Answer to reviewer and editor comments

Editors: *Inês Fragata and Claudia Bank*

This Perspective was reviewed by 3 external reviewers, with whom l agree that a criticism of the misleading use ot the term "costs of resistance" is a timely issue and that this manuscript can be of relevance for both empirical and theoretical studies.

Thanks for the comments and for handling this evaluation process.

However, all reviewers provided excellent suggestions that would allow for the manuscript to reach a larger target audience and that would improve its clarity. Most importantly, all reviewers suggest that the link to empirical studies and the implication for such studies needs to be more developed, and that there should be more concrete suggestions on on how to move beyond the term cost both theoretically and empirically. The reviewers also provide several interesting references that may complement the literature review provided by the authors.

We modified the last paragraph to develop concrete suggestions.

Although we appreciate Reviewer 3's concern that complementing the existing discussion via Fisher's Geometric model (FGM; which should indeed be defined to the "naive" reader) by a discussion of models based on dose-response curves may be illustrative and helpful for readers less familiar with FGM, we feel that this may go beyond the scope of the current manuscript. However, this alternative and commonly considered model of fitness effects across environments should be discussed.

We added a box to specifically make the correspondence between dose-response models and (the more general) fitness landscape models. We also added many comments in this box about ecotoxicological measures of resistance, as suggested by the reviewers.

For people unfamiliar with FGM, it could also be helpful to indicate the important aspects of the model in each figure directly, i.e., "Optimum AB-" instead of "O", etc.

We added details to the legends, following many suggestions of the reviewers. We also added an inset figure to figure 1 to illustrate more clearly the fitness mapping in 3D. We are unsure to understand what is meant here by “Optimum AB-“

Reviewer #1 *Danna Gifford*

In this Perspective, the authors tackle a juggernaut of the applied evolution literature, the “cost of antibiotic resistance”. Their essential argument is that describing resistance mutations as “costly” in the absence of antibiotics is an oversimplification that has lead the field astray—essentially, we ought to be considering fitness effects of resistance mutations in different environments, as we would for any other class of mutation.

I think this perspective is worthwhile, though perhaps a touch adversarial. The field at large does appear to be aware of the fact that resistance is not always costly (but they are “generally

costly"). I suppose I would like the authors to address why resistance isn't the default state for wild-type organisms, if resistance isn't generally costly.

Our whole point is that the term 'cost' is not very useful or helpful. Why resistance is not the default state is simply because, in absence of human intervention, environments are characterized by low (or zero) doses of pesticides (many antibiotics are naturally produced by some organisms, but are not present at massive doses). This is a simple Darwinian explanation, and "cost" play no part in it.

Similarly, the default state of peppered moths is to be light-colored in absence of pollution. It is adaptive to be light-colored in non-polluted areas, not because being dark-colored is 'costly', but because being light-colored confers better camouflage. In a polluted area, symmetrically, it is adaptive to be dark-colored, but, again, this is not because being light-colored is 'costly'. One could in principle argue that each phenotype is costly in the other environment, but then, why not just talk about the fitness effects of the traits in each environments directly? If both traits are costly, why introducing the term in the first place?

For human induced environmental changes, the 'default state' is likely to simply reflect the phenotypic adaptation corresponding to the human-free environment. In general, however, there is no clear reason to define adaptation in reference to past historical circumstances. According to (Reeve and Sherman 1993), adaptation is best defined here and now.

I think the section on considering the fitness effects of resistance in non-optimal genetic backgrounds is the most interesting and important aspect of this work. There are certainly specific situations where resistance mutations provide a benefit in non-optimal organisms (e.g. Kassen and Bataillon 2006, which the authors cite, although Bataillon et al. 2011 showed that many of the strains are not single mutants; and some examples in the rifampicin resistance literature e.g. Rodriguez-Verdugo et al. BMC Evol Biol doi:10.1186/1471-2148-13-50). The authors do an excellent job outlining this conceptually from FGM, but I do wonder about the strength of support from data---though one could certainly argue that the data are biased toward finding costs, due to the popularity of the concept of "costs of resistance" itself.

According to us, the problem is more fundamental than an ascertainment bias. It should not be surprising to find any value for "cost" (positive, zero or negative). This is not due to secondary mutations (although these mutations are important to consider). As we explain, all values are entirely plausible, as it mainly depends on whether the wild type (relative to which fitness measurements are made), is well adapted or not to the pair of environments considered.

The overall bias toward "positive costs" is simply caused by the choice of genotype and environment taken as references. These genotypes tend to be relatively well adapted to the environment without drug, so that mutating them tend to disrupt this adaptation.

In concrete terms, finding zero or even "negative" costs does not indicate an absence of trade-off across environments. This point is important as many researcher tend to think that an absence of "cost" indicates that resistance phenotype do not present any trade-off and should therefore evolve, regardless of any resistance management strategies. This is wrong, and one of the situations where the concept of cost is misleading. We now insist on this point in the last paragraph in discussion.

(A minor point, but the manuscript would benefit from minor copy editing to fix consistency in usage of e.g. single vs. double quotes, subject/verb agreement, grammar, etc.)

We did our best to check the text.

Section “Resistance mutations as beneficial mutations” The statement “In brief, resistance mutation are beneficial mutations.” is oversimplified and should be qualified with “in the presence of antibiotic”, otherwise it would seem to suffer from the point the authors are trying to make, that context is everything.

Of course, in presence of pesticides/antibiotic. We updated the text accordingly.

Section “The context dependence of fitness effects” “This selective advantage is not easy to estimate in the field, but is often thought to represent an inherent property of the mutation itself.” Nevertheless, there is parallelism in terms of the specific resistance alleles that are observed in clinical isolates, particularly notable in the TB rifamycin resistance literature, where recombination is thought to be low.

Parallel evolution can have multiple causes (see discussion in Lenormand, Chevin and Bataillon 2016 “Parallel evolution: what does it (not) tell us and why is it (still) interesting?”). Observing parallel evolution does not contradict the fact that the effect of mutations is context dependent. This can be explained using an example:

The G119S point mutations in *ace1* gene have been occurring independently in widely divergent groups of insects in response to treatments with organophosphorous insecticides. It represents one of the most astonishing case of parallel evolution. Yet, this mutation shows various fitness effects depending on the environment (e.g. it is “costly” in absence of treatment). Regarding “genetic context”, it appears that some essential functions (here nervous signaling), and corresponding genes (here *ace1*) are conserved at large phylogenetic distances and timescales. If this G119S mutation confers such resistance in such a diversity of species, it is because the genetic context is constant in this case (*ace1* gene carrying out the same function in all these species), not because the mutation has a context-independent effect.

What we mean by inherent property is something that would not depend on the environment or genetic background.

Section “Costs of resistance are not pleiotropic effects” After reading this section, I did not fully grasp the argument that costs of resistance are not due to pleiotropy. Certainly some resistance mutations are highly pleiotropic (e.g. rpoB resistance, which globally affects gene expression levels and patterns, Qi et al. 2014, doi:10.1128/mBio.01562-14).

This is illustrated on Fig 1: the cost depends on the segment RO, while the pleiotropic effects depend on the segment P_2R. The segments are distinct and therefore “Costs of resistance are not pleiotropic effects”. Imagine that a resistance mutation change the phenotype exactly in the direction of the new phenotypic optimum (along direction RO on Fig. 1). It would have a cost, but no pleiotropic effect. In other word, it would not change other traits than the one that corresponds to the direction of the new environmental optimum.

Figure 1: please identify P1 and P2 in the legend text.

We added the explanation in the legend.

Section “Resistance mutations do not have a cost” Might I suggest modifiying this to “Resistance mutations do not *always* have a cost”? Clearly sometimes resistance mutations do/can have a cost relative to the wild-type, as demonstrated in Figures 1 and 2.

We added italics and quotes on “*a cost*” to emphasize that the title is about the terminology. Saying that a resistance mutation has “*a cost*” conveys, according to us, two wrong ideas. We develop these two points in this paragraph, justifying this title (we reiterate these points below). Our point is not that resistance mutations do not *always* have a cost (they always have a fitness effect in all environments). We challenge the word “*a*” and the word “*cost*”. We challenge the terminology, not the fact that the fitness of resistance mutation in absence of drug can be lower, the same or higher than that of a particular wild type, chosen as reference (these measures are just data, and therefore should be uncontroversial).

The first issue is that the term “*a cost*” tends to convey the idea that a mutation has *one* fitness effect in absence of drug. This is of course wrong, as there are many possible environments without drug. The second idea is that “*a cost*” tends to indicate that the fitness of resistance mutation is necessarily lower than the wild type in absence of drug. This is implicit, because it makes no sense to talk about a ‘negative’ cost. In the usual economic sense, cost is necessarily positive, or it would not be a “*cost*” in the first place. Because of this terminology, people studying resistance mutations tend to have wrong expectations. For instance, they tend to be surprised when they find a resistance mutation with no “cost” or a “negative cost”. In this case, many researchers would conclude that the absence of “cost” reflects an absence of trade-off across treated and non-treated environments. As we explain, this conclusion is erroneous and, in our opinion, this error stems, in part, from the usage of the word “*cost*”. We added clarifications in the text to better explain this (Line 452-468)

Of course, nowhere do we mean that the fitness of a resistance mutation cannot be lower compared to wild-type, in absence of drug.

I think a point worth addressing would be, if resistance mutations do not generally have a cost, why is resistance not the default state for organisms?

The phenotypic state of an organism usually reflects its history of adaptation to its environment. There is no clear concept of ‘default state’ beyond this idea. For instance, many organisms on earth are resistant to oxygen, a toxic compound present in the atmosphere (but only at specific concentrations). In this case, resistance is the default state and it can be explained by the fact that organisms on earth evolved in the presence of O_2.

At the risk of appearing self-serving, the authors may find interest our work on the effects of rifampicin resistance mutations at a single locus under different environmental conditions (temperature, carbon source, Gifford et al. 2016, doi:10.1111/evo.12880), which I think demonstrates the authors’ essential point: resistance mutations are not necessarily deleterious, and that compensated resistance is not always fully restorative. The authors may also like a recent publication looking at genetically diverse E. coli from Basra et al. (2018, GBE, doi:10.1093/gbe/evy030).

Thanks for the references. We added the reference to Gifford et al 2016.

Figure 2: Please identify P1 and P3(?) in the figure legend.

Done

Section “What is a resistance mutation?” “If resistance mutations cannot be defined by the fact that they are beneficial in the treated environment”: it would be helpful to have a citation for this definition for resistance. I believe that a more commonly used definition is a mutation (or gene) that allows growth at a specific concentration of antibiotic, in the genetic background otherwise incapable of growing at that concentration.

The sentence “If resistance mutations cannot be defined by the fact that they are beneficial in the treated environment” refer to the explanation in the previous paragraph.

Indeed, we agree that most people would say that a resistance mutation “is a mutation (or gene) that allows growth at a specific concentration of antibiotic, in the genetic background otherwise incapable of growing at that concentration”. The problem is that such test does not guarantee that the mutation was specifically ‘dealing’ with the problem posed by the presence of the drug.

Let’s take an extreme example for trying to make this point more clearly: imagine that the environment used to ‘test’ the mutation presents two challenges (e.g. high salt concentration and some drug). ‘Growth’ might be obtained if the mutation improves the handling of osmotic stress without changing anything about the drug-related physiology. Of course, in this extreme example, it requires that the effect of the drug alone is not fatal, but that it was the combination drug+salt that was fatal to the wild type. This is exactly the point: it may not be *a priori* clear that a mutation conferring growth does so by only solving the physiological problem posed by the drug. Hence it is perhaps not so easy to define clearly what we mean by ‘resistance’ mutations, especially if the environment used for the test is rather artificial. This explanation repeats what is already indicated in the text.

Reviewer #2 *Anonymous*

The perspective by Lenormand and co-authors provides an interesting short historical overview of the term “cost of resistance”, the theoretical limitations of the term and discuss the academic usefulness of it. The authors present valid and clear arguments in a well-written manuscript and I enjoyed reading it. Moreover, I agree with the authors that the term “cost of resistance” has serious limitations. For instance, resistance mutations are indeed usually considered pleotropic (in traits) and the direct cost of the mutations is often forgotten. The cost of resistance in fact, is not the same entity as the “cost” of pleiotropy.

However, I feel that the perspective is somewhat biased. For instance, abolishing the term will not solve of the issues that fitness effects are dependent on the environment. Which environments are more relevant to study remains an open question since it is not possible to study fitness effects directly in some of the relevant environments (i.e.: the human host). Even though lots of problems arise when fitness costs are measured in vitro (i.e. in the test tube) and it is difficult to choose the right environment to study, it has led to the useful predictions that resistance mutations with lower costs should be more prevalent in clinical isolates of some pathogens. This was indeed observed for clinical isolates of Mycobacterium tuberculosis (Gagneux et al., Science 2006).

We are not sure to see where we disagree. Indeed, some test environments are more relevant than others. Indeed, we have a clear expectation that adapting to a new environment B, for a wild type well adapted to A, is likely to reveal a trade-off. Indeed, we expect that the first beneficial mutations to environment B will be likely to have negative pleiotropic effects that will be subsequently corrected. All these considerations are related to the theory of adaptation across multiple environments. They can be made very clear without having to use the concept of "cost of resistance". The concept of trade-off to adapt to different environments is sufficient to explain all this. In this perspective, we do not question the usefulness of the theory of adaptation (quite the opposite). We question the concept of "cost". As we previously mentioned, all the theory related to adaptation to different environments did not historically require any reference to this concept.

Having said this, we do not necessarily want to 'abolish' the term. It has become so widespread, that it may be difficult to entirely avoid it. We feel however that pointing its limitations is important given the sloppy usage and implicit (wrong) expectations it carries along. We added this more explicitly in the last paragraph of the discussion (line 419-426).

Moreover, the authors have not discussed the practical applications of the term "cost of resistance". For instance, this term has been helpful raising awareness to the problem of antibiotic resistance and it has influenced political decisions such as to halt the use of certain antimicrobials with observed decrease of antimicrobial resistance in clinical settings (Seppälä et al., 1997, N. Engl. J. Med. 337, 441–446; Enne et al., 2001 Lancet 357, 1325–1328; Bean et al., 2005 J. Antimicrob. Chemother. 56, 962–964; Gottesman et al., 2009 Clin. Infect. Dis. 49, 869–875).

Thanks for these references. They report the decline of antibiotic resistance (or the lack thereof) in microbes, following a decrease in antibiotic usage. We agree that the concept of cost was, at least initially, useful to raise awareness about the ecological conditions (dis)favoring resistance evolution. This is something we mention early on in our perspective:

"the [cost] concept **was helpful** to bring attention to the fact that a mutation could be both beneficial or deleterious, depending on circumstances, something well known in ecological genetics, but somewhat ignored in resistance studies. It **helped** introduce some ecology in the understanding of the fitness effect of resistance mutations. This can have **important consequences** as the cost of resistance is a powerful force that can keep resistance in check".

We also reiterate this point in the first sentence of the summary/conclusion: "taking into account the 'cost of resistance' has been a **major progress** because it is essential to distinguish the fitness effects of resistance mutations in treated versus non-treated environments"

(the emphasis was added here)

However, the relevant (and well-established) concept here is trade-off across environments. Raising awareness can be made, at least as convincingly, using this more robust concept. There are also possible confusing implications of using the "cost" terminology. For instance, it is frequent to see papers mentioning that resistance mutations with no (or negative) cost cannot be dealt with, in a resistance management context. This can be misleading, as an absence of cost does not necessarily indicate an absence of trade-off across environments, as

we explain. We developed ‘practical implications’ in our last paragraph to insist on this (lines 416-480).

Just to take an example from the most recent of the four papers mentioned (Gottesman et al 2009). Quoting their introduction: “If there is no fitness cost, then there is no force favoring the reversal of resistance, even when antimicrobial exposure is halted.” It is not specified in this sentence how the “cost” is measured. Even if zero cost is found in the lab, against some reference wild-type, it does not demonstrate the absence of fitness trade-off between treated and non-treated environments. This short-cut is typical, and the implication for resistance management should be made more cautiously.

Similarly, an absence of a frequency decrease in the field is not an indication that there is no cost. Frequency can change because of gene flow, or positive selection may still be present (e.g. due to treatments occurring in different contexts, e.g. in farms for antibiotics).

Demonstrating that the failure of a resistance management strategy is caused by an absence of fitness trade-off between treated and non-treated environments requires much more work than exhibiting that a resistance mutation has “no cost”. In our opinion, this absence of “cost” is an easy excuse in this context. The most likely reasons for the failure of management practices are that (1) the original drug or a drug showing cross-resistance, is still used, possibly secretly; (2) it takes comparatively more time for selection to operate and revert resistance than it took to select for resistance (this is simply because of the difference in magnitude of the selection coefficients in the treated and non- treated environments for the resistance mutation). We also added these points in discussion lines 464-468.

Reviewer #3 *Helen Alexander*

Thanks for this thorough review, it is not so frequent to see a review almost as long as the original paper.

This Perspective manuscript by Lenormand, Harmand and Gallet addresses the concept of ‘cost of resistance’, which is commonly defined as the reduction in fitness in the absence of drug (compared to the wild-type) associated with a drug resistance mutation. They trace how this concept developed historically and then highlight several of its shortcomings. My impression is that the authors have reviewed the literature quite thoroughly (with a few suggested additions noted later) and thus their perspective is well-grounded. They raise valid issues with the definition and usage of the term ‘cost of resistance’, and I believe this article can provide a worthwhile contribution to the literature by highlighting and bringing together several problems that, while not entirely new, have been underappreciated.
As a reader with a primarily theoretical background, I enjoyed reading this manuscript and was generally convinced by the authors’ arguments. However, my main concern is that this message may fail to reach a broader readership, including experimentalists, in its present form. The authors may choose to ignore this critique if this is not their target audience, but I believe their manuscript has the potential to have a much greater impact if they address this point. Specific suggestions are detailed below.

This is indeed a recurring concern, also with the other reviewers. We developed the ‘practical considerations’ in the last § (lines 416-480).

Secondly, while the geometric model is appealingly simple and useful for illustration, it comes across

rather dominantly in the current version of the manuscript. Although the ideas that the authors put forward are not specific to this model formulation, the generality of the points they raise could easily be lost due to the emphasis on the geometric model (including all 4 figures of the manuscript). More clearly conveying this generality could be achieved by adding some discussion and illustration using a model that deals directly with fitness, e.g. by linking to dose-response curves (see below).

We agree that dose-response models are useful and widely used in ecotoxicology, so we added a box 1, to make the correspondence and explain how a dose-response can be deducted from a fitness landscape model. However, we kept the fitness landscape model in the main text as they are more general and help dealing with several issues that could not be easily dealt with dose-response models (e.g. the occurrence of several possible optima without drug, the occurrence of several traits and pleiotropy).

In this box, we also mention a series of other topics (measures of fitnesses, relation to ecotoxicological measures). We hope this addition will help bridging the gap of generality.

I don't suggest to replace the geometric model, which is elegant and useful, but rather to complement this with some broader discussion.

Dose response models are, according to us, specific, more than the fitness landscape models we use. Fitness landscape models have been used to describe adaptation in a very broad sense and across many ecological situations, unlike dose-response models. See above and new Box 1.

Finally, I suggest to devote a bit more space on how to move forwards beyond 'cost'. Although the authors do touch on this point (including in the Abstract, where they propose "to study, measure and analyze the fitness effects of mutations across environments and to better distinguish those effects from 'pleiotropic effects' of those mutations"), the manuscript would benefit from a dedicated section providing concrete suggestions, including stronger links to experimental approaches.

We agree, and developed the last paragraph about 'practical implications'

As a technical point, it would be great if the authors could add page and line numbering to the next version of their manuscript, in order to facilitate the commenting process.

Sorry for the oversight. We included line numbering.

More specific comments on the content follow.

Model description:

* It would be useful to have a slightly more detailed introduction to the geometric model and its assumptions, as this forms a large part of the exposition but may not be familiar to all readers, even some theoreticians.
* In particular, although it is written that "we can assume that fitness declines with the distance from the peak in any given environment", please give the mapping from phenotype to fitness explicitly, and highlight the key assumption that fitness depends only on (Euclidean?) distance from the optimum in multi-dimensional trait space.

Thanks for pointing this out. Indeed, we agree that the text and the figure legends were not clear enough on the distance-to-fitness mapping. This is now mentioned several times in the text and in the figure legends (lines 172-180, 190-194, 315-317…). We also added a 3D inset in Fig. 1 so that it is clear to the reader that we are indeed considering a distance-to-fitness mapping.

Note that we do not require a very specific mapping to make the qualitative points we are making. We therefore avoided being too specific in order to avoid writing a too mathematically oriented paper. Here, and throughout, we only assume a monotonous mapping between Euclidian distance and fitness. Elaborate version of Fisher's geometric model, with explicit mapping, can be found in several papers (see references lines 157).

This mapping is needed in order to make deductions such as "The difference between these two distances [AR and AO in Fig. 1] scales with the selection
coefficient of the resistance mutation in the treatment environment" and that all points P_1 such that $|AR| = |AP_1|$ "confer the same benefit in the treated environment" (i.e. have the same fitness).

These statements simply require a monotonous mapping between Euclidian distance and fitness

* Since the geometric model is rather abstract, it would be helpful to include where possible any intuition or discussion of how it relates to measurable quantities – most importantly in this context, how the definitions of "resistance" and "cost" in this model relate to more commonly used empirical measures.

Discussing how to relate fitness landscape models to actual measures can be quite complicated (if one wants to specify the traits, see the extensive discussion in Martin and Lenormand 2006), or trivial (if one simply talk about fitness, which can be directly measured).

• In the geometric model, "resistance" is represented by a point in phenotypic space that is closer to the optimum in the treated environment than the wild type is (i.e. $|AR| < |OR|$ in Fig. 1). With a mapping from phenotype to fitness, this could be directly related to the common practical definition of resistance as an increase in minimum inhibitory concentration (MIC), meaning the resistant strain can still grow at higher drug concentration than the wild type.

See the new Box 1.

• Similarly, "The cost of resistance is shown by the distance OR, as it is defined as the fitness effect of the resistance mutation in the non-treated environment" (pp. 6-7). However, there is a subtle distinction to make here: this sentence first suggests that cost is in units of distance in trait space, but then that it is rather a difference in relative fitness. These are two different measures, which will be linked by the phenotype-to-fitness mapping assumed in the model. In practice, "cost" is usually measured directly in the currency of fitness (or rather, some proxy such as relative growth rate).

Absolutely. We indeed used a shortcut, which may be confusing. There is always a mapping between distance in the phenotypic space and fitness. The cost is measured in the currency of fitness. We modified the text and figure legends so that the mention of the monotonous mapping is made more clearly (in practice, this has no consequence for our qualitative argument).

* 1st paragraph on p. 7: "Should these [pleiotropic] effects be totally compensated, the phenotype would be in P_2 and it would indeed enjoy a greater fitness in both the treated and non-treated environments." How strongly does this conclusion (enhanced fitness in both environments) depend on the assumptions of the geometric model?

This conclusion simply reflects the definition of 'pleiotropy' used here. Because it is defined as all the 'unwanted changes' (i.e. in phenotypic directions other than the direction to the new optimum), it

necessarily implies that fitness is improved in both environments when pleiotropic effects are removed. With another definition of pleiotropy, it might be different, but it is unclear to us how to define pleiotropy in a better way in this context.

Additional conceptual links and references:
In general, these are not mandatory to add, but may be of interest to the authors and in my opinion would strengthen the manuscript!
* *The context dependence of fitness effects* (p. 3) points out that "thinking in terms of averages" is not always valid. This important point could usefully be expanded. Firstly, thinking in terms of averages can be misleading not only in terms of ecological conditions as the authors already mention, but also in terms of genetic background in the opposite case to that they describe, i.e. for the many relevant species that reproduce asexually or with limited recombination or horizontal gene transfer.

Indeed, asexuality brings in a lot of linkage disequilibria, which further complicate this averaging. This is well known, but there is little to add beyond this observation.

Furthermore, even if averaging (over genetic backgrounds or environments) might be considered reasonable for predicting long-term dynamics, in the initial establishment of rare resistance alleles when stochastic effects dominate, the context in which the allele first arises can be extremely important.

Indeed, this averaging can be made computing a probability of fixation. We agree, but this is a general (well known) point in population genetics, although not directly relevant to our topic, in our opinion.

Finally, while the authors mention that the spatial scale of dose variation (relative to dispersal) is relevant, an analogous point could be made for the temporal scale of dose variation (relative to generation time). A useful reference here would be Cvijović et al. (2015), *PNAS* E5021-E5028 (doi:10.1073/pnas.1505406112).

Yes, this averaging can also be done in fluctuating environments, and the relevant comparison with dispersal scale is generation time. We added a brief mention to fluctuating environments and added this reference (line 74-75).

* More generally, the authors refer to spatially heterogeneous models in several places. These points could often be extended to temporal heterogeneity, with links made to the extensive literature on pharmacokinetics and pharmacodynamics (PK/PD) of drug dosing.

We added a word on temporal averaging, see previous comment.

* What immediately sprung to my mind when reading this manuscript was the relationship to "doseresponse curves", which relate some demographic parameter (e.g. net population growth rate, which can be considered a measure of fitness) to drug dose.

See new Box 1

That these are never mentioned struck me as a glaring omission, particularly in the discussion of varying drug doses on pp. 11-12, which would provide a natural link. I think there are several reasons why it would be useful to bring these up:
• Using keywords such as "dose-response curve" (and perhaps "reaction norm", of which dose-response curves are an example) will catch the attention of more readers.

• Dose-response curves are more easily related to empirical measures than the geometric model, thus raising the interest for experimentalists.

If experimentalists measure fitness, all that we say apply to their results.

• By including a model that deals directly in the currency of fitness (or some measurable proxy), in addition to the geometric model, the authors would have the opportunity to highlight that the conceptual issues they raise are general.

It is quite striking that we consider geometric models of adaptation to be far more general than any dose-response model. If the message is that fitness should be measured in different environments (including different doses of drugs), we entirely agree, and this is actually what we say. Dose response models are usually quite specific, in that they are one dimensional (no pleiotropy), and often, do not involve several non-treated environments.

• Many of the authors' ideas could be very nicely illustrated by plotting dose-response curves, e.g. showing different "costs" of resistance mutations in different genetic backgrounds or environments, and illustrating the issues that arise at varying drug doses.

Thanks for the suggestion, but we are not sure to see exactly how to use such a figure. For instance, how defining pleiotropy with just a curve relating fitness to environmental variation? How deling with different environments without drug? Hopefully, the explanation made in Box 1 is sufficient now.

A figure or two like this could help provide some balance by indicating at a glance that the manuscript is not only about the geometric model.

We understand that this is a concern, and we do not want to give the impression that our presentation is too abstract (it is not, as we talk about fitness measures throughout). We added the Box 1 to address this concern. Thanks for insisting on this.

• There is a natural link between the authors' ideas and existing literature highlighting the limitations of using single-parameter measures of "resistance" or "fitness" (such as the MIC) to predict population dynamics, particularly in PK/PD models. A few relevant references:
- Regoes et al. (2004), Antimicrob. Agents Chemother. 48:3670 (doi: 10.1128/AAC.48.10.3670-3676.2004)
- Sampah et al. (2011), PNAS 108:7613-7618 (doi: 10.1073/pnas.1018360108)
- Gehring & Riviere (2013), Vet J 198:15-18 (doi: 10.1016/j.tvjl.2013.07.034)

Thanks for these references. These papers show indeed that simple metrics such as LD50, IC50, MIC, etc, are only partial summaries and imperfect fitness measures. We entirely agree. We now cite them (adding also Wen 2016, Scientific Reports 6: 37907). It is quite clear that in principle, e.g. two bacteria with the same MIC could have very different fitness at different doses below this MIC. Note that the arguments made in these papers could be taken a step further, by showing that demographic measures made in isolation need not correspond to fitness differences in competition. These points are now mentioned in Box 1.

* Another potential reference, which echoes the authors' point about the dependence of fitness on ecological context: Day, Huijben, Read (2015). *Trends Microbiol.* 23:126-133.

Thanks for mentioning this paper. We added the citation in box 1.

* Middle of p. 7, regarding compensation of pleiotropic effects: here one could also cite work by Dan Andersson's group (e.g. the Andersson & Hughes 2010 review already included in the references, and/or Andersson & Hughes 2012, *Drug Resistance Updates* 15:162-172.)

We tried to cite the original studies.

Discussion of varying drug doses (pp. 11-12):
* The idea that different drug concentrations could represent either different intensities of selection or different phenotypic optima is certainly interesting conceptually, and is clearly explained in the context of the geometric model. However, it could be more clearly explained in practical terms as well. Are there relevant empirical examples, e.g. for the statement that "it is fairly easy to imagine two mutations R_1 and R_2 that would qualify as resistance mutations, in each of the two environments, but not in the other"? My guess (although unfortunately I cannot offer a reference off the top of my head) is that there may well be such cases, e.g. where gene overexpression or amplification, an efflux pump, or enzymatic degradation of an antibiotic confers "resistance" at low doses, but only a target modification confers "resistance" at high doses.

We discuss this extensively in Harmand et al 2017, and in a newly published paper (Harmand et al 2018, Evolution Letters, in press "Evolution of bacteria specialization along an antibiotic dose gradient"). We added the reference.

* I think this whole discussion would flow better with a bit of reorganization. The argument (p. 12) that "it is difficult to conceive that adding a vanishingly small quantity of drug suddenly shifts away phenotypic requirements, and that further increases in dose only change the selection intensity" is convincing, and I think this could usefully be moved up front to where the idea of different optima versus different selection intensity is first introduced on p. 11. Likewise, the admittance that this has not been demonstrated empirically could be moved along with it. The discussion on benefits and costs at different drug doses would seem to flow more naturally afterwards. Indeed, the point that fitness is more generally a function of drug dose can be made without relying on the (rather abstract) distinction between distinct optima vs. different selection intensity.

Thanks for the suggestion. It does work better, reorganizing as suggested.

* "This [association of strong resistance with high costs] may well be true, but not necessarily" (p. 12). It would be great to back this up with empirical counterexamples, if available.

A counter example is available in e.g. Harmand et al 2017. We added the citation.

* Is the last sentence on p. 12 ("In any case, representing evolution of resistance as convergence to a phenotypic optimum has received some empirical support") specific to this section's discussion of varying drug doses, or is it more general? It sounds like a more general point that could provide an important connection to empirical literature, and thus could be made more prominent in the manuscript.

We moved this to the beginning of the paper where we introduce the fitness Landscape Model (line 160-163). Thanks for the suggestion.

Making the manuscript more accessible and relevant to a broader audience, including experimentalists:
* Several of the above suggestions already go towards increasing accessibility of the paper and relating it to empirical approaches. In addition, the authors could strengthen their case that the

concept/terminology of “cost of resistance” is a “hindrance” not only in models, but also in practice. I think they could readily argue that over-simplifying the fitness effects of “resistance” mutations will hinder prediction of the evolution of resistance, given that natural environments are multi-faceted and heterogeneous.

We developed several practical considerations in the last paragraph (lines 416-480).

* The authors could also devote more attention to what could be improved in practice (e.g. what should be measured in empirical studies). Here, it may be worth acknowledging that while their suggestions are clear and “simple” conceptually, they are not necessarily simple to implement: e.g. in the concluding section on p. 13, saying that “it may be safer in most cases to simply discuss and measure the fitness effects of mutation in different environments” sounds good in theory, but entails a lot of work (and decisions about which environmental factors to vary) in practice. Can the authors point to empirical studies that have made progress in this direction?

We developed several practical considerations in the last paragraph (lines 416-480).

Minor comments on wording:

* In a couple of places, the authors refer to the importance of “ecological conditions” when really environmental conditions, more broadly defined (both abiotic and biotic), are relevant. E.g. beginning of *The context dependence of fitness effects* (“the selective effects of mutations depend on ecological conditions”) and middle paragraph on p. 4 (“it helped introduce some ecology…”).

We are not sure to see the point. We do not make a fundamental difference between “ecological conditions” and “environmental conditions”. Ecological conditions encompass both biotic and abiotic conditions. In our view, ‘ecological’ is broader as ‘environmental’ could be understood as only reflecting abiotic factors.

* p. 3, 1st sentence under *The cost of resistance*: “This is where the concept of ‘cost of resistance’ becomes important” – I find this sentence too vague.

This transition sentence was rewritten.

* p. 4 (section *The cost of resistance*): “In particular, all the paper[s] on local adaptation, clines and all the field of ecological genetics developed without the need to refer to this concept [cost].” This statement, without further qualification on “all the paper[s]”, feels a bit too sweeping. Perhaps the authors could rephrase to mention the time frame they are referring to?

We added the time frame “before the 80s”. Excluding clines in hybrid zones (which are irrelevant to resistance evolution), most of the basic theory about clines was done before the 80s, i.e. before the concept of cost of resistance was even introduced. For instance, see Endler’s book on clines published in 1977. See also our answer below for references about ‘ecological genetics’.

It might also be more enlightening to summarize these topics as general models of adaptation in heterogeneous environments, if that is an accurate assessment. Again, could models involving temporal heterogeneity, as well as spatial heterogeneity, be included here?

The idea of trade-off and specialization was much more important with spatial than temporal variation, since polymorphism can be maintained in the former, but not easily with the later (see e.g. Hedrick, P.W. 1986. Genetic polymorphism in heterogeneous environments: a decade later. Ann. Rev. Ecol. Syst. 17: 535 566.). Our topic is clearly not to review model of polymorphism…

* pp. 4-5: In the paragraph on how cost arose from life history theory: I don't really see how it became "natural to think that the cost can evolve to be reduced, or even eliminated" only in later interpretations.

Because if cost simply arises from deleterious pleiotropic effects of the resistance mutation, it seems possible to reduce this cost to zero, by compensatory evolution eliminating these deleterious pleiotropic effects. Of course, this view is erroneous as costs are not pleiotropic effects (see Fig. 1).

* p. 5, 1st paragraph: "the best proof for this reasoning is that cost-free mutations are sometimes found…". This might be better reworded as "the best support for this reasoning is that apparently cost-free mutations are sometimes found".

No, these mutations are not 'apparently' cost free. They *are* "cost" free. The concept of cost is strange, not the actual measures that are wrong. It is perfectly fine and normal to find mutations with zero "cost".
It sounds odd simply because the term cost is used, but the Fig. 2 shows how it can simply happen. It just means that the resistance mutation is doing equally well when competing with a given 'wild type' in some environment without drug. It is not a direct measure of trade-off unless some precautions are taken (carefully chosen wild type, carefully chosen environment).

Moreover, given the direction of the authors' following
arguments, it would be helpful to follow up this statement with caveats or counter-examples, instead of ending this paragraph with the impression that evolution of reduced or eliminated costs is the norm.

We are not sure to understand this comment. We are not saying that reducing or eliminating the cost is the norm. We simply explain that we should not be surprised to find zero cost or negative costs.

* p. 5, 2nd paragraph: I don't understand what is meant by "an essentialization of mutation/genotypes".

We define the term at the beginning of the paper : "essentialize the properties of mutations or genotypes = that characterize mutational properties as intrinsic. This follows Oxford dictionary: "Characterize (a quality or trait) as fundamental or intrinsic to a particular type of person or thing."

* The section title "Costs of resistance are not pleiotropic effects" might be better worded as "Cost of resistance are not equivalent to pleiotropic effects".

We followed the suggestion

Similarly, the following section title
"Resistance mutations do not have a cost" could be misleading, and would be clearer if expanded to "Resistance mutations do not have a single, well-defined cost", or else modified to something like "'Cost of resistance' is poorly defined" or "'Cost of resistance' is a problematic/misleading term".

See answer to R1 on this issue. We added quote to "a cost" to indicate that we challenge the terminology.

* p. 7: In the discussion of compensation of pleiotropic effects ("amelioration") vs. cost evolution, it might be helpful to illustrate these two different processes on the figure, or at least refer back to the figure (e.g. amelioration of pleiotropic effects would correspond to moving from point R to point P_1

in Fig. 1, but this endpoint is still associated with a non-zero cost represented by the distance OP_1).

We followed the suggestion

* p. 8, 1st paragraph: “For instance, habitat quality varies and can even obscure the relationship between ‘absolute’ measures of fitness and environment variables…” This statement and its connection to the following E. coli example aren’t entirely clear. It might help to clarify that the finding that E. coli grows faster at temperatures slightly higher than 37 degrees was presumably obtained in lab conditions, where many other variables may also differ from the human host, which could result in the apparent non-optimality of the evolved wild type.

Yes, we added “in lab conditions”. Note that evolution is not supposed to optimize growth rate. This is not the place to discuss this further, but this is quite general topic that has been addressed many times in the literature.

* p. 8, 2nd paragraph: “Worse, this cost of resistance may not even actually be a ‘cost’.” This sentence is confusing and I think it would be better rephrased or left out, as the meaning is much clearer in the following sentence.

We rephrased it.

* p. 9: Is “phenotypic trade-off” a standard term? I am familiar with trade-offs in terms of fitness – while there is an implicit mapping from phenotype to fitness in this model, I’m not sure whether it is usual to refer to trade-offs directly on the phenotype level.

Fitness traded-offs have their origins in phenotypic trade-offs. This is for instance the case in one of the most well know “Y“ resource model, where resource allocated to trait A is at the expense of the resource allocated to trait B. This is purely phenotypic.

* Despite the authors’ emphasis earlier in the paper that mutations are not inherently beneficial or deleterious, but rather that their fitness effects depend on context, later on there are several instances where the authors refer only to a “beneficial mutation” without being clear about the corresponding environment. Specifically: on p. 10 (“the mutation R illustrated on Figure 2 would still be a beneficial mutation”);

We thought it was clear from context, but we added the information.

in the caption of Fig. 3 (“they are therefore both beneficial mutations”);

Here too

at the end of the 1st paragraph on p. 11 (“resistance mutations versus mere beneficial mutations”).

Here too

* p. 10: “This would be in general clearer and more insightful.” I find this sentence too vague. If it is meant as the introduction of a new, more useful definition of resistance, this should be made clearer and more prominent. This would be helpful later, e.g. on p. 12, when the authors continue to discuss “resistance mutations” and it is not always clear precisely how they are now being defined.

This is indeed saying that resistance mutations should be better defined this way. The following example in Fig 4 uses this definition (this is mentioned explicitly by saying that the wild-type is well-

adapted to the non-treated environment). In the conclusion, we use the term quite generally, as is used in the field, without referring to the more specific definition we suggest. The different conclusions apply to all cases.

* The authors should be careful to distinguish “wild-type” from “genotype that is optimal in a particular non-treated environment”. For example, in the last sentence on p. 10, the authors refer to both “a wild-type in O” and “a wild-type in B”.

Point O and B refer to positions in the phenotypic space. These positions can correspond to phenotypic optimum, or to the position of any phenotype. The sentences say “With such a definition, it is possible to distinguish mutation R and R’ on Figure 3 for instance. Both would be beneficial in both treated and non-treated environment, relative to a wild-type in O, but only R would be beneficial in the treated environment relative to a wild-type in B.” Here, we mention two possible positions of the wild-type phenotype: either in O or in B. This has nothing to do with the definition of the ‘wild-type’ (which we use as a synonym of ‘reference’ genotype/phenotype).

More accurately, there is only a single wild-type (at O, as previously used), whereas point B represents an optimum in a particular non-treated environment (e.g. that being tested in the lab, which may not represent the ancestral environment in which the wild-type evolved).

We are referring to a wild type that could be, hypothetically, in different experimental situations, either in O or B.

The same issue comes up in the captions to Figs. 3 & 4, and on p. 13
(“failing to measure costs relative to a well-adapted wild-type to the non-treated environment…”).

We agree that we don’t use the point O in Fig. 4. We let it there, for consistency with previous figure, but this is indeed perhaps confusing. We removed it since the legend mentions that we consider the wild type near B in this example.

* Beginning of p. 11: Here the authors seem to suggest that fitness is always measured relative to a wild-type. It is true that competition experiments are one common way of quantifying relative fitness in the lab. However, it is also common to measure absolute fitness of a given strain in isolation, as in dose-response curves where the “response” is e.g. the net rate of population growth or decline when exposed to the drug.

This is now mentioned in Box 1.

* Having debunked the concept of “cost of resistance”, what the authors mean when they nonetheless continue to refer to “cost” towards the end of the paper (pp. 12-13 and Figs. 3-4 captions) becomes blurry. It would be clearer to now avoid using the term “cost” all together, or else be careful to state precisely what is now meant by “cost”.

Throughout, the term cost refers to the fitness effect in one non-treated environment. We do not vary about this, and this is the conventional definition. We now state clearly whether the term should be avoided or not at the end.

* p. 12 minor wording clarification: “Studying ‘cost’ and ‘benefit’ at one particular dose may give the illusion…” (suggest adding the underlined part).

Added

* In Figs. 3 & 4, please illustrate “cost” in the plots if possible. Otherwise, the statements about cost do not belong in the figure captions.

We did not surcharge the figure, simply because this is the same idea than in the previous figures.

Citations:
* The reference list is actually quite extensive, including both theoretical and empirical references, and the authors’ perspective overall seems to be well-founded based on their broad reading. However, there are a several statements in the text that would be better justified by adding specific citations (at least as examples):
• section *Resistance mutations as beneficial mutations*: “classically, the fitness benefit of a resistance mutation…depends on the fraction of the population exposed to the drug…”, and later,

This is a trivial statement. If all resistance alleles survive, while all susceptible ones die upon treatment, then, if the fraction of the population exposed to the drug is f, the frequency p in the population after treatment is

$p' = p/ (1-f+f p)$ Eq (1)

hence, the change in frequency is

$p' – p = f p(1-p) + O(f)^2$ Eq (2)

which indicates clearly that the frequency change depends on the fraction of the population exposed (Eq 1), and that the selection coefficient favoring the resistance allele is equal to this fraction (as long as this fraction is O(1), Eq 2).

and if the
“This selective advantage…is often thought to represent an inherent property of the mutation itself.”

We did not want to single out few papers making this kind of claim, or at least implicitly using this simplification. There are so many papers doing this that it would be unfair to single them out. There is also no point in trying to make a review of the literature on this specific point and cite dozens of papers (this is not our topic). For instance, most papers in the field of molecular evolution are adopting this view, at least implicitly.

• bottom of p. 2 (near beginning of section *The context dependence of fitness effects*):
“Rather, it [the fitness effect of a mutation] depends on the ecological conditions, the genetic background, and on other alleles.”

Same answer as in the previous comment.

• middle of p. 4: “something well known in ecological genetics”

Again, doing a review here for this particular statement would be distracting. The field of ecological genetics was built on examples of local adaptation, where, obviously, the fitness effects of alleles differ among environments. There are so many papers that it would be unfair to single a few out. Bell’s book ‘Selection’ offers a nice recent reference; Ford’s 1971 book is historically important for the field of ‘ecological genetics’; Otherwise, original / historical references before the 90s would include

Haldane 1948;
Daday 1954;
Kettlewell and Berry 1961 ;
Barber 1965 ;
Jain et Bradshaw 1966 ;
Livingstone 1969 ;
Bishop 1972 ;
Koehn et al. 1980 ;
DiMichele and Powers 1982 ;
Kreitman 1983 ;
Simmons et al. 1989.

• top of p. 5: “resistance mutations are simply viewed [in recent interpretations] as pleiotropic”.
• 2nd paragraph of p. 7: “pleiotropic effects and the ‘cost of resistance’ are two different things… contrary to what is usually considered”. (It’s the part about what is usually considered that calls for citations.)

This is a very common practice to explain that the cost of resistance stems from the deleterious pleiotropic effects of resistance mutations. This is so widespread that we don’t want to either (unfairly) single out few studies or do an extensive and historical analysis across thousands of papers.

• end of 1st paragraph on p. 8: The *E. coli* example needs a citation. Subsequently, the claim “There are many other examples like this” should either be backed by citations, or simply cut.

We cut and added a reference for the E. coli example.