January 18, 2024

**Response to Reviewers**

*Manuscript MBE-23-1021*

Editorial Office

Molecular Biology and Evolution

Dear Editors,

Attached please find our revised manuscript, “*Aneuploidy can be an evolutionary diversion on the path to adaptation*”. We have made changes to the manuscript in response to the reviewers’ comments, which we hope you will find satisfactory. Details of the changes appear below, with our responses in blue. When quoting text from the paper, we use underline for insertions and strikethrough for deletions. In addition, we attach a version of the revised manuscript with tracked changes.

We have also added two co-authors who performed preliminary sequence analysis and were meant to be included in the original submission: Anna Selmecki and Pétra Vande Zande (University of Minnesota Medical School).

We are grateful to the editors and reviewers for their analysis and comments.

Sincerely,

Yoav Ram, PhD

Senior Lecturer

School of Zoology

Faculty of Life Sciences

Tel Aviv University

**Editor's comment:**

While the title "Aneuploidy can be an evolutionary detour on the path to adaptation" is clever, the work detour has two meanings- it could be a distraction or dead end as your work suggests or it could be a necessary side road that ultimately is required to reach the final goal. I suggest a simpler, although less clever title: Aneuploidy can impede rather than spur adaptation.

Thank you for noticing the double meaning of detour! We missed that it can also have a positive meaning like “bypass”. We appreciate your suggestions, but we decided to keep the original title structure while replacing “detour” with “diversion”, which we think only means “distraction or dead end” and not “necessary side road”. If you disagree, we are willing to consider the title you suggested.  
  
**Reviewer 1**  
As the equations I provided in the comments might not be displaying correctly, I have included my comments in the attached file for your reference.  
  
In a previous study, Yona et al. conducted an evolutionary experiment by evolving diploid yeast cells (2n) under high-temperature conditions (39 °C). Remarkably, over 95% of the evolved cells displayed aneuploidy (2n+1) by generation 450, with the "+1" indicating aneuploidy. Subsequently, in two distinct lineages, the majority of the evolved population transitioned to euploidy at generations 1700 and 2350, denoted as 2n\* where "\*" represents a beneficial mutation. Using mathematical modeling, the authors simulated the evolutionary dynamics, utilizing limited data such as fixation times and estimated fitness of genotypes as prior distributions, derived from the Yona’s group.  
  
The model was employed to test whether the evolved euploidy (2n\*) originated from aneuploidy (2n+1) or the ancestral cell (2n). The simulation outcomes suggested that the evolved euploidy (2n\*) descended from the ancestral cell (2n). This methodology empowers researchers to explore evolutionary dynamics, assess hypotheses, and formulate new hypotheses for designing experiments based on previous evolutionary trials. Notably, this approach is applicable not only to ongoing evolution experiments but also to those conducted in the past. It is particularly valuable for experiments predating the advent of next-generation sequencing, as it allows the investigation of evolution dynamics without relying on sequencing data.

We are happy that you find our approach valuable!  
  
However, there are certain concerns regarding the modeling results. These concerns stem from (1) the limited information available from the earlier study, which may not provide precise details for thorough analysis, and (2) potential issues related to the equations employed in the modeling process. Further details on these concerns are provided below.  
  
Major points  
1. Based on the literature, the rates of chromosome gain and loss exhibit an approximately tenfold difference, and the impact of heat stress on the transition between aneuploidy and euploidy remains unclear. The current model assumes equal rates for both gaining and losing a chromosome. If one were to consider disparate rates for chromosome gain and loss, would this variance affect the modeling results? Additionally, if the authors were to incorporate the assumption that the mutation rate increases in aneuploidy, would the modeling outcomes continue to align with the present findings?

We have tested a model that assumed increased rate of chromosome loss and it did not perform well, maybe because of the small amount of data we have for model fitting. We mention this in the main text (line 277): “A model that assumed increased rate of chromosome loss in aneuploid cells (as in Sheltzer et al.56) did not perform well, probably due to lack of statistical power, and was abandoned.” Additionally, as mentioned in the original manuscript, we could not find evidence of an increase in mutation rate in aneuploid cells (line 185).

We now consider the effects of disparate rates for gain and loss of chromosomes and an increased mutation rate in aneuploid cells. We summarize these effect in a new paragraph (line XXX) and supporting figure, Figure S11:

“Nevertheless, increasing the rate of chromosome loss (transitions from 2𝑛+1∗ to 2𝑛∗) without increasing the rate of chromosome gain (transitions from 2𝑛 to 2𝑛 + 1) increases 𝐹𝐴 (Figure S11B), but not to the same extent as increasing the rate of chromosome gain. In contrast, increasing the mutation rate in aneuploid cells can have a marked effect on the dynamics: when using the MAP parameter estimates, 𝐹𝐴 increases to 0.52 when the mutation rate in aneuploid cells increases 10-fold (Figure S11C).”



**Figure S11: Effect of genomic instability on genotype frequencies.** Genotype frequencies in the deterministic model without drift and with MAP parameter estimates (**A**), with 100-fold increase in rate of chromosome loss (transition from 2𝑛 + 1∗ to 2𝑛∗) (**B**), and with 10-fold increase in mutation rate in aneuploid cells (**C**). Corresponding 𝐹𝐴 values are 0.098, 0.223, and 0.519.

2. Yona's research team conducted evolution in a rich medium at 39°C for up to 2350 generations. During this process, they observed aneuploidy fixation (2n+1) at generation 450, followed by the reversion to euploidy (2n\*) between generations 1700 and 2350. Simultaneously, the group took the aneuploidy population from generation 450 and subjected it to further evolution in a minimal medium at 39°C for an additional 1000 generations. After 1000 generations in the minimal medium at 39°C, the aneuploidy population also transitioned to euploidy (2n\*, minimal medium). The authors derived prior distributions using fitness results between 2n\* and 2n+1 from Fig4A in Yona's work. However, it's noteworthy that the 2n\* in Fig4A does not originate from the rich medium at 39°C after generations 1700-2350 but rather from the minimal medium at 39°C after another 1000 generations. The timing of the cell's transition to euploidy in the minimal medium at 39°C is unclear, and the similarity in fitness between 2n\* in two different evolutionary environments is uncertain. Could these factors influence the model's performance? To address this, can the model be adapted to simulate the dynamics, considering the switch in evolutionary conditions?

The growth curves of 2n\* that evolved in minimal media (“refined” from Yona et al Fig. 4A) were only used in our *extended informative prior distribution*, which proved to be less useful (line 390), the full details are in the supplementary material, under the heading “extended informative prior distribution” (line 605).

We have thought about adapting the model to consider a switch in the evolutionary conditions, but this means extending the model with additional fitness parameters for growth in minimal media, and moreover, we do not have data about the timing of appearance/fixation of 2n\* in minimal media (as you have mentioned). Therefore, we cannot estimate the extra parameters. Note that the aneuploidy and mutation rates might also be affected by the difference in media.

3. Within the model, two parameters, represented by "b" and "c," denote the benefit and cost of

aneuploidy, respectively. In line 128, the equation specifies that " c=𝑤2n\* - 𝑤2n+1\*=0.003" However,

it seems that the cost should be a negative value, suggesting the equation should be adjusted to "

c= 𝑤2n+1\* -𝑤2n\* =-0.003" Consequently, the benefit of trisomy (b) can be calculated as "1.022(𝑤2n+1 )

- 1(𝑤2n) – (-0.03)(𝑐)=0.025." If my interpretation is correct, the model might be underestimating the

benefit of trisomy. It is suggested to rectify this equation in the model and rerun it to observe how this adjustment influences the results.

The model uses the parameters wi and therefore these are the values we inferred. The parameters s, b, and c were only presented in the text to facilitate interpretation of the model results. Therefore, there is no need to rectify the model and rerun it.

However, there was indeed an error in our presentation of the values for b, c, and s. After further considerations, we find that these interpretations introduce further assumptions (e.g., b is additive but 1-c is multiplicative) and therefore we decided to remove all mentions of b, c, and s from the text.

4. In line 286, there appears to be a potential typo. The term δf\_(2n+1)\_s  in the last equation in equation (2) f\_(2n\*)\_m=μf\_2n\_s+δf\_(2n+1)\_s+f\_(2n\*)\_s may need to be corrected to δf\_(2n+1\*)\_s. Kindly review the model to confirm whether this is the accurate equation employed in the modeling process.

Thank you for noticing this typo, we have corrected it in the text and confirmed that the correct equation is indeed employed in the model implementation.

5. The authors made an effort to incorporate the notion that mutation rates are higher in aneuploidy into their model. I commend the authors for considering this aspect, as aneuploidy has the potential not only to alter mutation rates but also to influence the type of mutations in the second mutation. This phenomenon aligns with the climbing Mount Probable theory, wherein the prior genotype affects the likelihood of the next mutation. Given that aneuploidy might enhance the probability of beneficial mutations and that ancestor cells might require at least two mutations to transition to the 2n\* stage (with the first mutation facilitating the occurrence of the second mutation), would the simulation results still substantiate the hypothesis that 2n\* originates from the ancestor 2n rather than from aneuploidy, if the model considers the transition from ancestor (2n) to 2n\* as a two-step mutation process and assigns a higher likelihood to aneuploidy reaching the 2n\* stage?

That is an interesting question. We have tested a more complex model in which more than a single mutation is required for adaptation. However, given the limited available data on timing of fixation and loss of aneuploidy, the model failed to fit to the data. We hope to explore such dynamics in future work with more extensive data.  
  
Minor point  
1. Beginning at line 103, the authors elucidated the significance of the estimated beneficial mutation rate in the model, drawing upon mutation rates from existing literature and supporting evidence in this paragraph. However, it is challenging to understand when the units of beneficial mutation rate and mutation rate differ. Aligning the units would facilitate a more straightforward illustration of the frequency distinctions between random and beneficial mutations. Please rephrase to enhance clarity.

We have revised the relevant section and hope it is clearer now (line 104): “The estimated beneficial mutation rate is 𝜇 = 2.965 · 10-6 [2.718 · 10-7-3.589 · 10-6] per genome per generation (that is, roughly 3 out of 106 cell divisions produce a mutant cell with a fitness advantage). From the literature, the mutation rate per base pair is roughly 2 − 3 · 10-10 (refs. 74, 35), but it may be higher under heat stress, as several stresses20, including heat22, may cause hypermutation in yeast. If we assume a 10-fold increase over the mutation rate reported in the literature, then the estimated beneficial mutation rate can be explained by a genomic target size of 1,000 base pairs (that is, 1,000 base pairs across the genome in which a mutation would provide a fitness advantage): 3 · 10-10 × 10 × 1, 000 = 3 · 10-6.”

2. In line 110, providing a direct description of the total tested Hsp90 variants in the reference would aid in comprehending the ratio difference between beneficial variants and the overall variants tested.

Flynn et al estimated the selection coefficient of 14,160 Hsp90 variants in 37°C. Out of these, Flynn et al determined 465 variants to be beneficial by comparison to synonymous mutations under the same conditions. The list of these beneficial mutants is in Fig. 5 of their manuscript. We revised the text to say (line 113): “Flynn et al.14 used a deep mutational scan of a single protein, Hsp90, to find 465 amino-acid variants (out of 14,160) that significantly increased growth rate in 37◦C.”

3. In line 297, the authors indicated a re-analysis of results from Yona’s group in the supplement. Please include pertinent details from this reanalysis, such as information on the generation-ploidy relationship in the reanalyzed data and what type of data is used for analysis.

We did not perform any re-analysis, sorry for the confusing phrasing. We have now replaced this text with a more detailed explanation: “In the original analysis of Yona et al.70, samples were routinely extracted from the evolving populations and tested for heat-shock tolerance. The first generation in which such indication was found was generation 200. Therefore, we determine that aneuploidy did not reach high frequency before generation 200.”

4. In lines, 606 and 620, the referred Figures S3L and S3H do not exist.

We now corrected this typo: the reference to Fig. S3L was supposed to be to Fig. S3H. The latter exists (bottom right panel in the figure) and shows a distribution of fitness values of 2n\* relative to 2n+1 inferred using *Curveball*.   
  
  
**Reviewer 2**  
Aneuploidy is an important topic in evolutionary genetics and is sometimes claimed to be a stepping stone for adaptation, despite its drawbacks. In this study, the authors used modeling to investigate this idea using published data on yeast evolving to high temperatures. They find that aneuploidy is not necessarily an intermediate state, but may be more immediately evident than other types of mutation owing to the fact that rates of chromosomal non-disjunction exceed rates of other types of beneficial mutation.  
  
In my opinion, this investigation provides useful insight into how population genetic forces can give rise to evolutionary outcomes in ways that can be counter-intuitive. I particularly appreciate that the authors have investigated an evolutionary scenario by taking into account the relative rates of different types of mutation, and not assuming that selection is the main driver of evolutionary dynamics. In this instance, the relative rates of point mutation versus non-disjunction appears to be critical, which is a genetic phenomenon that ought to be more widely considered.

Thank you! We are glad you found our study insightful.  
  
I think there are aspects of the manuscript that could be clarified or expanded.  
  
Areas for clarification:  
  
In previewing their results in the Introduction section, the authors conclude that euploid cells were not derived from aneuploid cells. The start of the Results section tells us that aneuploidy first became fixed, which seems contradictory (i.e., how could euploids not be derived from aneuploids if there were only aneuploids?). The solution is that the authors use the term “fixed” to refer to a frequency of >95% (line 93, line 296). Since this is a central element of the paper I recommend earlier clarification. (From a theoretical standpoint I usually consider “fixed” to mean an allele frequency of 1, with perhaps some very small frequency of mutant alleles, but I understand that this isn’t always the definition applied in empirical contexts, where allele frequency is measured with uncertainty).

You are right. Our definition follows that of Yona et al. (see their supplementary materials): “In all cases where we report aneuploidy, the estimated part of the population that is reported to either gain or eliminate a chromosome is at least 95%”. We have made several changes to the text to be clearer as to our use of the term ‘fixation’, either by expanding on it, or replacing it with another term:

* Line 62: Aneuploidy was therefore suggested to be a transient adaptive solution, because it can rapidly appear and ~~fixate in~~ take over the population
* Line 79: “Aneuploidy ~~fixed~~ reached high frequency (>95%) in all four experimental repetitions in the first 450 80 generations.”
* Line 94: “We fitted this model to the experimental results – time for fixation (frequency >95%) and for loss (frequency <5%) of aneuploidy”
* Line 299: “Aneuploidy fixed (frequency >95%) in all four population in the first 450 generations.”
* Line 316: “taking simulations in which 2n+1 fixed (reached >95%) before generation 450 but not before generation 200, and computing the fraction of such simulations 318 in which 2n\* did not fix by generation 1,700, and hence aneuploidy did not extinct (reach <5%) before generation 1,700.

The authors assume that the rate of chromosome gain (2n to 2n+1) is the same as the reverse rate. This is probably unavoidable, given the state of our knowledge around these basic rates, but we might reasonably expect trisomy to revert at a faster rate than it appears in the first place (at the very least because there are more chromosomes that can potentially mis-segregate). On line 275 the authors write that “A model that assumed increased aneuploidy rates in aneuploid cells… did not perform well,” but it’s not clear if this means additional aneuploidy or rather the reversion from aneuploidy to euploidy. The transition 2n+1 to 2n is not aneuploidy per se, but the reversion of it.

We are sorry for the confusion, we indeed meant “reversion from aneuploidy to euploidy” and not “additional aneuploidy”. We have now revised the text to (line 277) “A model that assumed increased rate of chromosome loss in aneuploid cells” to avoid this confusion.

It might be clearer to use “aneuploid” to refer to the karyotype, and “non-disjunction” to refer to chromosomal mis-segregation events. I would also encourage the authors to highlight the general scarcity of empirical data non-disjunction rates, which would allow for more detailed modeling.

We feel that the term “non-disjunction” is a bit technical and would prefer to stay with mis-segregation or aneuploidy; we think that the reader understand what we mean. We have added the word non-disjunction in line 97 when first referring to the rate of aneuploidy: “… inferring the model parameters: rates of aneuploidy (i.e., mis-segregation, non-disjunction) and mutation…” and similarly in line 122.

We now mention in line 244: “We estimate that the aneuploidy rate (i.e., number of chromosome gains per generation) is 1.7 · 10-3, higher than a previous estimate of 6.7 · 10-6 (ref 73). This may be due to genetic instability caused by heat stress5, but we note that there is a general scarcity of empirical data on aneuploidy rates.“

Line 375. I recommend more detail about fitness definitions in terms of s, c, and b.

Following a comment from reviewer 1, we decided to remove all mentions of s, c, and b. these parameters only served as interpretations for our estimated fitness values, and we found they are confusing and require additional assumptions.

It might be worth acknowledging that non-disjunction is arguably a form of mutation, but that in the context of this work “mutation” refers to point mutations or other events not relating to copy number variants.

We added a comment to the Results section that describes the evolutionary model (line 92): “Note that ‘mutation’ here refers to point mutations and other genetic variants unrelated to aneuploidy.”  
  
Areas for expanded discussion:  
  
Whether aneuploidy can serve as a true “stepping stone” might depend on the types of adaptive mutations that would ideally be acquired. In haploid yeast, a second chromosome copy could facilitate genomic rearrangements and gene duplication on that chromosome in particular, i.e., creating new types of available mutation rather than just potentially affecting the genome-wide mutation rate. I would have liked to see a bit more discussion on these more mechanical considerations.

As the experiments in question took place in diploid yeast, it’s possible that dominance plays a role. For recessive beneficial mutations, more time would be required for a mutation to first arise and then to become homozygous (spontaneous loss of heterozygosity is prevalent in yeast, with its own rate). The authors could suggest that future models consider this; since much of the data on these subjects come from yeast, in which experiments are performed in both haploids and diploids, it seems to me important to keep track of how our predictions might change with ploidy.

Another complication might arise is if key beneficial point mutations are on chrIII, i.e., the same chromosome in which aneuploidy is beneficial. This would mean that the loss of trisomy in 2n+1\* cells has a 1/3 chance of representing a change to 2n, rather than 2n\*. While chrIII is a small part of the genome, the fact that an extra copy of this chromosome is beneficial suggests there may be outsized potential for key adaptive point mutations here.

This is a good point. However, we neglect transitions to less fit genotypes (such as 2n+1\*->2n) because they are not expected to affect the evolutionary dynamics. Therefore, assuming that loss of chromosome also leads to loss of the beneficial mutation will effectively reduce the rate of chromosome loss by 33%. This rate is already very high: the 2n+1\* never increases in frequency as it is rapidly replaced by 2n\* (see gray and purple lines in Fig. S9). Thus, a 33% decrease would not change the dynamics.  
  
The authors discuss a possible increase in the point mutation rate in aneuploid cells (line 168), which has been found previously by Sheltzer et al. The authors’ discussion of heat stress as a factor makes sense, but it’s worth noting that Sheltzer studied disomic haploids, whereas this study concerns trisomic diploids. These may be quite different genomic contexts, e.g., in terms of gene dosage imbalance. Mutation accumulation experiments in which trisomy appears spontaneously have not detected evidence of a subsequent impact on the point mutation rate, as far as I know.