PNAS
www.pnas.org

Supplementary Information for

Quantifying the impact of treatment history on plasmid-mediated resistance evolution in human gut microbiota

Burcu Tepekule, Pia Abel zur Wiesch, Roger Kouyos, Sebastian Bonhoeffer

Burcu Tepekule.
E-mail: burcu.tepekule@env.ethz.ch

This PDF file includes:

Figs. S1 to S6
Tables S1 to S2
References for SI reference citations

Model parameters

The obtained growth rates are all positive, and consistent with the underlying biological assumptions, as well as the reported growth rates in (1). Values obtained for interaction terms are all negative except one inter-phyla interaction term, which is positive but small in magnitude. Hence, our numerical estimations for the interaction terms are dominated by competition, which is shown to improve gut microbiome stability and permit high diversity of species to coexist (2). Statistics on the parameter estimates are provided in Table S1 and Figures S1 and S2. Figure S1 shows that the phylum *Firmicutes* and *Actinobacteria* do not affect the abundances of *Proteobacteria* and *Bacteroidetes* strongly due to their low inter-phyla interaction rates. From the perspective of resistance evolution modeling, this indicates that the model can be reduced to two phyla including *Proteobacteria* and *Bacteroidetes*, since they are the only two phyla with the resistant variants. Dynamics of this reduced model are presented in Figure S3, where three random treatment courses are applied on the 4-phyla (full) and 2-phyla (reduced) models, and very similar results are observed. Note that this reduction is only possible in hindsight after the interaction parameters are estimated using the full model, and depends highly on the model scenario.

Estimated parameters led to steady state values of 2.92%, 59.77%, 32.48%, and 0.37% for *Proteobacteria* (C_0), *Bacteroidetes* (C_1), *Firmicutes* (C_2), and *Actinobacteria* (C_3), respectively. These values indicate that the most abundant 4 phyla represent the 95.5% of the microbial population, and are in agreement with the temporal mean of the time series data as well as reported values in the literature (3–9). Numerical values for conjugation frequencies (h_{inter}, h_{intra}), resistance costs (ρ_0, ρ_1), and missegregation fraction (γ) are jointly assigned to achieve a plausible decay rate for the plasmid, as explained in more detail in the Methods section. However, the time it takes for the plasmid-bearing population to go extinct (T_{ext}) is affected by the stochasticity in our simulations, and can be different for each realization of the system. This is demonstrated in Figure S4(A), where the distribution of T_{ext} is provided in the absence of treatment, calculated over 1000 hybrid deterministic - stochastic simulations. Figures S4(B) and S4(C) show the baseline dynamics of the system in the absence of any treatment, using a purely deterministic and a hybrid deterministic - stochastic realization, respectively.

Out of 10000 randomly sampled $\{h_{\text{intra}}, h_{\text{inter}}, \rho_0, \rho_1, \gamma\}$ sets, the set $\{10^{-15.831}\,\text{NR}^{-1} \times \text{day}, 10^{-15.923}\,\text{NR}^{-1} \times \text{day}, 0.157, 0.0122, 0.0131\}$ (NR : normalized reads) led to a reservoir extinction time of 720 days in the absence of treatment, which is the targeted extinction time of two years. Numerical values for death rates were obtained by employing the random sampling scheme described in Figure 6. Out of 10000 randomly sampled $\{\delta_0, \delta_1, \delta_2, \delta_3\}$ sets, $\{0.288 \times \text{day}^{-1}, 0.398 \times \text{day}^{-1}, 0.449 \times \text{day}^{-1}, 0.395 \times \text{day}^{-1}\}$ led to the minimum score and therefore used in the model. Standard error of the mean and 95% confidence intervals for $\{\delta_0, \delta_1, \delta_2, \delta_3\}$ and $\{h_{\text{intra}}, h_{\text{inter}}, \rho_0, \rho_1, \gamma\}$ are given in Table S2, and the corresponding parameter distributions are provided in Figures S5 and S6, respectively.

Information quantification

We first transformed each realization into a sequence of length $\max(T_L) + \max(T_{df}) = 1000 + 360 = 1360$ days, where the i^{th} day is denoted by b_i and set to

$$b_i = \begin{cases} 0, & \text{if } i < T_I \\ -1, & \text{if } i \geq T_I \text{ and } \mathbb{1}_T = 0 \text{ (there is no treatment)} \\ +1, & \text{if } i \geq T_I \text{ and } \mathbb{1}_T = 1 \text{ (there is treatment)} \end{cases}$$

so that the days before the colonization with the resistant reservoir ($i < T_I$), days without treatment after the colonization with the resistant reservoir ($i \geq T_I$, $\mathbb{1}_T = 0$), and days with treatment after the colonization with the resistant reservoir ($i \geq T_I$, $\mathbb{1}_T = 1$) has no (0), negative (−1), and positive (+1) impact on the prevalence of resistance P, respectively. After generating these sequences, we applied linear regression in the form of

$$\mathbf{P} = \alpha \cdot [\mathbf{1}\ \ \mathbf{b}], \qquad [1]$$

where $\mathbf{P}$ denotes the prevalence of resistance vector, α denotes the vector of regression coefficients including the intercept term, and $[\mathbf{1}\ \ \mathbf{d}]$ denotes the matrix of treatment history sequences with an additional column of ones for the intercept term. We used 600000 randomly subsampled realizations, 1200 realizations per $N = \{1, 2, \cdots, 20\}$ per $T_{df} = \{0, 15, 30, \cdots, 360\}$ value, and calculated the linear regression coefficients α_i for each day d_i. Figure 5 shows the absolute values of the linear regression coefficients given the day index i, and an exponential decay function that is fit to these data points.

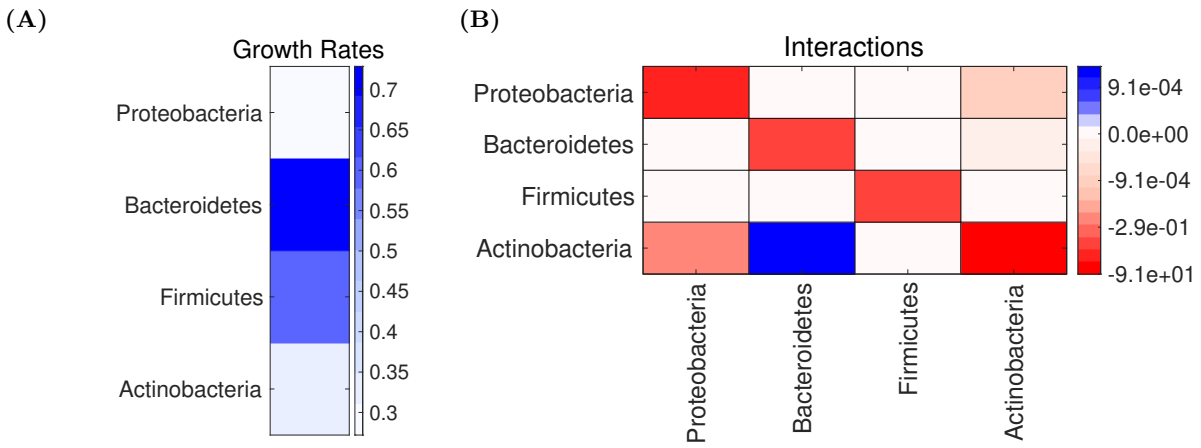

Fig. S1. Visualization of the **(A)** growth and **(B)** interaction parameter estimates provided in Table S1.

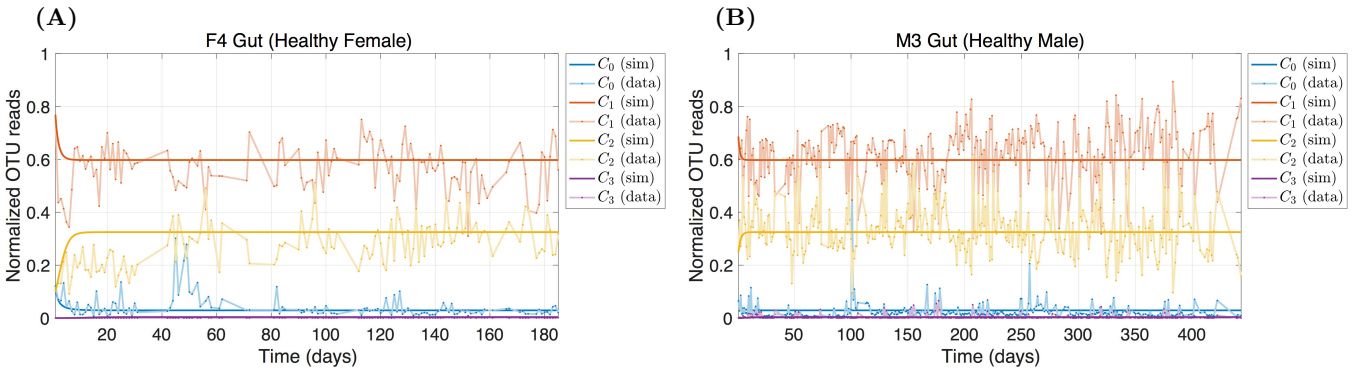

Fig. S2. Phylum-level time series data obtained from two healthy subjects' gut microbiota **(A)** F4 GUT and **(B)** M3 GUT provided in (10). C_0, C_1, C_2, and C_3 denote the phyla *Proteobacteria*, *Bacteroidetes*, *Firmicutes*, and *Actinobacteria*, respectively. Phylum name (data) denotes the data provided in (10), and phylum name (sim) denotes the realization of the system using the estimated parameters, starting from the same initial conditions with the data.

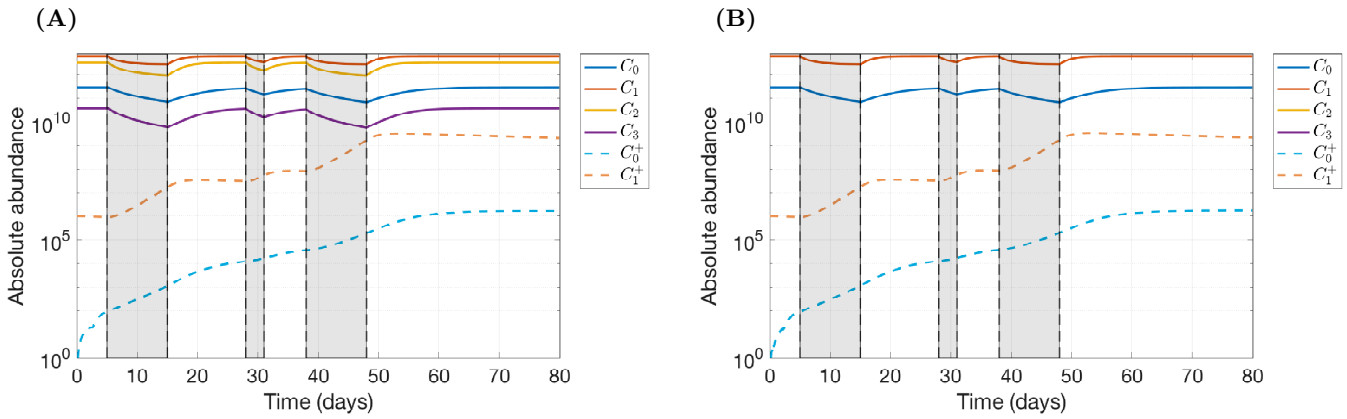

Fig. S3. **(A)** 4-phyla (full) model used in the main manuscript, including *Proteobacteria* (C_0), *Bacteroidetes* (C_1), *Firmicutes* (C_2), and *Actinobacteria* (C_3). **(B)** 2-phyla (reduced) model including only *Proteobacteria* (C_0) and *Bacteroidetes* (C_1).

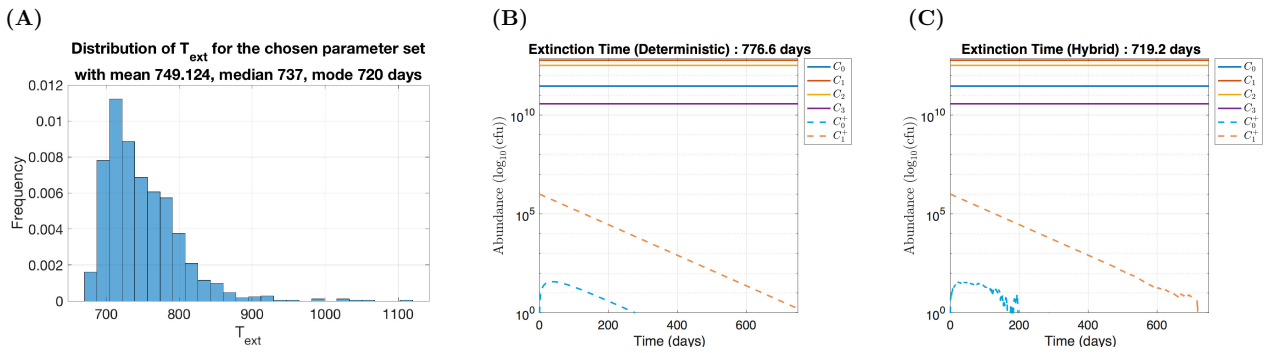

Fig. S4. (A) Distribution of the extinction time of the plasmid (T_{ext}), calculated over 1000 hybrid deterministic - stochastic simulations for the parameter set used in the model. Demonstration of the baseline dynamics of the system in the absence of any treatment, using **(B)** a purely deterministic and **(C)** a hybrid deterministic - stochastic realization.

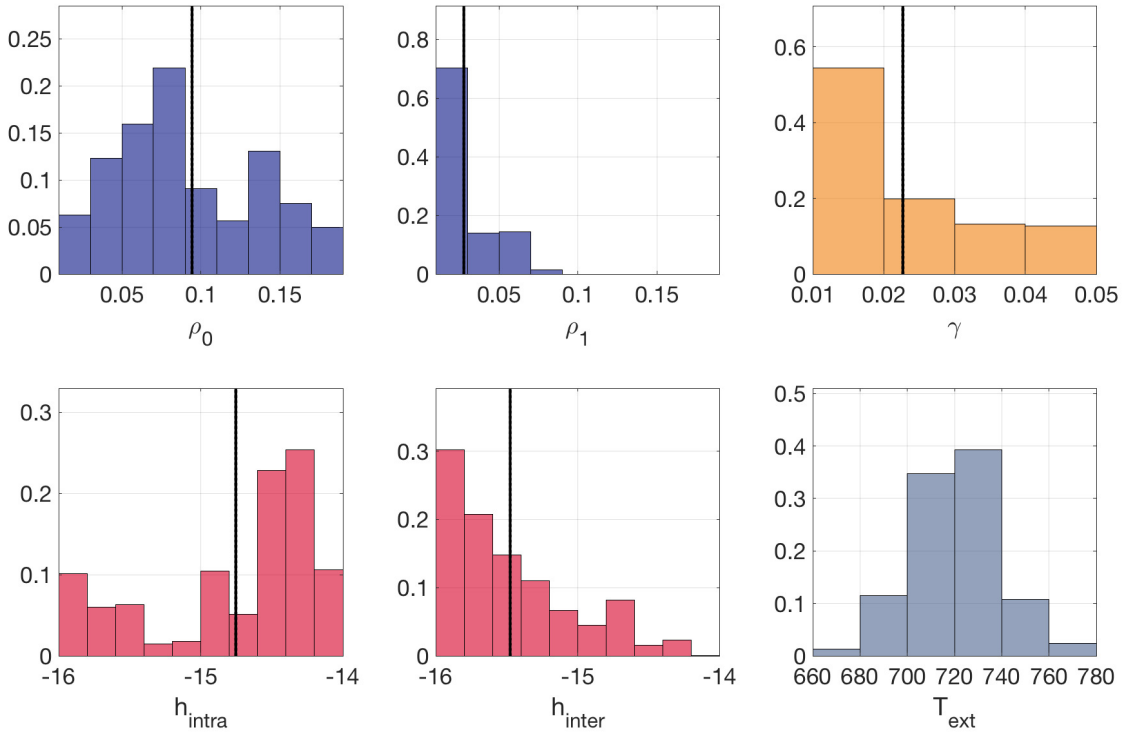

Fig. S5. Distributions of $\{h_{\text{intra}},\ h_{\text{inter}},\ \rho_0,\ \rho_1,\ \gamma\}$ leading to a normally distributed plasmid extinction time with mean 720 and standard deviation of 20 days.

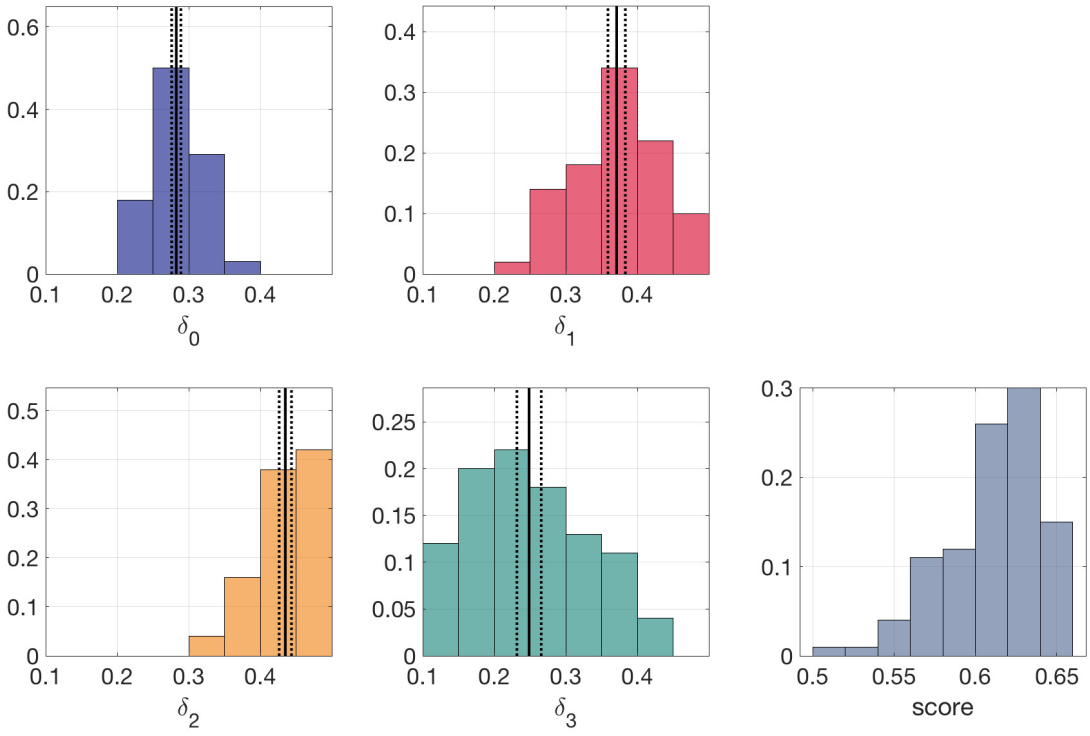

Fig. S6. Distributions of $\{\delta_0,\ \delta_1,\ \delta_2,\ \delta_3\}$ with the best 100 scores.

Table S1. Results of the parameter estimation for growth and interaction terms. Phylum-level time series data obtained from two healthy subjects' gut microbiota (referred as F4 GUT and M3 GUT) provided in (10) is used. To infer parameters invariant to the scale of the data, OTU reads are normalized across all phyla for each time point, and used the resulting quotients for *Proteobacteria*, *Bacteroidetes*, *Firmicutes*, and *Actinobacteria* as a proxy for their abundance in the gut. Bayesian variable selection algorithm in MDSINE (11) is adopted for the time series analysis of the data.

parameter type	source phylum	target phylum	value	significance	MCMC std
growth rate	NA	Actinobacteria	0.34119	0.073875	0.074441
growth rate	NA	Bacteroidetes	0.72859	0.073875	0.04815
growth rate	NA	Firmicutes	0.58842	0.073875	0.042861
growth rate	NA	Proteobacteria	0.27198	0.073875	0.040687
interaction	Actinobacteria	Actinobacteria	-93.2702	0.073875	8.3928
interaction	Actinobacteria	Bacteroidetes	0.011817	0.003255	0.2166
interaction	Actinobacteria	Firmicutes	0	0	0
interaction	Actinobacteria	Proteobacteria	-0.085337	0.0073875	1.0405
interaction	Bacteroidetes	Actinobacteria	-1.23E-05	4.44E-05	0.001839
interaction	Bacteroidetes	Bacteroidetes	-1.219	0.073875	0.078665
interaction	Bacteroidetes	Firmicutes	0	0	0
interaction	Bacteroidetes	Proteobacteria	0	0	0
interaction	Firmicutes	Actinobacteria	0	0	0
interaction	Firmicutes	Bacteroidetes	0	0	0
interaction	Firmicutes	Firmicutes	-1.8118	0.073875	0.12921
interaction	Firmicutes	Proteobacteria	2.88E-06	4.44E-05	0.00043126
interaction	Proteobacteria	Actinobacteria	-0.0010279	0.00026674	0.067505
interaction	Proteobacteria	Bacteroidetes	0	0	0
interaction	Proteobacteria	Firmicutes	0	0	0
interaction	Proteobacteria	Proteobacteria	-9.3	0.073875	0.86257

Table S2. Standard error of the mean and 95% confidence intervals for $\{\delta_0, \delta_1, \delta_2, \delta_3\}$ with the best 100 scores, and $\{h_{\text{intra}}, h_{\text{inter}}, \rho_0, \rho_1, \gamma\}$ leading to a normally distributed plasmid extinction time with mean 720 and standard deviation of 20 days.

parameter	mean	SEM	CI95
δ_0	0.283	0.00326	0.276-0.289
δ_1	0.371	0.00612	0.359-0.383
δ_2	0.435	0.00435	0.426-0.444
δ_3	0.249	0.00856	0.232-0.266
ρ_0	0.0945	4.66E-05	0.0944-0.095
ρ_1	0.0278	1.70E-05	0.0278-0.028
h_{intra}	-14.8	0.00059	-14.8–15
h_{inter}	-15.5	0.000443	-15.5–15
γ	0.0227	1.13E-05	0.0227-0.023

References

1. Stein RR, et al. (2013) Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota. *PLoS computational biology* 9(12):e1003388.
2. Coyte KZ, Schluter J, Foster KR (2015) The ecology of the microbiome: networks, competition, and stability. *Science* 350(6261):663–666.
3. Qin J, et al. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *nature* 464(7285):59.
4. Qin N, et al. (2015) Influence of h7n9 virus infection and associated treatment on human gut microbiota. *Scientific reports* 5:14771.
5. Eckburg PB, et al. (2005) Diversity of the human intestinal microbial flora. *science* 308(5728):1635–1638.
6. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* 489(7415):220.
7. Arumugam M, et al. (2011) Enterotypes of the human gut microbiome. *nature* 473(7346):174.
8. Huttenhower C, et al. (2012) Structure, function and diversity of the healthy human microbiome. *Nature* 486(7402):207.
9. Peris-Bondia F, Latorre A, Artacho A, Moya A, D'Auria G (2011) The active human gut microbiota differs from the total microbiota. *PloS one* 6(7):e22448.
10. Caporaso JG, et al. (2011) Moving pictures of the human microbiome. *Genome biology* 12(5):R50.
11. Bucci V, et al. (2016) Mdsine: Microbial dynamical systems inference engine for microbiome time-series analyses. *Genome Biology* 17(1):121.