Classification: Biological Sciences / Microbiology, Systems Biology

Title: Short-range interactions govern cellular dynamics in microbial multi-genotype systems

Authors: A. Dal Co[a*], S. van Vliet[a,b], D. J. Kiviet[a], S. Schlegel[a], M. Ackermann[a]

Affiliations:

[a]*Department of Environmental Systems Science, ETH Zurich, and Department of Environmental Microbiology Eawag, Ueberlandstrasse 133PO Box 611, 8600 Duebendorf, Switzerland*

[b]*Department of Zoology, University of British Columbia, 4200-6270 University Blvd., V6T 1Z4 Vancouver, Canada*

*Corresponding author: alma.dalco@gmail.com

Keywords: microbial interactions, synthetic communities, metabolic exchange

Abstract

Ecosystem processes result from interaction between organisms. When interactions are local, the spatial organization of organisms defines their network of interactions, and thus influences the system's functioning. This can be especially relevant for microbial systems, which often consist of spatially structured communities of cells connected by a dense interaction network. Here we measured the spatial interaction network between cells in microbial systems and identify the factors that determine it. Combining quantitative single-cell analysis of synthetic bacterial communities with mathematical modeling, we find that cells only interact with other cells in their immediate neighbourhood. This short interaction range impacts the functioning of the whole system by reducing its ability to perform metabolic processes collectively. Our experiments and models demonstrate that the spatial scale of cell-to-cell interaction plays a fundamental role in understanding and controlling natural communities, and in engineering microbial systems for specific purposes.

Significance Statement

Communities of interacting microbes perform fundamental processes on earth. We do not understand well how these processes emerge from the interactions between individual microbial cells. Our work investigates how strongly individual cells interact and how the strength of the interaction depends on the distance between cells. The discovery that individual cells 'live in a small world', i.e. they only interact with a small number of cells around them, changes our understanding of how cells in natural microbial communities are metabolically coupled and how their spatial arrangement determines emergent properties at the community level. Our quantitative single-cell approach allows to address central questions on systems composed of interacting genotypes and to increase our understanding and ability to control microbial communities.

Text

Microbial systems perform processes that ultimately sustain all of life (1, 2). These processes are often based on metabolic interactions between different types of organisms (3–5) and are thus an emergent property of microbial systems. Metabolic interactions are often essential because many microorganisms are unable to perform all anabolic functions required for life and thus need to obtain compounds from other

species (6–8). These metabolic interactions are expected to decay with the distance between individual cells (9), which thus determines whether cells can or and cannot interact. Little is known about the spatial scale at which cells interact and the biological parameters that determine this scale. The spatial scale of interaction is a fundamental property of any microbial system consisting of multiple genotypes because it defines the network of interactions that can occur.

Results and Discussion

The spatial scale over which individual cells interact can have a strong influence on cellular dynamics, as we show with a simple simulation of multiple genotypes that need to exchange essential metabolites in order to grow (Fig. 1a). This simulation reveals that the reproductive success of individuals is lower when interactions are limited to immediate neighbours (Fig. 1b) because often these neighbours are their kin. This effect becomes stronger when organisms depend on metabolites from more than one other cell type; a small range of interaction reduces the probability that all required partners are present in the interaction neighbourhood (Fig. 1b). These results indicate that collective metabolism might be hindered when interactions are short-ranged.

Our first goal was thus to quantify the spatial range of interaction between cells in a microbial system. We measured this range by growing a synthetic consortium in microfluidic chambers (Fig. 2b, Sup. Video S1) and quantifying how the growth rate of cells depends on the identity of their neighbours (Fig. 2c, Sup. Video S2). The synthetic consortium (Fig. 2a) is composed of two auxotrophic *Escherichia coli* strains that are unable to synthesize proline and tryptophan, respectively. Because cells naturally leak out amino acids (10, 11) a pair of auxotrophs can grow together by exchanging amino acids. In the absence of biological activity, amino acids would diffuse throughout the microfluidic chamber in less than a second to yield a homogenous environment; however cells can modify the local amino acid concentration and thereby influence their neighbours' growth.

To quantify the spatial range at which cells interact, we correlated the growth rates of individual auxotrophic cells with the presence of their complementary partner in the neighbourhood. Specifically, we measured the fraction of the complementary partner within a distance *d* from a cell's membrane, and determined the value of *d* that leads to the best prediction of the cell's growth rate. We call this distance the *interaction range*.

This analysis revealed that the interaction range is on the order of one cell length (Fig. 3a). This is found consistently across ten biological replicates (~15,000 cells analyzed in total). Specifically, the interaction range of the tryptophan auxotroph cells is 3.2±0.4 μm (mean ± standard error of the mean), while the interaction range of the proline auxotroph cells is 12.1±0.5 μm (significantly larger, $p<10^{-5}$, paired t-test, n=10). In other words, these cells live in a small world: they are only affected by a small group of individuals around them.

Our next goal was to identify the mechanisms that explain the small range of interaction. We addressed this question with an individual-based model (Fig. 4a). We assumed that the growth rate of auxotrophic cells is limited by the uptake of the amino acid that they need, and that cells take up amino acids actively and leak them passively. All model parameters are taken from literature, apart from the two leakage rates, which are estimated from the data (Supplementary Information). Our model predicts the individuals' growth given the combined activities of all the cells in the system and captures how growth depends on neighbourhood composition (Fig. 4c) and at what spatial scale (proline auxotroph: 15.0 μm model, 12.5 μm data, tryptophan auxotroph: 3.0 μm model, 3.2 μm data, Fig. 4b).

The model shows that the interaction range is set by fundamental biochemical parameters. In fact, the interaction range is directly proportional (Fig. 4e) to a second length scale, which we can calculate analytically: the *growth range*, the length scale describing the decrease in growth when the auxotrophs are arranged as in Fig. 4d. One can show that the growth range, and thus the interaction range between cells, depends more strongly on the amino acid uptake rate (square root dependence) than on the leakage rate (logarithmic dependence, Supplementary Information). The interaction range is generally small in systems where leakage is slow and uptake of

the exchanged compounds is fast compared to their diffusion (Fig. 4f). As the diffusion constant does not vary significantly for small molecules such as amino acids ($D_{\Delta trpC}/D_{\Delta proC} \sim 0.75$), the interaction range is primarily modulated by the uptake rates ($r^u_{\Delta trpC}/r^u_{\Delta proC} \sim 16$). Our modeling framework can be adapted to estimate the interaction range in other multi-genotype systems where the molecules mediating the interactions are taken up or degraded by individuals.

Next, we considered how this small interaction range affects community dynamics. Our communities show consistent dynamics: within about 25 hours all replicate communities reach a steady state composition where the tryptophan auxotroph is in minority (median fraction of total biomass = 0.23, Fig. 5a). This is in line with the growth dynamics that we measured at the single-cell level: to reach an equal growth rate of e.g. 0.1 per hour, the tryptophan auxotrophs need ~90% of the complementary partner within a small neighbourhood, while the proline auxotrophs need four times less within a much larger neighbourhood (Fig. 3c). The single cells' properties thus determine the community steady state.

Does the small interaction range that we measured in our experimental system limit the rate at which individual cells can grow? This question brings us back to our central hypothesis, that short-range interactions limit the exchange of resources and hinder collective metabolism. One would expect the individual cells in the multi-genotype system to grow faster if they would increase their mixing or their interaction range. We tested these predictions by applying our model to experimentally observed spatial arrangements. Specifically, the predicted average growth rate of individuals, and therefore the growth rate of the community as a whole, is higher when we randomize arrangements and disrupt kin clusters (Fig. 5b) or when we simulate an increase in the interaction range by lowering the uptake rates of amino acids (Fig. 5c).

While we expect that the spatial scale of interaction fundamentally affects the functioning and dynamics of any microbial multi-genotype system, the specific effects will depend on the nature of the interactions. For example, short-range interactions can stabilize the cooperative production of molecules, as they ensure that these molecules are only accessible to cells that also contribute to production, and inaccessible to non-producing individuals (12, 13). In contrast, short-range

interactions generally harm mutualistic cross-feeding communities (14), although they can have a stabilizing effect by preventing ecological invasion by non-contributing mutants (15).

The ecological and evolutionary outcome of cooperation and competition can change dramatically when interactions are limited to the local neighbourhood (9, 16, 17). Quantifying the interaction range and linking it to biochemical parameters of the system is therefore essential for understanding the functionality and dynamics of microbial multi-genotype systems. We found that interaction ranges are small whenever the uptake or degradation of the molecules mediating the interaction is fast compared to their diffusion in the environment, and when the density of individuals is high. We thus expect interactions to be local in dense microbial systems where cells interact through the exchange of cellular building blocks, quorum sensing molecules and metabolites that bind (18) or digest extracellular nutrients. Our work demonstrates that the spatial arrangement of organisms can significantly impact the functioning of microbial systems when the spatial scale of interaction is small. Knowing at which scale organisms interact is crucial for understanding and controlling natural communities, and for engineering microbial systems for desired purposes.

Methods

Strains

All experiments were performed using strains derived from MG1655; these strains are ΔtrpC-GFP (MG1655 *trpC*::frt, PR-*sfGFP*), ΔtrpC-RFP (MG1655 *trpC*::frt, PR-*mCherry*), ΔproC-GFP (MG1655 *proC*::frt, PR-*sfGFP*), and ΔproC-RFP (MG1655 *proC*::frt, PR-*mCherry*). The ΔtrpC strains are unable to produce proline due to a deletion in *proC*, the ΔtrpC are unable to produce tryptophan due to a deletion in *trpC*(19). The auxotrophic strains were made by transferring the respective kanamycin cassettes from the keio-collection (20) into TB204 and TB205 from the lab strain collection using lambda Red mediated recombination (21). TB204 and TB205 are *E. coli* MG1655 derivatives that constitutively express sfGFP or mCherry under the PR promoter from the chromosome. In brief, the kanamycin cassette including the homologous flanking regions were amplified by PCR from JW0377 (*proC*::*kan*) and JW1254 (*trpC*::*kan*) and transformed into TB204 and TB205 harbouring the pSim8 plasmid (kindly provided by Donald L. Court). Primer sequences used:

U_proC_fw: CAT AAA GTC ATC CTT TGT TGG G

D_proC_rv: CTT TAC GGA TTA GTG TGG GG

U_trpC_fw: AAC GTC GCC ATG TTA ATG CG

D_trpC_rv: GAA CTG AGC CTG AAA TTC AGG

The kanamycin cassette was transferred into a fresh strain of TB204 or TB205 using P1 mediated generalized transduction. Upon successful transduction, the phenotypes of the strains were confirmed (no growth without proline or tryptophan) and the kanamycin cassettes removed from the genome using the FLP-recombinase from plasmid pCP20(21). We confirmed the ability of our two auxotrophs to grow together by receiving the amino acid they cannot produce from their partner, as reported in previous work(19).

Media and growth condition

Monocultures of the two auxotrophs strains were started from a single colony taken from a LB-agar plate and were grown overnight at 37°C in a shaker incubator. The cells were growing overnight in M9 medium (47.76 mM Na_2HPO_4, 22.04 mM KH_2PO_4, 8.56 mM NaCl and 18.69 mM NH_4Cl) supplemented with 1mM $MgSO_4$, 0.1

mM $CaCl_2$, 0.2% glucose (all from Sigma-Aldrich), 50 µg/L of L-proline (434 mM) and 20 µg/L L-tryptophan (98 mM) and 0.1% Tween-20 (added to facilitate loading of cells in microfluidic device). Cells were loaded in stationary phase in a microfluidic device and grown in the same media. After approximately 10 hours, cells exit lag phase and started to fill the chambers. The medium was then switched to M9 medium + 0.2% glucose + 0.1% Tween-20 with no amino acids. The medium was fed at a flow rate of 0.5 ml/h for the whole duration of experiment (approximately three days). Imaging was started three hours before switching to this medium, to have a control of the cell's growth with amino acids in the medium.

Microfluidic experiment

The microfluidic devices consisted of chambers of 60x60 µm and 0.76 µm in height facing a feeding channel of 22 µm in height and 100 µm wide. Masks for photolithography were ordered at Compugraphics (Jena, Germany). A two-step photolithography was used to make SU8 molds on silicon wafers. Polydimethylsiloxane (PDMS, Sylgard 184 Silicone Elastomer Kit, Dow Corning) was mixed in a ratio of 1.5:10 and poured on the dust-free wafer, degassed in a desiccator for 30 minutes, and baked for around one hour at 80°C for curing. PDMS chips of approximately 2 cm×3.5 cm were cut out around the structures on the wafer. Holes for medium supply and outlet were punched (diameter of holes 1.2 mm). PDMS chips were bound to round (50 mm diameter) glass coverslips (Menzel-Gläser, Braunschweig, Germany) by treating them for 30 seconds at maximum power in a Plasma Cleaner (PDC-32G-2, Harrik Plasma, New York, USA), and left on a heated plate at 100°C for one minute for binding. Before an experiment, a small amount of medium was flushed into the channels using a pipette to wet the chambers. Then air was pushed through the main channel (medium remains in chambers). Cells in stationary phase, from overnight culture (14 h approximately) were concentrated approximately 100 times by centrifugation (5,000×*g*, 5 min.) and loaded into the chip using a pipette. Cells where pushed in the side chambers with the help of small air bubbles flowing through the main channel. Once a sufficient number of cells were pushed inside the chambers, fresh medium was pumped through the flow channel. For all experiments, syringe pumps (NE-300, NewEra Pump Systems) with 50 ml syringes containing the medium were used. Tubing (Microbore Tygon S54HL, ID 0.76 mm, OD 2.29 mm, Fisher Scientific) was connected to the syringes using 20G

needles (0.9 mm×70 mm), which were directly inserted into the tubing. Smaller tubing (Teflon, ID 0.3 mm, OD 0.76 mm, Fisher Scientific) was then inserted into the bigger tubing and directly connected to the inlet holes in the PDMS chip. Medium switches were performed by disconnecting the bigger tubing from the syringe and reconnecting it to new syringes. All experiments were run at a flow rate of 0.5 ml/h. The flow rate is high enough that amino acids do not accumulate in the feeding channel and are not exchanged via the main channel. In fact no growth was observed in chambers hosting only one of the two auxotrophs during the whole duration of the experiment.

Microscopy

Time-lapse microscopy was done using fully automated Olympus IX81 inverted microscopes (Olympus, Tokyo, Japan). Images were taken using a 100X NA1.3 oil objective (Olympus) with 1.6X manual auxiliary magnification and an ORCA-flash 4.0 v2 sCMOS camera (Hamamatsu, Hamamatsu, Japan). Fluorescent imaging was done using a X-Cite120 120 Watt high pressure metal halide arc lamp (Lumen Dynamics, Mississauga, Canada) and Chroma 49000 series fluorescent filter sets (N49002 for GFP and N49008 for RFP, Chroma, Bellows Falls, Vermont). Focus was maintained using the Olympus Z-drift compensation system and the entire setup was controlled with Olympus CellSens software. The sample was maintained at 37°C with a microscope incubator (Life imaging services, Basel, Switzerland). Several positions were imaged on the same microfluidic device and images were taken every ten minutes.

Image analysis

All image processing was done using Matlab (version 2016A and newer, MathWorks, Natick, Massachusetts) and Vanellus software (credit D. J. Kiviet, accessible at: http://kiviet.com/research/vanellus.php). Time-lapse frames were first registered and cells were then segmented using customized segmentation algorithms. Two different algorithms for segmentation were used: the '*segmentation of biomass algorithm'* and the '*segmentation of cells algorithm'*. The '*segmentation of biomass algorithm'* identifies the green and red biomass in the chamber: images were first cropped along the profile of the microfluidic chambers (up to 5 micrometers from the outlet), and biomass was segmented on the phase contrast image and assigned to its relative

252 colour after deconvolution; the algorithm was optimized to give the most accurate
253 estimation of the area occupied by cells of each type and not to segment the single
254 individuals. The '*segmentation of cells algorithm'* identifies individual cells for
255 subsequent single cell growth estimation (elongation rate). In this case, cells were
256 segmented on the green or the red fluorescent image, according to their fluorescence
257 colour. Single cell location was tracked using an optical flow based algorithm
258 (described below) and the tracking was manually corrected to prevent mistakes.
259 Subparts of the chambers were randomly selected for the single cell segmentation and
260 tracking and 250 cells per chamber were analyzed on average, giving a total of 15,475
261 cells across 61 chambers. The area close (within 5 μm) to the open end of the
262 chamber was not considered for analysis as amino acid concentrations in this area are
263 lower because they are washed out into the main flow channel. The tracking
264 algorithm based on optical flow can be described in three steps: 1) estimate vector
265 field of movement M between subsequent segmented images S_1 and S_2 using
266 Farneback(22) algorithm 2) back-transform the second image $S_{2,backtransformed} = - M \bullet$
267 S_2, to obtain a prediction of how S_1 should look like based on the vector field of
268 motion 3) for each cell in S_1 determine the area overlap with cells in $S_{2,backtransformed}$;
269 cells in S_1 are tracked to cells in S_2 based on maximum overlap area.
270

271 **Cell elongation rate**

272 Cell elongation rates were calculated by fitting the exponential curve $L(t)=L(0)\ 2^{\mu \cdot t}$ to
273 the cell length L over time. The fitting was done using a linear fit on the logarithm of
274 the cell length over a sliding time window of 5 time-points (40 minutes). Length of a
275 cell was measured as the length of the major axis of the ellipse that approximates the
276 cell, i.e. the ellipse that has the same normalized second central moments as the cell.
277

278 **Correlation analysis**

279 We quantified the composition of the neighbourhood of a focal cell as the fraction of
280 the other complementary partner present in that neighbourhood, e.g. in the case of the
281 tryptophan auxotroph we quantified the fraction ΔproC/(ΔproC+ΔtrpC). ΔtrpC and
282 ΔproC are the areas (in pixel) occupied by each auxotroph, therefore they are a
283 measurement of biomass and not of cell number. To calculate the fraction above, we
284 first identified biomass of the two types as described in the image analysis section;
285 then we calculated the area in pixel that each cell type occupies within increasing

distances from the focal cell's perimeter. For a given distance, we plotted the fraction (x-axis) against the growth rate (y-axis) for all cells and we calculated Spearman's rank correlation coefficient (no assumption on the functional relationship between variables). The correlation coefficient is maximal at a specific distance, which we call interaction range. We use linear regression to characterize the relationship between the growth rate of the cells and the fraction of the amino acid producing partner present within the estimated interaction range. For figure 3a, the correlation is calculated as Spearman ρ on 2,610 data points for proline auxotrophs and 2,162 for tryptophan auxotrophs, both from four biological replicates. The same correlation analysis performed when cells are fed amino acids shows that growth does not depend on the neighbours when amino acids are present in the medium (see Supplementary Data).

Individual-based model

We consider two cell types living on a on a 40x40 squared grid: Type A can only produce amino acid 1 while Type B can only produce amino acid 2. The growth of type A is thus limited by the supply of amino acid 2 leaked by type B cells and vice versa. We make the following assumptions:

a) Cells can maintain a constant internal concentration I of the amino acid they can produce.

b) Growth of a cell is limited only by the amino acid the cell cannot produce; growth is modeled using the Monod equation (23) $\mu = \mu^{max} I / (I + K)$, where K is the concentration at which cells grow at half maximum speed.

c) Both cell types have the same maximum growth rate μ^{max}.

d) Cells take up amino acids actively (10), and the process is approximated with linear kinetics: $inflow = r^{u} E$, where r^{u} is the uptake rate and E is the external concentration. Linear kinetics approximates Monod kinetics if the concentration of external amino acids E are low.

e) Cells leak amino acids through passive diffusion through the cellular membrane (10) $outflow = r^{l} (I - E)$, where r^{l} is the leakage rate.

f) Diffusion in the extracellular environment is modeled as diffusion in a crowded environment (24) $D^{eff} = D(1-\rho)/(1+\varrho/2)$, where D is the diffusion constant and ρ the cell density.

g) The ratio between the volume inside a cell and the available volume outside of a cell is constant and equal to $\alpha = \rho/(1-\rho)$.

With these assumptions, we can write the following equations for the internal concentration of amino acids for a cell of type A - which produces amino acid 1, and not amino acid 2:

$$\frac{\partial I_1}{dt} = 0$$

$$\frac{\partial I_2}{dt} = r_2^u E_2 - r_2^l \left(I_2 - E_2\right) - I_2 \, \mu^{max} I_2 / (I_2 + K_2)$$

and for type B -which produces amino acid 2 and not 1:

$$\frac{\partial I_2}{dt} = 0$$

$$\frac{\partial I_1}{dt} = r_1^u E_1 - r_1^l \left(I_1 - E_1\right) - I_1 \, \mu^{max} I_1 / (I_1 + K_1)$$

The external concentration of each amino acid is:

$$\frac{\partial E_i}{dt} = -\alpha \, r_i^u E_i + \alpha \, r_i^l \left(I_i - E_i\right) + D_i^{eff} \nabla^2 E_i$$

Some parameters can be eliminated by expressing the concentrations of amino acid *i* in units of K_i, time in units of $1/\mu^{max}$, and space in units of cell size. This gives a set of dimensionless equations with a reduced number of parameters (Supplementary Information). The other parameters are taken from literature or measured, with the exception of the leakage rates, which are estimated from data. These equations can be used to predict cells' growth rates in real or artificial arrangements of the two cell types. For detail about the numerical solution see Supplementary Information.

Cellular automaton

The cellular automaton models a system of two or more types of organisms that live on a grid and benefit from the presence of the other types. The model rests on two assumptions: 1) individuals place offspring close to themselves; 2) reproductive

success of individuals depends on the fraction of neighbours of the other type within the interaction range, the sole parameter in the model.

An operative description of the cellular automaton follows: individuals live in space, each occupying a site on a 40x40 grid; each site has 8 adjacent sites on the grid (Moore neighbourhood) and boundary condition wrap the grid into a Torus. For the two type communities, there are individuals of types 0 and 1. At every time step an individual dies at a random location on the grid and it is replaced with an individual of type 0 or 1. It will be of type 0 with probability P(0):

$$P(0) = \frac{\sum_i^{adjacent\ individuals} \delta_i\ Reproductive\ success_i}{\sum_i^{adjacent\ individuals} Reproductive\ success_i}$$

where δ_i is the Dirac delta function, which is one if grid site *i* contains type 0 and zero otherwise. The reproductive success of each individual *i* is:

$$Reproductive\ success_i = \frac{number\ of\ neighbors\ of\ the\ other\ type}{number\ of\ neighbors}$$

Individuals interact with all other individuals within a neighbourhood of range *R* (a square-shaped neighbourhood). For communities with more than two types, the reproductive success is the fraction of neighbours that is most rare in the neighbourhood:

$$Reproductive\ success_i = \frac{number\ of\ neighbors\ of\ the\ rarest\ type}{number\ of\ neighbors}$$

All the rest is easily extended from the two types community described above to communities of more than two types. To compare systems with a different number of types, the reproductive success is normalized by the reproductive success the systems has in well mixed conditions ($R \to \infty$), which is 1/2 for two types, 1/3 for three and 1/4 for four.

Starting from different initial configurations and varying proportions of the types, we let the system evolve and stop the simulation after the system has attained a dynamical equilibrium where the average reproductive success of individuals remains

approximately constant. Average steady state reproductive success result from 100 independent runs of the cellular automaton. The cellular automaton is implemented in C++.

Dataset and statistical analysis

The dataset consists of 15,475 cells, from 61 communities, grouped into ten biological replicates including both fluorescent label combinations. Four biological replicates were done with ΔtrpC-GFP and ΔproC-RFP (consortium 1) and six were done with ΔtrpC-RFP and ΔproC-GFP (consortium 2). Each biological replicate corresponds to one channel in a microfluidic chip and for each channel on average 6 chambers were analyzed (range: 3-9). Inside each chamber, on average 250 cells were tracked in time (see Image Analysis section). The experiments were performed in three different weeks (different microfluidic chips and different batch of media). The interaction range and relation between growth and neighborhood were estimated separately for consortium 1 and 2. The interaction ranges are consistent for the two consortia (Fig. 3a shows consortium 1, Fig. S5a shows consortium 2), but the fluorescent label affect the growth rate to some extent: the ΔtrpC-RFP grows generally slower than the ΔtrpC-GFP (Fig. 3b-c shows consortium 1, Fig. S5b-c consortium 2). To assess the variability of the estimate of the interaction range, we repeated the analysis for each replicate in isolation (results are shown in Fig. 3b).

Acknowledgements

We thank Andrea Cavagna for discussing the correlation analysis, Glen D'Souza for advise on the biological assumptions of the model, Gabriele Micali, Roman Stocker, and Michael Doebeli for comments on the model, and Charlotte Brannon for comments on the manuscript. The research was supported by funding from the Swiss National Science foundation (grants no. 31003A_149267 and 31003A_169978 to MA), by an ETH fellowship to DJK, by an Early Postdoc Mobility fellowship from the Swiss National Science Foundation to SVV (grant no. 175123), by ETH Zurich and Eawag.

Author Contributions

ADC and MA conceived the research, ADC performed the experiment with contributions of SVV, ADC analyzed the data with contributions of DJK and SVV,

SVV and ADC conceptualized the individual-based model and SVV implemented it, SS constructed the bacterial strains, DJK constructed the microfluidic device, ADC wrote the manuscript with contributions of MA and SVV.

Competing interest declaration

The authors declare no competing interests.

Data and materials availability

The data of this study are available from the corresponding author upon request.

References:

1. Horner-devine MC, Carney KM, Bohannan BJM (2004) An ecological perspective on bacterial biodiversity. *Proceeding R Soc* 271:113–122.
2. Falkowski PG, Fenchel T, Delong EF (2008) The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science (80-)* 320(5879):1034 LP-1039.
3. Cordero OX, Datta MS (2016) Microbial interactions and community assembly at microscales. *Curr Biol* (31):227–234.
4. Gilbert JA, Henry C (2015) Predicting ecosystem emergent properties at multiple scales. *Environ Microbiol Rep* 7(1):20–22.
5. Phelan V, Liu W-T, Pogliano K, Derrestein PC (2013) Microbial metabolic exchange - the chemotype-to-phenotype link Vanessa. *Nat Chem Biol* 8(1):1–22.
6. D'Souza G, et al. (2014) Less is more: Selective advantages can explain the prevalent loss of biosynthetic genes in bacteria. *Evolution (N Y)* 68(9):2559–2570.
7. Mee MT, Collins JJ, Church GM, Wang HH (2014) Syntrophic exchange in synthetic microbial communities. *Proc Natl Acad Sci U S A* 111(20):E2149-56.
8. Schink B (2002) *Schink B.. Synergistic interactions in the microbial world. Antonie Van Leeuwenhoek Int J Gen Mol Microbiol 81: 257-261*.
9. Nadell CD, Drescher K, Foster KR (2016) Spatial structure, cooperation and competition in biofilms. *Nat Rev Microbiol* 14:589.
10. D'Souza G, et al. (2018) Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Nat Prod Rep* 35:455–488.
11. Morris JJ (2015) Black Queen evolution : the role of leakiness in structuring

439 microbial communities. *Trends Genet* 31(8):475–482.
440 12. Drescher K, Nadell CD, Stone HA, Wingreen NS, Bassler BL (2014) Report
441 Solutions to the Public Goods Dilemma in Bacterial Biofilms. *Curr Biol*
442 24(1):50–55.
443 13. Hol FJH, et al. (2013) Spatial Structure Facilitates Cooperation in a Social
444 Dilemma : Empirical Evidence from a Bacterial Community. 8(10):2–11.
445 14. Muller MJI, Neugeboren BI, Nelson DR, Murray AW (2014) Genetic drift
446 opposes mutualism during spatial population expansion. *Proc Natl Acad Sci*
447 111(3):1037–1042.
448 15. Momeni B, Waite AJ, Shou W (2013) Spatial self-organization favors
449 heterotypic cooperation over cheating. *Elife* 2(2):1–18.
450 16. Mitri S, Xavier JB, Foster KR (2011) Social evolution in multispecies bio fi
451 lms. *PNAS* 108:10839–10846.
452 17. Oliveira NM, Niehus R, Foster KR (2014) Evolutionary limits to cooperation
453 in microbial communities. (7). doi:10.1073/pnas.1412673111.
454 18. Davies DG, Geesey GG (1995) Regulation of the alginate biosynthesis gene
455 algC in Pseudomonas aeruginosa during biofilm development in continuous
456 culture. *Appl Environ Microbiol* 61(3):860–867.
457 19. Marchal M, et al. (2017) A passive mutualistic interaction promotes the
458 evolution of spatial structure within microbial populations. *BMC Evol Biol*
459 17:106.
460 20. Baba T, et al. (2006) Construction of Escherichia coli K-12 in-frame, single-
461 gene knockout mutants: The Keio collection. *Mol Syst Biol* 2.
462 doi:10.1038/msb4100050.
463 21. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes
464 in Escherichia coli K-12 using PCR products. *Proc*
465 *Natl Acad Sci* 97(12):6640 LP-6645.
466 22. Farneback G (2003) Two-frame motion estimation based on polynomial
467 expansion. *Image Anal Proc* 2749(1):363–370.
468 23. Monod J (1949) The Growth of Bacterial Cultures. *Annu Rev Microbiol*
469 3(1):371–394.
470 24. Lebenhaft JR, Kapral R (1979) Diffusion-controlled processes among partially
471 absorbing stationary sinks. *J Stat Phys* 20(1):25–56.

Figures and Figure Legends

Fig. 1

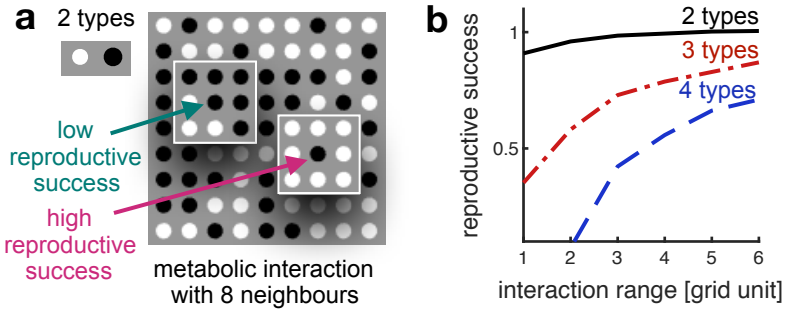


Fig. 1: Interacting locally lowers reproductive success. a Cells (black or white dots) exchange compounds to reproduce, and place offspring on adjacent sites. Reproductive success is maximal for cells that are surrounded by the other type. **b** The average reproductive success of individuals is lower in systems with smaller interaction range.

Fig. 2

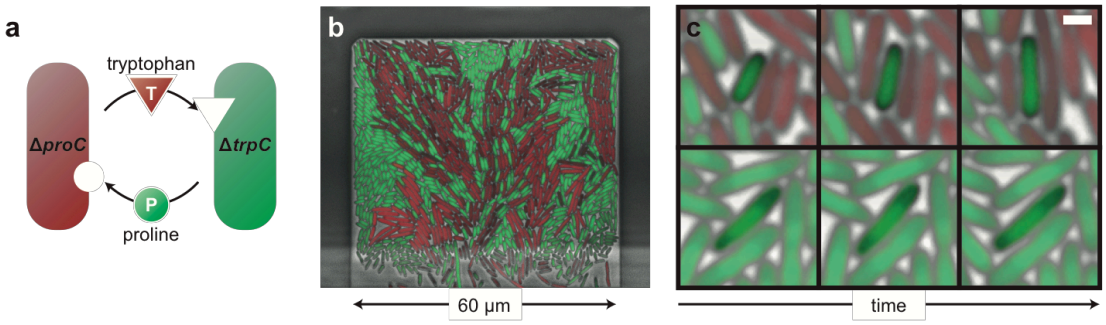
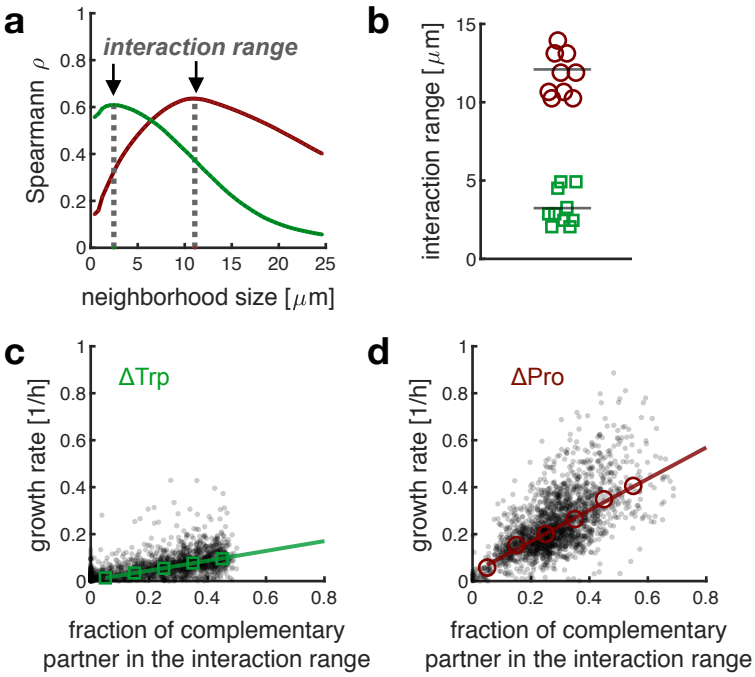


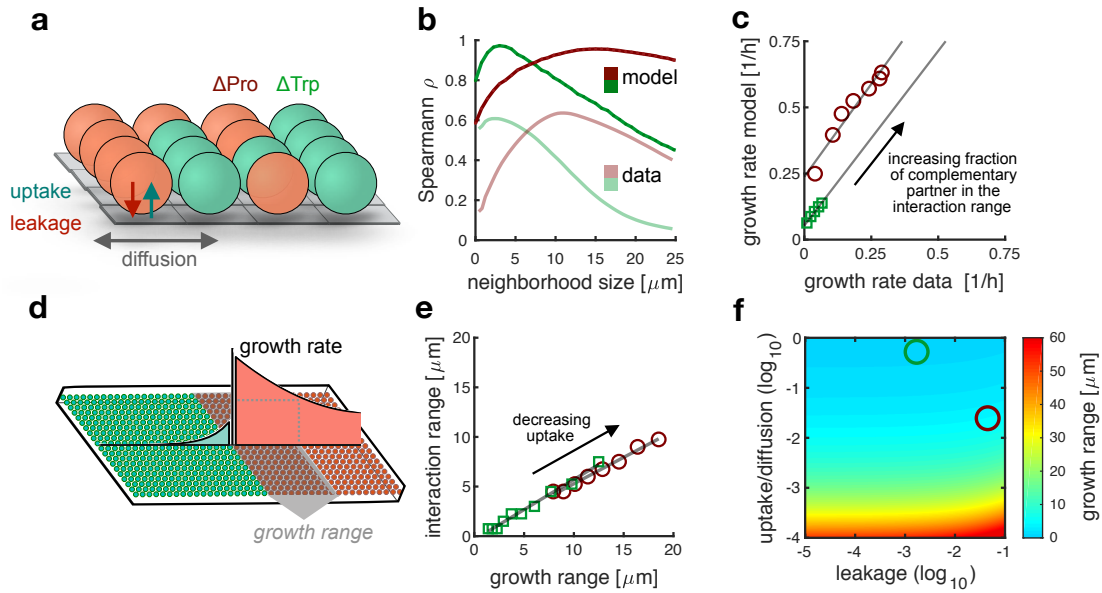
Fig. 2: Measuring the spatial range of cell-to-cell interaction. a Synthetic consortia of two auxotroph mutants of *E. coli* that depend on each other and that are labeled with constitutively expressed green and red fluorescent proteins. **b** Microfluidic chamber hosting ~1400 cells in a monolayer. Continuous flow of culture media at the bottom of the chamber removes cells as soon as they are pushed out of the chamber. **c** Auxotrophic cells surrounded by the complementary partner (top row) grow faster than auxotrophic cells surrounded by their own type (bottom row). Scale bar 1 µm.

Fig. 3



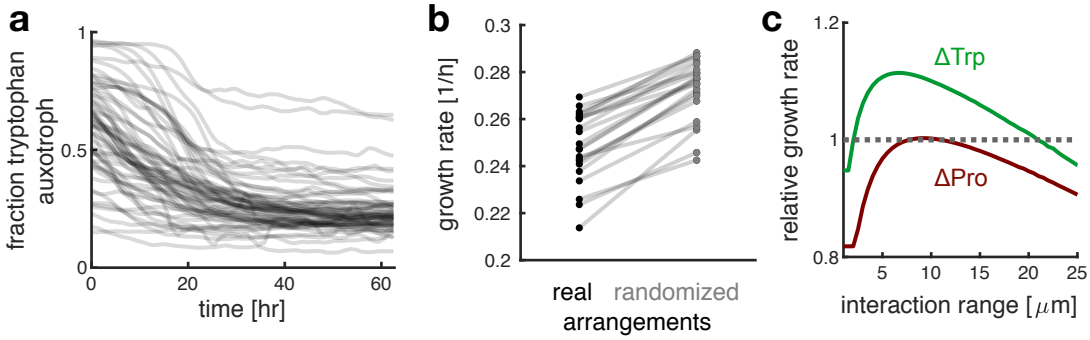
485 **Fig. 3: Individuals interact at a short spatial range**. **a** We calculated the correlation
486 coefficients between the growth rates of individual cells and the fraction of the
487 complementary partner in a given neighborhood size. When we plot the correlation
488 coefficient as a function of the neighborhood size, we observe that the strength of the
489 correlation is maximal for an intermediate neighborhood size (marked by dashed
490 lines); we call this neighborhood size the interaction range. **b** Proline (red) and
491 tryptophan (green) auxotrophs have different interaction ranges (10 biological
492 replicates, ~15,000 cells total). **c-d** Both auxotrophic cells grow faster when
493 surrounded by the complementary partner within the interaction range. Tryptophan
494 auxotrophs (c) generally have lower growth rates than proline auxotrophs (d), as
495 shown by the slopes of the linear regression (0.75 for ΔproC and 0.21 for ΔtrpC).
496 Black dots: single cells (2,610 ΔproC and 2,162 ΔtrpC cells); Open symbols: binned
497 median values; lines: linear regression on binned values.

Fig. 4



498 **Fig. 4: Mathematical model shows mechanism of local interactions. a** Individual-
499 based model where amino acids are actively taken up and passively leaked.
500 Tryptophan auxotrophs are shown in green and proline auxotrophs in red in all panels.
501 **b** The model (dark curves) predicts the empirical correlation analysis (light curves,
502 identical to those in Fig. 3a). **c** The predicted and experimentally measured growth
503 rates are strongly correlated. We grouped cells based on the fraction of the
504 complementary partner in their interaction range and for each group we compared the
505 measured growth rate (x-axis, same data as Fig. 3c-3d) to the predicted growth rates
506 (y-axis). Each symbol represents a single group. **d** For a symmetric arrangement with
507 a straight interface between the two types, we can analytically calculate the range
508 (shown in grey) in which the cellular growth rate is at least half of the maximal
509 growth rate observed at the interface; we call this range the growth range. The growth
510 range can be calculated from biochemical parameters. **e** The interaction range is
511 proportional to the growth range. When we decrease, in the model, the rate at which
512 cells take up amino acids, the growth range and the interaction range increase; circles
513 show combinations of growth range and interaction ranges for different values of the
514 uptake rate of amino acids. **f** Growth range as a function of leakage and uptake rate
515 (relative to maximum growth rate and diffusion constant respectively). Growth range
516 (and interaction range) of the tryptophan auxotroph (green circle) and proline
517 auxotroph (red circle) are small.

Fig. 5



518 **Fig. 5: The short interaction range reduces growth of the whole microbial**
519 **system.** **a** Communities equilibrate at compositions of 23% (median, n=61) of the
520 tryptophan auxotrophs; deviations are due to large clusters of tryptophan auxotrophs
521 in the back of the chamber. **b** Randomizing community arrangements leads to higher
522 mixing and higher predicted average growth rates (n randomization = 20, $p<10^{-5}$,
523 paired t-test, n=22); **c** The model predicts an increase in the average growth rate when
524 the interaction range increases (as a consequence of a decrease in the uptake rate).
525 When the uptake rate is very low and the interaction range therefore very large, amino
526 acids diffuse out of the chamber and the growth rate decreases.

Supplementary Information

Contents

1. Supplementary discussion 23
1.1. Independence of growth from neighbours when amino acids are supplied 23
1.2. Robustness of interaction range estimate to spatial arrangement 23
1.3. Limitation to the prediction of growth rates 24
1.4. Tradeoff in uptake rates of amino acids 25
1.5. Additional data 25
2. Individual-base model 26
2.1. Rescaled equations 26
2.2. Steady state equations 28
2.3. Numerical solution and boundary condition 29
2.4. Analytical limits 29
2.5. Discussion on the effect of parameters 34
3. Supplementary methods 35
3.1. Proportionality of analytical growth range and interaction range 35
4. Supplementary Tables 36
5. Captions for Movies 36
6. References 36

1. Supplementary discussion

1.1. **Independence of growth from neighbours when amino acids are supplied.** We verified that growth of auxotrophic cells does not depend on the identity of their neighbours when amino acids are externally supplied (Figure S2a and S2b).

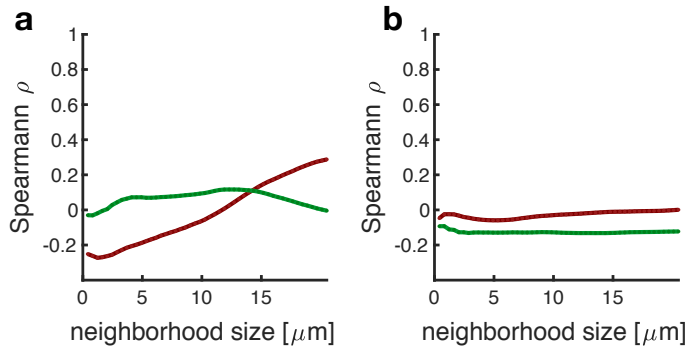


Figure S2. **Growth does not depend on the identity of neighbours when amino acids are fed.** When media is supplemented with proline and tryptophan, growth of auxotrophic cells does not depend on the presence of the complementary partner. The correlation is low for all distances analysed. Panel (a) shows results for consortia 1, panel (b) for consortia 2. Proline auxotroph (red), tryptophan auxotroph (green).

1.2. **Robustness of interaction range estimate to spatial arrangement.** The model allowed us to verify that the interaction range we measure does not arise from the spatial arrangements we analyse, but is rather a property of the system. Generally, inside the communities, the two cell types display different typical patch sizes, with the $\Delta trpC$ (the auxotroph that has the smaller interaction range) forming smaller patches. This observation raises the question: is our correlation analysis affected by patch size? We tested whether patch size affects our estimate of the interaction range by generating several synthetic datasets, each with 100 configurations of the two types arranging in patches of controlled sizes (Figure S3a); we analysed these synthetic datasets in the same way as our empirical dataset (Figure S3b and S3c). The results confirmed that the interaction range of each type, i.e. the location of the correlation peak in Figure 3a, is robust to changes in patch size (Figure S3d). In particular, we can show that the

location of the peak does not change more than 30% for a range of patch sizes which are typically observed in the data (visual inspection). This result supports that our analysis of correlation between growth rates of individuals and their neighbourhood composition is a valid method to determine how cells affect each other's growth within complex spatial arrangements.

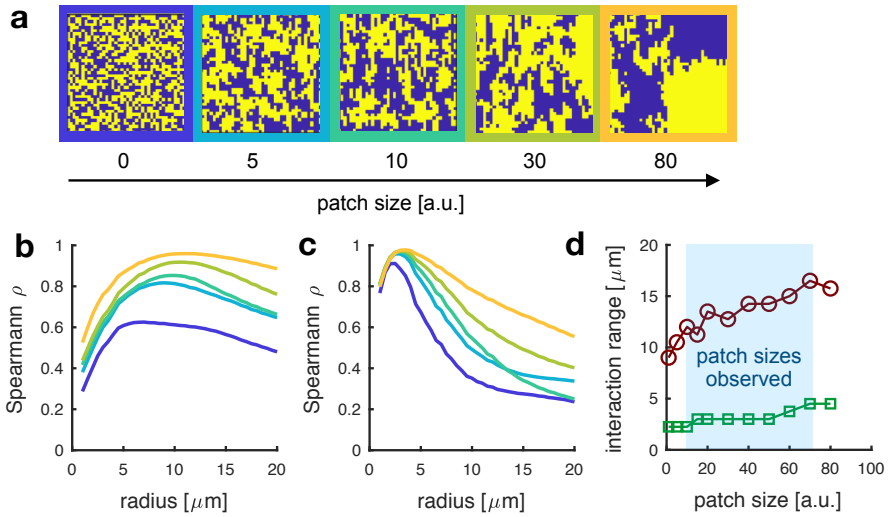


Figure S3. **Robustness of interaction range estimate to spatial arrangement of types.** (a) Examples of artificial arrangements with controlled patch size; dataset of 100 different arrangements per patch size were generated and analysed. The shape of the correlation curve changes for both tryptophan (b) and proline (c) auxotrophs but the interaction range changes only minimally (d) for the range of patch size typically observed in the data. Proline auxotroph (red circles), tryptophan auxotroph (green squares).

1.3. **Limitation to the prediction of growth rates.** The model recapitulates quantitatively the effect of spatial arrangement on growth (Fig. 4c - linear correlation between binned growth rates predicted by the model and measured in the data $R^2 > 0.96$), However, the model tends to overestimate the absolute growth rate of cells (Fig. 4c displays an intercept). In fact, the classical Monod equation does not consider that cells may need substrate (the limiting amino acid here) even when they do not grow. For this reason, the original Monod equation is often modified by introducing a term of maintenance (1).

A more refined model including a growth cost could improve the estimation of growth rates; we keep this for future studies.

1.4. **Tradeoff in uptake rates of amino acids.** High uptake rates of metabolites like amino acids seem advantageous for the growth of individual cells, however they are not for the whole community: high uptake rates hinder metabolic exchange between different genotypes and thus reduce overall growth. Our simulations (Fig. 5c) show that the average growth of the two auxotrophs increases when the interaction range increases (by lowering the uptake rates of amino acids). However, there is a tradeoff: when uptake rates are too low, the average growth of the auxotrophs decreases because amino acid diffuse out of the chamber (Fig. 5c). We can show that in closed systems, where amino acids cannot diffuse away, growth does not decrease (Fig S4)

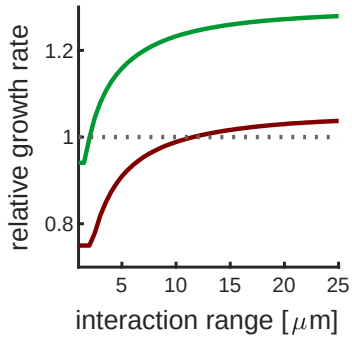


Figure S4. **Lower uptake rates lead to higher growth when amino acids cannot diffuse away from the community.** When simulating systems with closed boundaries, the growth rate of the two auxotrophs increases with increasing interaction range (decreasing uptake rates). Proline auxotroph (red), Tryptophan auxotroph (green).

1.5. **Additional data.** Figure S5 show the correlation analysis between single cells growth rates and fraction of the complementary partner in the interaction range for six biological replicates done with $\Delta trpC$-RFP and $\Delta proC$-GFP (consortium 2); figure 3 in the main text shows four biological replicates done with $\Delta trpC$-GFP and $\Delta proC$-RFP (consortium 1). The interaction ranges are consistent for the two consortia (Fig. 3a shows consortium 1, S52a consortium 2), but the fluorescent label affect the growth

rate to some extent: the $\Delta trpC$-RFP grows generally slower than the $\Delta trpC$-GFP (Fig. 3b-c shows consortium 1, Fig. S5b-c consortium 2).

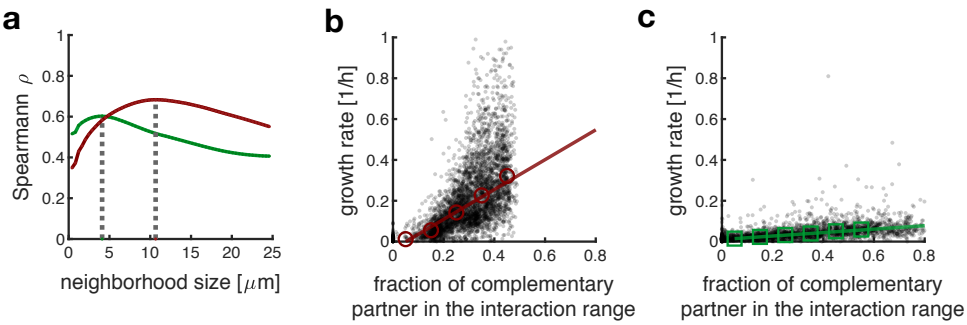


Figure S5. **Individuals interact at a small spatial range.** The plots shows data for consortia number 2 ($\Delta trpC$-RFP and $\Delta proC$-GFP), and complement Fig. 3 showing data from consortia number 1 ($\Delta trpC$ GFP and $\Delta proC$-RFP). (a) The cells' growth rate correlates maximally with the identity of their neighbours within the interaction range. (b-c) Both auxotrophic cells grow faster when surrounded by more complementary partners inside the interaction range. Tryptophan auxotrophs (b) achieve generally smaller growth rates then proline auxotrophs (c), as shown by the slopes of the linear regression (0.79 for $\Delta proC$ and 0.089 for $\Delta trpC$). Black dots: single cells (6,871 for $\Delta proC$ and 3,832 for $\Delta trpC$); red (green) open symbols: binned median values; lines: linear regression on binned values.

2. Individual-base model

2.1. **Rescaled equations.** The individual-based model describes two cell types on a squared grid: type A can only produce amino acid 1 while type B can only produce amino acid 2. The growth of type A is thus limited by the supply of amino acid 2 leaked by type B cells and vice versa. At every site, the internal and external concentration of amino acids are described by the following equations (definition of parameters in the

Methods).

$$\frac{\partial I_1}{\partial t} = 0 \qquad \text{if type is A} \tag{1}$$

$$\frac{\partial I_1}{\partial t} = r_1^u \cdot E_1 - r_1^l \cdot (I_1 - E_1) - \frac{\mu^{max} \cdot I_1}{K_1 + I_1} \cdot I_1 \qquad \text{if type is B} \tag{2}$$

$$\frac{\partial I_2}{\partial t} = r_2^u \cdot E_2 - r_2^l \cdot (I_2 - E_2) - \frac{\mu^{max} \cdot I_2}{K_2 + I_2} \cdot I_2 \qquad \text{if type is A} \tag{3}$$

$$\frac{\partial I_2}{\partial t} = 0 \qquad \text{if type is B} \tag{4}$$

$$\frac{\partial E_i}{\partial t} = -\alpha \cdot r_i^u \cdot E_i + \alpha \cdot r_i^l \cdot (I_i - E_i) + D_i^{eff} \nabla^2 E_i \tag{5}$$

We can reduce the number of parameters by rescaling units as follows:

- concentrations are measured in units of K_i
- time is measured in units of inverse growth rate: $\frac{1}{\mu^{max}}$
- space is measured in units of cell size: Δx.

This gives the following equations:

$$\frac{\partial \hat{I}_1}{\partial \hat{t}} = 0 \qquad \text{if type is A} \tag{6}$$

$$\frac{\partial \hat{I}_1}{\partial \hat{t}} = \hat{r}_1^u \cdot \hat{E}_1 - \hat{r}_1^l \cdot (\hat{I}_1 - \hat{E}_1) - \frac{\hat{I}_1}{1 + \hat{I}_1} \cdot \hat{I}_1 \qquad \text{if type is B} \tag{7}$$

$$\frac{\partial \hat{I}_2}{\partial \hat{t}} = \hat{r}_2^u \cdot \hat{E}_2 - \hat{r}_2^l \cdot (\hat{I}_2 - \hat{E}_2) - \frac{\hat{I}_2}{1 + \hat{I}_2} \cdot \hat{I}_2 \qquad \text{if type is A} \tag{8}$$

$$\frac{\partial \hat{I}_2}{\partial \hat{t}} = 0 \qquad \text{if type is B} \tag{9}$$

$$\frac{\partial \hat{E}_i}{\partial \hat{t}} = -\alpha \cdot \hat{r}_i^u \cdot \hat{E}_i + \alpha \cdot \hat{r}_i^l \cdot (\hat{I}_i - \hat{E}_i) + \hat{D}_i^{eff} \hat{\nabla}^2 \hat{E}_i \tag{10}$$

Where:

$$\hat{I}_i = \frac{I_i}{K_i} \qquad \hat{E}_i = \frac{E_i}{K_i} \qquad \hat{t} = \mu^{max} \cdot t$$

$$\hat{r}_i^u = \frac{r_i^u}{\mu^{max}} \qquad \hat{r}_i^l = \frac{r_i^l}{\mu^{max}} \qquad \hat{D}_i^{eff} = \frac{D_i^{eff}}{\mu^{max} \cdot \Delta x^2} = \frac{(1-\rho) \cdot D_i}{(1+\frac{\rho}{2}) \cdot \mu^{max} \cdot \Delta x^2} \qquad \alpha = \frac{\rho}{(1-\rho)}$$

In the remainder of this text we will omit the hats: all variables and parameters always refer to the rescaled ones.

2.2. **Steady state equations.** We want to obtain the steady state concentration $E_i(x, y)$ and $I_i(x, y)$ on every site (x, y) of the grid. Setting equation 7 or 8 to steady state and rewriting gives:

$$0 = r_i^u \cdot E_i - r_i^l \cdot (I_i - E_i) - \frac{I_i}{1 + I_i} \cdot I_i$$
$$0 = (1 + r_i^l)I_i^2 + \left(r_i^l - (r_i^u + r_i^l)E_i\right) I_i - (r_i^u + r_i^l)E_i$$

Which we can solve to get the internal concentration of the amino acid each cell cannot produce as function of the external concentration of that amino acid:

$$I_i(E_i) = f(E_i) = \frac{(r_i^u + r_i^l)E_i - r_i^l}{2(1 + r_i^l)} + \frac{\sqrt{\left(r_i^l\right)^2 + (r_i^u + r_i^l)^2 E_i^2 + (2r_i^l + 4)(r_i^u + r_i^l)E_i}}{2(1 + r_i^l)} \tag{11}$$

The internal concentration of the produced amino acid is kept constant (equations 6 and 9):

$$I_i(E_i) = I_i^C. \tag{12}$$

Setting equation 10 to steady state gives:

$$\nabla^2 E_i = \frac{\alpha \cdot (r_i^u + r_i^l)}{D_i^{eff}} E_i - \frac{\alpha \cdot r_i^l}{D_i^{eff}} I_i(E_i) \tag{13}$$

where $I_i(E_i)$ is given by equation 12 for grid sites where amino acid i is produced and by equation 11 otherwise. If we describe the spatial arrangement of the two cell types with the function $T(x, y)$:

$$T(x, y) = 0 \qquad \text{if site } (x, y) \text{ is occupied by type A} \tag{14}$$
$$T(x, y) = 1 \qquad \text{if site } (x, y) \text{ is occupied by type B}$$

then the external concentration of the two amino acid on each grid site is the solution of the equations:

$$\nabla^2 E_1(x,y) = \frac{\alpha(r_1^u + r_1^l)}{D_1^{eff}} E_1(x,y) - \frac{\alpha r_1^l}{D_1^{eff}} \cdot \left(T(x,y) \cdot f\left(E_1(x,y)\right) + [1 - T(x,y)] \cdot I_1^C\right) \tag{15}$$

$$\nabla^2 E_2(x,y) = \frac{\alpha(r_2^u + r_2^l)}{D_2^{eff}} E_2(x,y) - \frac{\alpha r_2^l}{D_2^{eff}} \cdot \left([1 - T(x,y)] \cdot f\left(E_2(x,y)\right) + T(x,y) \cdot I_2^C\right) \tag{16}$$

After solving for $E_i(x,y)$ we can obtain the growth profile $\mu(x,y)$:

$$\mu(x,y) = [1 - T(x,y)] \cdot \frac{I_2(x,y)}{1 + I_2(x,y)} + T(x,y) \cdot \frac{I_1(x,y)}{1 + I_1(x,y)} \tag{17}$$

$$I_i(x,y) = f\left(E_i(x,y)\right) \tag{18}$$

2.3. **Numerical solution and boundary condition.** We numerically solved equations 15 and 16 on the grid. On one edge of the grid we implement a Dirichlet boundary condition and set $E_i = 0$ to represent the flow-channel where all excreted amino-acids are washed away; on all other edges we implement Neumann no-flux boundary conditions to represents the solid wall of the growth chamber. We solved equations 15 and 16 by discretizing them using a second order finite difference scheme and solving them using a successive over-relaxation solver. To ensure numerical stability, we implemented a grid-refinement procedure: we first solved the equations on the 40x40 grid and then we iterated on refined grids (successively doubling the number of grid points in each dimension); we used the solution of the previous iteration as the initial state for the successive refined grid. The solution on the refined grid was downsampled to the 40x40 grid to calculate the growth rate for each cell and we continued this procedure until the maximum per cell change in growth rate was less than 1%. All code was implemented in Matlab.

2.4. **Analytical limits.** We derived several analytical approximations to understand the effect of each parameter, which we summarize in subsection 2.5. Here we will consider a single cell type at a time and follow the internal, I_L, and external, E_L, concentration

of the limiting amino-acid only:

$$\frac{\partial I_L}{\partial t} = r^u \cdot E_L - r^l \cdot (I_L - E_L) - \frac{I_L}{1 + I_L} \cdot I_L \tag{19}$$

$$\frac{\partial E_L}{\partial t} = -\alpha \cdot r^u \cdot E_L + \alpha \cdot r^l \cdot (I_L - E_L) + D^{eff} \nabla^2 E_L. \tag{20}$$

Where r^u, r^l, and D^{eff} always refer to the uptake, leakage, and effective diffusion constant of the growth limiting amino acid. It is useful to rewrite these equation in terms of $\epsilon = \frac{r^u + r^l}{r^l} E_L$:

$$\frac{\partial I_L}{\partial t} = r^l \cdot \epsilon - r^l \cdot I_L - \frac{I_L}{1 + I_L} \cdot I_L \tag{21}$$

$$\frac{\partial \epsilon}{\partial t} = \alpha \cdot (r^u + r^l) \cdot (I_L - \epsilon) + D^{eff} \nabla^2 \epsilon \tag{22}$$

2.4.1. *Limiting amino acid at steady state.* Setting the time derivatives to zero and solving 21 for I_L gives:

$$I_L(\epsilon) = \frac{(\epsilon - 1) r^l + \sqrt{(\epsilon - 1)^2 (r^l)^2 + 4\,(1 + r^l)\, r^l \epsilon}}{2\,(1 + r^l)} \tag{23}$$

2.4.2. *Maximum cell growth rate.* We derive the analytical expression for the growth rate of an auxotrophic cells surrounded by a large number of producing partners. If we assume that the single auxotroph has a negligible influence on the external concentration (i.e. all space is occupied by producers which have $I = I^C$), equation 22 gives the steady state external concentration of amino acids:

$$\epsilon_{max} = I^C \tag{24}$$

substituting ϵ_{max} for ϵ in eq. 23 we find:

$$\begin{aligned} I_L^{max} &= \frac{(I^C - 1) r^l + \sqrt{(I^C - 1)^2 (r^l)^2 + 4\,(1 + r^l)\, r^l I^C}}{2\,(1 + r^l)} \\ \mu^{max} &= \frac{I_L^{max}}{1 + I_L^{max}} \end{aligned} \tag{25}$$

This is the growth rate of a single auxotrophs surrounded by a large number of amino acid producing partners. We can simplify this expression if we make the following two assumptions:

Assumption 2.4.1. $I^C \gg 1$. *Biologically this means that a wild type (amino acid producing) cell can grow nearly as fast in the absence of amino acids ($\mu = \frac{I^C}{1+I^C}$) as in the presence of amino acids ($\mu = 1$). We verified experimentally that this assumption holds.*

Assumption 2.4.2. $r^l \ll 1$. *Biologically this means that in wild type cells the decrease in the concentration of limiting amino acid due to leakage (with rate r^l) is small compared to the decrease due to growth (with rate 1).*

With these assumptions the growth rate of an auxotroph surrounded by the producing partner (eq. 25) simplifies to:

$$\begin{aligned} I_L^{max} &\approx \tfrac{1}{2} I^C r^l \left(1 + \sqrt{1 + \tfrac{4}{I^C r^l}}\right) \\ \mu^{max} &= \frac{I_L^{max}}{1 + I_L^{max}} \end{aligned} \tag{26}$$

2.4.3. *Estimating leakage rates.* In this subsection we show how we estimated the leakage rates from the maximum empirical growth rates of the auxotrophs. In the limit of $r^l \ll 1$ and $I^C \gg 1$, the maximum growth rate (equations 26) depends on the product $I^C \cdot r^l$, and not on the two parameters separately. Both parameters are unknown but we expect $I^C \gg 1$ (assumption 2.4.1), which we arbitrarily set to 20 ($\mu_{producer} = 0.95$). With this, we can estimate the leakage rates r^l for each amino acid from μ^{max} of the corresponding auxotroph. Note that our result are robust to changes in value assigned to I^C as long as it is larger then one (we confirmed that our simulations depend only on the product $I^C \cdot r^l$ as long as $r^l \ll 1$ and $I^C \gg 1$). The maximum empirical growth rate μ^{max} is estimated for each auxotroph by performing a linear regression between the auxotroph's growth rate and the fraction of the producing partner within the interaction range (Fig. 3c-d); the maximum growth rate is the value extrapolated when the fraction is equal to one. We use linear regression because it is less sensitive to measurement noise than using the maximal observed growth rate, and because very few cells are found surrounded by the producing partner because of kin clustering.

2.4.4. *Analytical expression for the growth range.* In this subsection we show how we calculated the growth range, a quantity proportional to the interaction range measured experimentally (Fig.4e), from the parameters of the model. Let us consider a symmetric

spatial arrangement of types (Fig.4b), with all cells located at $x < 0$ producing the amino acid that cells at $x > 0$ require for their growth, and vice versa. This arrangement reduces the problem to one dimension and we can analytically calculate the cells' growth rates (we will do that for $x > 0$).

We can find the external amino acid concentration by solving equation 22 at steady state, where the internal concentration is $I = I^C$ for $x < 0$ and $I = I_L(\epsilon)$ for $x > 0$:

$$\frac{d^2\epsilon}{dx^2} = \begin{cases} \frac{1}{r_0^2}\left(\epsilon - I^C\right) & \text{if } x < 0 \\ \frac{1}{r_0^2}\left(\epsilon - I_L(\epsilon)\right) & \text{if } x > 0 \end{cases} \tag{27}$$

where $I_L(\epsilon)$ is given by equation 23 and

$$r_0 = \sqrt{\frac{D^{eff}}{\alpha(r^u + r^l)}}$$

For $x > 0$ the analytical solution cannot be found due to the non-linear term

$$I_L(\epsilon) = \frac{r^l}{2}(\epsilon - 1) + \tfrac{1}{2}\sqrt{(\epsilon - 1)^2(r^l)^2 + 4r^l\epsilon} \tag{28}$$

where we have simplified equation 23 using assumption 2.4.2 ($r^l \ll 1$). However it is easy to show that

$$I_L(\epsilon) < 1 + \epsilon r^l$$

and thus the non linear term $I_L(\epsilon)$ is negligible when $\epsilon \gg \frac{1}{1+r^l} \approx 1$. For $x \ll 0$ the external concentration is the steady state concentration as found in a region of producers only, i.e. $\epsilon(x \ll 0) = I^C \gg 1$ (see subsection 2.4.2). Close to the interface, we expect $\epsilon(x \approx 0)$ to be of the order of I^C for continuity. Thus, close to the interface we can solve a simplified linear ODE:

$$\frac{d^2\epsilon}{dx^2} \approx \frac{1}{r_0^2} \cdot \epsilon.$$

The full solution is thus given by:

$$\epsilon(x) = \begin{cases} C_1 \cdot e^{x/r_0} + I^C & \text{if } x < 0 \\ C_2 \cdot e^{-x/r_0} & \text{if } x > 0. \end{cases}$$

We can solve for C_1 and C_2 imposing continuity of concentration and flux at the interface:

$$C_1 \cdot e^{x/r_0} + I^C|_{x=0} = C_2 \cdot e^{-x/r_0}|_{x=0}$$
$$\tfrac{C_1}{r_0} \cdot e^{x/r_0}|_{x=0} = -\tfrac{C_2}{r_0} \cdot e^{-x/r_0}|_{x=0}.$$

From which we find that $C_1 = -\frac{I^C}{2}$ and $C_2 = \frac{I^C}{2}$. Thus the external concentration is given by:

$$\epsilon(x) = \begin{cases} I^C\left(1 - \frac{1}{2}\cdot e^{x/r_0}\right) & \text{if } x < 0 \\ \frac{I^C}{2}\cdot e^{-x/r_0} & \text{if } x > 0 \end{cases} \tag{29}$$

Thus within the consumer region the amino acid concentration ($E = \frac{r^l}{r^u + r^l}\epsilon$) decreases exponentially with scale factor r_0.

We are now interested in finding an analytical approximation for the *growth range* (GR), which is the distance from the interface where cells have 50% of the growth rate they have at the interface:

$$\mu(x = GR) = \frac{1}{2}\cdot\mu(x = 0) \tag{30}$$

as $\mu = \frac{I}{1+I}$ it follows that

$$I_L(x = GR) = \frac{I_L(x = 0)}{2 + I_L(x = 0)} \tag{31}$$

Where $I_L(x) \equiv I_L(\epsilon(x))$ is given by eq. 28. Substituting $\epsilon(x) = \frac{I^C}{2}e^{-GR/r_0}$ gives:

$$\begin{aligned} &\frac{r^l}{4}(I^C e^{-GR/r_0} - 2) + \tfrac{1}{2}\sqrt{(\tfrac{I^C}{2}e^{-GR/r_0} - 1)^2(r^l)^2 + 2r^l I^C e^{-GR/r_0}} \\ &\qquad = \frac{\frac{r^l}{2}(I^C - 2) + \sqrt{(\frac{I^C}{2} - 1)^2(r^l)^2 + 2r^l I^C}}{4 + \frac{r^l}{2}(I^C - 2) + \sqrt{(\frac{I^C}{2} - 1)^2(r^l)^2 + 2r^l I^C}} \end{aligned} \tag{32}$$

Which we can solve for GR to find:

$$GR = r_0 \cdot \ln\left[\frac{(I^C - 4)\left(r^l(I^C + 2) + \sqrt{(I^C - 2)^2(r^l)^2 + 8I^C r^l}\right) + 8I^C}{2(I^C(1 + r^l) - 4r^l)}\right] \tag{33}$$

and we can further simplify using assumption 2.4.1 ($I^C \gg 1$) to find:

$$GR = r_0 \cdot \ln\left[\tfrac{1}{2}r^l I^C\left(1 + \sqrt{1 + \tfrac{8}{r^l I^C}}\right) + 4\right] \tag{34}$$

Figure S6b shows the error of our analytical approximation of the growth range. As long as the growth range is smaller than 20, our simulations match the analytical result very well. As the growth range approaches 20 (Fig. S6a, this is same heat map as Fig 4f), the relative error increases because of the finite size (40x40) of the chamber in our

simulations; more specifically, the no-flux boundary conditions lead to overestimation of the growth range in the simulations compared to the analytical model.

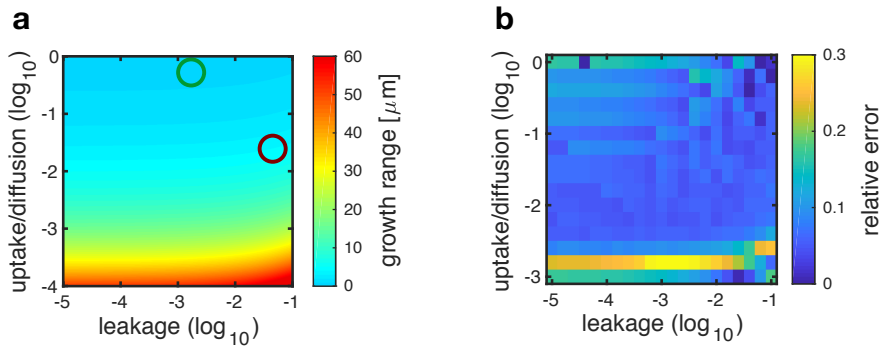


Figure S6. **Analytical approximation of growth range and simulations agree.** The heat map (b) shows the relative error between analytical approximation (eq. 34) for the growth range and the growth range estimated with simulations. The heat map (a) shows the analytical estimate of the growth range, and shows that the relative error in (b) is low when the growth range is below 20. Red circle is proline auxotroph and green circle is tryptophan auxotroph. Leakage is expressed in normalised units (units of μ^{max}).

2.5. **Discussion on the effect of parameters.** We have found two analytical approximations for the maximum growth rate of each auxotroph when surrounded by a large number of the amino acid producing partner and for the growth range:

$$\mu^{max} \approx \frac{r^l I^C \left(1+\sqrt{1+\frac{4}{r^l I^C}}\right)}{2+r^l I^C \left(1+\sqrt{1+\frac{4}{r^l I^C}}\right)} \tag{35}$$

$$GR \approx \sqrt{\frac{D^{eff}}{\alpha(r^u+r^l)}} \cdot \ln\left[\frac{1}{2} r^l I^C \left(1+\sqrt{1+\frac{8}{r^l I^C}}\right)+4\right] \tag{36}$$

We can make some observations:

- The maximum growth rate does not depend on the uptake rate of amino acids but only on the leakage rate.
- The growth range depends strongly (square-root) on the uptake rate and the diffusion constant and weakly (logarithmic) on the leakage rate.
- The cell density strongly affects the growth range by modulating the effective diffusion constant and ratio between *intra* to *extra* cellular environment.

To remind the reader: $\frac{D^{eff}}{\alpha}$ depends on the cell density ρ, which affects both the diffusion constant $D^{eff} = \frac{1-\rho}{1+\rho/2} \cdot D$ and the volume ratio of *intra* to *extra* cellular environment $\alpha = \frac{\rho}{1-\rho}$. So

$$\frac{D^{eff}}{\alpha} = \frac{2(1-\rho)^2}{\rho(2+\rho)} \cdot D \tag{37}$$

Figure S7 shows how D^{eff} and $\frac{D^{eff}}{\alpha}$ depend on density ρ.

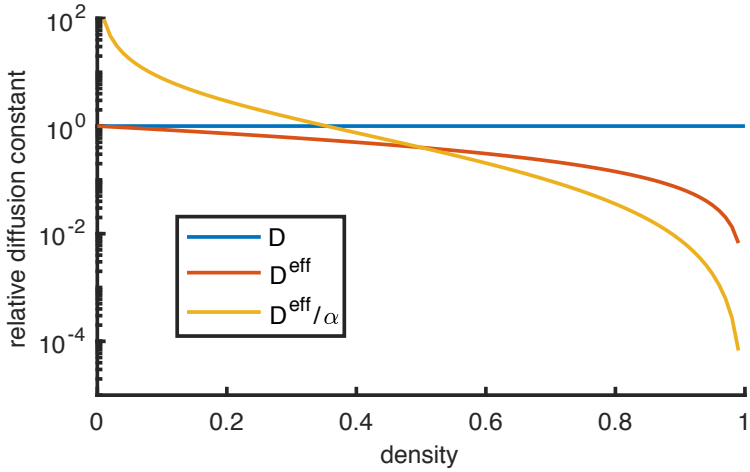


Figure S7. **Effect of density of cells on diffusion of molecules.** Dependence of D^{eff} and $\frac{D^{eff}}{\alpha}$ on density ρ. High cellular densities reduce the effective diffusion of molecules. In our microfluidic chambers cellular density is about 0.65.

3. Supplementary methods

3.1. **Proportionality of analytical growth range and interaction range.** Given a set of parameters, we calculated analytically the growth range using Equation 34 and we estimated the interaction range with the model. The interaction range was estimated as follows: we ran our model on experimentally observed spatial arrangements after downscaling the segmented images to a 40x40 grid. We used the model predicted growth rates and repeated the correlation analysis described in Methods to extract the predicted interaction range. Figure 4e (proportionality between growth range and interaction range) is made by changing the uptake of the amino acids and keeping all other parameters fixed (see Table S1).

4. Supplementary Tables

We list here all parameters of the individual-based model with their source.

Parameter	Description	Value	Source
r_1^u	uptake of proline	1.37 $1/sec$	Literature (2)
r_2^u	uptake of tryptophan	21.9 $1/sec$	Literature (3)
D_1	diffusion of proline	8.79*10^2 $\mu m^2/sec$	Literature (4)
D_2	diffusion of tryptophan	6.59*10^2 $\mu m^2/sec$	Literature (5)
μ^{max}	growth on M9 media + 0.2% glucose	1.29 $1/h$	Measured
r_1^l	leakage proline	1.61*$10^{-5}1/sec$	Fitted (see Sup. Eq.)
r_2^l	leakage tryptophan	6.08 *$10^{-7}1/sec$	Fitted (see Sup. Eq.)
ρ	density of cells	0.65	Measured
dX	grid (cell) size	1.5 µm	Estimated from number of cells per chamber

Table S1. **Parameters of individual-based model.** All parameters of the model are taken from literature or measured, apart from the two leakage rates, which are estimated as described in section 2.4.3

5. Captions for Movies

Supplementary Movie S1: Two auxotrophic strains of *Escherichia coli* growing in microfluidic chambers of 60x60 μm. Proline auxotrophic cells are shown in red, tryptophan auxotrophic cells in green. The auxotrophic cells grow faster when they are close to the complementary partner.

Supplementary Movie S2: The growth rate of auxotrophic cells depends on the identity of neighbours. Cells are coloured based on their growth rate (lighter colours indicate higher growth rates). Growth rates are higher for auxotrophic cells close to the complementary partner. Proline auxotrophic cells are shown in purple, tryptophan auxotrophic cells in yellow.

6. References

1. Kovarova-Kovar, K. Egli, T. (1998) Growth Kinetics of Suspended Microbial Cells: From Single- Substrate-Controlled Growth to Mixed-Substrate Kinetics. *Microbiol.*

Mol. Biol. Rev. 2:646-666 .

2. Grothe, S., Krogsrud, R. L., McClellan, D. J., Milner, J. L. Wood, J. M. (1986) Proline Transport and Osmotic Stress Response in Escherichia coli K-12. *Journal of Bacteriology*. 166:253-259

3. Piperno, J. R., Oxender, D. L. (1968) Amino Acid Transport Systems in Escherichia coli. *Journal. Biol. Chem.* 243:5914-5920 .

4. Wu, Y., Ma, P., Liu, Y. Li, S. (2001) Diffusion coefficients of l-proline , l-threonine and l-arginine in aqueous solutions at 25°C. *Fluid Phase Equilibria.* 186:27-38.

5. Longsworth, L. G. Diffusion (1953) Measurements, at 25°C, of Aqueous Solutions of Amino Acids, Peptides and Sugars. *Contrib. from Lab. Rockefeller Inst. Med.* Res. Nov. 20.