

Simulation and modeling of chromothripsis reveals new insights into ecDNA origination

Junxiao Yang¹, and Xing Liu^{1,2} ✉

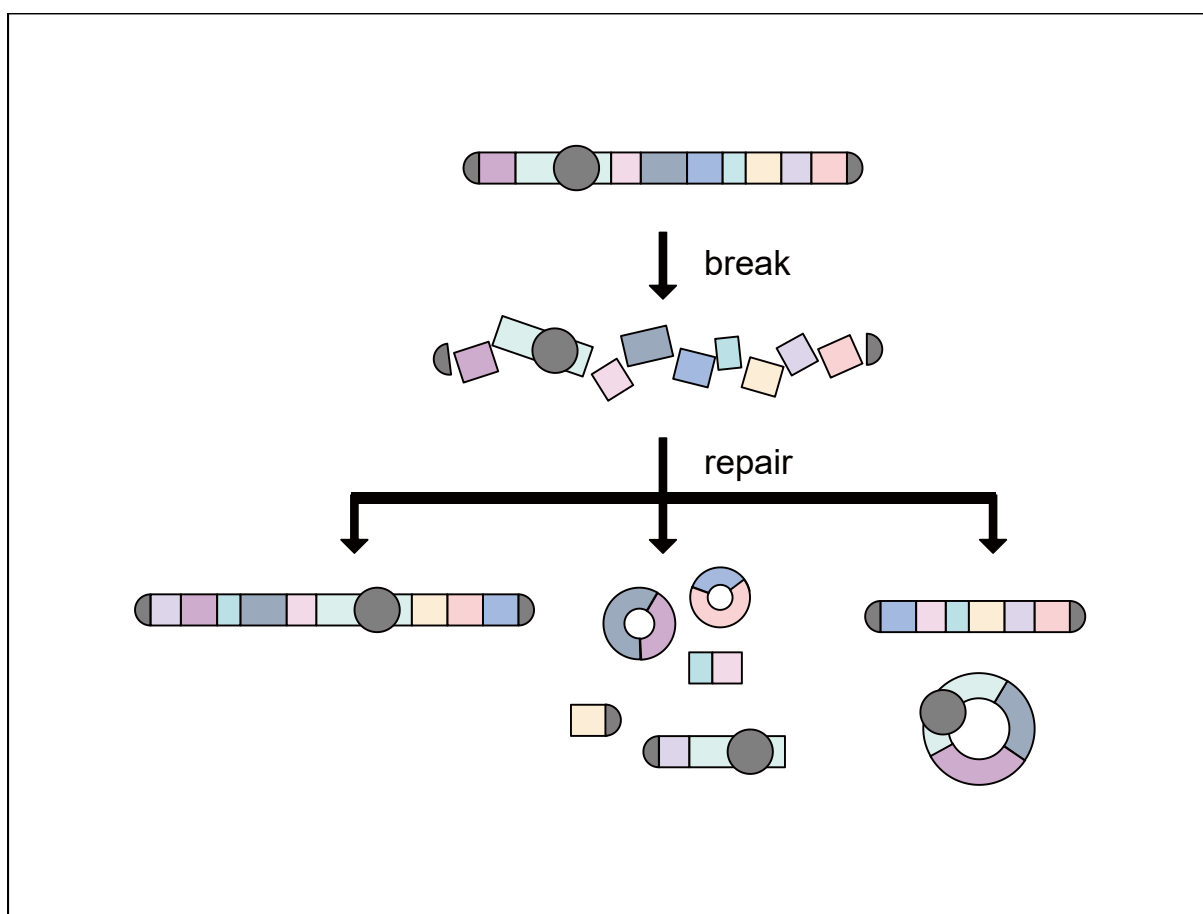
¹MOE Key Laboratory for Cellular Dynamics, Hefei National Research Center for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei 230026, China;

²Anhui Key Laboratory for Chemical Biology and New Quality Medicine, University of Science and Technology of China, Hefei 230027, China

✉Correspondence: Xing Liu, E-mail: xing1017@ustc.edu.cn

© 2026 The Author(s). This is an open access article under the CC BY-NC-ND 4.0 license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Graphical abstract



Modeling the process of chromothripsis.

Public summary

- We developed simulation programs with statistical algorithms to analyze batch-processed results, revealing how three key parameters (the compaction ratio, break frequency, and repair rate) modulate the outcomes of chromothripsis.
- Our modeling revealed that chromothripsis generates remarkably limited quantities of circular DNA (typically ≤ 4 molecules per event), with linear DNA consistently outnumbering circular DNA across most biological conditions.

Simulation and modeling of chromothripsis reveals new insights into ecDNA origination

Junxiao Yang¹, and Xing Liu^{1,2} 

¹MOE Key Laboratory for Cellular Dynamics, Hefei National Research Center for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei 230026, China;

²Anhui Key Laboratory for Chemical Biology and New Quality Medicine, University of Science and Technology of China, Hefei 230027, China

 Correspondence: Xing Liu, E-mail: xing1017@ustc.edu.cn

© 2026 The Author(s). This is an open access article under the CC BY-NC-ND 4.0 license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



Cite This: *JUSTC*, 2026, 56(3): 0303 (8pp)



Read Online

Abstract: Chromothripsis is a catastrophic chromosomal rearrangement that drives genomic instability and cancer evolution. Although extensive research has focused on biological mechanisms, the development of computational tools for modeling this process remains lacking. Here, we developed two novel programs: a simulation program that models the process of break and repair and outputs the rearrangement results and a computational program that quantitatively calculates the relative abundance of linear and circular DNA products. We employ these programs to analyze how the compaction ratio, break frequency, and repair rate influence the properties of both circular and linear DNA. Furthermore, we demonstrate that individual chromothripsis events produce exceptionally few circular DNA molecules (typically ≤ 4) under most conditions. Our modeling offers new insights into extrachromosomal DNA (ecDNA) biogenesis and a theoretical model for experimental validation.

Keywords: mitosis; chromothripsis; ecDNA; simulation program

CLC number: Q343.2

Document code: A

1 Introduction

Abnormal chromosomes can undergo various structural changes, such as deletions, duplications, inversions, and translocations. However, in recent decades, scientists have discovered that chromosomes can undergo complex structural changes, such as chromothripsis, chromoanasythesis, and chromoplexy^[1,2]. Among these, chromothripsis is the most common, occurring in approximately 50% of cancers. In contrast, chromoanasythesis and chromoplexy are far rarer, appearing in only 5% to 10% of cancer cases^[3,4]. Chromothripsis is a catastrophic chromosome rearrangement event in which a single chromosome (or several chromosomes) shatters into fragments and then reassembles into a complex new structure^[5]. Chromothripsis can be divided into four stages:

- **Formation of micronuclei.** The formation of micronuclei is principally driven by three mechanisms: DNA breaks, chromosome lagging during anaphase, and defective nuclear envelope assembly of chromosomes during telophase^[6]. When DNA breaks occur, a chromosome is divided into two distinct fragments—one containing a centromere and one without. The fragment with the centromere either undergoes normal segregation or forms a chromosome bridge, whereas the acentric fragment remains in the cytoplasm, where it subsequently acquires a nuclear envelope, forming a micronucleus. Chromosome lagging typically stems from kinetochore dysfunction, aberrant kinetochore–microtubule attachments, or defects in mitotic spindle assembly.

- **Rupture of micronuclear envelope.** Micronuclear membranes exhibit high instability and frequent rupture because of their small radius of curvature and defective nuclear lamina assembly^[7,8] (evidenced by studies showing that lamina-deficient micronuclei have increased rupture susceptibility). Upon rupture, micronuclei recruit the ESCRT-III complex, which unexpectedly exacerbates membrane damage rather than facilitating repair. This leads to a progressive decrease in the number of intact micronuclei and a concomitant accumulation of ruptured micronuclei as cancer cells transition from mitosis to interphase^[9]. Nevertheless, cellular quality control mechanisms, including but not limited to the cGAS-STING pathway, can involve the detection of some of these ruptured micronuclei and the initiation of their selective autophagic degradation^[10].

- **Fragmentation of DNA.** Micronuclear rupture has two major consequences: First, cytoplasmic nucleases gain access to the micronuclear compartment and cleave the exposed DNA; second, the collapse of nuclear–cytoplasmic compartmentalization leads to rapid depletion of nucleus-localized proteins, particularly those essential for DNA replication. The loss of these proteins induces severe replication defects, including premature replication fork stalling and collapse, ultimately leading to widespread chromosomal fragmentation during prophase chromatin condensation in subsequent mitosis^[11].

- **Repair of DNA breaks.** Despite fragmentation, ruptured micronuclear chromosomes remain clustered during mitosis through the action of MDC1, TOPBP1, and CIP2A^[12–14].

These damaged chromosomes face three distinct fates in the subsequent cell cycle: approximately 40% successfully reintegrate into the primary nucleus, where they undergo non-homologous end joining (NHEJ)^[15], and another 40% are reencapsulated into new micronuclei, whereas the remaining 20%, although initially remaining clustered, eventually fragment into smaller segments because of mitotic spindle-generated shear forces, which are ultimately distributed between the primary nucleus and cytoplasmic compartments^[16].

Chromothripsis can lead to several significant consequences. First, it can disrupt functional genes. When tumor suppressor genes are impaired, this promotes cancer development. Second, it can induce gene copy number variations, potentially causing either gene amplification or deletion. Third, it produces extrachromosomal DNA (ecDNA), small circular DNA molecules that accumulate in cancer cells and accelerate cancer evolution^[17,18]. To determine the harmful consequences of chromothripsis, we developed a Python-based chromothripsis simulation program. This program revealed many intriguing properties of circular DNA and linear DNA generated by chromothripsis, including the extremely limited quantity of ecDNA. To further investigate this phenomenon, we developed a mathematical model predicting the yields of circular and linear DNA for varying numbers of DNA breaks. All the programs can be obtained from https://github.com/nesjedi/chromothripsis_simulation_programs.

2 Materials and methods

2.1 Simulation of the chromothripsis process

We developed a Python program in which we simplified the process of chromothripsis into two phases: a breaking phase and a repairing phase. The breaking phase simulates random chromosomal breakage, as occurs in micronuclei, while centromeres and telomeres are intentionally excluded because of their cleavage-resistant nature. In this phase, every fragment is given a unique identifier, and their adjacent fragments are recorded to make the following processes much easier. The repairing phase simulates broken end connections on the basis of spatial positioning relationships.

Given that a broken end is typically spatially closer to its paired broken end on the same fragment than to other broken ends, we developed a mathematical model to quantify the distance distribution between paired broken ends within individual fragments, from which we derive length-dependent probability distributions for fragment circularization.

Assuming that DNA is a freely jointed chain with isotropic bending flexibility, the problem simplifies to a DNA fragment of length D μm divided into n child fragments (where n is sufficiently large), each with length $\delta = \frac{D}{n}$ and random angles between adjacent segments. By fixing one end (A) of the fragment as the spatial origin, we track the relative position of the other end (B). The position of B is determined by n successive steps of length δ from A, with each step's orientation randomly distributed in 3D space. This model is mathematically equivalent to a 3D random walk with a fixed step length δ ; therefore, the AB distance r approaches a Maxwell

distribution as $n \rightarrow \infty$:

$$f(r) = \sqrt{\frac{2}{\pi}} \frac{r^2}{\sigma^3} e^{-\frac{r^2}{2\sigma^2}}, \quad \sigma^2 = \frac{n\delta^2}{3}.$$

When $n > 100$ and $D < 5$, the truncation error $1 - f(x) < 10^{-60}$; therefore, the effects of finite r domain restrictions are negligible. $f(r)$ satisfies the following property:

$$\lim_{r \rightarrow 0^+} \frac{df}{dr} = 6 \sqrt{\frac{6}{\pi}} \frac{r}{D^{\frac{3}{2}} \delta^{\frac{3}{2}}}.$$

Assuming cyclization can only occur when the AB distance r is less than d ($d \ll D$), the cyclization probability $P(D) = 6 \sqrt{\frac{6}{\pi}} \frac{d}{D^{\frac{3}{2}} \delta^{\frac{3}{2}}}$. This equation captures the inverse relationship between D and P , where increasing fragment length D leads to decreased cyclization probability P . For $D > 5$ μm (exceeding nuclear scale), P becomes independent of length due to nuclear confinement, stabilizing at a constant value. For computational tractability, we normalize the probability distribution by setting $P(D = 5) = 1$ as the reference standard and obtain the normalized distribution:

$$P(D) = \begin{cases} 11.18034 \times D^{-\frac{3}{2}}, & 0 < D < 5; \\ 1, & D \geq 5. \end{cases}$$

Since the connection probability between ends on two separate fragments is not significantly different from that between ends of a single sufficiently long fragment, we normalize the relative interfragment connection probability to 1. This probability distribution enables quantitative analysis of how spatial positioning affects DNA repair. Both the breaking phase and the repairing phase are fully customizable through adjustable parameters: the break frequency controls the fragmentation intensity in the breaking phase, whereas the repair rate and compaction ratio determine the rejoining probability in the repairing phase.

2.2 Analysis of chromothripsis rearrangement products

To better understand the rearrangement products after chromothripsis, we must obtain key information, including whether circular DNA is present, the length of each rearrangement product, the number of fragments within each rearrangement product, and which fragments contain a centromere or a telomere, among other details. We developed a search algorithm to achieve this goal. In our search algorithm, the traversal status of every fragment is recorded so that the program can visit them exhaustively. The algorithm traverses every rearrangement product that is newly generated after simulated chromothripsis to calculate its length and the number of fragments it contains. Afterward, the results are visualized using separate line plots for linear and circular DNA data. To identify the newly generated DNA that contains centromeres or telomeres, the algorithm searches every fragment adjacent to the fragment that contains a centromere or telomere and ultimately outputs information about rearrangement products, including their topology (circular or linear), length, and number of constituent fragments.

On the basis of these algorithms, we developed a batch-

processing algorithm that runs for a set number of iterations while outputting the length and quantity information of both circular DNA and linear DNA from all cycles. Using these data, we analyzed and visualized the effects of the compaction ratio, break frequency, and repair rate on the properties of circular DNA and linear DNA and statistically analyzed the number of linear DNA fragments exceeding 50 Mb. All the simulations were parameterized with the most prevalent conditions in human cells.

2.3 Development of a mathematical model to predict circular DNA and linear DNA yields

To quantify circular DNA production from a single chromothripsis event, we developed a mathematical model. In this model, the complex factors affecting chromothripsis are reduced to two key parameters: the number of cuts (N_f) and the number of repairs (N_r). A critical constraint arises from telomere-containing fragments, which cannot undergo ligation with other DNA ends under normal conditions, significantly increasing model complexity. To address this, the model assumes that two telomeres can transiently ligate to each other, which will turn a linear chromosome into a circular form. Following the simulation, the temporary telomere ligation is reversed. This approach provides rapid simplification of the calculation.

This model initiates by simulating the break and repair process on circular DNA, producing two key metrics: the mean number of newly generated circular DNA (C) and the mean number of fragments contained in newly generated circular DNA (F). The model begins with N_f cuts on circular DNA, which will produce N_f fragments with two open ends. These fragments then undergo N_r iterative repair steps. C_i and F_i are recorded at each iteration. Considering that DNA can exist only in two forms—linear or circular—and that only linear DNA can participate in repair, the only way to generate a circular DNA is through the ligation of both ends of the same linear DNA (it will generate only a new linear DNA if two ends of two different linear DNAs are ligated together). When the i -th repair operation is being performed, each end has a $\frac{1}{2(N_f - i + 1) - 1}$ probability of being ligated to the other end on the same linear DNA. Therefore, it is easy to obtain C_i by the following formula:

$$C_i = \frac{1}{2(N_f - i + 1) - 1} \times \frac{2(N_f - i + 1)}{2(N_f - i + 1)} = \frac{1}{2N_f - 2i + 1}.$$

P_i can be directly calculated by multiplying F_i by the mean number of fragments carried on each linear DNA (E). Since E can be calculated by the formula

$$E = \frac{\text{the total number of fragments included in linear DNA}}{\text{the number of linear DNA}},$$

the total number of fragments included in linear DNA and the number of linear DNA can be calculated by the following formula:

$$\begin{aligned} \text{the total number of fragments included in linear DNA} = \\ N_f - F_{i-1}, \end{aligned}$$

$$\text{the number of linear DNA} = N_f - i + 1.$$

It is easy to obtain that:

$$\begin{aligned} F_i = F_{i-1} + C_i \times E = F_{i-1} + C_i \times \left(\frac{N_f - F_{i-1}}{N_f - i + 1} \right) = \\ F_{i-1} + \frac{(N_f - F_{i-1})}{(2N_f - 2i + 1)(N_f - i + 1)}, F_0 = 0. \end{aligned}$$

C_{total} can be derived by the formula: $C_{\text{total}} = \sum_{i=1}^{N_r} C_i$.

After the above steps are complete, we need to restore the assumed circular DNA to the original chromosome by breaking the ligated telomere. Since every fragment has the same possibility of being generated by the ligation of two telomere-containing fragments, every fragment is equally likely to be broken during the restoration process. If this fragment is contained in linear DNA, it will generate two new linear DNAs but will not affect the number of circular DNAs; if this fragment is contained in circular DNA, it will turn a circular DNA into linear DNA and reduce the number of circular DNAs by one. Therefore, the final mean number of circular DNAs produced in the process of chromothripsis (K) can be expressed as follows:

$$K = C_{\text{total}} - \frac{F_{N_r}}{N_f}.$$

The mathematical model was implemented in Python. This program calculates the theoretical mean number of circular DNA (using the formula above) and the theoretical mean number of linear DNA (equivalent to $N_f - N_r + 1$) and plots both values on a line graph with the number of cuts represented on the x -axis for comparative analysis. Since there are only two parameters—the number of cuts and the number of repairs—that need to be considered in this program and the number of cuts is set to be an x -axis variable while the number of repairs can be calculated from the degree of repair, users only need to input the maximum cut number (defining observation range) and the degree of repair.

3 Results and discussion

3.1 Developing a chromothripsis simulation program

To better understand the mechanism of chromothripsis, we developed a simulation program to model the process of chromothripsis. This program begins by simulating stochastic chromosomal fragmentation by *TREX1* exonuclease and premature chromatin condensation within micronuclei and then applying the Maxwell distribution to quantify the impact of spatial proximity between two ends of an individual fragment on repair probability, which enables more accurate simulation of DNA fragment repair within the nuclear space. Notably, the program assumes that repairs occur sequentially. This is because synaptic complexes colocalize broken DNA ends prior to the ligation process. Consequently, we can disregard the duration of the repair process and focus only on the formation of synaptic complexes. These synaptic complexes form too rapidly for multiple simultaneous occurrences to be possible.

Since chromothripsis naturally occurs not only on single

chromosomes but also frequently involves two or more chromosomes, we developed two separate simulation programs: one modeling single-chromosome chromothripsis events and another specifically designed for dual-chromosome cases. Schematic diagrams illustrating single-chromosome and dual-chromosome chromothripsis are presented in Fig. 1. Further validation of these simulations within biological systems, such as tissue culture and 3D organoids, would provide additional insights into the dynamics of chromothripsis and ecDNA formation.

In this model, the position of centromeres and telomeres after chromothripsis occurs should receive extra attention, as it significantly influences mitotic chromosome segregation. In single-chromosome events, the centromere can localize in one of three ways: on a fully reassembled linear chromosome with two telomeres, on linear DNA with one or two broken ends, or on circular DNA. This scenario becomes more complex in dual-chromosome chromothripsis, where centromeres originating from distinct chromosomes may become colocalized in either linear or circular DNA structures. Such abnormal chromosomes may lead to erroneous chromosome segregation during mitosis, contributing to chromosomal instability. Telomere distribution demonstrates two distinct patterns: DNA fragments containing two telomeres form stable linear molecules that resist further ligation events, whereas those with only one telomere are prone to breakage-fusion-bridge (BFB) cycles during mitosis because of their unprotected DNA ends. To systematically analyze these features, we developed a search algorithm that characterizes the topological state (circular or linear), molecular length, and quantity of the

constituent fragments. Additionally, our program supports queries of any chromosomal region of interest, allowing researchers to use gene position information to obtain detailed information about the rearranged fragment with a specific oncogene.

To map all the rearrangement products, the program classifies the DNA fragments into two topological categories: circular and linear. For each category, the program generates the distribution of DNA length and the distribution of constituent fragments per DNA fragment. The distribution of DNA length allows users to retrieve the quantity of DNA of each length after a chromothripsis event, and the distribution of fragments per DNA fragment can help users determine how many original fragments are included in each newly assembled DNA fragment. An example of this analysis is provided in Fig. 2.

3.2 The compaction ratio, break frequency and repair rate modulate circular DNA formation during chromothripsis

DNA compaction influences circularization probability through length modulation. To establish the quantitative relationship between the length of naked DNA and its physical length, we introduce a compaction ratio parameter. As shown in Fig. 3a–c, increasing the DNA compaction ratio leads to three coordinated effects: (i) elevated yields of circular DNA, (ii) a reduced number of constituent fragments per circular DNA, and (iii) a decreased length of circular DNA. This finding indicates that a higher compaction ratio increases the circularization probability of linear DNA with

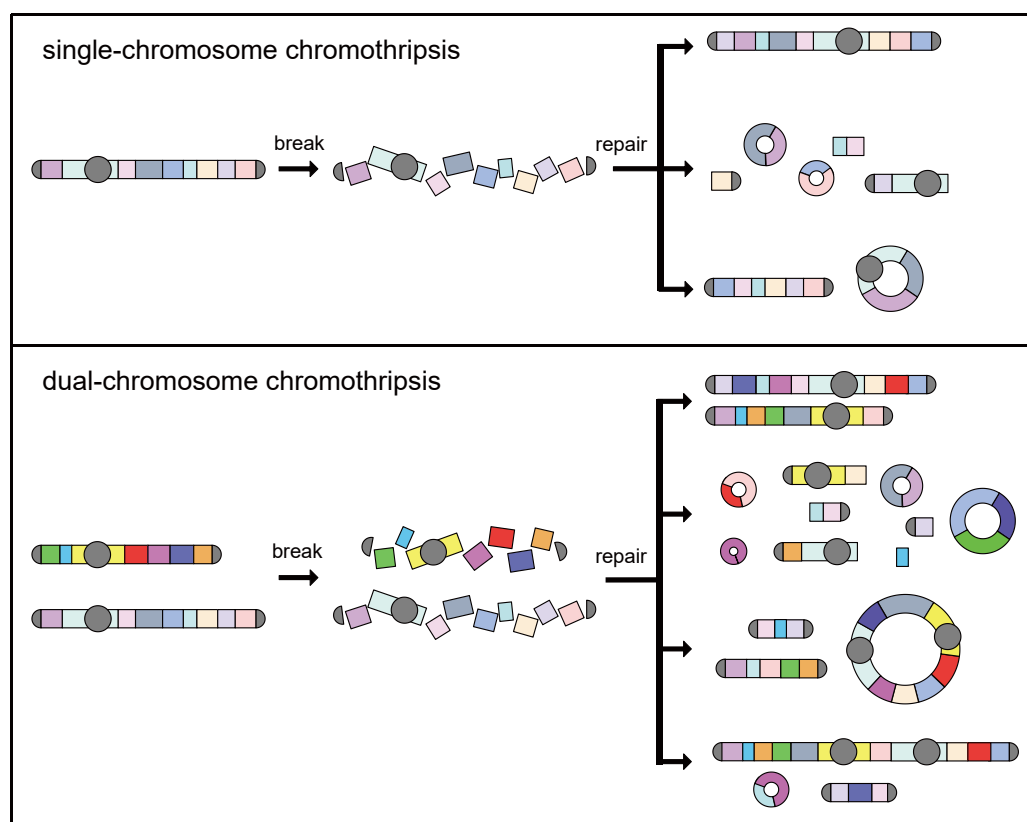


Fig. 1. Processes and outcomes of single-chromosome and dual-chromosome chromothripsis.

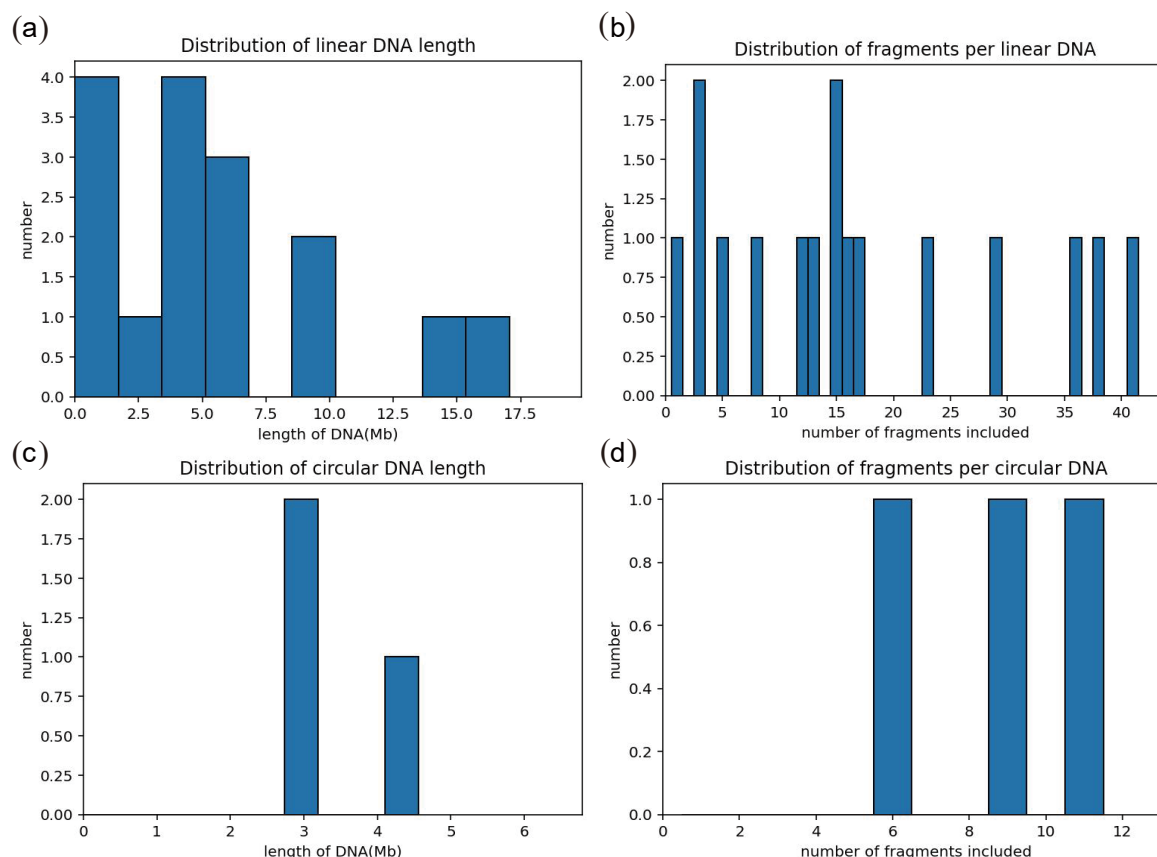


Fig. 2. Simulation results of simulated chromothripsis. Parameter input: chromosome length, 100.0 (Mb); centromere position, 30.0 (Mb); break frequency, 3.0 (breaks/Mb); repair rate, 95.0%; compaction ratio, 1.

fewer fragments. Reduced physical length can increase the probability of self-ligation of a single fragment's two ends. This phenomenon decreases the length of circular DNA while increasing its number. Considering the 7 : 1 compaction ratio of 5 nm nucleosomes and the 40 : 1 ratio of 30 nm solenoid fibers, chromatin under physiological conditions (with fiber diameters of 5–24 nm^[9]) has an intermediate compaction ratio ranging from 7 : 1 to 40 : 1. As shown in Fig. 3a, the mean number of circular DNA molecules remains below four within this physiological compaction range.

Break frequency can also affect circular DNA properties. While elevated break frequencies increase circular DNA yields (Fig. 3d), they concurrently reduce the number of constituent fragments per circular DNA (Fig. 3e). This phenomenon indicates that the reduced average length of fragments resulting from breakage increases the likelihood of circularization for linear DNA with fewer fragments than for those containing more fragments, which is consistent with the -1.5 scaling exponent in the fragment length–probability relationship. The combined effect of shorter fragment lengths and fewer constituent fragments led to a decrease in the circular DNA length at elevated break frequencies (Fig. 3f). Interestingly, the data in Fig. 3e and f reveal that while the maximum number of constituent fragments per circular DNA increases with increasing break frequency, the maximum length of these fragments has an inverse relationship. Notably, at the most prevalent break frequency of 1 break/Mb, the mean number of circular DNAs remained below four molecules

(Fig. 3d).

Furthermore, we investigated the impact of the repair rate on circular DNA properties. An increase in the repair rate nonlinearly increased the number of circular DNA, the number of fragments per circular DNA, and the length of the circular DNA, with effects intensifying progressively as the repair rate approached 100%, as shown in Fig. 3g–i. This phenomenon occurs because the circularization probability increases as the number of unfixed fragments decreases, demonstrating the critical role of final repair cycles in circular DNA formation. Notably, even at a 100% repair rate, the number of circular DNAs remains constrained to approximately four molecules.

3.3 The properties and quantity of linear DNA are influenced by the repair rate and break frequency

Following our circular DNA analyses, equivalent computational simulations were applied to linear DNA. As shown in Fig. 4a, b, increasing repair rates promote both the elongation of linear DNA and an increase in the number of constituent fragments per linear DNA. Compared with a 99% repair rate, a 100% repair rate significantly increased the length of linear DNA, demonstrating that the final few repair events play a crucial role not only for circular DNA but also for linear DNA. This phenomenon is further confirmed in Fig. 4c, where a significantly greater number of fragments exceeding 50 Mb are observed at a 99% repair rate than at a 90% repair rate. Surprisingly, a 99% repair rate generates more linear

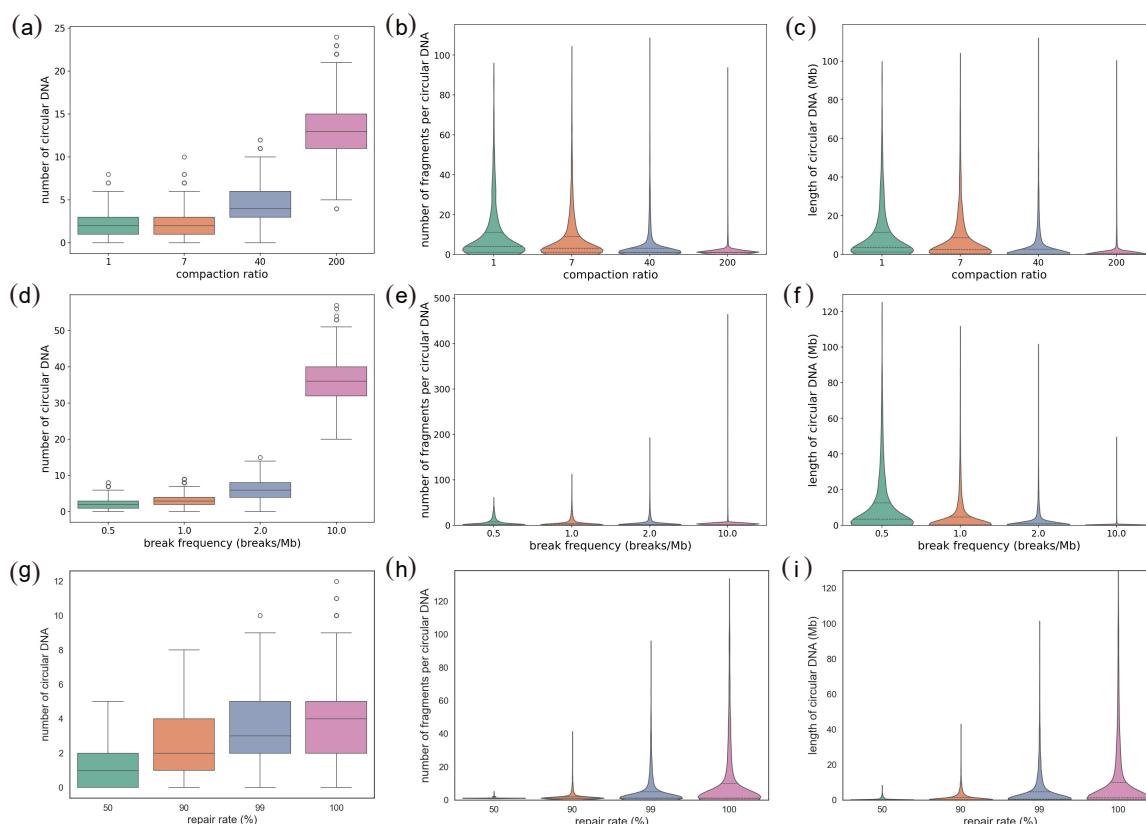


Fig. 3. The properties of circular DNA are affected by the compaction ratio, break frequency, and repair rate. (a–c) Illustration of how the compaction ratio influences the number of circular DNA molecules generated per chromothripsis event (a), the number of fragments per circular DNA (b), and the length of circular DNA (c), respectively. Sample size: 1000; parameters: chromosome length, 130.0 (Mb); break frequency, 1.0 (breaks/Mb); repair rate, 99.0 (%). (d–f) Illustration of how the break frequency influences the number of circular DNA molecules generated per chromothripsis event (d), the number of fragments per circular DNA (e), and the length of circular DNA (f), respectively. Sample size: 1000; parameters: chromosome length, 130.0 (Mb); compaction ratio, 25; repair rate, 99.0 (%). (g–i) Illustration of how the repair rate influences the number of circular DNAs generated per chromothripsis event (g), the number of fragments per circular DNA (h), and the length of circular DNA (i), respectively. Sample size: 1000; parameters: chromosome length, 130.0 (Mb); compaction ratio, 25; break frequency, 1.0 (breaks/Mb).

DNA fragments exceeding 50 Mb than a 100% repair rate does. This occurs because the 99% repair rate leaves some breakpoints unrepaired, allowing the persistence of multiple linear DNA fragments exceeding 50 Mb in cells. This implies that even 1% incomplete repair may serve significant biological functions.

Unlike repair effects, increasing the frequency of breaks increases the number of constituent fragments per linear DNA while reducing their average lengths and the number of linear DNA fragments exceeding 50 Mb (Fig. 4d–f). This finding indicates that although increased break frequency reduces linear DNA length, it increases the complexity of rearrangement fragments by increasing the repair event frequency.

3.4 Mathematical modeling of circular and linear DNA production in chromothripsis

Our chromothripsis simulation model consistently demonstrated this phenomenon: the mean number of circular DNAs consistently remained remarkably low under most conditions (Section 3.3). To elucidate this phenomenon, we established a simplified mathematical model to calculate the expected number of circular DNAs produced during a single chromothripsis event. This model disregards the enhanced single-fragment circularization propensity, given its minimal influ-

ence at a biologically relevant 7 : 1–40 : 1 chromatin compaction ratio. The model incorporates a key assumption of transient telomere–telomere ligation (details provided in Section 2) to accelerate computation without altering its results. On the basis of this model, we developed a Python program that generates a plot with the number of cuts on the x-axis and the mean number of circular DNA and linear DNA on the y-axis (Fig. 5). This plot allows for a comparative analysis of the number of circular DNAs and the number of linear DNAs and displays the trend of how both change with break frequency.

Two fundamental relationships related to the number of circular DNAs are shown in Fig. 5: a positive correlation with the number of cuts and an inverse correlation with the repair rate (this phenomenon can also be mathematically derived). Under conditions of maximal fragmentation (1000 breaks, although most chromothripsis events involve fewer than 1000 fragments) and perfect repair (100% efficiency), the expected number of circular DNA molecules reaches a maximum of 3.4 molecules per event. However, under typical conditions, the number of circular DNAs tends to be even lower, generally ranging from 0 to 2. Importantly, since extrachromosomal DNA (ecDNA) is a subset of circular DNA (specifically, short circular DNA containing functional regions), the

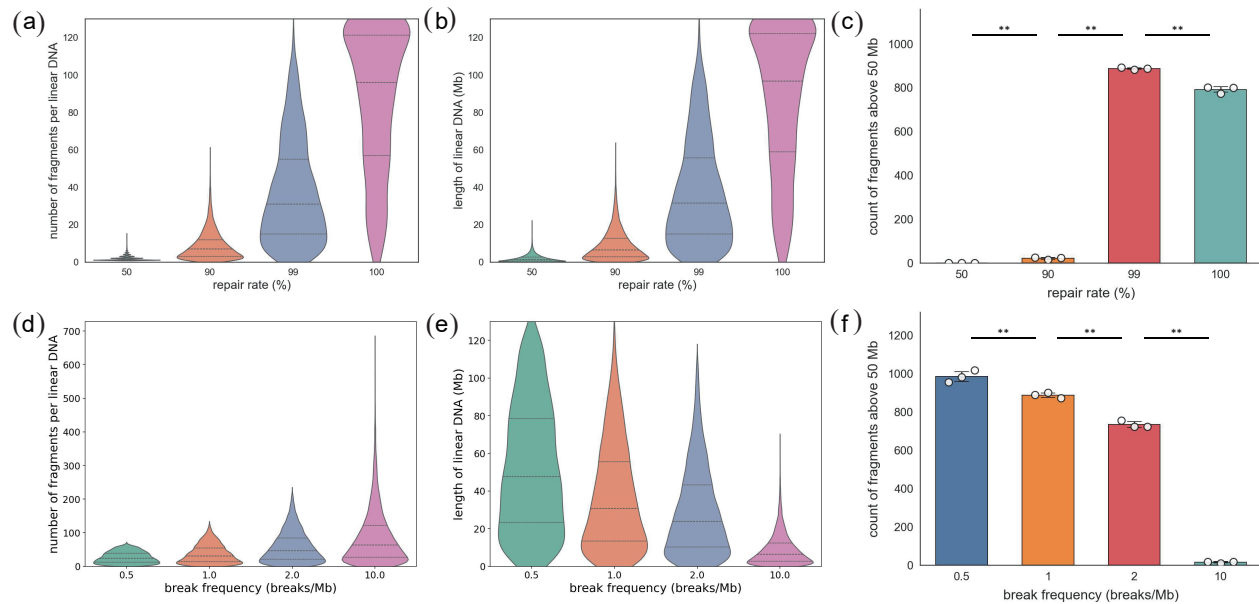


Fig. 4. The properties of linear DNA are affected by the repair rate, and the break frequency. (a, b) Illustration of how the repair rate influences the number of fragments per linear DNA (a) and the length of linear DNA (b), respectively. Sample size: 1000; parameters: chromosome length, 130.0 (Mb); compaction ratio, 25; break frequency, 1.0 (breaks/Mb). (c) The count of fragments above 50 Mb. Each data point represents the aggregate of 1000 simulation results, with three replicates per repair rate. Significant differences were observed between all adjacent groups ($p < 0.01$). (d, e) Illustration of how the break frequency influences the number of fragments per linear DNA (d) and the length of linear DNA (e), respectively. Sample size: 1000; parameters: chromosome length, 130.0 (Mb); compaction ratio, 25; repair rate, 99.0 (%). (f) Count of fragments above 50 Mb. Each data point represents the aggregate of 1000 simulation results, with three replicates per break frequency. Significant differences were observed between all adjacent groups ($p < 0.01$).

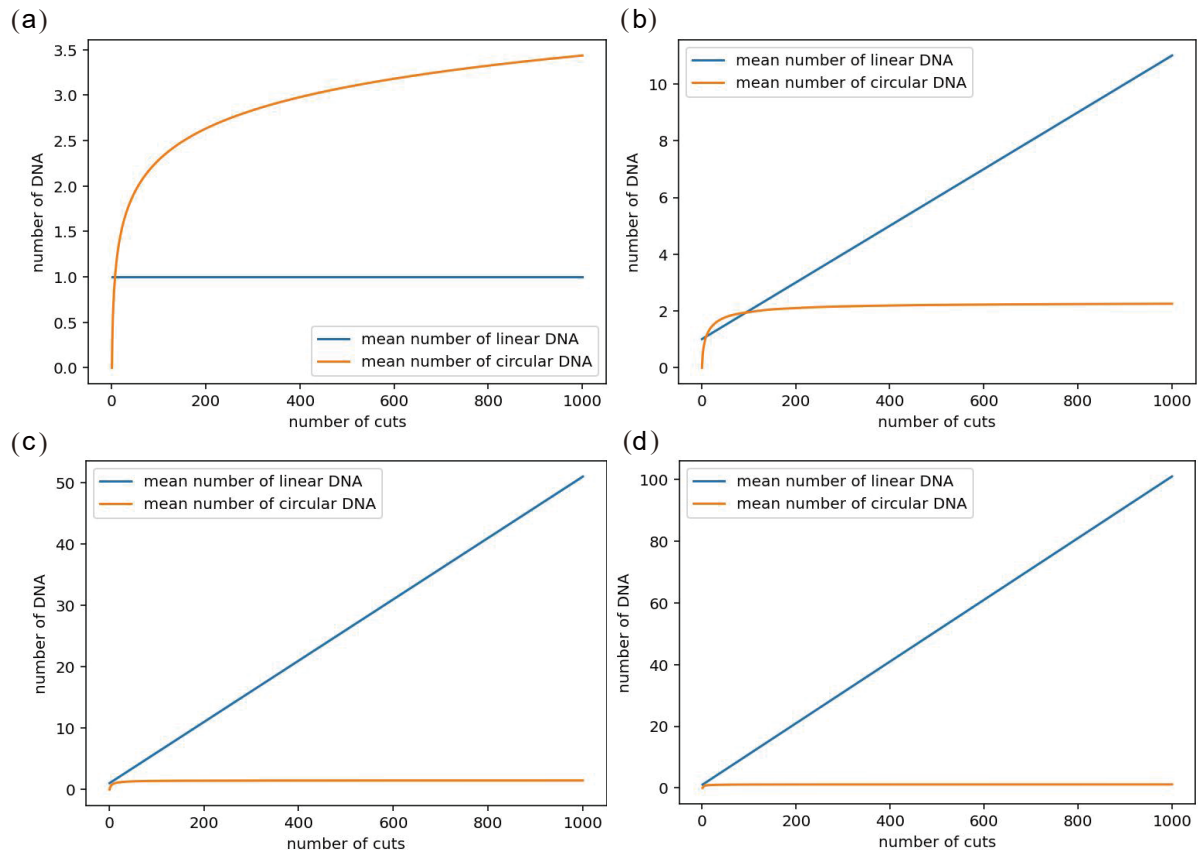


Fig. 5. The number of circular DNAs versus linear DNAs as a function of the number of cuts. These plots were generated by the program. The maximum cut number is set at 1000 but has different repair rates: (a) 100%; (b) 99%; (c) 95%; (d) 90%.

actual number of functional ecDNA molecules produced is even lower than the total circular DNA count. This phenomenon can be explained from two perspectives: First, only a very small fraction of ecDNA originates from chromothripsis, with the majority being generated through other mechanisms, such as the scarring excisional model or scarless excisional model^[20]. Second, this simulation model has several limitations: (i) the fragmentation process may not be fully random and preferentially produces shorter fragments, or (ii) it underestimates the condensation level of fragments, as they just enter interphase.

Another particularly significant observation is that when the repair rate falls below 99%, linear DNA fragments consistently outnumber circular DNA across most conditions. This quantitative dominance of linear DNA suggests that it may play an underappreciated role in tumorigenesis. We propose that these abundant linear DNA fragments could contribute to cancer progression through mechanisms such as activation of the cGAS-STING pathway^[21], potentially leading to PD-L1-mediated immune evasion^[22–24]. Consequently, our results suggest the need for greater research attention to linear DNA fragments in addition to ecDNA when studying chromothripsis in cancer. Understanding the implications of linear DNA in tumorigenesis could provide valuable insights into the mechanisms driving cancer evolution and the development of therapeutic strategies.

4 Conclusions

We successfully developed a chromothripsis simulation program, which effectively simulates the processes of chromothripsis, revealing key insights into its outcomes: (i) increased compaction ratios or break frequencies increase the number of circular DNA fragments by reducing the physical length of the fragment; (ii) increased repair rates increase circular DNA production by decreasing residual fragment quantity; (iii) increased repair rates and decreasing break frequencies increase the length of linear DNA while reducing its molecular count; and (iv) large linear DNA fragment generation strongly depends on a high repair rate. Another important discovery is the limitation in the production of circular DNA. These results were further verified by a newly developed computational program, challenging the prevailing assumption that chromothripsis is a major driver of ecDNA accumulation in cancers. Future research should explore the functional implications of these discoveries, especially concerning chromosomal instability and therapeutic resistance. Furthermore, these two programs highlight the underappreciated role of linear DNA fragments, which numerically dominate over circular DNA under most conditions. These fragments may contribute to tumorigenesis through mechanisms such as cGAS-STING pathway activation and immune evasion, suggesting their potential as therapeutic targets. While our models provide insights into the structural outcomes of chromothripsis, a key limitation is that they track only the rearrangement products and do not investigate the segregation of these fragments to daughter cells (or their loss). Consequently, they fail to capture the associated copy number (CN) changes. Subsequent research should consider these factors.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2022YFA0806800), the Plans for Major Provincial Science & Technology Projects of Anhui Province (202303a0702003), and the Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025XDXM505). We are grateful to Dr. Xuebiao Yao and Liangyu Zhang for critical reading.

Conflict of interest

The authors declare that they have no conflict of interest.

Biographies

Junxiao Yang is an undergraduate student at the School of Life Science and Medicine, University of Science and Technology of China. His current research focuses on the mechanism of chromothripsis and the development of programs to simulate chromothripsis.

Xing Liu is a Professor of MOE Key Laboratory for Cellular Dynamics, University of Science and Technology of China (USTC). She received her Ph.D. degree in Cell Biology from USTC in 2014. Her research career has been dedicated to solving the key scientific questions and challenges associated with cell fate decisions using chemical tools and new methods.

References

- [1] Krupina K, Goginashvili A, Cleveland D W. Scrambling the genome in cancer: causes and consequences of complex chromosome rearrangements. *Nature Reviews Genetics*, **2024**, 25: 196–210.
- [2] Pellestor F, Gatinois V. Chromoanasythesis: another way for the formation of complex chromosomal abnormalities in human reproduction. *Human Reproduction*, **2018**, 33: 1381–1387.
- [3] Cortés-Ciriano I, Lee J J K, Xi R, et al. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *Nature Genetics*, **2020**, 52: 331–341.
- [4] Liu P, Erez A, Nagamani S C S, et al. Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. *Cell*, **2011**, 146: 889–903.
- [5] Krupina K, Goginashvili A, Cleveland D W. Causes and consequences of micronuclei. *Current Opinion in Cell Biology*, **2021**, 70: 91–99.
- [6] Ejaz U, Dou Z, Yao P Y, et al. Chromothripsis: an emerging crossroad from aberrant mitosis to therapeutic opportunities. *Journal of Molecular Cell Biology*, **2024**, 16: mjae016.
- [7] Kwon M, Leibowitz M L, Lee J H. Small but mighty: the causes and consequences of micronucleus rupture. *Experimental & Molecular Medicine*, **2020**, 52: 1777–1786.
- [8] Lusk C P, King M C. Rotten to the core: Why micronuclei rupture. *Developmental Cell*, **2018**, 47: 265–266.
- [9] Hatch E M, Fischer A H, Deerinc T J, et al. Catastrophic nuclear envelope collapse in cancer cell micronuclei. *Cell*, **2013**, 154: 47–60.
- [10] Zhao M M, Wang F, Wu J H, et al. CGAS is a micronucleophagy receptor for the clearance of micronuclei. *Autophagy*, **2021**, 17: 3976–3991.
- [11] Terzoudi G I, Karakosta M, Pantelias A, et al. Stress induced by premature chromatin condensation triggers chromosome shattering and chromothripsis at DNA sites still replicating in micronuclei or multinucleate cells when primary nuclei enter mitosis. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, **2015**, 793: 185–198.

(Continued on page 0304–6)