Approximate Bayesian Computation for Parameter Estimation in an Agent-Based Model of Antibiotic Efflux by *Escherichia coli*

Lee Talman[1], Shayn M. Peirce, Ph.D.[1]

[1]Department of Biomedical Engineering, University of Virginia

MICROVASCULAR ENGINEERING LAB
UNIVERSITY OF VIRGINIA

UNIVERSITY of VIRGINIA
SCHOOL of ENGINEERING & APPLIED SCIENCE
Department of Biomedical Engineering

ABSTRACT

The global emergence of multidrug-resistant Gram-negative bacteria, such as *Escherichia coli*, is a growing threat to antibiotic therapy. Clinically-relevant drug efflux mechanisms in individual *E.coli* cells greatly contribute to antibiotic resistance and present a major challenge for antibiotic development, but their spatiotemporal effects on whole bacterial colonies remain elusive while key parameters for modeling these effects are unknown. We present an agent-based model of antibiotic efflux by *E.coli* that utilizes Approximate Bayesian Computation (ABC) to estimate key parameters for volume exclusion, growth, and efflux in individual cells.

- Bayesian Networks
- Bayesian Computation
- ***Better Parameter Estimation***

Building ABMs → Applying ABMs

- Random Forest and NN learning
- Multi-dimensional clustering
- ***Identifying emergent behavior***

AcrAB-TolC efflux pumps in *E.coli*

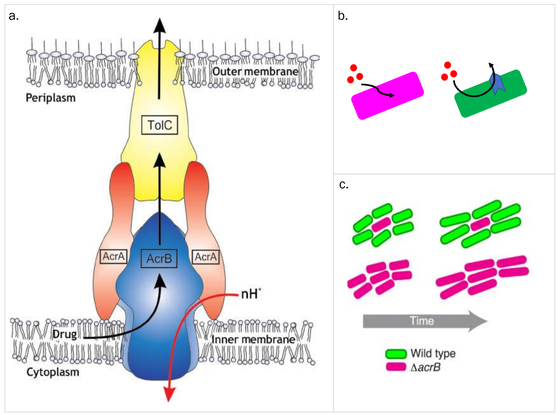


Figure 1 (a): A schematic representation of AcrAB-TolC efflux pumps expressed in *E.coli* and other gram-negative bacteria. **(b)** efflux pump expression confers multidrug resistance **(c)** Schematic showing when Δ*acrB* cells are surrounded by cells with AcrAB-TolC pumps, they grow more slowly than when surrounded by other Δ*acrB* cells when both groups are treated with antibiotics. This is due to the increased local concentration of antibiotic from neighbor cell efflux.

Spatial Agent-Based Rules

$$\frac{dC_{in}}{dt} = \frac{1}{6}\left(\sum_{j=1}^{all\ adjacent\ Neighbors}\left(\frac{1}{2}K_{out,j}+\frac{1}{2}K_{in}\right)C_{in,j}+\sum_{k=all\ adjacent\ Neighbors+1}^{4}K_{in}C_{out}\right)+\frac{1}{12}\left(\sum_{j=1}^{all\ diagonal\ Neighbors}\left(\frac{1}{2}K_{out,j}+\frac{1}{2}K_{in}\right)\times C_{in,j}+\sum_{k=all\ diagonal\ Neighbors+1}^{4}K_{in}C_{out}\right)-K_{out}C_{in}$$

$$\frac{dN}{dt}=\mu\cdot N\cdot\left(\frac{1}{1+\left(\frac{C_{in}}{K_c}\right)^{h_c}}\right)$$

Figure 2: *Wen et al.* (2018) proposes a model of two-dimensional, grid-based efflux antibiotic diffusion in *E.coli* colonies that accounts for changes in antibiotic concentration *C* due to uptake and efflux with a change in cell biomass *N*. We adapt this function for our agent-based model logic (see *Figure 3*)

X. Wen, A. M. Langevin, and M. J. Dunlop, "Antibiotic export by efflux pumps affects growth of neighboring bacteria," *Sci Rep*, vol. 8, no. 1, pp. 1–9, Oct. 2018.

MODEL DESIGN

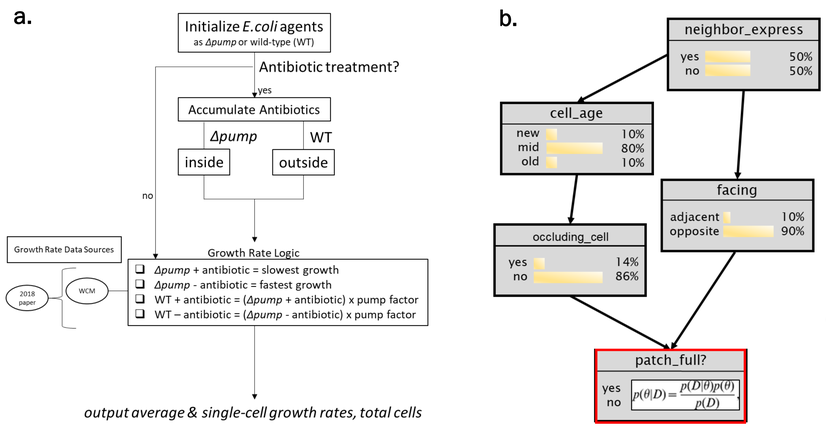


Figure 3 (left): (a) Schematic of the model base logic, which initializes *E.coli* cells as agents, then accumulates antibiotics in their local environment that are then either taken up by the cell or removed by pumps. **(b)** Model parameters as decision "nodes", with probabilities that indicate the likelihood of each agent acquiring a parameter value based on a node's probability and the previous nodes in the directional chart. **(c, below)** Nodes that cannot be derived directly from data are computed using Approximate Bayesian Computation (see panel below). A representative example is selected here – the *patch_full* node, which calculates the amount of cells that can fit in a given area of the model.

c.

i	θ_i	$S(y_0)$	Δ	Choice
1	**2**	**0.0091**	**0.0009**	**Accept**
2	5	0.0307	0.0086	Reject
3	1	0.0018	0.0082	Reject
4	8	0.1441	0.1341	Reject
5	**3**	**0.0121**	**0.0021**	**Accept**

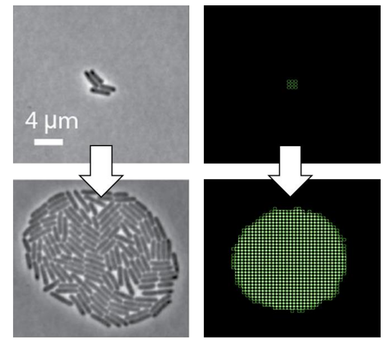
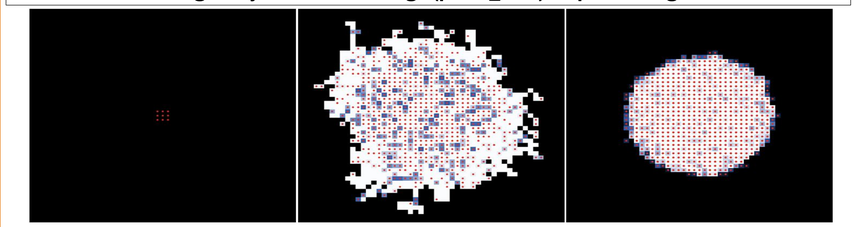


Figure 4 (above): Screenshot of the agent-based model running in NetLogo 6.0.4 compared to a plate culture of *E.coli* with color-based volume exclusion turned off.

RESULTS

Effect of single Bayesian node change (patch_full?) on predicted growth rate



no volume exclusion	default volume exclusion (1 patch holds 1 cell)	result of **patch_full?** network
Normalized growth rate (min^{-1})		**Published Growth Rate: 0.01**
0.0451	0.0018	0.0121

Figure 5 (above): A representative example of how ABC can assist in parameter estimation for agent-based models. In our model, the number of cells per "patch" (discrete 2D area) cannot be derived from real data, as the patches do not correlate with dimensions in real space. However, using the published growth rate of *E.coli* we can approximate a value for the distribution of *patch_full?*, *i.e.* we can estimate the probability that a patch will be available to store a new *E.coli* agent based on the number of cells that already exist in the patch, how old they are, where they are facing, and whether or not their neighbors express efflux pumps. Using the technique shown in the panel below, we can estimate the parameter without potentially intractable parameter sweeps.

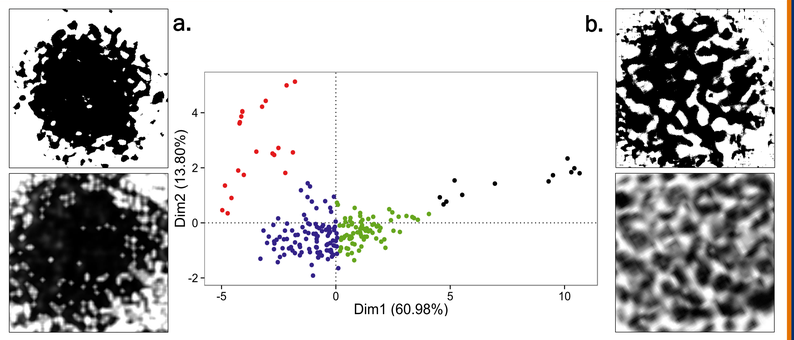


Figure 6 (above): Principle component analysis (PCA) of 150 model runs performed using different starting positions of *E.coli* cells with and without efflux pump expression. In all model runs, the number of cells expressing efflux pumps remains the same, but their starting locations are different. Clusters are separated by color. The red cluster is associated with the highest mean growth rate of all colored clusters, and shows circular arrangement of efflux-expressing cells (black) in both individual starting conditions (**a, top**) and an overlay of all 19 points in the cluster (**a, bottom**). Meanwhile, the other clusters show a more uniform distribution of efflux-expressing cells both in individual data points (**b, top**) and an overlay of all 131 points in the cluster (**b, bottom**).

Approximate Bayesian Computation (ABC)

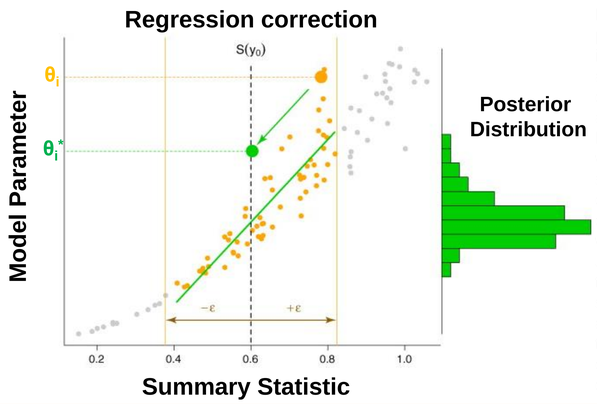


Figure 7: Sampling a parameter value θ_i from its prior distribution allows us to sample a dataset y_i using the agent-based model. From this simulated data we compute a summary statistic $S(y_i)$ (in our case, mean growth rate of colonies) and compare with an observed statistic $S(y_0)$. Values within tolerance range ε are sampled and adjusted according to a linear transform θ_i^* (green) to compute a posterior distribution of θ_i

CONCLUSIONS & FUTURE WORK

Conclusion: Approximate Bayesian Computation (ABC) utilized in a novel way for parameter estimation in Agent-Based Models (ABMs)

BNs let us use probabilistic models to fit parameters that are not measurable in literature (e.g. translating model space to real two-dimensional volume) – more consistent mathematical approach than traditional "best guess by experts"

Future Work: Machine learning using combined ABM and ABC models to develop image-based prediction tools

ML algorithms take the same data gained from parameter sweeps and adds clustering for validating known behavior & learning to identify emergent behavior. ML doesn't take away from existing workflow, only adds to it. We are using the PCA results in Figure 6 to train a random forest ML algorithm to read image data and predict overall growth rate of colonies based on efflux pump expression.

Takeaway: multiscale models that use mathematical approaches should also be built & designed in a mathematical way, and we are enabling that!

ACKNOWLEDGEMENTS & FUNDING

NIH
THE PAUL G. ALLEN FRONTIERS GROUP
ALLEN DISCOVERY CENTER at Stanford University