Co-colonization interactions drive explicit frequency-dependent dynamics among \( N \) types

Sten Madec\textsuperscript{a,c}, Erida Gjini\textsuperscript{b,d}
\textsuperscript{a}Institut Denis Poisson, University of Tours, Tours, France
\textsuperscript{b}Instituto Gulbenkian de Ciência, Oeiras, Portugal
\textsuperscript{c}sten.madec@univ-tours.fr
\textsuperscript{d}egjini@igc.gulbenkian.pt

Abstract

A fundamental question in ecology is the generation and maintenance of biodiversity. Classical approaches consider multi-species communities with Lotka-Volterra ODE models where inter- vs. intra-species interactions are key. Typically in high-dimensional systems, analysis is hard and model reduction is needed. Here, we describe and study a new system of multi-type interactions that arise in co-colonization, and develop a useful model reduction framework based on slow-fast dynamics under quasi-neutrality. We show that in a multi-type, Susceptible-Infected-Susceptible (SIS) system with co-colonization, neutral coexistence dynamics between \( N \) closely related strains occurs on a fast timescale, based on mean trait values, whereas the non-neutral selective forces act on a slow timescale, driven by the variance and asymmetries of the co-colonization interaction matrix. The explicit \( N \) equations for relative type frequencies on the slow manifold enable efficient computation of complex multi-strain dynamics, and provide analytical insight into high-dimensional community ecology, with potential applications to other multi-body systems.

Keywords: SIS model, slow-fast dynamics, weak selection, coexistence, multispecies community ecology, broken symmetries

Introduction

A fundamental question in ecology and evolutionary biology relates to how biodiversity is generated and maintained [1-4]. Many factors have been shown to be important, including inter- vs. intra-species interactions [5], population size [6], number and functional links with resources [7], and movement in space [8]. Most research in this field has considered multi-species interactions with classical Lotka-Volterra models or with evolutionary game theoretic approaches.

In this Letter, we describe and study a new system of multi-type interactions that arise in the epidemiological dynamics of co-colonizing strains of the same microbial species. An example of such a system could be polymorphic \textit{Streptococcus pneumonia} bacteria. We generalize a previously-introduced Susceptible-Infected-Susceptible (SIS) framework for 2 circulating strains with co-colonization [9,10], to the context of \( N \) strains. We show that an explicit reduced system emerges from a time scale separation, expressing the total dynamics as a fast plus a slow component, related to broken symmetries in co-colonization interactions.

In our system, microbial strains can infect a host simultaneously, and here we concentrate only on the case of up to 2 strains co-colonizing a host, whereby single colonization by a resident strain alters the susceptibility to incoming strains, (increasing or decreasing it) by a factor \( K_{ij} \), relative to uninfected hosts, without acquired immunity. The transmission and clearance rates of all strains are equal, except for the interactions in co-colonization given by an \( N \times N \) matrix. In earlier theoretical studies such interaction has been studied in a slightly different context in models considering arbitrary infection multiplicity [10]. Later evolutionary frameworks have also considered vulnerability to co-infection [11]. In the experimental literature, for example in microbiota studies, the phenomenon of colonization resistance has been reported [12]. In pneumococcal bacteria, more relevant to our model, carriage of one serotype has been shown to alter the acquisition rate of a second serotype [13]. Some studies have shown how this trait at the host-pathogen interface impacts disease persistence [14]. coexistence and vaccination effects [15,16], and contributes to diversity in other traits, e.g. virulence [17] and antibiotic resistance [18].

Yet, robust mathematical frameworks to analyze the full spectrum of eco-epidemiological patterns emerging exclusively from co-colonization interactions among a high number of strains remain undeveloped. Here, we fill this gap. We model an arbitrary number of closely-related strains, and show that in the \( N \)-strain system with altered susceptibilities in co-colonization, neutral coexistence dynamics between types occurs on a fast timescale, whereas the non-neutral selective forces act on a slow timescale, driven explicitly by trait variation in the co-colonization interaction space. These dynamics can be very complex, and are nonlinearly modulated by global parameters such as overall transmission intensity \( R_0 \), and mean interaction coefficient between strains. We derive a closed analytic solution for relative type frequencies over long time-scales in a changing fitness landscape.

\textit{N-strain SIS model with co-colonization}

We consider a multi-type infectious agent, transmitted via direct contact, following susceptible-infected-susceptible
(SIS) epidemiological dynamics, with the possibility of co-colonization. The number of potentially co-circulating strains is N. With a set of ordinary differential equations, we describe the proportion of hosts in several compartments: susceptibles, S, hosts colonized by one type I_i, and co-colonized hosts I_ij, with two types of each combination, independent of the order of their acquisition. We have:

\[
\begin{align*}
\dot{S} &= m(1-S) - S \sum_{j=1}^{N} F_j, \\
\dot{I}_i &= F_i S - m I_i - I_i \sum_{j=1}^{N} K_{ij} F_j, \quad 1 \leq i \leq N \\
\dot{I}_{ij} &= I_i K_{ij} F_j - m I_{ij}, \quad 1 \leq i, j \leq N
\end{align*}
\]

where \( F_i = \beta (I_i + \sum_{j=1}^{N} \frac{1}{2}(I_{ij} + I_{ji})) \) gives the force of infection of strain i. Remark that here we assume that hosts colonized with a mixture of two subtypes i and j transmit either with equal probability. For more information and interpretation of parameters see Table 1 (Supplementary materials 1). Notice that \( S = 1 - \sum (I_i + I_{ij}) \), thus the dimension of the system is indeed \( N + N^2 \). In practice, \( I_i \) and \( J_j \) cannot be distinguished so the dimension is effectively \( N + N(N - 1)/2 \). For the special case of \( N = 2 \), this model has been described and analyzed elsewhere \( [16, 9] \). We study pathogen diversity only in relation to how strains interact with each other upon co-colonization (\( K_{ij} \)), assuming equivalence in transmission \( \beta \) and clearance rate \( \gamma \) (although extension to further asymmetries in these traits is ongoing). The coefficients \( K_{ij} \) can be above or below 1, indicating facilitation or competition between any two types in co-colonization. In the above notation \( m = \gamma + r \), encapsulating both clearance rate \( \gamma \) of colonization episodes and recruitment rate of susceptible hosts \( r \), equal to the mortality rate from all compartments \( (r = d) \). For a model summary diagram see Fig 1a-b.

**Results**

**Re-writing the system in aggregated form**

In order to derive a reduced model for such a general N strain system, displaying complicated dynamics (Fig. 1b), we use the following notation:

\[
J_i = I_i + \sum_{j=1}^{N} \frac{1}{2}(I_{ij} + I_{ji}) \quad \text{and} \quad T = \sum I_i + \sum I_{ij} + \sum I_{ji}. \quad (2)
\]

for the prevalence of strain i in the population, and the total prevalence of all strains, respectively. Total prevalence satisfies \( T = \sum J_i \) and the forces of infection are: \( F_i = \beta J_i \). With these notations, the system (1) reads:

\[
\begin{align*}
\dot{S} &= m(1-S) - \beta S T, \\
\dot{T} &= \beta S T - m T, \\
\dot{I}_i &= \beta J_i S - m I_i - \beta I_i \sum_{j=1}^{N} K_{ij} J_j, \quad 1 \leq i \leq N \\
J_i &= (\beta S - m) I_i + \frac{\beta}{2} \sum_{j=1}^{N} (I_i K_{ij} J_j - I_j K_{ji} J_i), \quad 1 \leq i \leq N \\
\dot{I}_{ij} &= I_i K_{ij} J_j - m I_{ij}, \quad 1 \leq i, j \leq N
\end{align*}
\]

(3)

This system of \( 2 + N + N^2 \) equations admits a triangular structure: i) First we describe block of 2 equations \( (S, T) \). ii) Next, we study the block of \( 2N \) equations \( (I_i, J_i) \), which is the most complicated. iii) Lastly, we deal with the block of the \( N^2 \) equations of \( I_{ij} \), which is simple once the dynamics of \( I_i \) and \( J_i \) are known. In the following we describe briefly these three blocks.

**The 1st part \((S, T)\).**

\[
\begin{align*}
\dot{S} &= m(1-S) - \beta S T, \\
\dot{T} &= \beta S T - m T
\end{align*}
\]

Clearly, if the basic reproduction number \( R_0 = \frac{\beta}{m} > 1 \), then \( (S, T) \rightarrow \left( \frac{1}{R_0}, 1 - \frac{1}{R_0} \right) \). Hence, without loss of generality, we consider that \( (S, T) = \left( \frac{1}{R_0}, 1 - \frac{1}{R_0} \right) \).

**The 2nd block of equations \((I_i, J_i)\).** If we reduce the system to the invariant manifold \( (S, T) = (S, T) = \left( \frac{1}{R_0}, 1 - \frac{1}{R_0} \right) \), we obtain the following sub-system of \( 2N \) differential equations:

\[
\begin{align*}
\dot{I}_i &= m(J_i - I_i) - \beta I_i \sum_{j=1}^{N} K_{ij} J_j, \\
\dot{J}_i &= \frac{\beta}{2} \sum_{j=1}^{N} (I_i K_{ij} J_j - I_j K_{ji} J_i), \quad 1 \leq i \leq N
\end{align*}
\]

(4)

together with: \( \sum_{i=1}^{N} J_i = 1 - \frac{1}{R_0} = \frac{\beta}{m} \). The dynamics of (4) are complex and we use a slow-fast approach to analyze it.

**The co-colonization block \( I_{ij} \).** This is the simple system

\[
\dot{I}_{ij} = \beta I_i K_{ij} J_j - m I_{ij}, \quad 1 \leq i, j \leq N
\]

(5)

Once it is known that \( (I_i, J_i) \rightarrow (I^*_i, J^*_i) \), then (5) implies \( I_{ij} \rightarrow \frac{\beta}{2} I^*_i K_{ij} J^*_j \). Thus, once the dynamics (4)-(5) are explicit, so are the dynamics of (1).

---

![Figure 1: Model summary diagram](image-url)

**a)** Epidemiological model structure. The force of infection for each strain i is \( F_i \), and governs the transition from susceptible to singly-colonized state, and from singly-colonized to co-colonized state. Clearance happens at equal rate for single and co-colonization, back to the susceptible state. Co-colonization rate by strain j of singly-colonized hosts with i is altered by a factor \( K_{ij} \) relative to uncolonized hosts. b) Considering \( K_{ij} = k + \alpha x_{ij} \), the global epidemiological dynamics can be decomposed in a fast and slow component. On the fast time-scale (\( \alpha (1/\varepsilon) \)), strains follow neutral dynamics, driven by mean parameters, where \( S, I, \) and \( D \) are conserved. On a slow time-scale, \( \varepsilon t \), complex non-neutral dynamics between strains takes place, depicted here by the constituent variations within the blue and green.
Slow-fast representation of

Since we model N closely-related strains, we can write each co-colonization coefficient as: \( K_{ij} = k + e \alpha_{ij} \), where \( 0 \leq e \ll 1 \). Replacing these in (4), and re-arranging, we obtain:

\[
\begin{align*}
I &= m(T - I) - \beta k T I - e \beta \sum_{j=1}^{N} I_i \alpha_{ij} J_j \\
\dot{I}_i &= m(J_i - I_i) - \beta k T I_i - e \beta \sum_{j=1}^{N} \alpha_{ij} J_j, \\
J_i &= \frac{\beta k}{T} (I J_i - I_i T) + \frac{e \beta}{T} \sum_{j=1}^{N} (I_j \alpha_{ij} J_i - I_i \alpha_{ij} J_j),
\end{align*}
\]

where \( 1 \leq i \leq N \), and \( I = \sum I_i \).

Neutral system

If \( e = 0 \), then from above, we obtain the Neutral model

\[
\begin{align*}
I &= m(T - I) - (m + \beta k T) I \\
\dot{I}_i &= m(J_i - I_i) - (m + \beta k T) I_i, \quad 1 \leq i \leq N \\
J_i &= \frac{\beta k}{T} (I J_i - I_i T),
\end{align*}
\]

The first equation gives \( I(t) = I + e^{-\tau(m+\beta k T)}(I(0) - I) \rightarrow I := \frac{m(T - I)}{m + \beta k T}. \) Co-colonization prevalence is simply derived as: \( D = T - I \). Fixing \( I = I \), yields the (degenerate) linear system:

\[
\begin{pmatrix}
I_i \\
J_i
\end{pmatrix} = \begin{pmatrix}
-(m + \beta k T) & m \\
-\beta k T & \beta k T
\end{pmatrix} \begin{pmatrix}
I_i \\
J_i
\end{pmatrix} + A_0 \begin{pmatrix}
I_i \\
J_i
\end{pmatrix},
\]

Matrix \( A_0 \) has the two eigenvalues 0 and \( -\xi = tr(A_0) < 0 \), and eigenvectors \( V_0 = \begin{pmatrix} 1 \\ T \end{pmatrix}, \quad V_1 = \begin{pmatrix} 2T \\ D \end{pmatrix} \). We define \( H_i \) and \( z_i \) as:

\[
\begin{pmatrix}
H_i \\
z_i
\end{pmatrix} = \begin{pmatrix}
2T & 1 \\ D & T
\end{pmatrix}^{-1} \begin{pmatrix}
I_i \\
J_i
\end{pmatrix}.
\]

We have

\[
\begin{align*}
H_i &= -\xi H_i \\
z_i &= 0.
\end{align*}
\]

Thus

\[
H_i = \frac{IT}{2(T^2 - DI)} \begin{pmatrix}
I_i \\
J_i
\end{pmatrix}
\]

tends to zero, and

\[
z_i = \frac{(I_i)}{1} + \frac{2(T^2)}{2(T^2 - DI)} \begin{pmatrix}
J_i \\
I_i
\end{pmatrix},
\]

remains constant in a neutrally stable manner. \( H_i \) measures the difference between the part of strain \( i \) on the single infected versus the part of the strain \( i \) on the total number of infected. Thus, \( H_i \rightarrow 0 \) means that the proportion of strain \( i \) in single colonization, tends to be the same as the proportion of strain \( i \) in the total mass of all colonization. This implies it also tends to be equal to the proportion of strain \( i \) in co-colonization. Because on the slow manifold, we have \( H_i = 0 \), we can infer that during fast dynamics \( z_i \) tends to:

\[
z_i = \frac{J_i}{T} = \frac{I_i}{1},
\]

whereby on the slow manifold \( z_i \) describes exactly the frequency of strain \( i \) among all hosts that are colonized, with \( \sum z_i = 1 \).

The slow-fast dynamics

In (6), \( I \) plays a role in the order 0 term of the last two equations. Thus, we use the ansatz: \( I(t) = 1 + eX(t) + O(e) \), and rewrite (6) (for details see Supplementary Materials 2). After some algebraic manipulations, using newly defined variables we obtain the explicit slow-fast system:

\[
\begin{align*}
X &= -(m + \beta k T)X - \beta Q(H, z) + O(e) \\
\dot{H}_i &= -\xi H_i + O(e) \\
\dot{z}_i &= \epsilon \delta (g_{i,2}(H, z) + T(DH_i + Tz_i)X + O(e)),
\end{align*}
\]

for \( 1 \leq i \leq N \), where \( \delta = \frac{\beta}{2(T^2 - DI)} > 0 \) is a constant.

If \( e \rightarrow 0 \) then \( z_i \) remains constant and \( (X(t), H(t)) \rightarrow (\phi(z), 0) \) (shown in detail in S2). However, on the slow time scale \( t = eT \) in (11), \( z_i \) is not constant, and letting \( e \rightarrow 0 \), one obtains the slow dynamics on \( z_i \), given by the \( N \)-dimensional system:

\[
\dot{z}_i = \delta (g_{i,2}(0, z) + T^2 z_i \phi(z)), \quad 1 \leq i \leq N
\]

which reads explicitly

\[
\dot{z}_i = \Theta \sum_{j \neq i} \left[ \mu (\alpha_{ji} - \alpha_{ij}) + \alpha_{ij} z_j + \alpha_{ii} z_i - q(z) \right],
\]

for \( 1 \leq i \leq N \), where the constants \( \Theta, \mu > 0 \) are:

\[
\Theta = \frac{\beta T D}{2T^2 - DI} ; \quad \mu = \frac{1}{D} = \frac{1}{k(R_0 - 1)},
\]

and the term \( q(z) \) is given by:

\[
q(z) = \sum_{1 \leq j,k \leq N} \alpha_{jk} z_k z_j.
\]

The constant \( \Theta \) is the rate that sets the tempo of multi-strain "motion" on the slow manifold towards an equilibrium. Notice that \( \Theta \) depends specifically on the absolute transmission rate of the pathogen \( \beta \), but also on mean traits \( R_0 \) and \( k \), via the conserved aggregated quantities \( T, I, D \). The quantity \( \mu \) on the other hand represents the ratio between single colonization and co-colonization prevalence from the neutral system, a factor that amplifies the importance of type asymmetry.

The function \( q(z) \) represents the time-changing impact of all the strains on their environment, which in turn modifies the fitness landscape for each in an equal manner. A more explicit way to see \( q(z) \) is in terms of relative change in 'effective' mean interaction coefficient between all extant types in the system, which if negative, indicates a global trend toward more competition in co-colonization, and if positive, indicates a global trend toward more facilitation. At this stage, we can apply quasi-stationarity methods (see [19, 20, 21, 22]) to show that the solution of (11) tends to the solution of (12) as \( e \rightarrow 0 \) locally uniformly in time on \([t_0, +\infty)\) for some \( t_0 > 0 \).
Link with pairwise invasion fitness \( \lambda_i \)

Next, we provide another equivalent very useful representation of the model by using the notion of pairwise invasion fitness \( \lambda_i \). Let \( \lambda_i \) be the exponential growth rate of the strain \( i \) evaluated when introduced at the trivial endemic equilibrium of the strain \( j \) alone. If the fitness \( \lambda_i > 0 \), strain \( i \) will invade, and vice versa. By considering the rate of growth of strain \( i \) in an equilibrium set by \( j \) (\( dz_i/dt \) in the special case where all \( z_k = 0 \) for \( k \neq j \)), we find the exact formulation of pairwise invasion fitness is given by:

\[
\lambda_i = \mu(\alpha_{ji} - \alpha_{ij}) + \alpha_{ji} - \alpha_{ij}.
\] (13)

We can thus recast the system (12) using only these fitness notations (for details see Supplementary Material 3):

\[
\frac{d}{dt} z_i = \Theta z_i \cdot \left( \sum_{j \neq i} \lambda_i^j z_j - \sum_{k \leq i} \sum_{j \leq k} \lambda_i^k z_j z_k \right)
\] (14)

The above can be made more compact by denoting \( \Lambda = (\lambda_i^j)_{i,j} \), the pairwise invasion fitness matrix, and using vector notation:

\[
\frac{d}{dt} z = \Theta z \cdot (\Lambda z - z' \Lambda z).
\] (15)

It is this matrix \( \Lambda \) that defines all ‘edges’ of the rescaled interaction network between \( N \) strains. For \( N = 2 \), similar to the classical Lotka-Volterra model, in our model, as already shown [9], there are only four possible outcomes between 2 strains (edge linking 1 and 2): i) \( \lambda_1 > 0, \lambda_2 > 0 \) : stable coexistence of 1 and 2; ii) \( \lambda_2 < 0, \lambda_1 < 0 \) : bistability of 1-only and 2-only; iii) \( \lambda_1 > 0, \lambda_2 < 0 \) : 1-only competitive exclusion; iv) \( \lambda_1 < 0, \lambda_2 > 0 \) : 2-only competitive exclusion. Knowing all pairwise invasion fitnesses between each couple of strains, via (15) we can reconstitute the ultimate dynamics of the full system with \( N \) types and co-colonization.

Going back to the original \( N + N(N - 1)/2 \) variables

Recall the original system with \( N \) strains is given by the SIS model with co-colonization interactions (1). While the decomposition \( K_{ij} = k + \epsilon \alpha_{ij} \) is mathematically non-unique, and can be applied with respect to any reference \( k \), provided that the resulting \( \epsilon \) is small, a possible convenient choice is to define \( k \) as the average of the original interaction matrix entries \( K_{ij} \):

\[
k = \frac{\sum_{i,j} K_{ij}}{N^2}.
\] (16)

and to define deviation from neutrality, \( \epsilon \), as the root mean square distance of the \( K_{ij} \) from the mean interaction coefficient:

\[
\epsilon = \sqrt{\frac{\sum_{i,j} (K_{ij} - k)^2}{N^2}}
\]

thus representing the standard deviation of the \( K_{ij} \) traits. The direction of deviation from neutrality (bias) for the interaction between strain \( i \) and \( j \) is then obtained as:

\[
\alpha_{ij} = \frac{K_{ij} - k}{\epsilon}.
\]
In this study, we investigated how co-colonization interaction networks drive coexistence in a system with $N$ closely related types. We derived a slow-fast decomposition for the dynamics, which links explicitly variation in the interaction matrix with the slow time scale under which multi-type selective dynamics unfold. Model reduction for small deviations from neutrality allowed us to express strain frequencies on the slow manifold via only $N$ equations (instead of $N + N(N - 1)/2$ equations in the full system). These can be easily re-mapped to the more complex epidemiological variables of the original system.

The slow variables $z_i$ ($1 \leq i \leq N$), describing relative strain frequencies in the host population, necessarily equal in single and co-colonization, could be used as a practical test for quasi-neutrality, when strain prevalence data are available.

Our focus here was on presenting the timescale decomposition method for this system. The wider and more complete ecological picture, as well as the analysis of diversity-stability relationships in coexistence will be the focus of another paper [24]. We believe this work makes a crucial step towards the full characterization of conservative multi-type SIS dynamics.

The slow-fast framework allows to reduce the complexity of multi-strain systems, and understand transient and long-term coexistence on conserved manifolds near neutrality. We expect our results to stimulate further progress for the investigation of coexistence and evolution in multi-strain epidemiological systems such as the one of pneumococcal bacteria [25], where explicit mathematical results linking neutral and niche mechanisms, for $N$ large, are thusfar missing. Although motivated by such a pathogen with altered susceptibilities in co-colonization, the global transmission dynamics captured here have broader implications. As an example, information on the resilience of a group of strains may be derived from the sign of $q(z)$ in our system [12]: if $q(z) > 0$ then each strain is less competitive within the group but the overall community is more resistant to invasion by a new strain. Our model’s generality and closed analytical form for relative type dynamics provide a new bridge between epidemiology and community ecology, invite comparison with the classical Lotka-Volterra model, and offer extension to other multi-body problems in physics and genetics.

We acknowledge support by the University of Tours and EURAXESS for an invited researcher stay of Erida Gjini at IDP, in Tours during 2019.

**References**

[1] G. F. Gause, *The struggle for existence* (Williams and Wilkins, Baltimore, MD, 1934).

[2] R. H. MacArthur, *The theory of island biogeography*, vol. 1 (Princeton University Press, 1967).

[3] D. Tilman, Proceedings of the National Academy of Sciences USA 101, 10854 (2004).

[4] S. P. Hubbell, *The unified neutral theory of biodiversity and biogeography* (MPB-32), vol. 32 (Princeton University Press, 2001).
[5] D. Tilman, The American Naturalist 129, 769 (1987).
[6] C. Taylor, D. Fudenberg, A. Sasaki, and M. A. Nowak, Bulletin of mathematical biology 66, 1621 (2004).
[7] R. A. Armstrong and R. McGehee, Theoretical population biology 9, 317 (1976).
[8] M. A. Nowak and R. M. May, Nature 359, 826 (1992).
[9] E. Gjini and S. Madec, Theoretical Ecology 10, 129 (2017).
[10] F. R. Adler and R. C. Brunet, Theoretical population biology 40, 369 (1991).
[11] S. Alizon, Interface Focus (2013).
[12] R. A. Britton and V. B. Young, Trends in microbiology 20, 313 (2012).
[13] M. Lipsitch, O. Abdullahu, A. D’Amour, et al., Epidemiology 23, 510 (2012).
[14] M. Gavão, F. Dionisio, and E. Gjini, Bulletin of mathematical biology 79, 2068 (2017).
[15] M. Lipsitch, PNAS 94, 6571 (1997).
[16] E. Gjini, C. Valente, R. Sá-Leão, and M. G. M. Gomes, Journal of Theoretical Biology 388, 50 (2016).
[17] S. Alizon, J. C. de Roode, and Y. Michalakis, Ecology letters 16, 556 (2013).
[18] N. G. Davies, S. Flasche, M. Jit, and K. E. Atkins, Nature ecology & evolution 3, 440 (2019).
[19] A. N. Tikhonov, Matematicheskii sbornik 73, 575 (1952).
[20] C. Lobry and T. Sari, Electronic Journal of Differential Equations 19, 1 (1998).
[21] C. Lobry and T. Sari, Contrôle non linéaire et Applications 64, 155 (2005).
[22] F. Hoppensteadt, Transactions of the American Mathematical Society 123, 521 (1966).
[23] G. Meszéna, M. Gyllenberg, F. J. Jacobs, and J. A. Metz, Physical review letters 95, 078105 (2005).
[24] E. Gjini and S. Madec, in prep. (2019).
[25] S. Cobey and M. Lipsitch, Science (New York, NY) 335, 1376 (2012).
[26] K. Dietz, Statistical methods in medical research 2, 23 (1993).
SUPPLEMENTARY MATERIAL

Co-colonization interactions drive explicit frequency-dependent dynamics among $N$-types

by Sten Madec and Erida Gjini

1. Table of illustrative model parameters for a biological system (e.g. pneumococcus bacteria transmission)

| Parameter | Interpretation | Features | Illustrative value | Reference |
|-----------|----------------|----------|--------------------|-----------|
| $\beta$   | Transmission rate of pathogen | $\beta > 0$ | 2.3 (month)$^{-1}$ | [16] |
| $K_{ij}$  | Interaction coefficient between strains $i$ and $j$ in co-colonization | Alteration of susceptibility to $j$ when colonized by $i$ | 0.1 | [13] |
| $r$       | Susceptible recruitment rate | $r > 0$ | 0.02 (month)$^{-1}$ | [16] |
| $d$       | Natural host mortality rate | $d = r$ | 0.02 (month)$^{-1}$ | [16] |
| $\gamma$  | Clearance rate of colonization | $\gamma > 0$ | 0.8 (month)$^{-1}$ | [16] |
| $m$       | Net outflow rate from the colonization state | Inverse duration of any colonization episode | $m = r + \gamma$ | |
| $N$       | Number of strains in the system | Arbitrarily, $N > 1$ | 30-60 | [13] |
| $R_0$     | Basic reproduction number of the pathogen, $R_0 = \beta/m$, here assumed equal among strains | $R_0 < 1$ Disease extinction | \[7\] |
|           |                                | $R_0 > 1$ Persistence | 2-3 | [26], [16] |

2. Derivation details for the slow-fast dynamics

Using the ansatz: $I(t) = I + eX(t) + O(\varepsilon)$, system (4) reads

$$\dot{X} = -(m + \beta k T)X - \beta \sum_{j=1}^{n} \sum_{j=1}^{n} I_{i} \alpha_{ij} J_{j} + O(\varepsilon)$$

$$\begin{pmatrix} I_{i} \\
J_{j} \end{pmatrix} = A \begin{pmatrix} I_{i} \\
J_{j} \end{pmatrix} + e\phi \frac{X}{2} \sum_{j=1}^{n} I_{i} \alpha_{ij} J_{j} + e\beta W_{i}(I, J) + O(\varepsilon)$$

wherein we have denoted:

$$W_{i}(I, J) = \left( \frac{1}{2} \sum_{j} I_{i} \alpha_{ij} J_{j} \right)$$

The final slow fast form is given by using the new variables $H_{i}$ and $z_{i}$ defined in (9). Hence, define

$$\begin{pmatrix} h_{i,1}(I, J) \\
h_{i,2}(I, J) \end{pmatrix} = \begin{pmatrix} 2T & 1 \\
D & T \end{pmatrix}^{-1} \begin{pmatrix} -2 \sum_{j} \alpha_{ij} J_{j} \\
\sum_{j} I_{i} \alpha_{ij} J_{j} \end{pmatrix}$$

and let us introduce the $G_{i} = (g_{i,1}, g_{i,2})$ as

$$g_{i,s}(H, z) = (2(T)^{2} - ID)h_{i,s}(2TH + Iz, DH + Tz), \quad s = 1, 2.$$  

and

$$Q(H, z) = \sum_{i,j} \alpha_{ij} (2TH_{i} + Iz_{j})(DH_{j} + Tz_{j}).$$

we have

$$\dot{X} = -(m + \beta k T)X - \beta Q(H, z) + O(\varepsilon)$$

\[1\] Denoting $G = (G_{i})_{1 \leq i \leq N}$ and $W = (w_{i})_{1 \leq i \leq N}$, this simply means that $G = \phi^{-1} \circ W \circ \phi$ where $\phi(H, z) = (I, J)$ is the linear change of variables defined in (9).
Finally, using these new unknowns, one obtains the expression of the slow fast system (11):

\[
\begin{align*}
X &= -(m + \beta k T)X - \beta Q(H, Z) + O(\varepsilon) \\
\dot{H}_i &= -\mu H_i + O(\varepsilon), \quad 1 \leq i \leq N \\
\dot{z}_i &= \varepsilon \delta \left( g_{i,2}(H, Z) + T(DH_i + Tz_i)X + O(\varepsilon) \right)
\end{align*}
\]  

(17)

where \( \delta = \frac{\beta}{2(T)^2 - 1D} \) is a positive constant. If \( \varepsilon \to 0 \) then \( z_i \) remains constant and \( (X(t), H(t)) \to (\phi(z), 0) \) where:

\[
\dot{\phi}(z) = -\frac{\beta Q(0, z)}{m + \beta k T}T = -\frac{\beta IT}{m + \beta k T} \sum_{i,j} \alpha_{ij} z_i z_j.
\]

(18)

On the slow time scale \( \tau = \varepsilon t \) in (11), and letting \( \varepsilon \to 0 \), one obtains the slow dynamics

\[
\begin{align*}
0 &= -(m + \beta k T)X - \beta Q(H, z) \\
0 &= -\mu H_i, \quad 1 \leq i \leq N \\
\dot{z}_i &= \delta \left( g_{i,2}(H, z) + k(T)^2 z_i X \right)
\end{align*}
\]

(19)

Taking \( (X, H) = (\phi(z), 0) \) yields the slow dynamics reducted to the slow manifold

\[
\dot{z}_i = \delta \left( g_{i,2}(0, z) + \beta k T \phi(z) \right), \quad 1 \leq i \leq N
\]

where

\[
g_{i,2}(0, z) = IT \sum_{j=1}^{N} \left[ Dz_i \alpha_{i,j} z_j + T(z_i \alpha_{j,i} z_i - z_i \alpha_{i,j} z_j) \right]
\]

\[
= IT D \sum_{j=1}^{N} \left( D \alpha_{i,j} - D \alpha_{i,j} \right) z_i z_j
\]

(20)

\( \phi(z) \) is given by (18) and verifies

\[
kT^2 \phi(z) = -IT \frac{kT^2}{m + \beta k T} \sum_{i,j} \alpha_{ij} z_i z_j = -IT D \sum_{i,j} \alpha_{ij} z_i z_j
\]

Denoting \( \mu = I/D \) and \( \Theta = \frac{\beta IT D}{2(T)^2 - 1D} \), the explicit slow dynamics is then given by system (12) which may be rewritten as

\[
\frac{d}{dt} z = \Theta \cdot (\mu(A' - A)z + A'z - z'Az).
\]

(21)

Remark that we have:

\[
\frac{d}{dt} \sum_{j} z_j = \Theta \sum_{i,j} \alpha_{ij} z_i z_j \left( 1 - \sum_{j} z_j \right)
\]

(22)

which implies that the simplex \( \mathcal{P} = \{ z \in \mathbb{R}^N_+, \sum z_j = 1 \} \) is invariant under (12).

3. Characterization of the model via invasion fitnesses

We focus on the trivial solutions of (12), which are the extreme competitive exclusion equilibria \( E_i = (\delta_{ik})_{1 \leq k \leq n} \) where \( \delta_{ik} \) is the Kronecker delta: \( \delta_{ik} = \begin{cases} 1 & \text{if } i = k, \\ 0 & \text{if } i \neq k. \end{cases} \)

In particular, we note \( \lambda_i' \) the invasion fitness of the strain \( i \) with respect to the trivial solution \( E_j \). More precisely, \( \lambda_i' \) is the per-capita growth rate of the strain \( i \) evaluated at \( E_j \). Since the \( i \)-th line of the system reads \( \frac{d}{dt} z_i = z_i h_i(z_1, \cdots, z_n) \), we have the explicit formula:

\[
\lambda_i' = h_i(E_j)
\]

that is

\[
\lambda_i' = \mu (\alpha_{ii} - \alpha_{ij}) + \alpha_{ij} - \alpha_{jj}.
\]
We have that the line \( i \) of (21) is

\[
\dot{z}_i = z_i \left( \sum_{j \neq i}^N (\mu(\alpha_{ji} - \alpha_{ij}) + \alpha_{ji}) z_j + \alpha_{ii} z_i - \sum_{j=1}^N \sum_{k=1}^N \alpha_{kj} z_k z_j \right) .
\]

With the notations introduced above, this can be rewritten as:

\[
\dot{z}_i = z_i \left( \sum_{j \neq i}^N \lambda^i_j z_j + \sum_{j=1}^N \alpha_{jj} z_j - \sum_{j=1}^N \sum_{k=1}^N \alpha_{kj} z_k z_j \right) .
\] (23)

From \( \sum_{j=1}^N z_j = 1 \), we infer

\[
\sum_{j=1}^n \alpha_{jj} z_j - \sum_{k,j} \alpha_{kj} z_k = \sum_{j=1}^N z_j(\alpha_{jj} - \sum_{k=1}^N \alpha_{kj})
\]

\[
= \sum_{j=1}^N z_j(\alpha_{jj} - \sum_{k=1}^N \alpha_{kj})
\]

\[
= \sum_{j=1}^N \sum_{k=1}^N z_j(\alpha_{jj} - \alpha_{kj}) z_k
\]

finally, the fact that

\[
\sum_{j=1}^N \sum_{k=1}^N (\alpha_{jk} - \alpha_{kj}) z_j z_k = 0
\]

yields

\[
\sum_{j=1}^n \alpha_{jj} z_j - \sum_{k,j} \alpha_{kj} z_k = - \sum_{1 \leq k,j \leq N} \lambda^i_j z_j z_k .
\]

Plugging this into (23), we obtain a new version of system (21), entirely parametrized by the fitnesses \( \lambda^i_j \):

\[
\begin{cases}
\frac{d}{d\tau} z_i = \Theta z_i \cdot \left( \sum_{j \neq i}^N \lambda^i_j z_j - \sum_{1 \leq k,j \leq N} \lambda^i_j z_j z_k \right), & i = 1, \ldots, N \\
z_1 + \cdots + z_N = 1.
\end{cases}
\] (24)

or more shortly by noting \( \lambda^i_i = 0 \) and the pairwaise invasion fitness matrix \( \Lambda = (\lambda^i_j)_{1 \leq i,j \leq N} \):

\[
\begin{cases}
\frac{d}{d\tau} z = \Theta z \cdot (\Lambda z - z' \Lambda z) \\
z_1 + \cdots + z_N = 1.
\end{cases}
\] (25)

3. Supplementary Figures and Movies
**Supplementary Movie 1.** Here we illustrate the dynamic trajectory of the system (Figure 2 of the paper). A. Network of pairwise invasion fitnesses between strains, where edges can be of 4 types: coexistence (red), bistability (blue), exclusion of $i$, exclusion of $j$ (directed gray arrows toward the winner). The system trajectory $z_i$ is shown as a movement in a 6-dimensional space. B. Trajectory of the strain frequencies over time.

**Figure S1:** Illustration of more complex dynamics. Here we simulate the model in its reduced form for another set of parameters, leading to limit cycle coexistence between 5 strains in an $N = 6$ system. See associated movie S2.
Supplementary Movie 2. Here we illustrate the dynamic trajectory of the system (Figure 1 of the Supplements). A. Network of pairwise invasion fitnesses between strains, where edges can be of 4 types: coexistence (red), bistability (blue), exclusion of $i$, exclusion of $j$ (directed gray arrows toward the winner). The system trajectory $z_i$ is shown as a movement in a 6-dimensional space. B. Trajectory of the strain frequencies over time.

Figure S2: Supplementary Figure 2. Computation speed of the slow-dynamics approximation in detail. Here we plot the computation time as a function of number of strains $N$, in terms of number of seconds on a PC to simulate the slow-fast dynamic approximation (boxplots) for $T = 5000$ time units starting from the slow manifold. The line shows the mean over 20 random matrices for each $N$, where $r = 0.1$ and mean co-colonization coefficient $k = 1$ and $R_0 = 2$. 