Modeling horizontal gene transfer and persistance of plasmids : scaling limit, stationarity and ancestral lineages

Chaire MMB - Aussois 2026

École Polytechnique - CMAP
Inria - Merge

Mateo Deangeli Bravo - PhD Student

under the supervision of Sylvie Méléard and Viet-Chi Tran

Horizontal Transfer in a nutshell

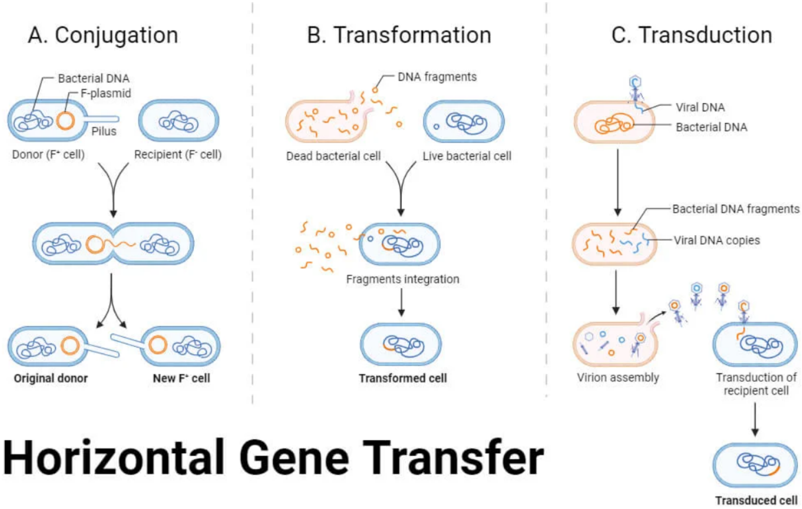
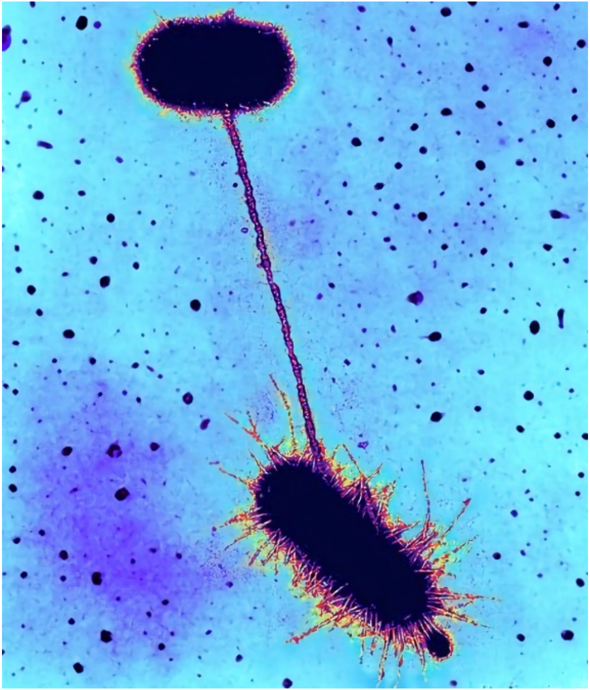


Figure: Three types of HT



Biological framework

Plasmid :

- Extrachromosomal DNA that can be transferred.
- Imposes a fitness-cost (in particular when not needed).

Why do we care ?

- In genetics, rarely modeled, compared to vertical transfer (= via reproduction)
- Antibiotics
- A significant component of natural selection

What do we want ? - Mathematical Framework

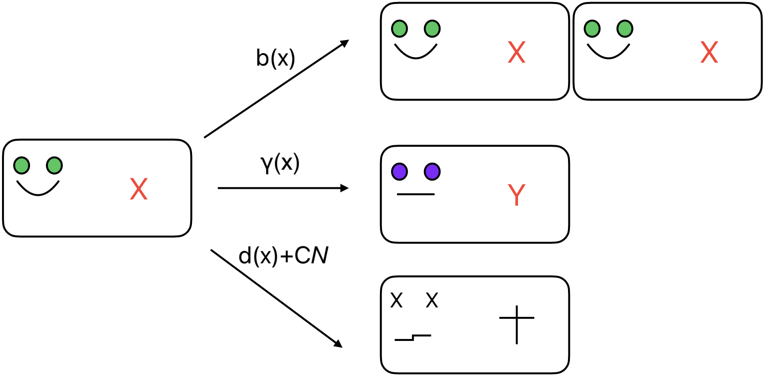
We want a random (Markov) process such that

- It describes a population of bacteria.
- Each bacterium is characterized by it's **amount of plasmid DNA**.
- It takes into account birth, death, competition, variation and transfer.
- **Neutral environnement : Plasmid = reduced fitness.**

(Could describe a structured population with a mimicking deleterious trait dynamic - tobacco consumption for example)

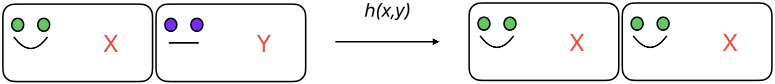
The model

- Birth
- Mutation
- Death (individual mortality and competition) - N =population size.



The model

- Horizontal transfer



- KEEP IN MIND ! Transfer is unilateral : $h(x, y) > 0$ if and only if $x > y$.
- $x = 0$ is fixed as the optimal trait (ie $\sup_{x \geq 0}(b(x) - d(x)) = b(0) - d(0) > 0$) and $b - d$ is decreasing.

Illustrations

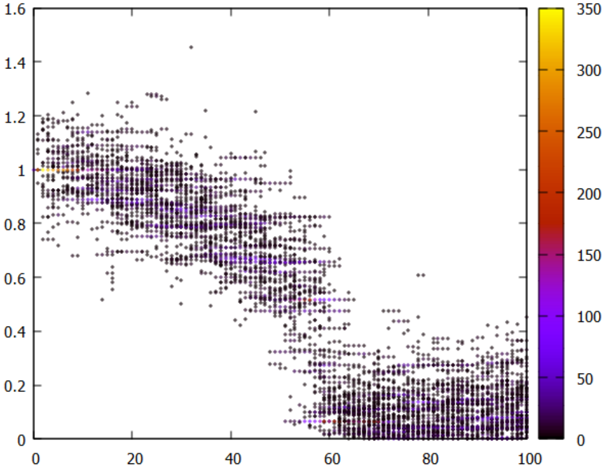


Figure: No transfer / Time in x-axis / Trait in y-axis

Illustrations

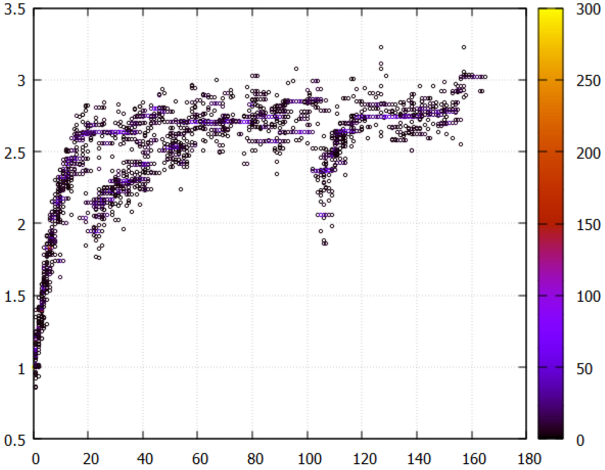


Figure: Transfer leads to fast extinction

Illustrations

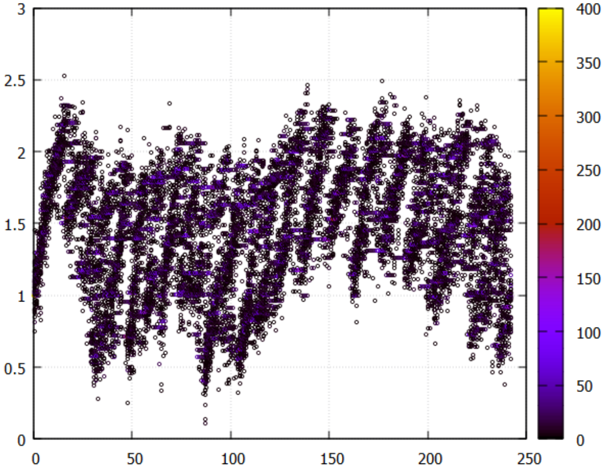


Figure: Transfer leads to cycles

Scaling limit

We set

$$Z_t^K = \frac{1}{K} \sum_{i \in V_t^K} \delta_{X_t^i} = \frac{\text{Population measure}}{\text{Scaling parameter}}$$

Under classical (and new !) assumptions, $Z^K \underset{K\to\infty}{\rightarrow} \xi$

$$\partial_t \xi_t(x) = (r(x) - C\rho_t)\, \xi_t(x) + \int_0^\infty \gamma(y)(\xi_t(y) - \xi_t(x)) m(y,x) dy$$

$$+ \frac{\tau}{\beta + \mu\rho_t} \left(\int_0^x \xi_t(y) dy - \int_x^\infty \xi_t(y) dy \right) \xi_t(x)$$

where

$$\rho_t = \int_{\mathbb{R}_+} \xi_t(y) dy, \qquad r(x) = b(x) - d(x).$$

Includes unbounded death, heavy tail natural variation.

Checkpoint

At this point, we have

- a random microscopic description Z
- a deterministic macroscopic approximation ξ

We want to answer the following questions

- Can bacteria survive despite the cost of plasmids ?
- If they survive, from which type of individuals do they descend ?

Numerical resolution

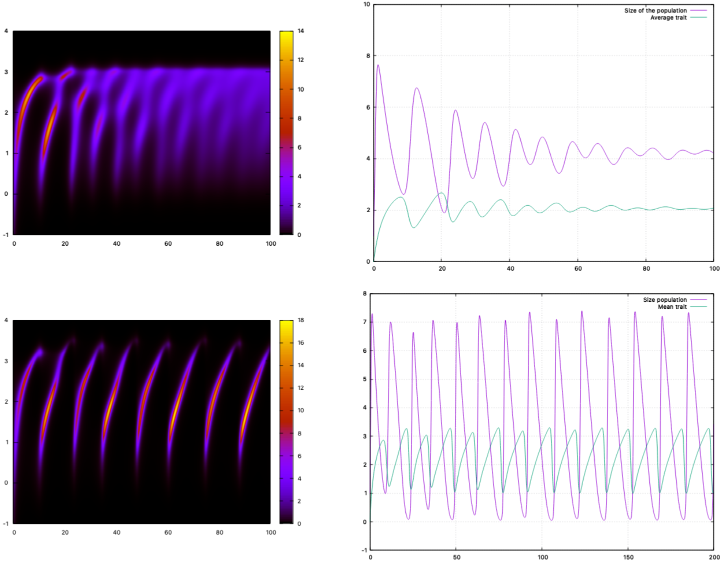


Figure: Simulation of the scaling-limit density : convergence and cycles

Existence of a stationary solution

We look at the stationary equation **without transfer**

$$r(x)N(x) + \int_0^\infty \gamma(y)(N(y) - N(x))m(y,x)dy = \lambda_0 N(x), \ x \in \mathbb{R}_+$$

$$N(x) > 0, \int_0^\infty N(x)dx = 1, \ \lambda_0 = \int_0^\infty r(x)N(x)dx > 0,$$

($\lambda_0 =$ fitness (or effective size) of the population)

and **with transfer**

$$r(x)u(x) + \int_0^\infty \gamma(y)(u(y) - u(x))m(y,x)dy + \mathcal{H}(x) = C\rho u(x)$$

$$\mathcal{H}(x) = u(x)\frac{\tau\rho}{\beta + \mu\rho}\int_0^\infty \operatorname{sgn}(x-y)u(y)dy$$

$$u(x) > 0, \int_0^\infty u(x)dx = 1, \rho = \frac{1}{C}\int_0^\infty r(x)u(x)dx > 0$$

($\rho =$ fitness (or effective size) of the population with transfer)

Existence of a stationary solution

Following *[Garriz, Leculier, Mirrahimi - 24]* approach :

Theorem

We have at least one non zero solution to the equation with transfer :

- *If* $\beta > 0, \mu \geq 0$ *and* $\lambda_0 > 0$*. In this case, the transfer fades when few individuals.*
- *If* $\beta = 0, \mu > 0$ *and* $\lambda_0 > \tau/\mu$*. In this case, the transfer is constant.*

The solution is continuous, bounded and L^1*.*

- $\lambda_0 =$ population global fitness without transfer.
- $\tau =$ frequency of transfer events.
- $\beta =$ mean time of encounter between two individuals.
- $\mu =$ mean time of transfer.

First Conclusion

- In the macroscopic approximation, we can have survival and stationarity even with transfer !
- An approximated condition : if $r(x) = r_0 - \alpha x^2$ and Gaussian mutations, then

$$\lambda_0 \underset{\sigma\sim 0}{\sim} r_0 - \sigma\sqrt{\frac{\alpha\gamma}{2}} > \frac{\tau}{\mu}\mathbb{1}_{\{\beta=0,\mu>0\}}$$

- *maximal reproduction rate* $>$ *combination of mutations and selection* $+$ *transfer intensity when immediate encounter of individuals*
- Uniqueness ???

Getting back to the microscopic description - Ancestral Lineages - Naive approach

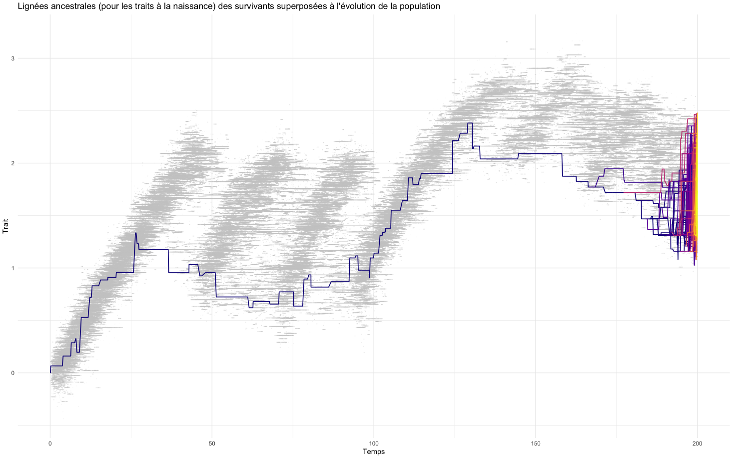


Figure: Ancestral lineages (in the random population process) - obtained with IBMPopSim Package

Obtaining the Feynman-Kac formula

We write the process

$$Z_t = \sum_{i \in V_t} \delta_{X_t^i}$$

X the random process giving the law of a simple lineage and $r(x) = b(x) - d(x)$.

$$\mathbb{E}\left[\sum_{i \in V_t} \varphi(X_t^i) \middle| Z_0 = \delta_x\right] = \mathbb{E}\left[e^{\int_0^t r(X_u)du} \varphi(X_t) | X_0 = x\right]$$

$$\mathbb{E}\left[\text{Whole Population}\right] = \mathbb{E}\left[\text{Size Bias} \times \text{Simple Lineage}\right]$$

where

$$V_t = \{\text{individuals alive a time } t\}$$

The evaluation of the population process can be given by a unique trajectory biased by size.

Nagasawa Theorem - Let's reverse time !

-Let $(t, X_t)_{t\in[0,T]}$ be a $[0, T] \times E$-valued Markov process, and P the corresponding semigroup.
-Let μ be an initial measure, for example $\mu = \delta_0 \otimes \delta_x$.
-Let η be the measure defined by

$$\eta(f) = \int \left(\int_0^\infty P_t f(s, x) dt \right) \mu(ds, dx)$$

-X^R the reversed-time process (and P^R its semigroup) defined by :

$$\forall t \in [0, T],\ X_t^R \overset{d}{=} X_{T-t}$$

Then we have the duality relation

$$\int P_t f(s, x) g(s, x) \eta(ds, dx) = \int f(s, x) P_t^R g(s, x) \eta(ds, dx)$$

Generator of the time-reversed spine

With the previous duality and Nagasawa's Theorem, we can find the generator of the time reversed spine process, given by

$$\begin{aligned} L^R\psi_s(x) &= -\partial_s\psi_s(x) \\ &\quad + \int_{\mathbb{R}_+} (\psi_s(z) - \psi_s(x)) \frac{\xi_s(z)}{\xi_s(x)} \gamma(z) m(z,x) dz \\ &\quad + \int_{\mathbb{R}_+} (\psi_s(y) - \psi_s(x)) \frac{\xi_s(y)}{\xi_s(x)} \xi_s(x) h(x,y,\xi_s) dy \end{aligned}$$

It gives the dynamic of the lineage, backward in time, when sampling an individual at time T.

Simulating the AL - Is that all ?

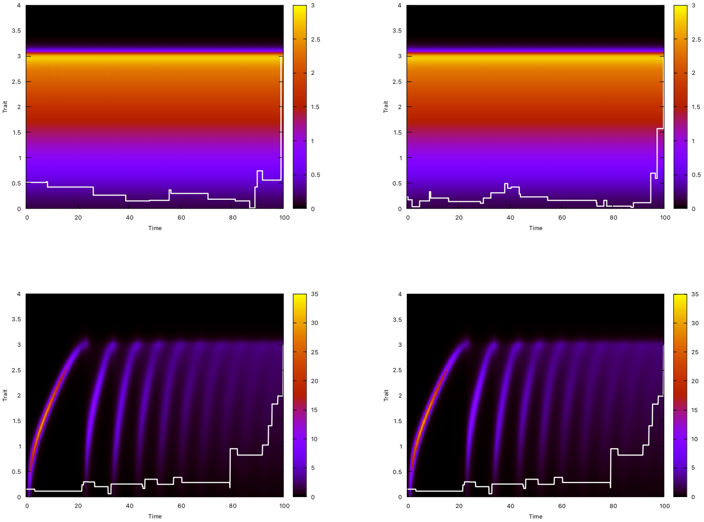


Figure: Simulation of the ancestral lineage process in the stationary and non-stationary cases, with gaussian mutations

Simulating the AL - Skyscraper mode

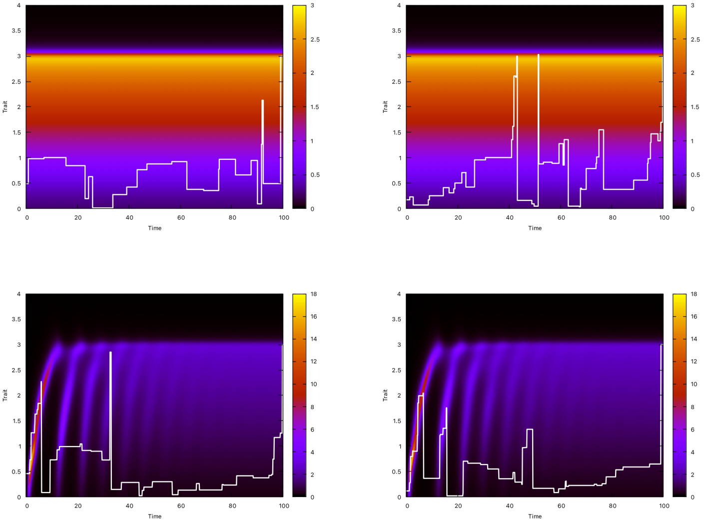


Figure: Simulation of the ancestral lineage process in the stationary and non-stationary cases, with fat-tail mutations

Final Conclusion

Mathematical results :

- Existence of a strong solution to the integro-differential equation and a (strong) non null stationary solution.
- A general method for the time-reversal of the spinal process (even with transport and diffusion terms).
- Exact efficient simulation of the Ancestral Lineages (in large-population limit).

Answers (given by the model) to our questions :

- In large population regime, possible survival and stationarity with horizontal transfer.
- Despite the survival, the survivors come from the lineages that were least exposed to transfer (when variations are not fat tail).

The End

Thank you for your attention and thank you to the organisers !
Any suggestions regarding the stationary solution are welcome !