

## PERSPECTIVE OPEN



# Quantitative systems-based prediction of antimicrobial resistance evolution

Daniel A. Charlebois <sup>1,2</sup> ✉

Predicting evolution is a fundamental problem in biology with practical implications for treating antimicrobial resistance, which is a complex system-level phenomenon. In this perspective article, we explore the limits of predicting antimicrobial resistance evolution, quantitatively define the predictability and repeatability of microevolutionary processes, and speculate on how these quantities vary across temporal, biological, and complexity scales. The opportunities and challenges for predicting antimicrobial resistance in the context of systems biology are also discussed. Based on recent research, we conclude that the evolution of antimicrobial resistance can be predicted using a systems biology approach integrating quantitative models with multiscale data from microbial evolution experiments.

npj Systems Biology and Applications (2023)9:40; <https://doi.org/10.1038/s41540-023-00304-6>

## INTRODUCTION

Predicting evolution is an important fundamental and practical problem in biology and medicine. The ability to quantitatively predict evolution will advance evolution from a descriptive theory to a predictive theory<sup>1,2</sup> that can tackle global health problems such as *antimicrobial resistance* (AMR)<sup>3</sup> (Box 1). The rapid evolution of pathogens in response to antimicrobial drugs motivates the need to transform evolution into a predictive, quantitative science to develop “evolution proof” drugs<sup>4</sup> and vaccines<sup>5</sup>, which will respectively enable effective treatments and prophylactics to combat AMR.

There is a pressing need to develop a predictive theory of AMR evolution to answer numerous outstanding questions. First, what do *evolutionary predictability* and *evolutionary repeatability* mean in the context of AMR evolution and which aspects of resistance can be forecasted for microbes? These questions are challenging given the multiple scales at which AMR occurs, ranging from the rapid emergence of resistance mutations inside of cells<sup>6,7</sup> to the longer-term establishment of complex resistance-conferring dynamics among microbial communities<sup>2</sup>, along with the fact that AMR is a complex system-level evolutionary process<sup>8</sup>. Next, can we predict evolutionary paths (trajectories) or are we restricted to evolutionary endpoints (outcomes)? Distinguishing between trajectories and outcomes is essential since evolutionary processes can follow different trajectories to arrive at the same evolutionary outcome or they can take similar trajectories but diverge to yield different evolutionary outcomes<sup>9</sup>. Last, which variables are important to develop predictive models of *micro-evolution*? This is critical for model development as well as the design of microbial evolution experiments, which can generate high-dimensional data (e.g., genetic sequence) or low-dimensional data (e.g., phenotypic growth rate).

Throughout this article, we tackle the questions posed above and explore the challenges and opportunities for predicting AMR evolution from a systems biology perspective, which incorporates the complex interactions between different scales and components of evolving microbial systems. We conclude that the evolution of AMR can be predicted (see “Resistance evolution

across the scales” for examples) using quantitative systems-based models informed by multiscale data from high-replicate or high-temporal resolution experiments on evolving microbial systems.

## Predictability versus repeatability in microbial evolution

Researchers lack a shared language to describe evolution<sup>2</sup>. We propose that evolving biological microbial systems be quantified based on their *evolutionary predictability* and their *evolutionary repeatability*. The predictability of an evolutionary process is ultimately a probabilistic statement about a biological system (or ensemble), which can be defined by the existence of a probability distribution (Fig. 1). If a probability distribution can be derived theoretically or obtained empirically, then an evolutionary process can be statistically predicted. For instance, if there is a known distribution of outcomes (e.g., of resistance mutations) for a given antimicrobial drug applied during an AMR evolution experiment or patient treatment, then the ensemble behavior of this system is predictable. Predictive distributions that change over time as biological systems evolve require a “dynamic” (time-dependent) model, as opposed to unchanging predictive distributions that can be described by a “steady-state” (time-independent) model<sup>10</sup>. However, the existence of a predictive distribution does not guarantee that any particular mutation (or other resistance traits such as gene expression level<sup>11</sup>) will be easily repeated or realized in an evolution experiment or during patient treatment. An analogy can be drawn to picking cards from a deck of playing cards; while the distribution of outcomes is predictable, picking a specific card is not easily repeatable.

Evolutionary repeatability is related to the likelihood of occurrence of individual events that constitute a statistical ensemble. Repeatability can be quantified using measures from statistical physics such as entropy (Fig. 1). If an evolutionary trajectory or outcome is highly uncertain, stochastic, or entropic, then it has low repeatability. This will reduce the realization of a specific trajectory or outcome in a microbial evolution experiment or during the treatment of a patient, such as a particular biological replicate or patient acquiring a specific sequence of drug

<sup>1</sup>Department of Physics, University of Alberta, Edmonton, AB T6G-2E1, Canada. <sup>2</sup>Department of Biological Sciences, University of Alberta, Edmonton, AB T6G-2E9, Canada.

✉email: [dcharleb@ualberta.ca](mailto:dcharleb@ualberta.ca)

resistance mutations. Shannon entropy<sup>12</sup>

$$H = E[-\log(p(x))] = -\sum_{i=1}^N p_i \log(p_i) \text{ or } -\int_{-\infty}^{\infty} p(x) \log(p(x)) dx \quad (1)$$

### Box 1 Terminology

**Antimicrobial Resistance (AMR):** AMR arises when bacteria, fungi, viruses, parasites, or other microbes no longer respond to antimicrobial drugs.

**Clade:** A group of genetically related organisms that evolved from a common ancestor.

**Clonal Interference:** Competition between beneficial mutations in different lineages with fixed genetic backgrounds in asexual populations.

**Deterministic:** Describes a process that evolves along the same trajectory to the same future state from a given initial state.

**Epistasis:** The phenotypic expression of a genetic mutation/gene is dependent on the expression of one or more other mutations/genes.

**Evolutionary Predictability:** A statement about the evolutionary trajectory/future state of a biological system, defined by the existence of a probability distribution (see "Limits on predicting antimicrobial resistance" for more details).

**Evolutionary Repeatability:** How likely a particular evolutionary event is to repeatedly occur in a biological system, defined by the likelihood, entropy, or other measures of variability (see "Limits on predicting antimicrobial resistance" for more details).

**Fitness:** The reproductive success of an organism. Growth rates are often used as a proxy for fitness in microbes.

**Fitness Landscape:** A visual representation of the relationship between genotype and fitness.

**Genetic AMR:** Drug resistance that is mediated by genetic mutations or through the acquisition of drug resistance genes.

**Genetic Drift:** The change in gene variant (allele) frequency in a population due to random chance.

**Macroevolution:** Large evolutionary changes such as the emergence of a novel pathogen over long timescales.

**Microevolution:** Change in allele frequency over a short timescale (as compared to macroevolution).

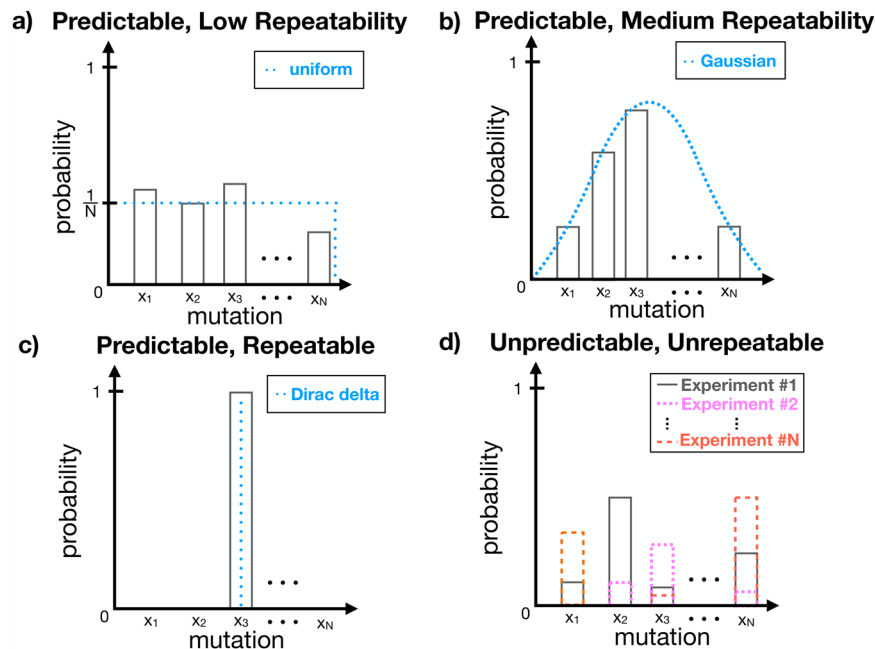
**Nongenetic AMR:** Drug resistance that arises among genetically identical microbes exposed to the same drug, which is not the result of genetic AMR.

**Stochastic:** Describes a process that randomly evolves along one of many possible trajectories to one of many possible future states from a given initial state.

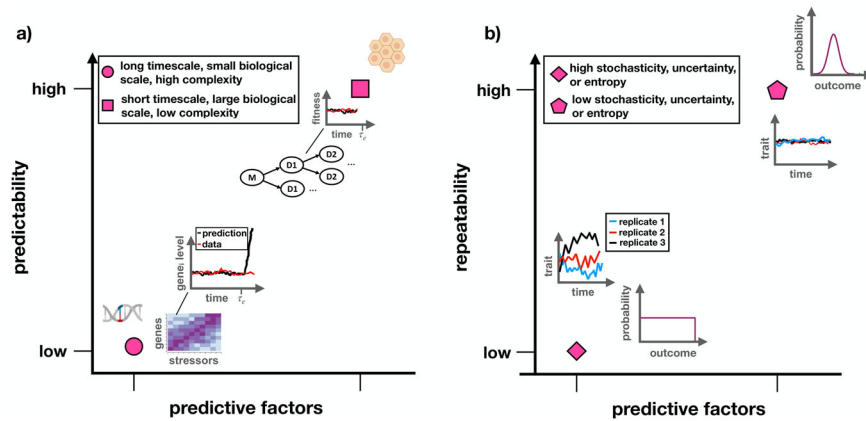
or other measures of variability can be used to quantify repeatability. Note that the entropy associated with a deck of cards is higher than that of a dice, due to the higher dimensionality (i.e., more equally probable outcomes) of the deck of cards compared to the dice. Here, if we consider an ensemble of drug resistance mutations that may appear during an evolution experiment or patient treatment, a uniform distribution represents maximum uncertainty or entropy (Fig. 1a), a Dirac delta function minimum uncertainty (Fig. 1c), and other distributions (e.g., Gaussian) intermediate uncertainty (Fig. 1b) in predictable evolutionary processes. An evolutionary process may also be unpredictable (no distribution describes it) and unrepeatable (the state space of possible outcomes changes each time) (Fig. 1d). AMR evolution is repeatable when different pathogen populations evolve similar resistance to an antimicrobial drug (e.g., acquired mutations in the same genes of the same pathogen infecting different patients)<sup>13</sup> and larger selection pressures are known to generate more repeatable evolution<sup>14,15</sup>. While AMR repeatability has been observed for some large-effect sequence changes<sup>16–18</sup>, it has been more broadly observed for molecular resistance phenotypes<sup>19,20</sup>.

### Limits on predicting antimicrobial resistance

There are fundamental and practical limitations to predicting AMR evolution (Fig. 2). Predictability is fundamentally constrained by random genome mutations and *genetic drift*<sup>21</sup>. This requires that a *stochastic* framework<sup>22</sup> be applied to analyze and make predictions based on data obtained from measurements of biological quantities (e.g., mutations, growth rates, etc.) at low concentrations during early stages or at late stages (if drug treatment is effective) of microbial evolution experiments and infections; genetic drift also plays a confounding role in this regime<sup>9</sup>. Recently, stochastic systems-based population models were proposed to guide drug therapies by providing predictions on resistance mutation appearance probabilities and first-appearance times, indicating timescales for substituting or combining



**Fig. 1 Predictability versus repeatability in antimicrobial resistance evolution.** Schematic of the probability of drug resistance mutations occurring during an evolution experiment or patient treatment. **a** Predictable: described by a uniform distribution. Low repeatability: maximum entropy ( $H_{\max} = \log N$ ). **b** Predictable: described by a Gaussian distribution. Medium repeatability: medium entropy ( $H = \frac{1}{2}(\log(2\pi\sigma^2) + 1)$ , where  $\sigma^2$  is the variance). **c** Predictable: described by a Dirac delta distribution (here  $\delta(x_k) = 1$  and  $\delta(x_{i \neq k}) = 0$ ). High repeatability: minimum entropy ( $H_{\min} = \log(1)$ ), as the same mutation emerges for every experiment or treatment. **d** Unpredictable: no distribution can be derived or fit empirically to the data. Unrepeatable: different outcomes for each experiment or treatment.



**Fig. 2 Limits on evolutionary predictability and repeatability.** **a** The predictability of microevolution decreases at long timescales (e.g., longer than the generation time of a microorganism), small biological scales (e.g., single-nucleotide polymorphisms), and high dimensionality/complexity (e.g., genomic profiles), whereas the predictability of microevolution increases at short timescales (e.g., shorter than the generation time of a microorganism), large biological scales (e.g., whole pathogen populations), and low dimensionality (e.g., phenotypic fitness measurements).  $\tau_e$  is the characteristic time for predicting evolution beyond which predictions become random. **b** The repeatability of evolution decreases as the stochasticity of evolutionary trajectories/outcomes increases, or as the uncertainty/entropy of the associated probability distribution increases. This schematic was made in part using BioRender.com (2023).

antimicrobial drug during patient treatment<sup>23</sup>. Random genetic drift distorts the impact of selective forces and decreases the ability of models to predict evolution<sup>2</sup>.

Other sources of biological stochasticity also pose a challenge for predicting the evolutionary path to resistance, including gene expression variability among genetically identical cells<sup>22,24</sup> and genetic drift<sup>25</sup>. The interplay between *nongenetic AMR* and *genetic AMR* is not fully understood<sup>26</sup>, and such knowledge gaps negatively impact our ability to predict resistance evolution. It has been hypothesized that nongenetic AMR facilitates evolution by allowing a fraction of the cell population to survive and subsequently evolve genetic resistance to antimicrobial treatment<sup>11,27</sup>. Systems-based stochastic population dynamics modeling has shed light on the nongenetic-genetic AMR enigma by incorporating resource competition between nongenetically resistant and genetically resistant subpopulations; these models predict that nongenetic resistance facilitates survival but slows down genetic AMR evolution<sup>23</sup>. Similar models can be inferred from coupled stochastic trajectories of gene expression and cell division<sup>28</sup>. The architecture of systems of interconnected genes comprising genetic networks modulates the amplitude and timescale of gene expression fluctuations, which enhances acute, reversible, nongenetic drug resistance<sup>11,29,30</sup> and facilitates prolonged, nonreversible, genetic drug resistance<sup>6,26</sup>.

The interactions and competition among mutations, along with data limitations, are other important considerations for predicting AMR evolution. *Epistasis* arises from a non-additive interaction between mutations that affect *fitness*<sup>31</sup>. Mutations that arise early in the evolution of AMR affect which mutations are subsequently selected<sup>32</sup>. Epistatic interactions generate a nonlinear *fitness landscape*<sup>33</sup>. This occurs because whether a mutation is beneficial, detrimental, or neutral in terms of selection depends on the genetic, extracellular, and environmental background<sup>1,34</sup>. Beneficial mutations may synergistically interact to decrease selective pressure and render evolution slower and more stochastic (since weaker selection is less *deterministic*), thus reducing repeatability. For instance, beneficial mutations in disjoint *clades* compete for fixation, whereas mutations in nested clades reinforce one another in *clonal interference*, a phenomenon that arises in large asexual microbial populations subject to strong selective pressures such as antimicrobial drug therapy<sup>34,35</sup>. Clonal interference may enhance predictability by reducing stochastic waiting times for fitter genetic variants and by ensuring that beneficial mutations that fix have a large fitness effect and are driven by natural selection,

not by genetic drift or environmental noise<sup>1</sup>. Clonal interference was found to be pervasive in an experimental model of *E. coli* in the mammalian intestine, and the targets of natural selection were similar in independently evolving bacterial populations, resulting in similar early-stage phenotypic changes<sup>36</sup>. Another fundamental challenge in predicting AMR evolution lies in the unknowns, complexities, and redundancies in mapping genotype to phenotype. Our ability to predict AMR evolution on fitness landscapes with considerable epistasis is highly dependent on the precision of the initial conditions data<sup>9</sup>. Epistasis can make predictive modeling more challenging, as the order of and interactions between mutations must be accounted for as well as the fact that epistasis facilitates the evolution of novel functions<sup>37</sup>; on the other hand, epistasis can constrain evolutionary trajectories to a given endpoint thus potentially increasing repeatability. Despite these challenges, metabolic fitness landscapes have been utilized to predict antibiotic resistance<sup>38</sup>. Data limitations further hinder our ability to predict resistance evolution, such as uncertainty in initial conditions for chaotic environmental systems in which microorganisms evolve<sup>39</sup>, even for evolution driven by deterministic natural selection<sup>9</sup>.

It has been argued that biology cannot be reduced to physics, as the complexity of living organisms and the biosphere is too vast to predict how life will evolve<sup>40</sup>. Others have argued that physics can be used to predict evolution on shorter timescales (Fig. 2)<sup>9,41</sup>. Evidence now exists supporting the predictability of AMR microevolution. For instance, in yeast mutations that emerged during evolution experiments on cells harboring synthetic drug resistance gene networks were computationally predicted beforehand based on the costs and benefits of expressing the synthetic gene network in a particular drug condition<sup>7</sup>. In bacteria, the same set of resistance mutations have been found to repeatedly fix in independently evolving populations<sup>18</sup>, suggesting that antibiotic resistance evolution is predictable and repeatable. Antiviral resistance mutations have also been predicted for SARS-CoV-2 using *in silico* mutational scanning inhibitor docking<sup>42</sup>. However, it is unlikely that *macroevolution* can be predicted for practical as well as fundamental reasons. Similar to the Lyapunov time (i.e., the characteristic timescale beyond which the chaotic trajectory of a dynamical system can no longer be predicted<sup>43</sup>), there exists a characteristic timescale for predicting evolving biological systems ( $\tau_e$ ) due to stochasticity in the data and errors in the predictive model<sup>1</sup>. Fluctuating selective pressures can also decrease repeatability due to the evolution on continuously changing fitness

landscapes; the selection timescale for beneficial mutations to emerge is reduced and the outcome of selection becomes more divergent<sup>44</sup>.

### Resistance evolution across the scales

Evolutionary predictability and repeatability vary across temporal, biological, and complexity scales (Fig. 2). Evolution has been shown to be more convergent at “higher” biological scales (e.g., the acquisition of drug resistance genes during treatment) than at “lower” biological scales (e.g., individual resistance mutations within a gene)<sup>45</sup>. For instance, the majority of single-nucleotide and amino acid changes occur in a single population in laboratory evolution experiments<sup>45–47</sup> (i.e., different populations follow divergent paths in sequence space<sup>1</sup>). Additionally, patients respond to an infection or vaccination by different immune receptor sequences<sup>48</sup>. However, even at the genetic scale notable examples exist where microevolution is highly repeatable, such as the establishment of the same mutation due to extensive standing variation<sup>49</sup> or strong selective pressure<sup>50</sup>.

AMR resistance evolution is more predictable when the dimensionality of the state space is reduced. For instance, the state space of all genotypes is high-dimensional, whereas the state space of fitness is one dimensional. Even despite the constraints imposed by the number of paths by selection, genome evolution is generally unpredictable as sequence space contains an effectively infinite number of possible evolutionary paths<sup>1</sup>. On the other hand, AMR evolution is more predictable than general genome evolution due to the finite number of AMR mutations and resistance mechanisms<sup>51</sup>. Phenotype space (e.g., fitness<sup>46,52</sup>) is also more predictable, as different mutations can have similar effects on fitness (i.e., different sequence changes map onto the same phenotypic effects). However, analyzing low-dimensional traits may be less informative than high-dimension traits, as low-dimensional traits are expected to yield a lower information gain relative to a prior, which can be quantified as the log ratio of predicted probability ( $q(t)$ ) to prior probability ( $p(t)$ ):  $I(t) = (q(t)/p(t))^{53}$ . Though few mutations have been found to be shared between replicates in high-replicate bacterial evolution experiments, adaptive convergence has been shown to emerge at higher levels of biological organization, including genes, operons, and functional complexes<sup>45,54</sup>.

An important consideration for predicting AMR evolution is that pathogens evolve within large, highly complex and interconnected ecosystems that are themselves evolving<sup>55</sup>. It remains unclear how ecological diversity in these biosystems, such as multiple pathogenic yeast species co-infecting a patient<sup>56</sup>, or diversity-generating processes such as negative frequency-dependent selection (which favors rare phenotypes over common ones)<sup>9,57</sup>, affect the predictability of AMR evolution. As the complexity of ecological interactions increase, adaptive tradeoffs to multiple selective pressures reduce our ability to predict evolution<sup>58,59</sup>. Though ecological-evolutionary feedbacks can increase our ability to predict evolution through negative frequency-dependent selection<sup>60</sup>.

### Opportunities and challenges for systems-based prediction of antimicrobial resistance

Machine learning is not yet prevalent in evolutionary studies, although it shows promise for predicting AMR. For instance, machine learning models have been trained to predict the evolutionary success of complex systems, such as the human influenza virus variants<sup>61</sup>. However, at present, machine learning models largely function as context-specific “black boxes” that do not elucidate the mechanisms underlying AMR evolution<sup>51</sup>, though this situation may change depending on progress made in the field of explainable artificial intelligence<sup>62</sup>. In contrast, quantitative systems-based models coupled with evolution

experiments on pathogens with well-cataloged resistance mutations would serve as powerful tools for elucidating the mechanisms underlying AMR evolution. It remains to be seen if quantitative systems-based modeling will possess the same predictive power as machine learning.

The performance of quantitative systems-based models trained on data generated from computationally or microbial evolution experiments can be evaluated using information measures<sup>63,64</sup>. This approach may prove valuable in identifying the characteristic time for predicting evolution (Fig. 2) over which evolutionary dynamics can be predicted by quantitative and machine learning models. Testing the predictions made by AMR evolution models in evolution experiments on microbes carrying genetically engineered drug resistance networks<sup>65</sup> would be particularly useful for the iterative “test, build, deploy” cycle that is common in synthetic biology. As drug resistance genes do not function or evolve in isolation, a systems approach rather than a reductionist approach has led to new targets against AMR mediated by genetic networks that control resistance genes<sup>6,29,30,65,66</sup>. For instance, therapeutically targeting regulator genes as opposed to resistance genes has been predicted to mitigate AMR<sup>65</sup>.

A promising avenue is to take a systems-based approach to predict AMR evolution based on properties that are not exhibited by individual resistance genes or microbial species, such as the emergent properties that are associated with interacting genes or microbes<sup>9,67</sup>. Models of rare, random evolutionary events could be validated by performing massively parallel simulations<sup>68</sup> or microbial evolution experiments with many replicates (ranging from tens to hundreds) and analyze a large collection of endpoint measurements<sup>65</sup>. Alternatively, researchers could perform *in silico* or *in vitro/in vivo* evolution experiments with a small number of replicates and track the evolutionary dynamics at a higher temporal resolution. Ideally, these experiments would measure evolutionary trajectories as well as endpoints at multiple scales (including growth assays, imaging, and high-throughput “omics” data generated from genomics, transcriptomics, proteomics, and metabolomics technologies). Another promising approach is to use DNA barcoding, which can detect mutations at low frequencies in cellular lineages and estimate their time of occurrence as well as fitness effects<sup>69,70</sup>. DNA barcoding is particularly helpful at the beginning of microbial evolution experiments when many mutations are present at low frequencies, in contrast to whole genome sequencing which is beneficial at the end of microbial evolution experiments when few resistance mutants have fixed in the population. Genome editing and deep mutational scanning open the possibility of quantifying the number AMR mutations and their individual fitness effects to predict drug resistance<sup>71</sup> and guide drug development<sup>72</sup>. The systems biology approach of combining different levels of omics technologies with genome-scale metabolic models can provide precision and robustness to AMR predictions<sup>8</sup>. Finally, systems-based epidemiological models of AMR evolution that include the structure of the host population, interactions between genetic loci, and integration of within- and between-host levels can predict strategies to limit the spread of drug-resistant pathogens<sup>73</sup>.

### DISCUSSION

Microbial evolution spans temporal, biological, and complexity scales, which differentially impact our ability to predict AMR. We can quantitatively define evolutionary predictability in terms of our ability to define probability distributions and evolutionary repeatability in terms of likelihood or entropy. Evolutionary predictability and repeatability can also be thought of in terms of information gain relative to a prior<sup>64</sup>, which is related to the entropy of the system<sup>74</sup>. The predictability of microevolutionary processes such as AMR increases on short timescales<sup>41</sup>, whereas



macroevolution, such as the emergence of a new pathogen, takes place on timescales well beyond the characteristic timescale for predicting evolution. AMR evolution is anticipated to be more predictable but less repeatable at larger biological scales, while at small biological scales it is less predictable but more repeatable. The complexity of biological systems presents practical challenges for making predictions but is also imposes constraints that can increase predictability. In particular, constraints arising from competing infectious agents<sup>23</sup> or from drug combinations<sup>75</sup> may aid the prediction of resistance evolution.

Predicting AMR evolution is an interdisciplinary endeavor that will require combining multiple fields including physics, mathematics, computer science, and the biological sciences to achieve breakthroughs. One promising strategy is to develop predictive models of evolution using a systems-based approach<sup>76</sup>, which incorporates dynamics across multiple scales ranging from the interactions between genes in a genetic drug resistance network<sup>29,30</sup> to species interactions in microbial communities<sup>9</sup>. The application of machine learning is another approach which promises to play an increasingly important role in drug design, patient diagnostics, and predicting AMR<sup>77</sup>. Ultimately, to predict AMR evolution we need to overcome two challenges: (1) develop quantitative, multiscale models of complex fast-resistance evolving biological systems and (2) generate high-replicate data that span time, space, and biological organization from systems-based evolution experiments to inform and test predictive models. Overcoming these challenges will enable clinicians to predict AMR evolution, revolutionizing the treatment of patients with drug-resistant infections.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Received: 13 June 2023; Accepted: 24 August 2023;

Published online: 07 September 2023

## REFERENCES

- Lässig, M., Mustonen, V. & Walczak, A. M. Predicting evolution. *Nat. Ecol. Evol.* **1**, 0077 (2017).
- Wortel, M. T. et al. Towards evolutionary predictions: current promises and challenges. *Evol. Appl.* **16**, 3–21 (2023).
- O'Neill, J. The review on antimicrobial resistance. *Tackling Drug-Resistant Infections Globally: Final Report and Recommendations*. 1–76 (Wellcome Trust, London, 2016).
- Bell, G. & MacLean, C. The search for 'evolution-proof' antibiotics. *Trends Microbiol.* **26**, 471–483 (2018).
- Micoli, F., Bagnoli, F., Rappuoli, R. & Serruto, D. The role of vaccines in combatting antimicrobial resistance. *Nat. Rev. Microbiol.* **19**, 287–302 (2021).
- Farquhar, K. S. et al. Role of network-mediated stochasticity in mammalian drug resistance. *Nat. Commun.* **10**, 2766 (2019).
- Gonzalez, C. et al. Stress-response balance drives the evolution of a network module and its host genome. *Mol. Syst. Biol.* **11**, 827 (2015).
- Francine, P. Systems biology: new insight into antibiotic resistance. *Microorganisms* **10**, 2362 (2022).
- Nosil, P., Flaxman, S. M., Feder, J. L. & Gompert, Z. Increasing our ability to predict contemporary evolution. *Nat. Commun.* **11**, 5592 (2020).
- Shahrezaei, V. & Swain, P. S. Analytical distributions for stochastic gene expression. *Proc. Natl. Acad. Sci. USA* **105**, 17256–17261 (2008).
- Charlebois, D. A., Abdennur, N. & Kaern, M. Gene expression noise facilitates adaptation and drug resistance independently of mutation. *Phys. Rev. Lett.* **107**, 218101 (2011).

- Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
- Lieberman, T. D. et al. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nat. Genet.* **43**, 1275–1280 (2011).
- Hawkins, N. J. & Fraaije, B. A. Contrasting levels of genetic predictability in the evolution of resistance to major classes of fungicides. *Mol. Ecol.* **30**, 5318–5327 (2021).
- van Dijk, T. et al. Mutation supply and the repeatability of selection for antibiotic resistance. *Phys. Biol.* **14**, 055005 (2017).
- Comas, I. et al. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat. Genet.* **44**, 106–110 (2011).
- Palmer, A. C. & Kishony, R. Understanding, predicting and manipulating the genotypic evolution of antibiotic resistance. *Nat. Rev. Genet.* **14**, 243–248 (2013).
- Rodríguez-Verdugo, A., Gaut, B. S. & Tenaillon, O. Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress. *BMC Evol. Biol.* **13**, 50 (2013).
- Schenk, M. F. & de Visser, J. A. G. M. Predicting the evolution of antibiotic resistance. *BMC Biol.* **11**, 14 (2013).
- Poelwijk, F. J., Kiviet, D. J., Weinreich, D. M. & Tans, S. J. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445**, 383–386 (2007).
- Gould, S. J. *The Structure of Evolutionary Theory* (Harvard University Press, 2002).
- Charlebois, D. A. & Kaern, M. What all the noise is about: the physical basis of cellular individuality. *Can. J. Phys.* **90**, 919–923 (2012).
- Guthrie, J. & Charlebois, D. A. Non-genetic resistance facilitates survival while hindering the evolution of drug resistance due to intraspecific competition. *Phys. Biol.* **19**, 066002 (2022).
- Makela, J., Kandavalli, V. & Ribeiro, A. S. Rate-limiting steps in transcription dictate sensitivity to variability in cellular components. *Sci. Rep.* **7**, 10588 (2017).
- Mahrt, N. et al. Bottleneck size and selection level reproducibly impact evolution of antibiotic resistance. *Nat. Ecol. Evol.* **5**, 1233–1242 (2021).
- Farquhar, K. S., Rasouli Koohi, S. & Charlebois, D. A. Does transcriptional heterogeneity facilitate the development of genetic drug resistance? *BioEssays* **43**, e2100043 (2021).
- Brock, A., Chang, H. & Huang, S. Non-genetic heterogeneity—a mutation-independent driving force for the somatic evolution of tumours. *Nat. Rev. Genet.* **10**, 336–342 (2009).
- Torres, A., Cockerell, S., Phillips, M., Balazsi, G. & Ghosh, K. MaxCal can infer models from coupled stochastic trajectories of gene expression and cell division. *Biophys. J.* **122**, 1–13 (2023).
- Charlebois, D. A., Balazsi, G. & Kaern, M. Coherent feedforward transcriptional regulatory motifs enhance drug resistance. *Phys. Rev. E* **89**, 052708 (2014).
- Camellato, B., Roney, I. J., Azizi, A. & Charlebois, D. A. Engineered gene networks enable non-genetic drug resistance and enhanced cellular robustness. *Eng. Biol.* **3**, 72–79 (2019).
- Miton, C. M. & Tokuriki, N. How mutational epistasis impairs predictability in protein evolution and design. *Protein Sci.* **25**, 1260–1272 (2016).
- Weinreich, D. M., Delaney, N. F., Depristo, M. A. & Hartl, D. L. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312**, 111–114 (2006).
- Salverda, M. L. M. et al. Initial mutations direct alternative pathways of protein evolution. *PLoS Genet.* **7**, e1001321 (2011).
- Neher, R. A. Genetic draft, selective interference, and population genetics of rapid adaptation. *Annu. Rev. Ecol. Syst.* **44**, 195–215 (2013).
- Huang, S. The molecular and mathematical basis of Waddington's epigenetic landscape: a framework for post-Darwinian biology? *BioEssays* **34**, 149–157 (2012).
- Barroso-Batista, J. et al. The first steps of adaptation of *Escherichia coli* to the gut are dominated by soft sweeps. *PLoS Genet.* **10**, e1004182 (2014).
- Metzger, B. P. H., Park, Y., Starr, T. N. & Thornton, J. W. Epistasis facilitates functional evolution in an ancient transcription factor. *eLife* **12**, RP88737 (2023).
- Pinheiro, F., Warsi, O., Andersson, D. I. & Lässig, M. Metabolic fitness landscapes predict the evolution of antibiotic resistance. *Nat. Ecol. Evol.* **5**, 677–687 (2021).
- Rego-Costa, A., Debarre, F. & Chevin, L.-M. Predictability of evolution in a changing environment. *Evolution* **72**, 375–385 (2018).
- Kauffman, S. A. *A World Beyond Physics: The Emergence and Evolution of Life* (Oxford University Press, 2019).
- Hall, R. & Charlebois, D. A. Book review on "A world beyond physics: the emergence and evolution of life" by Stuart A. Kauffman. *Quart. Rev. Biol.* **95**, 133–134 (2020).
- Sasi, V. M. et al. Predicting antiviral resistance mutations in SARS-CoV-2 main protease with computational and experimental screening. *Biochemistry* **61**, 2495–2505 (2022).

43. Cincotta, P. M., Giordano, C. M. & Shevchenko, I. I. Revisiting the relation between the Lyapunov time and the instability time. *Phys. D Nonlinear Phenom.* **430**, 133101 (2022).
44. Gorter, F. A. et al. Genomics of adaptation depends on the rate of environmental change in experimental yeast populations. *Mol. Biol. Evol.* **34**, 2613–2626 (2017).
45. Tenaillon, O. et al. The molecular diversity of adaptive convergence. *Science* **335**, 457–461 (2012).
46. Barrick, J. E. et al. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1247 (2009).
47. Lang, G. I. & Desai, M. M. The spectrum of adaptive mutations in experimental evolution. *Genomics* **104**, 412–416 (2014).
48. Thomas, N. et al. Tracking global changes induced in the CD4 T-cell receptor repertoire by immunization with a complex antigen using short stretches of CDR3 protein sequence. *Bioinformatics* **30**, 3181–3188 (2014).
49. Parts, L. et al. Revealing the genetic structure of a trait by sequencing a population under selection. *Genome Res.* **21**, 1131–1138 (2011).
50. Vazquez-Garcia, I. et al. Clonal heterogeneity influences the fate of new adaptive mutations. *Cell Rep.* **21**, 732–744 (2017).
51. Shankarnarayan, S. A., Guthrie, J. D. & Charlebois, D. A. Machine learning for antimicrobial resistance research and drug development. in *The Global Antimicrobial Resistance Epidemic—Innovative Approaches and Cutting-Edge Solutions* (ed. Tellez, G.) 143–144 (IntechOpen, 2022).
52. Lang, G. I., Botstein, D. & Desai, M. M. Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics* **188**, 647–661 (2011).
53. Luksza, M. & Lassig, M. A predictive fitness model for influenza. *Nature* **507**, 57–61 (2014).
54. Woods, R., Schneider, D., Winkworth, C. L., Riley, M. A. & Lenski, R. E. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **103**, 9107 (2006).
55. Lamberte, L. E. & van Schaik, W. Antibiotic resistance in the commensal human gut microbiota. *Curr. Opin. Microbiol.* **68**, 102150 (2022).
56. Hato, H. et al. Factor associated with oral candidiasis caused by co-infection of *Candida albicans* and *Candida glabrata*: a retrospective study. *J. Dent. Sci.* **17**, 1458–1461 (2022).
57. Healey, D., Axelrod, K. & Gore, J. Negative frequency-dependent interactions can underlie phenotypic heterogeneity in a clonal microbial population. *Mol. Syst. Biol.* **12**, 877 (2016).
58. Imhof, M. & Schlotterer, C. Microcosms indicate a tight link between predictability of ecosystem dynamics and diversity. *PLoS Genet.* **2**, e103 (2006).
59. Celiker, H. & Gore, J. Clustering in community structure across replicate ecosystems following a long-term bacterial evolution experiment. *Nat. Commun.* **5**, 4643 (2014).
60. Chevin, L. M., Gompert, Z. & Nosil, P. Frequency dependence and the predictability of evolution in a changing environment. *Evol. Lett.* **6**, 21–33 (2022).
61. Hayati, M., Biller, P. & Colijn, C. Predicting the short-term success of human influenza virus variants with machine learning. *Proc. R. Soc. B Biol. Sci.* **287**, 20200319 (2020).
62. Goebel, R. et al. Explainable AI: the new 42? in *Machine Learning and Knowledge Extraction* (eds. Holzinger, A. et al.) CD-MAKE 2018. Lectures Notes in Computer Science, 11015 (Springer, Cham, 2018).
63. Kolmogorov, A. N. Three approaches to the quantitative definition of information. *Probl. Inf. Transm.* **1**, 1–7 (1965).
64. Sarai, A. Molecular recognition and information gain. *J. Theor. Biol.* **140**, 137–143 (1989).
65. Farquhar, K. S., Flohr, H. & Charlebois, D. A. Advancing antimicrobial resistance research through quantitative modeling and synthetic biology. *Front. Bioeng. Biotechnol.* **8**, 583415 (2020).
66. de Vos, M. G., Poelwijk, F. J. & Tans, S. J. Optimality in evolution: new insights from synthetic biology. *Curr. Opin. Biotechnol.* **24**, 797–802 (2013).
67. Kitano, H. Systems biology: a brief overview. *Science* **295**, 1662–1664 (2002).
68. Charlebois, D. A., Intosalmi, J., Fraser, D. & Kaern, M. An algorithm for the stochastic simulation of gene expression and heterogeneous population dynamics. *Commun. Comput. Phys.* **9**, 89–112 (2011).
69. Blundell, J. R. & Levy, S. F. Beyond genome sequencing: lineage tracking with barcodes to study the dynamics of evolution, infection, and cancer. *Genomics* **104**, 417–430 (2014).
70. Bleuven, C. et al. A collection of barcoded natural isolates of *Saccharomyces paradoxus* to study microbial evolutionary ecology. *MicrobiologyOpen* **8**, e773 (2019).
71. Pines, G., Fankhauser, R. G. & Eckert, C. A. Predicting drug resistance using deep mutational scanning. *Molecules* **25**, 2265 (2020).
72. Dewachter, L. et al. Deep mutational scanning of essential bacterial proteins can guide antibiotic development. *Nat. Commun.* **14**, 241 (2023).
73. Blanquart, F. Evolutionary epidemiology models to predict the dynamics of antibiotic resistance. *Evol. Appl.* **12**, 365–383 (2018).
74. Gray, R. M. *Entropy and Information Theory*. Second Edition. (Springer, 2011).
75. Baym, M., Stone, L. K. & Kishony, R. Multidrug evolutionary strategies to reverse antibiotic resistance. *Science* **351**, aad3292 (2016).
76. Papp, B., Notebaart, R. A. & Pal, C. Systems-biology approaches for predicting genomic evolution. *Nat. Rev. Genet.* **12**, 591–602 (2011).
77. Sakagianni, A. et al. Using machine learning to predict antimicrobial resistance—a literature review. *Antibiotics* **12**, 452 (2023).

## ACKNOWLEDGEMENTS

The author acknowledges Prof. Michael Manhart and Prof. Meike Wortel for insightful discussions that inspired many of the ideas presented in this manuscript. The author also thanks Sanjina R. Aurin, Teresa Charlebois, Harold Flohr, and the two anonymous reviewers for helpful feedback on the manuscript, and Harold Flohr for help with the figures. Funding was provided by the Government of Canada's New Frontiers in Research Fund—Exploration program (2019-01208) and a Natural Sciences and Engineering Research Council of Canada Discovery Grant (RGPIN-2020- 04007).

## COMPETING INTERESTS

The author declares no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41540-023-00304-6>.

**Correspondence** and requests for materials should be addressed to Daniel A. Charlebois.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023