



HAL
open science

Inversion mutations attenuate complexity catastrophe in NK fitness landscapes

Leonardo Trujillo, Paul Banse, Guillaume Beslon

► **To cite this version:**

Leonardo Trujillo, Paul Banse, Guillaume Beslon. Inversion mutations attenuate complexity catastrophe in NK fitness landscapes. CCS2022, Oct 2022, Palma de Mallorca, Spain. 2022. hal-04096866

HAL Id: hal-04096866

<https://inria.hal.science/hal-04096866v1>

Submitted on 13 May 2023

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.

Public Domain

Inversion mutations attenuate complexity catastrophe in NK fitness landscapes

Leonardo Trujillo¹, Paul Banse¹, and Guillaume Beslon¹

¹Inria Beagle Team, Artificial Evolution and Computational Biology,
 Université de Lyon, Inria, ECL, INSA Lyon, Université Claude Bernard Lyon 1, Univ. Lumière Lyon 2,
 CNRS, LIRIS UMR 5205, F-69622, France
 {leonardo.trujillo, paul.banse, guillaume.beslon}@inria.fr

Context In the Kauffman’s NK model of rugged fitness landscapes [1], besides the genome length $N \in \mathbb{N}$, the integer $K \in \mathbb{Z}_{(0, N-1)}$, describes the epistatic interactions between loci and the contribution of each component to the total fitness—which depends on its own value as well as the values of K other loci. The total fitness $f \in [0, 1)$ for a binary sequence $\mathbf{x} \in \{0, 1\}^N$ is then defined as: $f(\mathbf{x}) := \frac{1}{N} \sum_{i=1}^N f_i(x_i, x_{i_1}, \dots, x_{i_K})$, with the fitness per locus f_i which depends on the state of locus x_i and K other loci x_{i_k} . It is usually assumed that the f_i ’s are given by independent and identically distributed random variables sampled from a given uniform probability distribution. Two popular schemes of interaction between loci are: the adjacent and random neighborhood models [1, p. 55]. On the other hand, evolution is simulated as adaptive walks, such that a starting genotype \mathbf{x} varies through point mutations resulting in a mutated sequence \mathbf{y} . Then, if $f(\mathbf{y}) > f(\mathbf{x})$, the mutated genotype is selected, otherwise other mutations on \mathbf{x} are tested until the fitness increases. The evolutionary dynamics stops when local maximum of fitness is reached. As can be appreciated in figure 1 the average of local fitness maxima $\langle f \rangle_K$ is not monotonic for different values of K . In particular for point mutations (red points in fig.1), Kauffman showed that the expected value of local fitness maxima is monotonically non-decreasing up to a certain value of K and then decreases such that $\lim_{K \rightarrow N-1} \langle f \rangle_K = 1/2$ [1]. This decrease of fitness values toward 0.5 is the well-known “complexity catastrophe” [1, p. 52]. That is, as K increases, the number of possible values of f_i is greater. But the number of conflicting constraints due to epistasis can also increase, implying that the accessible local maxima are lower and thus the expected fitness peaks decrease in height.

Results The adaptive dynamics in the NK model have been studied mainly with point mutations. Nevertheless in molecular evolution there are also structural mutations changing more than one nucleotide, for instance: duplications, inversions and other chromosomal rearrangements. These different types of structural variations are completely overlooked in conventional analysis of the NK model. We show that chromosomal inversions reveal interesting features of adaptive dynamics. In particular, we communicate results from adaptive walks simulations and show that the complexity catastrophe can be attenuated via chromosome inversions. This is sketched in figure 1 for a genome of size $N = 100$ and inversion mutations with sizes ranging from 1%—a single locus (i.e. inversions are like point mutations)—up to 100% of the genome size, for patterns of adjacent and random (inset plot in fig.1) epistatic interaction between loci. For the simplest case $K = 0$, with no epistatic interaction between neighboring loci, we can verify in fig.1, that the average fitness for all inversions size percentages are equal $\langle f \rangle_{K=0} \simeq 0.667$. In this case, the landscape is smooth with only a single peak. This result also agrees with the analyt-

ical result $\langle f \rangle_{K=0} = \frac{2}{3}$ [1, p. 55]. Then, for $K > 0$, the average fitness $\langle f \rangle_K$ increases with K until reaching a maximum value of fitness. After these maximum, the fitness values decrease as K increases. When $K \rightarrow 99$, the mean fitness converge to the same value, regardless of the type of epistatic interaction neighborhood. Now, what’s new is that for inversions with sizes greater than 1%, the average fitness values are higher than those for mutations of a single locus. Indeed, for $K \rightarrow N - 1$, the average fitness trend is very different from that of point mutations [1, 2]. The evolutionary process reaches higher expected fitness values $\langle f \rangle_{K=99} > 1/2$, for both random and adjacent epistatic interactions, and thus attenuating the complexity catastrophe. Finally, let’s note as a curious fact, that when the size of the inversion mutations is equal to 100% (blue points in fig.1), then $\langle f_I \rangle_{K=99} = 0.610$. We conjecture that for inversion mutations when $K \rightarrow N - 1$, the complexity catastrophe must converge to the reciprocal of the golden ratio $1/\varphi = 2/(1 + \sqrt{5}) \approx 0.618 \dots$

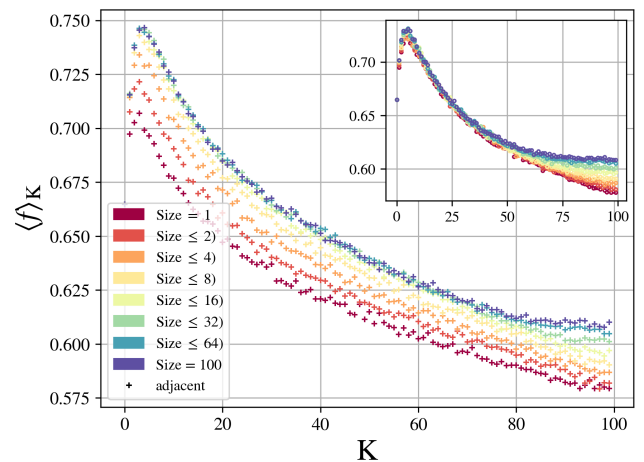


Fig. 1. **The average value of local fitness maxima show that the complexity catastrophe is attenuated by inversion mutations.** The behavior of the higher reached fitness vs the epistatic parameter K for point mutations (red) and inversions restricted to a certain size range (sizes values and colors nomenclature are displayed in the figures), averaged for 100 instances of adaptive walks simulations in NK landscapes with adjacent neighboring epistatic interactions. Inset: Display the results with random epistatic interactions.

- [1] Kauffman SA. The origins of order: Self-organization and selection in evolution. Oxford University Press, USA; 1993.
- [2] Solow D, Burnetas A, Tsai MC, Greenspan NS. Understanding and attenuating the complexity catastrophe in Kauffman’s NK model of genome evolution. Complexity. 1999;5(1):53–66.