Explaining the stable coexistence of drug-resistant and -susceptible pathogens: a simple model reveals the chaos beneath the calm

Pleuni S Pennings
San Francisco State University

Keywords: drug resistance; antibiotic resistance; coexistence; mutation-selection balance; HIV; E. coli.

Abstract

Drug resistance is a problem in many pathogens, including viruses, bacteria, fungi and parasites [1]. While overall, levels of resistance have risen in recent decades, there are many examples where after an initial rise, levels of resistance have stabilized [2–6]. The stable coexistence of resistance and susceptibility has proven hard to explain – in most evolutionary models, either resistance or susceptibility ultimately "wins" and takes over the population [2,3,7–9]. Here, we show that a simple stochastic model, mathematically akin to mutation-selection balance theory, can explain several key observations about drug resistance: (1) the stable coexistence of resistant and susceptible strains (2) at levels that depend on population-level drug usage and (3) with resistance often due to many different strains (resistance is present on many different genetic backgrounds). The model works for resistance due to both mutations or horizontal gene transfer (HGT). It predicts that new resistant strains should continuously appear (through mutation or HGT and positive selection within treated hosts) and disappear (due to the cost of resistance). The result is that while resistance is stable, which strains carry resistance is constantly changing. We used data on 37,000 *E. coli* isolates to test this prediction for a known resistance mutation and a resistance gene in the UK and found that the data are consistent with the prediction. Having a model that explains the dynamics of drug resistance will allow us to plan science-backed interventions to reduce the burden of drug resistance. Next, it will be of interest to test our model on data from different drug-pathogen combinations and to estimate model parameters for these drug–pathogen combinations in different geographic regions with varying levels of drug use. It will then be possible to predict the results of interventions, especially drug restriction policies and increase our understanding of whether and how fast such restrictions should result in reduced resistance [10,11].

Introduction

Drug resistance is common in viruses (HIV), bacteria (*S. aureus*, *E. coli, M. tuberculosis)*, parasites (*P. falciparum*) and fungi (*Candida* species), making it a major public health threat [1]. What can we do to prevent the rise of drug-resistant strains? Or to revert the trend if the level of resistance is already high? To plan successful science-backed interventions, we must understand what determines the level of drug resistance in a given pathogen. Therefore, a need exists for better models to explain current resistance levels, predict how resistance rates will change over time and identify which interventions would most reduce the burden of resistance. Here, we present a simple model that does all of these things.

Any model that might be used to predict drug resistance levels for different pathogens must be able to explain several observations that are currently not well explained [2,3,12]. First, while overall drug resistance has increased over the years, in many cases, an initial increase in resistance is followed by stable coexistence between drug-resistant and drug-sensitive strains. For an example, see **Figure 1A** on fluoroquinolone resistance in *E. coli* in Europe. Other examples of stable coexistence of resistant and susceptible strains can be found in *S. pneumoniae* [2,3], *S. aureus* [4] and HIV [5,6]). Second, a positive linear relationship has been observed between treatment intensity at the population level and drug resistance levels [13](**Figure 1B).** And third, many different resistant strains (i.e., resistance on different genetic backgrounds) often segregate at the same time in a pathogen population [5,14] (see **Figure 1C** for an example of resistance to the fluoroquinolone drug ciprofloxacin in *E. coli* in the UK; resistance appears repeatedly on the phylogenetic tree).

The second observation feels intuitive to most of us: If we use more drugs, we get more drug resistance. But the first observation, of stable coexistence, is puzzling: if there is selection *for* drug resistance, why doesn't the level of resistance keep increasing until it reaches 100%? The third observation, of resistance in many strains or genetic backgrounds, can be seen in many phylogenetic trees in articles about drug resistance, but usually does not receive a lot of attention. For example, in a large study of HIV isolates in Northern California, Rhee and colleagues found 82 phylogenetic clusters with one or more drug-resistance mutations [5]. In an *E. coli* dataset from the UK, 31 independent phylogenetic clusters with ciprofloxacin-resistant strains were found [15], and Casali and colleagues found 106 independent phylogenetic clusters with pyrazinamide resistance in *M. tuberculosis* [16]. These results show that resistance has evolved *de novo* or been acquired through horizontal gene transfer (HGT) many times [17,18].

Several different explanations have been put forward to explain the coexistence of resistance and susceptibility and the positive relationship between treatment and resistance levels [2,3]. These explanations have usually focused on mechanisms that require population structure or coinfection to create stabilizing selection [2,3,7–9]. We propose a simpler model akin to mutation-selection balance. Our model requires few assumptions, which means it can be applicable to a large variety of pathogens, and few parameters, which makes it easy to use.

We call the model "resistance acquisition and negative selection balance," and will show that it can explain all three key observations about drug resistance listed previously. Note that when resistance is due to HGT, we don't aim to explain why genetic elements that confer drug resistance exist or how they evolved, but rather why they are carried by a stable frequency of strains in pathogen populations.

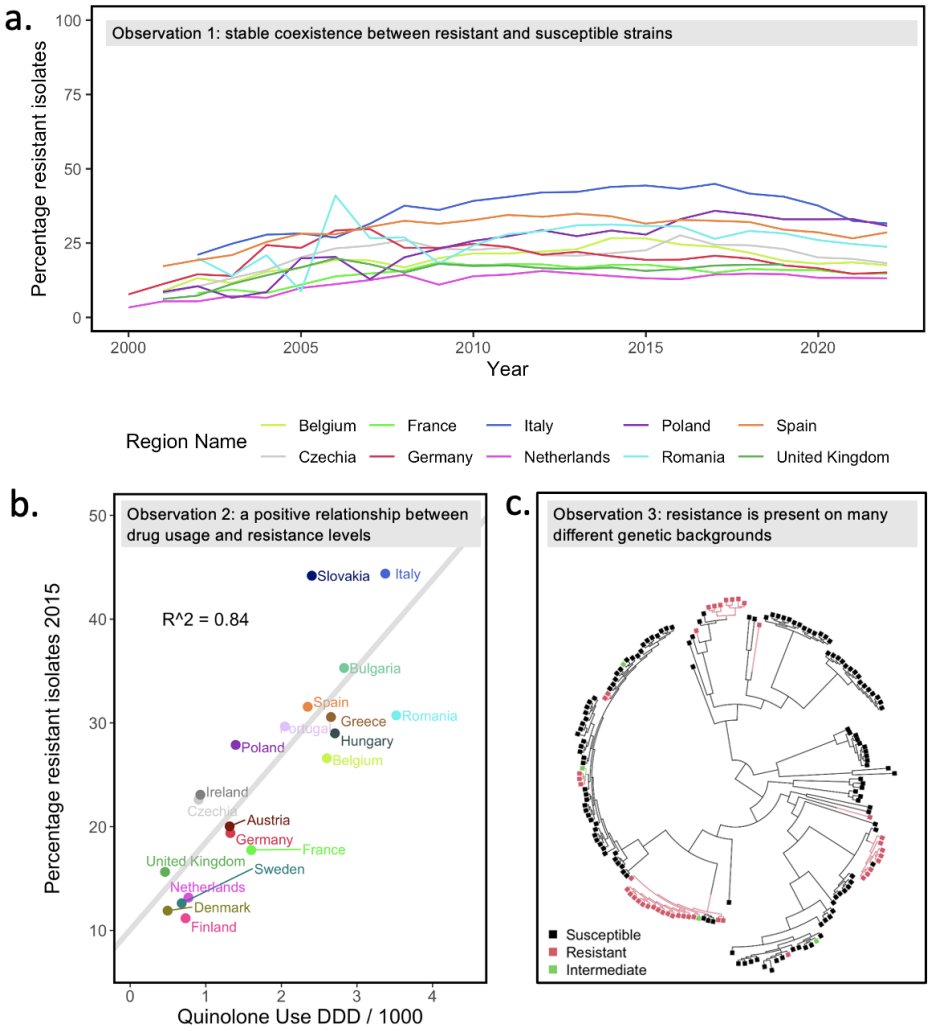

Figure 1. Three key observations about drug resistance, illustrated with data on resistance to quinolones in *E. coli*. *a.* *Observation 1. Typically, an initial increase in resistance is followed by stable coexistence between drug-resistant and drug-sensitive strains.* Each line depicts a *quinolone resistant E. coli strain over two decades in the 10 most populous countries (formerly) in the European Union (EU).* ***b.*** *Observation 2. Treatment intensity at the population level is correlated with drug resistance levels. A strong positive correlation exists between quinolone usage and quinolone resistance levels in 17 countries in the EU. Publicly available data from the Surveillance Atlas of Infectious Diseases were used to create Figures 1a and 1b.* ***c.*** *Observation 3. Many different resistant strains often segregate at the same time. Quinolone (specifically ciprofloxacine) resistance in bacteremia patients in the UK is due to many different origins of resistance (here phylogroup D and F are shown). Data from [13], figure reproduced from [14].*

The resistance acquisition – negative selection balance model

Our main goal is to determine whether a model based on the classic mutation-selection balance theory can explain three striking observations of drug resistance at the population level: (1) a stable coexistence of resistance and susceptibility, (2) a positive relationship between drug usage and resistance levels, and (3) the simultaneous existence of many drug-resistant strains. Here, we develop a mathematical model that

can explain the first two explanations for resistance due to *de novo* mutation or HGT, and then we run individual-based simulations that keep track of different resistant strains and explain the third observation. Finally, we analyze a large number of *E. coli* genomes over 15 years to visually compare predicted and observed patterns of resistance.

We base our model on classic mutation-selection balance theory because this model has already been used to explain a stable level of deleterious mutations and the existence of multiple alleles that carry a deleterious mutation [19,20]. Mutation-selection balance is a well-understood dynamic equilibrium, and we can borrow mathematical results from that theory directly here [21,22]. Mutation-selection balance has not received much attention as a potential explanation for drug resistance patterns because it is often believed to lead to resistance frequencies close to 0% (which is not usually observed in real life), and in the basic form, the model only considers negative selection (*against* resistance) and not positive selection (*for* resistance). Because it is clear that positive selection *for* drug resistance plays a role in the frequency of resistant strains (as evidenced by the second observation: more drugs lead to more resistance), we propose a version of mutation-selection balance that includes positive selection within treated hosts. It can explain stable resistant levels at any frequency.

In the basic mutation-selection balance model, mutation creates deleterious alleles at probability μ per individual per generation and these alleles reduce the fitness of the individual carrying them by selection coefficient s, see Table 1. Standard population genetic theory predicts that the deleterious allele frequency is μ/s [21]. Population genetic theory can also be used to predict the number of different mutations that segregate at any one time [19,20,23] and the time it takes for this equilibrium to be reached after the system is perturbed [24].

Table 1. Parameters for novel resistance acquisition-negative selection balance model vs classical mutation-selection equilibrium mode

Model parameter	**Mutation – selection balance**		**Resistance acquisition – negative selection balance**	
Generation of new variants	μ	mutation	$t{\cdot}e$	treatment and mutation or HGT plus positive selection due to the treatment
Selection against the variants	s	deleterious selection coefficient	c	cost of drug resistance in the general population
Equilibrium frequency of the variant	$f = \frac{\mu}{s}$		$f = \frac{t{\cdot}e}{c}$	
Model for individual-based simulations	One of the standard population genetic models (e.g., Wright-Fisher)		Standard epidemiological SIRS model with susceptible, infected and recovered hosts.	

To apply mutation-selection equilibrium theory to the problem of drug resistance, we have to make several adjustments and assumptions, but it is actually simple to recover the main mathematical result of the equilibrium level of resistance. First, we treat the population of the pathogen in an infected host as one pathogen individual. This assumption is valid if a strong transmission bottleneck occurs for each infection. Next, we assume that such "pathogen individuals" have a probability, $t{\cdot}e$, per time step, to be treated (t) and evolve to become resistant (e). The treatment parameter t is the probability that the host is

treated and the evolution parameter e is the probability that the within-host pathogen population evolves to become resistant through mutation or the acquisition of a genetic element, given that the host is treated. Mathematically, $t{\cdot}e$ plays the role of the mutation parameter μ in the standard model. Next, we assume that resistance comes with a mild cost, c, which can be understood as a reduction in R0 compared to the R0 of the susceptible pathogen strain. The cost parameter, c, determines how likely the resistant strain is to be transmitted to other hosts [25]. Mathematically, c plays the role of s in the standard model. We call the fraction of pathogen "individuals" that is resistant (equivalent to the fraction of resistant infections out of all infections): f. Only susceptible pathogen individuals ($1-f$ of the population) can evolve to become resistant. Under this model, the frequency of resistant strains, f, will increase at rate $t{\cdot}e*(1-f)$ through positive selection within hosts and be reduced by negative selection at rate $-c*f\,(1-f)$. Thus the change in frequency can be written as:

$$f' = (t \cdot e)(1 - f) - c \cdot f(1 - f).$$

An equilibrium will be reached when

$$(t \cdot e)(1 - f) = c \cdot f(1 - f),$$

which means that the predicted frequency of resistance at the host population level will be:

$f = (t \cdot e)/(c)$, see **Figure 2**. This result explains both the stability of the level of resistance and the linear relationship between treatment levels (through t) and the level of resistance.

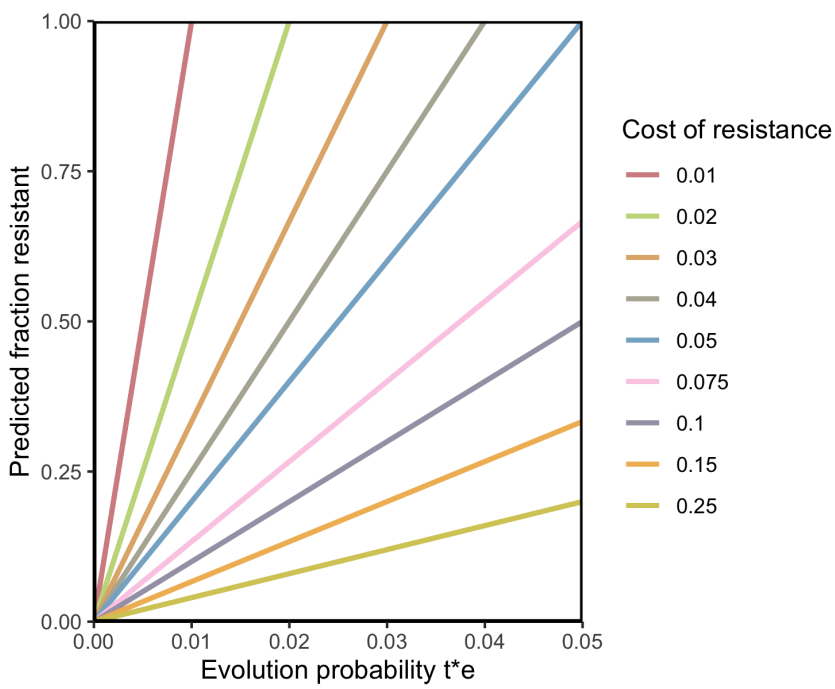

Figure 2: The level of resistance predicted by our Resistance Acquisition – Negative Selection Balance model for a population of pathogens is a function of the probability of resistance evolution per infected host, t*e, and the cost of resistance, c. *Higher levels of treatment lead to higher values of t*e, consistent with Key Observation 2 about drug resistance. Strains acquire resistance by mutation or through horizontal gene transfer. Resistant strains can be transmitted to other hosts, with an R0 that is reduced by (1-c) compared to the R0 of susceptible strains.*

Simulations in a SIRS model

To gain a better understanding of the resistance acquisition – negative selection balance model and to determine whether the results satisfied key observation #3, we ran individual-based simulations using a Susceptible, Infected, Recovered and Susceptible hosts (SIRS) model. A SIRS model assumes that recovered hosts are immune for a few time steps but then become susceptible again (hence the second S in SIRS) [26]. Initially, the population is seeded with only drug-susceptible pathogen strains and with a distribution of sequence types (STs) similar to what we find in the UK for *E. coli* [27]. After an initial period, treatment is started for randomly chosen infected hosts, which means that from there on, drug-susceptible strains have a probability $t \cdot e$ per time step to evolve to become drug-resistant due to mutation or acquisition of a genetic element. When a pathogen becomes resistant, it retains its ST identity. The cost of resistance is implemented as a reduction of the R0 value for drug-resistant strains relative to the R0 for the original susceptible strain, $R0_res = R0_sus*(1-c)$ [28]. For simplicity, we ignore back-mutation, co-infection and treatment while a host is not yet infected. We implemented the model as a fully stochastic, individual-based model in the R programming language.

Results

In the resistance acquisition – negative selection balance model, there is strong positive selection for drug resistance within hosts on treatment and mild negative selection against drug resistance in the majority of hosts that are not on treatment. The mathematical model can explain the first two observations laid out in **Figure 1**. First, drug resistance stably co-exists with susceptibility at the level of $t \cdot e/c$. Second, the level of resistance depends on the strength of positive selection and thus the usage of drugs at the population level, through the treatment parameter t. We then used individual-based simulations built on this model to confirm that the resistance acquisition – negative selection balance model can lead to many different resistant strains (i.e., resistance on different genetic backgrounds) in a population at any one point in time (see **Figure 3A**). The stable equilibrium that results from this model thus reflects the ongoing evolution of new resistant strains, of which some, due to chance, reach an appreciable level, but which ultimately all die out due to competition with the susceptible strain. This result leads to a clear prediction: Although the level of resistance is stable, "who" is resistant changes constantly. To determine whether this constant change is happening in reality we looked at temporal changes in the STs responsible for resistance in *E. coli* in the UK.

We analyzed a known resistance mutation (S83L in Gyrase A [29]), a known common resistance gene (Bla_{TEM-1B} [30]) in 37,000 *E. coli* genome assemblies from the UK gathered over 15 years and downloaded from Enterobase [27]. First, we noticed that for both of these resistance mechanisms, the level of resistance is quite stable (Key Observation 1; **Figure 3B-C**). Next, we examined the data at a more granular level in hopes of seeing evidence for the waxing and waning of resistant strains that had independently evolved (whether by mutation or by acquiring a gene). *E. coli* strains are traditionally assigned to different multi-locus STs. Because STs are typically stable, resistance in *different* STs must be due to independent mutation or HGT. In **Figures 3B and C** , we show the level of resistance in the population and which STs are represented among the resistant isolates in pie charts. As predicted by our model, at any one time, resistance is due to mutations or genes carried by many STs (i.e., many colors in each of the pie charts in **Figures 3B-C**; Key Observation 3). We also note that in both our simulations and

the empirical data, resistant strains appear and disappear dynamically. For example,Bla_{TEM-1B} was initially carried mostly by the 131 sequence type, but later by 245 (**3C**).

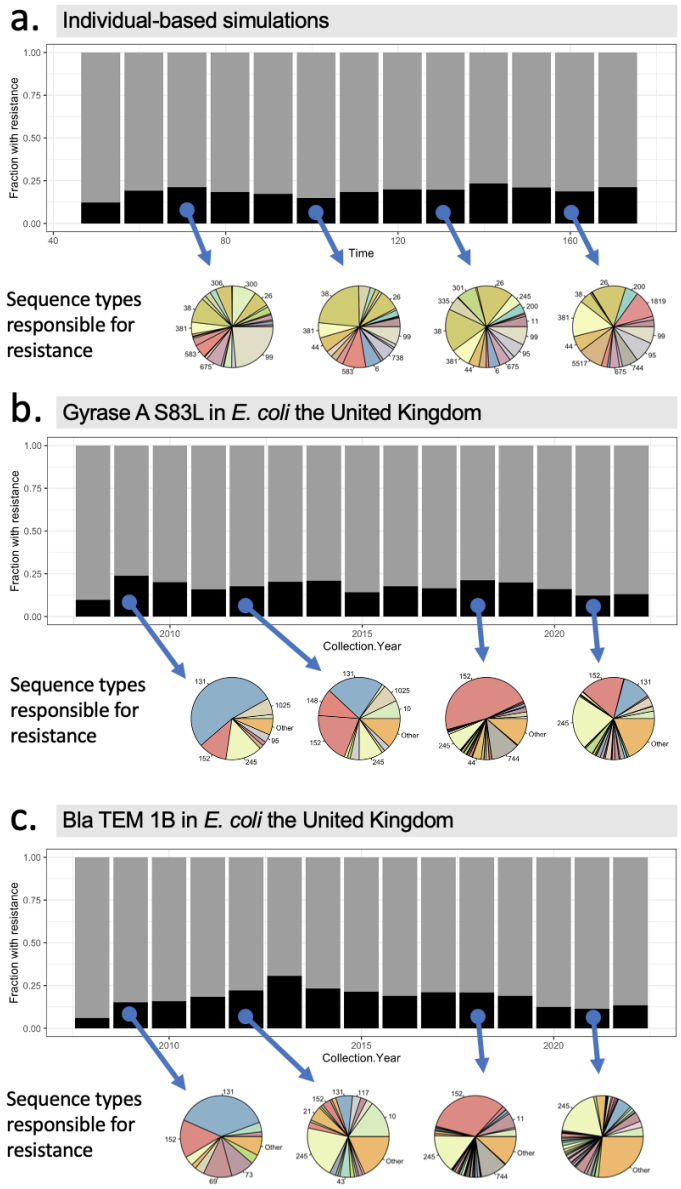

Figure 3: Dynamics of drug resistant strains in simulations and in *E. coli* in the UK.
a*: Simulations of the SIRS model with independently evolved drug resistant strains. Grey: isolates that don't carry resistance, black: isolates that carry resistance. Pie charts: for select time points, the sequence type diversity of isolates that carry resistance. Each color represents one of the 100 most common sequence types or all other sequence types (orange). Parameter values: c = 0.05, e = 0.01, host population size = 20,000, R0sus = 2. Figure* ***b****: As a, for the main ciprofloxacin resistance mutation (S83L in gyrase A) in 37,000 E coli samples from the UK. Figure* ***c****: as b, but for isolates that carry the* Bla_{TEM-1B} *gene. Data for 2B-C from Enterobase* [27].

Discussion

In order to understand, predict and influence levels of drug resistance in pathogens, we need mathematical models. Here we show that a simple stochastic model with strong positive selection for drug resistance within treated hosts and mild negative selection in the rest of the population can explain three key observations about drug resistance in a variety of pathogens (**Figure 1A**). First, the resistance acquisition – negative selection balance model we propose explains the stable coexistence of resistance and susceptibility in pathogen populations. Second, it explains the positive correlation between the amount of drugs used and the frequency of drug resistance (**Figure 2**). Finally, it recapitulates the simultaneous existence of many different strains that carry a resistance mutation or genetic element at any one time (**Figure 3A**). In addition to explaining these three key observations, our simulation results show that resistant strains increase and decrease in frequency, a behavior we observed in empirical data on ciprofloxacin resistance and two beta-lactamase genes in *E. coli* (**Figure 3B-C**).

A major strength of the resistance acquisition – negative selection balance model we propose is its simplicity. It relies on only a few assumptions and doesn't require consideration of co-infection or spatial structure of the host population. One benefit of this simplicity is that the model is more likely to be applicable to many different pathogen systems. The most important assumption we make is that resistance always comes with a cost for the pathogen, reducing its R0 – if that is not the case, resistant strains will outcompete susceptible strains [31]. A major difference between our model and some previous deterministic models to explain stable coexistence of drug resistance and susceptibility [3,4] is that ours is stochastic and includes the creation and dying out of resistant strains, even when genetic elements that confer resistance are stable. One reason why researchers have chosen deterministic models over stochastic ones is because it is assumed that drug-resistant strains are stable. However, stability at the phenotypic level doesn't necessarily mean stability at the genetic level. New strains may evolve resistance in the same way as previous strains, even acquiring genetic elements from previous strains. The result is that the genetic element or the mutation may be stably present, but "who" carries it may always be changing (**Figure 3**).

Another major strength of our model is that it makes clear and testable predictions. 1. It predicts that for a given drug, the higher levels of resistance in countries with high drug usage is due to the fact that resistance acquisition happens more often. We should therefore see more independently evolved resistant strains in countries with high levels of resistance. 2. The lifespan of resistant strains in a population should depend on the cost of the resistance for the pathogen. High-cost resistance should be associated with a shorter life span. 3. The time scale at which resistance levels change when treatment habits change should depend on the lifespans of resistant strains. Short life spans mean that a new equilibrium can be reached faster. These predictions can be used to determine for which pathogen systems the resistance acquisition – negative selection balance model works. Practically, the third prediction relating to how long it takes to reach a new equilibrium should be

Thanks to publicly available data, we were able to show that stability at the population level actually hides dynamic changes at lower levels (see **Figure 3B-C**). Although the resistance phenotype is present at a stable level (**Figure 1A**), and even the genetic elements that cause resistance are stable, the strains that carry the element are always changing (**Figure 3-C**). Although more detailed analysis is necessary, the *E. coli* data suggest that resistant strains may be around for several years before they disappear again. This observation is important because it indicates how long it would take for a new equilibrium to be established if treatment levels go up or go down. For example, if treatment with

fluoroquinolones was halted entirely, we would expect that no new resistant strains would evolve and based on the data analyzed in this article it would take a couple of years for the existing strains to disappear. This information can be used for planning and evaluating interventions [32].

If stable levels of resistance do indeed mask the ongoing appearance of new resistant strains, this will have implications for machine learning methods that are used to predict drug resistance phenotypes from genomic data. At any given year and location, a machine learning algorithm may learn to recognize certain STs responsible for much of the resistance burden in a species. Yet, over time, resistance may wane in these STs and increase in others, which means that we likely need to re-train such models for current and local situations [33].

Our study has several limitations. One limitation is that we modeled resistance to a single drug and did not consider interactions between drugs or resistances. For example, when plasmids carry multiple resistance genes, resistance to one drug can be acquired due to selection for resistance to another drug. Our model also doesn't explain why two seemingly unrelated resistances in some years occur on similar sequence types (**Figure 3B and 3C**). We also assumed that treatment does not affect transmission. If treatment itself is associated with strongly reduced transmission (as is the case in HIV), then this could in principle be modeled as a strong reduction of the R0 of resistant strains, but at least in the simulations it may be more realistic to include in the model the treatment state of hosts, where the transmission probability of hosts on treatment is significantly lower than those not on treatment [34,35]. The cost of resistance may also vary between strains and may be affected by compensatory mutations [31,36].

In conclusion, we propose a simple model that explains why drug resistance can stably co-exist with susceptible strains at levels that depend on drug usage. The model also explains why drug resistance is often present in many different strains (phylogenetic clusters, genetic backgrounds or sequence types) and predicts that resistant strains evolve and disappear again regularly. Data for ciprofloxacin resistance and beta-lactamase genes in *E. coli* are consistent with this prediction. This novel model has the potential to help us understand, predict and influence drug resistance levels and because of its simplicity, it can be readily tested in many different pathogen systems such as *S. pneumoniae* [2,3], *S. aureus* [4] and HIV [5,6]).

Acknowledgements

I am grateful for comments from and discussions with Kristin Harper, Joachim Hermisson, Sarah Cobey, Alison Feder, Nandita Garud, Rasmus Nielsen, Alison Hill, Scott Roy and the students in the SFSU CodeLab.

Bibliography

1. Murray CJL, Ikuta KS, Sharara F, Swetschinski L, Aguilar GR, Gray A, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet. 2022;399: 629–655. doi:10.1016/S0140-6736(21)02724-0
2. Krieger MS, Denison CE, Anderson TL, Nowak MA, Hill AL. Population structure across scales facilitates coexistence and spatial heterogeneity of antibiotic-resistant infections. PLOS Comput Biol. 2020;16: e1008010. doi:10.1371/journal.pcbi.1008010
3. Colijn C, Cohen T, Fraser C, Hanage W, Goldstein E, Givon-Lavi N, et al. What is the mechanism for persistent coexistence of drug-susceptible and drug-resistant strains of Streptococcus pneumoniae? J R Soc Interface. 2009;7: 905–919. doi:10.1098/rsif.2009.0400

4. Diekema DJ, Pfaller MA, Shortridge D, Zervos M, Jones RN. Twenty-Year Trends in Antimicrobial Susceptibilities Among Staphylococcus aureus From the SENTRY Antimicrobial Surveillance Program. Open Forum Infect Dis. 2019;6: S47–S53. doi:10.1093/ofid/ofy270
5. Rhee S-Y, Clutter D, Fessel WJ, Klein D, Slome S, Pinsky BA, et al. Trends in the Molecular Epidemiology and Genetic Mechanisms of Transmitted Human Immunodeficiency Virus Type 1 Drug Resistance in a Large US Clinic Population. Clin Infect Dis. 2019;68: 213–221. doi:10.1093/cid/ciy453
6. Rocheleau G, Brumme CJ, Shoveller J, Lima VD, Harrigan PR. Longitudinal trends of HIV drug resistance in a large Canadian cohort, 1996–2016. Clin Microbiol Infect. 2018;24: 185–191. doi:10.1016/j.cmi.2017.06.014
7. Cobey S, Baskerville EB, Colijn C, Hanage W, Fraser C, Lipsitch M. Host population structure and treatment frequency maintain balancing selection on drug resistance. J R Soc Interface. 2017;14: 20170295. doi:10.1098/rsif.2017.0295
8. Kouyos R, Klein E, Grenfell B. Hospital-Community Interactions Foster Coexistence between Methicillin-Resistant Strains of Staphylococcus aureus. PLOS Pathog. 2013;9: e1003134. doi:10.1371/journal.ppat.1003134
9. Blanquart F, Lehtinen S, Lipsitch M, Fraser C. The evolution of antibiotic resistance in a structured host population. J R Soc Interface. 2018;15: 20180040. doi:10.1098/rsif.2018.0040
10. Schuts EC, Boyd A, Muller AE, Mouton JW, Prins JM. The Effect of Antibiotic Restriction Programs on Prevalence of Antimicrobial Resistance: A Systematic Review and Meta-Analysis. Open Forum Infect Dis. 2021;8: ofab070. doi:10.1093/ofid/ofab070
11. Chatzopoulou M, Reynolds L. Role of antimicrobial restrictions in bacterial resistance control: a systematic literature review. J Hosp Infect. 2020;104: 125–136. doi:10.1016/j.jhin.2019.09.011
12. Davies NG, Flasche S, Jit M, Atkins KE. Within-host dynamics shape antibiotic resistance in commensal bacteria. Nat Ecol Evol. 2019;3: 440–449. doi:10.1038/s41559-018-0786-x
13. Goossens H, Ferech M, Stichele RV, Elseviers M. Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. The Lancet. 2005;365: 579–587. doi:10.1016/S0140-6736(05)17907-0
14. Nübel U, Roumagnac P, Feldkamp M, Song J-H, Ko KS, Huang Y-C, et al. Frequent emergence and limited geographic dispersal of methicillin-resistant Staphylococcus aureus. Proc Natl Acad Sci. 2008;105: 14130–14135. doi:10.1073/pnas.0804178105
15. van Nouhuijs F, Antony MA, Orcales F, Gozashti L, Roy SW, Ogbunugafor CB, et al. A phylogenetic approach to studying the roles of within-host evolution and between-host transmission of resistance for clinical *Escherichia coli* infections. Evolutionary Biology; 2022 Dec. doi:10.1101/2022.12.16.520823
16. Casali N, Nikolayevskyy V, Balabanova Y, Harris SR, Ignatyeva O, Kontsevaya I, et al. Evolution and transmission of drug-resistant tuberculosis in a Russian population. Nat Genet. 2014;46: 279–286. doi:10.1038/ng.2878
17. Nubel U, Roumagnac P, Feldkamp M, Song J-H, Ko KS, Huang Y-C, et al. Frequent emergence and limited geographic dispersal of methicillin-resistant Staphylococcus aureus. Proc Natl Acad Sci. 2008;105: 14130–14135. doi:10.1073/pnas.0804178105
18. Shin J, Choi M-J, Ko KS. Replicon sequence typing of IncF plasmids and the genetic environments of blaCTX-M-15 indicate multiple acquisitions of blaCTX-M-15 in Escherichia coli and Klebsiella pneumoniae isolates from South Korea. J Antimicrob Chemother. 2012;67: 1853–1857. doi:10.1093/jac/dks143
19. Wakeley J. Conditional Gene Genealogies under Strong Purifying Selection. Mol Biol Evol. 2008;25: 2615–2626. doi:10.1093/molbev/msn209
20. Hartl DL, Campbell RB. Allele multiplicity in simple Mendelian disorders. Am J Hum Genet. 1982;34: 866–873.
21. Baum DA, Futuyma DJ, Hoekstra HE, Lenski RE, Moore AJ, Peichel CL, et al. The Princeton Guide to Evolution. Princeton University Press; 2017.

22. Hartl DL, Clark AG. Principles of Population Genetics. Fourth Edition, Fourth Edition. Oxford, New York: Oxford University Press; 2006.
23. Pennings PS, Hermisson J. Soft Sweeps II—Molecular Population Genetics of Adaptation from Recurrent Mutation or Migration. Mol Biol Evol. 2006;23: 1076–1084. doi:10.1093/molbev/msj117
24. Pennings PS, Kryazhimskiy S, Wakeley J. Loss and Recovery of Genetic Diversity in Adapting Populations of HIV. PLOS Genet. 2014;10: e1004000. doi:10.1371/journal.pgen.1004000
25. Wagner BG, Garcia-Lerma JG, Blower S. Factors limiting the transmission of HIV mutations conferring drug resistance: fitness costs and genetic bottlenecks. Sci Rep. 2012;2: 320. doi:10.1038/srep00320
26. Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M. Absolute Humidity and the Seasonal Onset of Influenza in the Continental United States. PLOS Biol. 2010;8: e1000316. doi:10.1371/journal.pbio.1000316
27. Zhou Z, Alikhan N-F, Mohamed K, Fan Y, Group the AS, Achtman M, et al. The EnteroBase user's guide, with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and Escherichia core genomic diversity. Genome Res. 2020;30: 138–152. doi:10.1101/gr.251678.119
28. Andersson DI, Hughes D. Antibiotic resistance and its cost: is it possible to reverse resistance? Nat Rev Microbiol. 2010;8: 260–271. doi:10.1038/nrmicro2319
29. Everett MJ, Jin YF, Ricci V, Piddock LJ. Contributions of individual mechanisms to fluoroquinolone resistance in 36 Escherichia coli strains isolated from humans and animals. Antimicrob Agents Chemother. 1996;40: 2380–2386. doi:10.1128/aac.40.10.2380
30. Goussard S, Courvalin P. Updated Sequence Information for TEM β-Lactamase Genes. Antimicrob Agents Chemother. 1999;43: 367–370. doi:10.1128/aac.43.2.367
31. Cohen T, Murray M. Modeling epidemics of multidrug-resistant M. tuberculosis of heterogeneous fitness. Nat Med. 2004;10: 1117–1121. doi:10.1038/nm1110
32. Sarma JB, Marshall B, Cleeve V, Tate D, Oswald T, Woolfrey S. Effects of fluoroquinolone restriction (from 2007 to 2012) on resistance in Enterobacteriaceae: interrupted time-series analysis. J Hosp Infect. 2015;91: 68–73. doi:10.1016/j.jhin.2015.05.006
33. Moradigaravand D, Palm M, Farewell A, Mustonen V, Warringer J, Parts L. Prediction of antibiotic resistance in Escherichia coli from large-scale pan-genome data. PLOS Comput Biol. 2018;14: e1006258. doi:10.1371/journal.pcbi.1006258
34. Kühnert D, Kouyos R, Shirreff G, Pečerska J, Scherrer AU, Böni J, et al. Quantifying the fitness cost of HIV-1 drug resistance mutations through phylodynamics. PLOS Pathog. 2018;14: e1006895. doi:10.1371/journal.ppat.1006895
35. Stadler T, Kühnert D, Bonhoeffer S, Drummond AJ. Birth–death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). Proc Natl Acad Sci. 2013;110: 228–233. doi:10.1073/pnas.1207965110
36. Hinz A, Amado A, Kassen R, Bank C, Wong A. Unpredictability of the fitness effects of antimicrobial resistance mutations across environments in Escherichia coli. bioRxiv; 2023. p. 2023.10.25.563906. doi:10.1101/2023.10.25.563906
37. Kallonen T, Brodrick HJ, Harris SR, Corander J, Brown NM, Martin V, et al. Systematic longitudinal survey of invasive Escherichia coli in England demonstrates a stable population structure only transiently disturbed by the emergence of ST131. Genome Res. 2017. doi:10.1101/gr.216606.116
38. Bollback JP. SIMMAP: Stochastic character mapping of discrete traits on phylogenies. BMC Bioinformatics. 2006;7: 88. doi:10.1186/1471-2105-7-88
39. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things): *phytools: R package*. Methods Ecol Evol. 2012;3: 217–223. doi:10.1111/j.2041-210X.2011.00169.x

Methods

Code and data will be made available on github.

Figure 1A European surveillance data

Quinolone resistance over time in 10 European countries.
Data were downloaded from the European Surveillance Atlas for Infectious Diseases.
Link https://atlas.ecdc.europa.eu/public/index.aspx .
As health topic, we chose "Antimicrobial resistance", as subpopulation "*Escherichia coli*|" and "fluoroquinolones" and as indicator "percentage resistance". We then downloaded a csv file with data for all available years and regions. For Figure 1A we included the 10 most populous countries in the EU, including the United Kingdom: Germany, United Kingdom, France, Italy, Spain, Poland, Romania, Netherlands, Belgium, Czech Republic.

Figure 1B European surveillance data

For figure 1b, we used the percentage fluoroquinolone resistance in 2015 from the dataset described for Figure 1A, but now for the 20 most populous countries. This included in addition to the first 10: Sweden, Portugal, Greece, Hungary, Austria, Bulgaria, Denmark, Finland, Slovakia, Ireland.
For data on the use of fluoroquinolones in these countries, we downloaded data from https://www.ecdc.europa.eu/en/publications-data/antimicrobial-consumption-annual-epidemiological-report-2015 which included "Downloadable tables: Antimicrobial consumption - Annual Epidemiological Report for 2015".

Figure 1C *E. coli* phylogenetic tree Kallonen dataset

The phylogenetic tree in Figure 1C was reproduced from [15]. Data presented here include only E coli strains from phylogroup D and F. The genomes and phenotypes were described in [37]. Briefly, a core genome alignment was used to create a phylogenetic tree, then we simulated a single history of the resistance phenotype using the "simmap" function in Phytools [38,39], the simulated history is shown on the tree with resistant branches in red, susceptible branches in green and intermediate branches in green.

Figure 3A individual based simulations

Simulations are run with a custom made R script. A population of hosts is initiated with 20,000 hosts (PopSize). Hosts can be

- *uninfected and susceptible to infection ("0")*
- *infected with a susceptible pathogen strain ("1")*
- *infected with a resistant pathogen strain ("2")*
- *recovered and not susceptible to infection (temporary immune) ("-1")*

20% of the hosts are initially labeled as *infected with a susceptible pathogen strain* ("1") while the other 80% are labeled as *susceptible* ("0"). At the beginning of the simulation, the infected hosts are given a sequence type, randomly chosen from the 100 most common sequence types in the UK from the data described for Figures 3B and 3C. Each time step, three things happen:

1. All *recovered* ("-1") hosts have a chance of 20% to become *susceptible* ("0") again.
2. All *infected* ("1" and "2") hosts have a chance to infect *susceptible* ("0") hosts with their sequence type and resistance phenotype. For hosts that are infected with susceptible strains, the number of newly infected hosts is drawn from a poisson distribution with mean R0sus*(NumberSusceptible/PopSize), where R0sus is set to 2. For hosts that are infected with resistant strains, the R0 value is lower by the cost of resistance R0res = (1 - CostResistance)*R0sus. The cost of resistance is set to 0.05.
3. Starting from timestep 25, In each host that is newly *infected with a susceptible pathogen strain*, the pathogen can evolve to become resistant with probability RiskResistanceEvo, which is set to 0.01.

We keep track of the number of infected hosts, whether these infections are with susceptible and resistant pathogens and the sequence types of these infections.

Figure 3B

We started by going to https://enterobase.warwick.ac.uk/species/index/ecoli and downloading meta data for all E. coli from the UK. Importantly, we included the "Assembly Barcode (assembly stats)" and the "ST (Achtman 7-locus MLST)" for the multi-locus sequence type of the isolates through the experimental data tab. We then created a list of assembly barcodes (which end in _AS) and used it to download all genome assemblies to a local computer with a Python script.
Starting from the downloaded genome assemblies, we used Blast to identify the Gyrase A gene and determined the identity of the 83rd amino acid. We then calculated the fraction resistant as the number of isolates with L (7041 in total) divided by those carrying L or S (40445 in total) separated out for each year.

Figure 3C

Starting from the downloaded genome assemblies (as described for Figure 3B), we used the software Abricate (https://github.com/tseemann/abricate) to determine whether or not an isolate carried the Bla_{TEM-1B} gene or not. We then calculated the fraction resistant as the number of isolates with the Bla_{TEM-1B} gene (6741 in total) divided by those with or without the gene (38303 in total) separated out for each year.