The Mathematics of Microbial Evolution: Beyond the Limits of Classical Theory

Alexander, Helen (University of Edinburgh, United Kingdom),
Gandon, Sylvain (CNRS Montpellier, France),
Wahl, Lindi (Western University, Canada)

January 8 - 13, 2023

1 Overview

Microbial populations drive human health and disease, as well as ecosystem function and global biogeochemistry. With large population sizes and relatively short generation times, microbes can evolve rapidly: developing resistance to new antimicrobial drugs, infecting new host species (e.g. SARS-CoV-2), or emerging as highly pathogenic new strains (e.g. pandemic influenza). A deeper understanding of the dynamics of microbial populations, both around and within us, is urgently needed to control the spread of novel pathogens in a globally changing environment.

Historically, mathematical modelling has played a central role in the understanding of population dynamics and evolutionary processes [1, 2, 3, 4]. Models have largely been developed for organisms that reproduce sexually, carrying one maternal and one paternal copy of each gene, in populations of roughly constant size, in environments to which the population is well-adapted. Mathematically, this focus has facilitated progress in understanding evolutionary dynamics through the use of asymptotics and a standard tool-kit of simplifying assumptions. In microbial evolution, however, many of these key assumptions no longer hold. Mutation rates can be high [5] and vary according to time [6] and environment [7]. Single mutations, for instance resistance to lethal antimicrobials, can confer benefits that are orders of magnitude larger than those observed in higher organisms. Microbes reproduce both sexually and asexually, share genes promiscuously across organism and species boundaries [8, 9, 10], and can carry from one up to hundreds of copies of some or all genes [11, 12, 13]. Population sizes are large and highly variable in time. These complexities of microbial evolution not only break the standard asymptotic assumptions of evolutionary models, but demand entirely new approaches: both new models, and new tools for their analysis.

On the empirical side, technological developments (e.g. genetic sequencing, single-cell microscopy, genetic engineering) have facilitated a burgeoning understanding of microbial genetics, physiology, and 'lifestyle'. Understanding their consequences for microbial population dynamics and evolution calls for theoretical approaches, including development of new models and mathematical methods to address deviations from standard assumptions. However, these cutting-edge biological insights do not always reach mathematical modellers, due to the obstacles of disciplinary divides and specialist terminology. Conversely, experimental approaches provide powerful opportunities to test theoretical predictions.

A key aim of this workshop was therefore to foster a meaningful intersection between applied mathematics and recent experimental discoveries and methodological advances in microbiology. Towards this goal, we invited participants working across the theory-experiment spectrum to share their recent work and open

questions. The size and format of the workshop created an excellent venue for exchange between a diverse group of mathematicians/scientists who may not normally meet.

We also invited participants with attention to diversity and with the aim of including early-career researchers. While it can be difficult to identify appropriate ECRs at the proposal stage, one successful approach to expanding our participant list following workshop acceptance was to ask any originally proposed participants who declined the invitation (e.g. due to scheduling conflicts) to nominate a more junior scientist in their place. In the end, our 35 in-person workshop participants included 60% women and 29% graduate students or postdoctoral researchers. Longer talks (see next section) were selected by proposed title, blind to participant name, and turned out to reflect the gender balance of participants as a whole.

2 Workshop Structure

The workshop started with a series of 22 "lightning talks" (5 minutes) where all the participants not giving a longer talk were invited to introduce themselves and describe a single research topic they were working on. These lightning talks were grouped into three distinct topics, facilitating themed discussions:
(i) measurement and evolutionary consequences of the mutation rate,
(ii) life-history evolution of viruses,
(iii) infectious disease dynamics
plus several diverse talks on microbial evolution falling outside these categories.

By the end of the second day all participants (almost) had had a chance to present themselves and this was a great way to (1) meet everybody and (2) let open problems and topics emerge for the discussion groups (see below).

We also invited 14 longer presentations (30 minutes) that allowed some of the participants to spend more time describing a specific research project. The long talks and Q&A after the long talks were available to the 15 online participants who registered for the workshop.

Group discussions occurred after each group of lightning talks and after each long talk.

3 Break-out research groups

A unique feature of our meeting was that participants were invited to "pitch" topics and research ideas for break-out discussion groups. After these research questions were pitched to the group, participants chose break-out groups to join and engaged in two discussion sessions. Each group then reported back to the conference as a whole, before a second set of groups was formed (some research topics continued and some new topics were added at this stage).

Many participants commented in the exit survey that a number of new research ideas were formed during these break-out discussions, along with sharing of relevant papers, approaches and potential datasets. A Slack group was created for the workshop participants, with dedicated channels for some discussion groups to facilitate further post-conference interaction and resource-sharing.

The research break-out groups that emerged addressed the following topics at the intersection of microbial evolution and mathematical modelling:

1. Including bacterial recombination in theoretical population genetics
2. Biological differences between exponential growth versus stationary phase in bacteria and their implications for modelling
3. Interpretation of mutation rate measured in the lab versus molecular clock rate estimated by phylogenetic methods
4. How to measure mutation rates

5. Evolutionary consequences of noise versus plasticity in phenotypic trait expression
6. How to measure trait variability
7. Long read sequencing technology for viruses
8. Mobile genetic elements and plasmids
9. *In vivo* estimates versus *in vitro* predictions of antibiotic resistance

This final research question (9) generated sufficient interest that a longer, full-group discussion was devoted to this topic toward the end of the week. Following from this wide-ranging and enthusiastic group discussion, the plan is to write an opinion/perspectives article on the factors driving the probability of treatment failure due to evolution of resistance, discussing whether predictions based on *in vitro* measures align with *in vivo* observations of the emergence of resistance. In brief, theory predicts that the probability of treatment failure P_{TF} can be captured by:

$$P_{\mathrm{TF}} = P_{\mathrm{trans}} + (1 - P_{\mathrm{trans}})\left(P_{\mathrm{before}} + (1 - P_{\mathrm{before}})P_{\mathrm{after}} \right) \tag{1}$$

where P_{trans} is the probability that resistance to treatment is acquired from a transmission event (i.e., superinfection), $P_{\mathrm{before}} = 1 - e^{-N_{\mathrm{before}}P_e}$ is the probability resistance emerges before the implementation of treatment while $P_{\mathrm{after}} = 1 - e^{-N_{\mathrm{after}}P_e}$ is the probability resistance emerges after the implementation of treatment. These expressions depend on N_{before} and N_{after}, the expected number of mutation events producing the resistant strain before and after treatment, respectively. Also, we assume P_e measures the probability of establishment of a resistant strain (which may or not be the same before and after the start of the treatment).

This approach suggests a suite of open questions. Is it possible to estimate the various parameters that appear in the above expression? Does (1) provide accurate prediction *in vitro*? Can these approximations can be used to evaluate the risk *in vivo*? Can this analysis yield robust recommendations for more durable treatment strategies (i.e. treatment strategies with a lower risk of failure due to pathogen evolution)?

The mix of expertise at our workshop (theory/mathematical experts along with experts in bacterial, viral and fungal evolution) allowed us a uniquely broad view of these questions. At the end of the meeting, we sketched out a rough outline for sections of a perspectives paper and identified lead authors who will coordinate the writing of each section.

4 Feedback from the Meeting

A google form was used as an exit survey and 75% of participants completed the survey. Feedback was overwhelmingly positive. Participants commented on the workshop having led to new collaborations, new mentorship, and new ideas for research projects and/or grant proposals. New research connections between experimentalists and mathematicians were particularly highlighted. Quotations from the anonymous exit survey include:

"I have three specific ... research directions that have come out of conversations in this meeting that I attend to follow up with ... None of these research directions existed before the meeting."

"the main foci of my group's research for at least the next few years will be along the lines set out by this workshop"

"I foresee that I will start collaborations with at least two people and possibly organize a future workshop/conference with one of them on a topic that came up during discussions."

"Invaluable."

"I can't remember the last meeting/week where I feel like I'm walking away with so much new knowledge, paper references to read, and renewed perspectives from different people on some of the things I think about a lot."

"one of the best workshops I participated in recent years (including pre-COVID)."

"Wonderful and impressive leadership from junior and mid-career women throughout the entire conference - this is really unique amongst my experiences at conferences."

5 Conclusions

We feel that our meeting succeeded in its aim to bring together researchers across the theory-experiment spectrum, and constitutes an important first step towards advancing our understanding of microbial evolution "beyond the limits of classical theory". Some of the discussions that emerged will allow the participants to share their expertise and learn about the theory that has already been carried out on specific questions that emerged during the discussions. Other discussions identified open questions and we expect will contribute to developing new theoretical frameworks to tackle new biological questions. While it is too early to describe the output of these discussions, the format of the conference worked very well. The "mathematics of microbes" is a very active and dynamic field of research but there is little opportunity for interactions between biologists and mathematicians in a relatively small workshop. Given the high interest and enthusiasm for follow-on projects in this fast-developing field, we feel it would be particularly relevant to carry out a similar conference in 4-5 years. In this eventuality, the new organizers could build on the lessons learned about workshop format in this first iteration (including the organizers' observations and participants' feedback in the survey) to repeat what worked well and tweak what might be improved.

As noted in the previous section, the feedback from this meeting was overwhelmingly positive. New research ideas and collaborations emerged for many participants, and a group perspectives paper is in the works. Attending a mathematics conference at which the majority of participants and speakers (60%) were female was a unique experience for many, and was particularly important to our 29% graduate student and post-doctoral fellow participants. The facilities at the Banff Centre were ideal for fostering research interaction and collaboration; we are grateful for the amazing opportunity to host a workshop at BIRS.

References

[1] Fisher, R., The genetical theory of natural selection, (1930), Clarendon press.

[2] Crow, J. F. and Kimura, M., An introduction to population genetic theory, (1970), Harper and Row, New York.

[3] M. Nowak and R. May, Virus dynamics: mathematical principles of immunology and virology: mathematical principles of immunology and virology, (2000), Oxford University Press, UK.

[4] Diekmann, O., and Heesterbeek, J. A. P., Mathematical epidemiology of infectious diseases: model building, analysis and interpretation (2000). John Wiley & Sons.

[5] S. Duffy, Why are RNA virus mutation rates so damn high?, *PLOS Biology* **16** (2018), e3000003.

[6] Uphoff, S. (2018) Real-time dynamics of mutagenesis reveal the chronology of DNA repair and damage tolerance responses in single cells. *Proceedings of the National Academy of Sciences*, **115**, E6515–E6525.

[7] R. Krasovec, H. Richards, D. R. Gifford, C. Hatcher, K. J. Faulkner, R. V. Belavkin, A. Channon, E. Aston, A. J. McBain, and C. G. Knight, Spontaneous mutation rate is a plastic trait associated with population density across domains of life, *PLOS Biology* **15** (2017), e2002731.

[8] E. V. Koonin, K. S. Makarova, and L. Aravind, Horizontal gene transfer in prokaryotes: Quantification and classification, *Annual Review of Microbiology* **55** (2001), 709–742.

[9] Nazarian, P., Tran, F., and Boedicker, J. Q. (2018) Modeling multispecies gene flow dynamics reveals the unique roles of different horizontal gene transfer mechanisms. *Frontiers in Microbiology*, **9**, 2978.

[10] Touchon, M., de Sousa, J. A. M., and Rocha, E. P. (2017) Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. *Current Opinion in Microbiology*, **38**, 66–73.

[11] V. Pecoraro, K. Zerulla, C. Lange, and J. Soppa, Quantification of ploidy in proteobacteria revealed the existence of monoploid, (mero-)oligoploid and polyploid species, *PLOS One* **6** (2011), e16392.

[12] Y. Harari, Y. Ram, N. Rappoport, L. Hadany, and M. Kupiec, Spontaneous changes in ploidy are common in yeast, *Current Biology* **28** (2018), 825–835.e4.

[13] L. Sandegren and D. I. Andersson, Bacterial gene amplification: implications for the evolution of antibiotic resistance, *Nature Microbiology Reviews* **7** (2009), 578–588.