medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Explaining the stable coexistence of drug-resistant and -susceptible pathogens: the Resistance Acquisition Purifying Selection model

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 (Murray et al. 2022). While overall, levels of resistance have risen in recent decades, there are many examples where after an initial rise, levels of resistance have stabilized (Krieger et al. 2020; Colijn et al. 2009; Diekema et al. 2019; Rhee et al. 2019; Rocheleau et al. 2018). The stable coexistence of resistance and susceptibility has proven hard to explain (Krieger et al. 2020; Colijn et al. 2009; Cobey et al. 2017; Kouyos, Klein, and Grenfell 2013; Blanquart et al. 2018). 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 fitness cost of resistance). The result is that while resistance is stable, *which* strains carry resistance is constantly changing. We used data from a longitudinal genomic study on *E. coli* in Norway to test this prediction for resistance to five different drugs and found that, consistent with the model, most resistant strains indeed disappear quickly after they appear in the dataset.

Significance statement

It is unclear why drug-resistant and drug-susceptible strains of pathogens can co-exist. Basic evolutionary models would predict that one or the other would have higher fitness and thus take over the population. Here we propose the Resistance Acquisition - Purifying Selection model that can explain how resistant strains appear (due to treatment) and disappear (due to a fitness cost of resistance). We show how *E. coli* data are mostly consistent with the model. A better understanding of the balance between resistant and susceptible strains can help design interventions to reduce resistance levels.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

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 each 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 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 quinolone 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 quinolone drug ciprofloxacin in *E. coli* in the UK; resistance appears repeatedly on the phylogenetic tree [15,16]).

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* [17]. These results show that resistance has evolved *de novo* or been acquired through horizontal gene transfer (HGT) many times [18,19].

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 could be applicable to a large variety of pathogens, and few parameters, which makes it easy to use.

We call our model the "resistance acquisition and purifying selection" (RAPS) model and will show that it can explain all three key observations about drug resistance listed previously. We will also describe which observations cannot directly be explained by the RAPS model. 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.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

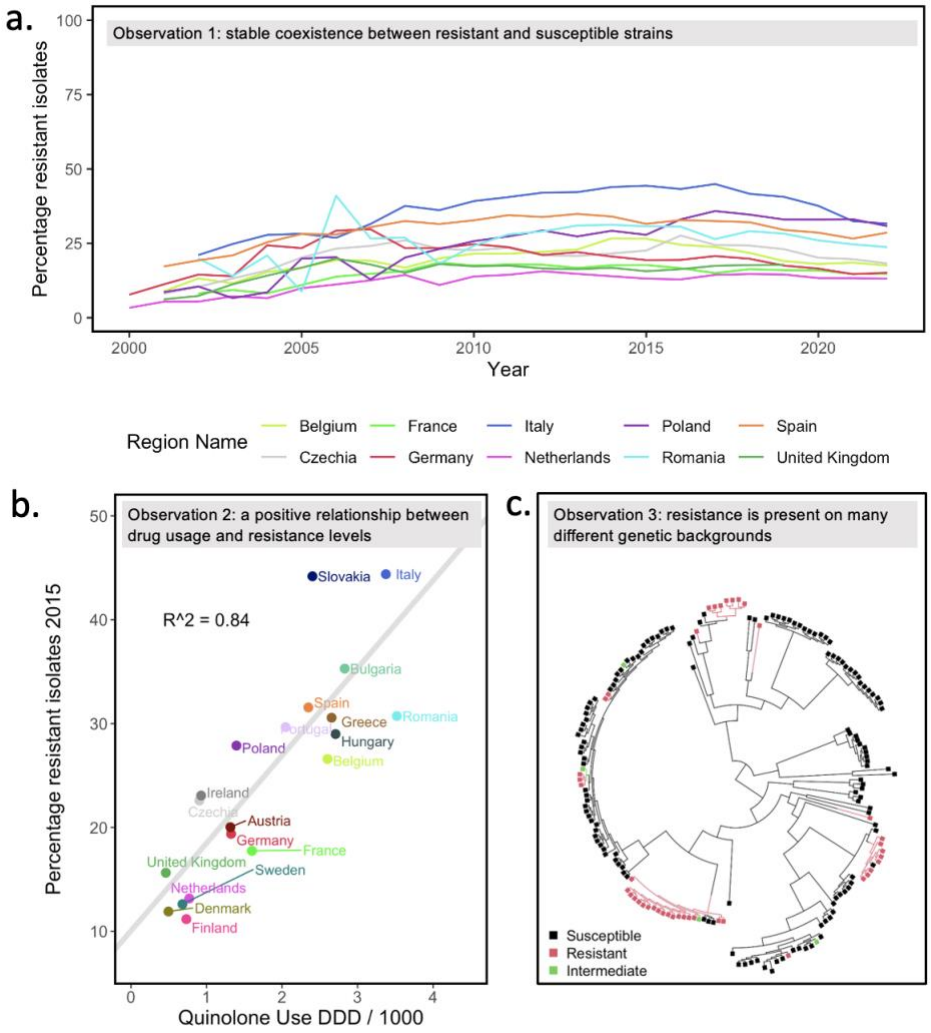


Figure 1. Three key observations about drug resistance, illustrated with data on resistance to quinolones in *E. coli*. *a.* *Observation 1. Typically, an initial increase in resistance is followed by stable coexistence between drug-resistant and drug-sensitive strains. Each line depicts the fraction of E. coli strains resistant to quinolones in a country over two decades in the 10 most populous countries (formerly) in the European Union (EU).* ***b.*** *Observation 2. Treatment intensity at the country level is correlated with drug resistance levels. A strong positive correlation exists between quinolone usage and quinolone resistance levels in 20 countries in the EU. Publicly available data from the Surveillance Atlas of Infectious Diseases were used to create Figures 1a and 1b* [20,21]. ***c.*** *Observation 3. Many different resistant strains often segregate at the same time. Quinolone (specifically ciprofloxacin) resistance in bacteremia patients in the UK is due to many different origins of resistance (here only phylogroup D and F are shown). Data from* [16], *figure reproduced from* [15].

The resistance acquisition – purifying selection (RAPS) model

Here we show that the RAPS model based on classic mutation-selection balance can explain the 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. Note that our goal here is not to determine whether the RAPS

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

model can better explain the data than the models proposed by others. Instead, we provide a proof of concept to show that the model can, in principle, lead to the main observations and should be considered as a possibility. The RAPS model works for resistance acquisition due to *de novo* mutation or HGT.

The RAPS model is based on the classic mutation-selection balance model, 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 [22,23]. Mutation-selection balance is a well-understood dynamic equilibrium, and we can borrow mathematical results from that theory directly here [24,25].

Mutation-selection balance has not received much attention as a potential explanation for drug resistance patterns. We see three main reasons for this omission: First, it is erroneously believed that mutation-selection balance can only lead to resistance frequencies close to 0%, whereas observed drug resistance frequencies are often much higher. Second, in the basic form, the model only considers negative selection (*against* mutations) and not positive selection (*for* mutations), whereas in the case of drug resistance, we know that positive selection plays an important role too. Third, a lot of focus in drug resistance research has been on well-established resistant strains. While well-established strains indeed require an explanation, it is important to realize that a lot of resistance is not due to these well-established strains.

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* per generation, see Table 1. Standard population genetic theory then predicts that the deleterious allele frequency is μ/s [24].

To apply mutation-selection equilibrium theory to the problem of drug resistance, we make several adjustments. First, as others have done, we treat the population of the pathogen as a metapopulation where each host is considered a "deme" [26]. Hosts can be infected (I), susceptible to infection (S) and recovered (R). Time is measured in units of duration of infection, so that R0 is the expected number of new infections (i.e., newly colonized demes) in a situation where all demes are susceptible [26] (alternatively, time could be measured per month or year, as we will do in the discussion). Next, we assume that demes have a probability, $t \cdot e$, per unit of time, to be treated (t) and to evolve to become resistant (e). The treatment parameter t is the probability (per unit time) 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 (see Table 1). Next, we assume that resistance comes with a mild fitness cost, c, which can be understood as a reduction in R0 compared to the R0 of the susceptible pathogen strain. The fitness cost parameter, c, determines how likely the resistant strain is to be transmitted to non-resistant hosts [27]. Mathematically, c plays the role of s (negative selection) in the standard mutation-selection model (see Table 1). We refer to the fraction of hosts that is infected with a resistant strain out of all infected hosts as f. Only hosts with susceptible strains ($1-f$ of the pathogen population) can evolve to become resistant.

Under this model, the frequency of resistant strains, f, will increase with $t \cdot e \cdot (1-f)$ per unit time through positive selection within hosts and will be reduced by negative selection with $-c \cdot f \cdot (1-f)$. Thus, the change in frequency (f') can be written as:

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

An equilibrium will thus be reached when

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

which means that the predicted equilibrium frequency of resistance will be:

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$f = (t \cdot e)/(c)$, see **Figure 2**. This equilibrium result can explain the stability of the level of resistance (*f*), the linear relationship between treatment levels (through *t*) and the level of resistance (*f*), while allowing for evolution of new resistant strains to occur.

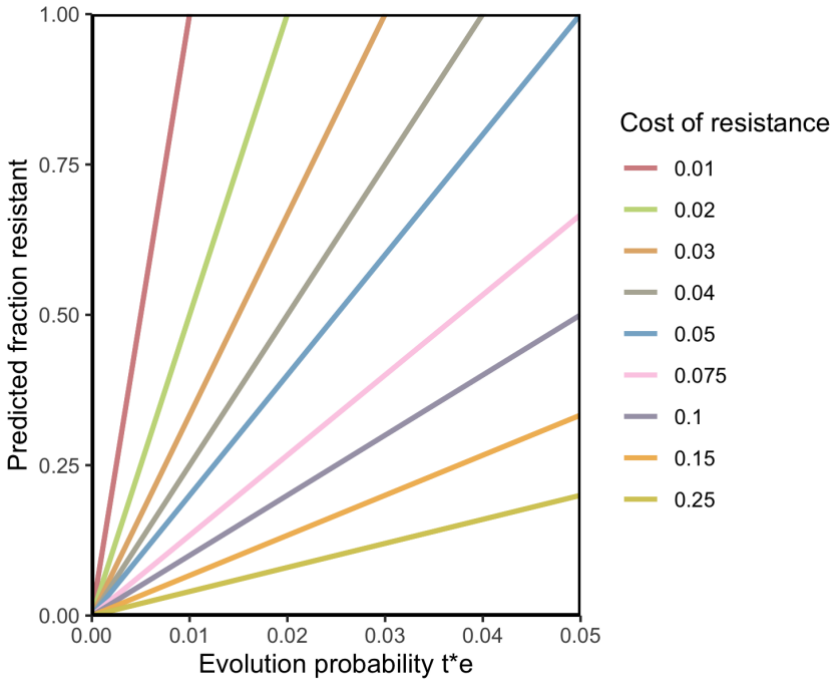


Figure 2: The level of resistance predicted by the Resistance Acquisition – Purifying Selection (RAPS) model for a population of pathogens is a function of the probability of resistance evolution per infection, $t \cdot e$, and the cost of resistance, *c*. *Higher levels of treatment lead to higher values of $t \cdot e$, consistent with Key Observation 2 about drug resistance. Strains acquire resistance by mutation or through HGT. Resistant strains can be transmitted to other hosts, with an R0 that is reduced by c compared to the R0 of susceptible strains.*

Table 1. Parameters for resistance acquisition-purifying selection (RAPS) model vs classical mutation-selection equilibrium model.

Model parameter	**Mutation – selection balance**		**Resistance acquisition – negative selection balance**	
Generation of new variants	μ	mutation per generation	$t \cdot e$	treatment and within-host evolution due to treatment per infection
Selection against the variants	*s*	deleterious selection coefficient per generation	*c*	cost of drug resistance in the general population per infection
Equilibrium frequency of the variant	$f = \frac{\mu}{s}$		$f = \frac{t \cdot e}{c}$	

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Genomic and phenotypic data

Using 16 years of publicly available data for resistance to several drugs in *E. coli* in Norway [28], we show that resistance acquisition is common and that most resistant strains are short-lived, which is consistent with the RAPS model. We used a dataset of *E. coli* genomes from bloodstream infections in Norway consisting of 3254 isolates from a surveillance program from 2002 - 2017. The fraction resistant strains in the dataset closely follows the reported fractions resistance from the European CDC surveillance atlas (see Figure 3A), which shows that even though the samples come from bloodstream infections, they are capturing important *E. coli* resistance dynamics in Norway.

Using the phylogenetic tree and the phenotypic resistance data provided by the authors of the original paper [28], we estimated that ampicillin resistance evolved over 400 times on the tree, and resistance to the other drugs (cefotaxime, ciprofloxacin, gentamicin and piperacillin-tazobactam) evolved between 43 and 93 times (Figure 3B) [see Methods]. Note that the big difference in the number of evolutionary origins between ampicillin and the other drugs is consistent with the fact that ampicillin is an antibiotic that is much more often prescribed than the others [20].

We found that for the five drugs we looked at, at least 75% of the strains were observed in only one year in the dataset (Figure 3C). The lifespan of individual resistant strains was estimated by determining, for each origin of resistance, in which years the resulting resistant strain was observed in the dataset. Next, we created an area plot that shows how most ciprofloxacin (quinolone) resistant strains are only seen for a short time span, but that also shows a few longer-lived strains and one resistant strain that is long-lived and reaches a high frequency (3D). While the many strains that appear and disappear again are consistent with the RAPS model, the model cannot explain the behavior of the strain that increases in frequency over several years. In fact, it is unlikely that this increase is driven by the resistance to ciprofloxacin, instead, as others have suggested, it could be that by chance the ciprofloxacin resistance mutations landed on a background with positively selected elements and are hitchhiking to high frequency [16].

In summary, our analysis of the Norwegian dataset shows that resistant strains to five different drugs evolve often, yet most resistant strains do not stay in the population for long, which is consistent with ongoing selection against resistance and with the RAPS model. Yet, there are a few long-lived resistance strains which are unexplained by the RAPS model which account for a significant portion of resistance in *E. coli*.

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

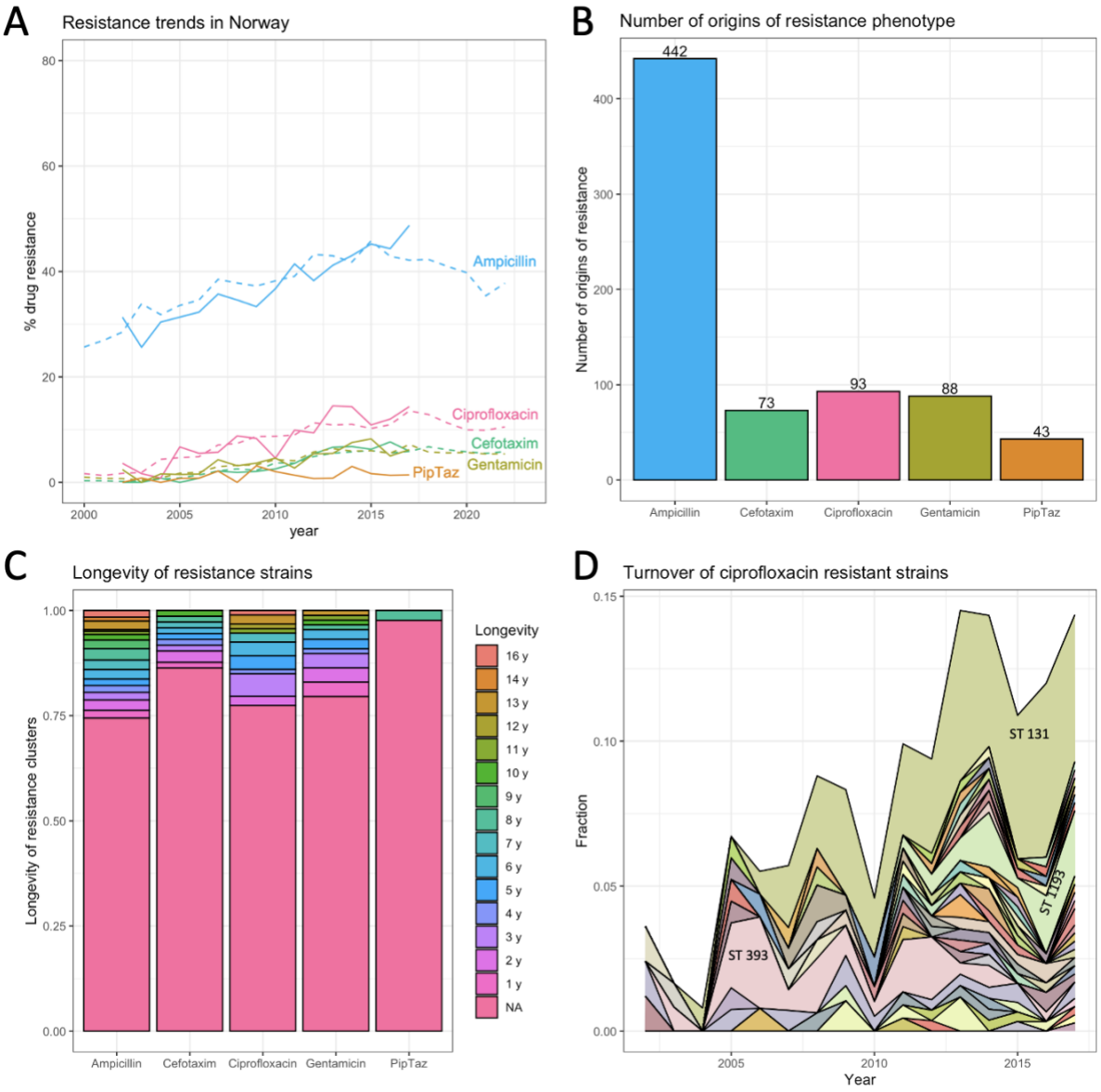


Figure 3: Dynamics of drug resistant strains in Norway. ***a.*** *Resistance trends in Norway for five drugs based on [20,28].* ***b.*** *Estimated number of origins of resistance to five drugs.* ***c.*** *Lifespans of resistant strains.* ***d.*** *Turnover of ciprofloxacin (quinolone) resistant strains*

Discussion

In order to understand, predict and influence levels of drug resistance in pathogens, we need mathematical models. Here we show that a simple model with strong positive selection for drug resistance within treated hosts and mild negative (purifying) selection in the rest of the population can explain three key observations about drug resistance in a variety of pathogens (**Figure 1**). First, the resistance acquisition – purifying selection (RAPS) model explains the stable coexistence of resistance and susceptibility in pathogen populations. Second, it explains the positive correlation between the amount of drug used and the frequency

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

of drug resistance (**Figure 2**). And finally, it recapitulates the simultaneous existence of many different strains that carry a resistance mutation or genetic element at any one time (**Figure 1C**).

A major strength of the RAPS model 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 different pathogen systems. The most important assumption we make is that resistance comes with a fitness cost for the pathogen, reducing its R0 – if that is not the case, resistant strains will outcompete susceptible strains [29]. A major difference between our model and some previous models that attempt to explain stable coexistence of drug resistance and susceptibility [2,3] is that the RAPS model includes the creation and dying out of resistant strains. This can happen even when the specific genetic elements that confer resistance are stably present in the population. One reason why researchers have often chosen to ignore the creation and dying out of resistant strains 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 3D**).

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 and, conversely, when resistance strains are generally short-lived this suggests that costs of resistance are high. 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, and interventions should pay off more quickly. Intervention studies where ciprofloxacin (quinolone) use was reduced seem to indicate that this influences resistance levels in a few months [30,31].

Thanks to publicly available data, we were able to show that stability at the population level hides dynamic changes at lower levels (see **Figure 3A-D**). Although the resistance phenotype is present at a stable level (**Figure 3A**), resistance is due to many evolutionary origins (**Figure 3B**) and 75% of resistant strains are only present for a year (**Figure 3C-D**). It will be of interest to see if more granular data (e.g., monthly surveillance from a hospital) would show the waxing and waning of resistant strains. A major open question is why there are some long-lived resistance strains while most are so short-lived. Are these long-lived strains hitchhiking with other genetic elements as others have suggested [16]? Do these strains occasionally lose the resistant phenotype, as would be expected if the resistance is a costly hitchhiker?

It is of interest to see if what we know about treatment levels and drug resistance evolution probabilities roughly fits with the model. We are aware of estimates for f (fraction resistant), t (treatment probability per year) and e (probability of evolution given treatment) for quinolone drugs. We can plug in those numbers in the equilibrium equation to get an (admittedly very rough) estimate for the cost of resistance per year. In 2015, the fraction of *E. coli* resistant to quinolones $\boldsymbol{f \approx 10\%}$ in Norway. We also know, for Norway, that the use of quinolones is around 0.5 doses per day per 1000 people [21] . If we assume that a course of quinolones is 5 days, this translates into $\boldsymbol{t \approx 3.3\%}$ probability per year of someone taking a course of quinolones ((0.5 DDD*365 days) / (5 days per course*1000 people)*100%). A recent study estimated that the probability of resistance evolution due to a course of quinolones is $\boldsymbol{e \approx 10\%}$ [32]. Because our model proposes that $f = (t \cdot e)/(c)$, we have $0.1 = 0.033 * 0.1 / c$, which means that the fitness cost should be around 0.033 or 3.3% per year. While we cannot currently know if this number is

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

correct, it is encouraging to see that the number is at least of an order of magnitude that seems plausible. It would be of interest to see whether phylogenetic approaches can be used to determine if resistant strains indeed have a lower R0 compared to non-resistant strains [33,34].

If stable levels of resistance mask the ongoing appearance of new resistant strains, this may 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 strains responsible for much of the resistance burden in a species. Yet, over time, these resistant strains may be replaced by others, which means that we likely need to re-train such models for current and local situations [35].

Our study has several limitations. One limitation is that the model considers only resistance to a single drug and not 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. 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 it would be useful if the model included the treatment state of hosts, where the transmission probability of hosts on treatment is significantly lower than those not on treatment [33,36]. The cost of resistance may also vary between strains and may be affected by compensatory mutations [29,37].

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 resistance to several drugs in *E. coli* in Norway are largely consistent with these predictions, though the model does not explain why some resistant strains are seen over many years. 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, Oana Carja, Rasmus Nielsen, Alison Hill, Scott Roy, Brandon Ogbunu, Ruth Hershberg 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

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

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. 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
17. 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
18. 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
19. 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
20. ECDC. Surveillance Atlas of Infectious Diseases. 28 Apr 2023 [cited 29 May 2024]. Available: https://www.ecdc.europa.eu/en/surveillance-atlas-infectious-diseases
21. ECDC. Antimicrobial consumption - Annual Epidemiological Report for 2015. 12 Jun 2018 [cited 29 May 2024]. Available: https://www.ecdc.europa.eu/en/publications-data/antimicrobial-consumption-annual-epidemiological-report-2015

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

22. Wakeley J. Conditional Gene Genealogies under Strong Purifying Selection. Mol Biol Evol. 2008;25: 2615–2626. doi:10.1093/molbev/msn209
23. Hartl DL, Campbell RB. Allele multiplicity in simple Mendelian disorders. Am J Hum Genet. 1982;34: 866–873.
24. Baum DA, Futuyma DJ, Hoekstra HE, Lenski RE, Moore AJ, Peichel CL, et al. The Princeton Guide to Evolution. Princeton University Press; 2017.
25. Hartl DL, Clark AG. Principles of Population Genetics. Fourth Edition, Fourth Edition. Oxford, New York: Oxford University Press; 2006.
26. Gordo I, Gomes MGM, Reis DG, Campos PRA. Genetic Diversity in the SIR Model of Pathogen Evolution. PLoS ONE. 2009;4: e4876. doi:10.1371/journal.pone.0004876
27. 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
28. Gladstone RA, McNally A, Pöntinen AK, Tonkin-Hill G, Lees JA, Skytén K, et al. Emergence and dissemination of antimicrobial resistance in Escherichia coli causing bloodstream infections in Norway in 2002–17: a nationwide, longitudinal, microbial population genomic study. Lancet Microbe. 2021;2: e331–e341. doi:10.1016/S2666-5247(21)00031-8
29. Cohen T, Murray M. Modeling epidemics of multidrug-resistant M. tuberculosis of heterogeneous fitness. Nat Med. 2004;10: 1117–1121. doi:10.1038/nm1110
30. 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
31. Gottesman BS, Carmeli Y, Shitrit P, Chowers M. Impact of Quinolone Restriction on Resistance Patterns of Escherichia coli Isolated from Urine by Culture in a Community Setting. Clin Infect Dis. 2009;49: 869–875. doi:10.1086/605530
32. Yaffe E, Dethlefsen L, Patankar AV, Gui C, Holmes S, Relman DA. A short course of antibiotics selects for persistent resistance in the human gut. bioRxiv; 2023. p. 2023.09.04.556257. doi:10.1101/2023.09.04.556257
33. 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
34. Helekal D, Keeling M, Grad YH, Didelot X. Estimating the fitness cost and benefit of antimicrobial resistance from pathogen genomic data. bioRxiv; 2022. p. 2022.12.02.518824. doi:10.1101/2022.12.02.518824
35. 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
36. 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
37. 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
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
40. Huelsenbeck JP, Nielsen R, Bollback JP. Stochastic mapping of morphological characters. Syst Biol. 2003;52: 131–158. doi:10.1080/10635150390192780

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Methods

All code and data are available on github https://github.com/pleunipennings/CoexistencePaper.

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 (named ECDC_surveillance_data_Antimicrobial_resistance_complete_DownloadApril2024.csv in the Github repository). 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 [20].

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" [21].

Figure 1C *E. coli* phylogenetic tree Kallonen dataset from Nouhuijs et al.

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 [16]. 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 Resistance trends in Norway for five drugs

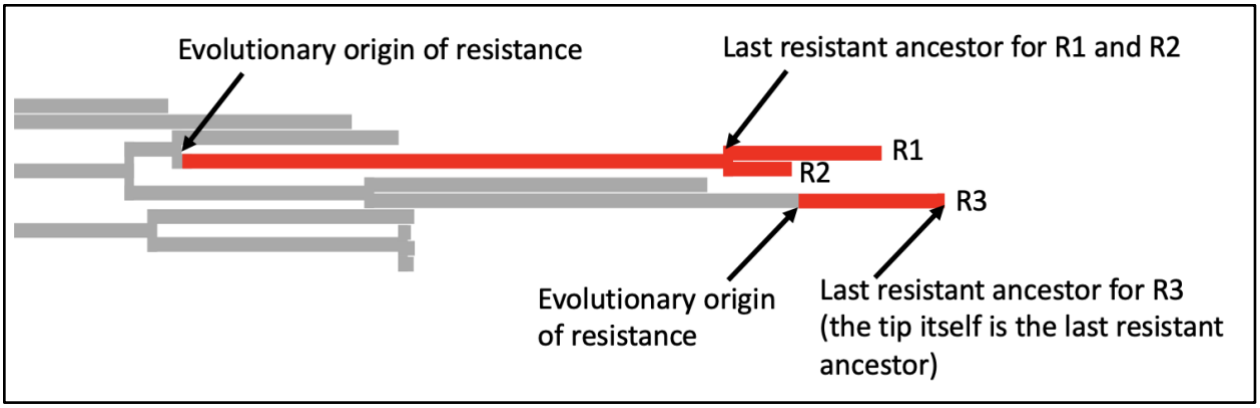
Data are from Gladstone, 2021 [28] and from the ECDC Surveillance Atlas [20].

Figure 3B Estimated number of origins of resistance to five drugs

We used the tree and metadata provided by Gladstone, 2021 [28]. This dataset consists of 16 years of surveillance data including 3254 *E. coli* samples from bacteremia patients in Norway between 2002 and 2017. Intermediate and susceptible strains were lumped in one category. We simulated a single history of the resistance phenotype using the "simmap" stochastic character mapping function in Phytools [39] which is based on methods by [38,40]. Next, we used a custom R script to identify the last resistant ancestor for

medRxiv preprint doi: https://doi.org/10.1101/2023.12.07.23299709; this version posted June 2, 2024. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

each resistant tip of the tree. If two tips share a last resistant ancestor, we assume that they are resistant due to the same evolutionary origin of resistance. This evolutionary origin may be due to a mutation, several mutations or horizontal gene transfer – our approach is based on the phenotypes, not the genotypes. Based on the simmap results, we count the number of times resistance was acquired on the tree.



Supplementary figure 1. In this example, there are three resistant tips. The branches are colored according to the simmap stochastic character mapping. R1 and R2 share a last resistant ancestor and are therefore assumed to be resistant due to the same evolutionary event in their history, we consider them part of the same resistant cluster or resistant strain. For R3, the last resistant ancestor is itself. It is the only member of the resistant cluster or strain.

Figure 3C Lifespans of resistant strains

Using the results from analysis described for 3B and the metadata from [28] to determine for each inferred resistance cluster for how many years it was observed in the data. For example, if R1 was a sample from 2005 and R2 was a sample from 2007, then we infer that the lifespan of this cluster was 3 years.

Figure 3D Turnover of ciprofloxacin (quinolone) resistant strains

Using the results from analysis described for 3B and the metadata from [28] to determine for each inferred ciprofloxacin resistance cluster what its frequency was in each year.