Heterogeneity, turnover rate and karyotype space shape susceptibility to missegregation-induced extinction

Gregory J. Kimmel, Richard Beck, Thomas Veith,
Samuel Bakhoum, Philipp M. Altrock, Noemi Andor

December 27, 2021

Abstract

The incidence of somatic copy number alterations (SCNAs) per base pair of the genome is orders of magnitudes larger than that of point mutations. This makes SCNAs phenotypically effective. One mitotic event stands out in its potential to significantly change a cell's SCNA burden – a chromosome missegregation. We present a general deterministic framework for modeling whole chromosome missegregations and use it to evaluate the possibility of missegregation-induced population extinction (MIE). The model predicts critical curves that separate viable from non-viable populations as a function of their turnover- and mis-segregation rates. Missegregation- and turnover rates estimated for nine cancer types are then compared to these predictions. The majority of tumors across all nine cancer types had missegregation- and turnover rates that were within viable regions of the parameter space. When a dependency of mis-segregation rate on ploidy was introduced, ploidy states associated with low mis-segregation rates acted as a stabilizing refuge, rendering MIE impossible when turnover rates are low. Intra-tumor heterogeneity, including heterogeneity in mis-segregation rates, increases as tumors progress. Our predictions suggest that this intra-tumor heterogeneity hampers the chance of success of therapies aimed at MIE.

1 Introduction

Aneuploidy, defined as a chromosome number that is not the exact multiple of the haploid karyotype, is common across several cancers, including non-small-cell lung, breast, colorectal, prostate cancer and glioblastoma [1–5]. The main driver of aneuploidy is chromosomal instability (CIN). CIN-induced genomic changes can be subdivided into two categories: the whole gain or loss of a chromosome (numerical CIN) or changes within localized regions of a chromosome (structural CIN).

Thompson and Compton used live cell imaging to evaluate the fidelity of chromosome segregation, finding missegregation rates ranging from 0.025-1 % per chromosome per mitosis [6]. We distinguish between stochastic and deterministic factors governing a cell's risk to missegregate. Stochastic events include DNA double-strand breaks (DSBs). Their location in the DNA is random, yet has been shown to influence the likelihood of mitotic delay and subsequent missegregation events [7–10]. This delay allows for DNA damage response (DDR) during mitosis and thus protects the genome from structural damage, but at the expense of increasing risk for numerical instability [11]. Deterministic factors that increase the incidence of missegregations include high ploidy [12] and suboptimal kinetochore-microtubule attachment stability. Tetraploid cells are more likely to fail to cluster centrosomes into two poles, leading to multipolar division. While multipolar divisions are likely lethal, multipolar mitosis can also cause the poles to coalesce leading to a pseudobipolar division and chromosome missegregations ([13]). Kinetochore-microtubule attachment stability must fall within a narrow permissible window to allow for faithful chromosome segregation [11, 14–16] and is influenced by both *cell-intrinsic* [11] and *extrinsic* factors. An example of extrinsic factors are Vinca alkaloids (e.g. vincristine, vinblastine), – a class of cytotoxic drugs which act directly upon the microtubule network [17], causing increased missegregation rates [18]. But even cytotoxic drugs not directly targeting the microtubule network have been shown to significantly impede segregation fidelity, through aforementioned stimulation of DDR and mitotic delay [19]. Drugs targeting the DDR are likely to induce numerical instability [19], suggesting that DNA-damaging therapies impart part of their cytotoxicity by interfering with chromosome segregation fidelity [20]. Changes of the tumor microenvironment such as glucose deprivation, hypoxia and acidification have also been linked to CIN [21, 22].

Anueploidy and CIN are often coupled and can create a positive-feedback

loop in which further structural or whole chromosome aberrations accumulate over time [23, 24]. But normal cells do not tolerate missegregations – aneuploid daughter cells are immediately cleared from the cell pool through apoptosis during G1 following a missegregation [25, 26]. The sudden genome-dosage imbalance caused by a missegregation event induces p53-mediated cellular senescence in the subsequent G1 phase [12, 25–27]. While cancer cells evolve to more proficiently avert missegregation-induced cell death, missegregations still activate p53 in the G1 phase, even among cancer cells [27], albeit less reliably [28]. High levels of CIN have been observed to be tumor suppressive in breast [29], ovarian, gastric, and non–small cell lung cancer [30]. The above suggests that a non-monotonic relationship between cell fitness and CIN likely exists, with a threshold of a critical level of CIN (which may be cancer type specific). A possible therapeutic avenue to target and exploit the degree of CIN in patients is therefore guided by the premise that a Goldilocks window exists for cancer to thrive.

Elizalde et. al. explored the phenotypic impact of CIN using a Markov-chain model and confirmed the existence of optimal chromosome missegregation rates [31]. The authors assumed that cells were not viable if they contained nullisomy (the loss of all copies of a chromosome), paired with a corresponding upper limit of eight copies. These assumptions were justified through a sensitivity analysis [32]. The main conclusion of the paper established that missegregation rates drove heterogeneity more than the age of the tumor. Under what circumstances missegregations lead to tumor extinction however remains unclear. Here we derive necessary conditions that drive a tumor population to nonviable karyotypes, typically through either nullisomy or the upper sustainable limit (e.g. eight copies [31]). We further refer to these conditions as missegregation-induced extinction (MIE).

The remainder of the paper is structured as follows. We first motivate the existence of MIE with a phenomenological equation of ploidy movement that takes the form of a diffusion-reaction equation. A heuristic argument comparing the time scales of net growth and missegregation will imply the existence of parameters that allow for MIE. Next, we develop our framework, which involves a general coupled compartment model. This model is then simplified to a version more amendable to theoretical analysis, while still retaining the qualitative behavior and form. Theoretical results are derived and presented on the existence of MIE and when it can be evaded. Finally, we use empirical values from the literature and surrogate biomarkers to estimate turnover- and missegregation rates and predict which cancer types are most

sensitive to MIE.

2 Methods

2.1 Motivating example

Consider a simple birth-death process on ploidy space, where, for the moment, we are interested only in the total amount of DNA content a cell contains. If the movement in ploidy content occurs only during mitosis, one can crudely approximate the total population $n(p)$ as a function of ploidy p by the following partial differential equation (PDE):

$$\frac{\partial n}{\partial t} = (\lambda - \mu)n + \beta\lambda \frac{\partial^2 n}{\partial p^2}, \tag{1}$$

where λ, μ are the birth, death rates, respectively and β is the missegregation rate. For boundary conditions, we make the assumption that there exists $p \in (p_{\min}, p_{\max})$, such that for ploidy outside this range, the population cannot survive. We also assume that $r = \lambda - \mu > 0$, that is, in the absence of missegregation, this population is favored to grow.

We now appeal to a heuristic argument of time scales. Let T_p be the time scale on which missegregation events occur, and T_r be the time scale on which the cell population grows. If $T_p \ll T_r$ then extinction via missegregation (hereby "missegregation-induced extinction" or MIE) is possible. We can estimate $T_p \sim L_p^2/(\beta\lambda)$, where L_p is the characteristic amount of DNA content shifted during a missegregation event. Similarly, one finds $T_r = 1/r$. The inequality implies that MIE can occur if

$$L_p^2 \ll \frac{\beta\lambda}{r} = \frac{\beta\lambda}{\lambda - \mu}. \tag{2}$$

An important takeaway from this simple argument is that the characteristic scale L_p can play a significant role and is tied to the typical change in copy number when a missegregation occurs. Smaller range of ploidy viability will lead to a weaker condition. Second, it is interesting to see that in theory, one does not need to increase missegregation, but rather can also increase birth and death rates in such a way that the quantity $(1 - \mu/\lambda)$ decreases. This suggests that tumors with high turnover rates may be more susceptible to MIE.

2.2 General discrete model

In practice, chromosomes are composed of discrete units of information. The granularity with which we choose to set the scale is dependent on the genomic resolution that we can access.

A general compartment model that describes the evolution of populations by their karyotypes would then be given by a M-dimensional vector $\vec{i} = (i_1, \dots i_M)$ which contains the number of copies $i_k \geq 0$ of the kth component. An example would be looking at the copy number of whole chromosomes (i.e. $M = 23$). Alternatively one can parse the chromosome into the p, q arm (i.e. $M = 46$). Movement between the compartments occurs via missegregation. We encapsulate this information in the tensor $\mathbf{q}$ with non-negative components $q_{\vec{i}\vec{j}}$, which is the rate at which a cell in compartment $\vec{i}$ moves to compartment $\vec{j}$. We require a conservation of chromosomal content (which will hold for any resolution). Let $\vec{i}$ be the parent and $\vec{j}^{(1)}$ and $\vec{j}^{(2)}$ be the offsprings, then it must be that

$$2i_k = j_k^{(1)} + j_k^{(2)}, \text{ for all } k. \tag{3}$$

This imposes structure on $\mathbf{q}$ since $0 \leq j_k \leq 2i_k$, we must have $q_{\vec{i}\vec{j}} = 0$ if $j_k > 2i_k$ for any k, thus $\mathbf{q}$ will be sparse for many applications.

We also want $\mathbf{q}$ to satisfy a flow conservation. That is, a dividing cell with chromosomal content $\vec{i}$ will give rise to two daughter cells with various chromosomal contents $\vec{j}$ such that:

$$\sum_{\vec{j}} q_{\vec{i}\vec{j}} = 1 + \beta_{\vec{i}}, \tag{4}$$

for all $\vec{i}$ and $\beta_{\vec{i}}$ is the missegregation rate. This is based on the biological condition that all chromosomal copies must be given to either one or the other daughter cell. Equations (3)-(4) are no longer true if we assume the possible formation of micronuclei. We investigated the implications of this in the discussion.

We are now in a position to write a general M-dimensional birth-death process:

$$\frac{dn_{\vec{i}}}{dt} = \underbrace{\lambda_{\vec{j}} n_{\vec{j}} q_{\vec{j}\vec{i}}}_{\text{Inflow}} - \underbrace{\lambda_{\vec{i}} n_{\vec{i}} (1 - q_{\vec{i}\vec{i}})}_{\text{outflow}} - \underbrace{\mu_{\vec{i}} n_{\vec{i}}}_{\text{death}}, \tag{5}$$

where $\lambda_{\vec{i}}, \mu_{\vec{i}}$ are the state-dependent birth and death rates, respectively. Einstein summation is implied over $\vec{j}$. Equation (3) enters into (5) with the flow rate $\mathbf{q}$.

We note that in the absence of missegregation, we require $q_{\vec{i}\vec{j}} = \mu_{\vec{i}\vec{j}}$ where we are using the vector Kronecker delta that is 1 if $\vec{i} = \vec{j}$ and 0 otherwise. This uncouples equation (5) to the classic deterministic birth-death process $dn_i/dt = (\lambda_i - \mu_i)n_i$ as expected.

One can instead suppose that a particular cancer cell line has a given missegregation rate β per chromosome copy per division and that these events are independent. Then one has the following function for $\mathbf{q}$ which is similar to a binomial distribution:

$$q_{\vec{i}\vec{j}} = \prod_k \binom{i_k}{z_k} \beta^{z_k} (1-\beta)^{i_k - z_k}, \tag{6}$$

where $z_k = |j_k - i_k|$. Note there is no inconsistency if we let $\beta \to 0$ since in that case we would have

$$q_{\vec{i}\vec{j}} = \begin{cases} 1 & \text{if } z_k = 0, \text{ for all } k, \\ 0 & \text{otherwise,} \end{cases}$$

which we recognize as the definition of the Kronecker delta defined above.

In summary, this is a flexible framework, offering a variety of biologically relevant parameter dependencies (e.g. missegregation rates may vary across compartments of different ploidy) and variable genomic resolutions. Finer genomic resolution leads to more compartments, thereby increasing computational resources required for numerical solutions.

2.3 Chromosomal aggregate model

The model given by equation (5) is complicated and cumbersome. A simpler model, amendable to analysis involves aggregating all chromosomal data into one index. Alternatively, it can be thought of as focusing on a dosage-sensitive chromosome, that must be present at copy numbers within a well defined range in order for a cell to survive. Mathematically, there are many ways to collapse our M-dimensional model to 1D, and one such way is to just sum over all indices to get the aggregated number of copies:

$$i = \|\vec{i}\|_1 = \sum_j i_j. \tag{7}$$

Then our system is given by:

$$\frac{dn_i}{dt} = \sum_j \lambda_j n_j q_{ji} - \lambda_i n_i (1 - q_{ii}) - \mu_i n_i. \tag{8}$$

We will further suppose that the parameters of the model (Supplementary Table 1) are not dependent on the size of compartments (e.g. λ_i is not dependent on any n_i, such as through a carrying capacity or Allee effect etc). This allows us to easily write the Jacobian, which is simply the coefficients of n_j in equation (8)

$$J_{ij} = \begin{cases} \lambda_i(2q_{ii} - 1) - \mu_i & \text{if } i = j, \\ \lambda_j q_{ji} & \text{if } i \neq j. \end{cases} \tag{9}$$

2.4 Numerical Simulations

Numerical simulations were performed for a range of input parameters (β, μ) to validate the predicted critical curves. All simulations had a uniform diploid population as initial condition. Each simulation ran until a quasi-steady-state (QSS) had been reached, considered to occur when the rate of change in karyotype composition was less than $0.1\% day^{-1}$ (although the cell population may still be growing or shrinking – therefore "quasi"). Hereby we define the *karyotype composition* of a population at a given time as the relative frequency of all possible copy number states in the population. Upon satisfaction of this condition, simulations with a positive rate of change in the total cell population were considered to be in the exponential growth regime.

Numerical simulations were performed for each pairwise combination of mis-segregation $B(i, \beta)$ and death functions $M(i, \mu, K)$, in order to determine the population average mis-segregation rates B_{pop} and death rates M_{pop} which are viable QSS's for the system. For each combination of functions, simulations were performed using large, manually curated ranges for the input parameters (β, μ), before B_{pop} and M_{pop} were calculated based on the QSS reached by the system.

Numerical simulations were performed using R, all code is available on Github.

2.5 Estimating missegregation rate from Interferon gene expression

Chromosome mis-segregations can trigger the formation of micronuclei. When micronuclei rupture, their genomic DNA spills into the cytosol. Cytosolic ds-DNA is sensed by the cGAS–STING pathway [33], leading to induction of type I interferon stimulated genes [34, 35]. Mis-segregations lead to the up-regulation of interferon production, which in turn subverts lethal epithelial responses to cytosolic DNA. This points to Interferon signaling as a potential surrogate measure for CIN [36].

We re-analyzed mis-segregation rates measured in 12 primary and 27 metastatic samples derived from the MDA-MB-231 cell line – a highly invasive triple-negative human Breast Invasive Carcinoma (BRCA) cell line [36]. Briefly, cells were injected into 21 mice and short-term cultures were derived from the resulting primary and metastatic tumors. All subsequent assays, including RNA-seq and immunofluorescence analysis were performed after a single passage from the primary sample as described in [36]. Mis-segregation positive cells were classified as cells with DNA and centromere staining in the middle of the anaphase plate. We integrated aforementioned BRCA-derived RNA-seq data [36] with RNA-seq data from 1,212 BRCA cases from TCGA [37], including 7 metastases. We also sampled 742 cases from 8 other cancer types (HNSC, COAD, ESCA, READ, CESC, SKCM, LUSC, DLBC) from TCGA [37, 38]. In order to isolate variation in expression across these two data sets from the biological variable of interest, we implemented a batch effect removal protocol. Using the Combat-seq function [39] in the R-package sva, we adjusted for two batches ([38] vs. [36]) as well as ten biological covariates (nine primary cancer types and one metastatic group for BRCA). We then performed Gene Set Variation Analysis (GSVA) [40] to quantify the expression activity of 1,093 REACTOME pathways [41], including 'Interferon signaling'. We fitted a linear regression on the breast cancer data from [36] as follows:

$$\beta = a\gamma + b, \tag{10}$$

where β is the log of observed percentage of cells with lagging chromosomes and γ the Interferon signaling activity as quantified with GSVA (adjusted R-square=0.32, p-value=2×10^{-4}; Supplementary Fig. 2). The resulting model was then used to predict β in the samples from [37, 38].

2.6 Estimating the distribution of cancer type specific turnover rates

We used median birth rates [42] and growth rates [43–51] reported for nine cancer types (Table 2), to estimate the distribution of turnover rates. For each cancer type we first estimated the standard deviation of birth- and growth rates as $L_\lambda/4$ and $L_r/4$, where L_λ, L_r are the ranges of birth and growth rates reported in literature respectively. We then used the 'rlnorm' function in the R package 'compositions' to obtain log normal distributions of birth- (λ) and growth rates (r), from the respective medians and standard deviations. We then calculated death rate as $\mu = 1 - \frac{r}{\lambda}$, giving us the distribution of turnover rates: $\frac{\mu}{\lambda}$.

3 Results

Let K (not necessarily finite) be the total number of compartments of the aggregate model (eq. (8)). The Jacobian (9) contains information on the local behavior of the system near the extinction state $n_i = 0$ for all i. If all the eigenvalues of the Jacobian at the extinction point are negative, then MIE occurs. The critical curve that separates MIE from exponential growth is when the maximum eigenvalue of J is 0. We discuss and provide conditions when MIE is possible.

3.1 MIE is possible when only two karyotypes are viable ($K = 2$).

Suppose now that a parental cell has content i during the G1 phase of the cell cycle, which implies $2i$ content will be available at mitosis. Assuming any uneven segregation is equally likely to occur leads to a $\mathbf{q}$ given by:

$$q_{ij} = \begin{cases} 1-\beta & \text{if } i = j, \\ \frac{\beta}{i} & \text{if } j \leq 2i, j \neq i, \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

where we have introduced a missegregation rate $0 \leq \beta \leq 1$.

Consider the case where we assume the only compartments capable of sustaining the population ($\lambda_i > 0$) are $i = k_1, k_2$, with $1 \leq k_1 < k_2 \leq 2k_1$.

The constraint on k_2 is required to ensure that k_1 has sufficient copies of the chromosome to reach k_2. When all parameters are identical and constant with $\lambda > \mu$ we arrive at the condition for MIE given by

$$\beta > \beta_c = \frac{2k_1k_2 + \sqrt{k_1k_2}}{4k_1k_2 - 1}\left(1 - \frac{\mu}{\lambda}\right). \tag{12}$$

For the general case, see the appendix. The minimum missegregation rate needed for a population with only two viable karyotypes to go extinct is:

$$\beta_c \geq \frac{1}{2}\left(1 - \frac{\mu}{\lambda}\right). \tag{13}$$

Note that even in the absence of death ($\mu = 0$), we have MIE possible provided that $\beta \geq 1/2$. Recall though that if only two karyotypes are viable, many missegregations put the daughter cell into a nonviable compartment, thereby acting as a pseudo-death rate.

3.2 MIE is impossible with the number of viable karyotypes $K \to \infty$.

Let us assume that only the nearest jump can be made (i.e. no more than one copy of a critical chromosome can mis-segregate during one cell division), then $\mathbf{q}$ has elements

$$q_{ij} = \begin{cases} 1 - \beta & \text{if } i = j, \\ \beta & \text{if } |i - j| = 1, \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

If all parameters are equal, we can show analytically that MIE is impossible with $K \to \infty$ (APPENDIX), confirming intuition. For large K the critical missegregation rate β_c is given by

$$\beta_c \sim \frac{2K^2}{\pi^2}\left(1 - \frac{\mu}{\lambda}\right). \tag{15}$$

Recall that $\beta \in [0, 1]$ and so for sufficiently large K, this will push us outside of this bound. Rearranging provides an illuminating result. Setting $\beta_c = 1$ (the maximum possible value), we see that MIE requires K to satisfy:

$$K \leq \lfloor K_c \rfloor = \left\lfloor \frac{\pi}{\sqrt{2}}\left(1 - \frac{\mu}{\lambda}\right)^{-1/2} \right\rfloor, \tag{16}$$

Table 1: **Sufficient conditions to avoid MIE for various biological scenarios:** 1) Small jumps with constant rates. 2) Birth-, death-, and/or mis-segregation rates vary across compartments. 3) Mis-segregation rate is constant.

Scenario	Sufficient condition to avoid MIE
1. $\forall \lvert i-j\rvert > 1: \quad q_{ij} = 0$, (μ, λ, β) constant	$\beta < \beta_c = \frac{1}{4}\left(1 - \frac{\mu}{\lambda}\right)$
2. (μ, λ, β) vary across compartments	$\forall i$: $\lambda_i(1 - 3\beta_i) - \mu_i - \lambda_{i+1}\beta_{i+1} > 0$
3. β is constant	$\forall i$: $\beta < \beta_c = \frac{1}{3+\lambda_{i+1}/\lambda_i}\left(1 - \frac{\mu_i}{\lambda_i}\right)$

where $\lfloor \cdot \rfloor$ is the floor function. This is highly restrictive. If $\mu/\lambda = 0.8$, then MIE is avoided if over four copies are permitted. With $\mu/\lambda = 0.99$, the maximum number of copies can be no larger than 22.

The above result is rigorous. To discuss the general case, a heuristic argument in the continuum limit shows that MIE is impossible if the measure of the set of viable regions tends to infinity (APPENDIX). This also includes the possibility of regions of karyotype viability interspersed between nonviable regions.

3.3 Ruling out MIE.

Here, we establish sufficient conditions for MIE to not occur based on Gershgorin's circle (GC) theorem [52]. The theorem bounds the locations of the eigenvalues in the complex plane for a given matrix $\mathbf{A}$, with elements a_{ij}. The GC theorem stipulates that the eigenvalues must be contained in the circles with centers a_{ii} and radii $R = \sum_{i \neq j} |a_{ij}|$.

This is a useful theorem, though the condition is very strong. Since MIE can be evaded if the maximum eigenvalue exceeds 0, a sufficient condition is that none of the GCs contain a part of the negative reals. Table 1 describes sufficient conditions to avoid MIE for various biological assumptions. For example in the case of small jump (section 3.2) with constant rates, we showed that MIE is impossible as the number of compartments tended to infinity. Of course, we know that for large K, *no* missegregation rate will lead to MIE. Table 1 just provides sufficient conditions.

The general problem for arbitrary $\mathbf{q}$ can be handled numerically, but an-

alytical conclusions can only be made for specific forms of $\mathbf{q}$. The conditions required are given by finding when the GC's are all contained in the positive half-plane:

$$\min_k \left[\lambda_k(2q_{kk} - 1) - \mu_k - \lambda_k \sum_{j \neq k} q_{kj} \right] > 0, \tag{17}$$

$$\min_k \left[\lambda_k(2q_{kk} - 1) - \mu_k - \sum_{j \neq k} \lambda_j q_{jk} \right] > 0. \tag{18}$$

As both of these need to be positive, we can find the minimum of these, which will provide sufficient condition to escape MIE.

3.4 Intra-tumor heterogeneity in mis-segregation- and death rates can render MIE impossible.

Given a fixed birth- and death-rate, can we predict at what mis-segregation rate a population will go extinct? Here we derive critical curves that separate viable from non-viable populations as a function of their turnover- ($\frac{\mu}{\lambda}$) and mis-segregation rates (β). Herein we make three assumptions: (i) the existence of a crucial chromosome imposes a hard limit on the number of viable compartments K (i.e. only cells with $\{1..K\}$ chromosome copies are viable); (ii) all missegregation events are possible (e.g. if parent after synthesis phase has $2i$ copies, then a daughter cell can have any integer in the range $[0, 2i]$); and (iii) constant turnover- and mis-segregation rates regardless of copy number.

To generate the critical curves conditional on these assumptions, we combined eqs. (6) and (8) and solved for β at various turnover rates $\frac{\mu}{\lambda}$. We employed an eigenvalue solver using Arnoldi iterations to find the largest real part eigenvalues, resulting in a critical mis-segregation rate β for a given turnover rate $\frac{\mu}{\lambda}$ (Fig. 2). We also simulated the ODE given by eq. (8) until the karyotype distribution reached the steady state (Supplementary Fig. 1A,C,E), assuming karyotypes with more than eight copies are not viable (i.e. $K = 8$). Numerical simulations confirmed that the predicted critical curve does separate exponential growth from population extinction (Fig. 2).

Above predictions assumed a cell population with homogeneous mis-segregation and turnover rates. Both are likely heterogeneous in reality. We therefore asked whether relaxing this assumption changes the critical curves.

Specifically, we assumed mis-segregation rate (B) increases when the number of chromosome copies $i \in \{1..K\}$ departs from the diploid state:

$$B(i, \beta) = 1 - \frac{1}{\beta + e^{\sqrt{|i-2|}}}, \tag{19}$$

with shape parameter β (Fig. 3A). Alternatively, we kept mis-segregation rate constant and instead varied death rate (M) as follows (Fig. 3B):

$$M(i, \mu) = \frac{[\mu i - \sin(1.7 - 3.2i)]^{0.8/i}}{K}, \tag{20}$$

Equation (20) models the relationship between karyotype and fitness accounting for two recent observations [53]: (i) the more a karyotype deviates from euploidy, the higher its fitness penalty (here: death rate) and (ii) the relative fitness penalty of aneuploidy decreases with increasing ploidy.

For each of the above biological scenarios, we again calculated the critical mis-segregation rate parameter (β in eq. (19)) and death rate parameter (μ in eq. (20)) respectively (Fig. 3C,D). We also simulated the ODE assuming the relation between copy number and mis-segregation or death rate is given by eqs. (19) or (20). The analytically predicted critical curves were again confirmed numerically (data not shown). Changes in karyotype composition over time for a single parameter combination along the critical curve is shown in Fig. 3E,F, indicating existence of a steady state karyotype composition (Supplementary Fig. 1B,D,F). We then calculated the weighted population-average mis-segregation and death rate at steady state for parameter combinations along the critical curve (Fig. 3G,H).

We repeated these calculations for various pairwise combinations of biological assumptions about the heterogeneity of mis-segregation and death rates (Supplementary Fig. 3). Intra-tumor heterogeneity (ITH) in mis-segregation and death rates did indeed shift the location of the critical curve with the direction of the shift dependent on the specific functional choice. Of note, the imposed upper limit of chromosome copies had a substantial effect on the location of the critical curve for all biological scenarios (Fig. 4A-I). As expected, the higher the mis-segregation rate the faster the populations reached their steady state karyotype composition in all cases (Supplementary Fig. 4).

Certain combinations of mis-segregation and death functions protected the population from MIE. I.e. the resulting steady state karyotypes were

guaranteed to lie outside MIE regions of the parameter space. We identified two mechanisms of protection against MIE, both being consequences of our choices of mis-segregation and death functions (Supplementary Fig. 3) or interactions between them: (i) selection favors karyotypes with lower mis-segregation rates (Fig. 4E-F,H) or (ii) mis-segregation rates steadily decrease as copy numbers approach the upper and lower limits (Fig. 4D). The latter effect is more pronounced as the upper limit K increases: at higher values of K mis-segregations depart the population further away from the $i = 3$ state (where B is highest), to $i = 16$ and beyond (where B is low). This movement to high i/low B states becomes a one-way street as K increases, because very low mis-segregation rates at $i \geq 16$ prevent the population from coming back to low i/high B states. This suggests divergent missegregation rates as a potential mechanism of protection against MIE, especially when turnover rates are low.

3.5 Predicting sensitivity to MIE across cancers.

At high turnover rates, when population growth is close to homeostasis ($M < 0.01$, further referred to as "near-homeostatic growth"), MIE remained possible in all aforementioned instances, even at the relaxed upper limit of $K = 32$ (Fig. 4). Turnover rates of several cancer types are reported to overlap with this near-homeostatic growth region (Table 2). We therefore compared the population-average critical curves with each other, focusing on the near-homeostatic growth region (Fig. 4), for $K = \{4, 8, 16, 32\}$ as the maximum number of chromosome copies. Compared to homogeneous death and homogeneous mis-segregation rates (Fig. 4A), ITH in either death or mis-segregation rates typically did not result in less restrictive critical curves (Fig. 4B,C,G). This remained true when both, ITH in mis-segregation and death rates was assumed simultaneously (Fig. 4F,H-I). A notable exception was the functional form of mis-segregation ITH shown in Fig. 4D,E, which resulted in less restrictive critical curves in near-homeostatic growth regions, compared to homogeneous mis-segregations.

We compared mis-segregation- and turnover rates for nine tumor types to the critical curves shown in 4A,D. Turnover rates were calculated from median birth and net growth rates reported in literature for the different cancers (Table 2, section: 2.6). Fig. 5 shows critical curves for $K = 32$ as the maximum number of chromosome copies, assuming mis-segregation rates are either heterogeneous (Fig. 4D) or homogeneous (Fig. 4A).

Table 2: ***In vivo* tumor dynamics across nine cancer types**. Birth rates taken from [42] and growth rates taken from multiple sources. The death rate is inferred. Many of the birth rates are comparable, but the net growth rates vary over orders of magnitude, which implies very different death rates.

Cancer type	birth rate (λ)	growth rate (r)	death rate ($\mu = \lambda - r$)
Head/Neck [54]	0.192 (0.015 - 1.667)	0.007 (0.003 - 0.033)	0.185 (0 - 1.664)
Esophageal [43]	0.207 (0.018 - 0.625)	0.154 (0.035 - 0.347)	0.053 (0 - 0.59)
Colorectal [44]	0.239 (0.019 - 1.111)	0.003 (0.0003 - 0.038)	0.236 (0 - 1.111)
Rectal [46]	0.303 (0.179 - 0.417)	0.005 (0.0004 - 0.013)	0.298 (0.166 - 0.416)
Breast [47]	0.122 (0.021 - 0.556)	0.003 (0.0003 - 0.016)	0.119 (0.005 - 0.555)
Cervix [48]	0.25 (0.159 - 0.323)	0.002 (0.0004 - 0.012)	0.248 (0.147 - 0.322)
Melanoma [49]	0.139 (0.024 - 0.286)	0.007 (0.002 - 0.014)	0.132 (0.01 - 0.284)
SC Lung [50]	0.122 (0.036 - 0.667)	0.008 (0.005 - 0.013)	0.114 (0.023 - 0.662)
D-LBCL [51]	0.063 (0.043 - 0.4)	0.023 (0.01 - 0.5)	0.04 (0 - 0.39)

For both assumptions, the majority of cancers did fall below the critical curves. Assuming heterogeneous mis-segregation rates resulted in a marginal increase in the proportion of near-homeostatic tumors that lie below the critical curve, in particular for colon cancer (Fig. 5). We note that the critical curves were chosen here for comparison with the data because they were the least restrictive (Fig. 4A,D). Other functional choices laid out in Fig. 4 would have left a larger proportion of cancers within regions predicted as unviable.

4 Discussion

We have presented a general deterministic framework for modeling whole chromosome missegregations. The theory has shown the existence of a potential mechanism of tumor control through the region in parameter space we have called MIE.

The model raises some important theoretical questions related to malignant and non-malignant cells. In particular, it is well known that normal cells maintain a level of homeostasis through a balanced turnover rate $\mu/\lambda \approx 1$. This seems to imply that *all* normal cells lay at the MIE boundary and are even more sensitive to MIE than malignant cells. Any finite missegregation rate would thus lead to the slow removal of normal cells over time. There are two potential explanations for this behavior. A likely explanation is given by our assumption that the birth rate is independent of population size. It is easy to see that introducing dependency on the total populations size (e.g. carrying capacity) could alleviate this issue as cell death would increase the birth rate in order to return to homeostasis. An alternative explanation is that this is just another natural aging mechanism through which normal cells are slowly displaced. Transformed cells often lose the homeostatic control mechanisms and so are likely less susceptible to contact inhibition.

Our model does not explicitly account for several biological mechanisms which are relevant to karyotype evolution, including WGD, mis-segregation induced apoptosis in the subsequent G1 phase of the cell cycle, and the formation of micronuclei. Though WGDs are technically feasible in the model (through an i parent giving birth to $2i$ and 0 copy number offspring), this ignores the biological mechanisms for how WGD events may occur. There are two mechanisms for WGD [55, 56] – endoreplication (i.e. completely skipping mitosis) and cytokinesis failure which often generates binucleated whole genome doubled cells. Chromosomes in the two nuclei can then merge the subsequent mitosis after nuclear envelope breakdown. We can modify equation (8) to account for WGD by introducing terms

$$\frac{dn_i}{dt} = \sum_j \lambda_j n_j q_{ji} - \lambda_i n_i (1 - q_{ii}) - \mu_i n_i - \theta_i n_i + \theta_{i/2} n_{i/2},$$

where $\theta_{i/2} = 0$ if $i/2$ is not an integer.

Mounting evidence suggests that a cell which mis-segregates many chromosomes has a higher risk to die than a cell mis-segregating only one chromosome [27, 28]. Our current model structure however does not distinguish between these scenarios. To account for this intolerance, we can amend the inflow term by introducing a jump length coefficient ϕ_{ij}. One possible form $\phi = (1 + |i - j|)^\alpha$ with $\alpha < 0$, where only relative distance is important. Other forms that take the parental state into account could also be consid-

ered. Combining this with the above, leads to

$$\frac{dn_i}{dt} = \sum_j \phi_{ji}\lambda_j n_j q_{ji} - \lambda_i n_i(1 - q_{ii}) - \mu_i n_i - \theta_i n_i + \theta_{i/2} n_{i/2}.$$

The only requirement is that ϕ_{ij} must fall off with increasing distance between states.

Yet another assumption, which could be eventually relaxed is the copy number conservation in equation (3). The errors during mitosis could lead to the formation of micronuclei which are not directly involved in further downwind mitosis [18, 36, 57–59]. This amounts to replacing equation (3) with $2i_k \geq j_k^{(1)} + j_k^{(2)}$ for all k.

Despite these simplifications, the majority of tumors across all nine cancer types studied had missegregation- and turnover rates that were compatible with our viability predictions (Fig. 5), albeit under the assumption that cells can tolerate a high number of chromosome copies ($K > 16$). This remained true when a dependency of mis-segregation- and death rates on ploidy was introduced. While representing only a small sample from the space of possible interactions between mis-segregation, death and ploidy that may exist in-vitro/in-vivo, our results emphasize that populations can evolve to protect themselves from MIE, especially as they depart further away from homeostasis. That each tumor consists of cells with heterogeneous mis-segregation rates, the measurement being just the population-average rate, is a likely scenario supported by recent results [60, 61]. Copy number states associated with low mis-segregation rates act as a stabilizing refuge, protecting the population from extinction. Intra-tumor heterogeneity, including heterogeneity in mis-segregation rates, increases as tumors progress. Our predictions suggest that this intra-tumor heterogeneity hinders the chance of success of therapies aimed at MIE. On the other hand, inducing missegregations in more homogeneous, early stage cancers, is likely not possible without also increasing intra-tumor heterogeneity, hampering the overall potential for MIE as a therapeutic strategy.

Future applications of this model will include studying the fitness costs and benefits of high ploidy. Coexistence of cancer cells at opposite extremes of the ploidy spectrum occurs frequently in cancer and missegregations are a major contributor to heterogeneous ploidy states within a population. Our model can help understand how much robustness high ploidy confers to the sudden genome-dosage imbalance caused by a missegregation event [12] and

can help quantify the energetic requirements of high ploidy cells. The translational potential of this knowledge lies in the ability to successfully tailor MIE therapies to tumors that cannot afford high ploidy.

5 Acknowledgements

The results shown here are in part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

6 Figures

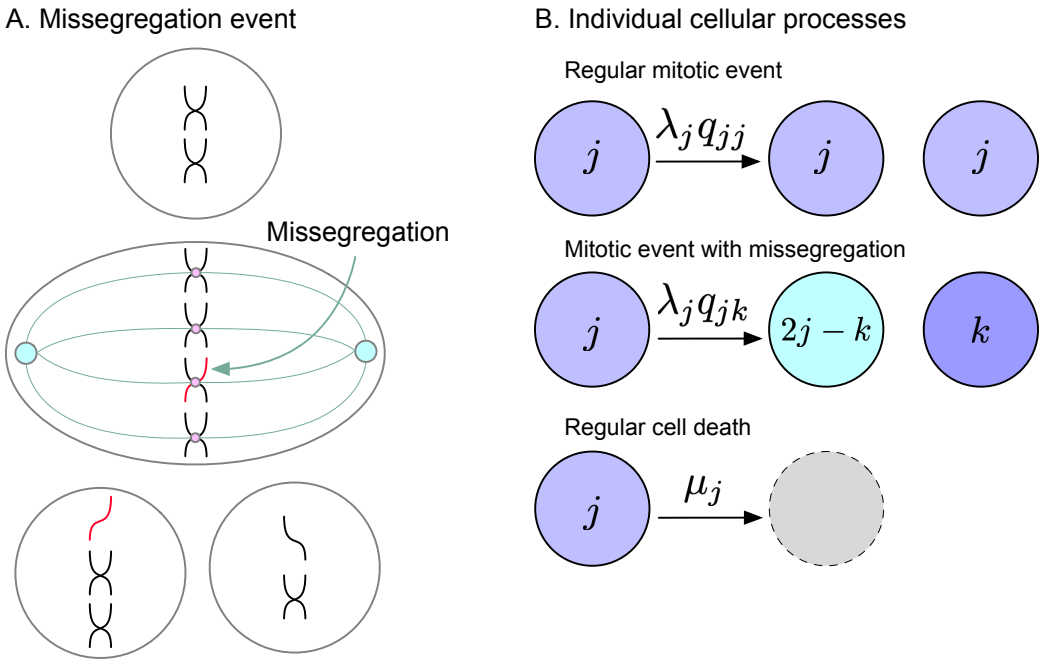

Figure 1: **Mathematical modeling of chromosome missegregations.** (**A**) Missegregation event. During anaphase, one daughter cell improperly takes both chromosomes leading to aneuploidy. Note the ploidy conservation assumption here. (**B**) Model's individual cellular processes. The tensor q_{jk} encodes the probability at which a cell of karyotype j may produce offspring of karyotype k (and also $2j-k$ by ploidy conservation). Thus karyotype j goes through anaphase with faithful chromosome segregations at a rate $\lambda_j q_{jj}$ and missegregates into karyotype k (and $2j-k$) at a rate $\lambda_j q_{jk}$. Cell death occurs with rate μ_j.

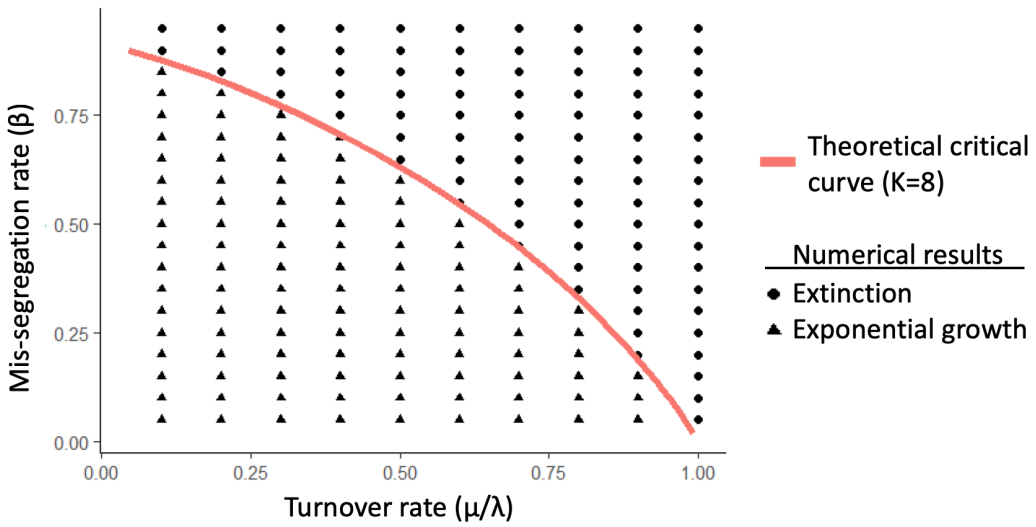

Figure 2: **Predicting MIE as a function of homogeneous mis-segregation and turnover rates**. Critical curve was obtained by finding $(\beta, \frac{\mu}{\lambda})$ for which the maximum eigenvalues of the Jacobian for eq. (8) is 0. Herein we assumed existence of a critical chromosome x for which the viable number of copies is $0 < i_x < K := 8$ and that both mis-segregation and turnover rates are constant and independent of copy number state i_x. Note that the size of the Jaccobian is a function of K. Numerical simulations confirm that critical curve separates viable from non-viable states.

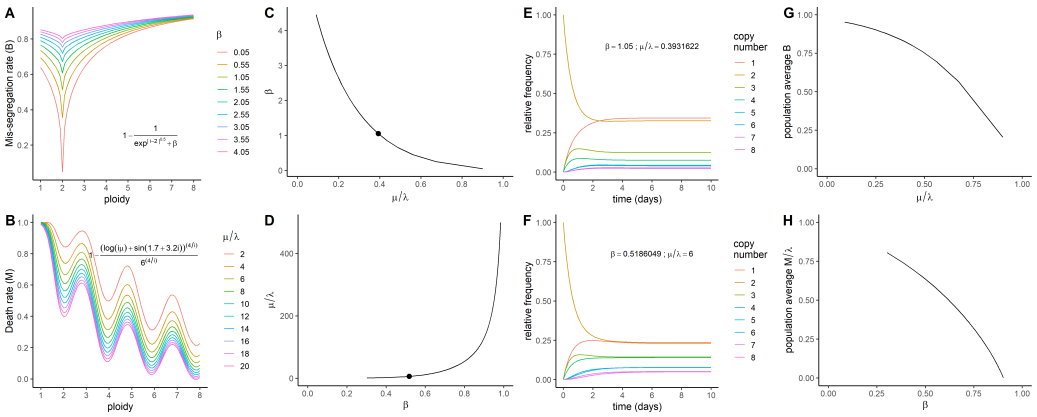

Figure 3: **Predicting MIE when mis-segregation and turnover rates are heterogeneous**. (A) Intra-tumor heterogeneity in mis-segregation rate is modeled as a function of ploidy (x-axis). Varying parameter β (color code) yields different mis-segregation rates (B). (B) Same as (A), but here intra-tumor heterogeneity in death rates is assumed and μ is the parameter that governs death rate (M). (C-H) $\lambda := 1$ in all calculations, i.e. cells divide once per day. (C) Critical curve was obtained for (A) by finding $(B, \frac{\mu}{\lambda})$ for which the maximum eigenvalues of the Jacobian for eq. (8) is 0. (D) Critical curve was obtained for (B) in the same manner as in (C), but here equations were solved for $(\beta, \frac{M}{\lambda})$. (E,F) We used the parameters highlighted in (C,D) to simulate copy number evolution until the population composition reached the steady state. (G,H) The eigenvectors corresponding to the eigenvalues found in C/D are the steady state population compositions. These are used in conjunction with the functions in A/B to determine the population average mis-segregation (G) and turnover rates (H).

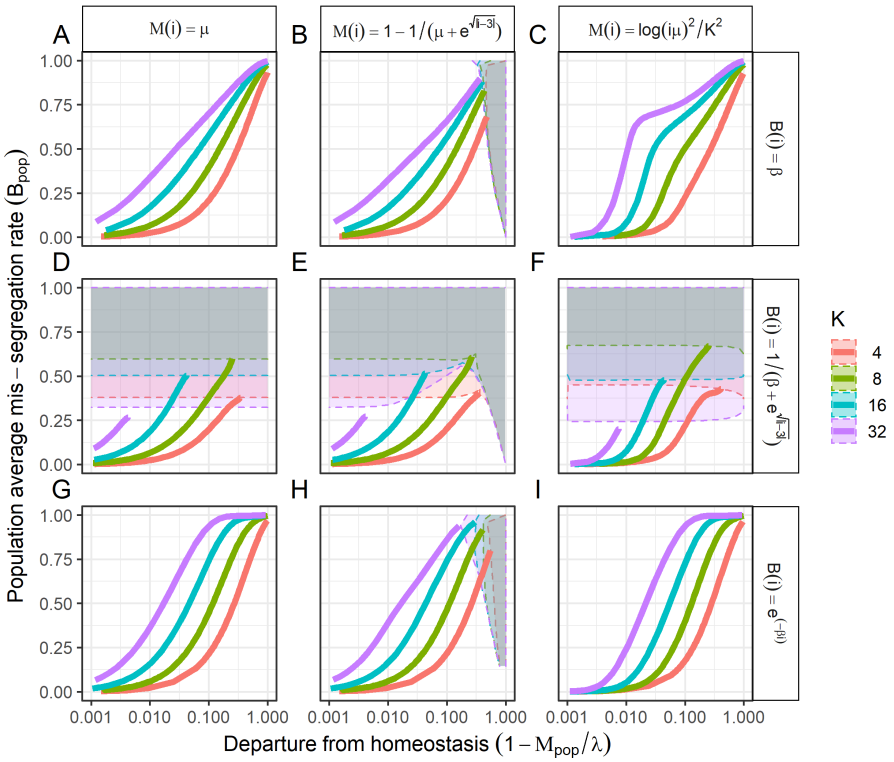

Figure 4: **Heterogeneous mis-segregation and death rates can render MIE impossible at high turnover rates**. The phenotypic consequence of homogeneous (A) and heterogeneous mis-segregation (D-I) and turnover rates (B-C,E-F,H-I) are compared with each other. Population average missegregation rate (B_{pop}) is defined as fraction of divisions in which a missegregation occurs (i.e. $1-(1-B)^i$). Clear areas represent possible quasi-steady states for the system (i.e. the population distribution has reached a steady state, so average mis-segregation or turnover rates are constant - although the cell population may be growing or shrinking). Within these clear areas, critical curves separate exponential growth from population extinction. Population steady state never lies within the shaded areas, i.e. the population average mis-segregation rate at steady state is guaranteed to lie outside MIE regions. This is a consequence of the respective inbuilt dependency of mis-segregation rates (B) and/or death rates (M) on copy number. Because the MIE curves rely on mis-segregation rate at steady state (see Fig. 3), they cannot be computed in shaded areas. Note that death rate (M_{pop}) and maximum copy number (K), both influence what steady state the population evolves into.

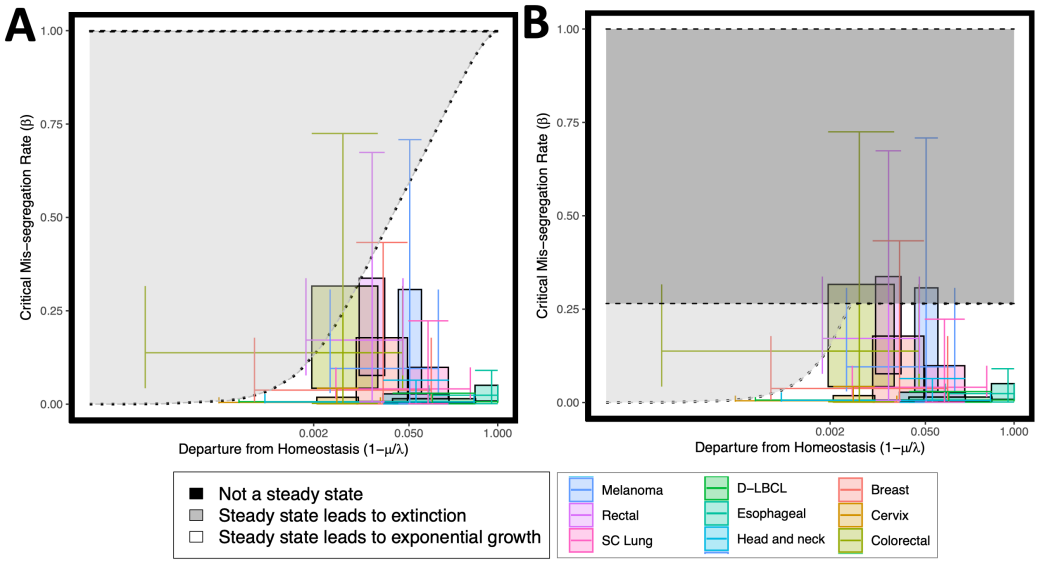

Figure 5: **Predicting sensitivity to MIE across cancers.** Predicted critical curves separate MIE (light gray shaded areas) from exponential growth (clear areas). Steady state karyotype composition never lies within the dark shaded areas. Critical curves calculated assuming either mis-segregations are independent of ploidy (i.e. homogeneous; A) or that mis-segregation rate increases with increasing departure from the triploid state (B). Both curves assume existence of a critical chromosome x for which the viable number of copies is $0 < i_x < 32$. Mis-segregation- (β), and turnover rates (μ/λ) for nine tumor types (colors) are displayed as two independent 1D boxplots alongside the predictions. 2D boxplots imply the assumption that mis-segregation rates are independent of turnover rates [62]. Assuming homogeneous mis-segregation rates, the majority of tumors lie below the critical curves in eight of the nine cancer types. The predicted curve fails to explain high mis-segregation rates observed for the near-homeostatic growth of colon cancer. Intra-tumor heterogeneity in mis-segregation rates explains slightly more of this near-homeostatic growth, at the expense of missing a higher proportion tumors that grow further away from homeostasis.

References

[1] D Choma, JP Daures, X Quantin, and JL Pujol. Aneuploidy and prognosis of non-small-cell lung cancer: a meta-analysis of published data. *British journal of cancer*, 85(1):14–22, 2001. Publisher: Nature Publishing Group.

[2] Mallika Siva Donepudi, Kasturi Kondapalli, Seelam Jeevan Amos, Pavithra Venkanteshan, and others. Breast cancer statistics and markers. *Journal of cancer research and therapeutics*, 10(3):506, 2014. Publisher: Medknow Publications.

[3] Axel Walther, Richard Houlston, and Ian Tomlinson. Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis. *Gut*, 57(7):941–950, 2008. Publisher: BMJ Publishing Group.

[4] Maximilian Lennartz, Sarah Minner, Sophie Brasch, Hilko Wittmann, Leonard Paterna, Katja Angermeier, Eray Öztürk, Rami Shihada, Mingu Ruge, Martina Kluth, and others. The combination of DNA ploidy status and PTEN/6q15 deletions provides strong and independent prognostic information in prostate cancer. *Clinical Cancer Research*, 22(11):2802–2811, 2016. Publisher: AACR.

[5] Daniel Stieber, Anna Golebiewska, Lisa Evers, Elizabeth Lenkiewicz, Nicolaas HC Brons, Nathalie Nicot, SÚbastien Bougnaud, Frank Hertel, Rolf Bjerkvig, Laurent Vallar, and others. Glioblastomas are composed of genetically divergent clones with distinct tumourigenic potential and variable stem cell-associated phenotypes. *Acta neuropathologica*, 127(2):203–219, 2014. Publisher: Springer.

[6] Sarah L. Thompson and Duane A. Compton. Examining the link between chromosomal instability and aneuploidy in human cells. *The Journal of Cell Biology*, 180(4):665–672, February 2008.

[7] Samuel F. Bakhoum, Lilian Kabeche, John P. Murnane, Bassem I. Zaki, and Duane A. Compton. DNA-damage response during mitosis induces whole-chromosome missegregation. *Cancer Discovery*, 4(11):1281–1289, November 2014.

[8] Makoto T. Hayashi, Anthony J. Cesare, James A. J. Fitzpatrick, Eros Lazzerini-Denchi, and Jan Karlseder. A telomere-dependent DNA damage checkpoint induced by prolonged mitotic arrest. *Nature Structural & Molecular Biology*, 19(4):387–394, March 2012.

[9] Rune Troelsgaard Pedersen, Thomas Kruse, Jakob Nilsson, Vibe H. Oestergaard, and Michael Lisby. TopBP1 is required at mitosis to re-

duce transmission of DNA damage to G1 daughter cells. *The Journal of Cell Biology*, 210(4):565–582, August 2015.

[10] Sheroy Minocherhomji, Songmin Ying, Victoria A. Bjerregaard, Sara Bursomanno, Aiste Aleliunaite, Wei Wu, Hocine W. Mankouri, Huahao Shen, Ying Liu, and Ian D. Hickson. Replication stress activates DNA repair synthesis in mitosis. *Nature*, 528(7581):286–290, December 2015.

[11] Samuel F. Bakhoum, Lilian Kabeche, Duane A. Compton, Simon N. Powell, and Holger Bastians. Mitotic DNA Damage Response: At the Crossroads of Structural and Numerical Cancer Chromosome Instabilities. *Trends in Cancer*, 3(3):225–234, March 2017. Publisher: Elsevier.

[12] Jason M. Sheltzer, Julie H. Ko, John M. Replogle, Nicole C. Habibe Burgos, Erica S. Chung, Colleen M. Meehl, Nicole M. Sayles, Verena Passerini, Zuzana Storchova, and Angelika Amon. Single-chromosome Gains Commonly Function as Tumor Suppressors. *Cancer Cell*, 31(2):240–255, February 2017.

[13] Neil J Ganem, Susana A Godinho, and David Pellman. A mechanism linking extra centrosomes to chromosomal instability. *Nature*, 460(7252):278–282, 2009. Publisher: Nature Publishing Group.

[14] Norman Ertych, Ailine Stolz, Albrecht Stenzinger, Wilko Weichert, Silke Kaulfuß, Peter Burfeind, Achim Aigner, Linda Wordeman, and Holger Bastians. Increased microtubule assembly rates influence chromosomal instability in colorectal cancer cells. *Nature Cell Biology*, 16(8):779–791, August 2014.

[15] Samuel F. Bakhoum, Sarah L. Thompson, Amity L. Manning, and Duane A. Compton. Genome stability is ensured by temporal control of kinetochore-microtubule dynamics. *Nature Cell Biology*, 11(1):27–35, January 2009.

[16] Samuel F. Bakhoum and Duane A. Compton. Kinetochores and disease: keeping microtubule dynamics in check! *Current Opinion in Cell Biology*, 24(1):64–70, February 2012.

[17] Darcy Bates and Alan Eastman. Microtubule destabilising agents: far more than just antimitotic anticancer drugs. *British Journal of Clinical Pharmacology*, 83(2):255–268, February 2017.

[18] Paola Leopardi, Francesca Marcon, Gabriella Dobrowolny, Andrea Zijno, and Riccardo Crebelli. Influence of donor age on vinblastine-induced chromosome malsegregation in cultured peripheral lymphocytes. *Mutagenesis*, 17(1):83–88, January 2002.

[19] Hee-Sheung Lee, Nicholas C. O. Lee, Natalay Kouprina, Jung-Hyun Kim, Alex Kagansky, Susan Bates, Jane B. Trepel, Yves Pommier, Dan Sackett, and Vladimir Larionov. Effects of Anticancer Drugs on Chromosome Instability and New Clinical Implications for Tumor-Suppressing Therapies. *Cancer Research*, 76(4):902–911, February 2016.

[20] Samuel F. Bakhoum, Lilian Kabeche, Matthew D. Wood, Christopher D. Laucius, Dian Qu, Ashley M. Laughney, Gloria E. Reynolds, Raymond J. Louie, Joanna Phillips, Denise A. Chan, Bassem I. Zaki, John P. Murnane, Claudia Petritsch, and Duane A. Compton. Numerical chromosomal instability mediates susceptibility to radiation treatment. *Nature Communications*, 6:5990, January 2015.

[21] Chunyan Dai, Feifei Sun, Chunpeng Zhu, and Xun Hu. Tumor environmental factors glucose deprivation and lactic acidosis induce mitotic chromosomal instability–an implication in aneuploid human tumors. *PLoS One*, 8(5):e63054, 2013. Publisher: Public Library of Science.

[22] Miyako Kondoh, Noritaka Ohga, Kosuke Akiyama, Yasuhiro Hida, Nako Maishi, Alam Mohammad Towfik, Nobuo Inoue, Masanobu Shindoh, and Kyoko Hida. Hypoxia-induced reactive oxygen species cause chromosomal abnormalities in endothelial cells in the tumor microenvironment. *PloS one*, 8(11):e80349, 2013. Publisher: Public Library of Science.

[23] Sally M. Dewhurst, Nicholas McGranahan, Rebecca A. Burrell, Andrew J. Rowan, Eva Grönroos, David Endesfelder, Tejal Joshi, Dmitri Mouradov, Peter Gibbs, Robyn L. Ward, Nicholas J. Hawkins, Zoltan Szallasi, Oliver M. Sieber, and Charles Swanton. Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. *Cancer Discovery*, 4(2):175–185, February 2014. 00068.

[24] Takeshi Fujiwara, Madhavi Bandi, Masayuki Nitta, Elena V. Ivanova, Roderick T. Bronson, and David Pellman. Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature*, 437(7061):1043–1047, October 2005.

[25] Mario Vitale. Intratumor BRAFV600E heterogeneity and kinase inhibitors in the treatment of thyroid cancer: a call for participation. *Thyroid: official journal of the American Thyroid Association*, 23(4):517–519, April 2013. 00000.

[26] Samuel F. Bakhoum and Duane A. Compton. Chromosomal instability and cancer: a complex relationship with therapeutic potential. *The Journal of Clinical Investigation*, 122(4):1138–1143, April 2012.

[27] Sarah L. Thompson and Duane A. Compton. Proliferation of aneuploid human cells is limited by a p53-dependent mechanism. *The Journal of Cell Biology*, 188(3):369–381, February 2010.

[28] Stefano Santaguida, Amelia Richardson, Divya Ramalingam Iyer, Ons M'Saad, Lauren Zasadil, Kristin A. Knouse, Yao Liang Wong, Nicholas Rhind, Arshad Desai, and Angelika Amon. Chromosome mis-segregation generates cell cycle-arrested cells with complex karyotypes that are eliminated by the immune system. *Developmental cell*, 41(6):638–651.e5, June 2017.

[29] Rebecca Roylance, David Endesfelder, Patricia Gorman, Rebecca A Burrell, Jil Sander, Ian Tomlinson, Andrew M Hanby, Valerie Speirs, Andrea L Richardson, Nicolai J Birkbak, Aron C Eklund, Julian Downward, Maik Kschischo, Zoltan Szallasi, and Charles Swanton. Relationship of extreme chromosomal instability with long-term survival in a retrospective analysis of primary breast cancer. *Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 20(10):2183–2194, October 2011. 00039 PMID: 21784954.

[30] Nicolai J Birkbak, Aron C Eklund, Qiyuan Li, Sarah E McClelland, David Endesfelder, Patrick Tan, Iain B Tan, Andrea L Richardson, Zoltan Szallasi, and Charles Swanton. Paradoxical relationship between chromosomal instability and survival outcome in cancer. *Cancer research*, 71(10):3447–3452, May 2011. 00068 PMID: 21270108.

[31] Sergi Elizalde, Ashley M. Laughney, and Samuel F. Bakhoum. A Markov chain for numerical chromosomal instability in clonally expanding populations. *PLoS computational biology*, 14(9):e1006447, 2018.

[32] Ashley M. Laughney, Sergi Elizalde, Giulio Genovese, and Samuel F. Bakhoum. Dynamics of Tumor Heterogeneity Derived from Clonal Karyotypic Evolution. *Cell Reports*, 12(5):809–820, August 2015.

[33] Lijun Sun, Jiaxi Wu, Fenghe Du, Xiang Chen, and Zhijian J. Chen. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science (New York, N.Y.)*, 339(6121):786–791, February 2013.

[34] Yuk Yuen Lan, Diana Londoño, Richard Bouley, Michael S. Rooney, and Nir Hacohen. Dnase2a deficiency uncovers lysosomal clearance of damaged nuclear DNA via autophagy. *Cell Reports*, 9(1):180–192, October 2014.

[35] Karen J. Mackenzie, Paula Carroll, Carol-Anne Martin, Olga Murina, Adeline Fluteau, Daniel J. Simpson, Nelly Olova, Hannah Sutcliffe, Jacqueline K. Rainger, Andrea Leitch, Ruby T. Osborn, Ann P. Wheeler, Marcin Nowotny, Nick Gilbert, Tamir Chandra, Martin A. M. Reijns, and Andrew P. Jackson. cGAS surveillance of micronuclei links genome instability to innate immunity. *Nature*, 548(7668):461–465, August 2017.

[36] Samuel F. Bakhoum, Bryan Ngo, Ashley M. Laughney, Julie-Ann Cavallo, Charles J. Murphy, Peter Ly, Pragya Shah, Roshan K. Sriram, Thomas B. K. Watkins, Neil K. Taunk, Mercedes Duran, Chantal Pauli, Christine Shaw, Kalyani Chadalavada, Vinagolu K. Rajasekhar, Giulio Genovese, Subramanian Venkatesan, Nicolai J. Birkbak, Nicholas McGranahan, Mark Lundquist, Quincey LaPlant, John H. Healey, Olivier Elemento, Christine H. Chung, Nancy Y. Lee, Marcin Imielenski, Gouri Nanjangud, Dana Pe'er, Don W. Cleveland, Simon N. Powell, Jan Lammerding, Charles Swanton, and Lewis C. Cantley. Chromosomal instability drives metastasis through a cytosolic DNA response. *Nature*, 553(7689):467–472, January 2018.

[37] Daniel C. Koboldt, Robert S. Fulton, Michael D. McLellan, Heather Schmidt, Joelle Kalicki-Veizer, Joshua F. McMichael, Lucinda L. Fulton,

David J. Dooling, Li Ding, Elaine R. Mardis, Richard K. Wilson, Adrian Ally, Miruna Balasundaram, Yaron S. N. Butterfield, Rebecca Carlsen, Candace Carter, Andy Chu, Eric Chuah, Hye-Jung E. Chun, Robin J. N. Coope, Noreen Dhalla, Ranabir Guin, Carrie Hirst, Martin Hirst, Robert A. Holt, Darlene Lee, Haiyan I. Li, Michael Mayo, Richard A. Moore, Andrew J. Mungall, Erin Pleasance, A. Gordon Robertson, Jacqueline E. Schein, Arash Shafiei, Payal Sipahimalani, Jared R. Slobodan, Dominik Stoll, Angela Tam, Nina Thiessen, Richard J. Varhol, Natasja Wye, Thomas Zeng, Yongjun Zhao, Inanc Birol, Steven J. M. Jones, Marco A. Marra, Andrew D. Cherniack, Gordon Saksena, Robert C. Onofrio, Nam H. Pho, Scott L. Carter, Steven E. Schumacher, Barbara Tabak, Bryan Hernandez, Jeff Gentry, Huy Nguyen, Andrew Crenshaw, Kristin Ardlie, Rameen Beroukhim, Wendy Winckler, Gad Getz, Stacey B. Gabriel, Matthew Meyerson, Lynda Chin, Peter J. Park, Raju Kucherlapati, Katherine A. Hoadley, J. Todd Auman, Cheng Fan, Yidi J. Turman, Yan Shi, Ling Li, Michael D. Topal, Xiaping He, Hann-Hsiang Chao, Aleix Prat, Grace O. Silva, Michael D. Iglesia, Wei Zhao, Jerry Usary, Jonathan S. Berg, Michael Adams, Jessica Booker, Junyuan Wu, Anisha Gulabani, Tom Bodenheimer, Alan P. Hoyle, Janae V. Simons, Matthew G. Soloway, Lisle E. Mose, Stuart R. Jefferys, Saianand Balu, Joel S. Parker, D. Neil Hayes, Charles M. Perou, Simeen Malik, Swapna Mahurkar, Hui Shen, Daniel J. Weisenberger, Timothy Triche Jr, Phillip H. Lai, Moiz S. Bootwalla, Dennis T. Maglinte, Benjamin P. Berman, David J. Van Den Berg, Stephen B. Baylin, Peter W. Laird, Chad J. Creighton, Lawrence A. Donehower, Gad Getz, Michael Noble, Doug Voet, Gordon Saksena, Nils Gehlenborg, Daniel DiCara, Juinhua Zhang, Hailei Zhang, Chang-Jiun Wu, Spring Yingchun Liu, Michael S. Lawrence, Lihua Zou, Andrey Sivachenko, Pei Lin, Petar Stojanov, Rui Jing, Juok Cho, Raktim Sinha, Richard W. Park, Marc-Danie Nazaire, Jim Robinson, Helga Thorvaldsdottir, Jill Mesirov, Peter J. Park, Lynda Chin, Sheila Reynolds, Richard B. Kreisberg, Brady Bernard, Ryan Bressler, Timo Erkkila, Jake Lin, Vesteinn Thorsson, Wei Zhang, Ilya Shmulevich, Giovanni Ciriello, Nils Weinhold, Nikolaus Schultz, Jianjiong Gao, Ethan Cerami, Benjamin Gross, Anders Jacobsen, Rileen Sinha, B. Arman Aksoy, Yevgeniy Antipin, Boris Reva, Ronglai Shen, Barry S. Taylor, Marc Ladanyi, Chris Sander, Pavana Anur, Paul T. Spellman, Yiling Lu, Wenbin Liu, Roel R. G. Verhaak, Gordon B. Mills, Rehan Akbani, Nianxiang Zhang, Bradley M.

Broom, Tod D. Casasent, Chris Wakefield, Anna K. Unruh, Keith Baggerly, Kevin Coombes, John N. Weinstein, David Haussler, Christopher C. Benz, Joshua M. Stuart, Stephen C. Benz, Jingchun Zhu, Christopher C. Szeto, Gary K. Scott, Christina Yau, Evan O. Paull, Daniel Carlin, Christopher Wong, Artem Sokolov, Janita Thusberg, Sean Mooney, Sam Ng, Theodore C. Goldstein, Kyle Ellrott, Mia Grifford, Christopher Wilks, Singer Ma, Brian Craft, Chunhua Yan, Ying Hu, Daoud Meerzaman, Julie M. Gastier-Foster, Jay Bowen, Nilsa C. Ramirez, Aaron D. Black, Robert E. Pyatt, Peter White, Erik J. Zmuda, Jessica Frick, Tara M. Lichtenberg, Robin Brookens, Myra M. George, Mark A. Gerken, Hollie A. Harper, Kristen M. Leraas, Lisa J. Wise, Teresa R. Tabler, Cynthia McAllister, Thomas Barr, Melissa Hart-Kothari, Katie Tarvin, Charles Saller, George Sandusky, Colleen Mitchell, Mary V. Iacocca, Jennifer Brown, Brenda Rabeno, Christine Czerwinski, Nicholas Petrelli, Oleg Dolzhansky, Mikhail Abramov, Olga Voronina, Olga Potapova, Jeffrey R. Marks, Wiktoria M. Suchorska, Dawid Murawa, Witold Kycler, Matthew Ibbs, Konstanty Korski, Arkadiusz Spychała, Paweł Murawa, Jacek J. Brzeziński, Hanna Perz, Radosław Łaźniak, Marek Teresiak, Honorata Tatka, Ewa Leporowska, Marta Bogusz-Czerniewicz, Julian Malicki, Andrzej Mackiewicz, Maciej Wiznerowicz, Xuan Van Le, Bernard Kohl, Nguyen Viet Tien, Richard Thorp, Nguyen Van Bang, Howard Sussman, Bui Duc Phu, Richard Hajek, Nguyen Phi Hung, Tran Viet The Phuong, Huynh Quyet Thang, Khurram Zaki Khan, Robert Penny, David Mallery, Erin Curley, Candace Shelton, Peggy Yena, James N. Ingle, Fergus J. Couch, Wilma L. Lingle, Tari A. King, Ana Maria Gonzalez-Angulo, Gordon B. Mills, Mary D. Dyer, Shuying Liu, Xiaolong Meng, Modesto Patangan, The Cancer Genome Atlas Network, Genome sequencing centres: Washington University in St Louis, Genome characterization centres: BC Cancer Agency, Broad Institute, Brigham & Women's Hospital & Harvard Medical School, Chapel Hill University of North Carolina, University of Southern California/Johns Hopkins, Genome data analysis: Baylor College of Medicine, Institute for Systems Biology, Memorial Sloan-Kettering Cancer Center, Oregon Health & Science University, The University of Texas MD Anderson Cancer Center, Santa Cruz/Buck Institute University of California, NCI, Biospecimen core resource: Nationwide Children's Hospital Biospecimen Core Resource, Tissue source sites: ABS-IUPUI, Christiana, Cureline, Duke University

Medical Center, The Greater Poland Cancer Centre, ILSBio, International Genomics Consortium, Mayo Clinic, MSKCC, and MD Anderson Cancer Center. Comprehensive molecular portraits of human breast tumours. *Nature*, 490(7418):61–70, October 2012. Bandiera_abtest: a Cc_license_type: cc_y Cg_type: Nature Research Journals Number: 7418 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Breast cancer;Cancer genetics;Molecular biology Subject_term_id: breast-cancer;cancer-genetics;molecular-biology.

[38] Marcel Ramos, Ludwig Geistlinger, Sehyun Oh, Lucas Schiffer, Rimsha Azhar, Hanish Kodali, Ino de Bruijn, Jianjiong Gao, Vincent J. Carey, Martin Morgan, and Levi Waldron. Multiomic Integration of Public Oncology Databases in Bioconductor. *JCO Clinical Cancer Informatics*, (4):958–971, November 2020. Publisher: Wolters Kluwer.

[39] Yuqing Zhang, Giovanni Parmigiani, and W. Evan Johnson. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR genomics and bioinformatics*, 2(3):lqaa078, September 2020.

[40] Sonja Hänzelmann, Robert Castelo, and Justin Guinney. GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics*, 14(1):7, January 2013.

[41] David Croft, Antonio Fabregat Mundo, Robin Haw, Marija Milacic, Joel Weiser, Guanming Wu, Michael Caudy, Phani Garapati, Marc Gillespie, Maulik R. Kamdar, Bijay Jassal, Steven Jupe, Lisa Matthews, Bruce May, Stanislav Palatnik, Karen Rothfels, Veronica Shamovsky, Heeyeon Song, Mark Williams, Ewan Birney, Henning Hermjakob, Lincoln Stein, and Peter D'Eustachio. The Reactome pathway knowledgebase. *Nucleic Acids Research*, 42(Database issue):D472–D477, January 2014.

[42] DA Rew and GD Wilson. Cell production rates in human tissues and tumours and their significance. Part II: clinical data. *European Journal of Surgical Oncology (EJSO)*, 26(4):405–417, 2000. Publisher: Elsevier.

[43] Karin Haustermans, Lucien Vanuytsel, Karel Geboes, T Lerut, J Van Thillo, J Leysen, Willy Coosemans, and E Van Der Schueren. In vivo cell kinetic measurements in human oesophageal cancer: what can be learned from multiple biopsies? *European Journal of Cancer*, 30(12):1787–1791, 1994. Publisher: Elsevier.

[44] Seung Joon Choi, Hyung-Sik Kim, Su-Joa Ahn, Yu Mi Jeong, and Hye-Young Choi. Evaluation of the growth pattern of carcinoma of colon and rectum by MDCT. *Acta Radiologica*, 54(5):487–492, 2013. Publisher: SAGE Publications Sage UK: London, England.

[45] Mitsunobu IDE, Minoru Jimbo, Masaaki Yamamoto, Yutaka Umebara, Shinji Hagiwara, and Osami Kubo. Growth rate of intracranial meningioma: tumor doubling time and proliferating cell nuclear antigen staining index. *Neurologia medico-chirurgica*, 35(5):289–293, 1995. Publisher: The Japan Neurosurgical Society.

[46] S Bolin, E Nilsson, and R Sjödahl. Carcinoma of the colon and rectum–growth rate. *Annals of surgery*, 198(2):151, 1983. Publisher: Lippincott, Williams, and Wilkins.

[47] RG Margolese and Buchholz TA HG. Natural history and prognostic markers. *Holland-Frei Cancer Medicine. 6th edition ed. Hamilton (ON): BC Decker*, 2003.

[48] GM Zharinov and VA Gushchin. The rate of tumor growth and cell loss in cervical cancer. *Voprosy onkologii*, 35(1):21–25, 1989.

[49] J Andrew Carlson. Tumor doubling time of cutaneous melanoma and its metastasis. *The American journal of dermatopathology*, 25(4):291–299, 2003. Publisher: LWW.

[50] Kassem Harris, Inga Khachaturova, Basem Azab, Theodore Maniatis, Srujitha Murukutla, Michel Chalhoub, Hassan Hatoum, Thomas Kilkenny, Dany Elsayegh, Rabih Maroun, and others. Small cell lung cancer doubling time and its effect on clinical presentation: a concise review. *Clinical Medicine Insights: Oncology*, 6:CMO–S9633, 2012. Publisher: SAGE Publications Sage UK: London, England.

[51] Katja Roesch, Dirk Hasenclever, and Markus Scholz. Modelling lymphoma therapy and outcome. *Bulletin of mathematical biology*, 76(2):401–430, 2014. Publisher: Springer.

[52] Richard S Varga. *Geršgorin and his circles*, volume 36. Springer Science & Business Media, 2010.

[53] Uri Ben-David and Angelika Amon. Context is everything: aneuploidy in cancer. *Nature Reviews. Genetics*, 21(1):44–62, January 2020.

[54] Jake C Forster, Michael JJ Douglass, Wendy M Harriss-Phillips, and Eva Bezak. Simulation of head and neck cancer oxygenation and doubling time in a 4D cellular model with angiogenesis. *Scientific reports*, 7(1):1–11, 2017. Publisher: Nature Publishing Group.

[55] Neil J Ganem, Zuzana Storchova, and David Pellman. Tetraploidy, aneuploidy and cancer. *Current opinion in genetics & development*, 17(2):157–162, 2007. Publisher: Elsevier.

[56] Qinghua Shi and Randall W. King. Chromosome nondisjunction yields tetraploid rather than aneuploid cells in human cell lines. *Nature*, 437(7061):1038–1042, October 2005.

[57] Karen Crasta, Neil J Ganem, Regina Dagher, Alexandra B Lantermann, Elena V Ivanova, Yunfeng Pan, Luigi Nezi, Alexei Protopopov, Dipanjan Chowdhury, and David Pellman. DNA breaks and chromosome pulverization from errors in mitosis. *Nature*, 482(7383):53–58, 2012. Publisher: Nature Publishing Group.

[58] Cheng-Zhong Zhang, Alexander Spektor, Hauke Cornils, Joshua M Francis, Emily K Jackson, Shiwei Liu, Matthew Meyerson, and David Pellman. Chromothripsis from DNA damage in micronuclei. *Nature*, 522(7555):179–184, 2015. Publisher: Nature Publishing Group.

[59] Emily M Hatch, Andrew H Fischer, Thomas J Deerinck, and Martin W Hetzer. Catastrophic nuclear envelope collapse in cancer cell micronuclei. *Cell*, 154(1):47–60, 2013. Publisher: Elsevier.

[60] Salehi S, Kabeer F, Ceglia N, Andronescu M, Williams Mj, Campbell Kr, Masud T, Wang B, Biele J, Brimhall J, Gee D, Lee H, Ting J, Zhang Aw, Tran H, O'Flanagan C, Dorri F, Rusk N, de Algara Tr, Lee Sr, Cheng Byc, Eirew P, Kono T, Pham J, Grewal D, Lai D, Moore R, Mungall Aj, Marra Ma, undefined, McPherson A, Bouchard-Côté A, Aparicio S, and Shah Sp. Clonal fitness inferred from time-series modelling of single-cell cancer genomes. *Nature*, 595(7868):585–590, June 2021.

[61] Noemi Andor, Billy T. Lau, Claudia Catalanotti, Anuja Sathe, Matthew A. Kubit, Jiamin Chen, Susan M. Grimes, Carlos Jose Suarez,

and Hanlee P. Ji. Joint single cell DNA-seq and RNA-seq of gastric cancer cell lines reveals rules of in vitro evolution. *NAR Genomics and Bioinformatics (in press)*, March 2020.

[62] Sean Becketti and William Gould. Rangefinder Box Plots: A Note. *The American Statistician*, 41(2):149–149, May 1987. Publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/00031305.1987.10475467.