EXTENDED MATHEMATICAL METHODS: STOCHASTIC POPULATION DYNAMICS MODEL

Plasmids promote antimicrobial resistance through insertion-sequence–mediated gene inactivation

This document provides the full mathematical description of the stochastic population dynamics model used in the study. Details the biological processes included in the simulations, the corresponding reaction propensities, and the parameterization procedures used to match experimental measurements. All numerical simulations were implemented in Python using a Google Colab notebook, which includes the complete code for the stochastic model, parameter initialization, and analysis routines. The complete and reproducible Colab notebook is publicly available in the accompanying GitHub repository.

1. Overview of the modeling framework

To capture the inherent randomness of mutation, plasmid loss, and horizontal transfer in small bacterial populations, we implemented a stochastic model based on the Gillespie algorithm. Unlike deterministic ODE frameworks, this formulation resolves events one reaction at a time and therefore captures low-probability evolutionary changes such as insertion-sequence (IS) transpositions, segregational loss of plasmids, and conjugative transfer between species.

The model integrates three hierarchical layers of biological processes:

(i) **Cellular processes:** resource-limited growth and antibiotic-induced mortality, both of which depend on local environmental conditions;
(ii) **Plasmid dynamics:** stochastic plasmid segregation at cell division and mass–action conjugative transfer between plasmid-bearing donors and plasmid-free recipients;
(iii) **Evolutionary events:** accumulation of single-nucleotide polymorphisms (SNPs) and IS-mediated transpositions, which move cells between mutation classes and gradually diversify each lineage.

All processes are represented as discrete stochastic reactions. Each reaction has a propensity defined by the current state of the system, including strain-specific parameters (maximum growth rate μ_i, maximum death rate δ_i, conjugation coefficient γ_{ij}, segregation probability σ_i, SNP mutation rate $\mu_{m,i}$, and IS transposition rate τ_i), the instantaneous resource level R, the antibiotic concentration A, and the size of each subpopulation. Reaction propensities are recalculated at every simulation step, ensuring that growth, death, transfer, and mutational events reflect the current ecological and evolutionary state of the population.

2. State variables and Monod kinetics

Each bacterial species i is represented by two subpopulations: plasmid-free ($B_{i,0}$) and plasmid-bearing ($B_{i,p}$). In the simulation, each of these subpopulations is further divided into classes indexed by the number of accumulated SNP and IS mutations. These classes are tracked by two counters (m,t), where m is the SNP level and t is the IS transposition level. The model therefore maintains variables $B_{i,0}(m,t)$ and $B_{i,p}(m,t)$. (0)

Resource-limited growth follows Monod kinetics,

$$\mu_i(R) = \mu_{i,\max} \frac{R}{K_{S,i} + R}, \tag{S2}$$

where $\mu_{i,\max}$ and $K_{S,i}$ are strain-specific parameters, where the effective birth rate for each class is updated at every step as $\mu_{i,\max} R/(K_{S,i} + R)$.

Antibiotic-mediated mortality is modeled by a saturating function,

$$\delta_i(A) = \delta_{i,\max} \frac{A}{K_{A,i} + A}, \tag{S3}$$

where $\delta_{i,\max}$ and $K_{A,i}$ are obtained from the experimental dose–response fits. In our implementation, the effective death rate applied to each (m,t) class is recalculated as $\delta_{i,\max} A/(K_{A,i} + A)$.

Resource consumption is proportional to the number of birth events. Each division reduces the available resource R by a fixed amount c, consistent with the per-division consumption rate used in the Python code.

3. Reaction system and propensities

Each biological process is represented as an elementary reaction with an associated propensity a_k. In the implementation, propensities are recalculated at every simulation step for every strain, plasmid state, and mutation class (m,t), using the current values of R, A, and the corresponding subpopulation size. For a population $B_{i,s}(m,t)$ of strain i in state $s \in \{0, p\}$:

$$\text{Birth:} \quad a_{\text{birth},i,s}(m,t) = \mu_i(R)\, B_{i,s}(m,t), \tag{S4}$$

$$\text{Death:} \quad a_{\text{death},i,s}(m,t) = \delta_i(A)\, B_{i,s}(m,t), \tag{S5}$$

$$\text{Segregation (plasmid loss):} \quad a_{\text{seg},i}(m,t) = \sigma_i B_{i,p}(m,t), \tag{S6}$$

$$\text{Mutation (SNP):} \quad a_{\text{mut},i,s}(m,t) = \mu_{m,i} B_{i,s}(m,t), \tag{S7}$$

$$\text{Transposition (IS):} \quad a_{\text{IS},i,s}(m,t) = \tau_i B_{i,s}(m,t). \tag{S8}$$

At each simulation step, propensities are evaluated independently for every (m,t) class, and the number of events for each reaction is sampled from a Poisson distribution with mean $a_k \Delta t$. This ensures that births, deaths, segregation events, mutations, and IS insertions scale with the current abundance of each class.

Plasmid conjugation reactions, $a_{\text{conj},ij}$ (S9), are modeled as mass-action interactions between plasmid-bearing donors and plasmid-free recipients, as described below. Each conjugation event converts a plasmid-free cell of strain j in class (m',t') into a plasmid-bearing cell of the same class, preserving total cell number.

All reaction events update the state vector of strain i

$$\mathbf{X}(t) = \{B_{i,0}(m,t), B_{i,p}(m,t), R, A\},$$

and these updates are applied immediately after each stochastic time step.

To efficiently simulate these reactions, we use a small fixed time step Δt and sample the number of events of each reaction from independent Poisson distributions with mean $a_k \Delta t$. This corresponds to a standard tau-leaping approximation of the Gillespie algorithm. Because Δt is chosen to be much smaller than the characteristic timescales of growth, death, mutation, and conjugation, this approach preserves the statistical properties of the full Gillespie direct method while remaining computationally tractable for multispecies communities.

4. Plasmid conjugation and segregation dynamics

Plasmid loss occurs during cell division with probability σ_i for strain i. This event converts a cell bearing a plasmid $B_{i,p}(m,t)$ into a cell without plasmid $B_{i,0}(m,t)$ of the same mutation class and is applied independently to every (m,t) class. The expected number of segregation events in each time step is proportional to $\sigma_i B_{i,p}(m,t)$, ensuring that plasmid loss scales with the abundance of plasmid-bearing cells.

Conjugation is represented as a mass-action interaction between plasmid-bearing donors and plasmid-free recipients. A conjugation event between the donor strain i and the recipient strain j converts a single recipient cell into a plasmid-bearing cell while leaving the donor unchanged:

$$B_{i,p} + B_{j,0} \xrightarrow{\gamma_{ij}} B_{i,p} + B_{j,p}.$$

This process is evaluated for all pairs of donor and recipient mutation classes (m,t) and (m',t'). The expected number of transfer events is proportional to

$$\gamma_{ij}\, \varepsilon_{ij}\, B_{i,p}(m,t)\, B_{j,0}(m',t'),$$

and each event decreases $B_{j,0}(m',t')$ by one and increases $B_{j,p}(m',t')$ by one. This formulation preserves the total population size while allowing the horizontal plasmid to spread across strains under the constraints imposed by the conjugation network.

5. Multispecies extension

The model generalizes to M species, each with a strain-specific parameter set

$$\Theta_i = \{\mu_{i,\max}, K_{S,i}, \delta_{i,\max}, K_{A,i}, \gamma_{ij}, \sigma_i, \tau_i, \mu_{m,i}\},$$

where γ_{ij} denotes the conjugation coefficient from donor i to recipient j. Species differ in their growth and mortality parameters, their rates of mutation and transposition, and in the cost of plasmid carriage, which is reflected in a reduced maximum growth rate for plasmid-bearing cells ($\mu_{i,p} < \mu_{i,0}$).

For each species, the population is divided into plasmid-free and plasmid-bearing groups, and each of these is further subdivided into mutation classes (m,t) as described above. The full state of the system at time t is therefore

$$\mathbf{X}(t) = [B_{1,0}(m,t), B_{1,p}(m,t), \ldots, B_{M,0}(m,t), B_{M,p}(m,t), R], \tag{S10}$$

with sums over all mutation levels implied.

This formulation allows all species to grow, die, mutate, undergo transposition, lose plasmids through segregation, and exchange plasmids through conjugation. All processes are evaluated for every species and for every allowed donor–recipient pair, ensuring that the full ecological and evolutionary interactions within the community are captured.

6. Network of horizontal gene transfer

Horizontal transfer between donors i and recipients j is controlled by the adjacency matrix ε, which specifies whether plasmid transfer between each ordered pair is allowed:

$$\varepsilon_{ij} = \begin{cases} 1, & \text{if conjugation } i \to j \text{ is permitted,} \\ 0, & \text{otherwise.} \end{cases}$$

The conjugation propensity between donor i and recipient j is

$$a_{\text{conj},ij} = \gamma_{ij}\,\varepsilon_{ij} \left(\sum_{m,t} B_{i,p}(m,t) \right) \left(\sum_{m',t'} B_{j,0}(m',t') \right),$$

where γ_{ij} is the strain-specific conjugation rate coefficient, $\sum_{m,t} B_{i,p}(m,t)$ is the total number of plasmid-bearing cells of strain i, and $\sum_{m',t'} B_{j,0}(m',t')$ is the total number of plasmid-free cells of strain j. This structural matrix determines which donor–recipient pairs can exchange plasmids, but it does not set the rate at which transfer occurs. Conjugation dynamics are governed by the transfer coefficients γ_{ij}, which scale the frequency of donor–recipient encounters.

In the simulations we consider two related settings. First, to study how transfer efficiency alone affects plasmid persistence, we fix a fully connected network with $\varepsilon_{ij} = 1$ for all $i \neq j$ and vary a common conjugation coefficient over a range of values (10^{-11} to 10^{-8} cell^{-1} h^{-1}). Second, to examine the role of connectivity, we fix the conjugation rate and progressively remove a random subset of links by setting selected ε_{ij} entries to zero. In this way, the overall connectivity of the conjugation network is controlled by the fraction of active links, while the strength of each remaining link is kept constant. Three representative topologies were considered:

- **Fully connected:** $\varepsilon_{ij} = 1$ for all $i \neq j$, allowing transfer between every ordered donor–recipient pair.
- **Sparse:** only a random fraction p of possible links is active, generated by Bernoulli sampling of each ε_{ij}.

7. Calibration of strain-specific model parameters

The stochastic model relies on strain-specific growth, mutation, and antibiotic susceptibility parameters that were obtained directly from experimental measurements of *Klebsiella pneumoniae* strains grown in isolation. Model calibration proceeded in three steps: (i) fitting growth parameters, (ii) estimating mutation and transposition rates, and (iii) inferring antibiotic susceptibility parameters.

Bacterial growth parameters. Strain-specific parameters were estimated from monoculture growth assays (see Table S1). For each strain, the Monod parameters $\mu_{\max}$ and K_S were inferred by fitting a

simplified, single-strain version of the model to optical-density growth curves. This reduced model included resource-limited growth and death, but excluded conjugation, segregation, and mutation processes. Parameter optimization was carried out using the L-BFGS-B algorithm, minimizing the squared deviation between simulated and experimental trajectories (Figure 1).

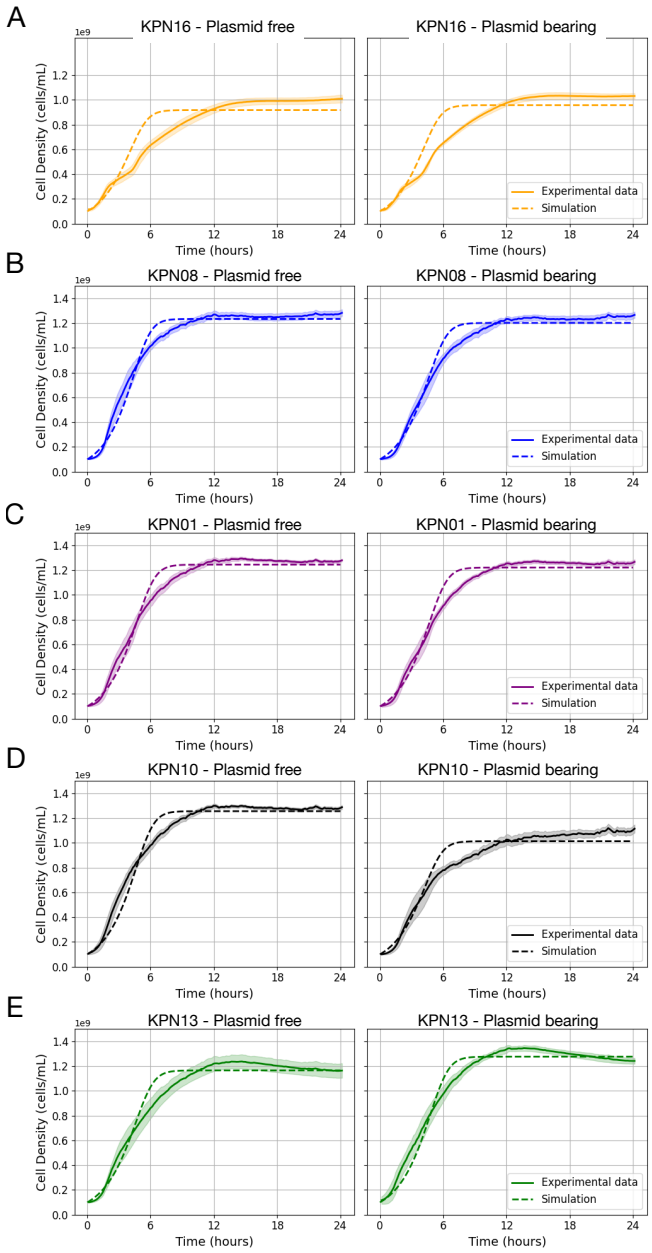

Figure 1. Model parameterization. Comparison between simulations of the stochastic model (dotted lines) and experimental growth data (solid lines) for *K. pneumoniae* strains grown in isolation. Simulations were performed using best-fit Monod parameters estimated independently for each strain. The left column shows plasmid-free populations, and the right column shows plasmid-bearing populations. Panels correspond to different strains: (A) KPN16 (orange), (B) KPN08 (blue), (C) KPN01(purple), (D) KPN10 (black), and (E) KPN13 (green).

Mutation rates. SNP and IS transposition rates were estimated from fluctuation assays followed by whole-genome sequencing. Independent replicate cultures were initiated from single ancestral cells and grown under nonselective conditions to stationary phase. For each lineage, the total number of *de novo* SNPs and IS insertion events was divided by the chromosomal length L_{chr} to obtain per-base mutation and transposition rates (μ_{SNP} and τ_{IS}). These values were then normalized by the duration of the assay to yield per-base, per-hour rates, which were used as the mutation and transposition parameters in the stochastic simulations. Fluctuation assays were performed for both plasmid-free (B_0)

and plasmid-bearing (B_p) isogenic lineages to determine whether plasmid carriage altered baseline mutation or transposition rates. The resulting parameters $(\mu_{\mathrm{SNP},0}, \mu_{\mathrm{SNP},p}, \tau_{\mathrm{IS},0}, \tau_{\mathrm{IS},p})$ were assigned to the corresponding subpopulations. These mutation and transposition rates were then combined with growth and mortality parameters to obtain the full strain-specific parameter set used in the simulations.

Drug susceptibility parameters. Antibiotic-dependent mortality parameters ($\delta_{\max}$ and K_A) were inferred by fitting dose–response curves to measured IC_{90} values for each strain. All strains were susceptible to colistin, and the fitted K_A values were similar across isolates; a common K_A was therefore applied to all species.

Horizontal transmission. To evaluate how the conjugation rate (ε) influences plasmid persistence and community composition, we simulated multispecies communities across a range of transfer rates from 10^{-11} to 10^{-8} cell^{-1} h^{-1}. This range spans values for which plasmid transfer is low up to values where horizontal transmission can compensate for the fitness cost of plasmid carriage. As ε increases, the simulations show a clear threshold-like transition in community behavior. At very low conjugation rates, plasmids are rapidly lost because segregational losses and the growth disadvantage of plasmid-bearing cells are not balanced by horizontal transfer. Near the threshold, sporadic transfer events are sufficient to maintain plasmids in a subset of strains, leading to mixed communities with unstable plasmid presence. Above this threshold, frequent transfer events allow multiple species to acquire and retain the plasmid, producing stable coexistence of plasmid-bearing lineages despite the associated growth cost.

8. Numerical implementation and serial-transfer protocol

All simulations were implemented in `Python` using standard scientific libraries (`NumPy`, `pandas`, `matplotlib`). The stochastic resource-explicit model was simulated using a fixed–time step tau-leaping scheme with $\Delta t = 0.1$ h, as described above. At each time step, the reactions governing growth, death, mutation, transposition, segregation, and conjugation were sampled from independent Poisson distributions with mean $a_k \Delta t$, and the population state was updated accordingly.

To reproduce the experimental serial-transfer protocol, each *in silico* evolution experiment consisted of a sequence of 24 h growth cycles followed by dilution. During each cycle, the model was integrated for 24 h using the stochastic update rules above. At the end of the cycle, all populations were diluted by a fixed factor d (1:100), and the resource level was reset to its initial concentration R_0. This growth–dilution process was repeated for up to $n_{\mathrm{transfers}} = 60$ consecutive days in simulations exploring long-term plasmid persistence. All code and data necessary to reproduce these simulations are available in the GitHub repository: https://github.com/jorgEVOplasmids/MDR_ISs.

Table 1. Parameter values used in the computer simulations.

Parameter	Meaning / where set	Units	Value
M	Number of species	–	5
T	Simulation time	h	24
Δt	Recording interval (`dt`)	h	0.1
A	Antibiotic concentration	IC_{90} units	0.0, 0.1, 0.2
σ	Plasmid segregation rate	h^{-1}	1×10^{-7}
γ	Conjugation rate sweep	$cell^{-1}\,h^{-1}$	$10^{-11} \ldots 10^{-7}$
c	Resource consumption per division	res. units·$cell^{-1}$	*from fit*
μ_{max}	Birth/growth rate	h^{-1}	*from fit (matrix)*
K_S	Half-saturation (resource)	res. units	*from fit (matrix)*
K_A	Half-saturation (antibiotic)	IC_{90}	*from fit (0/p matrices)*
μ_m	SNP mutation rate	h^{-1}	*from fit (matrix)*
τ	IS transposition rate	h^{-1}	*from fit (matrix)*