Evolutionary escape from local fitness peaks through inversion mutations
Leonardo Trujillo, Paul Banse, Guillaume Beslon

To cite this version:
Leonardo Trujillo, Paul Banse, Guillaume Beslon. Evolutionary escape from local fitness peaks through inversion mutations. 2021, pp.1-1. hal-03426022

HAL Id: hal-03426022
https://inria.hal.science/hal-03426022
Submitted on 11 Nov 2021

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.
Evolutionary escape from local fitness peaks through inversion mutations

Leonardo Trujillo, Paul Banse & Guillaume Beslon

Inria Grenoble - Rhône-Alpes / Inria Sophia Antipolis - Méditerranée, 2004 route des Lucioles, 06902 Sophia Antipolis, France

Abstract

Adaptive molecular evolution is often conceptualized as adaptive walks on rugged fitness landscapes that are driven by changes in the genome (mutations).

The problem

Once a local fitness peak is reached there are no more selectively accessible pathways by the mechanism of single nucleotides (point) mutations. Then, the evolutionary dynamics get stuck in a suboptimal peak, hindering any possibility to reach other higher peaks.

The model

We simulate artificial molecular evolution of digital organisms (replicates), each of which contains a single piece of DNA. We engineer a computational method to carton the double-stranded structure of DNA, and simulate inversion-like mutations: a permutation of the complimentary strands, which are then exchanged with the main strand within the segment where the inversion occurs. We model digital genotypes made up of a binary alphabet [0, 1]. In analogy with DNA primary sequence, the bits represent (artificial) binary nucleotides. We also assume that genomes are arranged in circular strings with constant number of base-pairs, i.e., in that sense, our model does not consider non-coding regions, mimicking the molecular evolution of some viruses and mitochondrial DNA with compact genomes and closed double-stranded DNA circles. It is important to emphasize that our computational model simulates inversion-like mutations. We do not include recombination, so we are modelling asexual replication.

Model of molecular evolution

Adaptive molecular evolution is often conceptualised as adaptive walks on rugged fitness landscapes that are driven by changes in the genome (mutations).

The problem

Once a local fitness peak is reached there are no more selectively accessible pathways by the mechanism of single nucleotides (point) mutations. Then, the evolutionary dynamics get stuck in a suboptimal peak, hindering any possibility to reach other higher peaks.

The model

We simulate artificial molecular evolution of digital organisms (replicates), each of which contains a single piece of DNA. We engineer a computational method to carton the double-stranded structure of DNA, and simulate inversion-like mutations: a permutation of the complimentary strands, which are then exchanged with the main strand within the segment where the inversion occurs. We model digital genotypes made up of a binary alphabet [0, 1]. In analogy with DNA primary sequence, the bits represent (artificial) binary nucleotides. We also assume that genomes are arranged in circular strings with constant number of base-pairs, i.e., in that sense, our model does not consider non-coding regions, mimicking the molecular evolution of some viruses and mitochondrial DNA with compact genomes and closed double-stranded DNA circles. It is important to emphasize that our computational model simulates inversion-like mutations. We do not include recombination, so we are modelling asexual replication.

Model of molecular evolution

Adaptive molecular evolution is often conceptualised as adaptive walks on rugged fitness landscapes that are driven by changes in the genome (mutations).

Enumeration of accessible mutants

The enumerated accessible mutants for inversion mutations form a set of integers: 1, 2, 3, 5, 7, 8, 9, 10, 11, 12, 20, 21, 23, 24, 30, 32, 33, 34, 48, 52, 53, 54, 100. Such set is isomorphic to the hypercube H(10, 2), i.e. the hypercubes with 10 dimensions in which each vertex corresponds to a binary string of 10 bits. The relationship between each vertex and its adjacent vertices is defined by Hamming distance, which is the number of positions at which the corresponding bits are different. The vertices where the Hamming distance is equal to 1 correspond to adjacent vertices, while vertices where the Hamming distance is equal to 2 correspond to random vertices.

Enumerated accessible mutants

Example of accessible mutants (genotypes with N=4)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Accessibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>0</td>
</tr>
<tr>
<td>0001</td>
<td>1</td>
</tr>
<tr>
<td>0010</td>
<td>2</td>
</tr>
<tr>
<td>0011</td>
<td>3</td>
</tr>
<tr>
<td>0100</td>
<td>5</td>
</tr>
<tr>
<td>0101</td>
<td>6</td>
</tr>
<tr>
<td>0110</td>
<td>7</td>
</tr>
<tr>
<td>0111</td>
<td>8</td>
</tr>
<tr>
<td>1000</td>
<td>9</td>
</tr>
<tr>
<td>1001</td>
<td>10</td>
</tr>
<tr>
<td>1010</td>
<td>11</td>
</tr>
<tr>
<td>1011</td>
<td>12</td>
</tr>
<tr>
<td>1100</td>
<td>13</td>
</tr>
</tbody>
</table>

The model

We simulate artificial molecular evolution of digital organisms (replicates), each of which contains a single piece of DNA. We engineer a computational method to carton the double-stranded structure of DNA, and simulate inversion-like mutations: a permutation of the complimentary strands, which are then exchanged with the main strand within the segment where the inversion occurs. We model digital genotypes made up of a binary alphabet [0, 1]. In analogy with DNA primary sequence, the bits represent (artificial) binary nucleotides. We also assume that genomes are arranged in circular strings with constant number of base-pairs, i.e., in that sense, our model does not consider non-coding regions, mimicking the molecular evolution of some viruses and mitochondrial DNA with compact genomes and closed double-stranded DNA circles. It is important to emphasize that our computational model simulates inversion-like mutations. We do not include recombination, so we are modelling asexual replication.

Algorithm 2: AdaptiveWalk (x, p, N, tmax)

input: x ∈ {0, 1}^N, p ∈ (0, 1)
1: t 0
2: f 0 = f(x)
3: ... 0
15: else
16: f_{t+1} = f_t
17: end if
18: t t +1
19: until t = tmax
output: y ∈ {0, 1}^N, f(y) ∈ [0, 1)

Example of accessible mutants (genotypes with N=4)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Accessibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>0</td>
</tr>
<tr>
<td>0001</td>
<td>1</td>
</tr>
<tr>
<td>0010</td>
<td>2</td>
</tr>
<tr>
<td>0011</td>
<td>3</td>
</tr>
<tr>
<td>0100</td>
<td>5</td>
</tr>
<tr>
<td>0101</td>
<td>6</td>
</tr>
<tr>
<td>0110</td>
<td>7</td>
</tr>
<tr>
<td>0111</td>
<td>8</td>
</tr>
<tr>
<td>1000</td>
<td>9</td>
</tr>
<tr>
<td>1001</td>
<td>10</td>
</tr>
<tr>
<td>1010</td>
<td>11</td>
</tr>
<tr>
<td>1011</td>
<td>12</td>
</tr>
<tr>
<td>1100</td>
<td>13</td>
</tr>
</tbody>
</table>

Appendix

This work provided computational evidence that once an adaptive walk reaches a local fitness peak, then inversion mutations can redirect the evolutionary process towards genotypes with higher fitness values. The way our model was conceived, also proves that escape from a local peak of fitness can occur in constant environments without contingencies. Our model for inversion mutations not only elucidated an escape mechanism, but have also made it possible to uncover interesting aspects about the combinatorics of inversion mutations. We have been able to contrast the isomorphism of mutational graphs for point mutations and their isomorphism with the accessibility networks in layered representation. Node colours correspond to fitness values, increasing from left to right. The shortest paths from lowest to highest fitness genotypes are highlighted in red. The top row shows the cases for point mutations, while the bottom row corresponds to inversion mutations. The landscapes ruggedness goes from single peaks (N=0), intermediate ruggedness (K=1, K=2, and the full rugged case (K=3).

Future works

This work provided computational evidence that once an adaptive walk reaches a local fitness peak, then inversion mutations can redirect the evolutionary process towards genotypes with higher fitness values. The way our model was conceived, also proves that escape from a local peak of fitness can occur in constant environments without contingencies. Our model for inversion mutations not only elucidated an escape mechanism, but have also made it possible to uncover interesting aspects about the combinatorics of inversion mutations. We have been able to contrast the isomorphism of mutational graphs for point mutations and their isomorphism with the genotypes space. Finally, it was possible to verify that for a very simple model of population genetics in the strong weak mutation selection limit—with rugged landscapes simulated with the well known NK model—our adaptive walks simulations showed that when the evolutionary dynamics is driven by inversions, higher fitness values can be reached.

Bibliography