

Studying Insertion Sequences through the artificial life software AEVOL

Juliette LUISELLI

INRIA

Team Beagle

Under the supervision of Guillaume BESLON

5th June to 23rd July 2019

IS in Biology

Insertion Sequences dynamics

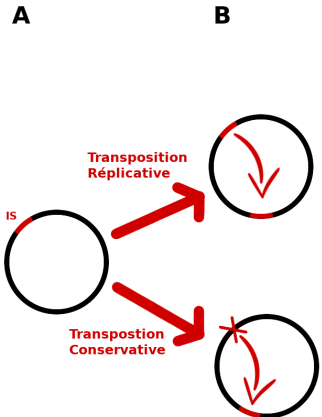


Figure: Insertion sequences dynamic.
(A) Time 0 (B) One time step.

Insertion Sequences dynamics

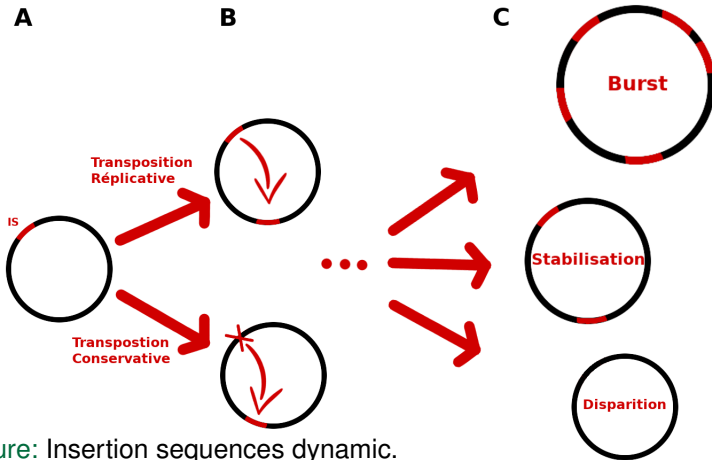


Figure: Insertion sequences dynamic.

(A) Time 0 (B) One time step. (C) Possible long term dynamics

How do IS maintain themselves in the genome ?

Several Hypothesis:

- Sometimes beneficial (changing environment)
- Transpose fast enough
- Domestication

How do IS maintain themselves in the genome ? - Domestication

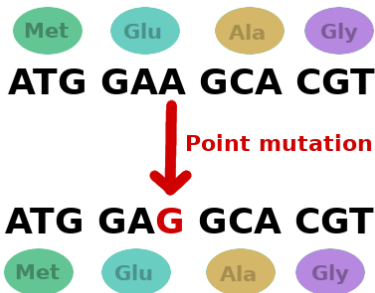


Figure: Proteic and DNA sequences before and after a silent mutation

Implementing IS into AEVOL

Overview of the artificial life software AEVOL

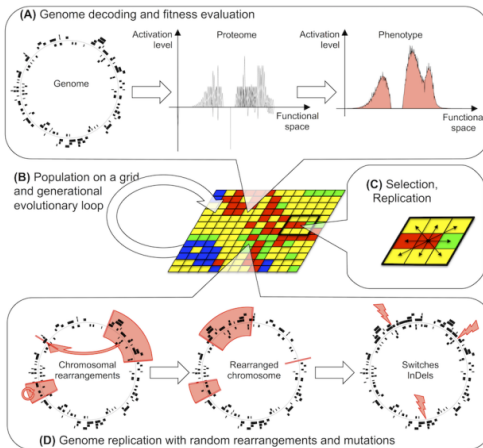


Figure: The Aevol model.(from LIARD2018^[1])

Algorithm to manage IS

- 4 bases model for AEVOL → genetic code redundancy:
 $0 \equiv 2$ and $1 \equiv 3$ so that the rest of the software is unchanged
- Managing IS :
 - Stored as an array of two lists
 - Update lists at each mutation
 - Undergo a specific mutation for transposition
- Test the model for biases

Parameters calibration

Exploring the parameters space

Many parameters to be fitted:

- Mutation rate:

10^{-5} , $5 \cdot 10^{-6}$, 10^{-6} mutation/base/generation

- Transposition rate:

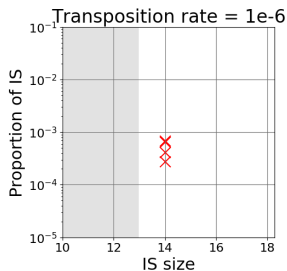
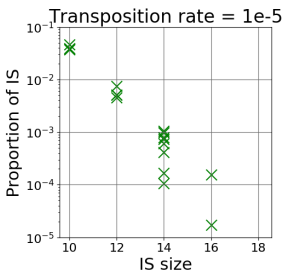
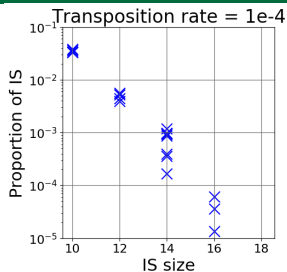
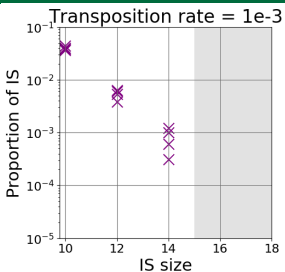
10^{-3} , 10^{-4} , 10^{-5} , 10^{-6} transposition/IS/generation

- IS size:

10, 12, 14, 16, 18 base pairs (bp)
(Hamming distance from consensus ≤ 4)

- Target sequence size: 5

Calibration results - Proportion of IS



Chosen parameters

Mutation rate 5.10^{-6}

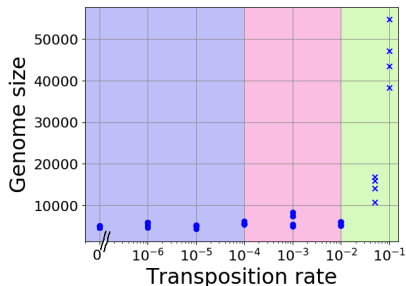
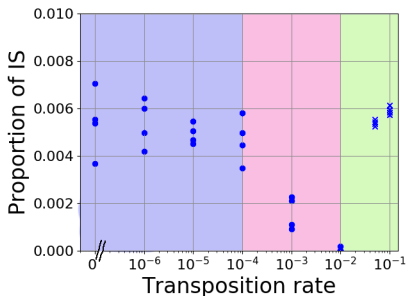
Size: 12

Transposition rate: No significant effect?

⇒ Further exploration for the transposition rate, with fixed size and mutation rates.

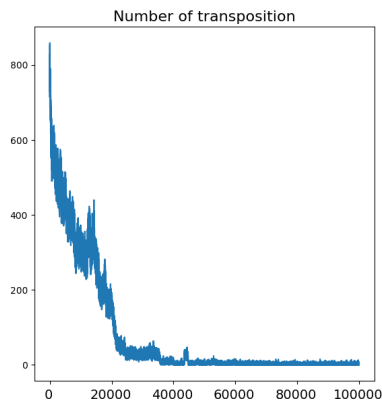
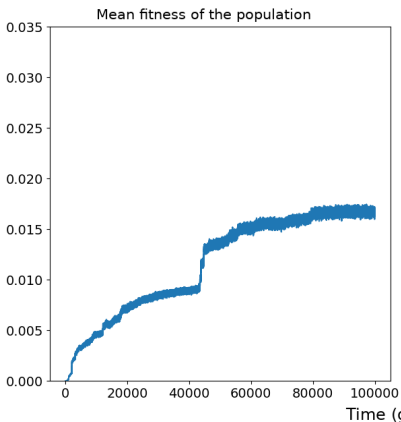
Insights into IS dynamics

IS dynamics as a function of transposition rate

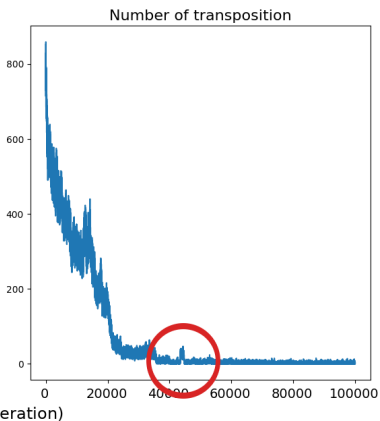
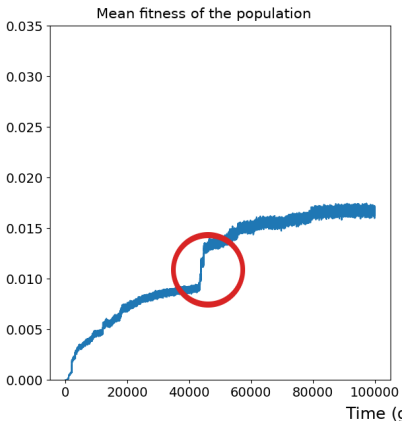


⇒ **3 main dynamics observed**

IS bursts while they are repressed?

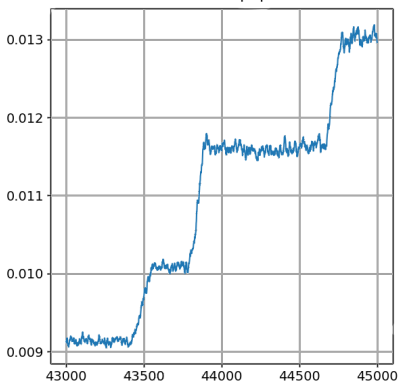


IS bursts while they are repressed?

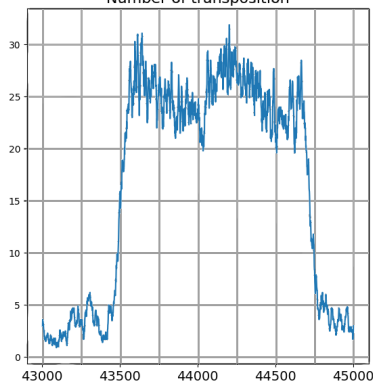


IS bursts while they are repressed? - zoom

Mean fitness of the population



Number of transposition

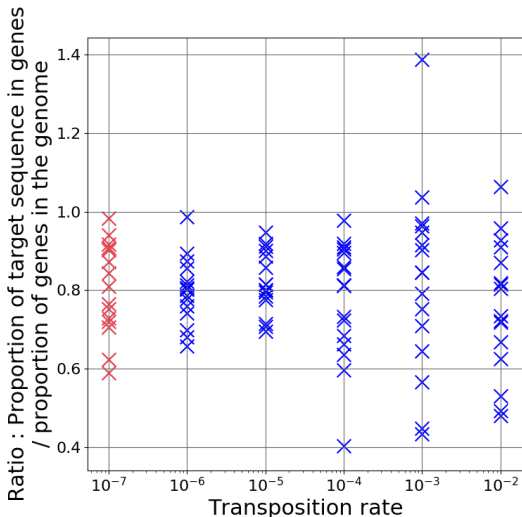


Proportion of target sequences in the coding genome

Hypothesis: Counter selection against target sequences in coding genome

Proportion of target sequences in the coding genome

Hypothesis: Counter selection against target sequences in coding genome

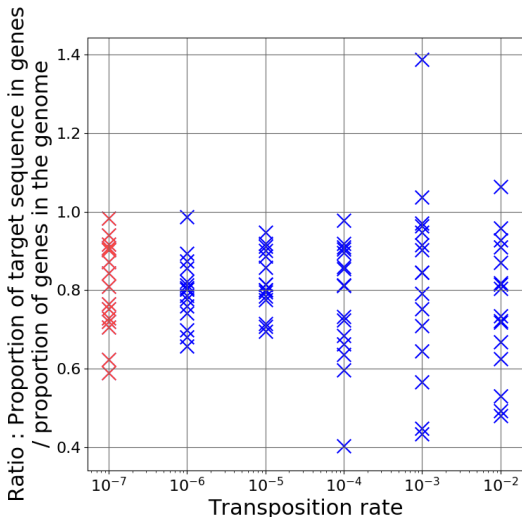


Counter selection of target sequences ?

Proportion of target sequences in the coding genome

Hypothesis: Counter selection against target sequences in coding genome

Conclusion: ?



Beneficial effect in changing environment ?

Are IS beneficial in a context of changing environment

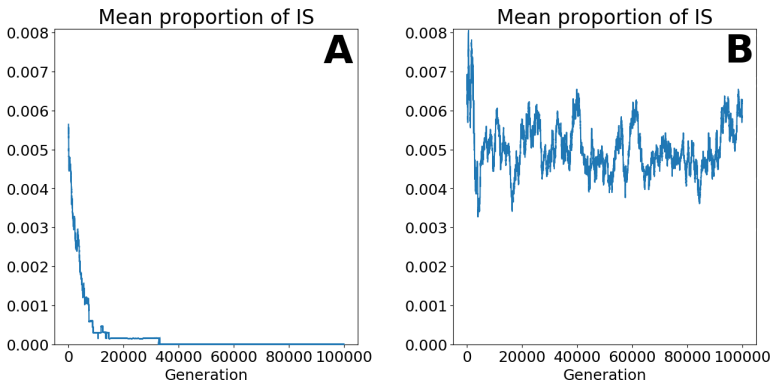


Figure: Proportion of IS in a stable (A) or changing environment (B) with a constant transposition rate of 10^{-2} transposition/IS/generation

Conclusion & Perspectives

Conclusion

- Prototype for IS successfully developed:
 - Model for IS developed
 - Algorithm implemented into AEVOL
 - Main values calibrated
- Three kind of dynamics observed : little to no activity, strong repression of IS or exploding genomes

Perspectives

Many fields to be explored:

- Constant **environmental changes**
- Follow the **lineages** in the population to better understand the link between the fitness and transposition rises.
- Combining with **recombination based on homologous sequences?**

Thanks

I would like to thank my supervisor, Guillaume Beslon, and my coworkers on Aevol: Jonathan Rouzaud-Cornabas, Paul Banse & Laurent Turpin.

Special thanks to all the interns for the wonderful discussions and breaks!



Bibliography I



V. LIARD, D. PARSONS, J. ROUZAUD-CORNABAS & G. BESLON: *The Complexity Ratchet: Stronger than selection, weaker than robustness.*

In *The 2018 Conference on Artificial Life*, pp. 250–257. MIT Press, Tokyo, Japan, 2018.

[doi:10.1162/isal_a_00051](https://doi.org/10.1162/isal_a_00051).



D. PARSONS, C. KNIBBE & G. BESLON: *Aevol : un modèle individu-centré pour l'étude de la structuration des génomes.*

p. 8.

An example : Computing the fitness from the phenotype

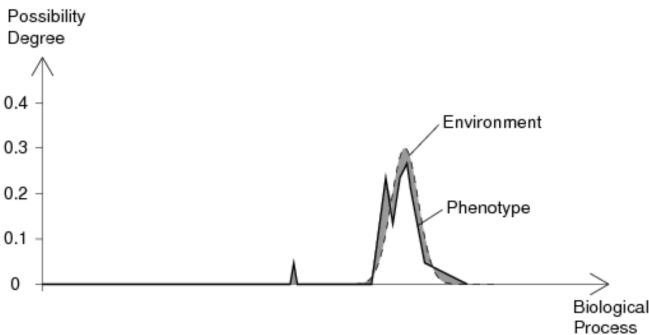


Figure: The phenotype of an individual (from PARSONS2010^[2])

Preliminary results - Effect on fitness

