Testing evolution predictability using the aevol software

Guillaume Beslon, Vincent Liard, Santiago F Elena

To cite this version:

Guillaume Beslon, Vincent Liard, Santiago F Elena. Testing evolution predictability using the aevol software. 16th international meeting of the European Society of Evolutionary Biology (ESEB 2017), Aug 2017, Groningen, Netherlands. 2017. hal-01577115

HAL Id: hal-01577115
https://hal.archives-ouvertes.fr/hal-01577115

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.
Testing evolution predictability using the aevol software

Guillaume Bleson\(^1\), Vincent Liard\(^1\), Santiago F. Elena\(^2,3\)

1: INRIA-Beagle team (INSA-Lyon), Lyon, France
2: IBMCP (CSIC-UPV), Valencia, Spain; 3: Santa Fe Institute, Santa Fe NM, USA

Abstract
Motivated by RNA virus’ genome biology, we used the aevol software to simulate the evolution of compacted genomes under high mutation rates. 30 independent wild-type (WT) genomes were generated after 200,000 generations of evolution under similar conditions. Then, each of these WTs was cloned 30 times and we let evolution to continue for 30,000 extra generations. By comparing these clones, we aimed to reveal the extent of evolutionary predictability for such compacted genomes. Results show that: (i) WTs are not equivalent in terms of evolutionary potential: some WTs are more prone than the others to increase their fitness during the last 30,000 generations. (ii) Evolution frequently occurs in bursts which imply that the probability to fix a mutation is increased after fixation of another mutation. Moreover these bursts are often initiated by chromosomal rearrangements (mainly duplications) because these rearrangements open new evolutionary pathways in the fitness landscape. Indeed, we quantified the “evolvability potential” of every clone after each mutation and found that the bursts are triggered by a strong increase of evolvability that quickly leads to point substitutions and indels fixation.

Methods: in silico experimental evolution (ISEE) with the aevol platform
Aevol (www.aevol.fr) is an ISEE platform that models bacteria at the genomic level with explicit mutation and selection processes. ISEE mimics experimental evolution with in silico organisms subjected to variation and selection in a computer.

Results (1): Virus-like genomes
The high mutation rates lead to the evolution of virus-like organisms: small genomes, around 10 coding sequences, few mRNAs, overlapping genes, and almost no coding base.

Results (2): Lineage analysis shows that innovation occurs in bursts:
• In 178/215 innovators 80% of the fitness gain occurs in less than 20% of the time (including 68 clones that gain 80% of the fitness in a single mutational event).
• Fixed mutation are not randomly distributed along time (bursts of mutations).

Local sampling of the fitness landscape shows that innovation bursts correspond to very specific regions:
• Areas of low robustness (generally due to genome size increase).
• Areas of very high evolvability: innovations are triggered by mutational events that strongly increase the evolvability of lineages.

Discussion: the combinatorics of fitness landscapes
In compacted genomes, like viral ones, the combinatorics of point mutations is quickly exhausted. Yet, the combinatorics of rearrangements is much larger and cannot be fully explored in a reasonable time. Duplications open new paths in the fitness landscape and allows for new favorable mutations to occur, leading to bursts of mutations. Hence, the fitness landscape must be considered as more connected at large scale than at short scale. Moreover, considering only point mutations is likely to miss many important evolutionary phenomena, as large-scale rearrangements allow for evolvability bursts that lead to innovations.

Acknowledgement: Supported by the European Commission 7th Framework (FP7-ICT-2013.9.6- FET Proactive: Evolving Living Technologies) EvoEvo project (ICT-610427).

Drift vs innovation in the 900 clones:
• 163 clones have lost fitness between generations 200,000 and 225,000.
• 251 kept their ancestral fitness.
• 271 marginally improved their fitness.
• 215 significantly improved fitness; 50% of these “innovators” come from WTs.

Experimental design
Local sampling of the fitness landscape:
• For each ancestral genotype of the 900 lineages we generated 10,000,000 offspring.
• From these offspring we can estimate the evolution of robustness, anti-robustness and evolvability along the evolutionary path.

Results (3): triggering events
The 215 innovations were triggered by:
• 104 duplications (93 favorable, 10 deleterious, 1 neutral).
• 33 small insertions (19 favorable, 14 deleterious).
• 10 small deletions (all deleterious).
• 11 point substitutions (all deleterious).
• 57 uncharacterized events due to time resolution.

Innovations are often triggered by duplications

Degree in the fitness landscape of WT2C1
Point mutations 521
Small insertions 65,446
Small deletions 3,126
Duplications 241,149,320
Large deletions 270,920
Translocations 140,607,480
Inversions 270,920

Fixed events Mean fitness improvement Mean evolvability improvement Mean time from previous mutation Mean time to next mutation
Substitutions 1172 7.16 \times 10^-4 -4.41 \times 10^-3 818 407
Small insertions 916 9.99 \times 10^-4 -2.98 \times 10^-3 1176 840
Small deletions 29 2.57 \times 10^-4 -1.47 \times 10^-3 682 250
Duplications 168 4.10 \times 10^-4 2.18 \times 10^-3 1176 836
Large deletions 29 2.57, \times 10^-4 -1.47 \times 10^-3 682 250
Translocations 168 4.10 \times 10^-4 2.18 \times 10^-3 1176 836
Inversions 4 1.70 \times 10^-4 2.37 \times 10^-3 1263 1746

Discussion: the combinatorics of fitness landscapes
In compacted genomes, like viral ones, the combinatorics of point mutations is quickly exhausted. Yet, the combinatorics of rearrangements is much larger and cannot be fully explored in a reasonable time. Duplications open new paths in the fitness landscape and allows for new favorable mutations to occur, leading to bursts of mutations. Hence, the fitness landscape must be considered as more connected at large scale than at short scale. Moreover, considering only point mutations is likely to miss many important evolutionary phenomena, as large-scale rearrangements allow for evolvability bursts that lead to innovations.

Acknowledgement: Supported by the European Commission 7th Framework (FP7-ICT-2013.9.6- FET Proactive: Evolving Living Technologies) EvoEvo project (ICT-610427).