Collateral sensitivity is contingent on the repeatability of evolution

Daniel Nichol\textsuperscript{1,2*}, Joseph Rutter\textsuperscript{3}, Christopher Bryant\textsuperscript{4}, Peter Jeavons\textsuperscript{1}, Alexander RA Anderson\textsuperscript{5}, Robert A Bonomo\textsuperscript{3,4,6,7}, Jacob G Scott\textsuperscript{7,8,9*}

1 Department of Computer Science, University of Oxford, Oxford, UK
2 Centre for Evolution and Cancer, The Institute of Cancer Research, London, UK
3 Research Service, Louis Stokes Department of Veterans Affairs Hospital, Cleveland, OH, USA
4 Department of Medicine, Case Western Reserve University School of Medicine, Cleveland, OH, USA
5 Department of Integrated Mathematical Oncology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, USA
6 Departments of Biochemistry, Microbiology, Molecular Biology and Pharmacology, Case Western Reserve University School of Medicine, Cleveland, OH, USA
7 Center for Proteomics and Bioinformatics, Case Western Reserve University School of Medicine, Cleveland, OH, USA
8 Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford, Oxford, UK
9 Departments of Translational Hematology and Oncology Research and Radiation Oncology, Cleveland Clinic, Cleveland, OH, USA

* daniel.nichol@icr.ac.uk (DN); scottj10@ccf.org (JGS)
Antibiotic resistance represents a growing health crisis that necessitates the immediate discovery of novel treatment strategies. One such strategy is the identification of sequences of drugs exhibiting collateral sensitivity, wherein the evolution of resistance to a first drug renders a population more susceptible to a second. Here, we demonstrate that sequential multi–drug therapies derived from in vitro evolution experiments can have overstated therapeutic benefit – potentially suggesting a collaterally sensitive response where cross resistance ultimately occurs. The evolution of drug resistance need not be genetically or phenotypically convergent, and where resistance arises through divergent mechanisms, the efficacy of a second drug can vary substantially. We first quantify the likelihood of this occurring by use of a mathematical model parametrised by a set of small combinatorially complete fitness landscapes for Escherichia coli. We then verify, through in vitro experimental evolution, that a second–line drug can indeed stochastically exhibit either increased susceptibility or increased resistance when following a first. Genetic divergence is confirmed as the driver of this differential response through targeted sequencing. These results indicate that the present methodology of designing drug regimens through experimental collateral sensitivity analysis may be flawed under certain ecological conditions. Further, these results suggest the need for a more rigorous probabilistic understanding of the contingencies that can arise during the evolution of drug resistance.

The emergence of drug resistance is governed by Darwinian dynamics, wherein resistant mutants arise stochastically in a population and expand under the selective pressure of therapy [23]. These evolutionary principles underpin resistance to the presently most effective therapies for bacterial infections [4], cancers [8], viral infections [2] and disparate problems such as the management of invasive species and agricultural pests [14]. Biological mechanisms of drug resistance often carry a fitness cost in the absence of the drug and further, different resistance mechanisms can interact with one another to produce non–additive fitness effects, a phenomenon known as epistasis [20]. These trade–offs can induce rugged fitness landscapes, potentially restricting the number of accessible evolutionary trajectories to high fitness [21, 25] or rendering evolution irreversible [24].

Identifying evolutionary trade-offs forms the basis of an emerging strategy for combating drug resistance; prescribing sequences of drugs wherein the evolution of resistance to the first induces susceptibility to the next [10, 12, 17]. Where this occurs, the first drug is said to induce collateral sensitivity in the second. Conversely, where the first drug induces increased resistance in the second, collateral (or cross) resistance has occurred. Recently, in vitro evolution experiments have been performed, in both bacteria [5, 10, 16] and cancers [7, 29], to identify drug pairs or
sequences exhibiting collateral sensitivity. These experiments proceed by culturing a population in increasing concentrations of a drug to induce resistance and then assaying the susceptibility of the resultant population to a panel of potential second-line therapies. From these experiments, sequences or cycles of drugs in which each induces collateral sensitivity in the next have been suggested as potential therapeutic strategies to extend the therapeutic efficacy of a limited pool of drugs [7, 10]. For some cancer therapies, which often have severe side-effects and high toxicity, such sequential therapies may be the only way combine the use of multiple drugs.

We argue that collaterally sensitive drug pairs identified from a small number of in vitro evolutionary replicates likely do not always induce collateral sensitivity. This hypothesis arises from the observation that evolution is not necessarily repeatable; resistance to a drug can arise through multiple different mechanisms, as has been observed in cancers [28] and bacteria [1]. An a priori reason to assume that these different mechanisms will have correlated fitness effects under a second drug is not evident – just like the grade school lesson of convergent evolution: bats and birds can both fly, but their predators often differ. Indeed, one mutation may confer resistance to a second drug, whilst another may induce increased susceptibility (in comparison to the susceptibility of the wild-type), as was recently demonstrated in a drug screen of over 3000 strains of Staphylococcus aureus [11]. The potential impact of such divergent evolution can be conceptualised in the classical fitness landscape model of Wright [26], wherein genotypes are projected onto the two dimensional $x - y$ plane and fitness represented as the height above this plane. Evolution can be viewed as a stochastic ‘up–hill’ walk in this landscape wherein divergence can occur at a saddle. Figure 1 shows such a schematic fitness landscape annotated to demonstrate the capacity for divergent evolution and the potential effects on collateral sensitivity.

Previous studies have attempted to empirically determine the structure of the fitness landscape for a number of organisms and under different drugs [6]. In these studies, a small number of mutations associated with resistance are first identified. Strains are engineered corresponding to all possible combinations of presence and absence of these mutations and the fitness of each strain is measured by a proxy value, for example minimum inhibitory concentration (MIC) of a drug or average growth rate under a specific dose. These measurements are combined with the known genotypes to form a fitness landscape. However, to derive fitness landscapes through this method, the number of strains that must be engineered grows exponentially with the number of mutations of interest. Thus only small, combinatorially complete, portions of the true fitness landscape can be measured, for example consisting of 2-5 alleles [6] [19] [25]. Nevertheless, these restricted fitness landscapes can provide valuable insight into the evolution of drug resistance.

Mira et al. [15] derived fitness landscapes for E. coli with all combinations of four fitness conferring mutations (M69L, E104K, G238S and N276D) in the TEM gene and measured fitness under 15 different $\beta$-lactam antibiotics (See Supplementary Table 1), using the average growth
Figure 1. Evolutionary saddle points can drive divergent collateral response. A) A schematic fitness landscape model in which divergent evolution can occur. Following Wright [26], the \( x - y \) plane represents the genotypes and the height of the landscape above this plane represents fitness. Two evolutionary trajectories, both starting from a wild-type genotype (yellow circle), are shown. These trajectories diverge at an evolutionary saddle point (blue triangle) and terminate at distinct local optima of fitness (purple pentagon, green star). As the saddle point exists, evolutionary trajectories need not be repeatable. B) Schematic landscapes for a potential follow-up drug are shown, the collateral response can be (i) always cross-resistant, (ii) always collaterally sensitive or (iii) dependent on the evolutionary trajectory that occurs stochastically under the first drug. C) A potential evolutionary branching point in the TEM gene of \( E. coli \) identified in the fitness landscape for cefotaxime derived by Mira et al. [15].

rate (over 12 replicates) as a proxy of fitness. Of these 15 landscapes, 14 were identified as having multiple local optima of fitness, indicating the potential for the divergence of evolutionary trajectories. We utilised these landscapes, coupled with a previously published mathematical model [17] (see Methods), to estimate the likelihood of the different evolutionary trajectories from a wild-type genotype (denoted 0000) to each of the fitness optima. Using this model, we performed \textit{in silico} assays for collateral sensitivity mirroring the approach taken Imamovic and Sommer [10] (Figure 2). For each drug, we first stochastically simulated an evolutionary trajectory from the wild-type genotype to a local fitness optimum genotype and then, for all other landscapes, compared the fitness of this local optimum genotype to that of the wild-type. A schematic of this simulation is shown in Figure 2(A). Figure 2(B) shows an example of two evolutionary trajectories, which are modelled as sequences of randomly arising fitness conferring substitutions achieving fixation, that can arise stochastically in the fitness landscape.
for Ampicillin as derived by Mira et al. [15].

We exhaustively enumerated all tables of collateral response that can arise under this model. Figure 2(C) shows the best case (most susceptible following evolution), worst case (highest resistance following evolution) and most likely (mean and median values arising for each pair) collateral response tables that arose. In these tables, columns indicate the drug landscape under which the evolutionary simulation was performed and rows indicate the follow-up drug for which fold-change from wild-type susceptibility was measured. This analysis shows the remarkable variation in collateral response that can arise from divergent evolution under a first drug. Indeed, we find a total of 82,944 unique tables can arise, of which the most likely occurs with probability 0.0023. Amongst the 225 ordered drug pairs, only 29 show a guaranteed pattern of collateral sensitivity, whilst a further 94 show a pattern of guaranteed cross resistance. For 88 pairs, the first drug can induce either collateral sensitivity or cross resistance in the second as a result of divergent evolution under the first drug. Critically, if a table of collateral response table is generated by stochastic in silico simulation of the methodology of Imamovic and Sommer [10], and a collaterally sensitive drug pair chosen at random from this table, then the first of these two drugs will induce cross resistance in the second with probability 0.52.

The mathematical model used to derive these results represents a simplification of biological reality, owing to the assumptions of a monomorphic population and a parametrisation using only small fitness landscapes. To experimentally validate our predictions, we verified the existence of divergent collateral response through experimental evolution. Mirroring previously experimental approaches [5, 11, 10, 16, 29], we performed in vitro evolution of E. coli in the presence of the β-lactam antibiotic cefotaxime. The bacterial populations were grown using the gradient plate method with concentrations of cefotaxime varying between 0.06 µg/ml and 256 µg/ml over a course of 10 passages lasting 24 hours (See Figure 3(A) and Methods for details). For each replicate, and after every second passage, aliquots were taken such that the minimum inhibitory concentration (MIC) for a panel of second line drugs could be determined. A time-series for the MIC of the 12 replicates under cefotaxime is shown over the 10 passages in Figure 3(B) indicating that each replicate exhibited increased drug resistance after the 10th passage, although with varying magnitude and trajectory.

For each of a panel of 40 second-line antibiotics, the MIC for the strains X1-X4 was determined following passage 10, in addition to the MIC for the wild-type strain (Supplementary Table 2). From these MIC values, a smaller panel of second-line antibiotics appearing to exhibit divergent collateral response was identified and the MIC of these drugs calculated for each of the 12 evolutionary replicates. Figure 3(C) shows the table of collateral response for this restricted panel following the final passage in the experiment. As predicted, we identify divergent collateral response for the commonly prescribed antibiotics piperacillin (PIP), ticarcillin/clavulanic acid....
Figure 2. Mathematical modeling predicts highly variable collateral response. A) A schematic of the model used to derive collateral response. Sequential mutations are simulated to fix in the population until a local optimum genotype arises. The fitness of this resultant genotype is compared to the fitness of the wild-type genotype for each of the panel of antibiotics. B) The landscape for ampicillin derived by Mira et al. [15] represented as a graph of genotypes. Arrows indicate fitness conferring mutations between genotypes represented as nodes. Cyan nodes indicate genotypes from which evolution can stochastically diverge, grey nodes indicate genotypes from which there is only a single fitness conferring mutation. Squares indicate local optima of fitness with colour indicating the ordering of fitness amongst these optima (darker red indicates higher fitness). Two divergent evolutionary trajectories, in the sense of the model shown schematically in A, are highlighted by coloured arrows. C) The best, worst, median and mean tables of collateral response derived through stochastic simulation of the experimental protocol. Columns indicate the drug landscape under which the simulation was performed and rows indicate the follow-up drug under which the fold-change from wild-type susceptibility is calculated. Bar charts indicate, for each labelled first drug, the number of follow-up drugs exhibiting collateral sensitivity (blue) or cross resistance (red) in each case.
(TIC) and ampicillin/sulbactam (AMS). The patterns of collateral response exhibited between these drugs are not identical, for example, the replicate X12 exhibits increased sensitivity to PIP and AMS but increased resistance to TIC whilst X2 exhibits increased sensitivity to all three drugs and X1 exhibits increased resistance to all three drugs.

Differential patterns of drug resistance could be driven by the different strains having experienced different numbers of sequential mutations along a single trajectory wherein each induces a shift in response (temporal collateral sensitivity [29]), by competition along individual trajectories [18], by evolutionary divergence at a saddle point in the landscape or by non-genetic mechanisms of resistance. To elucidate the underlying mechanism, we performed targeted sequencing of the gene SHV for each of the 10 passage time points and the 12 evolutionary replicates (Figure 3(B)). We identified five variants of SHV amongst the 12 replicates. X1, X5, X7-X9 and X11 all exhibit wild-type SHV, X2 exhibits the substitution G242S, X3 exhibits G238C, X4 and X6 both exhibit G238A, and X10 and X12 both exhibit G238S. Our analysis revealed no evidence of double substitutions in SHV, although mutations to genes other than SHV could not be excluded. Such a mutation might explain the different drug sensitivity of the replicates X10 and X12 (both of which harbour G238S) to PIP and AMS. This analysis identifies a minimum of four fitness conferring substitutions that can occur in SHV during exposure to cefotaxime, indicating the existence of a multi-dimensional evolutionary saddle point in the fitness landscape. Further, the sensitivity of the population to a second drug is dependent on which of these substitutions occurs (Figure 3(C)). For example, G238C (replicate X3) induces increased susceptibility to TIC whilst G238A (replicates X4 and X6) induces a slight increase in resistance.

To conclude, we have shown the existence of an evolutionary saddle point in the fitness landscape of cefotaxime that can induce divergent evolution and differential collateral response in second-line antibiotics. Further, through a mathematical model of evolution parametrised by small, combinatorially complete fitness landscapes, we have highlighted the extent and importance of this phenomenon of evolutionary divergence. Specifically, modelling highlights that divergent collateral response is likely common (occurring in 14/15 drugs for which empirical landscapes were derived) and further, that even where collateral sensitivity is reported from small number of evolutionary replicates, cross-resistance can still occur with high likelihood.

Taken together, our results highlight the potential advantage of reporting tables of collateral response derived from evolutionary experiments with many replicates. In the worst case, where too few replicates of evolutionary replicates are performed, the reported tables of collateral response may indicate an effective, collaterally sensitive, drug pair where in fact the first can induce substantial cross-resistance in the second. Rather than give up entirely on the concept of collateral sensitivity between drugs, we propose that collateral sensitivity likelihoods (CSLs) are
Figure 3. Experimental evolution reveals divergent collateral response. A) A schematic of the evolutionary experiment. *E. coli* were grown using the gradient plate method and passaged every 24 hours for a total of 10 passages. Twelve replicates of experimental evolution were performed. B) The MIC of each strain (X1-X12) under cefotaxime was measured following passages 0, 2, 4, 6, 8 and 10, these values are plotted. The gene SHV was sequenced following each passage. Geometric shapes indicate distinct mutations at the earliest time point they were detected. C) A partial collateral response matrix showing the fold–change in susceptibility for the twelve replicates at passage 10 under exposure to four antibiotics: piperacillin (PIP), ticarcillin/clavulanic acid (TIC) and ampicillin/sulbactam (AMS) and ceftolozane/tazobactam (CFL). Differential collateral response is observed for PIP, TIC and AMS. Inset stars indicate value is a lower bound for fold-increase in resistance.
Figure 4. Collateral sensitivity likelihood. (Left) The table of collateral sensitivity likelihoods (CSLs) derived from the mathematical model. Each entry indicates the likelihood that the first drug (rows) induces increased sensitivity in the second (columns). (Right) The CSL table thresholded for drugs with \( p = 1.0 \) (top) and \( p \geq 0.75 \) probability of inducing collateral sensitivity.

Instead reported. For example, Figure 4 shows an example table of collateral sensitivity likelihoods derived from the \textit{in silico} evolution model. By looking for drug pairs with a high likelihood \((p > 0.75)\) of collateral sensitivity (instead of guaranteed \( p = 1.0 \)) we see that the number of potentially effective drug pairs is increased and further, the inherent risk associated with each pair is explicitly stated. To empirically derive CSLs will likely require novel experimental approaches. We propose two here: firstly, high throughput \textit{in vitro} evolution experiments, likely facilitated by automation of the experimental process \cite{27}. Secondly, as drug sequences are frequently prescribed in the clinic, we propose the distributed collection of matched pre– and post–therapy drug sensitivity assays, possibly coupled with genomic sequencing, to permit the derivation of CSLs. A similar approach is already employed in the treatment of HIV to monitor the evolution of drug resistance \cite{9, 13}. Regardless of the approach taken to derive CSLs, what is clear is that we must move beyond the present methodology of designing drug sequences through low–replicate–number experimental evolution, and towards an evolutionarily informed strategy that explicitly accounts for the inherent stochasticity of evolution.
Methods

Mathematical Modelling of Evolution

The probability for evolutionary trajectories through the empirically derived fitness landscapes were calculated from a previously described mathematical model [17]. Briefly, the population is assumed to be isogenic and subject to Strong Selection Weak Mutation (SSWM) evolutionary dynamics. Thus, the population genotype (taken from domain \{0, 1\}^4) is modelled as periodically replaced by a fitter (as determined by the landscape) neighbouring genotype (defined as any genotype whose Hamming distance from the population genotype is equal to one). This process is stochastic and the likelihood of a genotype, \(j\), replacing the present population genotype, \(i\), is given by

\[
P(i \rightarrow j) = \begin{cases} 
\frac{\sum_{g \in \{0, 1\}^4, \text{Ham}(i, g) = 1} (f(g) - f(i))^r}{(f(j) - f(i))^r} & \text{if } f(j) > f(i) \text{ and Ham}(i, j) = 1 \\
0 & \text{otherwise}
\end{cases} \quad (1)
\]

Where no such fitter neighbour exists, the process is terminated. The value of \(r\) determines the extent to which the fitness benefit of a mutation biases the likelihood that it becomes the next population genotype. We take \(r = 0\), corresponding to fixation of the first arising resistance conferring mutation, but our results are robust to changes in \(r\).

For the simulations of in vitro evolutionary experiments, we assume an initial genotype of \(g_0 = 0000\) and determine the final population genotype by sampling from the model until termination at a local optimum of fitness, say \(g^*\). Simulated collateral response was calculated as the fold difference between \(g_0\) and \(g^*\) in a second fitness landscape.

Experimental Adaptation to Cefotaxime

All 12 evolutionary replicates were derived from with \textit{E. coli} DH10B carrying phagemid pBC SK(-) expressing the \(\beta\)-lactamase gene SHV-1 [22]. All evolutionary experiments were performed in Mueller-Hinton agar.

Using a spiral plater, cefotaxime solution was applied to Mueller Hinton (MH) agar plates in a continuously decreasing volume equivalent to a thousand-fold dilution. \textit{E. coli} DH10B pBCSK(-) \textit{bla}_{SHV-1} colonies were suspended to a concentration of 7log10 CFU/ml in MH broth. Antibiotic plates were then swabbed along the antibiotic gradient with the bacterial suspension. Plates were incubated overnight at 37°C. The most resistant colonies, as measured by the distance of growth along the gradient, were resuspended and used to swab a freshly prepared antibiotic
plate. The process was repeated for a total of 10 passages. The entire experiment was completed 12 times using the same parent strain to generate the cefotaxime resistance strains X1–X12.

Determination of Minimum Inhibitory Concentration

The minimum inhibitory concentration of each of the antibiotics in Figure 3 was determined for both the parent strain and the 12 cefotaxime resistant strains X1-X12 according to guidelines outlined by the Clinical and Laboratory Standards Institute [3]. MICs were calculated in triplicate and the mean value used in for analysis. The maximum concentration considered was, 4096µg/ml, where the MIC exceeded this concentration the precise value was not calculated and a MIC of 8192µg/ml was used in the analysis. In this case the associated collateral response value is delineated to indicate that it is a lower bound. For X1-X4 the MICs were calculated for an extended panel of antibiotics (Supplementary Table 2).

Collateral Sensitivity Analysis

Collateral sensitivity (or resistance), as depicted in Figure 3, was determined from the mean MIC values for the parent and passage 10 adapted strains (X1 - X12). For strains $x = 1 \ldots 12$ the collateral response to an antibiotic, $d$, is calculated as

$$ CR = \log_2 \left( \frac{MIC_d(X_x)}{MIC_d(parental)} \right) $$

(2)

Sequencing and Analysis

Plasmid DNA was isolated using the Wizard Plus Minipreps DNA purification systems (Promega). Sequencing of the SHV gene was performed using M13 primers (MCLab, Harbor Way, CA).

References


