweg, 2012, 2014; Batut et al., 2016a, 2016b; Fischer et al., 2014). Fitness increase continues during streamlining. Through variation in genome size, mutation rate is an evolvable property. Reducing genome size, and therefore mutation rate, is a mechanism to increase robustness relative to the mutation that actually happens, and is commonly seen in long-term evolution of such systems. For measuring changes in robustness in the simple evolutionary systems discussed above, a fixed number (generally 1, rarely 2) of point mutations is applied. However, doing this in case of flexible genomes the increased robustness due to lower mutation rates is ignored, leading to an underestimate of evolution of robustness. On the other hand, if instead we apply the actual mutation protocol, one should be aware that many “mutants” will not be mutated, leading to a large overestimation of neutrality of mutations. Nevertheless, decreasing genome size by itself is a strategy to increase robustness.
- Various mutational operators will have different effects on fitness, and these effects will be shaped by their presence/absence in the evolutionary history. Although we can, of course measure the robustness relative to each of them, only by using the original mix of mutations can we assess the evolution of robustness relative to the pressures it experienced.

*We will see that evolution can and will tune the occurrence of different types of mutations, and this is a powerful strategy to maintain fitness, at the level of individuals and/or populations*

## **3.2. Evolutionary strategies to cope with high mutation rates: Shaping the mutational neighborhood**

### **3.2.1. Evolution genome structure in mutator strains**

In the LTEE a subset of the population evolved so called “mutator” strains in which mutation rate increased hundred fold by the loss of a repair pathway. Contrary to classical theoretical predictions, these strains did not lose fitness in the short term by these high mutation rates and even gained fitness slightly faster in the long term than the non-mutator strains. We studied the effect of hundred fold increase in mutation rate in the “coli-like” wild-types we evolved (see milestone MS10) using the aevol platform, by calculating the mutational neighborhood of the wild type strains and 'mutator strains' (Rutten et al., 2016a, 2016b). Unlike in the in vitro experiments we can directly compare the evolutionary trajectory of a wild-type and a mutator clone starting with an identical genome (in the in vitro experiment all clones derive from the same ancestor, but these have diverged when the mutator evolves and we do not have non-mutator from that time-point onwards). We observe the following:

- The immediate effect of the higher mutation rate is a decline of the average fitness of the population. This is to be expected because of the higher mutational load, without change in the coding structure (GP mapping).
- The fitness of the ancestor lineage also declines rapidly: the prevalent mutational neighborhood cannot sustain high individual fitness through selection at the high mutational load.
- However in the long run the mutator ancestor lineage regains a similar (and sometimes even higher) fitness than the matching wild-type lineage. However in doing so the average fitness of the population declines, whereas over the same period both the lineage and the average fitness increase.
- The population fitness of the mutator strain can be understood in terms of changes in the mutational neighborhood. There is a U shape mutational neighborhood in both the wild-type and the mutator populations. However there is a subtle change in its shape: whereas the degree of neutrality remains virtually the same (despite the higher mutation rate, and therefore less 'non-mutations') the number of slightly deleterious mutations decreases and the number of highly deleterious mutations increases drastically, i.e. the U shape 'deepens', and gets more skewed toward deleterious mutations. Therewith the ancestor strain is stronger selected relative to its mutants: it can therefore retain high fitness, but at the cost of lower population fitness.
- Interestingly, and unexpectedly, the mutator lineage fitness regains fitness by an increase in genome size. As mentioned above an increase in genome size increases per genome mutation rate even further, and would therefore be expected to increase the harmful effects of higher mutation rates. However increase of genome size differentially increases the amount of different types of mutations: while the number of point mutations as well as LCR events increase linearly with genome size, the impact of LCR increases super linearly as the size of the stretch duplicated or deleted scales with genome size as well
- Scrutinizing the evolved genomes shows that the size of the non-coding genome increases whereas that of the coding genome decreases (coding is here both protein coding and promoter coding).
- The increase of non-coding genome increases the proportion of highly deleterious mutations and therewith selection strength. The decrease of the size of coding part helps to preserve the neutrality.
- In contrast the wild-types evolve during the same period to increase the coding part and decrease the non-coding part of the genome. The U shape is deepened 'symmetrically', and both average fitness and lineage fitness increase.
- The mutational neighborhood of de novo evolved strains with mutation rates equal to the mutator is very similar to that of the wild-type, also quantitatively, and deepens symmetrically in late evolution.

*Thus the evolutionary strategy of coping with suddenly increased mutation rates is skewing the U shaped mutational neighborhood to more deleterious mutations, therewith increasing selection pressure. This is accomplished through differential increase of the size of the genome, and therewith the ratio of the different types of mutations. In long-term evolution, independent of mutation rate, the U shape deepens symmetrically, increasing robustness by both increasing neutrality and selection, i.e. increasing both individual and population based robustness.*

### 3.2.2. Functional differentiation of mutational neighborhood

So far we have only considered the evolution of mutational neighborhood in terms of the fitness profile of mutants, and as only interaction within a population competition, mediated by fitness differences.

However, in more open-ended systems, evolution can shape the mutational neighborhood even more profoundly, selecting not only a fitness profile but also a functional profile of mutants (Colizzi & Hogeweg, 2014) We have shown this in our studies of the RNA world, a presumed stage in early evolution. How to cope with high mutation rates is a long-standing question in the study of early evolution. The presumably high mutation rates should limit the amount of information that can be maintained (i.e. limits robustness) and limits the amount of information which can be accumulated evolutionary, i.e. should limit evolvability.

We studied the evolution to high mutation rates in a RNA world, consisting replicators with as genotype an RNA sequence, and as phenotype RNA secondary structure. A particular secondary structure conveyed replicator activity, i.e. catalyzed the template based replication of RNA's it could bind to its 5'primed end. An arbitrary RNA sequence with replicator capability survives only moderate mutation rates. However by increasing mutation rates gradually RNA strains were evolved which could cope with much higher mutation rates (contrary to the previous one, only point mutations were allowed in this study). At lower mutation rates a multi species community evolves (Takeuchi & Hogeweg, 2008) In contrast at high mutation rate a very homogeneous population survives.

Studying the mutational neighborhood we found the following:

- With respect to fitness, the mutational neighborhood of the most frequent and 'best' (i.e. most fit in competition experiments) evolved strain has an extremely skewed U shape: almost all mutants are not viable (i.e. cannot be replicated), and there is very low neutrality ( $\lambda$ ), i.e. the local landscape is extremely steep.
- Strikingly, the non-viable mutants are very 'special', in the sense of being very atypically relative to arbitrary replicators, and actually help the survival of the population.
- There is a strong over-representation of 'helpers' that catalyze the replication of the replicator, and there are no 'parasites' (cheaters). At a somewhat larger mutational distance there is a strong over-representation of 'stallers', which harm the replication of replicators.
- In this spatial extended system, mutational neighborhood is reflected in the spatial neighborhood, and the evolved mutational neighborhood in fact creates a supportive niche (with helpers and no parasites) around the ancestor lineage, with a defensive layer around it (with stallers), keeping competitors and parasites away.

*Thus, at high mutation rates a steep and highly atypical, functionally differentiated mutant profile evolved which is essential for the survival (robustness) of the population.*

### 3.2.3. Evolvable mutations

Above we have seen that mutation rate, and the ratio of different mutational operators can change through changes in genome size, and that this can be used to attain and regain fitness. In natural

population we see in addition evolved mechanisms that influence both the rate and the mix of mutational operators. Recently an interesting example has been described in yeast. High transcription rates induce a high mutation rate: this appears to be an unavoidable mechanical effect. In yeast an intricate genome organization and regulatory system has evolved with respect to ribosomal genes, which accomplishes two things: (1) the increase in mutation load due to transcriptional load is molded such that they affect InDel rate much stronger than point mutation rate. (2) High resource availability in combination with transcriptional load moreover biases the InDels toward duplications rather than deletions. The latter leads to fast increase of the number of ribosomal genes under high resource condition. This increase does, however not lead to higher fitness, and therefore these genome changes are not under direct selection.

In a special purpose model we show that these observed features in yeast appear to have evolved to accomplish long term evolutionary robustness: a clear-cut example of evolution of evolution (Colizzi & Hogeweg, 2016c).

- At a low basal mixture of point mutations and InDels long-term evolution leads to low fitness populations with large genomes, containing many pseudo genes (non-functional genes).
- When an increase in mutational load due to transcriptional load is implemented, things get even worse, as should be expected.
- However, if the ratio of these transcription-induced mutations is allowed to evolve, they evolve towards exclusively InDels, (point mutations) remain at the basal level).
- This results in high fitness populations with fairly small genomes, containing almost no pseudo-genes.
- If in addition we bias basal InDels to deletions (as is the case in yeast), we see moreover a bias of these induced mutations toward duplications.
- Like in yeast the evolved large arrays of ribosomal genes can be reduced drastically without loss of fitness.

As already discussed in D3.1, different mutational operators affect the evolutionary dynamics differently. InDels and LCR appear to be 'good for evolution'

*Like in the case of the mutator strains, this yeast example shows that an increase in mutation rate biased toward InDels/LCR increases robustness, both on a short timescale and on long evolutionary timescales.*

## **4. Evolutionary strategies to cope with Variable environments**

---

### **4.1. Introduction**

In variable environments not only mutational robustness but also environmental robustness must be considered. In considering variable environments an important distinction is whether the environmental change can or cannot be sensed by the evolving system.

Various 'solutions' to cope with a variable environment are tabulated in table 1 in terms of the relation between genotype, phenotype and fitness.



Table 1: Robustness strategies in variable environments

Genotype	Phenotype	Fitness	
Constant	Constant	Constant	(evolved) independence
Constant	Variable	Constant	Regulatory response (plasticity)
Variable	Variable	Constant	Evolvability
Constant	Constant	Variable	Frequency of strains/species; Population/ecosystem robustness

The 'how and when' of these possibilities are discussed below.

## 4.2. Externally imposed environmental variability

### 4.2.1. Regulation and/or evolvability

In previous work we studied, in the virtual cell model, mechanisms and consequences of robustness to externally imposed environmental fluctuations of more than 3 orders of magnitude, using homeostasis as fitness criterion (Cuypers and Hogeweg, 2012, 2014). By using homeostasis as fitness criterion we strongly selected for regulatory robustness. If fluctuations were fast enough, i.e. an individual experienced typically more than one environment in its lifetime, strict homeostasis evolved in about half of the evolutionary runs. Once such regulation is achieved, the environment appears constant for the ongoing evolution. Indeed with respect to mutational robustness, we observe the same trends as in a constant environment: a U shaped mutational neighborhood evolves, which deepens during ongoing evolution.

We next studied how these evolved homeostatic virtual cells could cope with environmental changes at a much longer timescale (10-1000 generations) (Cuypers et al., 2016). We used as proxy for these additional environmental changes, resetting non-evolvable parameters of the model (such as degradation rate, conversion stoichiometry, diffusion, and set-point for the homeostasis). Note that the first three change the internal state of the cell, and therefore can be 'sensed', whereas the latter cannot be sensed. This means that, in principle coping to the former changes by regulation is possible, whereas to the latter it is not. We observe the following:

- Recovery of fitness (i.e. homeostasis) is remarkably fast and most often achieved.
- This is also true for set-point changes, and is achieved by evolution.
- Repeating the long-term environmental changes, improves the 'robustness' to these changes, i.e. evolvability improves.
- Focusing on sensible changes, we studied the influence timescale of these changes, and observe:

- For short term changes (10-20 time-steps) only regulatory robustness evolved, i.e. full fitness is kept despite the environmental change, by changing expression levels of the genes.
- However evolving such regulation required a large number of repeated changes.
- In contrast, coping with these change through evolvability occurred faster, often needing only a few environmental switches.
- At intermediate rates of change (30 time-steps) both regulatory and evolvability evolved, from the same initial conditions, and the same switching protocol. The average fitness over time was very similar, for both types of solution, but was reached earlier for the evolvability strategy.
- In the long run the evolvable strategy evolved in some cases into the regulatory strategy, in which, of course fitness is continuously maintained.
- During the evolution of regulation or evolvability, the mutational robustness, i.e. the U shaped mutational neighborhood is virtually unchanged.

*We conclude that evolving evolvability is 'easier' than evolving regulatory robustness to environmental change. Moreover we conclude that regulatory robustness to some environmental changes enhances the evolution to evolvability to other repeated environmental changes, and indeed to cope with novel condition.*

Many in vitro evolutionary experiments to novel environmental conditions indeed show very fast evolutionary changes to adapt to these condition, and often involve the same genes/pathways which are also part of regulatory responses (e.g. Ferea et al., 1999).

#### **4.2.2. Population variability and regulation in periodical refreshed environments**

In the LTEE experiment very regular environmental variation is imposed. Every 24 hours a subset of the population is inoculated into a fresh medium. In the first few hours this population grows exponentially, exhausting most resource. It then switches to stationary phase. In some of the experiments the population split into 2 strains: one using the primary resource, whereas the other uses secondary resource produced by the first population.

We mimicked the experimental protocol in two platforms, the integrated evolutionary model developed to study EvoEvo (Evo<sup>2</sup>Sim) (Rocabert et al., 2015, 2016a) and in the virtual microbe model (VM) (van Dijk and Hogeweg, 2016, van Dijk et al., 2016).

The two modeling frameworks both model a diverse metabolic environment, and a pearl on a string-like genome. Important differences are a predefined fitness criterion in Evo<sup>2</sup>Sim, and an open-ended metabolic and enzymatic universe, whereas in the VM model the metabolic universe is restricted, and enzymes cannot unlimitedly switch specificity. On the opposite, in the VM model there is no predefined fitness, growth rate depending on the production of certain building blocks but (longer term) fitness being achievable by low growth rate.

We observe the following:

- In the Evo<sup>2</sup>Sim model two subspecies reliably evolved, one using the primary resource and the second the 'waste' products of the other species (Rocabert et al., 2016a, b).



- The first species dominates the population early in the cycle, whereas the second does later on. The population size as a whole remains constant after the growth phase, i.e. it is robust to low resource concentrations.
- The speciation depends crucially on the regularity of the protocol: a population is transferred to a fresh environment, containing only the primary resource at regular interval, just like in the LTEE experiments, but varying the time-span prohibits the stable maintenance of the two lineages.
- Both these lineage are variable and keep evolving but their evolution are independent (i.e. mutants invading one of the lineages does not invade the other)
- Sustained regulation was never observed in the Evo<sup>2</sup>Sim model in those conditions.
- In the VM model cells were pre-evolved in a mildly temporal and spatially variable environment. Transferring to the LTEE protocol, the population-based solution is rarely observed, but in the majority of cases a regulatory solution evolved.
  - In the first part of the cycle exponential growth occurs, in the latter part cells switch to a stationary phase, with low expression levels of pump and enzymes.
  - During evolution this regulation is tuned to the periodicity of the environment.
  - Initial growth rate is increased, and resources are exhausted more rapidly, as is also observed in the LTEE.
  - Cell size increases over evolutionary time in this growth phase (like LTEE).
  - Cell size decreases in stationary phase (like LTEE)
  - Population size decreases during stationary phase (unlike LTEE), and this decline is precisely tuned to the period: in some cases decreasing to the transfer size, so that the whole population is transferred.

*In both cases we see that the populations are precisely tuned to the periodicity of the environmental change. This implies long-term information integration. It also implies that the resulting populations are sensitive (non robust) to changes in the periodicity. This leads to the interesting question ether or not this is also the case for the LTEE populations.*

### **4.3. Population induced spatial-temporal variation**

Above we discussed how populations cope with externally imposed environmental variation. However, in spatial multiple species/multiple strains eco-evolutionary systems, as the system discussed in the previous subsection are, major environmental variation is induced by the populations themselves, whether or not there is externally induced variation. Moreover the environment experienced by individuals is deviates from the average environment. Thus in order to assess mutational robustness in a manner discussed in the single species systems discussed above, we have to take this 'self-generated' spatial neighborhood into account.

Moreover, if no external fitness criterion is used, only competition with the locally occurring specimen can be used to assess the fitness profile of the mutational neighborhood. We did this in a model of a bacterial population competing (fighting) through the production of toxins and antitoxin. Apart from InDels, horizontal gene transfer (HGT) was present as mutational operators. Focusing on gaining/losing toxin and antitoxin genes by HGT we see, like above, that a (skewed) U shape mutational neighborhood is tuned during evolution, and is quite different relative to the local or the

global (average) competing strains. Indeed structuring the local neighborhood is an important means for long-term evolutionary robustness, as discussed in the next section.

## 5. Evolutionary robustness

---

### 5.1. Introduction

From our experiments a novel concept emerged: evolutionary robustness, which unlike previous discussed concepts cannot be expressed in terms of mutational neighborhood, because it operates at longer evolutionary timescale, and more importantly, may be caused by self-destructive selection pressures.

Long-term evolutionary deterioration that can lead in the long run to extinction, is a well known phenomenon, which occurs when selection is not strong enough relative to mutational load. The former is known as Muller's ratchet in finite populations and as the error threshold in infinite population. As we have seen above, evolution of the mutational neighborhood, increasing both neutrality as well as selection pressure (deepening of the U shaped mutational neighborhood) can help to prevent/postpone this. Moreover, we have seen that an increase in InDels/LCR relative to point mutations is a powerful strategy to avoid such mutation-induced meltdown.

Moreover we have seen that in yeast genome structure and regulatory mechanisms have evolved to 'repair' mutations that are in the short term neutral. Our models show that this prevents the long-term evolutionary deterioration these mutations will cause. Thus these mechanisms evolved to ensure *evolutionary robustness*.

Evolution towards extinction through 'destructive' selection is maybe the best-known systems which depend on their survival on costly cooperation, i.e. which require 'altruistic' behavior. Preeminent example are replicator systems in which there is negative selection for giving catalysis, e.g. the RNA world discussed above, or communities which depend on the production of an expensive common good (where the phenomenon is known as the tragedy of the commons). Increasing short-term individual fitness will lead to a decrease of cooperation and/or the emergence of parasitic lineages, and extinction. This is true for well-mixed systems. In contrast in spatial systems with local interactions such extinction may happen locally, but the system as a whole can survive. This is due to the emergence of spatial patterns 'with a dynamics of their own'.

### 5.2. Spatial self-organization and multilevel selection

Spatial pattern formation leads to multilevel selection: the spatial patterns themselves act as a unit of selection, as was already shown in a spatial hypercycle model (Boerlijst & Hogeweg, 1991) and more extensively in the minimal replicator model (RP model) (Takeuchi & Hogeweg, 2008, 2009; Colizzi & Hogeweg, 2016a) as well as a classical common good model (Colizzi & Hogeweg, 2016b). Selection pressure ('fitness') (as defined as survival – mediated by birth and death rate) of the spatial patterns can be opposite of that of the individual replicators, and can prevent evolution toward extinction of the individual replicators, and enable the evolution of more complex, more efficient individuals and more robust individuals.

Here we mention three results on the consequences of such multilevel organization on the robustness and evolvability of the system as a whole, as well as the individuals.

### 5.2.1. Mutation rate and the emergence and persistence of diverse ecosystems

In the HGT toxin-antitoxin systems mentioned above, three levels of selection operate: the gene level, the individual level and the ecosystem level. We studied the evolution of HGT rates (van Dijk and Hogeweg, 2015). Despite the occurrence of 'killers', i.e. individuals harboring toxin genes, which could kill anybody in the population, a robust, diverse ecosystem persists, as, indeed is also observed in nature (Cordero et al., 2012). This robustness depends on all three levels of selection:

- In well mixed systems the 'killer' species would drive the system to extinction, only due to spatial pattern formation this is prevented.
- HGT rates of toxin genes evolve to high levels, whereas those of anti-toxin genes are minimized
- High HGT rates are needed for the survival of toxin genes: due to spatial pattern formation they mostly harm and only rarely benefit the individual which harbors.
- These individuals, and therewith the genes they harbor tend to be out-competed.
- If HGT rates of both toxin and antitoxin genes are set to the same value, whether this value is high or low, and do not evolve, only a single species ecosystem survives in the long run (in absence of external influx of genes/species).
- This single species has neither toxin nor antitoxin genes, i.e. has a minimal genome.

*We conclude that multilevel selection leads to a robust ecosystem, which ensures the survival of relative complex individuals*

### 5.2.2. Multilevel selection and the evolution of DNA in the RNA world

In earlier work (Takeuchi et al., 2011) on the evolution of DNA in the RNA world is a preeminent example of evolution of mechanisms for evolutionary robustness. It works as follows:

- The RNA replicators have a selection pressure to minimize replication of other molecules, down to a minimum level for survival.
- By evolving DNA the replication cycle becomes longer (DNA is transcribed in RNA, which catalyzes the replication of DNA), and therefore slower.
- However, the evolved DNA-RNA interactions establish division of labor between information storage (DNA) and information usage (RNA).
- There is no selection pressure to lower the amount of catalysis for DNA, as it does not perform the catalysis itself, only stores the information.
- Information flow is strictly from DNA to RNA (no reversed transcription). in accordance with 'the central dogma' (Crick, 1971).
- Consequently, the RNA, transcribed from the DNA, maintains high catalytic activity.
- When DNA is locally lost, the remaining RNA based system initially out-competes the DNA-based systems, but subsequently loses catalysis strength and is out-competed by the DNA-based system in the longer run

*Thus multilevel selection enhances evolutionary robustness at the lower level, and can explain the important evolutionary transition from the RNA world to the DNA based world we know today*

### **5.2.3. Spatial self-organization and ‘open ended’ evolution in StringMol**

StringMol (Hickinbotham et al., 2016) is a ‘RNA-like’ computational system, which, like in the RNA world discussed above, consists of replicators that replicate others by complimentary binding. However, unlike the RNA world model, this replication is not automatic upon binding given the correct structure, but it is accomplished by the program coded (as a string) in the replicator. Upon mutation, this program changes, and therewith the replication process may change, introducing new types of mutations (e.g. InDels, hybridization, etc). Thus, unlike all models discussed previously it evolves novel mutational operators, instead of just tuning their frequency.

Unfortunately previous experiments with StringMol suffered heavily from systematic extinction through parasites, before long-term evolution could happen. Implementing the system in space (Hickinbotham & Hogeweg, 2016) shows that also in this case spatial self-organization is sufficient to solve the parasite problem. Because replication happens in this system one operation at the time, this result is quite remarkable, despite its ubiquitous occurrence in previous studied systems, because parasites can evolve not only strong binding, but can also be much shorter, and therewith replicate an order of magnitude faster.

Therefore they could be much more destructive. Nevertheless ecosystem level robustness emerges, in which evolving replicators escape parasite onslaught and recolonize empty space created by local extinctions. In the robust but very dynamic ecosystem selection pressures vary over time and long-term (and so far ‘open-ended’) evolution of the replicators unfolds, leading in turn to novel selection pressures.

*We conclude that emergence of higher-level selection by spatial self-organization and the resulting ecosystem level robustness and dynamics is essential for open-ended evolution of evolution at the individual level.*

## **6. Conclusions**

---

Because of the complexity of biological systems, every model should make important simplification. In order to derive general conclusions have studied a variety of modeling platforms and models, which incorporate different simplifications and constraints. Here we summarize common features of the evolution of robustness and its consequences which emerge in several of these platforms.

The platforms used are:

- Aevol, incorporating a fine-grained genomic representation, including coding and non-coding regions, a fixed GP map and a fixed fitness criterion, and using a wide range of mutational operators.
- Integrated evolutionary model (Evo<sup>2</sup>Sim), which combines features of AEVOL and the pearl on a string encoding, and incorporating a very rich metabolic universe, consisting of natural

numbers. Influxing just 1 metabolite, cells gain fitness in terms of the number of prime number metabolites they contain.

- Virtual cell (VC) and virtual microbe (VM) model, modeling a virtual cell, which use a coarse grained pearl on a string genome representation and a rich set of mutational operators. They evolve their GP map, and are studied in a variable environment. In the virtual cell model fitness in varying environments is predefined (homeostasis) whereas in the more elaborate virtual microbe model, no prior fitness criterion is defined except the ability to grow and survive.
- RNA world model, using the RNA GP mapping (earlier described as the ideal evolvable molecule), where replication occurs as a cooperative event: a replicase replicates other molecule, if it can bind to it. Embedding in space is a requirement for viability. No predefined fitness criterion except the ability to grow and survive. Only point mutations allowed.
- HGT model, a simple spatial toxin/antitoxin model in which horizontal gene transfer is added to the genetic operators.
- StringMol, which implements a RNA-like world, in the form of computer programs encoded in a string, which actually perform the copying, one instruction at the time. The evolving program, evolve new types of mutation and new types of interaction between the strings.

We derive the following general conclusions:

- The classical way of describing robustness in terms of at or steep landscape should be replaced and refined.
  - Abstracting from the landscape metaphor, a useful concept is mutational neighborhood that can be applied in a variable environment, a variable genome size, and a variable set of mutational operators.
  - The concept mutational neighborhood refers to the profile of fitness effects of mutants, and can be extended to other properties of the mutants apart from fitness.
  - A 'U shape' mutational neighborhood evolves, and enhances fitness by combining high neutrality ('flatness') and high selection ('steepness') Such U shapes mutational neighborhoods have been measured in viruses and yeast. Moreover it evolves in Aevol in a constant environment, in the virtual cell in a variable environment, and in the RNA world without explicit fitness criterion, as well as in the HGT model. For technical reasons, the U-shape has not been studied in the other models but we expect it to evolve there as well.
- An important mechanism for attaining and retaining robustness under various mutational regimes is the fine scale adjustment of the depth and skewness of the U shape. Indeed, at high (point) mutation rates the U becomes more skewed to deleterious mutations (Aevol mutator strains, RNA world at high mutation rates). Independent of mutation rate the U shape deepens during evolution.
- Evolving and tuning the U shape can occur by simply seeking the appropriate portion of a fixed fitness landscape as was seen in the RNA world, or by deformation of the local landscape by InDels and LCR.

- Moreover, the U shape can be modified by tuning the frequency of different mutational operators: InDels and LCR tend to have other (often more disastrous) effect than point mutations.
- Tuning of the frequency and of mutational operators happens automatically by changes in genome size. By adjusting genome size and the ratio of coding and non-coding parts, stability and fitness can be retained over a large range of point mutation rates.
- In nature many mechanisms have evolved to regulate mutation rate, as well as the occurrence of different mutations. We have shown in a special purpose yeast model that the evolved structure ensures long-term evolutionary stability, whereas it does not affect fitness directly. In other words it has evolved for evolutionary robustness: a preeminent example of evolution of evolution.
- In multispecies spatial evolutionary systems multiple levels of selection automatically emerge through spatial pattern formation.
- Spatial self-organization is a powerful mechanism to reverse, or at least contain, self-destructive evolutionary trends.
- The ecosystem level robustness resulting from self-organization, allows the evolution of novel genetic mechanisms for robustness to evolve.
- Persistence of genes, individuals, populations and spatial patterns are intrinsically linked, should be considered together.

In conclusion:

*Robustness and Evolvability are tightly linked: Robustness is an evolved property, one of the mechanisms for robustness is evolvability and one of the consequences of robustness is evolvability.*

## 7. References

---

B Batut, G Beslon and C Knibbe (2016a) Unexpected genome inflation and streamlining in variable environments. *Proceedings of Jobim 2016, Lyon (France), June 2016*, pp. 320-322

B Batut, G Beslon and C Knibbe (2016b) Genome inflation under variable environments. *In prep.*

MC Boerlijst and Hogeweg (1991) Spiral wave structure in pre-biotic evolution: Hypercycles stable against parasites. *Phys D Nonlin Phenom*, 48(1):17-28

S Ciliberti, OC Martin and A Wagner (2007) Innovation and robustness in complex regulatory gene networks. *PNAS*, 104(34):13591-13596

ES Colizzi and P Hogeweg (2014) Evolution of functional diversification within quasispecies. *Genome biology and evolution*, 6(8):1990-2007

ES Colizzi and P Hogeweg (2016a) Parasites Sustain and Enhance RNA-Like Replicators through Spatial Self-Organisation. *PLoS Computational Biology*, 12(4): e1004902.

ES Colizzi and P Hogeweg (2016b) High cost enhances cooperation through the interplay between evolution and self-organisation. *BMC Evolutionary Biology*, 16:31

ES Colizzi and P Hogeweg (2016c) Mutational load is ameliorated by increased transcriptional load-associated mutations, if these are biased towards duplications and deletions. *2<sup>nd</sup> EvoEvo Workshop, Amsterdam (NL), September 2016*, 5 p.

OX Cordero, H Wildschutte, B Kirkup, S Proehl, L Ngo, F Hussain, F Le Roux, T Mincer and MF Polz (2012) Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. *Science*, 337(6099):1228-1231

F Crick (1971) Central dogma of molecular biology. *Tsitologiia* 13(7):906-910.

TD Cuypers and P Hogeweg (2012) Virtual genomes in flux: an interplay of neutrality and adaptability explains genome expansion and streamlining. *Genome biology and evolution*, 4(3):212-229

TD Cuypers and P Hogeweg (2014) A synergism between adaptive effects and evolvability drives whole genome duplication to fixation. *PLoS computational biology*, 10(4): e1003547

TD Cuypers, JP Rutten and P Hogeweg (2016) Mutate or Regulate: evolutionary strategies along a continuum of ecological time scales. (*in prep*).

B van Dijk and P Hogeweg (2015) Multi-level evolution, differential gene mobility, and the persistence of population diversity, *Proceedings of the European Conference on Artificial Life 2015, York (UK), July 2015*, pp. 9-11

B van Dijk and P. Hogeweg (2016) In silico gene-level evolution explains microbial population diversity through differential gene mobility. *Genome biology and evolution*, 8(1):176-188

B van Dijk, T Cuypers and P Hogeweg (2016) Evolution of r- and K-selected species of Virtual Microbes: a case study in a simple fluctuating 2-resources environment. *2<sup>nd</sup> EvoEvo Workshop, Amsterdam (NL), September 2016*, 5 p.

M Eigen and P Schuster (1979) The hypercycle: a principle of natural self-organization. Berlin (Germany): Springer-Verlag; 1979.

M Eigen, J McCaskill and P Schuster (1989) The Molecular Quasi-Species, In: *Advances in Chemical Physics*, Volume 75 (eds I. Prigogine and S. A. Rice), John Wiley & Sons, Inc., Hoboken, NJ, USA.

SF Elena, CO Wilke, C Ofria and RE Lenski (2007) Effects of population size and mutation rate on the evolution of mutational robustness. *Evolution*, 61(3):666-674.

TL Ferea, D Botstein, PO Brown and RF Rosenzweig (1999) Systematic changes in gene expression patterns following adaptive evolution in yeast. *PNAS*, 96(17): 9721–9726

E Ferrada and A Wagner (2008) Protein robustness promotes evolutionary innovations on large evolutionary time-scales. *Proc R Soc B*, 275(1643): 1595–1602

S Fischer, S Bernard, G Beslon and C Knibbe (2014) A model for genome size evolution. *Bulletin of mathematical biology*, 76(9):2249-2291

- S Hickinbotham, E Clark, A Nellis, S Stepney, T Clarke and P Young (2016) Maximizing the Adjacent Possible in Automata Chemistries. *Artificial Life*, 22(1):49-75
- S Hickinbotham and P Hogeweg (2016) Evolution towards extinction in replicase models: inevitable unless... *2<sup>nd</sup> EvoEvo Workshop, Amsterdam (NL), September 2016*, 5 p.
- MA Huynen and P Hogeweg (1994) Pattern generation in molecular evolution: exploitation of the variation in RNA landscapes. *Journal of Molecular Evolution*, 39(1):71-79
- MA Huynen (1996) Exploring phenotype space through neutral evolution. *Journal of molecular evolution*, 43(3):165-169
- MA Huynen, PF Stadler and W Fontana (1996) Smoothness within ruggedness: the role of neutrality in adaptation. *PNAS*, 93(1):397-401
- DC Krakauer and JB Plotkin (2001) Redundancy, antiredundancy, and the robustness of genomes. *PNAS*, 99(3):1405-1409
- E van Nimwegen, JP Crutchfield and M Huynen (1999) Neutral evolution of mutational robustness. *PNAS*, 96(17):9716-9720
- C Rocabert, C Knibbe and G Beslon (2015) Towards an Integrated Evolutionary Model to Study Evolution of Evolution. *First EvoEvo Workshop, York (UK), July 2015*, 15 p.
- C Rocabert, C Knibbe, J Consuegra, D Schneider and G Beslon (2016a) Beware Batch Culture: Seasonality and Niche Construction Predicted to Favor Bacterial Adaptive Diversification. *Under review*.
- C Rocabert, C Knibbe, J Consuegra, D Schneider and G Beslon (2016b) In-silico experimental evolution highlights the influence of environmental seasonality on bacterial diversification. *2<sup>nd</sup> EvoEvo Workshop, Amsterdam (NL), September 2016*, 3 p.
- JP Rutten, P Hogeweg and G Beslon (2016a) Evolution of mutator populations in constant environments. *2<sup>nd</sup> EvoEvo Workshop, Amsterdam (NL), September 2016*, 5 p.
- JP Rutten, P Hogeweg and G Beslon (2016b) Adapting the engine to the fuel: hypermutator populations can escape the mutational burden by reorganizing their genome structure. *In prep*.
- R Sanjuan, A Moya and SF Elena (2004) The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *PNAS*, 101(22):8396-8401
- W Scharloo (1991) Canalization: genetic and developmental aspects. *Annual Review of Ecology and Systematics*. 22:65-93
- N Takeuchi and P Hogeweg (2008) Evolution of complexity in RNA-like replicator systems. *Biology Direct*, 3:11
- N Takeuchi and P Hogeweg (2009) Multilevel selection in models of prebiotic evolution II: a direct comparison of compartmentalization and spatial self-organization. *PLoS Computational Biology*, 5(10):e1000542



N Takeuchi, P Hogeweg and EV Koonin (2011) On the origin of DNA genomes: evolution of the division of labor between template and catalyst in model replicator systems. *PLoS Computational Biology*, 7(3):e1002024

A Wagner (2005) *Robustness and evolvability in living systems*. Princeton University Press.

A Wagner (2008) Robustness and evolvability: a paradox resolved. *Proc R Soc B*, 275(1630):91-100

DM Wloch, K Szafraniec, RH Borts and R Korona, (2001) Direct estimate of the mutation rate and the distribution of fitness effects in the yeast *Saccharomyces cerevisiae*. *Genetics*, 159(2):441-452

S Wright (1932) The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the Sixth International Congress on Genetics*.