# Modeling the contributions of chromosome segregation errors and aneuploidy to *Saccharomyces* hybrid sterility

Primrose J. Boynton[1,*], Thijs Janzen[2,3], Duncan Greig[4]

1) Environmental Genomics Group, Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany

2) Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany

3) Institute for Biology and Environmental Sciences, Carl von Ossietzky University, 26111 Oldenburg, Germany

4) Department of Genetics, Evolution, and Environment, University College London, London WC1E 6BT, United Kingdom

* Corresponding author:

Primrose J. Boynton

Max Planck Institute for Evolutionary Biology

August-Thienemann-Str. 2

24306 Plön, Germany

+49 4522 763-237

pboynton@evolbio.mpg.de

Running head: *Saccharomyces* hybrid meiosis

Keywords: Hybrid, Meiosis, Postzygotic reproductive isolation, Disomy, Ascospore

Abstract

Errors in meiosis can be important postzygotic barriers between different species. In *Saccharomyces* hybrids, chromosomal missegregation during meiosis I produces gametes with missing or extra chromosomes. Gametes with missing chromosomes are inviable, but we do not understand how extra chromosomes (disomies) influence hybrid gamete inviability. We designed a model predicting rates of missegregation in interspecific hybrid meioses assuming several different mechanisms of disomy tolerance, and compared predictions from the model to observations of sterility in hybrids between *Saccharomyces* yeast species. Sterility observations were consistent with the hypothesis that chromosomal missegregation causes hybrid sterility, and the model indicated that missegregation probabilities of 13-50% per chromosome can cause observed values of 90-99% hybrid sterility regardless of how cells tolerate disomies. Missing chromosomes in gametes are responsible for most infertility, but disomies may kill as many as 11% of the gametes produced by hybrids between *S. cerevisiae* and *S. paradoxus*.

Introduction

The yeast *Saccharomyces cerevisiae* and its close relatives are models for studying postzygotic mechanisms of reproductive isolation. Haploid cells from different *Saccharomyces* species readily fuse to form F1 diploids in laboratory crosses, but the hybrid diploids have low fertility: when they undergo meiosis, 81-99% of the resulting gametes fail to form visible colonies and are considered inviable (Hunter *et al.*, 1996; Greig *et al.*, 2002a; Delneri *et al.*, 2003; Libkind *et al.*, 2011; Xu and He, 2011; Almeida *et al.*, 2014). Chromosomal missegregation during hybrid meiosis, caused by meiosis I non-disjunction, is a key cause of *Saccharomyces* hybrid sterility (Chambers *et al.*, 1996; Hunter *et al.*, 1996;

Greig *et al.*, 2002b), but we do not know how much of observed hybrid sterility is caused by chromosomal missegregation and how much by other mechanisms.

When diploid *Saccharomyces* cells are starved, they undergo meiosis to produce haploid gametes (Neiman, 2005). Under normal conditions, a yeast cell duplicates each of its chromatids before meiosis, the four chromatids from both parents align with one another during meiotic prophase I, and homologous chromosomes cross over (Figure 1a). Crossovers hold the chromatids in position relative to one another before the meiotic spindle apparatus can pull two of each chromatid to opposite poles of the dividing cell (reviewed in Petronczki *et al.*, 2003). The daughter cells then undergo a second division, meiosis II, and produce four haploid gametes (ascospores) enclosed in a sac (ascus, plural asci) derived from the mother cell (Petronczki *et al.*, 2003; Neiman, 2005). Healthy gametes germinate and can grow mitotically, producing more haploid cells, and each haploid cell can fuse with another compatible haploid cell to form a diploid, completing the *Saccharomyces* life cycle.

Meiotic crossing-over is hindered in interspecific *Saccharomyces* hybrids, resulting in meiosis I non-disjunction, chromosome missegregation, and the production of aneuploid gametes (gametes with non-haploid chromosome complements). *Saccharomyces* species have the same number (sixteen) of mostly syntenic nuclear chromosomes (Greig *et al.*, 2002a; Kellis *et al.*, 2003; Scannell *et al.*, 2011; Liti *et al.*, 2013; Baker *et al.*, 2015). However, anti-recombination proteins, including the mismatch repair system, inhibit crossing over between homeologous chromosomes (homologous chromosomes from different parental species) where sequence homology is low (Hunter *et al.*, 1996). Without crossovers, chromosomes can missegregate during meiosis I: the meiotic spindle apparatus can pull all four of a chromosome pair's chromatids to one side of the dividing cell (Hunter *et al.*, 1996; Greig *et al.*, 2003) (Figure 1b). Following meiosis II, all four of the resulting gametes are aneuploid, inheriting either zero or two copies of the missegregated chromosome. Evidence for this

mechanism comes from the low recombination rates and high frequencies of extra chromosomes (disomies) observed in rare viable hybrid gametes, and from the fact that knocking out anti-recombination proteins increases both crossing over and overall hybrid gamete viability (Hunter *et al.*, 1996).

Aneuploid gametes lacking a chromosome are inviable because all chromosomes are essential, and aneuploid gametes with disomies may have decreased viability and fitness (Mulla *et al.*, 2013; Santaguida and Amon, 2015). If disomic gametes mate, the resulting F2 zygotes will also be aneuploid, containing combinations of disomic, trisomic, and tetrasomic chromosomes. Many F2 hybrids are reproductively isolated from both parental species, perhaps because they have different karyotypes from their parents (Greig *et al.*, 2002b). It is not known whether missegregation only kills gametes lacking one or more chromosomes or whether disomies also reduce gamete viability. Whilst some hybrid gametes are viable despite carrying disomic chromosomes (Hunter *et al.*, 1996; Greig *et al.*, 2002b), this does not rule out the possibility that hybrid gametes can be killed by disomies.

The consequences of disomies have been more thoroughly studied in non-hybrid *S. cerevisiae* than in interspecies *Saccharomyces* hybrids. For example, *S. cerevisiae* gametes carrying large numbers of disomies can be generated by inducing meiosis in triploid cells, so that six copies of each homologue must segregate into four gametes (Charles *et al.*, 2010; Zhu *et al.*, 2012). Although *S. cerevisiae* triploids do have reduced fertility, producing only about 50% viable gametes, it is not clear whether gamete inviability is due to disomies or something else. While the number of disomic chromosomes per gamete initially has a random distribution in viable gametes, disomic chromosomes have high mitotic instability and aneuploid gametes rapidly evolve back to euploidy after germination (St. Charles *et al.*, 2010).

Changes in protein stoichiometry and number appear to be the main causes of aneuploidy-related stress in *Saccharomyces* (Torres *et al.*, 2007; Santaguida and Amon, 2015). Proteins coded by extra chromosomes can overwhelm a cell's protein processing machinery and either misfold or not degrade at the appropriate moment, and extra proteins can form toxic aggregates (Oromendia *et al.*, 2012). Cells prevent protein toxicity by chaperoning proteins as they fold and degrading extra or misfolded proteins (reviewed by Dobson, 2003; Goldberg, 2003), but these pathways can be limited if too many excess mRNAs are transcribed (Santaguida and Amon, 2015). Aneuploid cells can also produce unbalanced ratios of proteins in a pathway or protein complex (Torres *et al.*, 2007; Makanae *et al.*, 2013). These dosage incompatibilities negatively impact fitness and viability when unused proteins are toxic or when unbalanced subunit ratios prevent a protein complex from correctly assembling (Papp *et al.*, 2003; Veitia *et al.*, 2008). A classic example of dosage incompatibility in a two-protein system is the interaction between α- and β-tubulin, which together form microtubules (McKean *et al.*, 2001). When β-tubulin is overexpressed, as in a cell disomic for Chromosome VI, microtubules do not form efficiently, excess β-tubulin subunits form aggregates, and cells die (Burke *et al.*, 1989; Weinstein and Solomon, 1990).

Given the reduced fertility of *S. cerevisiae* triploids and the negative effect of many disomies on growth of non-hybrid *Saccharomyces* cells, it is likely that missegregation generates inviable disomic gametes in addition to gametes lacking essential chromosomes. The ideal way to investigate the relative contributions of extra or lacking chromosomes to gamete inviability would be to directly measure the chromosome missegregation rate. Unfortunately, missegregation rates cannot be measured directly because the chromosome complement of inviable gametes cannot easily be determined. Further, inferring the chromosome complements of dead gametes from the chromosome complements of surviving gametes is also problematic, if, as in *S. cerevisiae*, viable hybrid aneuploids rapidly revert to

euploidy under mitotic growth. We therefore developed a mathematical model to determine the possible contributions that disomies could make to hybrid sterility at a given chromosome missegregation rate, under different models of disomy tolerance.

The model predicts the proportion of surviving offspring and their karyotypes after a population of genetically identical diploid F1 hybrids undergoes meiosis. Given a per-chromosome missegregation probability, it first calculates the frequencies of numbers of disomies per cell after meiosis I. It then predicts the proportions of surviving cells with each disomy number for a given set of assumptions about how cells tolerate disomies. We included three hypotheses modeling increased protein toxicity with increasing numbers of extra chromosomes with stepwise, additive, or multiplicative relationships between disomy number and gamete death, and one hypothesis modeling protein dosage incompatibility where an imbalance in protein dosages in two-protein systems is completely lethal to cells.

The model was designed with interspecific crosses between *Saccharomyces* yeasts in mind, but it can predict gamete inviability for any meiosis, as long as chromosomal missegregation is the only cause of gamete inviability. We expect the model to be most useful for researchers interested in hybrid meiosis because chromosomal missegregation is hypothesized to be a major contributor to hybrid infertility, but it may also be useful for other researchers interested in missegregation. The model does not account for other potential causes of postzygotic reproductive isolation. For example, chromosomal rearrangements, incompatible parental gene combinations (Dobzhansky-Muller incompatibilities), and reductions in gamete mitotic growth rates are not modeled (Fischer *et al*., 2000; Greig *et al.*, 2002a; Greig, 2007; Kao *et al.*, 2010; Xu and He, 2011; Hou *et al.*, 2014). The intention of the model is to identify realistic missegregation parameters and disomy intolerance mechanisms. Its predictions can then be used as a starting point for investigating further reproductive isolation mechanisms in *Saccharomyces* and other taxonomic groups. The

missegregation model is only for diploid cells undergoing meiosis to produce haploid gametes. It quantifies the frequency of aneuploid cells produced that are inviable due to missing chromosomes and the frequency of aneuploid cells produced that are inviable due to extra chromosomes (disomies).

## Materials and Methods

### *Missegregation model*

Assuming all chromosomes missegregate with equal probability and independently of one another (Campbell *et al.*, 1981), the probability M(*k, n, p*) that a meiosis produces *k* missegregations out of *n* total chromosomes, given a missegregation probability *p*, follows a binomial distribution (see Table 1 for a summary of variable definitions):

$$\mathrm{M}(k,n,p) = \binom{n}{k} p^k (1-p)^{n-k}$$

After meiosis I is complete, each missegregated chromosome will either be missing from a daughter cell or present as two copies, with a 50% probability of each outcome. Each properly segregated chromosome will be present as a single copy. The proportion of meiotic offspring with *k* missegregations with one or two copies of every chromosome (*i.e.*, a full complement of chromosomes) is the probability of every missegregated chromosome migrating into the same daughter cell:

$$\mathrm{S}(k) = 0.5^k$$

We assume that a cell with a missing chromosome is inviable, so all cells without a full complement of chromosomes die. After meiosis, the proportion of gametes in the population with a full complement of chromosomes is the sum of proportions of cells containing each number of disomies:

$$s_{meiosis} = \sum_{k=0}^{n} \mathrm{M}(k,n,p) \times \mathrm{S}(k)$$

where every surviving cell has *k* disomic chromosomes and (*n* - *k*) single copy chromosomes. If disomies do not decrease gamete viability, then $s_{meiosis}$ is the proportion of surviving gametes in the population.

If disomies do decrease gamete viability, the distribution of disomy numbers in the population will change before colonies are detected. The total number of surviving gametes will still be the sum of proportions of cells with each disomy number, but disomy numbers will be adjusted to account for gamete deaths due to extra chromosomes:

$$s = \sum_{k=0}^{n} \mathrm{M}(k, n, p) \times \mathrm{S}(k) \times \mathrm{T}(k, n, a)$$

Where T(*k*, *n*, *a*) is the proportion of gametes with *k* disomies out of *n* chromosomes that can survive to germination, assuming a disomy tolerance function with parameter *a*.

We model disomy tolerance below based on four hypotheses: 1) cells may tolerate only a certain number or fewer disomies (step disomy tolerance, Figure 2a); 2) each disomy may impart a fixed survival cost to a cell (additive disomy tolerance, Figure 2b); 3) each disomy may impart a survival cost to a cell relative to the total number of disomies (multiplicative disomy tolerance, Figure 2c); or 4) disomy tolerance may be a function of dosage-dependent interactions between genes on different chromosomes (dosage incompatibility disomy tolerance, Figure 2d).

For all hypotheses, we assume that cells tolerate disomies in a symmetric manner with regard to euploidy. For example, we assume the probability of germination and survival is the same when *k* = 1 and *k* = (*n*-1) because both values of *k* are one chromosome different from a full set of haploid or diploid chromosomes. To fulfill this requirement for some disomy tolerance models, we defined a transformation of *k* to reflect the number of chromosomes that would need to be added to or subtracted from a cell's chromosome complement in order for the cell to have a full set of chromosomes:

$$k_{sym} = \begin{cases} k \; if \; k \leq \frac{n}{2} \\ n - k \; if \; k > \frac{n}{2} \end{cases}$$

*Step disomy tolerance*

The simplest disomy tolerance assumption is that disomies are completely tolerated up to a given threshold ($a_{step}$), after which they are lethal (Figure 2a). This assumption would be realistic if the cellular pathways compensating for extra or misfolded proteins function perfectly until a certain threshold of extra proteins is reached, at which point they can no longer prevent cell death. Under this assumption, all cells with $a_{step}$ or fewer disomies survive, while all cells with more than $a_{step}$ disomies die. In other words:

$$\mathrm{T}_{step} = \begin{cases} 1 \; if \; k_{sym} \leq a_{step} \\ 0 \; otherwise \end{cases}$$

Note that any cell tolerating half the chromosome number or more ($a_{step} \geq (n/2)$) will tolerate all possible numbers of disomies.

*Additive and multiplicative disomy tolerance*

Instead of a cell only tolerating a fixed number of disomies, each disomy may have a fixed cost ($a$) to a cell's survival probability. The "additive" and "multiplicative" disomy intolerance hypotheses (Figures 2b, c) assume that each extra chromosome causes some proteotoxic stress that decreases a gamete's survival probability by a fixed amount. For the additive hypothesis, the stresses are different for each chromosome, and for the multiplicative hypothesis, the stresses are independent of one another and can overlap.

When costs are additive, the cost ($a_{add}$) of each additional disomy is relative to the survival probability of a completely euploid cell (Figure 2b). Additive disomy tolerance can be described by the equation:

$$T_{add} = \begin{cases} 1 - (a_{add} \times k_{sym}) \; if \; 1 - (a_{add} \times k_{sym}) > 0 \\ 0 \; otherwise \end{cases}$$

When costs are multiplicative, the cost ($a_{mult}$) of each additional disomy is relative to the survival probability of a cell without the disomy (Figure 2c). In other words, the cost of a cell's $k^{th}$ disomy is relative to the survival probability of a cell with ($k$-1) disomies. Multiplicative disomy tolerance can be described by the equation:

$$T_{mult} = (1 - a_{mult})^{k_{sym}}$$

*Dosage incompatibility disomy tolerance*

Finally, disomy may be lethal when interacting genes lie on chromosomes present in different numbers in a cell (Makanae *et al.* 2013). If the genome contains a few pairs of genes with lethal dosage-dependent gene incompatibilities, the probability of gamete death depends on the probability of one gene in a pair lying on a disomic chromosome while the other gene lies on a single copy chromosome. If both genes in a pair lie on disomic chromosomes, or both on single copy chromosomes, then the relative dosages of the two genes is unchanged from those of a fully euploid cell. We assume that each gene in every dosage-dependent incompatibility pair is randomly and independently placed on one of $n$ chromosomes, but never on the same chromosome, and that all instances of dosage incompatibility are lethal. For $a_{inc}$ pairs of dosage-dependent genes (Figure 2d):

$$T_{inc} = (P(both\ genes\ lie\ on\ disomic\ chromosomes) + P(both\ genes\ lie\ on\ single\ copy\ chromosomes))^{a_{inc}}$$

$$T_{inc} = \left(\frac{k}{n} \times \frac{(k-1)}{(n-1)} + \frac{(n-k)}{n} \times \frac{(n-k-1)}{(n-1)}\right)^{a_{inc}}$$

$$T_{inc} = \left(\frac{(n-k)(n-k-1) + k(k-1)}{n(n-1)}\right)^{a_{inc}}$$

*Model fitting*

We attempted to validate the model by comparing its predictions against published observations of hybrid gamete inviability and surviving gamete karyotypes. While the total proportion of surviving *Saccharomyces* hybrid gametes ranges from 1% to 19% depending on the species crossed, most interspecific *Saccharomyces* F1 hybrids produce close to 1% viable gametes, especially in well-studied crosses between *S. cerevisiae* and *S. paradoxus* (Hunter *et al.* 1996; Greig *et al.* 2002a; Delneri *et al.* 2003; Libkind *et al.* 2011; Xu and He 2011; Almeida *et al.* 2014). We report model predictions below for 1% and 10% gamete viabilities.

The model also can be fit to distributions of numbers of disomies per surviving gamete, but there are no data available in the literature accurately reporting the karyotypes of interspecific *Saccharomyces* hybrid gametes. These data are difficult, if not impossible, to collect because we expect extra chromosomes to be rapidly lost during mitosis before karyotypes can be measured. With these caveats in mind, we did fit the model to one published dataset of gamete karyotypes produced from a *S. cerevisiae*-*S. paradoxus* F1 hybrid (Xu and He, 2011) to help establish a lower bound of reasonable missegregation rates and to help identify realistic and unrealistic disomy tolerance hypotheses.

The authors of the dataset genotyped colonies produced by 94 haploid gametes from a single interspecies cross at 93 loci distributed among all sixteen chromosomes. We inferred a disomy when the authors detected alleles from both parents at at least one locus on a chromosome. In some cases, the authors detected alleles from both parents at one locus, but only a single allele at another locus on the same chromosome. We assumed that these cases represented disomies where either the authors did not detect both alleles at all loci on the chromosome or the chromosome did experience crossing-over and recombination, but the crossing-over did not prevent missegregation. Regardless, we expect the data to represent a considerable underestimate of overall gamete aneuploidy. The authors reported ~99% gamete

inviability in the entire population, and we therefore assumed that 9306 gametes did not survive the mating.

We used Maximum Likelihood estimation to infer the model parameters *p* and *a* given the data in (Xu and He, 2011). We obtained the likelihood of the model given the data, and we assumed that errors in the data were Poisson distributed because numbers of observed disomies are count data. Comparing the frequency of surviving cells with *k* disomies between the data and the model, the log likelihood is:

$$LL = \sum_{k=1}^{n} s_k \log \lambda_k - \lambda_k - \log \Gamma(s_k + 1)$$

Where $\lambda_k$ is the proportion of surviving cells with *k* disomies in the data and $s_k$ is the proportion of surviving cells with *k* disomies in the model (*i.e.*, $s_k = \mathrm{M}(k,n,p) \times \mathrm{S}(k) \times \mathrm{T}(k,n,a)$); the log likelihood represents the probability of observing $s_k$ surviving cells in the model, given $\lambda_k$ surviving cells in the data. For $\lambda_k = 0$, the log likelihood is 0 by default. Using the log likelihood, we calculated Aikaike's Information Criterion (AIC) (Aikaike, 1974) of each model:

$$\mathrm{AIC} = 2df - 2LL$$

The degrees of freedom for all disomy tolerance functions were two: the misssegregation parameter *p* and the aneuploidy tolerance parameter *a*.

*Model implementation*

The model and model fitting procedures were implemented in R version 3.3.1 (R Core Development Team, 2016). Figures were produced using the stats, graphics, ggplot2, and colorspace packages (Wickham, 2009; Ihaka *et al.*, 2015; R Core Development Team, 2016). Functions implementing the model are included in the supporting information (File S1).

Results

Both missegregation probability (*p*) and disomy intolerance (*a*) influence the proportion of surviving gametes (Figure 3). Here we report predictions for meiosis with a haploid chromosome number of ($n$ = 16), the same as that of *Saccharomyces spp*. The model predicts that the proportion of surviving gametes (*s*) is higher when disomies are well tolerated than it is when disomies are poorly tolerated. When disomies are well tolerated, the proportion of surviving gametes is higher at low probabilities of missegregation than at high probabilities of missegregation. Survival probability always decreases as missegregation probability increases from zero to one-half. Missegregation probabilities above one-half are unlikely to occur because homeologous chromosomes would have to migrate to the same pole of a dividing cell more frequently than would be expected by chance. We include predictions for high missegregation probabilities for the sake of completeness only (Figure 3), and note that when cells tolerate very few disomies, survival probability can increase slightly at very high missegregation probabilities (above about 0.7). Most of these surviving meiosis I products would be diploid or nearly diploid cells, which are tolerated as well as haploid or nearly haploid cells.

At survival probabilities similar to those observed in interspecific *Saccharomyces* hybrids (1-10%), most sterility is due to missing chromosomes, rather than additional chromosomes (Figure 4). If disomies are never tolerated, the rate of missegregation required to give 1% spore viability is just 0.25, with 88% of gametes dying because of missing chromosomes and 11% dying due to disomies. If disomies are always tolerated, the rate of missegregation must increase to 0.50 and all cell death is due to missing chromosomes.

We parameterized the model based on a published dataset of hybrid gamete-derived colony karyotypes (Xu and He, 2011). As discussed above, the reported karyotypes are likely underestimates of disomy frequencies in surviving gametes because chromosomes were

likely lost after gamete germination; as a result, our parameter estimates are also likely underestimates of missegregation probability and disomy tolerance. With these caveats in mind, the best-fitting model predicts multiplicative disomy tolerance (*i.e.*, the cost of each additional disomy is relative to the survival probability of a cell without the disomy) with a missegregation probability ($p$) of 0.35 and a proportional disomy cost ($a_{mult}$) of 0.42 (supporting information Figure S1, Table S1). In other words, 35% of all chromosomes are predicted to missegregate and a gamete's viability is predicted to decrease by 42% per additional disomy. In crosses with 16 haploid chromosomes, a missegregation probability of 0.35, and a survival proportion of 1%, 95.4% of all gametes are inviable due to missing chromosomes and an additional 3.6% of all gametes are inviable due to disomy intolerance (Figure 4a). Given the fitted data, the estimate of missegregation probability was robust to our disomy tolerance assumptions, and was between 0.35 and 0.37 for the three best-fitting models (supporting information Table S1).

## Discussion

### *Predictions of hybrid gamete viability*

Empirical observations of hybrid gamete viability and aneuploidy are consistent with *Saccharomyces* reproductive isolation through meiotic chromosomal missegregation. Given observed hybrid gamete viability values of 1-10%, our missegregation model predicts that missing chromosomes are responsible for most gamete inviability (Figure 4). Missegregation probabilities between 0.13 and 0.50 result in the deaths of 90-99% of gametes; this range probably encompasses missegregation probabilities for most *Saccharomyces* hybrid meioses in laboratory crosses. If chromosome missegregation in meiosis I were completely random (*i.e.*, equal to ½), our model would predict 1% gamete viability or less regardless of how disomies are tolerated (Figures 3, 4). In laboratory crosses with 1% gamete viability, either

chromosome migration is completely random and disomies are always tolerated or chromosomes sometimes missegregate and disomies are sometimes deadly for gametes (Figure 4a).

Our model is similar to a previously published model predicting the number of surviving gametes per ascus in intraspecific *S. cerevisiae* crosses (Chu *et al.*, 2016). Like us, these authors based their predictions on a binomial distribution of chromosomal missegregations; they also assumed that gametes tolerated disomies in the same way we did with our multiplicative disomy intolerance hypothesis. Unlike us, they focused on causes of spore inviability in intraspecific crosses where spore viability is high (>75%) relative to interspecific crosses (<10%). The goal of their study was to contrast the relative contributions of missegregation and random gamete death to gamete inviability in crosses with low inviability. They did not explicitly contrast the relative contributions of missing chromosomes and disomies to cell death, nor did they explore how different hypotheses about disomy tolerance influence gamete inviability. Indeed, our model demonstrates that a given missegregation probability can result in different rates of gamete inviability, depending on how a gamete tolerates disomies (Figures 3, 4).

*Disomy intolerance in hybrid gametes*

Of the four disomy intolerance hypotheses investigated in the model, we suspect that the multiplicative and dosage incompatibility hypotheses are more realistic than the step and additive hypotheses. The multiplicative and dosage incompatibility hypotheses were better fits to empirical observations of surviving aneuploid gametes, although as discussed above and below, the empirical observations are problematic. If fits to the empirical data accurately represent the biological processes responsible for disomy intolerance, then realistic parameters for disomy intolerance include either about a 42% reduction in probability of

gamete survival per disomy or about four pairs of genes on different *Saccharomyces* chromosomes that are completely lethal when expressed at different levels.

The quality of the predictions of each disomy tolerance hypothesis depends on the accuracy of its assumptions and the cellular processes hypothesized. For example, the step hypothesis would be consistent if cells have a fixed amount of protein processing machinery (e.g., a fixed number of protein chaperones) that do not scale with chromosome number, even as harmful extra proteins do scale with chromosome number. Both the step and additive hypotheses would be consistent if all harmful extra proteins in a disomic cell are equivalent to one another and accumulate in a quantitative manner. In contrast, if each extra protein in a disomic cell had an independent effect on a cell function, the multiplicative hypothesis would be more realistic. Future work studying how protein toxicity may cause cell death in aneuploid cells is needed to understand which of our disomy tolerance hypotheses is most realistic.

The dosage incompatibility disomy tolerance model assumes that cell inviability is a result of pairs of proteins where a dosage incompatibility is completely lethal. A recent empirical study has shown that fitness costs of aneuploidy are likely to be caused by many gene dosage incompatibilities with small fitness costs (Bonney *et al.*, 2015). A more realistic dosage incompatibility model would model large numbers of sets of incompatible genes with small effects on cell viability, although we chose not to include such a model because it would require more than one disomy intolerance parameter. In reality, inviability of disomic gametes is most likely the result of a combination of protein toxicity and gene dosage incompatibilities.

*Data and model limitations*

We fit the model's predictions to the only available dataset reporting disomy numbers and percent inviability for a large number of gametes (Xu and He, 2011), but we expect this dataset to be biased against cells with many disomies. The authors of the dataset did not intend to study aneuploidy; instead, chromosome copy number data were incidentally collected during an investigation of hybrid recombination rate and gene incompatibility. The amount of time between gamete germination and chromosome copy number inference was likely long enough for the population of cells to lose disomies: karyotypes in the dataset were assayed after at least four days of mitotic growth and at least one post-germination single-cell bottleneck. In addition to not detecting disomies lost during mitotic growth, this dataset probably underestimates disomy frequencies because the authors inferred aneuploidy based on markers located throughout the *Saccharomyces* genome instead of whole-genome sequencing, and because they discarded all gametes disomic for chromosomes X and XVI. A better dataset to fit to our model would explicitly measure karyotypes of gametes after as few mitotic divisions as possible, or, ideally, before mitosis begins, but to our knowledge no such data have been published.

The model reported here also makes simplifying assumptions about missegregation and disomy tolerance. We assumed that every chromosome in a cell has the same probability of missegregation and every disomy has the same influence on viability. It seems likely that these assumptions are an oversimplification. For example, missegregation has correlated with chromosome length in some studies but not others (Chu *et al.*, 2016). When we relaxed the assumption of equal missegregation probabilities for every chromosome, we found that the proportion of surviving gametes was less for cells with large among-chromosome variations in missegregation probabilitiy compared to cells with small or no variation. However, the relationship between average missegregation probability and gamete survival had a similar shape regardless of missegregation probability variation (supporting information Figure S2).

In the empirical data fit to the model, some disomies were overrepresented (Xu and He, 2011) (supporting information Figure S3). However, we do not know if differences in individual chromosomes' incidence of disomy are due to nonrandom missegregation and disomy tolerance, or due to nonrandom chromosome loss during mitosis.

*Other mechanisms of reproductive isolation*

The purpose of this model was to establish realistic predictions about the results of meiosis when missegregation causes reproductive isolation between *Saccharomyces* species, but missegregation is not the only possible mechanism of *Saccharomyces* reproductive isolation. Other postzygotic and prezygotic mechanisms may be equally important for maintaining *Saccharomyces* species in nature, and it is possible that multiple mechanisms work together to reduce hybrid fertility. For example, because the missegregation frequency must be 50% or less, additional postzygotic isolation mechanisms must come into play if hybrid gamete inviabilities above 99% are observed.

In addition to chromosomal missegregation, chromosomal rearrangements, Dobzhansky-Muller gene incompatibilities, and low hybrid mitotic viability and fitness are postzygotic mechanisms that can isolate *Saccharomyces* species. Chromosomal rearrangements are rare but not absent among *Saccharomyces* species; collinearity is mostly conserved, but some species contain inversions or reciprocal translocations with respect to one another (Fischer *et al.*, 2000; Kellis *et al.*, 2003; Liti *et al.*, 2006; Boynton and Greig, 2014). The gametes of crosses between parents with some chromosomal rearrangements are inviable, but can be rescued by genetic manipulations restoring collinearity (Delneri *et al.*, 2003). Additionally, Dobzhanksy-Muller incompatibilities (*i.e.*, incompatibilities between genes from different parents) have been detected between *Saccharomyces* nuclear and mitochondrial genes which can prevent sporulation, so that hybrid gametes cannot be

produced (Lee *et al.*, 2008; Chou *et al.*, 2010). Dobzhansky-Muller incompatibilities between *S. cerevisiae* nuclear genes have also evolved in an experimental system: hybrid diploids from parents evolved under different stressful conditions had lower fitness and produced fewer gametes than nonhybrids, although gamete viability was not affected (Dettman *et al.*, 2007; Anderson *et al.*, 2010). Low hybrid fitness can also come about when hybridization breaks up locally adapted gene combinations (Edmands, 2002). Conversely, hybridization can bring genes together that increase hybrid fitness relative to parents, especially in novel environments (Stelkens *et al.*, 2014; Clowers *et al.*, 2015; Bernardes *et al.*, 2017). For example, extra chromosomes have the potential to mask Dobzhansky-Muller incompatibilities if both parental copies of a gene are present (Greig, 2007).

While *Saccharomyces* gametes are not strongly prezygotically isolated in laboratory crosses, prezygotic isolation might occur in their natural forest and fermentation habitats. Different sympatric *Saccharomyces* species have different growth rates at different temperatures, and researchers have speculated that they are unlikely to be metabolically active at the same time and may therefore never meet and mate (Sweeney *et al.*, 2004; Sampaio and Gonçalves, 2008). Mate choice can also isolate species: when confronted with compatible gametes from multiple species, *Saccharomyces* gametes fuse with gametes of their own species more frequently than they do gametes of another *Saccharomyces* species (Maclean and Greig, 2008).

Under natural conditions, reproductive isolation most likely comes about through a combination of missegregation and other mechanisms. Despite strong reproductive isolation among *Saccharomyces* species, hybrid *Saccharomyces* have been observed outside of the laboratory, especially in domesticated fermentations (Lopandic *et al.*, 2007; González *et al.*, 2008; Sipiczki, 2008). The question remains as to how these hybrids come about and how they are maintained in the population. Further natural history observations are needed to

understand interactions among prezygotic isolation mechanisms, postzygotic isolation mechanisms, and hybrid fitness in maintaining these natural hybrid populations. For example, observations of interactions between *Saccharomyces* and insects indicate that insects may help *Saccharomyces* overcome prezygotic isolation by bringing species together in their guts (Reuter and Greig, 2007; Stefanini *et al.*, 2016). Here, we showed that observed postzygotic isolation of *Saccharomyces* species is consistent with missegregation as a mechanism of reproductive isolation, and that missegregation can isolate species regardless of how gametes tolerate extra chromosomes. We hope that predictions from our model (particularly predictions of biologically realistic missegregation probabilities), predictions from other models, and laboratory experiments can be combined with natural history observations to further understand the natural circumstances that promote and discourage interspecific *Saccharomyces* hybridization.

Acknowledgements

Thank you to Xionglei He for clarifications on the empirical dataset used to fit the model. This research was funded by the Max Planck Society.

References

Akaike H. 1974. A New Look at the Statistical Model Identification. *IEEE Trans Automat Contr* **19**: 716-723.

Almeida P, Gonçalves C, Teixeira S, *et al.* 2014. A Gondwanan imprint on global diversity and domestication of wine and cider yeast *Saccharomyces uvarum*. *Nat Commun* **5**: 5044.

Anderson JB, Funt J, Thompson DA, *et al.* 2010. Determinants of divergent adaptation and Dobzhansky-Muller interaction in experimental yeast populations. *Curr Biol* **20**: 1383-1388.

Baker E,Wang B, Bellora N, *et al.* 2015. The genome sequence of *Saccharomyces eubayanus* and the domestication of later-brewing yeasts. *Mol Biol Evol* **32**: 2818-2831.

Bernardes JP, Stelkens RB, Greig D. 2017. Heterosis in hybrids within and between yeast species. *J Evol Biol* **30**: 538-548.

Bonney ME, Moriya H, Amon A. 2015. Aneuploid proliferation defects in yeast are not driven by copy number changes of a few dosage-sensitive genes. *Genes Dev* **29**: 898-903.

Boynton PJ, Greig D. 2014. The ecology and evolution of non-domesticated Saccharomyces species. *Yeast* **31**: 449-462.

Burke D, Gasdaska P, Hartwell L. 1989. Dominant effects of tubulin overexpression in *Saccharomyces cerevisiae*. *Mol Cell Biol* **9**: 1049-1059.

Campbell D, Doctor JS, Feuersanger JH, Doolittle MM. 1981. Differential mitotic stability of yeast disomies derived from triploid meiosis. *Genetics* **98**: 239-255.

Chambers SR, Hunter H, Louis EJ, Borts RH. 1996. The mismatch repair system reduces meiotic homeologous recombination and stimulates recombination-dependent chromosome loss. *Mol Cell Biol* **16**: 6110-6120.

Chou J, Hung Y, Lin K, *et al.* 2010. Multiple molecular mechanisms cause reproductive isolation between three yeast species. *PLoS Biol* **7**: e1000432.

Chu DB, Burgess SM. 2016. A computational approach to estimating nondisjunction frequency in *Saccharomyces cerevisiae*. *G3* **6**: 669-682.

Clowers KJ, Will JL, Gasch AP. 2015. A unique ecological niche fosters hybridization of oak-tree and vineyard isolates of *Saccharomyces cerevisiae*. *Mol Ecol* **24**: 5886-5898.

Delneri D, Colson I, Grammenoudi S, *et al.* 2003. Engineering evolution to study speciation in yeasts. *Nature* **422**: 68-72.

Dettman JR, Sirjusingh C, Kohn LM, Anderson JB. 2007. Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. *Nature* **447**: 585-588.

Dobson CM. 2003. Protein folding and misfolding. *Nature* **426**: 884-890.

Edmands S. 2002. Does parental divergence predict reproductive compatibility? *Trends Ecol Evol* **17**: 520-527.

Fischer G, James SA, Roberts IN, *et al.* 2000. Chromosomal evolution in *Saccharomyces*. *Nature* **405**: 451-454.

Goldberg AL. 2003. Protein degradation and protection against misfolded or damaged proteins. *Nature* **426**: 895-899.

González SS, Barrio E, Querol A. 2008. Molecular characterization of new natural hybrids of *Saccharomyces cerevisiae* and *S. kudriavzevii* in brewing. *Appl Environ Microbiol* **74**: 2314-2320.

Greig D, Borts RH, Louis EJ, Travisano M. 2002a. Epistasis and hybrid sterility in *Saccharomyces*. *Proc R Soc Lond B* **269**: 1167-1171.

Greig D, Louis EJ, Borts RH, Travisano M. 2002b Hybrid speciation in experimental populations of yeast. *Science* **298**: 1773-1775.

Greig D, Travisano M, Louis E, Borts R. 2003. A role for the mismatch repair system during incipient speciation in Saccharomyces. *J Evol Biol* **16**: 429-437.

Greig D. 2007. A screen for recesive speciation genes expressed in the gametes of F1 hybrid yeast. *PLoS Genet* **3**: e21.

Hou J, Friedrich A, de Montigny J, Schacherer J. 2014. Chromosomal rearrangements as a major mechanism in the onset of reproductive isolation in *Saccharomyces cerevisiae*. *Curr Biol* **24**: 1153-1159.

Hunter N, Chambers SR, Louis EJ, Borts RH. 1996. The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. *EMBO J* **15**: 1726-1733.

Ihaka R, Murrell P, Hornik K, *et al.* 2015. *colorspace: Color Space Manipulation*. R package (version 1.2-6) URL http://CRAN.R-project.org/package=colorspace

Kao KC, Schwartz K, Sherlock G. 2010. A genome-wide analysis reveals no nuclear Dobzhansky-Muller pairs of determinants of speciation between *S. cerevisiae* and *S. paradoxus*, but suggests more complex incompatibilities. *PLoS Genet* **6**: e1001038.

Kellis M, Patterson N, Endrizzi M, *et al.* 2003. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**: 241-254.

Lee H, Chou J, Cheong L, *et al.* 2008. Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. *Cell* **135**: 1065-1073.

Libkind D, Hittinger CT, Valéro E, *et al.* 2011. Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proc Natl Acad Sci U S A* **108**: 14539-14544.

Liti G, Barton DBH, Louis EJ. 2006. Sequence diversity, reproductive isolation and species concepts in *Saccharomyces*. *Genetics* **174**: 839-850.

Liti G, Ba ANN, Blythe M, *et al.* 2013. High quality *de novo* sequencing and assembly of the *Saccharomyces arboricolus* genome *BMC Genomics* **14**: 69.

Lopandic K, Gangle H, Wallner E, *et al.* 2007. Genetically different wine yeasts isolated from Austrian vine-growing regions influence wine aroma differently and contain putative hybrids between *Saccharomyces cerevisiae* and *Saccharomyces kudriavzevii*. *FEMS Yeast Res* **7**: 953-965.

Maclean CJ, Greig D. 2008. Prezygotic reproductive isolation between *Saccharomyces cerevisiae* and *Saccharomyces paradoxus*. *BMC Evol Biol* **8**: 1.

Makanae K, Kintaka R, Makino T, *et al.* 2013. Identification of dosage-sensitive genes in *Saccharomyces cerevisiae* using the genetic tug-of-war method. *Genome Res* **23**: 300-311.

McKean PG, Vaughan S, Gull K. 2001. The extended tubulin superfamily. *J Cell Sci* **114**: 2723-2733.

Mulla W, Zhu J, Li R. 2013. Yeast: a simple model system to study complex phenomena of aneuploidy. *FEMS Microbiol Rev* **38**: 201-212.

Neiman AM. 2005. Ascospore formation in the yeast *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **69**: 565-584.

Oromendia AB, Dodgson SE, Amon A. 2012. Aneuploidy causes proteotoxic stress in yeast. *Genes Dev* **26**: 2696-2708.

Papp B, Pál C, Hurst LD. 2003. Dosage sensitivity and the evolution of gene families in yeast. *Nature* **424**: 194-197.

Petronczki M, Siomos MF, Nasmyth K. 2003. Un ménage à quatre: the molecular biology of chromosome segregation in meiosis. *Cell* **112**: 423-440.

R Core Team. 2016. R: A language and environment for statistical computing (version 3.3.1). R Foundation for Statistical Computing: Vienna, Austria. URL https://www.R-project.org/

Reuter M, Bell G, Greig D. 2007. Increased outbreeding in yeast in response to dispersal by an insect vector. *Curr Biol* **17**: R81-R83.

Sampaio JP, Gonçalves P. 2008. Natural populations of *Saccharomyces kudriavzevii* in Portugal are associated with oak bark and are sympatric with *S. cerevisiae* and *S. paradoxus*. *Appl Environ Microbiol* **74**:2144-2152.

Santaguida S. Amon A. 2015. Short- and long-term effects of chromosome mis-segregation and aneuploidy. *Nat Rev Mol Cell Biol* **16**: 473-485.

Scannell DR, Zill OA, Rokas A, *et al.* 2011. The awesome power of yeast evolutionary genetics: new genome sequences and strain resources for the *Saccharomyces sensu stricto* genus. *G3* **1**: 11-25.

Sipiczki M. 2008. Interspecies hybridization and recombination in *Saccharomyces* wine yeasts. *FEMS Yeast Res* **8**: 996-1007.

St. Charles J, Hamilton ML, Petes TD. 2010. Meiotic chromosome segregation in triploid strains of *Saccharomyces cerevisiae*. *Genetics* **186**: 537–550.

Stefanini I, Dapporto L, Berná L, *et al.* 2016. Social wasps are a *Saccharomyces* mating nest. *Proc Natl Acad Sci U S A* **113**: 2247-2251.

Stelkens RB, Brockhurst MA, Hurst GD, *et al.* 2014. The effect of hybrid transgression on environmental tolerance in experimental yeast crosses. *J Evol Biol* **27**: 2507–2519.

Sweeney JY, Kuehne HA, Sniegowski PD. 2004. Sympatric natural *Saccharomyces cerevisiae* and *S. paradoxus* populations have different thermal growth profiles. *FEMS Yeast Res* **4**: 521-525.

Torres EM, Sokolsky T, Tucker CM, *et al.* 2007. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science* **317**: 916-924.

Veitia RA, Potier MC. 2015. Gene dosage imbalances: action, reaction, and models. *Trends Biochem Sci* **40**: 309-317.

Weinstein B, Solomon F. 1990. Phenotypic consequences of tubulin overproduction in *Saccharomyces cerevisiae*: differences between alpha-tubulin and beta-tubulin. *Mol Cell Biol* **10**: 5295-5304.

Wickham H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag: New York.

Xu M, He X. 2011. Genetic incompatibility dampens hybrid fertility more than hybrid viability: yeast as a case study. *PLoS One* **6**: e18341.

Zhu J, Pavelka N, Bradford WD, *et al.* 2012. Karyotypic determinants of chromosome instability in aneuploid budding yeast. *PLoS Genet* **8**: e1002719.

The authors declare that there is no conflict of interest.

Table 1: Summary of model variables

| variable | explanation |
|---|---|
| $n$ | number of haploid chromosomes in the organism |
| $k$ | number of missegregation events a dividing cell has experienced, number of disomies in a gamete |
| $p$ | probability of a single chromosome pair missegregating |
| $s_{meiosis}$ | proportion of cells surviving meiosis if all disomies are tolerated |
| $s$ | proportion of cells surviving meiosis |
| $k_{sym}$ | difference between the number of disomies in a cell and a full haploid or diploid set of chromosomes |
| $a$ | disomy tolerance; see the text for interpretations of different disomy tolerance parameters |
| $M(k,n,p)$ | probability of meiosis having $k$ missegregation events |
| $S(k)$ | proportion of cells with $k$ missegregations that have a full complement of chromosomes |
| $T(k,n,a)$ | proportion of cells with $k$ disomies that survive to germination |
| $s_k$ | model prediction of proportion of surviving cells with $k$ disomies |
| $\lambda_k$ | observed proportion of surviving cells with $k$ disomies (Xu and He, 2011) |

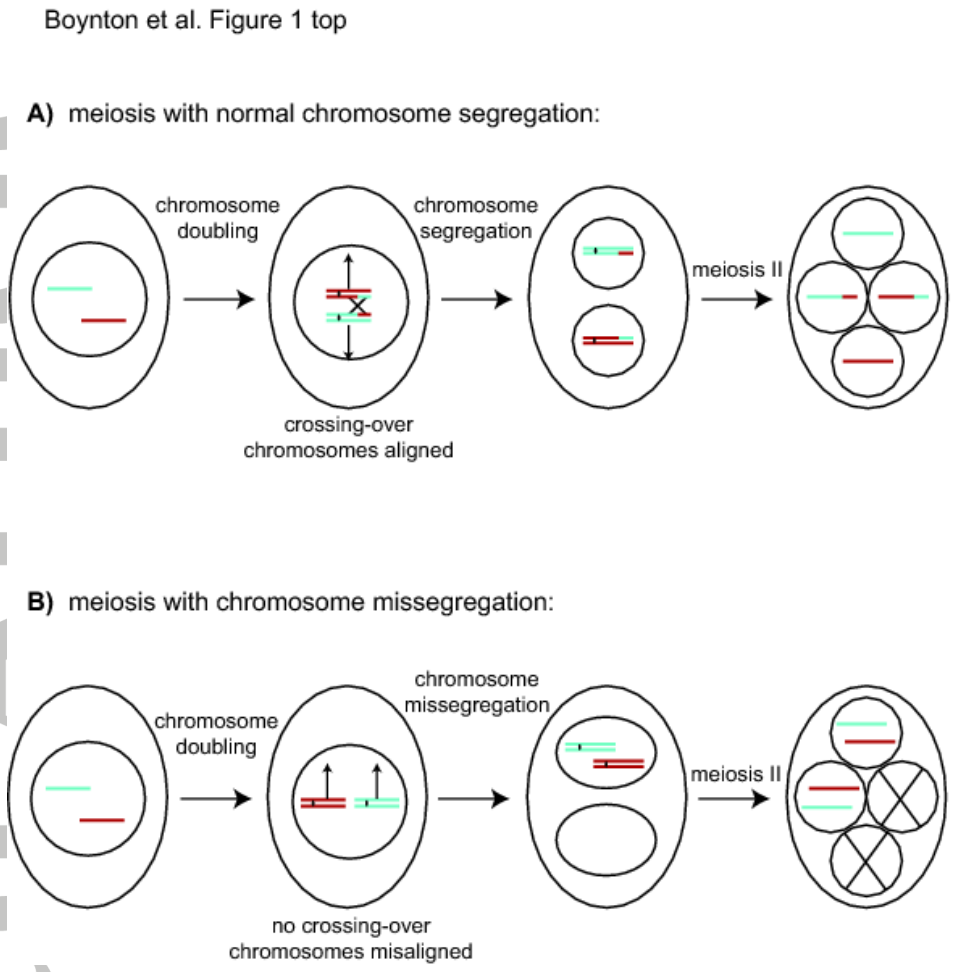


Figure 1: Schematic illustrating meiosis of a diploid cell with a chromosome number of one with A) normal segregation and B) missegregation of the chromosome. Chromosomes from different parents are different colors.

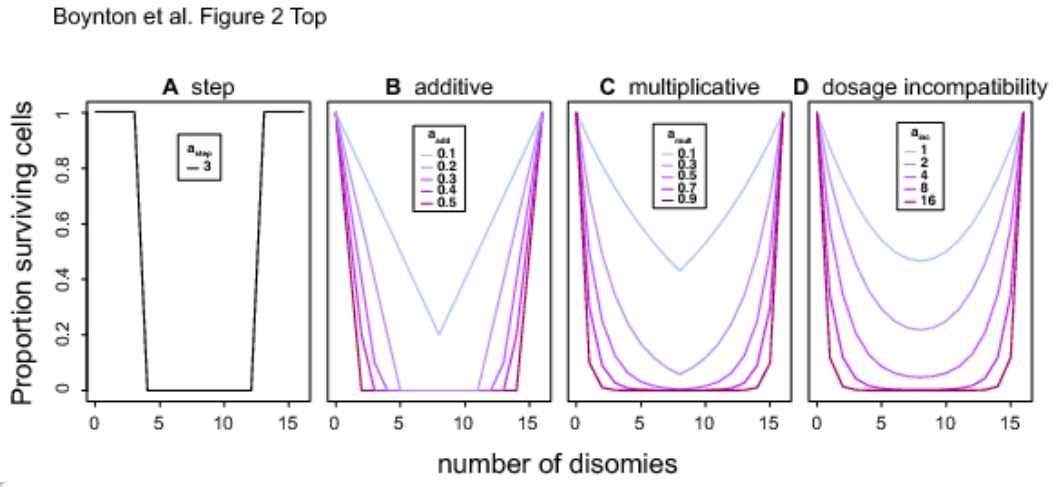


Figure 2: Example relationships between disomy numbers and proportion of cell survival for all four disomy tolerance hypotheses and several disomy parameter values. Haploid chromosome number (*n*) is 16 for all examples.

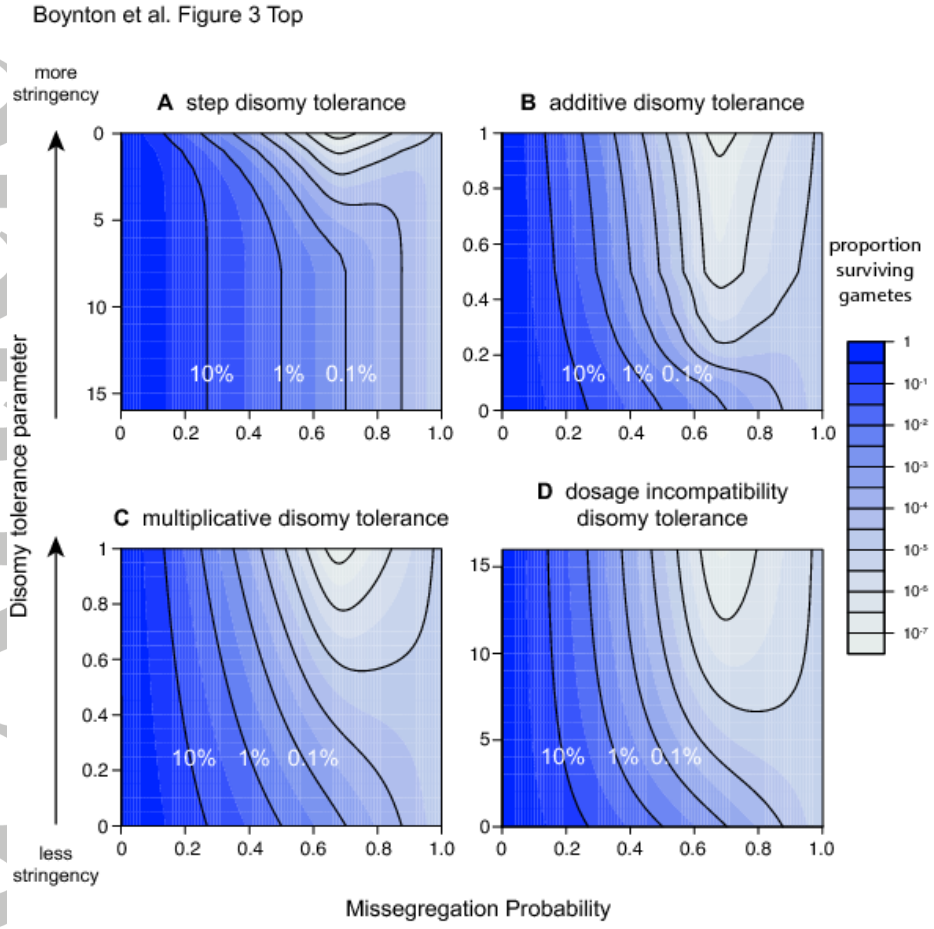


Figure 3: Predicted model meiotic daughter cell survival probabilities for several combinations of missegregation probability and disomy tolerance parameter given each of four disomy tolerance hypotheses. Haploid chromosome number (*n*) is 16, and all possible parameter combinations are depicted for A) step disomy tolerance, B) additive disomy tolerance, and C) multiplicative disomy tolerance. For D) dosage incompatibility disomy tolerance, more dosage-dependent gene pairs than depicted are possible. Black lines depict parameter combinations leading to survival probabilities equal to exponents of 10, and survival probability = 10%, 1%, and 0.1% are indicated in white text.

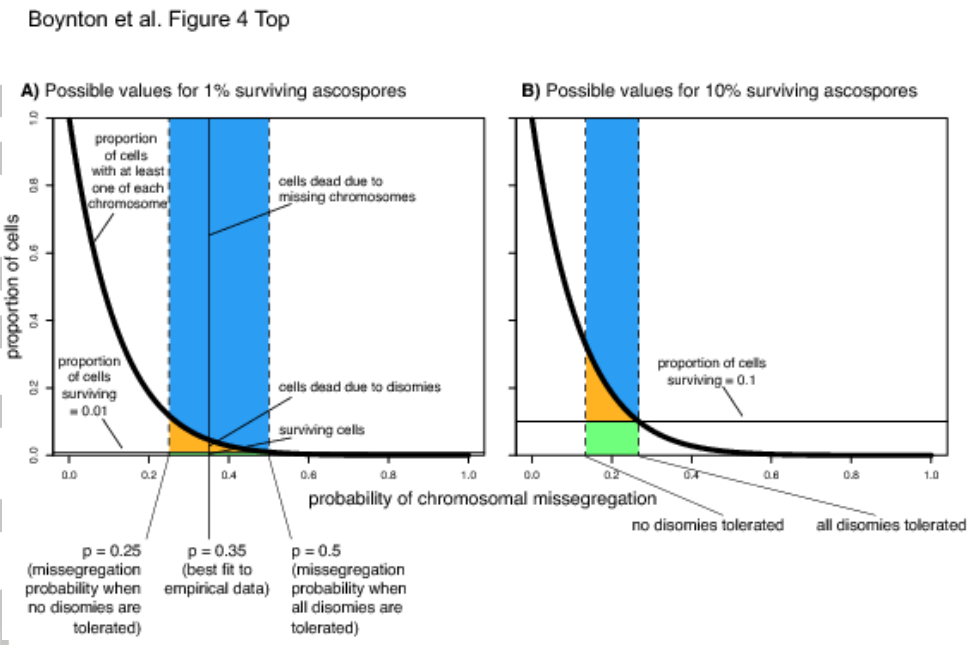


Figure 4: Summary of possible gamete fates and missegregation parameter values when A) 1% and B) 10% of all gametes survive and the haploid chromosome number (*n*) is 16. Possible proportions of cells dead due to missing chromosomes, cells dead due to disomies, and surviving cells are indicated in blue (top shaded region), orange (middle), and green (bottom), respectively. The heavy black line is the proportion of cells with at least one of each chromosome, the thin horizontal line is the proportion of cells surviving, and the vertical dotted lines indicate the ranges of possible missegregation probabilities. The thin vertical line in A) is the best fit to the data of Xu and He (2011) (supporting information Table S1).