Supporting Information

Chia *et al.* 10.1073/pnas.0804962105

SI Text: Details of Simulation Results in Figs. 3 and 4

Simulating the dynamics of the MGE-mediated phase variation follows closely along the lines described in *Minimal Model for Mobile Genetic Element Mediated Phase Variation* in the main text. In particular, the simulations were discrete time, finite population size implementations of master equations, which reproduce Eqs. **1–4** in the infinite population limit. While the more analytically tractable Eqs. **1–4** are more appropriate for describing continuous time, infinite population dynamic, the simulations were carried out with a finite resource limit and discrete cells to more directly represent the discrete interactions acting on individual cells. Individual microbes are represented by a genome containing two loci—an active locus and an inactive locus. The gene in the active locus determines the phenotype of the cell, which is represented as a 0 or 1. Likewise, genes are also represented by a 0 or 1, corresponding to the resulting phenotype when the gene is in the active locus. The population of microbes is represented by an array of genomes where the size of the array represents the carrying capacity or maximum size of the system. Each element of the population array can take one of five states: 00, 01, 10, 11, or "empty." The first four states represent the genes in the two loci genome starting with the gene active locus. The last state indicates that there is no cell occupying this position in the array.

Simulations were conducted by iterating the following steps:

1. Growth. Cells for replication are selected at random. The base rate of replication depends on the lesser of two quantities, either the number of cells, or the difference between the carrying capacity of the system and the number of cells. In any case, half the lesser number determines the base rate of replication. These rates correspond to exponential growth in the small population limit, and logarithmic growth up to the carrying capacity of the system in the large population limit. The base rate is then multiplied by a factor that controls the size of the individual time steps, in this case 0.25. For a non-integer number of replication events, an extra replication event occurs sometimes according to the appropriate probability. In Fig. 4, where the cell types exhibit different attachment rates, this is simulated by using unequal growth rates for each phenotype. In this case, a fraction of the randomly selected slower growers do not replicate according to a set probability determined by the ratio of the attachment rates between the two phenotypes. For equal growth rates, as in Fig. 3, this is unnecessary and all randomly selected cells replicate.
2. Gene transfer. A random gene is selected from the population of microbes. This could represent a direct exchange between two microbial cells. Likewise, this also could be considered an approximation of genetic uptake from a communal gene pool, whose composition approximately reflects the genes within the cellular genomes. The selected gene then is paired with a randomly selected cell. The gene then overwrites the active locus according to the probability q_0 or q_1 or the inactive locus with a probability of $1 - q_0$ or $1 - q_1$, depending on whether the gene is a 0-gene or a 1-gene. The number of these gene transfer events is determined according to the number of cells in the population, multiplied by a rate of gene transfer (chosen to be ¾ for Figs. 3 and 4) and the same factor of 0.25 that indicates the size of each time step in the simulation.

To make the gene transfer rate the only unconstrained parameter of this model, we use two pieces of data. First, we used the growth curve in Fig. 3 to determine the rate of replication. For the experiment in Fig. 4, where we lacked this data, we used a similar overall growth rate as used in Fig. 3, but with a 0-phenotype to 1-phenotype growth rate ratio of 0.85. Anchoring the growth rate to experimental data allows us to frame the gene transfer rate in terms of transfers per replication in the exponential growth phase. To set the other parameters, we choose q_0 and q_1 such that they are consistent with the coexistence ratio of the two phenotypes 0 and 1 for the final experimental time point, i.e., we match the steady state ratio of phenotypes, in Fig. 3. Note that there may be multiple values of q_0 and q_1 that result in the same phenotype coexistence ratio. However, simulation results remain remarkably similar for any choice made matching the steady state constraint.

Last, initial conditions are chosen based on the simulated steady state distribution of genotypes. That is to say, for Fig. 3, where the culture is inoculated by a single phenotype, the ratio of genotypes in the initial population is given by the ratio 00:01. For Fig. 4, where the innoculum contains both 0 and 1 phenotypes, the ratio of genotypes for each phenotype is given according to the steady state ratios 00:01 and 10:11, respectively.