Chromosomal duplication can be an evolutionary detour on the path to adaptation

Ilia Kohanovski[1,2,*], Martin Pontz[1,*], Avihu H. Yona[3], and Yoav Ram[1,†]

[1]School of Zoology, Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel

[2]School of Computer Science, Reichman University, Herzliya, Israel

[3]Institute of Biochemistry, Food Science and Nutrition, Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Israel

[*]These authors contributed equally to this work

[†]Corresponding author: yoav@yoavram.com

September 20, 2022

Classifications. Biological Sciences: Evolution, Genetics, Microbiology, Population Biology.

Keywords: aneuploidy, evolutionary model, adaptive evolution

Abstract

Aneuploidy is common in eukaryotes, often leading to decreased growth and fitness. However, evidence from yeast and fungi, as well as human tumour cells, suggests that aneuploidy can be beneficial under stressful conditions and facilitate adaptation. In a prominent example, an evolutionary experiment with yeast, populations evolving under heat stress had become aneuploid, only to later revert back to euploid after genetic mutations have accumulated. It has therefore been suggested that aneuploidy serves as a "stepping stone" on the path to adaptation. Here, we test this hypothesis. First, we apply DNA sequencing to show that mutant alleles common in aneuploid cells are uncommon in the evolved euploid population. Second, we develop an evolutionary model with both aneuploidy and mutation, and fit it to the results of the experiment using a Bayesian inference framework. We then predict the genotype frequency dynamics during the experiment, demonstrating that the majority of the evolved euploid population likely did not descend from aneuploid cells, but rather directly from the euploid wild-type population. Together, our results suggest that aneuploidy is an evolutionary "detour" rather than a "stepping stone": it delays rather than facilitates the adaptation of the population, and cells that become aneuploid leave less descendants compared to cells that remain diploid.

Introduction

Aneuploidy is an imbalance in the number of chromosomes in the cell: an incorrect karyotype. Evidence suggests aneuploidy is very common in eukaryotes, e.g. animals[? ? ?], and fungi[? ? ? ?]. Aneuploidy has been implicated in cancer formation, progression, and resistance[? ? ? ?]. It is also common in protozoan pathogens of the Leishmania genus, a major global health concern[?], and contributes to the emergence of drug resistance[?] and virulence[?] in fungal pathogens, which are under-studied[?], despite infecting a billion people per year, causing significant morbidity in >150 million and death in >1.5 million people per year[? ?].

Experiments with human and mouse embryos found that aneuploidy is usually lethal. It is also associated with developmental defects and lethality in other multicellular organisms[?]. For example, aneuploid mouse embryonic cells grow slower than euploid cells[?]. Similarly, in unicellular eukaryotes growing in benign conditions, aneuploidy usually leads to slower growth and decreased overall fitness[? ? ? ? ? ?], in part due to proteotoxic stress caused by increased expression in aneuploid cells[? ? ?] and hypo-osmotic-like stress[?].

However, aneuploidy can be beneficial under stressful conditions due to the wide range of phenotypes it can produce, some of which are advantageous[? ?]. Thus, aneuploidy can lead to rapid adaptation in unicellular eukaryotes[? ? ? ?], as well as to rapid growth of somatic tumour cells[? ?]. For example, aneuploidy in *Saccharomyces cerevisiae* facilitates adaptation to a variety of stressful conditions like heat and pH[?], copper[? ?], salt[?], and nutrient limitation[? ? ?], with similar results in *Candida albicans*[?]. Importantly, aneuploidy can also lead to drug resistance in pathogenic fungi such as *C. albicans*[? ? ?] and *Cryptococcus neoformans*[?], which cause candidiasis and meningoencephalitis, respectively.

? demonstrated experimentally the importance of aneuploidy in adaptive evolution. They evolved populations of *S. cerevisiae* under strong heat stress. The populations adapted to the heat stress within 450 generations, and this adaptation was determined to be due a duplication of chromosome III. Later on, after more than 1,500 generations, the populations reverted back to an euploid state, while remaining adapted to the heat stress. Aneuploidy was therefore suggested to be a *transient adaptive solution*, because it can rapidly appear and fixate in the population under stressful conditions, and can then be rapidly lost when the cost of aneuploidy outweighs its benefit—after the stress is removed, or after "refined" beneficial mutations appear and fixate[?]. Furthermore, it has been suggested that aneuploidy is an evolutionary "stepping stone" that facilitates future adaptation by genetic mutations, which require more time to evolve[? ?].

Here, we test the hypothesis that aneuploidy is a an *evolutionary stepping stone* that facilitates adaptive evolution by genetic mutations. First, we analyze previously unpublished sequencing data from the original experimental populations of **?** to assess if the evolved euploid population is descended from the aneuploid population. Second, we develop an evolutionary genetic model and fit it to the experimental results of **?** in order to predict the genotype frequency dynamics in the experimental populations, thereby estimating the frequency of evolved euploid cells that descended from aneuploid cells. Our results show that aneuploidy reached high frequencies in the experimental populations, but nevertheless, the majority of cells in the evolved euploid populations likely did not descend from aneuploid cells, but rather directly from wild-type euploid cells. These results suggest that aneuploidy can be an evolutionary detour, rather than a stepping stone, on the path to adaptation.

Results

In the heat-stress experiment of **?**, four populations of *S. cerevisiae* evolved under 39 °C. Aneuploidy fixed in all four experimental repetitions in the first 450 generations. Two of the repetitions, marked $H2$ and $H4$, carried no large-scale duplications other than a chromosome III trisomy. These two repetitions continued to evolve under the same conditions, wherein aneuploidy was eliminated by generation 1,700 and 2,350.

Empirical frequencies of mutant alleles. For each of two evolved populations ($H2$ and $H4$) we sequenced the ancestral population (generation 0), the aneuploid population (generation 450), and the evolved euploid population (generation 1,700 or 2,350) to estimate the mutant allele frequencies (Tables S1 and S2). Overall, between 100 and 173 mutant alleles were detected with at least a single read in the six populations that were sampled. Disregarding 45 and 40 alleles that were present in the ancestral populations at a frequency >10%, the aneuploid and euploid populations carried a large number of mutant alleles: 82 and 95, respectively, in repetition $H2$, and 60 and 66 in repetition $H4$.

Surprisingly, out of all these mutant alleles, none was present at a frequency >20% in both the aneuploid and the evolved euploid populations. Furthermore, a high mutant allele frequency in the aneuploid population implies a low frequency in the evolved euploid population, and vice-versa (Spearman's correlation coefficient $\rho = -0.64$ and -0.66 in the two experimental repetitions; Figure 1), such that mutant alleles frequent in the aneuploid populations decreased in frequency when aneuploidy was lost. Moreover, for the 18 mutant alleles with high frequency in the aneuploid populations (>20%), the highest frequencies in the euploid populations were 15.4%, 16%, 16.3% and 19.6% (the rest were below

15%). Similarly, for the 48 mutant alleles with high frequency in the evolved euploid populations, the highest frequencies in the aneuploid populations were 2.7%, 7.7%, and 11.1% (the rest were below 1%). These results suggest evolved euploid cells are unlikely to descend from aneuploid cells.

Evolutionary genetic model. To explore the dynamics during the evolutionary experiments, we developed an evolutionary genetic model, fitted the model to empirical data, and used it to predict the genotype frequency dynamics, or specifically, the fraction of the evolved euploid population descended from aneuploid cells.

The model includes the effects of natural selection, genetic drift, aneuploidy, and mutation, and follows a population of cells characterized by both their genotype: euploid wild-type, *2n*, is the ancestral diploid genotype; euploid mutant, *2n**, has a diploid karyotype and a single beneficial mutation; aneuploid wild-type, *2n+1*, has an extra chromosome due to a chromosome duplication event; and aneuploid mutant, *2n+1**, has and extra chromosome and a beneficial mutation. Fitness values of the different genotypes are denoted by w_{2n}, w_{2n*}, w_{2n+1}, and w_{2n+1*}, and the rate of mutation and aneuploidy are denoted by μ and δ. See Figure 2 for an illustration of the model.

We fitted this model to the experimental results**?** – time for fixation (>95%) and for loss (<5%) of aneuploidy – using approximate Bayesian computation with sequential Monte Carol (ABC-SMC)**?** , thereby inferring the model parameters: rates aneuploidy and mutation and the fitness of all genotypes. We then sampled posterior predictions for the genotype frequency dynamics using the estimated parameter values and compared different versions of the model to test additional hypotheses about the evolutionary process.

Estimated rates and fitness effects of aneuploidy and mutation. We inferred the posterior distribution of model parameters (Figure 3). We report parameter estimates using the MAP (maximum a posteriori) and providing the 50% HDI (highest density interval) in square brackets. See Supplementary Material for sensitivity analysis.

The estimated mutation rate, $\mu = 2.965 \cdot 10^{-6}$ $[2.718 \cdot 10^{-7} - 3.589 \cdot 10^{-6}]$, corresponds to a mutation target size of $\sim 10^4$, assuming the mutation rate per base pair is roughly $2 \cdot 10^{-10}$ (ref.**?**) or $3.3 \cdot 10^{-10}$ (ref.**?**). The estimated aneuploidy rate, $\delta = 1.72 \cdot 10^{-3}$ $[1.47 \cdot 10^{-3} - 2.786 \cdot 10^{-3}]$ is higher than in previous studies: for chromosome III in diploid *S. cerevisiae*, **?** estimated $6.7 \cdot 10^{-6}$ chromosome gain events per generation, and **?** estimate $3.0 - 4.3 \cdot 10^{-5}$ chromosome loss events per generation (95% confidence interval). The estimated fitness values are $w_{2n+1} = 1.022$ $[1.021 - 1.023]$, $w_{2n+1*} = 1.025$ $[1.024 - 1.026]$, $w_{2n*} = 1.028$ $[1.026 - 1.029]$, all relative to the fitness of *2n*, which

is set to $w_{2n} = 1$. Thus, we can infer that the cost of trisomy is $c = w_{2n*} - w_{2n+1*} = 0.003$ (or 0.3%) and the benefit of trisomy is $w_{2n+1} - 1 - c = 0.019$ (1.9%), whereas the benefit of the beneficial mutation is $w_{2n*} - 1 = 0.028$ (2.8%).

If we allow for transitions (mutation, chromosome loss and gain) to less-fit genotypes (e.g., *2n** to *2n+1**), then we infer similar but slightly different values, see Supplementary Material.

Model comparison and goodness-of-fit. Our model fits the data well: in simulations using the MAP parameter estimates, *2n** fixed in 61% of simulations by generation 1,700 and in 100% of simulations by generation 2,350 (Figure 4B).

However, a model without aneuploidy (where the aneuploidy rate is fixed at zero, $\delta = 0$), fails to explain the experimental observations (Figure 4). The estimated mutation rate without aneuploidy is $\mu = 7.98 \cdot 10^{-9}$ $[7.906 \cdot 10^{-9} - 8.138 \cdot 10^{-9}]$, much lower compared to a model with aneuploidy and suggesting a target size of just 40. The fitness of the mutant is also much lower at $w_{2n*} = 1.013$ $[1.012 - 1.013]$. This is because, without aneuploidy, a high mutation rate or fitness effect will lead to faster appearance and fixation of *2n** than in the experimental observations.

We also checked a model in which aneuploidy occurs but is adaptively neutral compared to the wild-type, that is, $w_{2n+1} = w_{2n}$ and $w_{2n+1*} = w_{2n*}$ but $\delta > 0$. This model fits the data better than the model with no aneuploidy (in which $\delta = 0$), but worse than a model with positive selection for aneuploidy, in which $w_{2n} < w_{2n+1} < w_{2n+1*} < w_{2n*}$ (Figure 4).

Model predictions of genotype frequency dynamics. We simulated 50 replicate genotype frequency dynamics using the MAP estimate parameters. Figure 5A shows the simulated frequencies of the four genotypes (*2n*, *2n+1*, *2n+1** and *2n**), as well as the frequencies of *2n** cells that arose from either *2n+1* cells via a sequences of mutation and chromosome loss events ($2n^*_A$), or directly from *2n* cells via a mutation event ($2n^*_M$). We find that *2n+1** never reaches substantial frequency as it is quickly replaced by *2n** in a process similar to *stochastic tunneling*[??] .

To test the hypothesis that aneuploidy facilitates adaptation, we estimated F_A, the expected frequency of *2n** that arose from *2n+1*, computed as the average frequency of such $2n^*_A$ cells at the end of simulations using the MAP estimate parameters. Surprisingly, we observe that the majority of *2n** cells are $2n^*_M$, a product of a direct mutation in *2n* cells, rather than descending from *2n+1* cells ($F_A^{MAP} = 0.106$, Figure 5A). This is despite the fact that the *2n+1* genotype reaches high frequencies in the population (at least 0.98, Figure 5A).

This result is not unique to the MAP parameter estimate. We simulated genotype frequency dynamics using parameter samples from the posterior distribution (Figure 3), and computed the posterior distribution of F_A (Figure 5B). The mean F_A was just 0.1673 [0.0154-0.370 95% CI] and only in 489 of 100,000 posterior samples (0.489%) F_A was larger than 0.5 (see Supporting Material for results when transitions to less-fit genotypes are allowed, such as *2n** to *2n+1**). Thus, if we sample a random cell from the evolved *2n** population, it is more likely to have descended directly from an euploid cell than from an aneuploid cell. The probability of *2n** descending from *2n+1* (F_A) increases with the aneuploidy rate, δ, and decreases with the mutation rate, μ, and population size N (Figure 5C,D). In some cases it can also be affected by the fitness parameters (Figure S10).

Genetic instability in aneuploid cells. It has been suggested that aneuploidy increases genetic instability[? ?] . Therefore, we inferred model parameters under the assumption that the mutation rate increases in aneuploid cells by a factor τ = 1, 33/32 (due to an additional chromosome), 2, 5, 10, or 100 (due to genetic instability). We found that the posterior distribution was similar for τ =1, 33/32, 2, and 5 (Figure S4). With τ = 100, the estimated mutation rate was about 7-8-fold lower compared to τ = 1 ($\mu = 4.094 \cdot 10^{-7}$ [$6.252 \cdot 10^{-8} - 6.046 \cdot 10^{-7}$]) and the aneuploidy rate was about 2-3-fold lower ($\delta = 0.744 \cdot 10^{-3}$ [$0.506 \cdot 10^{-3} - 1.827 \cdot 10^{-3}$]). With τ = 10, the estimated mutation rate was only slightly lower compared to τ = 1 ($\mu = 1.67 \cdot 10^{-6}$ [$2.836 \cdot 10^{-8} - 2.245 \cdot 10^{-6}$]). WAIC (lower is better, see Methods) is lowest for $\tau = 33/32$ and τ = 1 (Table S3). Therefore, evidence does not support an increase in mutation rate in aneuploid cells, and moreover, unless the increase is strong ($\tau \geq 10$), it does not seem to affect our inference. We also checked the differences in genotype frequency dynamics for different τ values. We observe τ = 100 could be distinguished if accurate data was available for the waiting time until the frequency of *2n* to decrease below 95% (Figure S5A) or for waiting time for the frequency of *2n+1* to either reach or go below 95% (Figure S5B).

Discussion

In a landmark study on the role of chromosome duplication in adaptive evolution, **?** found that a chromosome III trisomy was acquired by *S. cerevisiae* populations evolving under heat stress, only to be later replaced by euploid mutant cells that carry "refined" solutions to the stress. Additionally, such a replacement also occurred when they initiated evolutionary experiments with a population in which all cells carry a chromosome III trisomy. They hypothesized that the euploid mutant cells evolved by heat-resistance mutations in aneuploid cells followed by reversion of trisomy due to a chromosome loss event.

If indeed the evolved euploid population is descended from the aneuploid population, then mutant alleles that were common in the aneuploid populations should also be common in the evolved euploid population. However, we found that this is not the case (Figure 1): mutant allele frequencies in the aneuploid and euploid populations are negatively correlated, such that common alleles in the former are rare in the later. Furthermore, we developed an evolutionary genetic model of adaptive evolution by aneuploidy and mutation (Figure 2), fitted it to the experimental results of **?**, and used it to predict the genotype frequency dynamics. The model predicted that only about 10-15% of the evolved euploid population descended from aneuploid cells–that is, the majority of the euploid population are not descended from aneuploid cells, but rather are direct descendants of the ancestral wild-type population (Figure 5).

This happens despite aneuploidy reaching a high frequency in the population (>95%). Conventional wisdom might suggest that once the aneuploid genotype *2n+1* reaches high frequency, it will have a better chance at producing "refined" solutions via mutations, and its descendants will come to dominate the population: the frequency of $2n^*_A$ (which arises from *2n+1**) will be higher than the frequency of $2n^*_M$ (which arises directly from *2n*).

So how does $2n^*_M$ prevail? Initially, the supply rates of *2n+1* and $2n^*_M$ are $N\delta \approx 11,000$ and $N\mu \approx 19$, respectively (assuming MAP parameter estimates). Therefore, both genotypes are expected to appear immediately at the beginning of the experiment (Figure S9). However, *2n+1* appears at a much higher frequency as $\delta \gg \mu$ by 2-3 orders of magnitude. After they first appear, $2n^*_M$ has higher fitness. But as long as the frequency of *2n* is high, the supply rate of *2n+1* is higher than that of $2n^*_M$, again due to $\delta \gg \mu$. However, supply rates of both genotypes decreases with the frequency of *2n*. Therefore, when the latter decreases, mainly due to the increase in the frequency of *2n+1*, both supply rates diminish. At this stage, the higher fitness of $2n^*_M$ comes into play and it starts to take over the population, which is mainly composed of *2n+1*. For the aneuploid lineage to compete with the mutant lineage, it must produce $2n^*_A$ via a mutation followed by chromosome loss. Although this is a stochastic process (due to drift), our results show that the time until $2n^*_A$ reaches a frequency of 0.1% is roughly 450 generations, without much variation (intersection of purple lines and vertical dashed line in Figure S9). However, by that time $2n^*_M$ is already at a roughly 10-fold higher frequency (1.86%), and since both mutants have the same fitness, their relative frequency remains roughly the same until the end of the experiment.

Predictions for small populations. We examined the effect of the population size, N, on the frequency of *2n+1* descendants in the evolved population, F_A. We found that F_A is expected to decrease

as the population size increases (Figure 5D), ranging from about 90% when the population size is 10,000, to about 10% when the population size is above 1,000,000 (less than the experimental population size, which was 6,425,000). Thus, our model provides a testable prediction: if the experiment was repeated under a lower population size (via stronger daily dilutions or in a smaller volume), than the fraction of the population descending from aneuploid cells would be much higher.

Aneuploidy delays rather than facilitates adaptation. An additional interesting result of our study is that aneuploidy increases, rather than decreases, the adaptation time (Figure 5E). This happens despite the fact that the mean fitness initially increases faster in the presence of aneuploidy (Figure 5E). This is because once *2n+1* is common, selection for the mutant strain (*2n+1** or *2n**) is weaker compared to when *2n** competes directly with *2n*.

Rate and fitness effect of aneuploidy and mutation. We inferred the rates of aneuploidy and mutation and their effects on fitness. We estimate that the aneuploidy rate (i.e., number of chromosome gains per generation) is $1.7 \cdot 10^{-3}$, higher than a previous estimate of $6.7 \cdot 10^{-6}$ **?** . This may be due to genetic instability caused by heat stress **?** . In addition, we find no evidence for increased mutation rates in aneuploid cells. Previous empirical studies have suggested that genetic instability (e.g., elevated mutation rates) in aneuploid cells is due to stress associated with the aneuploid state **? ? ? ?** . However, in the experiment of **?**, both the wild-type and the aneuploid were under heat stress, which may explain why we did not find evidence for an increased mutation rate.

Conclusions. Here, we tested the hypothesis that aneuploid cells are an evolutionary "stepping stone", or adaptive intermediate, between wild-type euploid cells and mutant euploid cells. Our results suggest that, although it seems the population goes from euploid to aneuploid and back, this is not the case at the individual level. We estimate that only about 10-15% of the euploid cells descended from aneuploid cells, whereas the rest are direct descendants of the wild-type euploid cells. This surprising result reinforces the importance of models when making interpretations on evolutionary processes, and emphasizes the unintuitive outcomes of clonal interference during adaptive evolution.

Limitations and future directions The number of measurement of the frequency of aneuploidy in *S.cervisiae* is limited to three time points (0, $\tilde{4}50$ and $\tilde{2}000$). A more thorough time series could improve inference of model parameters.

Add caveats and possible extension to ABC-SMC

The mathematical model is rather simplistic. It assumes one mutation of fixed effect, at a single locus

that accounts for all the benefit. In the original experiment, the replicates had fluctuating population sizes due to dilution. We use the harmonic mean to determine a single number, on which all of the inference and simulations is based. Extending the model and the simulations to include fluctuating population sizes and adding other features (e.g. dominance, more than one locus) should be done in the future as it is beyond the scope of the current paper.

Models and Methods

DNA sequencing. BLA BLA BLA

Evolutionary genetic model. We model the evolution of a population of cells using a Wright-Fisher model[?] , assuming a constant effective population size N, non-overlapping generations, and including the effects of natural selection, genetic drift, aneuploidy, and mutation. We focus on beneficial genetic modifications, neglecting the effects of deleterious and neutral mutations or karyotypic changes. The model allows for a single aneuploid karyotype (e.g., chromosome III duplication) and a single mutation to accumulate in the genotype. Thus, the model follows four genotypes (Figure 2): euploid wild-type, *2n*, the initial genotype; euploid mutant, *2n**, with the standard karyotype and a single beneficial mutation; aneuploid wild-type, *2n+1*, with an extra chromosome, i.e., following chromosome duplication; and aneuploid mutant, *2n+1**, with and extra chromosome and a beneficial mutation.

Transitions between the genotypes occur as follows (Figure 2): Beneficial mutations from *2n* to *2n** and from *2n+1* to *2n+1** occur with probability μ, the mutation rate. We neglect back-mutations (i.e., from *2n** to *2n* and from *2n+1** to *2n+1*). Aneuploidy is formed by chromosome mis-segregation, so that cells transition from *2n* to *2n+1* and from *2n+1** to *2n** with probability δ, the aneuploidy rate. That is, we assume chromosomes are gained and lost at the same rate, and we neglect events that form a less-fit genotype (i.e., *2n+1* to *2n* and *2n** to *2n+1**).

In the experiment by **?**, the population was grown every day from $1.6 \cdot 10^6$ cells until reaching stationary phase and then diluted 1:120. Thus, we set the population size to $N = 6.425 \cdot 10^6$, the harmonic mean of $\{2^k \cdot 1.6 \cdot 10^6\}_{k=0}^{7}$[?] . The initial population has N cells with genotype *2n*. The effect of natural selection on the frequency f_i of genotype $i = 2n, 2n+1, 2n+1^*$, or $2n^*$ is given by

$$f_i^s = \frac{f_i w_i}{\bar{w}} , \tag{1}$$

where w_i is the fitness of genotype i and $\bar{w} = \sum_j f_j w_j$ is the population mean fitness. The effect of

mutation and aneuploidy on genotype frequencies is given by

$$\begin{aligned} f^m_{2n} &= (1-\delta-\mu) f^s_{2n} \,, \\ f^m_{2n+1} &= \delta f^s_{2n} + (1-\mu) f^s_{2n+1} \,, \\ f^m_{2n+1^*} &= \mu f^s_{2n+1} + (1-\delta) f^s_{2n+1^*} \,, \\ f^m_{2n^*} &= \mu f^s_{2n} + \delta f^s_{2n+1} + f^s_{2n^*} \,. \end{aligned} \tag{2}$$

Finally, random genetic drift is modeled using a multinomial distribution[?] ,

$$\mathbf{f}' \sim \frac{1}{N} \cdot Mult(N,\, \mathbf{f^m}) \,, \tag{3}$$

where $\mathbf{f^m} = (f^m_{2n}\,,\ f^m_{2n+1}\,,\ f^m_{2n+1^*}\,,\ f^m_{2n^*})$ are the frequencies of the genotypes after mutation and aneuploidy, $\mathbf{f}'$ are the genotype frequencies in the next generation, and $Mult(N,\ \mathbf{f})$ is a multinomial distribution parameterized by the population size N and the genotype frequencies $\mathbf{f}$. Overall, the change in genotype frequencies from one generation to the next is given by the transformation $f_i \to f'_i$.

Empirical data for model inference. We use the results of evolutionary experiments reported by **?**. In their heat-stress experiment, four populations of *S. cerevisiae* evolved under 39 °C. Aneuploidy fixed in all four population in the first 450 generations. Hereafter, fixation or elimination of a genotype *by generation t* means that more than 95% or less than 5% of the population carry the genotype at generation t, and possibly earlier. From re-analysis of data not published in the original paper, aneuploidy did not fix before at least 200 generations elapsed. The experiment continued with two populations, in which aneuploidy was eliminated by generation 1,700 and 2,350 while still under the same conditions of elevated heat (39 °C).

Likelihood function. Because our model, just like the Wright-Fisher model, is non-linear and stochastic, computing the distribution of fixation time $T(g)$ of genotype g for use in the likelihood function is intractable (it is even hard to use a diffusion-equation approximation due to the model having multiple genotypes, rather than just two). We overcome this problem by approximating the likelihood using simulations. We simulate 1,000 experiments per parameter vector $\theta = (\mu, \delta, s, b, c)$, resulting in a set of simulated observations $\tilde{\mathbf{X}} = \{\tilde{X}_i\}_{i=1}^{1000}$. We then compute the approximate likelihood,

$$\begin{aligned} \mathcal{L}(\theta) = P^4(200 \le T(2n+1) \le 450) \cdot \Big[1 - \\ P^4_{\tilde{\mathbf{X}}}\big(!\{T(2n^*) < 1700\} \mid 200 \le T(2n+1) \le 450\big) - \\ P^4_{\tilde{\mathbf{X}}}\big(!\{1700 < T(2n^*) < 2350\} \mid 200 \le T(2n+1) \le 450\big) + \\ P^4_{\tilde{\mathbf{X}}}\big(!\{T(2n^*) < 1700\} \wedge !\{1700 < T(2n^*) < 2350\} \mid 200 \le T(2n+1) \le 450\big) \Big] \,, \end{aligned} \tag{4}$$

where !{...} is the "logical not" operator, $P^4(\ldots)$ is the 4th power of $P(\ldots)$, and all probabilities $P_{\tilde{\mathbf{X}}}(\ldots)$ are approximated from the results of the simulations $\tilde{\mathbf{X}}$. For example, $P_{\tilde{\mathbf{X}}}\big(!\{T(2n^*) < 1700\} \mid 200 \le T(2n+1) \le 450\big)$ is approximated by taking simulations in which *2n+1* fixed before generation 450 but not before generation 200, and computing the fraction of such simulations in which *2n** did not fix by generation 1,700, and hence aneuploidy did not extinct before generation 1,700. Figure S1 compares results with less and more simulated experiments, demonstrating that 1,000 simulations are likely sufficient.

For a model without aneuploidy (that is, when the aneuploidy rate is fixed at zero, $\delta = 0$), we disregard the increased expression in chromosome III and the growth advantage measured in generation 450, and focus on the growth advantage measured in later generations, presumably due to a beneficial mutation. Therefore, the likelihood is approximated by

$$\begin{aligned}\mathcal{L}_!(\theta) = 1 &- P^4_{\tilde{\mathbf{X}}}\big(!\{T(2n^*) < 1700\}\big) - \\ &P^4_{\tilde{\mathbf{X}}}\big(!\{1700 < T(2n^*) < 2350\}\big) + \\ &P^4_{\tilde{\mathbf{X}}}\big(!\{T(2n^*) < 1700\} \wedge !\{1700 < T(2n^*) < 2350\}\big)\ .\end{aligned} \tag{5}$$

Parameter inference. To infer model parameters, we use approximate Bayesian computation with a sequential Monte-Carlo scheme, or ABC-SMC[?], implemented in the `pyABC` Python package[?] pyabc.readthedocs.io. This approach uses numerical stochastic simulations of the model to infer a posterior distribution over the model parameters. It is a method of likelihood-free, simulation-based inference[?], that is, for estimating a posterior distribution when a likelihood function cannot be directly computed. It is therefore suitable in our case, in which the likelihood function can only be approximated from simulations, and cannot be directly computed.

The ABC-SMC algorithm employs sequential importance sampling over multiple iterations[? ? ?]. In iteration t of the algorithm, a set of parameter vectors, $\{\theta_{i,t}\}_{i=1}^{n_t}$, also called *particles*, are constructed in the following way. A proposal particle, θ^*, is sampled from a proposal distribution, and is either accepted or rejected, until n_t particles are accepted. The number of particles, n_t, is adapted at every iteration t using the adaptive population strategy[?] pyabc.readthedocs.io. For $t = 0$, the proposal particle is sampled from the prior distribution, $p(\theta)$. For $t > 0$, the proposal particle is sampled from the particles accepted in the previous iteration, $\{\theta_{i,t-1}\}_{i=1}^{n_{t-1}}$, each with a probability relative to its weight $W_{t-1}(\theta_{i,t-1})$ (see below). The proposal particle is then perturbed using a kernel perturbation kernel, $K_t(\theta^* \mid \theta)$ where θ is the sample from the previous iteration. Then, a set of synthetic observations $\tilde{\mathbf{X}}^*$ is simulated, and the proposal particle θ^* is accepted if its approximate likelihood (eq. (4)) is high enough, $\mathcal{L}(\theta^*) > 1 - \epsilon_t$ (or more commonly, if $1 - \mathcal{L}(\theta^*) < \epsilon_t$), where $\epsilon_t > 0$ is the *acceptance*

threshold, as higher values of ϵ_t allow more particles to be accepted. The acceptance threshold ϵ_t is chosen as the median of the $1 - \mathcal{L}(\theta)$ of the particles accepted in the previous iteration, $t - 1$, and $\epsilon_0 = 0.01$. For each accepted particle $\theta_{i,t}$ a weight $W_t(\theta_{i,t})$ is assigned: for $t = 0$, $W_0(\theta_{i,0}) = 1$, and for $t > 0$, $W_t(\theta_{i,t}) = p(\theta_{i,t}) / \sum_{i=1}^{n_{t-1}} W_{t-1}(\theta_{i,t-1}) K_t(\theta_{i,t}, \theta_{i,t-1})$, where $p(\theta)$ is the prior density of θ and $K_t(\theta', \theta)$ is the probability of a perturbation from θ to θ'. $K_t(\theta' \mid \theta)$ is a multivariate normal distribution, fitted at iteration t to the particles from the previous iteration, $\{\theta_{i,t-1}\}_{i=1} n_{t-1}$, and their weights, $\{W(\theta_{i,t-1})\}_{i=1}^{n_{t-1}}$.

Acceptance is determined according to the approximate likelihood (eq. (4)), which has a maximum value of $\mathcal{L}_{max} = 0.875$ (giving a minimal value of $\epsilon_{min} = 0.125$). We terminated the inference iterations when the change in ϵ value from one iteration to the next was small. With our standard prior and model, we reached $\epsilon = 0.13$ (or $\mathcal{L} = 0.87$) after six iterations, with $n_6 = 982$ accepted parameter vectors and effective sample size ESS=651 (Figure S2). Running the inference algorithm with different initialization seeds and less or more simulations for approximating the likelihood produced similar posterior distributions (Figure S1).

After producing a set of weighted particles from the the posterior distribution using the above ABC-SMC algorithm, we approximate the posterior using kernel density estimation (KDE) with Gaussian kernels. We truncate the estimated posterior to avoid positive posterior density for values with zero prior density. The MAP (maximum a posteriori) estimate is computed as the the maximum of the estimated joint posterior density. We then draw $5,000,000$ samples from the posterior distribution to compute the HDI (highest density interval) and draw $50,000$ samples to visualize the posterior distribution with histograms.

Model comparison. We examine several versions of our evolutionary models, e.g. without aneuploidy or with increased mutation rate in aneuploid cells, as well as several different prior distributions (see below). To compare these, we plot posterior predictions: for each model we execute $10,000$ simulations using the MAP parameter estimates and plot the distributions of time to fixation of $2n*$, one of key properties of the model likelihood. These plots visualize the fit of each model to the data. Also, for similar models we plot the marginal and joint posterior distributions of the parameters; if these are similar, we consider the models interchangeable. We validate this by comparing HDI (highest density interval) of posterior distributions.

Where posterior plots are very similar and the number of parameters is the same, we use WAIC, or

the widely applicable information criterion [?], defined as

$$WAIC(\theta) = -2\log\mathbb{E}[\mathcal{L}(\theta)] + 2\mathbb{V}[\log\mathcal{L}(\theta)] \quad (6)$$

where θ is a parameter vector, and $\mathbb{E}[\cdot]$ and $\mathbb{V}[\cdot]$ are the expectation and variance taken over the posterior distribution, which in practice are approximated using 50,000 samples from the posterior KDE. We validated that upon resampling WAIC values do not significantly change and that differences in WAIC between models are preserved. WAIC values are scaled as a deviance measure: lower values imply higher predictive accuracy[?].

Prior distributions. We used informative prior distributions for $w_{2n+1} = 1 - c + b$, $w_{2n+1*} = (1+s)(1-c) + b$ and $w_{2n*} = 1 + s$, which we estimated from growth curves data from mono-culture growth experiments previously reported by **?** , Figs. 3C, 4A, and S2. We used `Curveball`, a method for predicting results of competition experiments from growth curve data[?] curveball.yoavram.com. Briefly, `Curveball` takes growth curves of two strains growing separately in mono-culture and predicts how they would grow in a mixed culture, that is, it predicts the results of a competition assay. From these predictions, relative fitness values can be computed. Because `Curveball` uses a maximum-likelihood approach to estimate model parameters, we were able to estimate a distribution of relative fitness values to be used as a prior distribution by sampling 10,000 samples from a truncated multivariate normal distribution defined by the maximum-likelihood covariance matrix (Figure S3).

We used growth curves of *2n* and *2n+1* in 39 °C to estimate an informative prior distribution for w_{2n+1} (Figure S3-D, assuming $w_{2n} = 1$). In this prior distribution, we used the same prior for w_{2n+1*} and w_{2n*}. To increase computational efficiency, we also assumed $w_{2n*} > w_{2n+1*} > w_{2n+1} > w_{2n}$; running the inference without this assumption produced similar results. See *supporting material* for an extended informative prior distribution that uses growth curves of *2n** and *2n+1* growing in 39 °C; this prior distribution proved to be less useful.

As a control, we tested an uninformative uniform prior with $U(1,6)$, for (i) all w_{2n+1}, w_{2n+1*}, w_{2n*}, or (ii) only for w_{2n+1*}, w_{2n*}, using the above informative prior for w_{2n+1}. In these cases the inference algorithm failed to converge.

For the mutation rate, μ, and aneuploidy rate, δ, we used uninformative uniform priors, $\mu \sim U(10^{-9}, 10^{-5})$ and $\delta \sim U(10^{-6}, 10^{-2})$. A wider mutation rate prior, $\mu \sim U(10^{-9}, 10^{-3})$, produced similar results.

Acknowledgements

We thank Yitzhak Pilpel, Orna Dahan, Lilach Hadany, Judith Berman, David Gresham, Shay Covo, Martin Kupiec, Uri Obolski, Daniel Weissman, and Tal Simon for discussions and comments. This work was supported in part by the Israel Science Foundation (ISF, YR 552/19), the US–Israel Binational Science Foundation (BSF, YR 2021276) Minerva Stiftung center for Lab Evolution (YR), Minerva Stiftung short-term research grant (MP).

Supplementary Material

Supplementary Analysis

Sensitivity analysis. Changing a single parameter while keeping the rest fixed at the MAP estimate produces a worse fit to the data (Figure S6). Furthermore, we fitted models with a mutation rate fixed at $\mu = 10^{-5}$, 10^{-6} and 10^{-7}. We inferred similar parameters estimates for the model with $\mu = 10^{-6}$ compared to the model with a free μ parameter, in which the inferred mutation rate is $\mu \approx 3 \cdot 10^{-6}$. Inference assuming $\mu = 10^{-5}$ or $\mu = 10^{-7}$ produced similar estimates except that the estimated aneuploidy rate, δ, was higher, and assuming $\mu = 10^{-7}$, the estimated fitness of *2n+1* was lower (Figure S7).

Extended informative prior distribution. In an extended informative prior distribution, we used additional growth curves of *2n** (*refined* strain from **?**) and *2n+1* in 39 °C to estimate w_{2n*}/w_{2n+1} (Figure S3L). The same distribution was used for w_{2n*}/w_{2n+1*}. Thus, our main informative prior uses a single prior distribution for fitness values of *2n+1*, *2n+1**, and *2n**, whereas the extended informative prior uses one distribution for *2n+1*, and another distribution for both *2n+1** and *2n**.

We estimated the parameters under this extended informative prior. Inference took much longer to run but the posterior distribution seemed to converge, as it did not change much in the final iterations. The posterior predictive plot shows that inference with this extended prior produces a posterior distribution that fails to explain the empirical observations (pink in Figure 4). However, the inferred posterior distribution is considerably narrower (compare Figures 3 and S8) and therefore parameter estimates are less variable. The estimated mutation rate was much lower compared to the main informative prior, with $\mu = 2.474 \cdot 10^{-9}$ $[2.423 \cdot 10^{-9} - 2.612 \cdot 10^{-9}]$. Other parameter estimates are: $\delta = 2.705 \cdot 10^{-3}$ $[2.094 \cdot 10^{-3} - 3.094 \cdot 10^{-3}]$, $w_{2n+1} = 1.022$ $[1.021 - 1.024]$, $w_{2n+1^*} = 1.052$ $[1.05 - 1.054]$, $w_{2n^*} = 1.053$ $[1.051 - 1.055]$, the latter two being much higher compare to the main informative prior. Notably, the mode of the posterior ratio $w_{2n*}/w_{2n+1} = 1.0009$ is much lower than the mode of the prior ratio of 1.033 (Figure S3H) and closer to the ratio of 1 that we assume in the main informative prior. Together with the posterior predictive results, we conclude that the main informative prior is preferable over the extended informative prior.

Model with transitions to less-fit genotypes We also estimated the parameters of a version of the model that includes transitions (mutation, chromosome loss and gain) to less-fit genotypes (e.g., *2n**

to *2n+1**),

$$
\begin{aligned}
f^m_{2n} &= (1-\delta-\mu) f^s_{2n} + \delta f^s_{2n+1} + \mu f^s_{2n^*} \,, \\
f^m_{2n+1} &= \delta f^s_{2n} + (1-\delta-\mu) f^s_{2n+1} + \mu f^s_{2n+1^*} \,, \\
f^m_{2n+1^*} &= \mu f^s_{2n+1} + (1-\delta-\mu) f^s_{2n+1^*} + \delta f^s_{2n^*} \,, \\
f^m_{2n^*} &= \mu f^s_{2n} + \delta f^s_{2n+1^*} + (1-\delta-\mu) f^s_{2n^*} \,.
\end{aligned}
\tag{7}
$$

The inferred values are slightly different. The estimated mutation rate, $\mu = 1.036 \cdot 10^{-7}$ [$8.01 \cdot 10^{-8} - 1.339 \cdot 10^{-7}$], corresponds to a mutation target size of $\sim 300 - 500$, assuming the mutation rate per base pair is roughly $2 \cdot 10^{-10}$ (ref.**?**) or $3.3 \cdot 10^{-10}$ (ref.**?**). The estimated aneuploidy rate, $\delta = 2.358 \cdot 10^{-4}$ [$1.766 \cdot 10^{-4} - 2.837 \cdot 10^{-4}$] is 5-35-fold higher than in previous studies: for chromosome III in diploid *S. cerevisiae*, **?** estimated $6.7 \cdot 10^{-6}$ chromosome gain events per generation, and **?** estimate $3.0 - 4.3 \cdot 10^{-5}$ chromosome loss events per generation (95% confidence interval). The estimated fitness values are $w_{2n+1} = 1.024$ [1.023 – 1.025], $w_{2n+1^*} = 1.025$ [1.024 – 1.026], $w_{2n^*} = 1.032$ [1.031 – 1.033], all relative to the fitness of *2n*, which is set to $w_{2n} = 1$. Thus, we can infer that the cost of trisomy is $c = w_{2n^*} - w_{2n+1^*} = 0.007$ (or 0.7%) and the benefit of trisomy is $w_{2n+1} - 1 - c = 0.017$ (1.7%), whereas the benefit of beneficial mutation is $w_{2n^*} - 1 = 0.032$ (3.2%).

We simulated genotype frequency dynamics using parameter samples from the posterior distribution, and computed the posterior distribution of F_A. The mean F_A in this case is just 0.0189 [0.0004 - 0.1214 95% CI], lower than without the transitions to less-fit genotypes. Here, F_A is the sum of frequencies of both $2n^*_A$ and $2n + 1^*_A$, which reaches a frequency of 0.0007. Out of 100,000 posterior samples, none had F_A above 0.05 (i.e., 5% of the population).

Supplementary Figures & Tables

Table S1: Mutant alleles in population $H2$.

Mutant alleles identified in the ancestor (generation 0), aneuploid (generation 450), and evolved (generation 1,700) of population $H2$. See supplementary file.

Table S2: Mutant alleles in population $H4$.

Mutant alleles identified in the ancestor (generation 0), aneuploid (generation 450), and evolved (generation 2,350) of population $H4$. See supplementary file.

Table S3: WAIC values for different τ values.

Model	WAIC
$\tau = 1$	-9
$\tau = 33/32$	-9
$\tau = 2$	-8
$\tau = 5$	-12
$\tau = 10$	-9
$\tau = 100$	-12

WAIC defined in eq. (6).

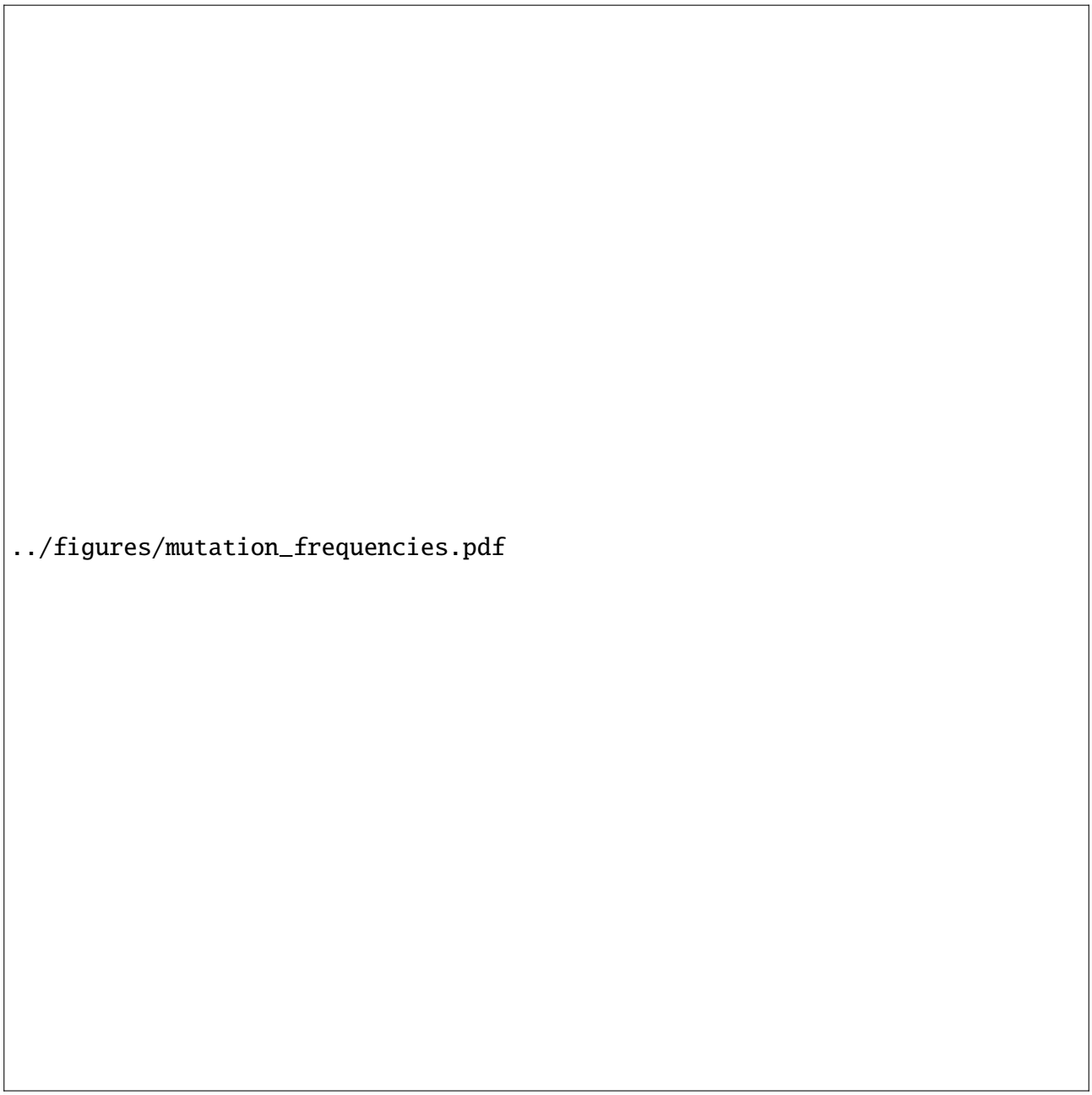


Figure 1: Frequencies of mutant alleles in the experimental populations are negatively correlated. Frequencies of mutant alleles when trisomy was widespread in the population (y-axis) and after it was eliminated (x-axis) in two experimental repetitions (circles for $H2$ and triangles for $H4$) from **?**. Mutant alleles with >20% in the aneuploid population were <20% in the euploid population, and vice versa (the upper-right quadrant is empty), suggesting that the majority of evolved euploid cells did not descend from the most common aneuploid genotypes. Alleles with frequency below and above 10% in the ancestral populations are in black and gray, respectively. Solid black line is a linear orthogonal distance regression line (slope=-0.559, intercept=0.164; a regression through alleles that reach at least 20% in one of the populations has slope=-0.645 and intercept=0.297). Dashed vertical and horizontal lines show allele frequencies 20%. Error bars show standard error of the mean accounting for the number of reads.

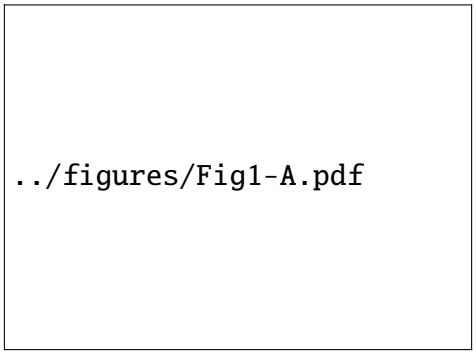


Figure 2: Model Illustration. There are four genotypes in our model: euploid wild-type, *2n*; euploid mutant, *2n**; aneuploid wild-type, *2n+1*; and aneuploid mutant, *2n+1**. Overall there are two possible trajectories from *2n* to *2n**. Arrows denote transitions between genotypes, with transitions rates μ for the beneficial mutation rate and δ for the aneuploidy rate.

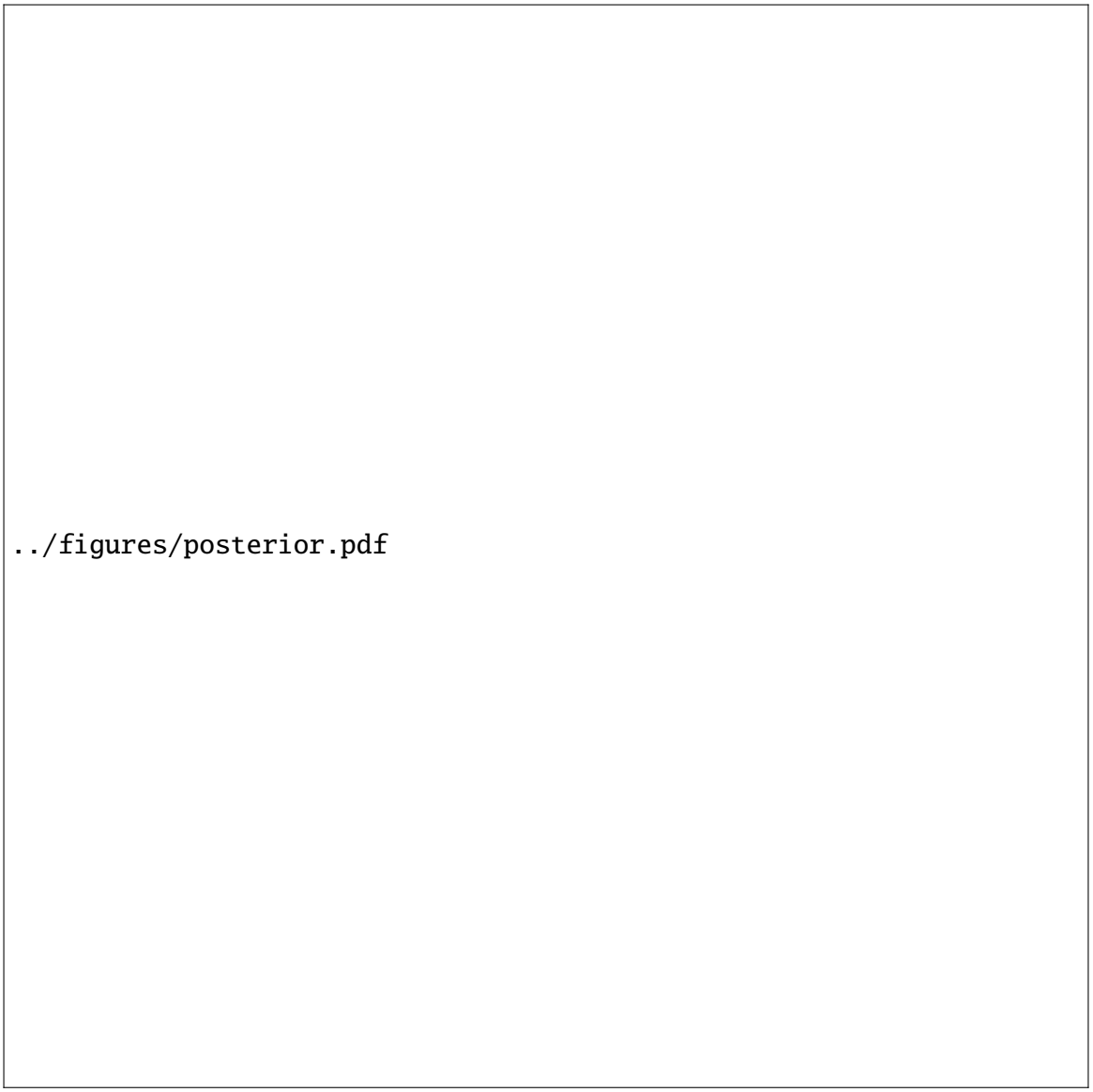


Figure 3: Posterior distribution of model parameters. On the diagonal, the marginal posterior distribution of each model parameter. Below the diagonal, the joint posterior distribution of pairs of model parameters (dark purple and bright yellow for low and high density, respectively). Red markers and orange lines for the joint MAP estimate (which may differ from the marginal MAP, as the marginal distribution integrates over all other parameters).

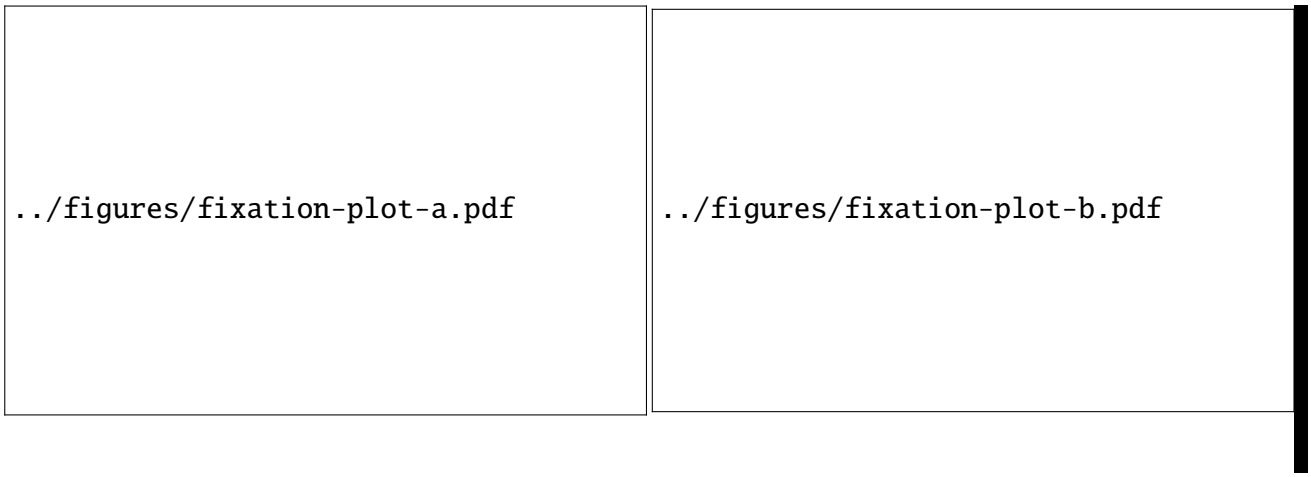


Figure 4: Model fit with and without aneuploidy. The distribution of time to fixation of $2n*$ (i.e., adaptation time) in 10,000 simulations using MAP parameters of the model with beneficial aneuploidy (blue; $\delta > 0,\ w_{2n} < w_{2n+1} < w_{2n+1*} < w_{2n*}$) compared to alternative models: a model with the same parameter values but without aneuploidy (gray, $\delta = 0$, concentrated at $t = 450$); a model fitted to the data assuming no aneuploidy (green, $\delta = 0$); a model fitted to the data assuming neutral aneuploidy (yellow, $\delta > 0,\ w_{2n+1} = w_{2n},\ w_{2n+1*} = w_{2n*}$); and a model with beneficial aneuploidy and an extended prior distribution (pink). In the experiment by **?**, one population lost aneuploidy by generation 1,700 and another by generation 2,350 (dashed lines) but not before generation 450. Thus, the blue distribution has a better fit compared to the other distributions (the gray distribution has a particulary poor fit). The MAP likelihood (eq. (4)) is 0.84, 0.78, 0.67, and 0.14 for the models represented by blue, yellow, green, and pink distributions, respectively. **(A)** Histogram of the time to fixation of $2n*$. The last bin contains all values equal or greater than 3,000. **(B)** Cumulative distribution of the time to fixation.

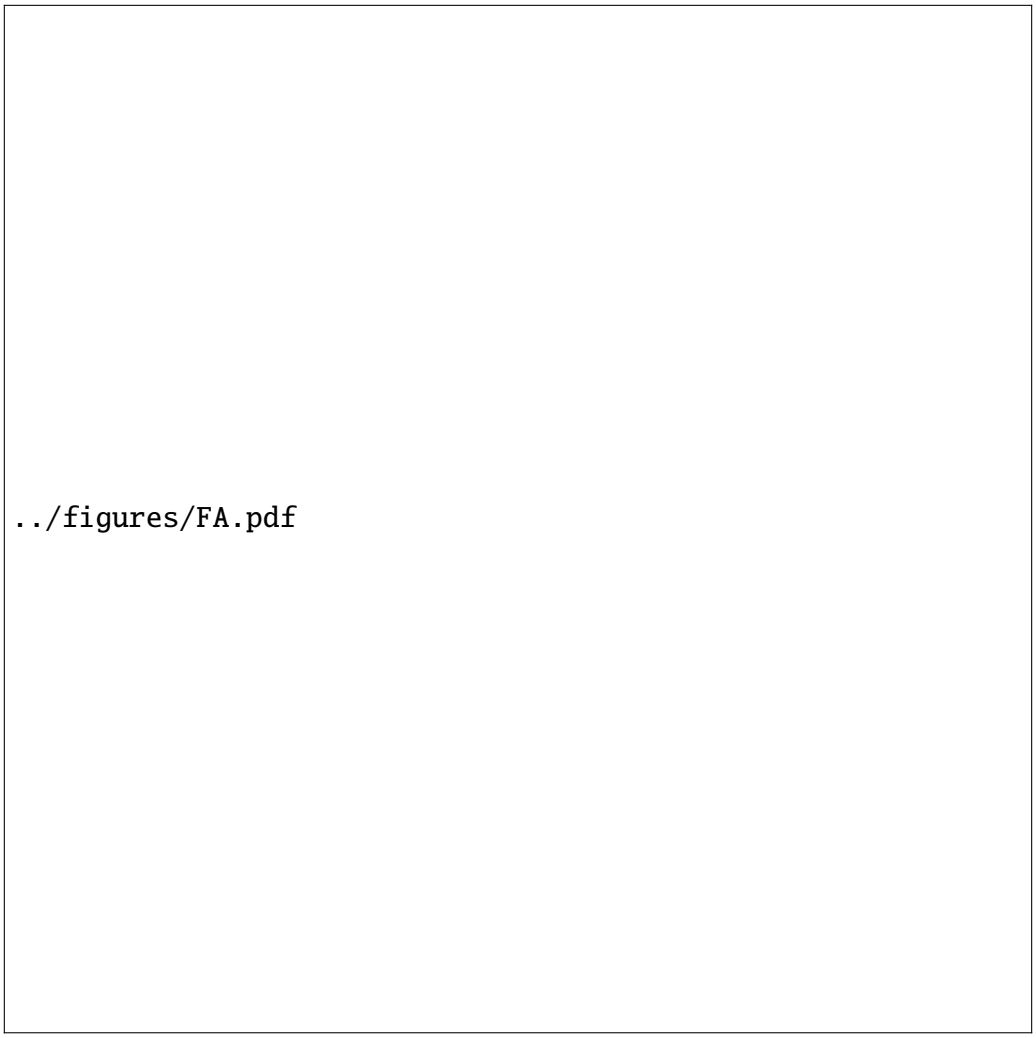


Figure 5: Predicted frequency of aneuploid-descended cells. **(A)** Posterior predicted genotype frequencies over time, including the source of *2n**: $2n^*_A$ arose from $2n + 1$, whereas $2n^*_M$ arose directly from $2n$. Colored curves are 50 simulations using the MAP estimate parameters. Black dashed curves are the expected genotype frequencies without genetic drift (from a deterministic model). See Figure S9 for log-log scale, in which the sequence of events is easier to observe. **(B)** Posterior distribution of F_A, the expected frequency of *2n** cells descended from *2n+1* cells, computed as the average frequency at the end of 100 simulations for 100,000 samples from the parameter posterior distribution. Solid and dashed lines show the mean and 95% CI. **(C)** F_A values (color coded) from panel B, with their corresponding mutation rate μ on x-axis and aneuploidy rate δ on the y-axis. White star shows the MAP estimate. See also Figure S10. **(D)** F_A as a function of the population size, N, in posterior predictions with MAP parameters. Markers show F_A in 250 simulations per population size. Error bars show mean F_A with 95% CI (bootstrap, $n = 10,000$). Vertical dashed line for population size in the experiment, $6.425 \cdot 10^6$. Horizontal line for $F_A^{MAP} = 0.106$. **(E)** Population mean fitness in a model without drift using MAP estimate parameters. Solid lines for mean fitness with aneuploidy ($\delta > 0$), where the population reaches adaptation (mean fitness at 99.99% of maximum value) at generation 1,802. Dashed lines for mean fitness without aneuploidy ($\delta = 0$), where the population adapts much earlier, at generation 535.

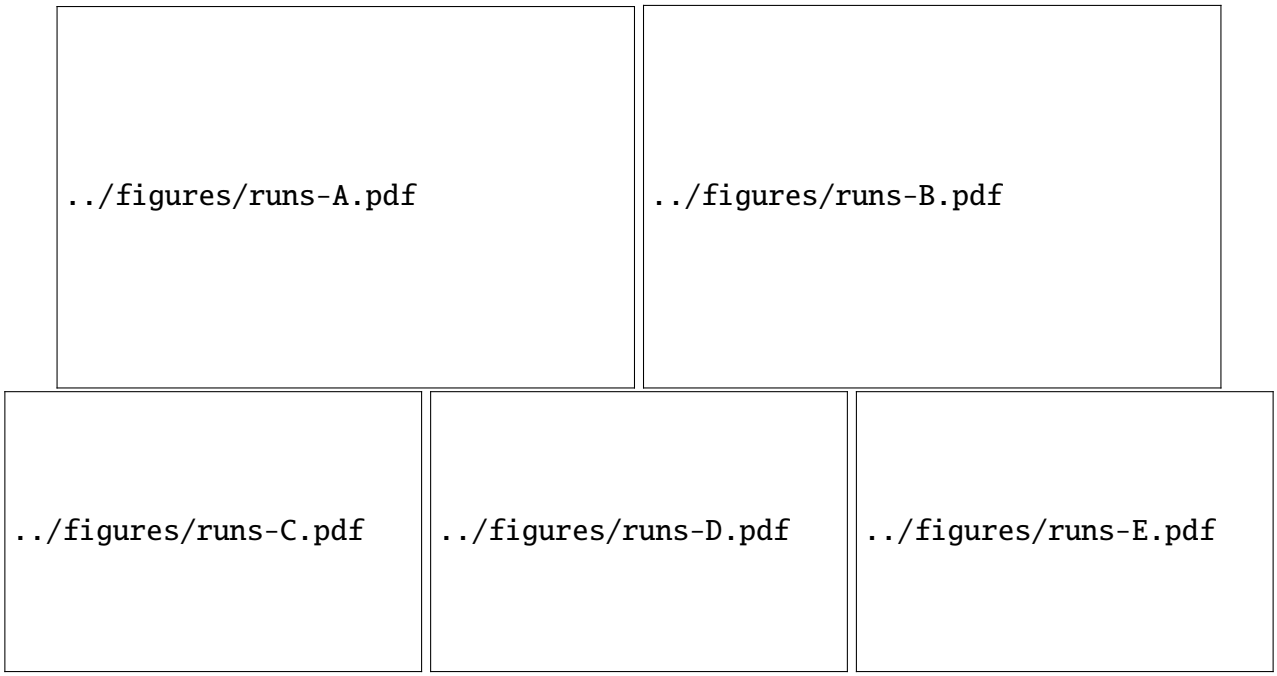


Figure S1: Posterior distribution validation. The posterior distribution of model parameters is roughly the same regardless of the number of simulations (4-10,000 replicates) used to approximate the likelihood (eq. (4)).

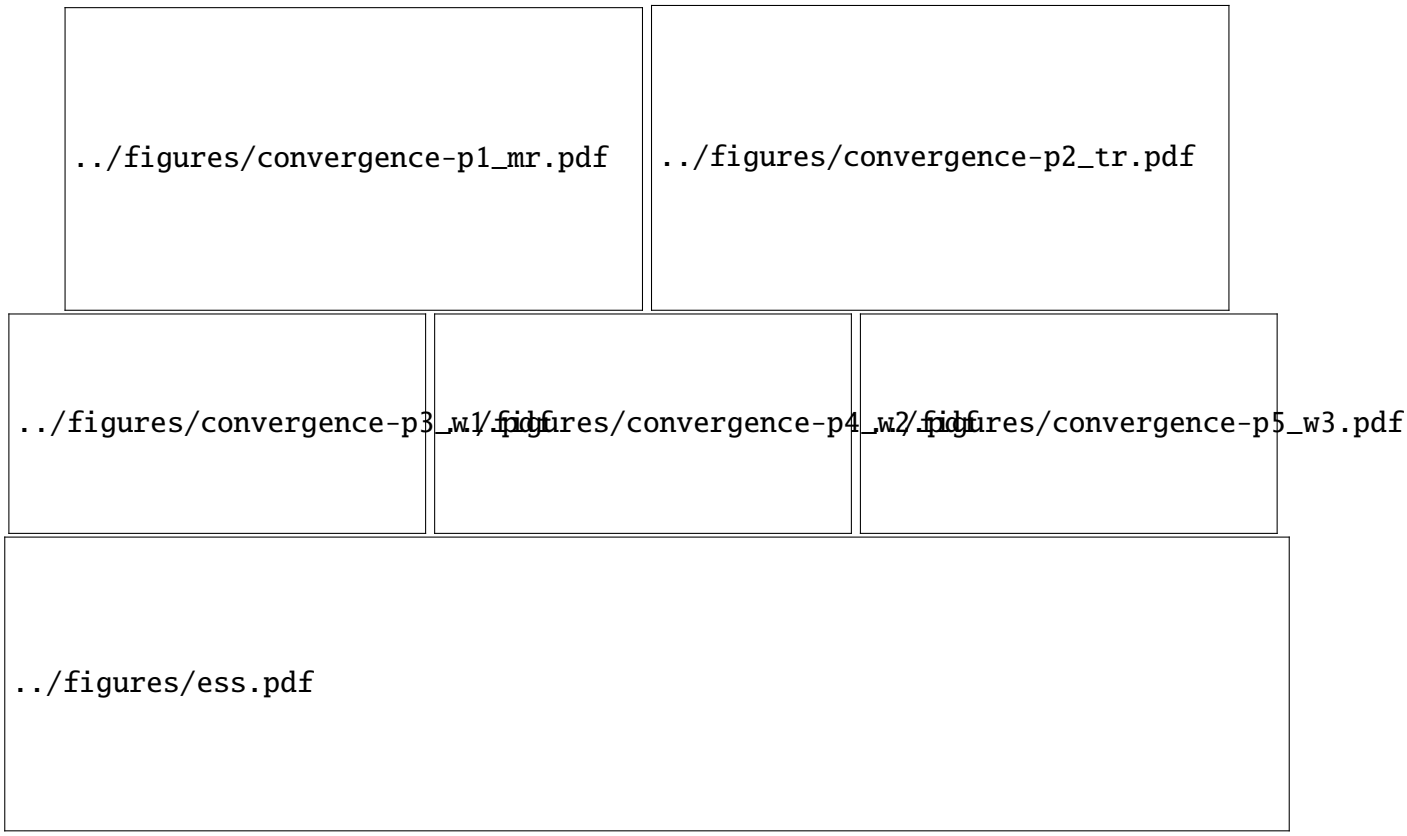
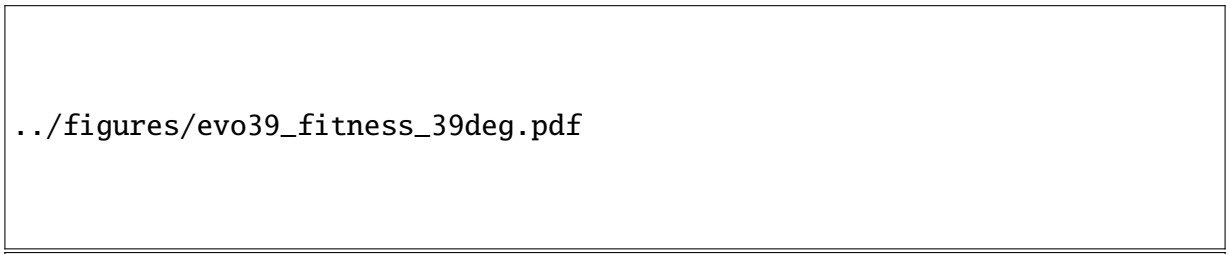


Figure S2: Inference convergence. The ABC-SMC algorithm was used to infer the model parameters. **(A-E)** The approximate posterior distributions of model parameters at each iteration of the ABC-SMC algorithm demonstrates convergence, as the posterior did not significantly change after the first iteration, $t = 1$. **(F-I)** ABC-SMC measures of convergence. After iteration number 6, the acceptance threshold was $\epsilon = 0.13$ (i.e., $\mathcal{L} = 0.87$, eq. (4)), the acceptance rate was 0.018, the number of particles was 982, and the effective sample size ESS=651.



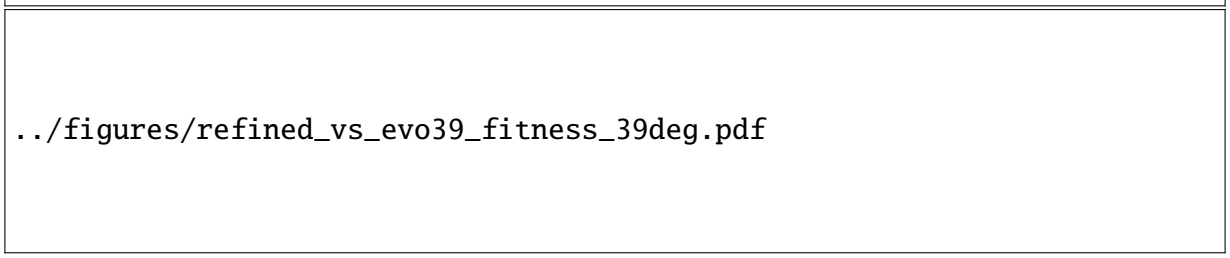


Figure S3: Fitness estimation from growth curves. (A-D) Fitness estimation from growth curves of *2n* and *2n+1* at 39 °C. w_{2n+1}/w_{2n}=1.024 (95% CI: 0.959 - 1.115). `Curveball` **(E-H)** Fitness estimation from growth curves of *2n+1* and *2n** at 39 °C. w_{2n*}/w_{2n+1}=1.033 (95% CI: 1.027 - 1.041). Growth curves previously described in **?** , Figs. 3C, 4A, and S2. Fitness estimated from growth curves using `Curveball`, a method for predicting results of competition experiments from growth curve data[?] curveball.yoavram.com. See *Models and Methods, Prior distributions* for more details. **(A,B;E,F)** Mono-culture growth curve data (markers) and best-fit growth models (lines). **(C,G)** The mixed-culture prediction for the strains from A,B and E,F respectively, 6,375 generated curves. **(D,H)** The relative fitness distribution for *2n+1* relative to *2n* (panel D) and *2n** relative to *2n+1* (panel H). Figures generated by `Curveball`.

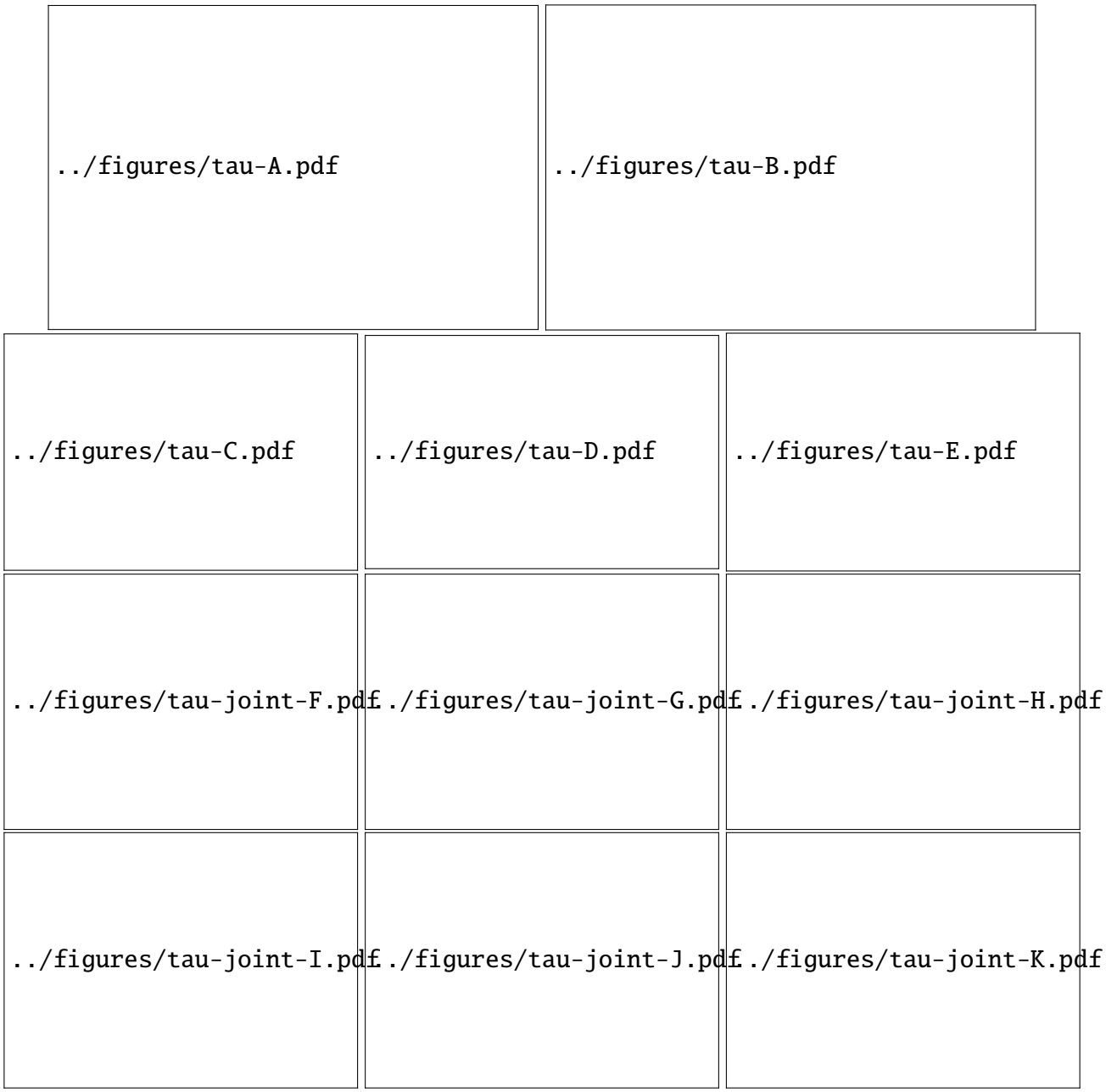


Figure S4: Model with elevated mutation rate in aneuploid cells. (A-E) The inferred posterior distributions for models with different values of τ, the fold-increase in mutation rate in aneuploid cells (*2n+1* and *2n+1**). Vertical dashed lines represent the MAP (maximum a posteriori) of each distribution. When the increase in mutation rate is high, $\tau = 10$ and $\tau = 100$, the inferred mutation (A) and aneuploidy (B) rates tend to be lower. **(F-K)** The inferred joint posterior distribution of mutation rate (μ) and aneuploidy rate (δ) with different τ values (dark purple and bright yellow for low and high density, respectively).

Figure S5: Genotype fixations for models with increased genetic instability. We estimated the parameters for different models, each assuming a different value of τ, the fold-increase in mutation rate in aneuploid cells. We then generated 10,000 simulations using the MAP estimate of each model and evaluated the fraction of simulations in which the frequency of genotype *2n* (**A**), *2n+1* (**B**), and *2n** (**C**) is above 95% (y-axis) at each generation (x-axis). Note that *2n+1** did not fix. We can see that $\tau = 100$ can be distinguished if the waiting time for $f_{2n} < 95\%$ is known (panel A) or if the waiting time for $f_{2n+1} > 95\%$ or $f_{2n+1} < 95\%$ is known (panel B). It is harder to distinguish between $1 \leq \tau \leq 10$.

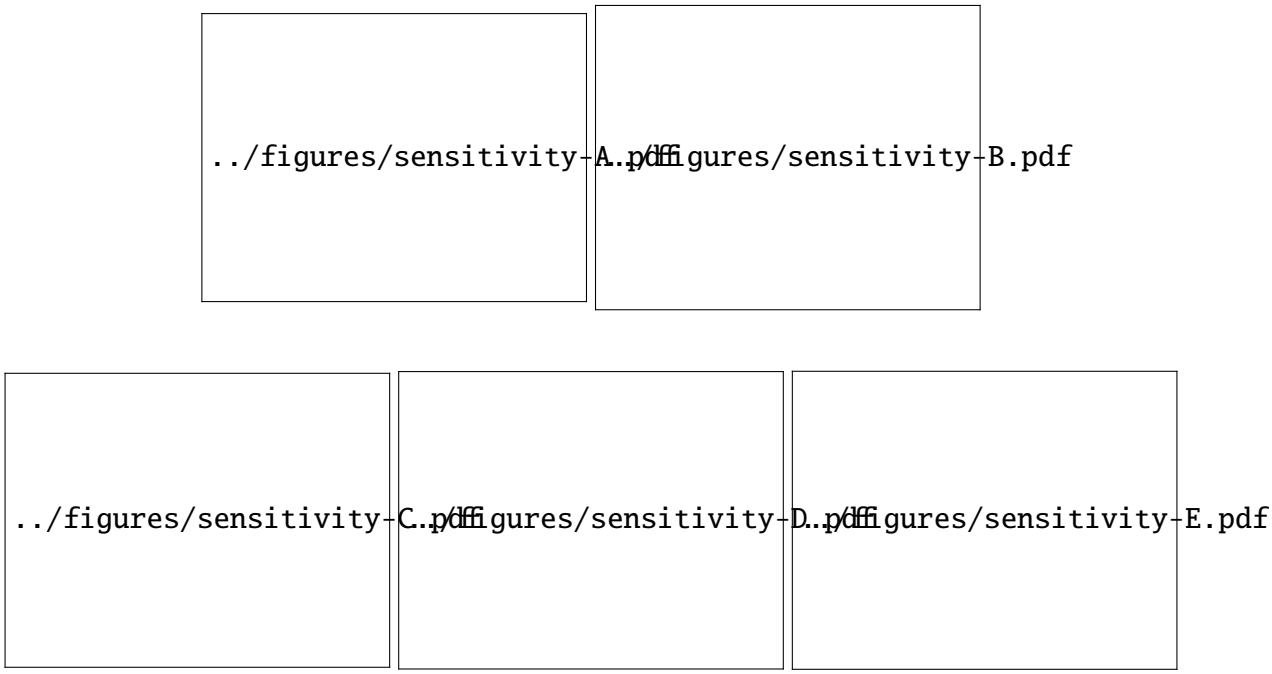


Figure S6: Likelihood profiles. Sensitivity of the model approximate likelihood, $\mathcal{L}(\theta)$, to changing a single parameter while the other parameters remain fixed at their MAP estimates. Dashed vertical line represents the MAP value. The prior distributions for the mutation rate and aneuploidy rate are $\mu \sim U(10^{-9}, 10^{-5})$ and $\delta \sim U(10^{-6}, 10^{-2})$, respectively.

Figure S7: Model with fixed mutation rate. (A-D) The inferred posterior distributions for models with free and fixed mutation rate, μ. The MAP (maximum a posteriori) and 50% HDI (highest density interval) for each model are: **free μ, run 1:** $\delta = 1.720 \cdot 10^{-3}$ $[1.470 \cdot 10^{-3} - 2.786 \cdot 10^{-3}]$, $w_{2n+1} = 1.022$ $[1.021 - 1.023]$, $w_{2n+1^*} = 1.025$ $[1.024 - 1.026]$, $w_{2n^*} = 1.028$ $[1.026 - 1.029]$; **free μ, run 2:** $\delta = 2.129 \cdot 10^{-3}$ $[1.334 \cdot 10^{-3} - 2.695 \cdot 10^{-3}]$, $w_{2n+1} = 1.022$ $[1.02 - 1.023]$, $w_{2n+1^*} = 1.025$ $[1.023 - 1.026]$, $w_{2n^*} = 1.028$ $[1.026 - 1.029]$; **$\mu = 10^{-5}$:** $\delta = 2.903 \cdot 10^{-3}$ $[2.399 \cdot 10^{-3} - 3.156 \cdot 10^{-3}]$, $w_{2n+1} = 1.022$ $[1.021 - 1.023]$, $w_{2n+1^*} = 1.024$ $[1.023 - 1.025]$, $w_{2n^*} = 1.027$ $[1.026 - 1.028]$; **$\mu = 10^{-6}$:** $\delta = 1.917 \cdot 10^{-3}$ $[9.624 \cdot 10^{-4} - 2.447 \cdot 10^{-3}]$, $w_{2n+1} = 1.022$ $[1.02 - 1.023]$, $w_{2n+1^*} = 1.025$ $[1.023 - 1.026]$, $w_{2n^*} = 1.028$ $[1.027 - 1.029]$; **$\mu = 10^{-7}$:** $\delta = 2.901 \cdot 10^{-3}$ $[2.139 \cdot 10^{-3} - 3.671 \cdot 10^{-3}]$, $w_{2n+1} = 1.019$ $[1.017 - 1.02]$, $w_{2n+1^*} = 1.025$ $[1.024 - 1.026]$, $w_{2n^*} = 1.027$ $[1.026 - 1.029]$.

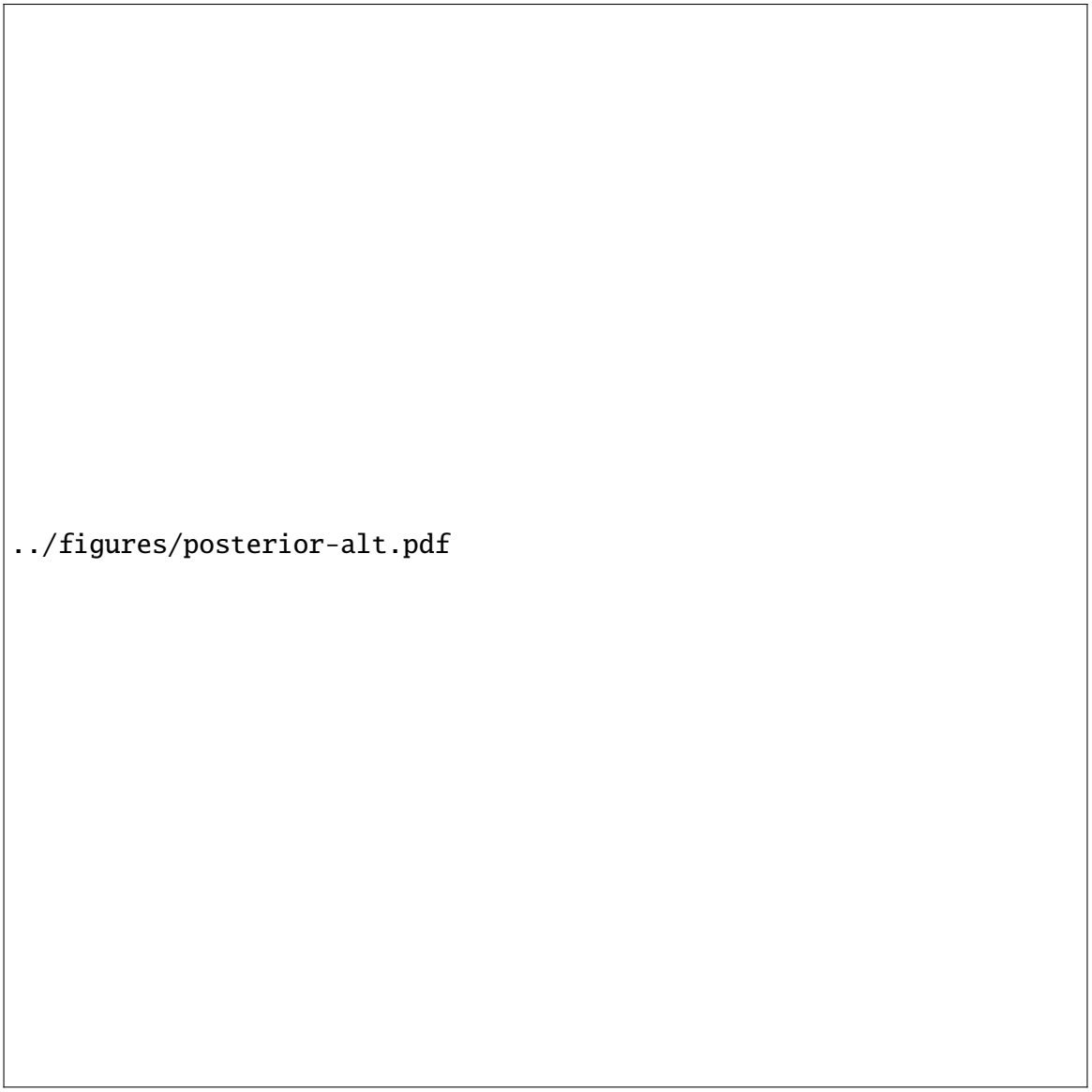


Figure S8: Posterior distribution of parameters inferred with the extended prior distribution. On the diagonal, the inferred posterior distribution of each model parameter. Below the diagonal, the inferred joint posterior distribution of pairs of model parameters (dark purple and bright yellow for low and high density, respectively). Red markers and orange lines for the joint MAP estimate (which may differ from the marginal MAP, as the marginal distribution integrates over all other parameters).

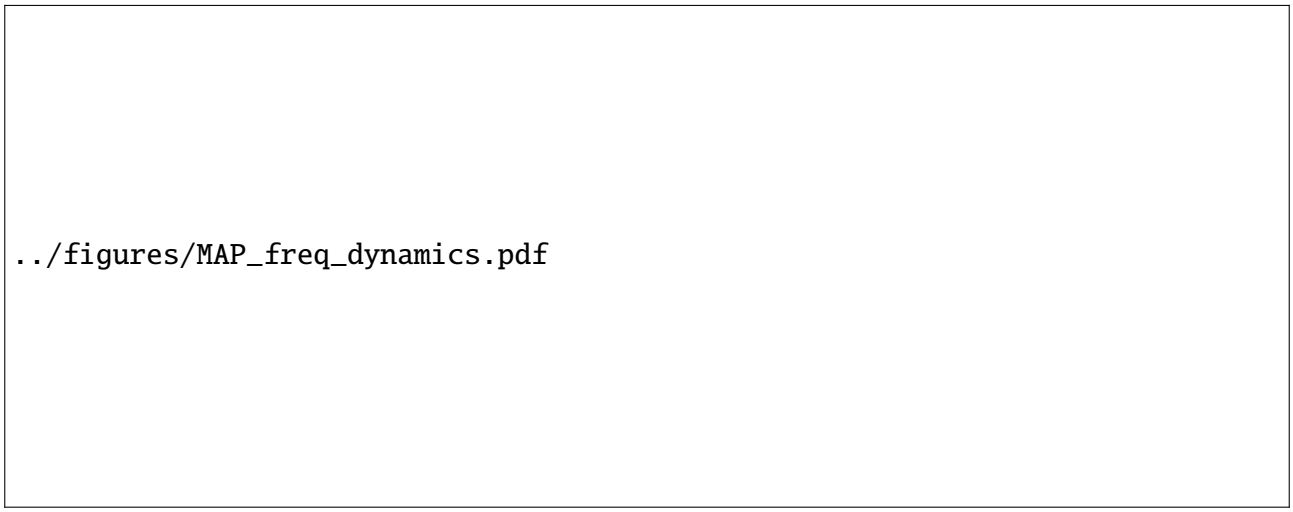


Figure S9: Posterior predicted genotype frequencies in log-log scale. Frequency dynamics of the different genotypes with MAP parameter estimates, same as Figure 5A, but in log-log scale. Black dashed curves for a deterministic model without genetic drift. Clearly, appearance of *2n+1* and $2n^*_M$ is deterministic. Appearance of *2n+1**, and therefore $2n^*_A$, is stochastic, however, the frequency dynamics are deterministic above a frequency of roughly 0.001.

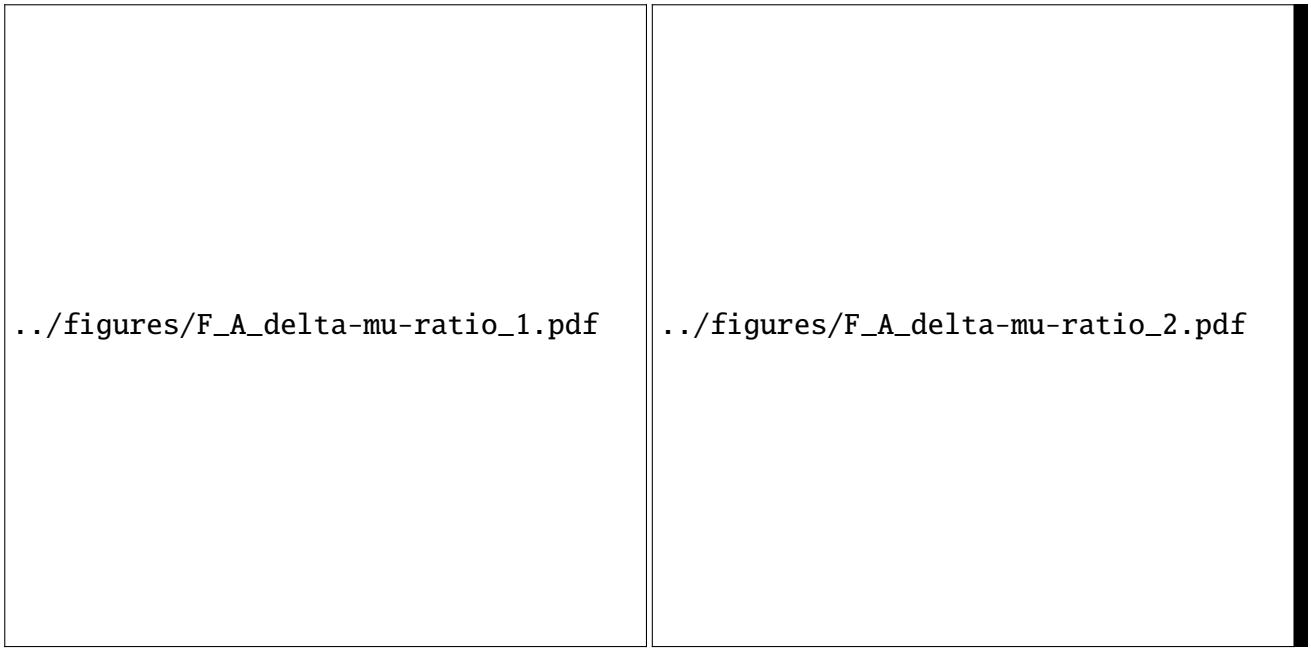


Figure S10: Posterior distribution of F_A. (A,B) F_A values (color coded) as in Figure 5 for different parameter choices on the x- and y-axes. White star denotes the MAP estimate.