Dear editor,

We feel that the last version of our manuscript has not been clear enough, as there has been misunderstanding about several of our assumptions and models.

Specifically: we have not assumed the same s for all mutations, but rather *s\_beneficial* > *s\_deleterious*; our main results do depend on U, and *U* << *s* is assumed only for first order approximations; In the current version we explain these points better, as well as our mean fitness model (of course all individuals can receive both deleterious and beneficial mutations, see below) and the reason for the improved adaptation rate of SIMe in comparison with CM. We also included simulations in which 0.1<U/s<1.

Altogether, it seems to us that we have a good answer for each of the comments raised by the editor on the last round. Thus we wondered whether evolution journal would consider the revised manuscript before we submit it to a different journal. We would appreciate your opinion on that.

**Response to editor:**

This version of the ms is much clearer, which is very helpful and allows me to make a proper assessment. The main advance of this paper is to measure rates of adaptation with SIM vs constant mutagenesis. (They also show results regarding the evolution of SIM as I had requested but I also now remember that was done more thoroughly in their earlier paper, Ram & Hadany 2012). The question is whether the work on adaptation rates warrant publication in Evolution. On the one hand, the work is quite interesting. On the other hand, I am not sure they have explored this question deeply enough.

1. The assumptions the authors have made are not wholly unreasonable but nor are they entirely reasonable either.
   1. The analysis assumes *s* is the same for all mutations. Theoreticians often assume constant *s* for background selection but it is a more limiting assumption to claim that the beneficial effects have the same *s* too. If *s\_beneficial* > *s\_deleterious* (or *s\_specific* > *s\_non-specific*) then hitchhiking is possible (i.e., adaptation from a loaded genome).

Actually, we did assume *s\_beneficial* > *s\_deleterious:* in our model, *s\_beneficial*=*sH* and *s\_deleterious=s*, and we use *H=2*. Therefore, in our model indeed *s\_beneficial > s\_deleterious* by *3s*. We have shown that in the explored parameter space, hitchhiking is possible but has a minor effect on the adaptation process – see the difference between the simulation results (with hitchhiking) and the approximations (without hitchhiking) in Figure 1 and S1. We have also considered higher values of *U*; see Figure 3 in the revised ms.

* 1. However, the authors have also assumed *U* << *s*. As a consequence almost all individuals are unloaded (so they can assume things like *Ne-U/s* = *N*). The model has the appearance of being reasonably sophisticated by allowing for adaptation as well as background selection but in reality, the background selection effects are assumed away.

This is why all main results are independent of *U* (eq. 5-7). Essentially all of the analysis that includes *U* is more or less unnecessary because they assumed *U << s* so there are almost no loaded genotypes anyway.

There must have been confusion here. First, in our revised manuscript U does appear in eqs. 5-8, lines 205-6 (note that *q* in equation 5 refers to an expression including *U*). Second, we make the assumption that *U<<s* only for the purpose of making the first order approximations (*e-U/s ~ 1-U/s* etc.) to get a simple form for the equations, but the analytic model is valid whenever the MSB distribution of deleterious mutations is Poisson; this occurs as long as the expected number of individuals in the fittest class is at least one - *Ne-U/s > 1* or *U<s\*logN* - which is much more relaxed than *U<<s* (see Gessler 1995). Following the comment, we now clarify the assumptions required for each result in the text (line 125 in the revised ms). Finally, we also ran simulations in which *0.1<U/s<1* and the analytic approximation for CM is the only thing that breaks in this case: the difference between SIM and CM remains the same and the analytic approximation for SIM is very close to the simulation results. These results are shown in Figure 3 of the revised ms.

With these assumptions, the results boil down to a very simple answer. Under the assumption of the model (mutation-limited evolution), the rate of adaptation is proportional to the mutation rate. Mutational strategies that increase the mutation rate (CM, SIM) thus have higher rates of adaptation. In the case of constitutively higher mutation rates (CM) the rate is increased by the square of τ (the ratio of mutational increase) because it depends on the frequency of the single mutant at mutation-selection balance AND on the appearance of the epistatic second locus (eq. 7). For SIM it depends only on τ because the elevated mutation rate only applies to the latter step by way of assumption. (More generally the answer to eq. 8 would be υ*SIM* ≈ τ*ab* τ*Ab\_or\_aB* υ*NM*; in essence they assumed τ*ab* = 1).

Surely, there are a large fraction of mutations in *E. coli* with *s* << *U*. In flies, experimental data would suggest mean *s* is ~2% (similar to what is used here) but inferences from sequence analysis suggest that most mutations have a much, much weaker effect.

Our model deals with haploid asexual organisms (*i.e.* microbes). We therefore make no claim on how adaptation proceeds in sexual, multi-cellular organisms such as files. This is an interesting direction for future research.

In the Supplementary Material, we find results indicating that results do not hold too well when *U*τ gets large. However, no results are shown for *s ≤ U*.

We added simulation results for 0.1<*U/s*<1 and the only thing that doesn't hold is that the analytic approximation misses the simulation results (Figure 3 in revised ms). These results were not shown before because we limited our scope for asexual organisms in which *U<s* is reasonable.

1. The main text (ln 299) says that the advantage of SIMe for probability of fixation over CM was verified by simulations shown in Fig. S2. I cannot see how the data shown there support this claim.

The green lines, representing SIMe, are always higher than the red and blue lines representing CM and SIM, respectively. We now describe that in the Si. Both the approximations (dashed lines) and simulation results (solid lines representing logistic regression on the results) support the advantage of SIMe.

1. Moreover, while the probability of fixation may (or may not) be correct, I do not think the rate of adaptation is correct (eq.10 & ln 630). CM would have more single mutants at MSB just prior to environmental change (*N*µτ*/s* compared to *N*µ*/s* for CMe and SIMe). Thus, there would be more individuals on which double mutants could arise. This effect would far outweigh the difference in probability of fixation.

The simulations explicitly allowed for the population to achieve MSB before allowing for adaptation. If standing variation was significant then the simulation results would show that CM adapts faster; however, the results show that SIMe is as fast as CM if not faster. The reason for that is that the waiting time for appearance of a double mutant is relatively long (>1000 generations even with the highest *τ* , see Fig. S1). Thus SIMe population quickly approaches a new MSB after the environmental change (For example, with *s=0.*05 and *U=0.*0004, the average number of deleterious mutations is *0.99\*U/s* after 90 generations - see Gordo and Dionisio 2005 for details on how to compute this). This means that standing variation is quickly generated in SIMe populations and doesn't give CM populations much of an advantage. We include that explanation in the revised ms (line 318).

1. Mean fitness model seems wrong. Individuals with less than X deleterious alleles can only receive deleterious mutations and those with more than X can only get beneficial ones. This probably makes little difference if beneficials are rare but it still is not strictly correct.

Incorrect. In our mean fitness model (SI), all individuals can receive both deleterious and beneficial mutations, except for mutation-free individuals (see definition of *mx,y* in line 64-70 of the SI). This is similar to the model described in Ram and Hadany 2012, which you already know.

Summary: After thinking about this a fair bit, I feel the assumptions highlighted in (1) simplify the model to an extent that the results, while interesting, are not enough to warrant an Evolution paper, given the large and increasing number of submissions. The authors have considered a relevant but too narrow part of parameter space. At the very least more simulations need to be done, allowing them to thoughtfully consider the effects of background selection and genetic hitchhiking. I am sorry I did not point these issues out on the original submission but I (and the reviewers) did not understand what was being done.