

# **Master thesis title**

Pablo Lechón Alonso

Imperial College London

Word count: 10

## **Abstract**

# Contents

<b>1</b>	<b>Introduction notes</b>	<b>2</b>
<b>2</b>	<b>Introduction</b>	<b>4</b>
<b>3</b>	<b>Methods and Results</b>	<b>5</b>
3.1	Consumer resource model with cross feeding interactions . . . . .	5
3.2	Community Assembly . . . . .	8
3.3	A metric of community cohesion . . . . .	11
3.4	$\Theta$ predicts the outcome of community coalescence experiments. . .	12
<b>4</b>	<b>Discussion</b>	<b>12</b>
<b>5</b>	<b>Things to do in the future</b>	<b>14</b>
<b>6</b>	<b>Appendix</b>	<b>16</b>
6.1	Reversible enzyme kinetics . . . . .	16
6.2	Table of parameter values and meaning . . . . .	18

# 1 Introduction notes

- State my aims/hypotheses/questions by the end of the introduction.
- Make sure I explain everything adequately. Provide more background in the introduction and methods that
- What is the problem I am tackling? you end up looking forward to read more
- the general way the problem has been approached.
- Clearly define my aims of the research project.
- Why is it interesting? Why don't we know the answer?
- Build from the most general and fundamental hypotheses to the most refined or tenuous ones.
- How will I go about testing my hypothesis.

Following, what I want to talk about

- Microbial Community Coalescence in communities with mutualisms ie, cross-feeding.
- Thermodynamic constraints
- What drives a community to be successful in a coalescence event?
- Will a coalesced community be more persistent than a naive one upon a event if it has a history of coalescence?
- How cohesive a community is when we analyze it in terms of cohort and dominant species?
- Experimental studies report cohesiveness in community coalescence events.

- 23     • talk about why this model is good with methanogenic communities, strong  
24       cross feeding
- 25     • say at some point how many possible networks there are.
- 26     • talk about coselection
- 27     • measure community productivity and check that the correlation breaks  
28       down because there is co-selection.
- 29     • stress importance of dominant community interaction types
- 30     • microbial ensembles
- 31     • Theoretical efforts for community coalescence (Tikhonov 2016, Livingston  
32       et al. 2013, Toquenaga 1997, Gilpin 1994)
- 33     • Experimental efforts to understand community coalescence (Lu et al. 2018,  
34       Sierocinski et al. 2017)
- 35     • Community coalescence is just another expression of the fight competition  
36       vs facilitation...

## 37 2 Introduction

38 Microbial communities are widespread throughout our planet, from the deep  
39 ocean to the human gut, and they play a critical role in natural processes  
40 ranging from animal development and host health (Huttenhower et al. 2012)  
41 to biogeochemical cycles (Falkowski et al. 2008). These communities are very  
42 complex, often harboring hundreds of species, making them hard to character-  
43 ize. Recently, DNA sequencing has facilitated a high-resolution mapping of these  
44 consortia, opening a niche for ambitious theorists and experimentalists to col-  
45 laboratively disentangle the complexity of these systems (Marsland et al. 2019,  
46 Goldford et al. 2018, Goyal & Maslov 2018, Friedman et al. 2017, Costello et al.  
47 2012). One of the problems yet to solved is community assembly – the process  
48 by which species come together and interact to establish a community. Contrary  
49 to what is found in the macroscopic world, in microbial ecology it is usual that  
50 whole communities move to a region where they encounter another community.  
51 The process by which two or more communities that were previously separated  
52 join and reassemble into a new community has been termed ”community coales-  
53 cence” (Rillig et al. 2015). This type of event happens repeatedly in nature due  
54 to abiotic (wind, tides or river flow), biotic (animal courtship, parent-offspring  
55 interactions or leaves falling), and anthropogenic (industrial anaerobic digestion,  
56 agriculture, between-human contact) factors (Castledine et al. 2020). Despite the  
57 frequency and importance of microbial community coalescence, the mechanisms  
58 responsible for the community structure and function resulting from coalescence  
59 events remain poorly understood.

60 Early mathematical models of community-community invasion revealed that  
61 when two communities previously separated by a barrier merge due to its re-  
62 moval, asymmetrical dominance of one community over the other one is likely  
63 to occur (Gilpin 1994, Toquenaga 1997). As an explanation for this observa-  
64 tion, it was argued that, because communities have been assembled through a

65 history of competitive exclusion, they are likely to compete with each other as  
 66 coordinated entities, rather than as a random collection of species. More recent  
 67 theoretical work uses consumer-resource models to show that coalescing micro-  
 68 bial communities exhibit an emergent cohesiveness (Tikhonov 2016, Tikhonov  
 69 & Monasson 2017). =====These, have been hypothesized to be essential  
 70 for the formation of Metabolically Cohesive [microbial] Consortium (MeCoCos)  
 71 (Pascual-García et al. 2020)===== . Recent results from coalescence exper-  
 72 iments of methanogenic communities suggest that during a coalescence event  
 73 between two communities, multiple taxa from the same community act as co-  
 74 hesive units and are selected together (ecological co-selection) (Sierocinski et al.  
 75 2017). Further experimental evidence of co-selection in community coalescence  
 76 has been reported in Lu et al. (2018), where it was shown that successful col-  
 77 lective invasions are accompanied by strong community-level interactions. The  
 78 microbial communities used in these experiments are characterized by complex  
 79 cross-feeding interactions (Hansen et al. 2007, Lawrence et al. 2012, Embree  
 80 et al. 2015). Furthermore, the type of trophic interactions present in a commu-  
 81 nity has been suggested as a factor that might affect the outcome of community  
 82 coalescence (Castledine et al. 2020). Yet, theoretical models used in community  
 83 coalescence studies so far have considered competition between species as the  
 84 only force driving community assembly.

85 In this work, I explore the role of other types of interactions, which appear to be  
 86 ubiquitous in microbial communities. Specifically, I propose a metric of commu-  
 87 nity cohesion that accounts for both competitive and mutualistic interactions. I  
 88 then use a consumer-resource model that includes both facilitation of metabo-  
 89 lites via by-product secretion, and competition for substrates, to simulate many  
 90 instances of community assembly . Finally, I measure the cohesion level on the  
 91 simulated communities and use it to predict the outcome of microbial community  
 92 coalescence events.

### 93 3 Methods and Results

#### 94 3.1 Consumer resource model with cross feeding interactions

95 In order to simulate communities with cross-feeding interactions, I use a consumer-  
 96 resource model inspired in the the work of (Marsland et al. 2019) that PhD stu-  
 97 dent Jacob Cook has developed. I modify parts of this model to make it more  
 98 suitable for my purposes.

99 I consider the population dynamics of  $s$  consumers (eg. bacterial strains) that  
 100 feed on  $m$  resources. In this model, a species is defined by its metabolic stratetgy  
 101 to harvest energy from the environment. Let  $G_\alpha(\mathcal{M}, \mathcal{N})$  be the metabolic net-  
 102 work of species  $\alpha$ , where  $\mathcal{M}$  is a set of nodes  $\mathcal{M} = \{x : x \text{ is an integer from}$   
 103  $\text{the interval } [1, m] \text{ labeling the metabolite}\}$  and  $\mathcal{N}$  a set of uni-directed edges  $\mathcal{N}$   
 104  $= \{(x, y) : x \in \mathcal{M}, y \in \mathcal{M} \text{ and } x < y \text{ (x and y are product and substrate of}$   
 105  $\text{a chemical reaction})\}$ . The growth power of species  $\alpha$ ,  $J_\alpha^{grow}$  will be given by  
 106 the product of the amount of generated energy  $\eta_i$  and rate  $q_i$  of each reaction,  
 107 summed across all reactions in  $\mathcal{N}$ .

$$108 \quad J_\alpha^{grow} = \sum_{i=1}^{|\mathcal{N}|} q_i \eta_i \quad (1)$$

109 where  $|\cdot|$  denotes cardinality of a set. Refer to subsection 6.1 for specifications  
 110 on  $q$  and  $\eta$ .

111 Every species has a maintenance cost  $\chi_\alpha$  that represents the required energy to  
 112 sustain life, and I model as

$$113 \quad \chi_\alpha = \chi_0 \sum_{\mathcal{N}} (y - x) \quad (2)$$

114 where  $\chi_0$  is the average cost per reaction, and the summatory term adds up  
 115 the metabolite gap of all reactions. Therefore, the more reactions a speciess  
 116 has or/and the more energetically they are, the higher is the maintenance cost.



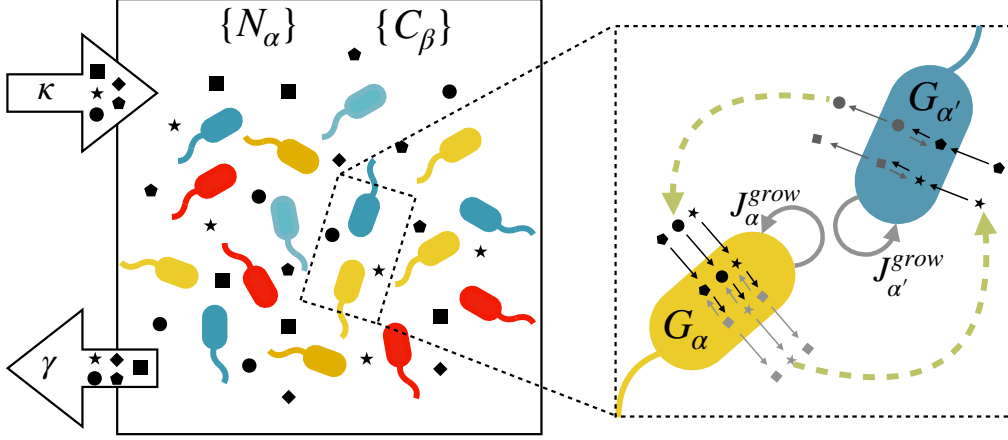


Figure 1: **Schematic of the model.** (left) All  $m$  metabolites are steadily supplied into a chemostat with different bacterial strains at rate  $\kappa$  and diluted at rate  $\gamma$ . (right) Bacteria use the metabolites in the environment,  $\{C_\beta\}$  through their reaction networks  $G_\alpha$ , to obtain the necessary power  $J_\alpha^{grow}$  to increase their abundance  $\{N_\alpha\}$  by replicating. The green arrows in the magnified portion emphasize that species  $\alpha$  (yellow) facilitates metabolites to species  $\alpha'$  (blue) and viceversa. The double arrows in the reactions happening inside the cells stand for the reversible enzyme kinetics considered by this model.

117 The cost function (Eq. 2) ensures that neither generalists, nor specialists, are  
 118 systematically favored during the community assembly.

119 Under this parametrization, the time evolution of the population of species  $\alpha$   
 120 can be written as

$$121 \quad \frac{dN_\alpha}{dt} = g_\alpha N_\alpha [J_\alpha^{grow} - \chi_\alpha] \quad (3)$$

122 where  $g_\alpha$  is a proportionality constant relating energy to abundance of strain  $\alpha$   
 123 The dynamics of the resources depend on the incoming and outgoing resource  
 124 fluxes due to the biochemical reactions taking place inside bacteria, as well as  
 125 the resource external dynamics. The incoming resource flux of metabolite  $\beta$   
 126 generated by strain  $\alpha$  is its rate of consumption due to all the biochemical reac-  
 127 tions possessed by  $\alpha$  in which  $\beta$  is a substrate. The outgoing flux is that due to

128 reactions in which  $\beta$  is a product.

$$\begin{aligned}
 v_{\alpha\beta}^{in} &= \sum_{\mathcal{S}} q & \text{with } \mathcal{S} &\equiv N \cap \{(x = \beta, y)\} \\
 v_{\alpha\beta}^{out} &= \sum_{\mathcal{P}} q, & \text{with } \mathcal{P} &\equiv N \cap \{(x, y = \beta)\}
 \end{aligned}
 \tag{4}$$

131 The external resource dynamics are modelled as a supply rate minus a dilution  
 132 rate that depends on the resource concentration to ensure convergent dynamics.

$$133 \quad h_{\beta} = \kappa - \gamma C_{\beta} \tag{5}$$

134 Therefore, the variation with time of the concentration of metabolite  $\beta$  has the  
 135 form

$$136 \quad \frac{dC_{\beta}}{dt} = h_{\beta} + \sum_{\alpha=1}^s (v_{\alpha\beta}^{in} - v_{\alpha\beta}^{out}) N_{\alpha} \tag{6}$$

137 Thus, the model consists of  $s + m$  coupled differential equations completely  
 138 specified by Eqs. 3 & 6.

### 139 **3.2 Community Assembly**

140 Armed with this model I now simulate  $n_s = 2 \cdot 10^3$  community assembly events  
 141 of  $s = 10$  species in an environment with  $m = 15$  substrates that are steadily  
 142 supplied.

143 The values of the of the parameters of the model (subsection 6.2, table 1) remain  
 144 constant throughout all simulations. The reason for this is that, my aim is not  
 145 to parametrize the model to reveal large-scale patterns found in experiments  
 146 (although that would be a fruitful endeavour because of the rich parameter space  
 147 of this model). Rather, I use it to simulate a set of microbial communities with  
 148 cross-feeding interactions that will be later used in the community coalescence  
 149 experiments.

150 In order to do so, I first create  $s \cdot n_s$  random reaction networks,  $G_{\alpha}(\mathcal{M}, \mathcal{N})$  using

the following procedure. Consider, the  $m \times m$  adjacency matrix  $A_{ij}^\alpha$ , whose elements represent the edges  $(i, j)$  of  $G_\alpha$ . Since the reaction network is hierarchical ( $i < j$ , subsection 6.1), the adjacency matrix is an upper triangular matrix with zeros in the main diagonal, and the reactions possessed by strain  $\alpha$  can be expressed as  $(i, i + k)$ , where  $k$  represents the  $k^{th}$  diagonal of  $A$  ( $k = 1, \dots, m - 1$  with  $k = 0$  being the main diagonal), and  $i$  is the row number of one of its elements ( $i = 1 \dots m$ ). The reaction network  $G_\alpha$  is constructed by sampling  $n_r$  pairs  $(i, k)$  according to the algorithm summarized below.

1. Choose  $n_r$  by sampling it from a uniform distribution  $U(1, m)$
2. Choose  $k$  by sampling one value from truncated normal distribution  $N(1, \sqrt{m - 1})$  with limits  $[1, m - 1]$ , and rounding it to the closest integer.
3. Sample  $i$  from a uniform distribution of integers  $U(0, m - k)$ .
4. The reaction  $(i, i + k)$  is stored, and the process is repeated until  $n_r$  reactions have been sampled.

Some notes about this algorithm are, first, sampling  $k$  from a truncated normal distribution ensures that high metabolite gaps (very energetic reactions) are not likely to happen. This introduces a bias against the presence of super-organisms with few and very energetic reactions, which are rare in microbial communities. Second, the truncation limits in step 2 have been chosen to respect the hierarchical character of the network:  $k \neq 0$  to avoid reactions of the form  $(i, i)$ . Third, the upper limit of the uniform distribution from which  $i$  is sampled is bounded by  $k$ , the diagonal we are sampling from.

When the sampling of reaction networks is completed, equations 6 and 3 are integrated using a Runge Kutta method (Dormand & Prince 1980) with initial conditions  $N_\alpha(0) = 2$  and  $C_\beta(0) = 0$ . Relevant results stemming from the simulations of community assembly events are plotted in figure 2

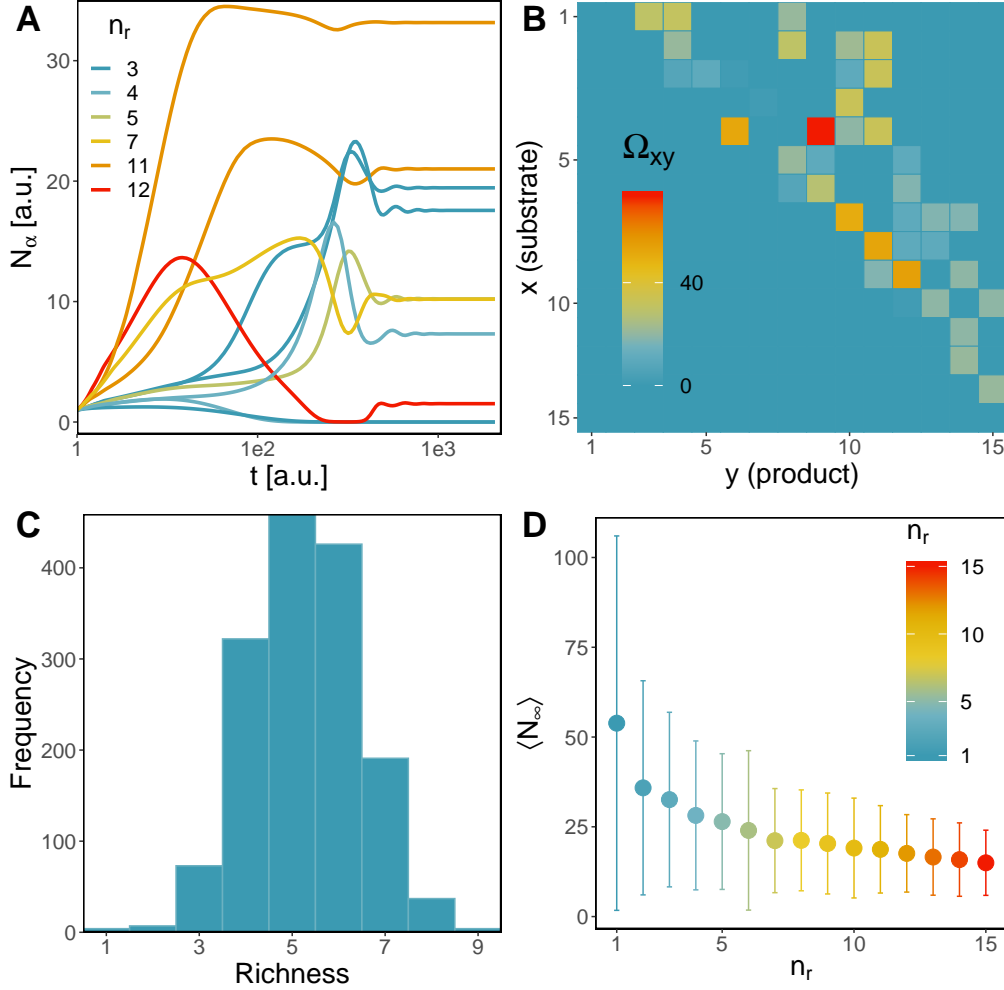


Figure 2: **Results from community assembly simulations.** Plots (A) and (B) exemplify one community assembly event and (C) and (D) convey results across simulations. (A) Time variation of species' abundance for one instance of community assembly with  $m = 15$ ,  $s = 10$ , and a set of  $s$  randomly generated reaction networks. Time (x-axis) and population (y-axis) are measured in arbitrary units. Each time series is coloured according to  $n_r$ , the number of reactions possessed by the reaction network of each strain. (B) Community reaction network, obtained by summing the adjacency matrices of all species weighted by their respective carrying capacity:  $\Omega = \sum_{k=1}^s N_\infty^k A_k$ . (C) Histogram of richness of the  $n_s$  simulations. (D). Mean value of carrying capacity  $\pm \sigma$  (error bars) against the number of reactions  $n_r$ : species with less number of reactions (specialists) are more abundant than those with higher  $n_r$  (generalists); several specialists deplete all resources through their combined action more efficiently than one generalist.

### 178 3.3 A metric of community cohesion

179 Following the ensemble of many synthetic communities, I postulate a metric  
180 of community cohesion that is later used to predict the outcome of community  
181 coalescence.

182 Let  $s_1$  and  $s_2$  be two sequences of integers labeling metabolites. I am interested  
183 in measuring their *overlapping degree*  $\xi(s_1, s_2)$ , ie, the proportion of metabolites  
184 of  $s_1$  that intersect with  $s_2$  summed with the proportion of metabolites of  $s_2$   
185 that intersect with  $s_1$ , normalized to 1.

$$186 \quad \xi(s_1, s_2) = \frac{1}{2} \sum_{k \in s_1 \cap s_2} \left( \frac{D_{s_1}(k)}{|s_1|} + \frac{D_{s_2}(k)}{|s_2|} \right) \quad (7)$$

187 Here,  $k$  takes the values in the set that results from intersecting  $s_1$  and  $s_2$ .  $D_s(k)$   
188 is the number of elements from the sequence  $s$  that are equal to  $k$ . The purpose  
189 of all denominators in equation 7 is to normalize  $\xi$  to 1. Note that fractions  
190 inside the summation term are equal only if the number of elements in both se-  
191 quences  $s_1$  and  $s_2$  are the same.

192 The cohesion of a community – the degree to which species in a community  
193 collaborate with each other, can be thought of as the difference *cohesion* = *facil-*  
194 *itation* – *competititon*. One way to capture the facilitation of a community is by  
195 calculating its facilitation matrix  $F$ , which is composed of the facilitation indices  
196 of all possible ordered pairs  $i, j$  of species in the community. Precisely, the facili-  
197 tation index  $f_{ij}$  of species  $i$  towards species  $j$ , is given by the overlapping degree  
198 of the sequence of products of species  $i$ ,  $y_i$ , with the sequence of substrates of  
199 species  $j$ ,  $x_j$ . Equivalently, the competition matrix  $C$  gathers the competition  
200 level of the community. The competition index between species  $i$  and  $j$ ,  $c_{ij}$  is  
201 given by the overlapping degree of the sequence of substrates of species  $i$ ,  $x_i$  and

the sequence of substrates of species  $j$ ,  $x_j$ . Thus,

$$f_{ij} = \begin{cases} \xi(y_i, x_j) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \quad c_{ij} = \begin{cases} \xi(x_i, x_j) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \quad (8)$$

Note that facilitation is directional but competition is not. This implies that  $f_{ij} \neq f_{ji}$  and  $F$  is not symmetric, but  $c_{ij} = c_{ji}$  and  $C$  is symmetric.

Community-level cohesion can now be formally defined as the mean value of the matrix  $F - C$

$$\Theta = \langle F - C \rangle \quad (9)$$

### 3.4 $\Theta$ predicts the outcome of community coalescence experiments.

Following the community assembly and consequent measure of cohesion in all communities, I perform community coalescence experiments to test the predictive power of my metric.

First, I select the  $N$  communities with 5 species, and perform all  $\binom{N}{2}$  community coalescence events in which a resident community  $\mathcal{C}_R$  is mixed with an invading one  $\mathcal{C}_I$ . At each event, calculate the similarity of between post-coalescence and resident communities as the normalized scalar product of their species abundance vector at stable state.

$$S(\mathcal{C}_R, \mathcal{C}_P) = \frac{\vec{N}_\infty^R \cdot \vec{N}_\infty^P}{\sqrt{|\vec{N}_\infty^R|} \sqrt{|\vec{N}_\infty^P|}} \quad (10)$$

## 4 Discussion

- Although it certainly seems an exaggeration to view the community as such a tightly regimented entity, it is also perilous to ignore the fact that

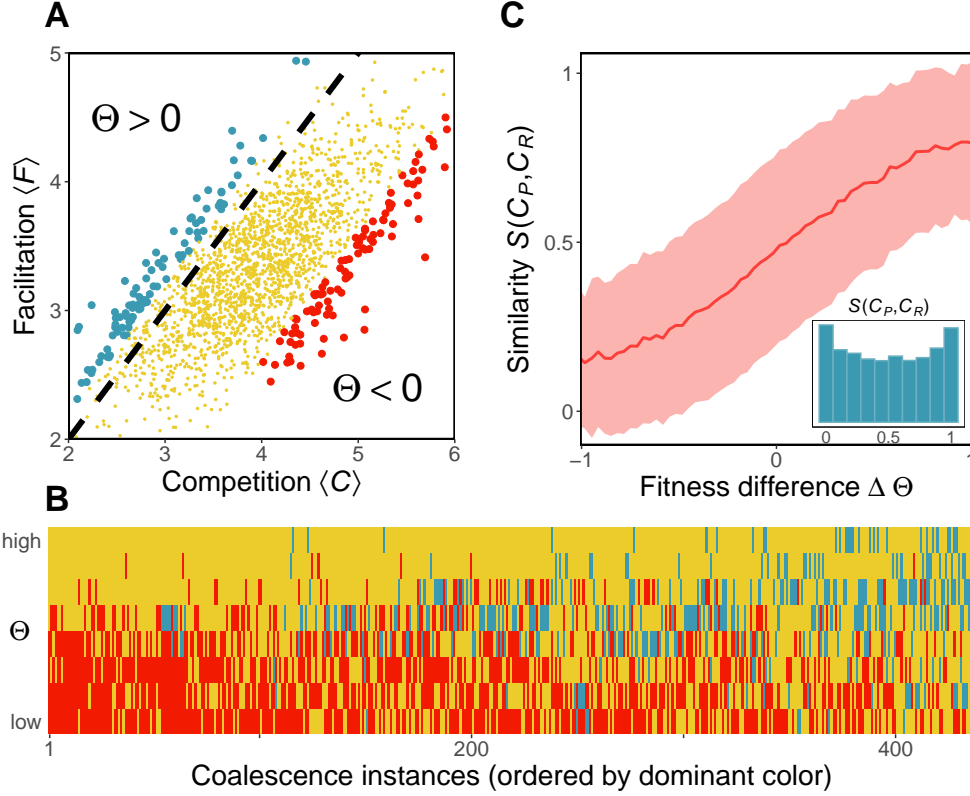


Figure 3: **Results from community coalescence experiments.** (A) Each simulated community is plotted in a competition-facilitation diagram. Communities in blue (upper region) have  $\Theta > 0$ ; they are in the facilitation-dominated regime, and communities in red (lower region) have  $\Theta < 0$ ; they belong to the competition-dominated regime. Coalescence experiments where one community from the blue group mixes with one community from the red group are performed, only for communities of richness 5. Dashed line  $\langle C \rangle = \langle F \rangle$  is plotted for reference. (B) Altruistic communities outperform competitive communities in the latter experiments. In this elimination assay, each column represents one coalescence instance, and each element in a column is a species. Extinctions are coloured to match the group in plot A to which the extinct species belonged (blue, or red). There is a higher proportion of extinct species from the red group (more red tiles than blue tiles). (C) The outcome of community coalescence is predicted by community-level cohesion. The similarity between the post-coalescence community and the resident community,  $S(C_P, C_R)$  is plotted as a function of the community cohesion difference  $\Delta\Theta$  between them, for all possible coalescence events between 2 communities of richness 4. The greater the difference in community-level cohesion, the more similar is  $C_P$  to its more cohesive parent community. (C, inset) Histogram of similarity showing that monodominance of one community after coalescence ( $S = 0$ ,  $S = 1$ ) is more frequent than a perfect mixing ( $S = 0.5$ )

- 224 coevolutionary processes can play an important role in communities.
- 225 • Cooperation reduces community stability, tho increases community fit-  
226 ness... (Coyte et al. 2015)
  - 227 • This thesis addresses the question of what are truly the mechanisms ex-  
228 plaining what experiments show? An alternative measure of cohesiveness  
229 that stems from a more realistic modeling of microbial ecosystems is able  
230 to reproduce this results, and thus is closer to uncover what are the real  
231 mechanisms behind community cohesion.
  - 232 • community coalescence is a way to explicitly show and test the cohesiveness  
233 of microbial communities while asking questions about how these commu-  
234 nities came to be.
  - 235 • use Pascual-García et al. (2020) for the discussion
  - 236 • discuss why I chose  $m$  as the upper limit for the number of reactions that  
237 a strain can posses.
  - 238 • Discuss possible refinements of the measure of cohesion: instead of aver-  
239 aging, I could count the number of closed loops.
  - 240 • Discuss why the traditional fitness (how fast resources are consumed) doesn't  
241 correlate with what I call fitness: community cohesion. Show that in the  
242 case of pure competition, it does (Tikhonov 2016), but in the case of purely  
243 facilitation, it doesn't
  - 244 • My measure of cohesion is an aproximate one. Does facilitation help the  
245 same degree that competition bothers?

## 246 5 Things to do in the future

- 247 • Ask Emma about papers of hierarchy of metabolites



- 248     • Find a paper that says that organisms with few and very energetic reactions  
249       are rare.
- 250     • Should I include a page at the end specifying what things did I do, and  
251       what things didn't I do, and that way I don't have to do it during the  
252       paper?
- 253     • make a nice looking table of the paramters of the model.
- 254     • plot mean abundance as a function of number of reactions for both assebled  
255       communities
- 256     • List of plots I want to make: 1. The s plot with richness == 5. 2. The  
257       histogram of similarity with richness == 4. 3. The elimination asssay. 4.  
258       The cloud. 5. The evolution of cooperation. 6. The community reaction  
259       network.
- 260     • Revise the cohesion of my thesis as a whole: are sections well separated?  
261       Do they link well with each other?

## 262 6 Appendix

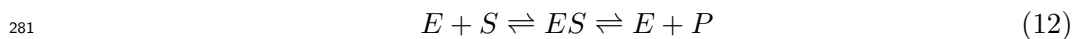
### 263 6.1 Reversible enzyme kinetics

264 Outside the the bacterial cell, the energy resides in the form of chemical potential  
265  $\mu$  held by the metabolites, and biochemical reactions inside the cell produce  
266 energy due to a difference in the chemical potentials of substrate and product. I  
267 assigned chemical potentials to each metabolite according to

$$268 \quad \mu_\beta = E \left( 1 - \sqrt{\frac{\beta - 1}{m - 1}} \right) \quad (11)$$

269 where  $\beta = 1, \dots, m$  and  $E$  is the energy of the most energetic metabolite. I have  
270 chosen this chemical potential function because I hope to find papers where they  
271 explain that there is a hierarchy on the metabolite energetic spectrum. This  
272 means that the energy produced by a reaction of the type  $(\beta, \beta + 1)$  decreases as  
273 you go down the hierarchy. Reactions involving metabolites situated higher in  
274 the hierarchy are more energetic than reactions that involve those lower in the  
275 hierarchy.

276 The rate at which a given chemical reaction transforms substrate into product  
277 is modeled using reversible Michaelis-Menten enzyme kinetics. Thus, the model  
278 considers chemical reactions where a substrate  $S$  binds to an enzyme  $E$  to form  
279 an enzyme-substrate complex  $ES$ , which in turn produces a product  $P$ , and  
280 recovers enzyme  $E$ .



282 The choice of fully reversible enzyme kinetics, instead of the traditional assump-  
283 tion of irreversibility in the second reaction, aims to capture more accurately the  
284 nature of biochemical reactions taking place in microbial communities. In these  
285 reactions the Gibbs energy change  $\Delta G$  is not always high, which implies that  
286 the reaction of product formation can reach equilibrium at a similar time scale  
287 as the formation of the complex (Keener & Sneyd 2008). In this case, the tradi-

288 tional irreversible Michaelis-Menten scheme breaks down, and more elaborated  
 289 frameworks, like the fully reversible one that this model offers, need to be used.  
 290 To comply with  $2^{nd}$  law of thermodynamics, the network  $G_\alpha$  is completely hier-  
 291 archical, ie. the edges are unidirectional ( $x < y$ ), going from the more energetic,  
 292 to the less energetic metabolite. Thus, for the reaction scheme in 12 and the  
 293 imposed thermodynamic constraint only reactions where  $\Delta G^0 = \mu_P - \mu_S < 0$   
 294 can take place.  
 295 With all the above considerations, the expression for the rate of reaction  $i$  poss-  
 296 esed by strain  $\alpha$  is given below. A formal derivation of equation 13 can be found  
 297 in Hoh & Cord-Ruwisch (2000)

$$298 \quad q_{\alpha i} = \frac{q_m^{\alpha i} S_\alpha (1 - \theta_\alpha)}{K_S^{\alpha i} + S_\alpha (1 + k_R^{\alpha i} \theta_\alpha)} \quad (13)$$

299 Here,  $\theta_\alpha$  measures how far is the reaction from equilibrium (0 being the furthest,  
 300 and 1 being equilibrium).

$$301 \quad \theta = \frac{[P]}{[S]K_{eq}} \quad (14)$$

302 where  $[ ]$  denote concentration and  $K_{eq}$  is the equilibrium constant

$$303 \quad K_{eq} = \exp \left( \frac{-\Delta G^0 - \eta \Delta G_{ATP}}{RT} \right) \quad (15)$$

304 The energy produced by the reaction is then stored in the form of ATP molecules.  
 305 In the model,  $\eta$  represents the moles of ATP molecules produced per mole of  
 306 reaction. For a given reaction  $(x, y)$  eta I calculate eta as

$$307 \quad \eta = \frac{y - x}{m} \quad (16)$$

308 which represents the normalized metabolite gap between substrate and product  
 309 of the reaction. Therefore, the higher the gap, the more energy will be stored.

310

311 **6.2 Table of parameter values and meaning**

Parameter	Meaning	Value
m	Number of metabolites	100
s	Number of strains	10
$\Delta G_{ATP}$	ATP Gibbs energy	$7.5 \cdot 10^4$
$\mu_0$	Most energetic metabolite	$3 \cdot 10^4$
nATP	$\max\left(\frac{\Delta G_{S \rightarrow P}^0}{\Delta G_{ATP}}\right)$	4
$\eta$	Moles of ATP energy per reaction	0.5
$q_m$	Maximum reaction rate	1
$K_S$	Saturation constant	0.1
$k_r$	Reversibility constant	10
g	Growth factor	1
m	Maintenance factor	$0.2 \cdot J_{grow}$
$\kappa$	Externally supplied resource	1
$\gamma$	Dilution rate	0.5
$N_0$	Populations initial conditions	(1, 1, ..., 1)
$C_0$	Concentrations initial condition	(0, 0, ..., 0)

Table 1: Parameter meanings and their values

## References

- Castledine, M., Sierocinski, P., Padfield, D. & Buckling, A. (2020), ‘Community coalescence: An eco-evolutionary perspective’.
- Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J. & Relman, D. A. (2012), ‘The application of ecological theory toward an understanding of the human microbiome’.
- Coyte, K. Z., Schluter, J. & Foster, K. R. (2015), ‘The ecology of the microbiome: Networks, competition, and stability’, *Science* .
- Dormand, J. R. & Prince, P. J. (1980), ‘A family of embedded Runge-Kutta formulae’, *Journal of Computational and Applied Mathematics* .
- Embree, M., Liu, J. K., Al-Bassam, M. M. & Zengler, K. (2015), ‘Networks of energetic and metabolic interactions define dynamics in microbial communities’, *Proceedings of the National Academy of Sciences of the United States of America* .
- Falkowski, P. G., Fenchel, T. & Delong, E. F. (2008), ‘The microbial engines that drive earth’s biogeochemical cycles’.
- Friedman, J., Higgins, L. M. & Gore, J. (2017), ‘Community structure follows simple assembly rules in microbial microcosms’, *Nature Ecology and Evolution* .
- Gilpin, M. (1994), ‘Community-level competition: Asymmetrical dominance’, *Proceedings of the National Academy of Sciences of the United States of America* .
- Goldford, J. E., Lu, N., Bajić, D., Estrela, S., Tikhonov, M., Sanchez-Gorostiaga, A., Segrè, D., Mehta, P. & Sanchez, A. (2018), ‘Emergent simplicity in microbial community assembly’, *Science* .

- Goyal, A. & Maslov, S. (2018), ‘Diversity, Stability, and Reproducibility in Stochastically Assembled Microbial Ecosystems’, *Physical Review Letters* .
- Hansen, S. K., Rainey, P. B., Haagenzen, J. A. & Molin, S. (2007), ‘Evolution of species interactions in a biofilm community’, *Nature* .
- Hoh, C. Y. & Cord-Ruwisch, R. (2000), ‘A practical kinetic model that considers endproduct inhibition in anaerobic digestion processes by including the equilibrium constant’, *Biotechnology and Bioengineering* .
- Huttenhower, C., Gevers, D., Knight, R. & Al., E. (2012), ‘Structure, function and diversity of the healthy human microbiome’, *Nature* .
- Keener, J. P. & Sneyd, J. (2008), ‘Mathematical Physiology’, *Book* .
- Lawrence, D., Fiegna, F., Behrends, V., Bundy, J. G., Phillimore, A. B., Bell, T. & Barraclough, T. G. (2012), ‘Species interactions alter evolutionary responses to a novel environment’, *PLoS Biology* .
- Livingston, G., Jiang, Y., Fox, J. W. & Leibold, M. A. (2013), ‘The dynamics of community assembly under sudden mixing in experimental microcosms’, *Ecology* .
- Lu, N., Sanchez-gorostiaga, A., Tikhonov, M. & Sanchez, A. (2018), ‘Cohesiveness in microbial community coalescence’, *bioRxiv* .
- Marsland, R., Cui, W., Goldford, J., Sanchez, A., Korolev, K. & Mehta, P. (2019), ‘Available energy fluxes drive a transition in the diversity, stability, and functional structure of microbial communities’, *PLoS Computational Biology* .
- Pascual-García, A., Bonhoeffer, S. & Bell, T. (2020), ‘, which can be found in subsection 3.3 (page 11).’.

- Rillig, M. C., Antonovics, J., Caruso, T., Lehmann, A., Powell, J. R., Veresoglou, S. D. & Verbruggen, E. (2015), ‘Interchange of entire communities: Microbial community coalescence’.
- Sierocinski, P., Milferstedt, K., Bayer, F., Großkopf, T., Alston, M., Bastkowski, S., Swarbreck, D., Hobbs, P. J., Soyer, O. S., Hamelin, J. & Buckling, A. (2017), ‘A Single Community Dominates Structure and Function of a Mixture of Multiple Methanogenic Communities’, *Current Biology* .
- Tikhonov, M. (2016), ‘Community-level cohesion without cooperation’, *eLife* .
- Tikhonov, M. & Monasson, R. (2017), ‘Collective Phase in Resource Competition in a Highly Diverse Ecosystem’, *Physical Review Letters* .
- Toquenaga, Y. (1997), ‘Historicity of a simple competition model’, *Journal of Theoretical Biology* .