# Genetic Load and Potential Mutational Meltdown in Cancer Cell Populations

Yuezheng Zhang,[†,1] Yawei Li,[†,1,3] Tao Li,[†,1,3] Xu Shen,[†,1,4] Tianqi Zhu,[5,6] Yong Tao,[1] Xueying Li,[7] Di Wang,[1] Qin Ma,[1,3] Zheng Hu,[1] Jialin Liu,[1] Jue Ruan,[1] Jun Cai,[1,3] Hurng-Yi Wang,[*,8,9] and Xuemei Lu[*,1,2,3]

[1]Key Laboratory of Genomics and Precision Medicine, Beijing Institute of Genomics, Beijing, China

[2]CAS Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

[3]University of Chinese Academy of Sciences, Beijing, China

[4]State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-Sen University, Guangzhou, China

[5]National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

[6]Key Laboratory of Random Complex Structures and Data Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

[7]School of Life Sciences, Peking University, Beijing, China

[8]Graduate Institute of Clinical Medicine, National Taiwan University, Taipei, Taiwan

[9]Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei, Taiwan

[†]These authors contributed equally to this work.

[*]**Corresponding authors:** E-mails: hurngyi@ntu.edu.tw; luxm@big.ac.cn.

**Associate editor:** Jeffrey Townsend

## Abstract

**Large genomes with elevated mutation rates are prone to accumulating deleterious mutations more rapidly than natural selection can purge (Muller's ratchet). As a consequence, it may lead to the extinction of small populations. Relative to most unicellular organisms, cancer cells, with large and nonrecombining genome and high mutation rate, could be particularly susceptible to such "mutational meltdown." However, the most common type of mutation in organismal evolution, namely, deleterious mutation, has received relatively little attention in the cancer biology literature. Here, by monitoring single-cell clones from HeLa cell lines, we characterize deleterious mutations that retard the rate of cell proliferation. The main mutation events are copy number variations (CNVs), which, estimated from fitness data, happen at a rate of 0.29 event per cell division on average. The mean fitness reduction, estimated reaching 18% per mutation, is very high. HeLa cell populations therefore have very substantial genetic load and, at this level, natural population would likely face mutational meltdown. We suspect that HeLa cell populations may avoid extinction only after the population size becomes large enough. Because CNVs are common in most cell lines and tumor tissues, the observations hint at cancer cells' vulnerability, which could be exploited by therapeutic strategies.**

*Key words:* genetic load, Muller's ratchet, HeLa cell line, copy number variation.

## Introduction

It is now generally accepted that somatic mutations promoting cell proliferation is the ultimate driving force of tumor evolution. In addition to the driver mutations, which provide cells a fitness advantage, mutations of no fitness consequences (or neutral "passenger" mutations) are believed to be common as well (Ling et al. 2015; Tao et al. 2015; Li et al. 2017). Intriguingly, the most common type of mutations in organismal evolution, namely, deleterious mutations (Li 1997), receives little attention in the cancer biology literatures despite the strong signal in cancer genome sequences (Wu et al. 2016). As aforementioned, the driver/passenger nomenclature excludes this class of mutations, which may be referred to as negative drivers (Wu et al. 2016). Although negative driver mutations may be several times more common than positive drivers among single nucleotide variants (SNVs; Wu et al. 2016), SNVs in coding regions occur rather infrequently, at no >0.01 per genome per cell division. In this study, we aim to detect deleterious copy number variation (CNV) mutations, which occur at a much higher rate than SNVs.

Under most circumstances, deleterious mutations are eliminated by natural selection shortly after their emergence and have little evolutionary consequences. However, there are cases that natural selection cannot purge the cells carrying deleterious mutations fast enough as they accumulate. The best-known circumstances are large genomes with high mutation rates, small population sizes, and in the absence of recombination (Muller 1964; Charlesworth 1978; Haigh 1978; Lynch and Gabriel 1990; Bachtrog and Gordo 2004). Human tumor cells meet all these conditions. First, the

human cancer genomes, larger than those of most unicellular organisms, have correspondingly higher mutation rates per cell. Second, during the tumorigenesis the cell population size is usually very small (Wu et al. 2016; Chen et al. 2018). Lastly, recombination allows natural selection to operate efficiently in removing individual deleterious mutations. Somatic cells do not recombine and selection has to operate at the level of whole genome, thus, canceling out selection on deleterious and advantageous mutations of the same cell. This phenomenon is often referred to as the Hill–Robertson effect (Hill and Robertson 1966; Felsenstein 1974; Hartl and Clark 2007).

In populations of nonrecombining genomes, deleterious mutations will accumulate as no genome can reduce the mutation number via recombination. A population's fitness peak is thus determined by its least loaded class; that is, the individuals (or cells) with the smallest number of deleterious mutations. Evolution does not reverse the accumulation of mutations and is compared with a ratchet mechanism (Muller 1964; Felsenstein 1974; Haigh 1978). When the least loaded class has one more deleterious mutation than in previous generations, it is said that Muller's ratchet has advanced one notch (Gordo and Charlesworth 2000a; Charlesworth 2012). If the ratchet continues to advance, outpacing the ability of natural selection to reign in the mutation accumulation (MA), a possible consequence is the so-called mutational meltdown (Lynch and Gabriel 1990; Gabriel et al. 1993; Lynch et al. 1993) whereby the population would evolve to extinction.

Direct measurements of the effects of individual mutations in vivo are challenging. Nevertheless, an assessment of their collective action is possible, given a system that generates such variants at an appreciable rate. HeLa cells present such a system because it has been extensively cultured and exhibits a short doubling time and aggressive growth. In this study, HeLa cells are not used to reveal the process of tumorigenesis but mainly a model for addressing the underlying evolutionary forces, which needs to be powerful enough to be measured in laboratory settings.

We examined variation in growth rates among individual HeLa cells by monitoring clones from a common ancestral population. To interpret the variation, we performed whole genome sequencing of the clones. The genomic variation reveals that the growth rate is associated with the total length and number of CNVs. We perform computer simulations of cell growth to estimate deleterious mutation rate and the average fitness decrease per mutation. Our results suggest that human cells that have been cultured for a sufficiently long period of time still generate deleterious mutations in the form of CNVs at a high rate and a high-intensity level. For such systems, the operation of Muller's ratchet leading to mutational meltdown might be plausible.

## Results

### Genetic Variation in Growth Rate in a HeLa Cell Population

To ensure genomic homogeneity of the initial population, we first established a HeLa cell line (E6) derived from an ancestral cell line (JF; supplementary fig. S1, Supplementary Material online). When E6 population size reached $\sim 5 \times 10^4$ cells (after 15–16 divisions), five single-cell clones were generated from E6 and established in culture. When the clone cultures reached $10^6$–$10^7$ cells, we measured cell proliferation rate of these clones using the MTT assay. The B8 and G3 clones proliferated faster than E3, E7, and G2, suggesting that cells within E6 were heterogeneous (fig. 1A). Because these clones were all descendants of E6, which originated from a single cell and experienced only 15–16 divisions, our results indicate that heterogeneity in cell growth can be generated in a very short period of time in cancer cells.

To test whether variation in growth rates among clones is heritable, we isolated 39 cells from B8 (fast-growing clone) and 40 cells from E3 (slow-growing clone), and monitored their growth from a single cell for seven days. Approximately 23% of B8 and 50% of E3 cells died out within seven days (supplementary fig. S2, Supplementary Material online), due to either damage caused during cell isolation or genetic defects. Furthermore, most cell lines with growth rates < 0.6 died within 2 months. Thus, only 60% of B8 and 27% of E3 cells survived for more than two months.

Whereas the growth rates of cells from a single origin varied greatly, the progeny of the fast-growing clone (B8) grew faster on average than those of the slow-growing clone (E3; fig. 1B; P-value = 0.004, KS test). Mean cell number among the B8 progeny was 62.0 at day 7, whereas it was 17.3 among E3-derived cell lines at the same time. The estimated heritability was 0.24 (P-value = 0.003), implying a significant genetic contribution to the growth rate phenotype. We further drew $\sim 100$ cells from one fast- (progeny D) and one slow- (progeny C; supplementary fig. S2, Supplementary Material online) growing derivative of B8 and monitored growth of each cell clone for seven days. As expected, cells from the fast-growing progeny had a higher mean growth rate than that from the slow-growing descendants (P-value = $4 \times 10^{-11}$; fig. 1C) and the heritability was 0.32 (P-value $< 3 \times 10^{-11}$).

If mutations that slow cell proliferation frequently arises in cancer cell populations, we would expect a decrease in proliferation rates as the population is maintained. To test this, we monitored growth of single cells that were randomly drawn from cell populations at different time points. We first set up six single-cell clones from B8 (supplementary table S1, Supplementary Material online). After cell numbers reached 100–500 (8–10 cell divisions), 20–30% of the cells from each population were randomly drawn and isolated into single cells. The growth of these isolates was monitored for eight days. This step was repeated when the size of the six populations exceeded 5,000 cells. In all six cases, the growth rates of cells drawn at the first time point were higher than those from the second time point (P-value = $1.08 \times 10^{-8}$, ANOVA). In addition, in four out of six cases, growth rate was higher at the first than at the second time point (supplementary fig. S3, Supplementary Material online; t-test).

To test whether the slow-growing cells would be outcompeted by their fast-growing counterparts, we performed a competition assay. The fast- and slow-growing cells with the stable expression of different fluorescent proteins,
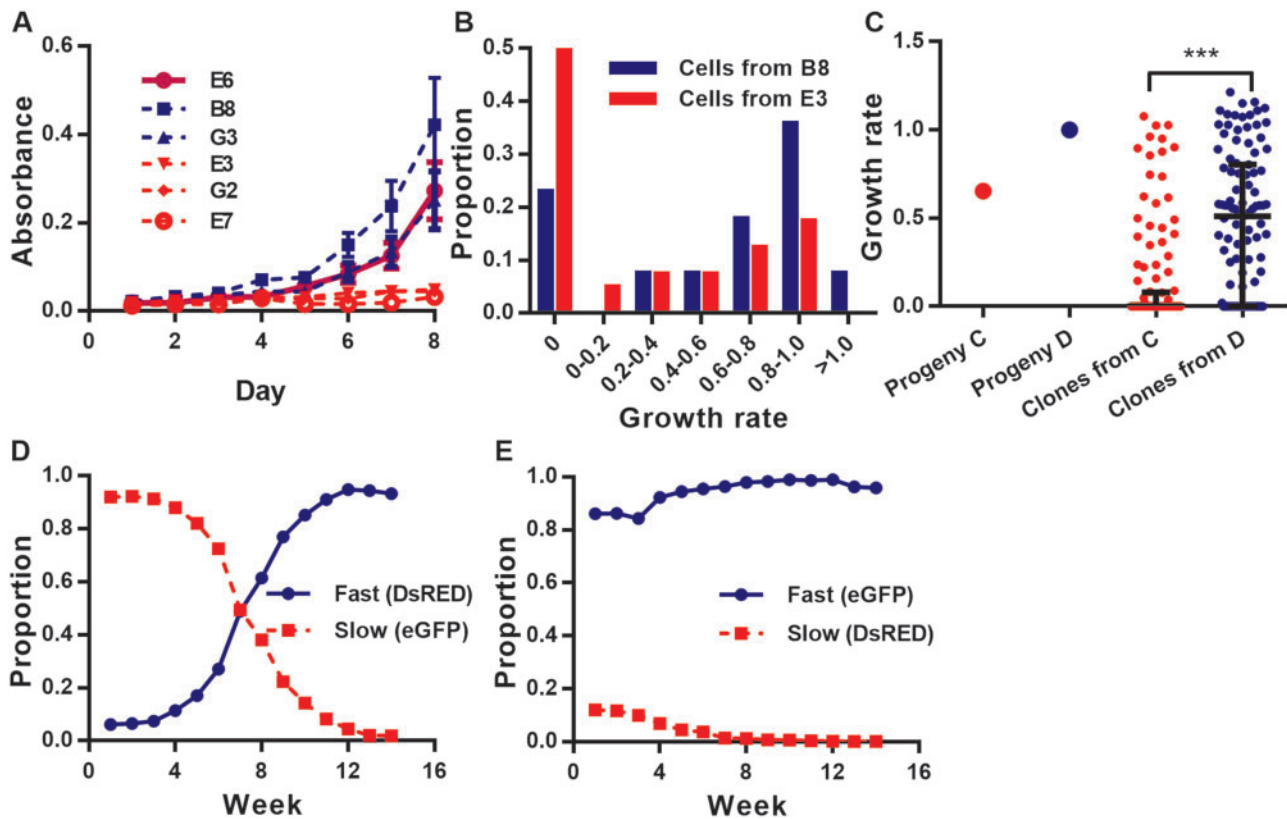
**FIG. 1.** Growth rate, survival rate and fitness of fast- and slow-growing cells. (A) Growth graphs of five descendant clones (B8, G3, E3, E7, G2) and the ancestral clone (E6). The mean MTT assay read-out (the values of absorbance, y-axis) taken from day 1 to day 8 were plotted for each clone (8–12 replicates). Blue, red, and magenta lines represent growth graphs of fast, slow, and ancestral clones, respectively. Error bars represent 95% confidence interval. (B) Proportion of progeny with different growth rates. The growth rates of 39 B8 (red) and 40 E3 (blue) single cell progeny were monitored and calculated for seven days. The growth rates were represented by the slope value that calculated based on cell number counting from 7 days' cell culture (see Materials and Methods). (C) Growth rate of one slow- (C; red) and one fast- (D; blue) growing descendant of B8 and its single-cell clones. Average growth rate of the cell clones from D (blue dots) is significantly higher than that of the clones from C (red dots) (P-value $= 4 \times 10^{-11}$, KS test). (D and E) Competition assay between slow- and fast-growing cells. The proportion (y-axis) of fast- (blue) and slow-growing (red) cells in a mixed population was measured from week 1 to week 8 (x-axis) by flow cytometry.

*Discosoma* sp. red fluorescence protein (DsRed) or enhanced green fluorescence protein (eGFP), were constructed, and cocultured in different ratios. The expression of eGFP and DsRed was used to distinguish the two types of cells in the competition assay. Their proportions in populations were monitored weekly over time using flow cytometry. We ran the experiment for eight weeks. Regardless of the initial proportions, the fast-growing cells always outgrew the slow-growing cells in our coculture assays over time (fig. 1D and E), suggesting that fast cells indeed possessed higher fitness than the slow ones.

Figure 1 demonstrates that a cancer cell population can generate heterogeneity in growth rate within several cell replications, even starting from a single cell. Moreover, the majority of these changes is heritable and reduce the fitness of cancer cells (defined as proliferation rate), suggesting that the fitness reduction we observed is largely genetically determined.

## Cell Growth Rate Is Correlated with the Total Length and Number of CNVs

To study the genetic basis of growth rate heterogeneity among our cancer cell lines, we assessed genomic variation in E6 and five of its descendant clones by constructing a digital copy-number profile based on next generation sequencing reads. We focused on CNVs because single-nucleotide mutation rates are too slow to produce significant sequence variation during our short-duration culturing experiments.

Because the sequence depth of these clones was only 10X, CNVs found in each clone were presumably originated from ancestral clone (see Materials and Methods). The number of CNVs ranged from 2 (B8 and G3) to 9 with the average of 5.2 (supplementary table S2, Supplementary Material online). The three slow-growing clones showed a clear increase in the number and the total length of CNVs compared with the E6 parental line (fig. 2, supplementary table S2, Supplementary Material online), whereas the two fast-growing clones were more similar to E6, suggesting that that most CNVs are deleterious. As E6 experienced only 15–16 divisions before its five descendant clones were generated, above results also indicate that CNVs can be generated in a very short period of time. To test this hypothesis, 11 clones derived from B8 with different growth rates between day 1 and 8 were picked for further analyses (fig. 3A).
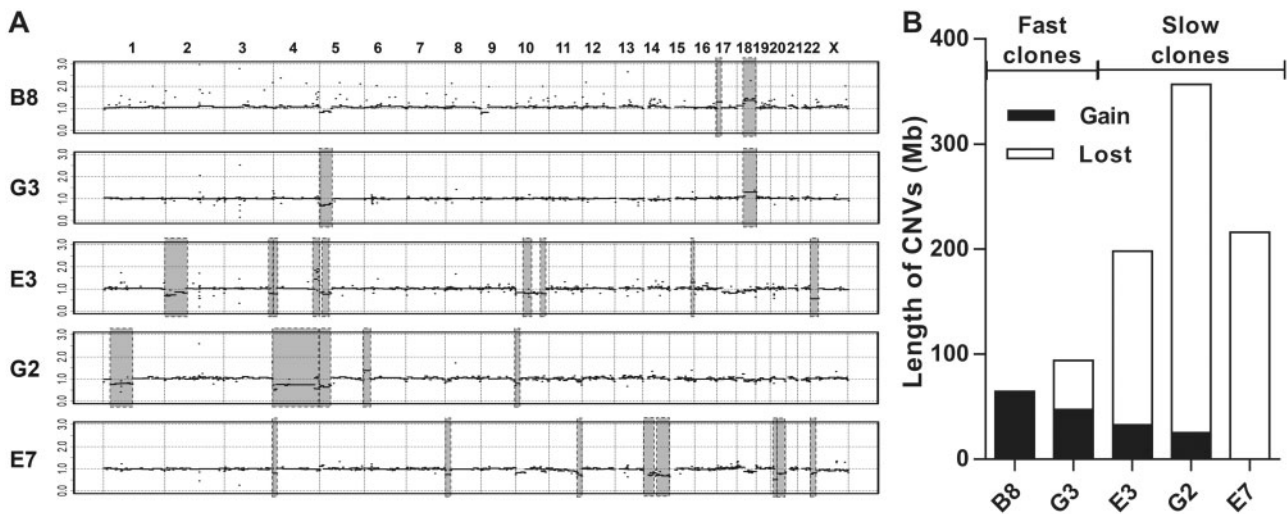
**Fig. 2.** Copy number variation (CNVs) in five clones derived from E6. (*A*) The CNV regions in fast- (B8 and G3) and slow-growing (E3, E7, and G2) clones are highlighted with grey rectangles. The *y*-axis is the ratio of copy number in a descendant clone divided by copy number in the ancestor, E6. (*B*) Summary of CNV gain and loss among five descendant clones (see also supplementary table S2, Supplementary Material online).

The growth rate of each clone was measured again by RTCA iCELLigence when the population reached $\sim 10^6$ cells (after 20–30 cell divisions). The results were highly correlated ($R^2 = 0.713$, *P*-value $< 0.01$; fig. 3*B*) with the previous eight-day measurements, demonstrating that variation in growth rates among clones was consistent and not due to stochastic fluctuation at different stages. We picked $\sim 20$ cells from each of the 11 clones and counted their chromosome numbers. The karyotypes ranged from 38 to 113 chromosomes, with most (72%) cells harboring between 55 and 70 chromosomes (fig. 3*C*), indicating that they are triploid as shown in a previous study (Adey et al. 2013). Therefore, despite single-cell origin, the progeny quickly generated aneuploidy within only 20–30 cell divisions, again illustrating frequent cytogenetic change in cancer cells.

Although chromosome numbers varied among clones, their distribution by combing all clones reconstituted the chromosome distribution of their ancestor B8. The average number of chromosomes in the fast-growing group (62.5) was similar (*P*-value = 0.56, *t*-test) to the B8 clone (61.1), whereas the slow-growing group (66.5) showed significantly greater chromosome numbers than the fast-growing (*P*-value $< 0.01$) and the ancestral (*P*-value = 0.04) clones (fig. 3*D*). Consequently, although the number of observations is still small, figures 2 and 3 suggest that cancer cells exhibit a very high rate of CNV generation and most of these CNVs are deleterious, reducing the cells' proliferative ability.

## Modeling Population Growth Attenuation

To understand MA rates in tumors, we need quantitative estimates of deleterious mutation rates and effect distributions. Therefore, we constructed and applied a simple model of cell growth and the MA process (see Materials and Methods). Let $N_t$ be the population size at day $t$, where $N_0 = 1$, and $R_t$ be the population growth rate at day $t$. For each generation, a proportion of cells ($\mu$) generate new mutations that decrease their growth rate. The average deleterious effect of a mutation is $s$. We have

$$N_t = R_0^t (1 - \mu s)^{t(t-1)/2}, \qquad (1)$$

where $R_0$ is the growth rate at day 0.

To estimate parameters of this model ($R_0$, $\mu$, and $s$) we randomly drew 18 single cells from B8 and monitored their growth for 7–8 days (supplementary table S3, Supplementary Material online). We divided these newly derived cell lines into fast- (cell number $> 100$) and slow-growing (cell number $< 100$) groups. We then conducted computer simulations to evaluate the parameters that best fit the observed data (see Materials and Methods).

Using fast-growing cell data, we estimate posterior mean of $R_{0,fast} = 2.37$ ($R_{0,fast} \in [2.22, 2.52]$), $\mu_{fast} = 0.29$ ($\mu_{fast} \in [0.26, 0.30]$), and $s_{fast} = 0.18$ ($s_{fast} \in [0.17, 0.20]$) (supplementary table S3, Supplementary Material online). Estimates from slow-growing cells $R_{0,slow} = 2.00$ ($R_{0,slow} \in [1.66, 2.21]$), $\mu_{slow} = 0.29$ ($\mu_{slow} \in [0.25, 0.33]$), and $s_{slow} = 0.18$ ($s_{slow} \in [0.16, 0.20]$) are similar, except that the initial growth rate is slower, as expected. The deleterious mutation rate ($\mu$) of 0.29 (0.25, 0.33) suggests that there is approximately one deleterious mutation for every 3–4 cell divisions. Since $\mu$ is scaled per cell division, the product of $\mu$ and $s$ reflects fitness change per generation. Our estimates indicate that the HeLa cells experience a 5% (4%, 6%) reduction in fitness for every generation (Crow 1958). Using point estimates of $\mu$ (0.29) and $s$ (0.18), we fitted our model to the growth rates observed in a range of experimental data (supplementary fig. S2, Supplementary Material online) and estimated initial growth rates ($R_0$). Only cell lines that showed monotonic increase in cell numbers were considered. The estimation of $R_0$, ranging from 1.64 to 2.54 in 43 sets of experimental data (supplementary table S4, Supplementary Material online), suggested that the ancestral cell of slowest

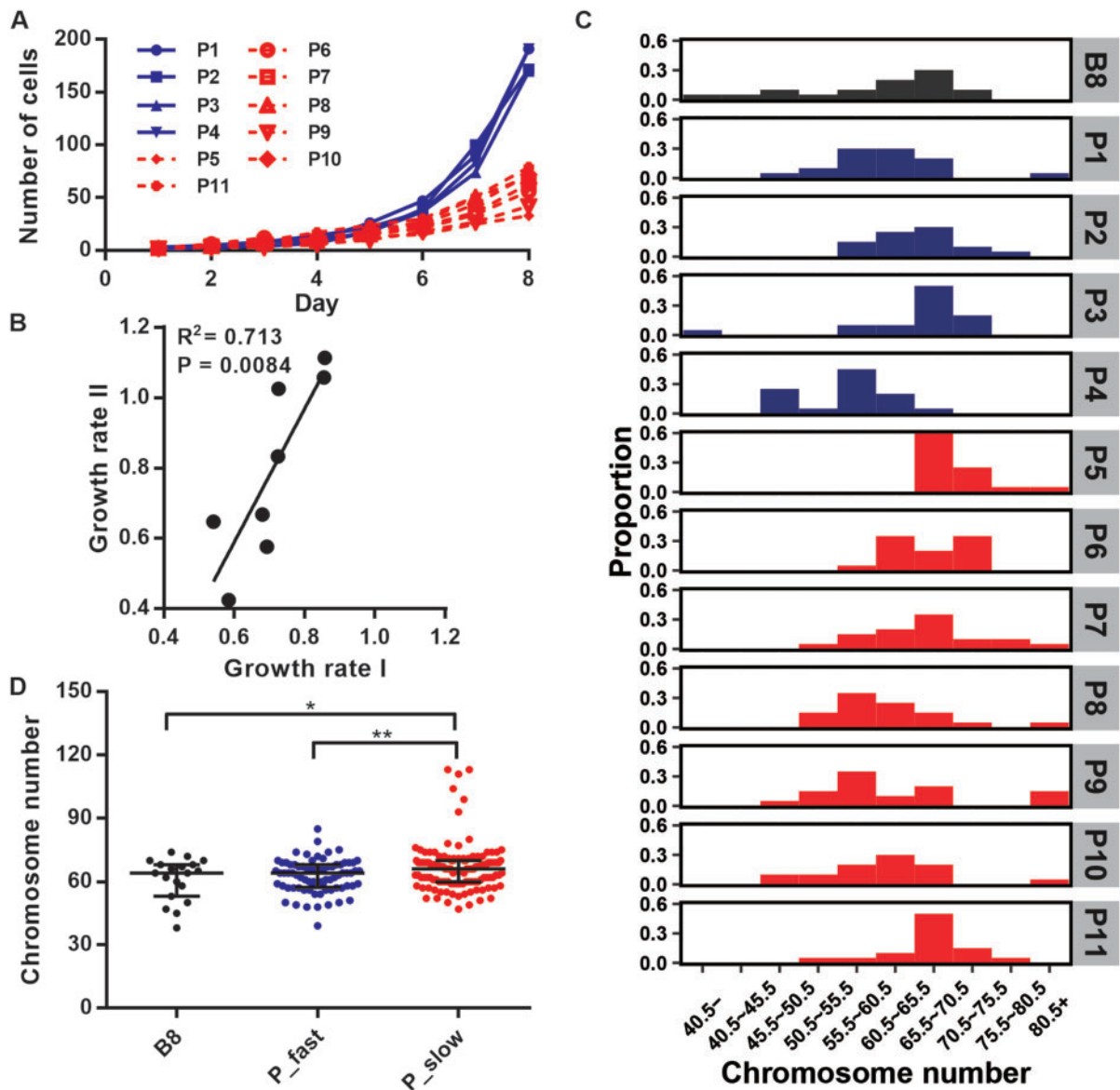**FIG. 3.** Growth rate and chromosome number variation among single-cell clones generated from B8. (*A*) Cell numbers of single-cell clones from day 1 to day 8. The clones labeled in blue grow faster than the clones in red. (*B*) Correlation between growth rate I and II of single-cell clones. The growth rate I (*x*-axis) was measured in the first eight days and the growth rate II (*y*-axis) measured by RTCA iCELLigence method when the cell populations reached $\sim 10^6$ cells. Each dot represents a single cell clone; only eight out of 11 clones were measured at the second time point. (*C*) Distributions of chromosome numbers in cells from ancestral and descendant clones. Chromosome numbers (*x*-axis) of 20–30 cells randomly drawn from each clone were counted. The black, blue, and red histograms each represent cells from the ancestral, fast, and slow progeny clones. (*D*) Chromosome number in cells from the ancestral (B8) and the progeny clones. The black, blue, and red dots represent cells from the ancestral, fast (P_fast), and slow (P_slow) progeny clones, respectively. *$P < 0.05$, **$P < 0.01$.

growing lineage had accumulated $\sim 2.2$ [log (1.64/2.54)/log (1−0.18)] more deleterious mutations than the ancestor of fastest growing lineage.

To further empirically test the estimates of $\mu$, we picked 12 cells from B8, allowed them to divide once, and isolated each daughter cell in a well of a 96-well plate. We then monitored the growth of each of the 12 pairs of cells for four days (supplementary fig. S4, Supplementary Material online). Four cell pairs lost at least one of the daughters in the first four days, probably due to injury during preparation. Of the remaining eight pairs, two (h and i, supplementary fig. S4, Supplementary Material online) exhibited different

proliferation rates between the daughter cell lineages. The result suggests that there is approximately one deleterious mutation in every four cell divisions in this cancer cell line, which is very close to parametric estimates. Such a high deleterious mutation rate implies that many cells at day four would carry deleterious mutations that reduce cell proliferation. To test this, each descendant of clone f was harvested and resuspended into a 96-well plate. We monitored growth of these progeny for another four days. As expected, the majority of the surviving clones exhibited slower growth than their ancestor f (supplementary fig. S5, Supplementary Material online).

**Table 1.** Fitness Reduction Per Cell Division in Different Cancer Cell Lines.

| Cell Line | Cancer Type | Number of Clones[a] | $R_0$[b] | $\mu$[c] | $s$[d] | $\mu \times s$ |
|---|---|---|---|---|---|---|
| Hela | Cervical cancer | 18 | 1.66–2.48 | 0.25–0.33 | 0.17–0.20 | 0.04–0.06 |
| PC3 | Prostate cancer | 10 | 1.68–2.11 | 0.26–0.31 | 0.16–0.20 | 0.04–0.06 |
| A204 | Rhabdomyosarcoma | 7 | 1.81–2.07 | 0.28–0.33 | 0.17–0.22 | 0.05–0.07 |
| A375 | Melanoma | 17 | 2.14–2.64 | 0.26–0.34 | 0.16–0.23 | 0.05–0.08 |

NOTE.—To estimate parameters in equation (1), single cells were randomly chosen from different cancer cell lines (supplementary table S5, Supplementary Material online) and their growth from day 0 to day 7 was monitored. Fitness reduction was evaluated based on the parameters estimated by applying computer simulations, which are described in Results and Materials and Methods.
[a]The number of single cell clones generated from different cancer cell lines.
[b]Growth rate of cell population at day 0.
[c]Deleterious mutation rate.
[d]Average deleterious effect of a mutation.

## Rapid Accumulation of Deleterious Mutations in Different Types of Cancer Cell Lines

Interestingly, while $R_0$ shows substantial variation among Hela cell clones, estimates of $\mu$ and $d$ are similar, possibly indicative of an intrinsic property of rapid cancer cell division. To further test whether the property in the accumulation of deleterious mutations is consistent in cancer cells, we estimated the $R_0$, $\mu$, and $d$ in tumor cell lines from different cancer types by the same method (table 1). Although the $R_0$ are considerably variable, $\mu$ and $d$ are generally consistent among cell lines, implying that a fitness reduction ($\mu \times d$) in a variety of cancer cells is close to our initial estimate of 5% per generation.

## Modeling Genetic Load in Long-term Cancer Growth

High deleterious mutation rate would raise an impression that the HeLa cell lines may have gone extinct long ago. To address this issue, we modeled the dynamic of "long-term" cell growth starting from a single cell and monitored MA within populations. The aim is to evaluate the level of genetic load that affects cancer growth accompanied by continuous accumulation of deleterious mutations. In a total of 200 simulations, each population was originated from one single cell and the mutation status of all cells within population was recorded until the population size of the clone reached 5,000 (survived) or 0 (defected; fig. 4A, see Materials and Methods).

Nineteen clones (9.5%; fig. 4B) were still survived after 180 generations (days). We then randomly chose one clone and let it grew to the population size of $10^6$ and the mutation spectrum of this clone was obtained by randomly sampling of 1,000 cells (fig. 4C). As expected, the majority of cells (86.9%) had accumulated at least one deleterious mutation compared with their ancestral cell. The results also show that the slowest growing cells have accumulated five more deleterious mutations than the fastest growing cells, indicating a large diversity within the clone, a result similar to our observations in figures 1 and 3. Although most of the cells accumulated deleterious mutations and were worse than the ancestral cells, there were still 13.1% of cells that were mutation-free. According to the least-loaded theory (Gordo and Charlesworth 2000a), these growing number of mutation-free cells can prevent the population from extinction.

## Discussion

The increase in mutation rate and genome instability is one of the hallmarks during tumorigenesis, which not only generates advantageous mutations (drivers), but also accumulates large number of passengers that are neutral or deleterious. However, it remains unknown, first, how many deleterious mutations may accumulate in cancer cell populations. As the fixed and high frequency mutations that calculated in cancer genomic studies are usually advantageous or neutral for the cell fitness, deleterious mutations are hardly detected in empirical studies. The detection of a moderate proportion of negative selective mutations in TCGA pan-cancer data indicates the ignorance of deleterious mutations (Wu et al. 2016). Second, deleterious mutations may arise at a high rate and cannot be efficiently purged by natural selection in populations with no recombination, such as cancer cells (Haigh 1978; Maynard Smith 1978; Kondrashov 1982; Charlesworth 1990; Lynch and Gabriel 1990; Lynch et al. 1993; Wu et al. 2016). We, herein, evaluate whether high mutation rate and the irreversible accumulation of deleterious mutations might lead to mutational decay of cancer cell populations. Performing multiple single-cell progeny assays and extensively genotyping in the clonal populations of HeLa cell lines, we estimate the rate and effect of deleterious CNV mutations in cancer cell populations. The results show that deleterious CNV mutations appear at the rate that falls in the range of 0.26–0.33 per genome per generation with the mean deleterious effect of 0.18 ([0.175, 0.187], 95% confidence interval), indicating the high level of fitness reduction in cancer cell populations (Fischer et al. 2011; McFarland et al. 2013; Lynch 2016).

While our observation that majority of mutations was deleterious in cancer cell lines is similar to previous results derived mainly from MA experiments in multi-cellular organisms (Willi 2013; Katju et al. 2015; Lovell et al. 2017), these two types of studies differ in two major aspects. First, genomes of multi-cellular organisms evolve to ensure the survival and reproduction of the whole organisms. Cancer cells evolve from multi-cellular state toward a single-cell state that is distinct from their longstanding role. In that sense, similarities between our results and those in microbes (Kibota and Lynch 1996) may be unsurprising. In addition, Cancer cells usually exhibit a higher mutation rate than normal cells (Wu et al. 2016). While increased mutation rate likely imposes a cost, it also permits cancer cells to evolve new properties, such as
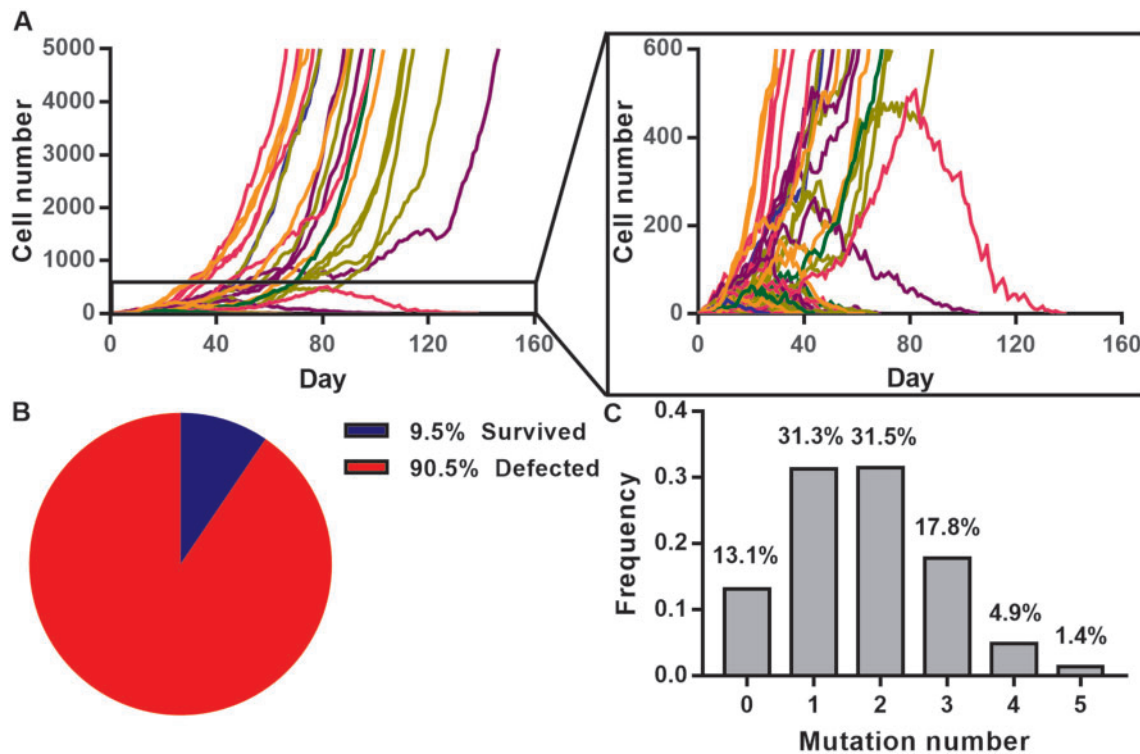
**Fig. 4.** Dynamics of cancer cell growth and their mutation distribution. (*A*) The growing trajectory of 200 tumor clones in silico. Each curve represents the dynamic of one single cell clone from day 0. (*B*) Among 200 replicates, only 19 clones survived successfully whereas most of the clones were defected (clone extinction). (*C*) The distribution of deleterious mutation number per cell from 1,000 selected cells. The demographic parameters are $R_0 = 2.0$, $\mu = 0.3$, and $s = 0.18$.

poly- and aneuploidy and other CNVs than normal cells can (Molaro and Malik 2017; Xu et al. 2017). A second major difference is that most MA experiments have been conducted under constant small population size ($Ne = 1$ or 10), including a strong bottleneck effect under which mutations with small effect, either deleterious or beneficial, can be stochastically fixed within populations. This stochasticity may explain mixed results that have been observed in a few MA studies (Shaw et al. 2000; Dickinson 2008; Woodruff 2013).

The mechanism of Muller's ratchet is determined by the population size ($N$), the mutation rate ($\mu$), and the selection coefficient ($s$) (Haigh 1978; Gordo and Charlesworth 2000a). Although we observed the high level of genetic load and the population decay in cancer cell lines, further sophisticated analysis, taken considering the key parameters, must be performed to reveal the process and the consequences in the tumors in patients. It may help reveal the process and the consequences in the tumors in patients and develop a new therapeutic regime for cancer.

The increase of $N$ and $s$ will let up the ratchet. Accordingly, as the tumors grow, the operation of Muller's ratchet may slow down, leading to the absent of mutational meltdown. Thus, tumor size would be positively correlated with the malignancy which has been shown in many cancer types, including breast, kidney, pancreatic and lung cancers (Agarwal et al. 2008; Thompson et al. 2009; Narod 2012; Zhang et al. 2015). On the contrary, the $s$ could be small in a large tumor. In solid tumors, the local cell populations do not migrate and are highly structured. Selection can be more inefficient in solid

tumors than in nonsolid tumors such as leukemia, because the cell populations are so spatially constrained that a selection sweep may not occur in solid tumors (Ling et al. 2015). Other factors may interfere the process of Muller's ratchet and warrant further attention. For example, the occurrence of advantageous mutations can effectively compensate the fitness decay from Muller's ratchet (Bachtrog and Gordo 2004).

Practically, mutation rate can be increased by targeting the DNA repair system (Yaglom et al. 2014) or by regulating DNA replication stress (Burrell et al. 2013). The increase of mutation rate will accelerate the ratchet (Gordo and Charlesworth 2000a, 2000b; Lynch et al. 2016) which would in turn reduce tumor fitness and make tumors more sensitive to anticancer drugs (Pavel and Korolev 2017). For example, it has been suggested that high mutation rate may associate with a higher occurrence of neo-epitope formation and tumor immunogenicity. Thus, increase sensitivity to immune checkpoint inhibitors (McGranahan et al. 2016; Palmieri et al. 2017), such as anticytotoxic T-lymphocyte antigen-4 (Van Allen et al. 2015) and programmed cell death-1 (PD-1; Le et al. 2015) antibodies.

However, for these strategies to be effective, an additional layer of cancer biology needs to be considered. Muller's ratchet also explains the paradoxical relationship between chromosomal instability and prognosis in ovarian, gastric, and nonsmall cell lung cancer, with poorest outcome in tumors with intermediate, rather than extreme chromosomal instability (Birkbak et al. 2011). Cancer cells may benefit from generating genetic variation paving the way for tumor

progression. The tumor populations, thus, must strike a delicate balance to maintain heterogeneity and at the same time curb relentlessly increasing mutations which may be used to develop a new therapeutic regime for cancer. Therapies that directly target cell proliferation should be considered. In order to maintain high proliferation rates, rapidly proliferating cells need to increase their translational capacity and are dependent on high rates of ribosome biogenesis (Ruggero and Pandolfi 2003; Gentilella et al. 2015). Thus, inhibition of ribosome biogenesis could be a selective approach to cancer therapy (Bywater et al. 2012; Cortes et al. 2016). More importantly, this effect is enhanced in cells with higher proliferation rates, but less so in cells with lower proliferation rates.

## Materials and Methods

### Cell Culture and Single-cell Isolation
Cancer Cell lines, including HeLa cell lines, Human prostate cancer cell line PC3, Human Rhabdomyosarcoma cell line A204, and Human malignant melanoma cell line A375, were obtained from the Culture Collection of the Chinese Academy of Sciences, Shanghai, China. Cells were cultured in DMEM Medium for HeLa cell, PC3 and A375, in medium of McCoy's 5A for A204 (Life Technologies) and supplemented with 10% fetal bovine serum (FBS; Gibco, Life Technologies), 100 U/mL of penicillin, and 100 $\mu$g/mL of streptomycin at 37 °C with 5% $CO_2$. Cells were trypsinized using 0.05% trypsin at room temperature for 3 min. Cell counting was performed with a hemocytometer. The cell suspensions were diluted with medium to the final concentration of 1 cell per 100 $\mu$L. Single cells were seeded into each well of a 96-well cell culture plate and maintained at 37 °C with 5% $CO_2$. After incubating for 12 h, during which the cells attached to the wells, the single-cell isolation was visually confirmed by microscopy. The number of cells in each well was counted every day. When the cell colony became sufficiently large, cells were transferred to 24-well plates, and subsequently transferred to 6-well plates and 10-cm culture dishes.

### Karyotype Analysis
The cultured cells in a stage of active division were treated with colchicine (200 ng/mL for 1 h at 37 °C, then harvested and resuspended in 0.07 M KC1 for 30 min, and slowly added 10 drops of Carnoy's fixative (3:1, methanol:acetic acid). The suspended cell was dropped onto a slide, dried the slide rapidly, and stained with 4% DAPI for 5 min. At least twenty cells were spread out in metaphase for karyotyping. Evaluation of interphase nuclei was performed by OLYMPUS BX51 fluorescence microscope. Photographs were taken by CCD camera with 40 or 100 times objective. Image-Pro Plus software was used for digital image acquisition in the TIFF format, pseudocoloring and merging.

### MTT Assay
Cells were suspended and seeded at the concentration of ~700 cells/100 $\mu$l/well in 96 well plate. A volume of 20 $\mu$l-dissolved MTT was pipetted into each well. After being incubated for 4 h at 37 °C in a humidified $CO_2$ incubator, the media was removed and 200 $\mu$l sterile DMSO was added to each well. The absorbance values were then read at 570 nm with a microplate spectrophotometer. The proliferation of living cells was monitored based on absorbance values.

### Real-time Cell Analysis Assay (RTCA)
Cells in 10% FBS media were seeded at a density of 5,000 per well into a 16-well E-plate (ACEA) and incubated for 72 h at 37 °C in a humidified $CO_2$ incubator. The impedance for each well was monitored by the RTCA iCELLgence system (ACEA Biosciences Inc.) for every 70 min. Relative impedance signal level (represented as "cell index" in manufacturer's software) that indicated the number of cells was analyzed using the RTCA Software 1.2 program (Roche Diagnostics). The cell growth curves were automatically recorded on the iCELLigence system in real time. The cell index was followed for 3 days.

### Measurement of Growth Rate
Population growth rate was estimated using the equation:

$$\frac{dN(t)}{dt} = kN(t), \tag{2}$$

where $N(t)$ is the total number of cells at time $t$, and $k$ is a constant coefficient. In MTT assay and RTCA, $N(t)$ represents the absorbance value (OD) and "cell index," respectively. To obtain a linear function, the cell numbers were converted to logarithms to the base of Euler's number $e$. The least-squares method (LSM) was used to fit the linear regression in which the slope ($k$) of the regression line would be the growth rate.

### Whole Genome Sequencing and Analysis of CNV
Genomic DNA was extracted from $10^6$ HeLa cells in E6 and each of the E6 descendant clones (including G3, G2, B8, E7, and E3) using QiagenAllpre DNA/RNA Mini Kit (Qiagen). The genomic DNA (3 $\mu$g) was fragmented by Adaptive Focused Acoustics on a Covaris E120 (Covaris Inc.). The range of product size was from 250 bp to 350 bp. The fragmented DNA was purified by Qiaquick PCR purification column and quantified on 2100 Bioanalyzer by using the Agilent DNA 1000 kit (Agilent Technologies, Palo Alto, CA, USA). The whole genome libraries were constructed by IlluminaTruseq DNA sample preparation kit according to the manufacturer's instructions. Whole genome sequencing was performed using the Illumina Hiseq 2000 in the Beijing Institute of Genomics (BIG, Beijing, China). After quality control, the average read depth was 10×. Reads were mapped to the human reference genome (hg18) using BWA software (Version 0.4.9) with default parameters (Li and Durbin 2009).

The aligned reads were used as input of the Control-FREEC software (Control-FREEC v10.3; http://bioinfo-out.curie.fr/projects/freec/) for characterizing the large-scale chromosomal and segmental duplication and deletion events (>10 MB) in five descendant clones against E6. Control-FREEC was run with the following parameters: window size, 100 kb; step size, 50 kb; contaminationAdjustment = TRUE; noisyData = TRUE; BAF calculation inactivated. To determine significant CNA calls, Wilcoxon test and

Kolmogorov–Smirnov test (P-value < 0.01) were performed using the script assess significance. Since HeLa is triploid and some regions may contain 4–5 DNA copies (Adey et al. 2013), the ratio for changes from three to two copies is 0.67 and from three to four copies is 1.33. Similarly, the ratio for changes from four to three copies is 0.75 and from four to five copies is 1.25. Consequently, we set the relative ratios of >1.2 for copy number gain and <0.8 for copy number lose in the "MedianRatio" (fig. 2A, and supplementary table S2, Supplementary Material online).

## eGFP and DsRed Transfection

Cells with the stable expression of eGFP and DsRed were constructed by transfecting with lentivirus vectors, plenti6.3-MCS-IRES-eGFP and plenti6.3-MCS-IRES-DsRed (Life Technologies), which contained a blasticidin resistance gene and an eGFP or a *Discosoma* sp. red fluorescent protein (DsRed) sequence. The expression of eGFP and DsRed were used for labeling and distinguishing the two cells in the competition assay. The vectors were packaged into the lentivirus particles with infectious activity (Invitrogen). Before transfection, $2 \times 10^5$ cells per well were incubated with DMEM in a 6-well plate. After incubating for 24 h, the medium was replaced by the transfection medium that contained the lentivirus particles and polybrene with the concentration of 8 μg/mL. The multiplicity of infection (MOI) value was 3. After transfecting for 24 h, the cells were washed three times with PBS, and cultured in the DMEM medium with blasticidin of 10 μg/mL for at least 4 weeks in order to select cells that stably express eGFP and DsRed.

## Isolation of Fast-growing and Slow-growing Cells

The CellTrace Violet Cell Proliferation Kit, for flow cytometry (Thermo Fisher, C34557) was used for isolating the fast-growing and slow-growing cells. Cells were labeled by the cell-dye following the CellTrace Violet Cell Proliferation Kit workflow after cell cycle synchronization arrested at G1/S phase (double thymidine block). The labeled cells were continuously cultured for 7 days. At the 7th day after labeling, the cells were detached by trypsin–EDTA solution and suspended on culture medium. The BD Influx flow cytometer (BD) was used to isolate the fast-growing and slow-growing cells. The top 10% cells with strong fluorescence signals and bottom 10% cells with weak fluorescence signal were sorted as the slow growing and fast-growing cells respectively. Analysis was completed using the BD Influx flow cytometer with 405 nm excitation and a 450/40 nm bandpass emission filter.

## Competition Assay for Slow-growing Cell and Fast-growing Cells

Fast-growing cells with stable eGFP expression (or DsRed expression) and the slow-growing cells with DsRed expression (or eGFP expression) were mixed and cocultured at different initiation ratios (1:1, 2:8, and 8:2). The proportions of the two cell populations in the mixture were monitored by fluorescent intensity of DsRed and eGFP every 3 days by using flow cytometry.

## Modeling Population Dynamic of Cell Expansion

We construct and apply a simple model of cell growth and MA process. We define $N_t$ as population size at the day t, where $N_0 = 1$, and $R_t$ as the growth rate of population at day t. In this case, $R_t > 1$ represents population increase and $R_t < 1$ denotes population decrease. For each division, a proportion of cells ($\mu$) generates new mutations which decrease growth rate of the cells. The average deleterious effect of a mutation is s. $N_t$ is denoted by

$$N_t = R_{t-1} N_{t-1}. \tag{3}$$

$R_t$ and $R_{t-1}$ are given by a recursive function

$$R_t = (1-\mu)R_{t-1} + \mu R_{t-1}(1-s) = R_{t-1}(1-\mu s). \tag{4}$$

Iterating equation (4) t times, $R_t$ can be expressed as

$$R_t = R_{t-1}(1-\mu s) = \ldots = R_0(1-\mu s)^t, \tag{5}$$

where $R_0$ is the growth rate at day 0. Substitute equation (5) into equation (3) yields:

$$N_t = N_{t-1} R_0 (1-\mu s)^{t-1}$$
$$= N_0 \prod_{i=1}^{t} R_0 (1-\mu s)^{i-1}$$
$$= N_0 R_0^t (1-\mu s)^{\frac{t(t-1)}{2}} \tag{6}$$

Since the experiments start from single cell isolation ($N_0 = 1$), $N_t$ can be derived as

$$N_t = R_0^t (1-\mu s)^{t(t-1)/2}, \text{ for } t = 1, \ldots, 7. \tag{7}$$

## Estimate $R_0$ and Fitness Reduction $\mu s$

In equation (7) $N_t$ ($t = 1, \ldots, 7$) can be obtained from observed data, so $R_0$ and $\mu s$ can be expressed by $N_t$ ($t = 1, \ldots, 7$). We then divide $N_{t-1}$ by $N_t$ and divide $N_t$ by $N_{t+1}$:

$$\begin{cases} \dfrac{N_t}{N_{t-1}} = R_0(1-\mu s)^{t-1} \\ \dfrac{N_{t+1}}{N_t} = R_0(1-\mu s)^t \end{cases} . \tag{8}$$

Combine two formulas in equation (8) and eliminate $R_0$, the fitness reduction $\mu s$ can be calculated as

$$\mu s = 1 - \frac{N_{t+1} N_{t-1}}{N_t^2}. \tag{9}$$

Using linear regression model with the cell number from 8 days, the fitness reduction $\mu s$ can be approximated as

$$\mu s = 1 - \sum_{i=1}^{7} \frac{\frac{N_{t+1} N_{t-1}}{N_t^2}}{7}. \tag{10}$$

To calculate $R_0$, we substitute equation (9) into equation (8). Then $R_0$ can be expressed as

549

$$R_0 = \frac{N_{t+1}}{N_t(1 - \mu s)^t}. \qquad (11)$$

## Cell Growth Simulation

We extended the previously described genetic load to simulate the cell growth with accumulation of deleterious mutations. The simulation begins with a single cell. Each day, cells attempt to divide and produce additional cells. The initial growth rate, $R_0$ ($R_0 \in (1.0, 3.5]$), follows a Poisson distribution. For each cell replication at day $t$, the number of growth rate ($R_t$) which is randomly generated followed by mean growth rate $R_t$. $R_t > 1$ represents cell proliferation, whereas $R_t = 0$ denotes cell death. During replication, a proportion of cells ($\mu$) generates deleterious mutations. The mutations accumulated in the cells decreases the growth rate. The average deleterious effect of a mutation is $s$. Therefore, the growth rate of a cell population is reducing as a function of time. The computational model is qualitatively similar to our mathematical model, but considers much more complicated conditions, that is, the fluctuations of the parameters ($R_t$, $\mu$, and $s$) caused by their distributions.

When $R_0$ is fixed and $\mu$ and $s$ are given, we can monitor the cell number every day using the simulation model. In summary, we cultured $10^6$ cell lines in the simulation with different combinations of $R_0$, $\mu$, and $s$. Comparing the $10^6$ results with observed data with HeLa cell lines, we used the Approximate Bayesian Computation (ABC) method to select optimal groups. The mean number of the selected groups represent the estimated $R_0$, $\mu$, and $s$.

## Inference of $R_0$, $\mu$, and $d$ by ABC Method

Because of the complexity of the parameter space, we used ABC method by comparing the simulated $10^6$ cell lines with observed data with HeLa to estimate $R_0$, $\mu$, and $s$. ABC algorithms are a group of methods for performing Bayesian inference without the need for explicit evaluation of the model likelihood function. The algorithms can be used with implicit computer models that generate sample data sets rather than likelihoods (Beaumont et al. 2002; Wilkinson 2013). By using ABC we can compute the posterior probability distribution of a multivariate parameter $\Theta$ ($\Theta = \{R_0, \mu, s\}$). A parameter value $\Theta_i$ is sampled from its prior distribution $f(\Theta)$ to simulate a data set $S(\Theta_i)$, for $i = 1, \ldots, 10^6$. A set of summary statistics, the value that calculated from the data to represent the maximum amount of information in the simplest possible form, $S(\Theta_i)$ is computed from the simulation. By using a distance function $\rho$, we calculated the distance between $S(\Theta_i)$ and observed data $S_{obs}$. If the distance between $S(\Theta_i)$ and $S_{obs}$ is less than a given threshold, the parameter value $\Theta_i$ is accepted. In order to set the threshold that which simulations are accepted, we provide the tolerance rate $\epsilon$, which is defined as the percentage of accepted simulation (Csillery et al. 2010; Csilléry et al. 2015).

The ABC inference scheme is:

(1) Sample a candidate parameter $\Theta_i = \{R_0, \mu, s\}$ from the prior distribution $f(\Theta)$;

(2) Simulate the growth process of the cell line $\Theta_i$ and calculate the summary statistics $S_i$;

(3) Compare the simulated data set $S(\Theta_i)$, with the observed data $S_{obs}$, using a distance function $\rho$ and tolerance rate $\epsilon$, if $\rho(S(\Theta_i), S_{obs}) < \varepsilon$, accept $\Theta_i$;

(4) Go to 1.

Here are summary statistics $S(\Theta_i) = \{$simulated cell numbers of 7days$\}$. The observed summary statistics $S_{obs} = \{$observed 7days' cell numbers from experiment$\}$. The prior distribution $f(\Theta)$ in our model was $R_0 \sim Uniform[1.0, 3.5]$, $\mu \sim Uniform[0.05, 0.55]$ and $s \sim Uniform[0.01, 0.40]$. The distance function $\rho$ we set was the Euclidean distance and the tolerance rate $\varepsilon$ we set was 0.1%. As a whole 1,000 groups of $\Theta_i$ were accepted.

By using different clones' data from B8 as $S_{obs}$, we finally calculated the parametric ranges of rapid proliferated cells were $R_{0.fast} \in [2.22, 2.52]$, $\mu_{fast} \in [0.26, 0.30]$, and $s_{fast} \in [0.17, 0.20]$. Meanwhile, the parametric ranges of defected cells were $R_{0,slow} \in [1.67, 2.21]$, $\mu_{slow} \in [0.25, 0.33]$, and $s_{slow} \in [0.16, 0.20]$.

## Modeling Genetic Load in Long-term Cancer Growth

To assess how this high genetic load will affect the growth of HeLa cell lines in a long run, we modeled the dynamic of long-term cell growth in silicon. The simulation started from a single cell. The growth rate ($R_0 = 2.0$), mutation rate ($\mu = 0.3$ per genome per cell division) and mutational effect ($d = 0.18$) are from our estimation above. As the clone grew, we recorded the mutation number of each cell every day. Because, in our previous simulation, clone seldom went extinction when its size reached 5,000, the simulation was stopped when the clone reached 5,000 (survived) or went extinction (defected). The simulation was repeated 200 times. Next, we randomly chose one survived clone and let grow to $10^6$ cells. A total of 1,000 cells were randomly picked from this clone and their mutation profiles were recorded.

## Data Availability

The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Wang et al. 2017) in BIG Data Center (Members 2017), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under accession numbers PRJCA000406 that are publicly accessible at http://bigd.big.ac.cn/gsa.

## Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

## Acknowledgments

## References

Adey A, Burton JN, Kitzman JO, Hiatt JB, Lewis AP, Martin BK, Qiu R, Lee C, Shendure J. 2013. The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line. Nature 500(7461):207–211.

Agarwal B, Correa AM, Ho L. 2008. Survival in pancreatic carcinoma based on tumor size. Pancreas 36(1):e15–e20.

Bachtrog D, Gordo I. 2004. Adaptive evolution of asexual populations under Muller's ratchet. Evolution 58(7):1403–1413.

Beaumont MA, Zhang W, Balding DJ. 2002. Approximate Bayesian computation in population genetics. Genetics 162(4):2025–2035.

Birkbak NJ, Eklund AC, Li QY, McClelland SE, Endesfelder D, Tan P, Tan IB, Richardson AL, Szallasi Z, Swanton C. 2011. Paradoxical relationship between chromosomal instability and survival outcome in cancer. Cancer Res. 71(10):3447–3452.

Burrell RA, McClelland SE, Endesfelder D, Groth P, Weller MC, Shaikh N, Domingo E, Kanu N, Dewhurst SM, Gronroos E, et al. 2013. Replication stress links structural and numerical cancer chromosomal instability. Nature 494(7438):492–496.

Bywater MJ, Poortinga G, Sanij E, Hein N, Peck A, Cullinane C, Wall M, Cluse L, Drygin D, Anderes K, et al. 2012. Inhibition of RNA polymerase I as a therapeutic strategy to promote cancer-specific activation of p53. Cancer Cell. 22(1):51–65.

Charlesworth B. 2012. The effects of deleterious mutations on evolution at linked sites. Genetics 190(1):5–22.

Charlesworth B. 1978. Model for evolution of Y chromosomes and dosage compensation. Proc Natl Acad Sci U S A. 75(11): 5618–5622.

Charlesworth B. 1990. Mutation-selection balance and the evolutionary advantage of sex and recombination. Genet Res. 55(3):199–221.

Chen B, Shi Z, Chen Q, Shibata D, Wen H, Wu C-I. 2018. Quasi-neutral molecular evolution—when positive and negative selection cancel out. bioRxiv 330811. doi:https://doi.org/10.1101/330811.

Cortes CL, Veiga SR, Almacellas E, Hernandez-Losa J, Ferreres JC, Kozma SC, Ambrosio S, Thomas G, Tauler A. 2016. Effect of low doses of actinomycin D on neuroblastoma cell lines. Mol Cancer. 15:1.

Crow J. 1958. Some possibilities for measuring selection intensities in man. Hum Biol. 30(1):1–13.

Csilléry K, Lemaire L, François O, Blum M. 2012. Approximate Bayesian Computation (ABC) in R: a vignette. http://cran.r-project.org/web/packages/abc/vignettes/abcvignette.pdf.

Csillery K, Blum MG, Gaggiotti OE, Francois O. 2010. Approximate Bayesian Computation (ABC) in practice. Trends Ecol Evol. 25(7):410–418.

Dickinson WJ. 2008. Synergistic fitness interactions and a high frequency of beneficial changes among mutations accumulated under relaxed selection in Saccharomyces cerevisiae. Genetics 178(3):1571–1578.

Felsenstein J. 1974. The evolutionary advantage of recombination. Genetics 78(2):737–756.

Fischer A, Greenman C, Mustonen V. 2011. Germline fitness-based scoring of cancer mutations. Genetics 188(2):383–393.

Gabriel W, Lynch M, Burger R. 1993. Muller's ratchet and mutational meltdowns. Evolution 47(6):1744–1757.

Gentilella A, Kozma SC, Thomas G. 2015. A liaison between mTOR signaling, ribosome biogenesis and cancer. Biochim Biophys Acta. 1849(7):812–820.

Gordo I, Charlesworth B. 2000a. The degeneration of asexual haploid populations and the speed of Muller's ratchet. Genetics 154(3):1379–1387.

Gordo I, Charlesworth B. 2000b. On the speed of Muller's ratchet. Genetics 156:2137–2140.

Haigh J. 1978. The accumulation of deleterious genes in a population–Muller's Ratchet. Theor Popul Biol. 14(2):251–267.

Hartl DL, Clark AG. 2007. Principles of population genetics. Sunderland (MA): Sinauer Associates.

Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. Genet Res. 8(3):269–294.

Katju V, Packard LB, Bu L, Keightley PD, Bergthorsson U. 2015. Fitness decline in spontaneous mutation accumulation lines of Caenorhabditis elegans with varying effective population sizes. Evolution 69(1):104–116.

Kibota TT, Lynch M. 1996. Estimate of the genomic mutation rate deleterious to overall fitness in E. coli. Nature 381(6584):694–696.

Kondrashov AS. 1982. Selection against harmful mutations in large sexual and asexual populations. Genetical Res. 40(3):325–332.

Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, et al. 2015. PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med. 372(26):2509–2520.

Li C, Hou Y, Xu J, Zhang A, Liu Z, Qi F, Yang Z, Chen K, Liu S, Huang H, et al. 2017. A direct test of selection in cell populations using the diversity in gene expression within tumors. Mol Biol Evol. 34(7):1730–1742.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25(14):1754–1760.

Li W-H. 1997. Molecular evolution. Sunderland (MA): Sinauer Associates.

Ling S, Hu Z, Yang Z, Yang F, Li Y, Lin P, Chen K, Dong L, Cao L, Tao Y, et al. 2015. Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution. Proc Natl Acad Sci U S A. 112(47):E6496–E6505.

Lovell JT, Williamson RJ, Wright SI, McKay JK, Sharbel TF. 2017. Mutation accumulation in an asexual relative of Arabidopsis. PLoS Genet. 13(1):e1006550.

Lynch M. 2016. Mutation and human exceptionalism: our future genetic load. Genetics 202(3):869–875.

Lynch M, Ackerman MS, Gout JF, Long H, Sung W, Thomas WK, Foster PL. 2016. Genetic drift, selection and the evolution of the mutation rate. Nat Rev Genet. 17(11):704–714.

Lynch M, Burger R, Butcher D, Gabriel W. 1993. The mutational meltdown in asexual populations. J Hered. 84(5):339–344.

Lynch M, Gabriel W. 1990. Mutation load and the survival of small populations. Evolution 44(7):1725–1737.

Maynard Smith J. 1978. The evolution of sex. Cambridge [Eng.], New York: Cambridge University Press.

McFarland CD, Korolev KS, Kryukov GV, Sunyaev SR, Mirny LA. 2013. Impact of deleterious passenger mutations on cancer progression. Proc Natl Acad Sci U S A. 110(8):2910–2915.

McGranahan N, Furness AJ, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, Jamal-Hanjani M, Wilson GA, Birkbak NJ, Hiley CT, et al. 2016. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science 351(6280):1463–1469.

Members BDC. 2017. The BIG Data Center: from deposition to integration to translation. Nucleic Acids Res. 45:D18–D24.

Molaro A, Malik HS. 2017. Sex chromosome evolution culture shock. Elife 6:e33312.

Muller HJ. 1964. The relation of recombination to mutational advance. Mutat Res. 106:2–9.

Narod SA. 2012. Tumour size predicts long-term survival among women with lymph node-positive breast cancer. Curr Oncol. 19(5):249–253.

Palmieri G, Colombino M, Cossu A, Marchetti A, Botti G, Ascierto PA. 2017. Genetic instability and increased mutational load: which diagnostic tool best direct patients with cancer to immunotherapy? J Transl Med. 15(1):17.

Pavel AB, Korolev KS. 2017. Genetic load makes cancer cells more sensitive to common drugs: evidence from Cancer Cell Line Encyclopedia. Sci Rep. 7(1):1938.

Ruggero D, Pandolfi PP. 2003. Does the ribosome translate cancer? *Nat Rev Cancer.* 3(3):179–192.

Shaw RG, Byers DL, Darmo E. 2000. Spontaneous mutational effects on reproductive traits of *Arabidopsis thaliana. Genetics* 155(1):369–378.

Tao Y, Hu Z, Ling S, Yeh S-H, Zhai W, Chen K, Li C, Wang Y, Wang K, Wang H-Y, et al. 2015. Further genetic diversification in multiple tumors and an evolutionary perspective on therapeutics. *bioRxiv* 025429. doi:https://doi.org/10.1101/025429.

Thompson RH, Kurta JM, Kaag M, Tickoo SK, Kundu S, Katz D, Nogueira L, Reuter VE, Russo P. 2009. Tumor size is associated with malignant potential in renal cell carcinoma cases. *J Urol.* 181(5):2033–2036.

Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, Sucker A, Hillen U, Foppen MHG, Goldinger SM, et al. 2015. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 350(6257):207–211.

Wang Y, Song F, Zhu J, Zhang S, Yang Y, Chen T, Tang B, Dong L, Ding N, Zhang Q, et al. 2017. GSA: genome Sequence Archive*. *Genomics Proteomics Bioinformatics.* 15(1):14–18.

Wilkinson RD. 2013. Approximate Bayesian computation (ABC) gives exact results under the assumption of model error. *Stat Appl Genet Mol Biol.* 12(2):129–141.

Willi Y. 2013. Mutational meltdown in selfing *Arabidopsis lyrata. Evolution* 67(3):806–815.

Woodruff RC. 2013. An extreme test of mutational meltdown shows mutational firm up instead. *Genetica* 141(4–6):185–188.

Wu CI, Wang HY, Ling S, Lu X. 2016. The ecology and evolution of cancer: the ultra-microevolutionary process. *Annu Rev Genet.* 50:347–369.

Xu J, Peng XX, Chen YX, Zhang YZ, Ma Q, Liang L, Carter AC, Lu XM, Wu CI. 2017. Free-living human cells reconfigure their chromosomes in the evolution back to uni-cellularit. *Elife* 6:e28070.

Yaglom JA, McFarland C, Mirny L, Sherman MY. 2014. Oncogene-triggered suppression of DNA repair leads to DNA instability in cancer. *Oncotarget* 5(18):8367–8378.

Zhang J, Gold KA, Lin HY, Swisher SG, Xing Y, Lee JJ, Kim ES, William WN. 2015. Relationship between tumor size and survival in non-small-cell lung cancer (NSCLC): an analysis of the surveillance, epidemiology, and end results (SEER) registry. *J Thorac Oncol.* 10(4):682–690.