An agent-based simulator for microbial communities' evolution

Massimo Bellato[1], Marco Cappellato[1], Andrea Calzavara[1], Sara Rebecca[1], Alessandro Lucchiari[1], Barbara Di Camillo[1]

[1]Department of Information Engineering, University of Padova, Padova.

Email of Corresponding authors: massimo.bellato@unipd.it, barbara.dicamillo@unipd.it

Abstract

In this work, we blueprint a Dashboard which allows users to simulate bacterial community's evolution through a intuitive GUI. The underlying Python-coded simulator implements an agent-based model of bacterial species, nutrients and environment, allowing full customization and upgradability of the tool, due to its intrinsic modularity. Specifically, the model aims to represent discretized spaces, hosting certain number of bacteria for each species and a defined amount of nutrients characterizing the surrounding environment. Bacteria can migrate from a spatial unit into another, looking for different nutrients (i.e., metabolites) across the whole space path. Their growth and survival are governed by their metabolism, which is in turn function of the metabolites present in each specific spatial unit at a certain time point. Thus, the simulator shows how bacteria consume and produce metabolites, following species-specific metabolism rules, letting the system dynamically evolve through bacterial growth, death, spatial migration and continuous updates of the available metabolites pool.

Introduction

Human gut microbiota composition and behavior has been proved to be fundamental for our health; thus, proper modeling of microbial ecology is a hot topic in computational, systems and synthetic biology [1-3].

Classical ecological models, such as Generalized Lotka Volterra models (GLVm), describe the absolute abundance of each community mainly as function of intrinsic growth rate and pairwise interactions with each community members; parameters need to be inferred by co-occurrence networks leveraging on bioinformatic approaches, i.e., obtained from metagenomic data and investigated via topological analysis [4]. The inherent biases associated with each experimental method adopted to retrieve data for training the models, motivates the need for standardization across different studies [5-8]. Additionally, these models usually rely on constant parameters to describe microbe-microbe interactions, which are however determined, for example, by the temporally changing biotic and abiotic environment, failing in capturing community-level behaviors when the environment changes significantly over time. Summarizing, those methods still lack in generalizability, poorly considering information on molecular mechanisms generating the observed scenarios as well as environmental-related variables. Nevertheless, spatial and temporal features are completely neglected [9-11].

Another promising approach relies on flux balance analysis (FBA) methods; in this case, information on bacteria metabolisms is included in the analysis through genome-scale models (i.e., structured information of possible reactions happening in a bacteria species based on annotated transcriptomes through metabolic networks), which are used to implement an optimization problem with biomass production as objective function [12]. While several groups are adopting FBA-derived to develop predictive tools, the lack of sufficiently annotated metabolic models is still a major issue; additionally,

similarly to GLVm, this approach need nesting in more complex models to provide spatial information.
Microbe-effector models and, more specifically, agent-based modeling is an inspiring solution to implement efficient and robust methods to describe systems counting for billions of bacterial cells from hundreds of different species. This approach allows to develop mechanistic-level descriptions of simple agents, dynamically interacting with the surrounding environment that can be arbitrarily modeled; then, from agents decoupled behaviors, the complex microbial ecologies should emerge without the need of describing interaction network, which are hard to identify [13-16].
The technology to characterize bacterial metabolisms in a cheap, reliable and high-throughput fashion is still under development. However, we believe that the intrinsically modular and incremental structure of our simulator is generalizable enough to be already used for roughly defined communities, but also to be directly used in a near future, when all the required data for a full characterization of these types of models will be available.

Methods

The space is abstracted as a series of boxes called *cells*, each one containing a specific set of metabolites, connected with the closest neighborhood cells. This allows representing four different possible scenarios (currently, only the first two scenarios have been implemented):

- single cell: a well-mixed space where bacteria and metabolites are equally distributed (e.g., a stirring bioreactor);
- chain of cells: a linear tract where fluxes of metabolites and bacteria occur in 1 dimension (e.g., representing the intestinal tract with peristaltic movements);
- plane of cells: a surface where movements occur in 2 dimensions (e.g., a Petri dish with chemotaxis-driven propagation);
- 3D geometry of cells: a volume discretized in several cells (e.g., an organoid with random-walk expansion).

Bacterial agents (i.e., bacteria of different species) can process the metabolites in a *cell*, grow and move into a connected *cell*, if existent. This is implemented in Python through two classes:

The class *Cell*, whose instances implement the spatial unity, is defined by:

- an array of metabolites available at the current iteration
- a matrix of metabolites produced by the bacteria at the current iteration
- a dictionary of the bacteria in the cell, with the number of bacterial agents for each *Bact* instance

The class *Bact*, whose instances implement bacteria species with specific metabolism, is defined by:

- a string with the species names
- a vector with the metabolisms, indicating for each metabolite if it is consumed or produced
- a vector indicating which metabolites lead to toxic effects
- two integers describing the maximum growth rate and the toxicity level, to be adopted in the Hill equations (Equation 1) for the population growth and death.

$$\mu_{i,j} = \frac{\mu_i^{max}}{1+\left(\frac{k_i \cdot n_{i,j}^{bac}}{f_{i,j}^{tot}(\mathbf{f_j},\mathbf{m_i})}\right)^{\eta_i}}$$

Equation 1: Bacteria growth rate with:
$\boldsymbol{f}_j$*:* j^{th}-cell nutrients; $\boldsymbol{m}_i$*:* i^{th}*-bacterial* species metabolism; $f_{i,j}^{tot}$: consumed metabolites; $n_{i,j}^{bac}$: *number of bacteria;* μ_i^{max}, k_i, η_i*: Hill parameters.*

It is worth to note that, while the adoption of a Hill's equation to model bacterial growth as a function of the processed metabolites is purely empirical, the modularity of the class allows to easily change this function into an alternative one.

For example, considering a linear tract representing an intestine tract: at each iteration, bacteria can migrate from the i^{th} to the $(i+1)^{th}$ cell (to mimic the direction of chime), grow, die, metabolize the available metabolites, and produce a new set of metabolites for the next iteration (Figure 1). In addition to the accessor and set methods – which are fundamental for the interface between the simulator model and the GUI – the *Cell* class includes methods to update the concentration of metabolites and the number of bacteria in each cell at each iteration. Analogously, methods for the class *Bact* include methods to compute the variation in each species' abundance, depending on the nutrients available in a certain cell and the presence of toxic metabolites. The food update involves the sampling of the nutrients from a multivariate hypergeometric distribution so to model the propensity of remaining nutrients to be consumed by the most abundant and highly metabolizing species. Lastly, for what concern the chain of cell implementation, a method describing the flux of microbes at each interaction from one cell to the next one is implemented by fixing the number of bacteria flowing as proportional to the species abundance and used to realize the system evolution. Again, thanks to the modularity of the implementation, this method can be easily adapted to realize different spatial driving forces or constraints, such as chemotaxis or random walk.

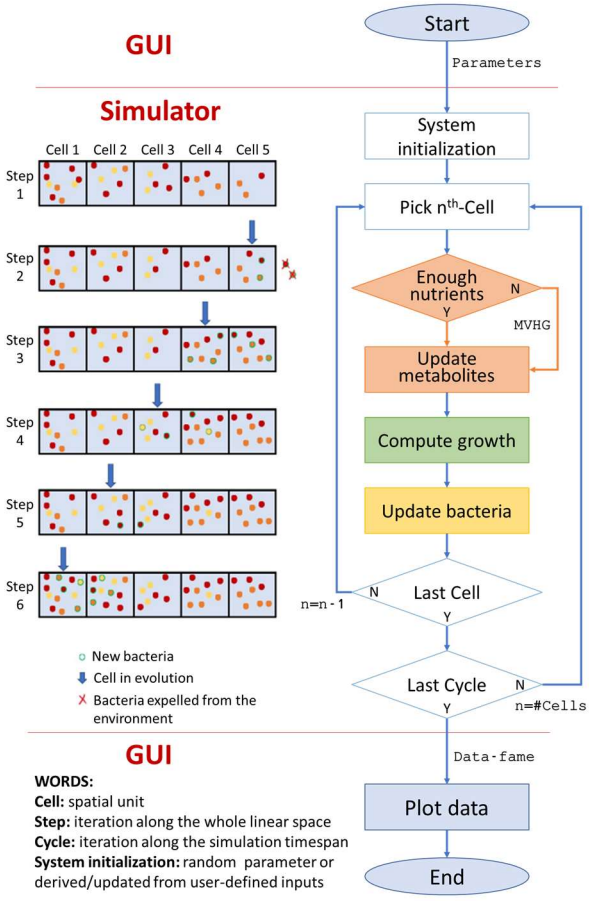

Figure 1: *Example of the simulation framework for a chain of cells representing an intestine tract with monodirectional migration due to peristalsis.*

Results

The GUI allows simple visualization of the running simulation parameters, and both nutrients and bacterial composition evolution over time can be observed on the Dashboard (Figure 2)

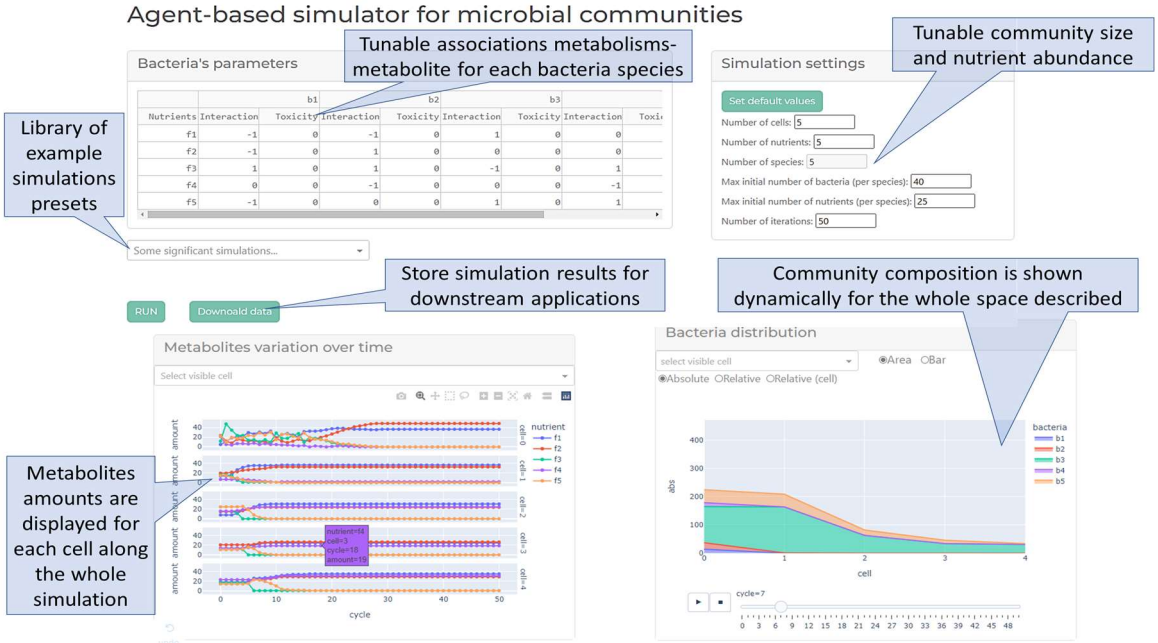

Figure 2: *Example of a toy-simulation implementing a chain of cells representing and intestine tract.*

Starting from randomly generated data (metabolites concentration and bacterial abundance), the model was able to simulate a number of different conditions (Figure 3) such as:

- Perfect fitness: most of the nutrients in a cell are highly metabolized by a specific species due to a favorable environment.
- Bad fitness: no nutrient in a cell is metabolized by any of the species causing a particularly hostile environment.
- Commensalism: most of the metabolites produced by a certain species are metabolized by another specific one and vice versa.

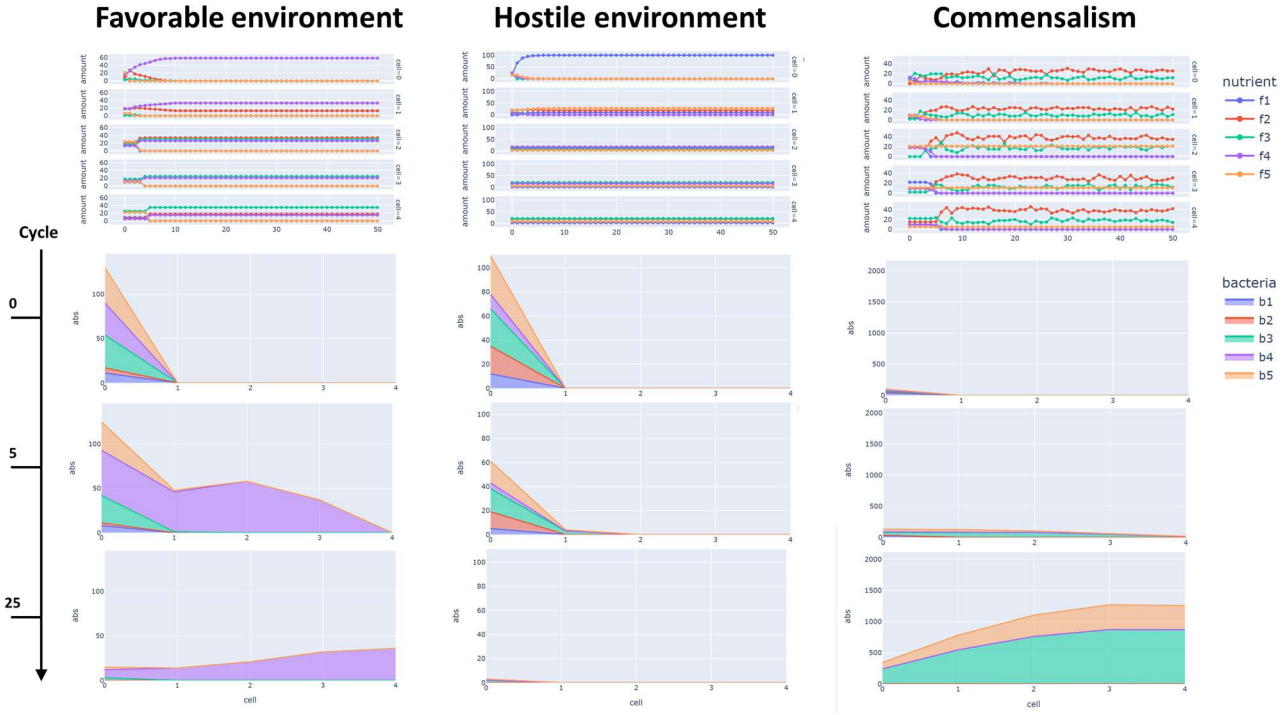

Figure 3: *Results for three biologically relevant community behaviors*

Conclusion

Despite the simplicity of the current implementation, which is still undergoing significant enrichments and refinements, the simulator is already able to provide meaningful results; indeed, biologically sound characteristic community behaviors emerge from random metabolic patterns and nutrient distribution, mimicking a natural evolution.

In addition to spatial evolution refinements, ongoing activities are focused to integrate information on specific nutrient lists and presets of known media composition from curated databases [17-19]. Moreover, information on actual bacterial species and related metabolisms are being added on the basis of data contained in the genome-scale metabolic reconstruction of human gut microbe's dataset [20]. In this way, starting from stoichiometric matrices and filtering on the above-mentioned nutrient lists, it is possible to parametrize nutrient consumption and production.

Finally, further information integration will be easily manageable by users via *ad-hoc* plugins that will be added to the GUI. Indeed, we expect that as soon as further -omics data will be available for species beyond the human-associated ones (e.g., soil and water communities as well as veterinary and agricultural relevant microbiotas), our simulator will allow boosting bacterial community studies, especially for those fields where the complexity of bacterial co-culture methods and media composition variability severely hamper the advance of the research.

Our model allows the description of bacterial communities growing on linear space (e.g., a gut section) Future developments will include the occurrence of mutations (i.e., biologically acceptable small alterations in the metabolisms).

Acknowledgements

This work was partially funded by: (i) Fondazione Cariparo grant "Bando Ricerca Scientifica di Eccellenza 2021 n 59576" ReActing Restoring Antibiotic sensitivity in Bacteria: a synthetic biology approach; (ii) Department of Information Engineering of the University of Padova grant n DI_C_BIRD2020_01 RECENTRE tRajectoriEs of baCtErialNeTwoRks from hEalthy to disease state and back.

Related literature

[1] Davar D, et al. Fecal microbiota transplant overcomes resistance to anti–pd-1 therapy in melanoma patients. Science, 371(6529):595–602, 2021.
[2] Turnbaugh PJ, et al. A core gut microbiome in obese and lean twins. Nature, 457(7228):480–484, 2009.
[3] Zheng P, et al. Gut microbiome remodeling induces depressive-like behaviors through a pathway mediated by the host's metabolism. Mol Psychiatry, 21(6):786–796, 2016.
[4] Bucci V, et al. MDSINE: Microbial Dynamical Systems INference Engine for microbiome time-series analyses. Genome Biol, 17:1-17, 2016.
[5] Barlow JT, et al. A quantitative sequencing framework for absolute abundance measurements of mucosal and lumenal microbial communities. Nat Commun, 11(2590), 2020.
[6] Jian C, et al. Quantitative PCR provides a simple and accessible method for quantitative microbiota profiling. PLoS One, 15: e0227285, 2020.
[7] Tourlousse DM, et al. Synthetic spike-in standards for high-throughput 16S rRNA gene amplicon sequencing. Nucleic Acids Res, 45(e23), 2017.
[8] Vandeputte D, et al. Quantitative microbiome profiling links gut community variation to microbial load. Nature, 55:507-511, 2017.
[9] Faust K, et al. Microbial Co-occurrence Relationships in the Human Microbiome. PLoS Comp Biol 8(7), 2012.
[10] McGregor K, et al. MDiNE: a model to estimate differential co-occurrence networks in microbiome studies. Bioinformatics, 36(6), 2020.
[11] Ren Z, et al. Bacterial Communities Present Distinct Co-occurrence Networks in Sediment and Water of the Thermokarst Lakes in the Yellow River Source Area. Front Microbiol, 2021.
[12] Edwards J, et al. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. Nat Biotech, 19(2):125–130, 2001.
[13] Qian Y, et al. Towards a deeper understanding of microbial communities: integrating experimental data with dynamic models. Cur Op Microbiol, 62:84–92, 2021.
[14] Bauer E, et al. Bacarena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. PLoS comp biol, 13(5):e1005544, 2017.
[15] Jayathilake R, et al. A mechanistic individual-based model of microbial communities. PloS one, 12(8):e0181965, 2017.
[16] Marsland R, et al. The community simulator: A python package for microbial ecology. Plos one, 15(3):e0230430, 2020.
[17] Overbeek R, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res, 2014
[18] Brettin T, et al. RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep, 2015
[19] Söhngen C, et al. BacDive--the Bacterial Diversity Metadatabase. Nucleic Acids Res, 42. 2014
[20] Magnúsdóttir S, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. Nat Biotechnol, 2017