A Novel Method to Investigate the Influence of H3K9-methylated Heterochromatin on Genetic Mutations

（遺伝子変異に対する H3K9 メチル化ヘテロクロマチンの影響を調査する新しい方法）

Ola Mahmoud Ibrahim Abdalla

Graduate School of Chemical Sciences and Engineering

Hokkaido University

2024

A Novel Method to Investigate the Influence of H3K9-methylated Heterochromatin on Genetic Mutations

Doctoral Dissertation Presented by

Miss Ola Mahmoud Ibrahim Abdalla

Graduate School of Chemical Science and Engineering

Student ID: 28215027

Supervisor`s Name: Prof. Ishimori Koichiro

Acknowledgments

I, Ola Abdalla, would like to express my deepest gratitude to **Prof. Ishimori Koichiro and Prof. Sada Kazuki** for their unwavering support. Their exceptional professionalism and guidance have been instrumental in my academic achievements, serving as a source of profound inspiration for me to complete my doctoral degree. I am truly thankful for the invaluable contributions they have made to my academic journey.

I acknowledge the significant contribution of Hokkaido University in providing lab facilities and resources. Finally, I am thankful for the collaboration with the University of Auckland in New Zealand and appreciate their contribution to this study.

Contents

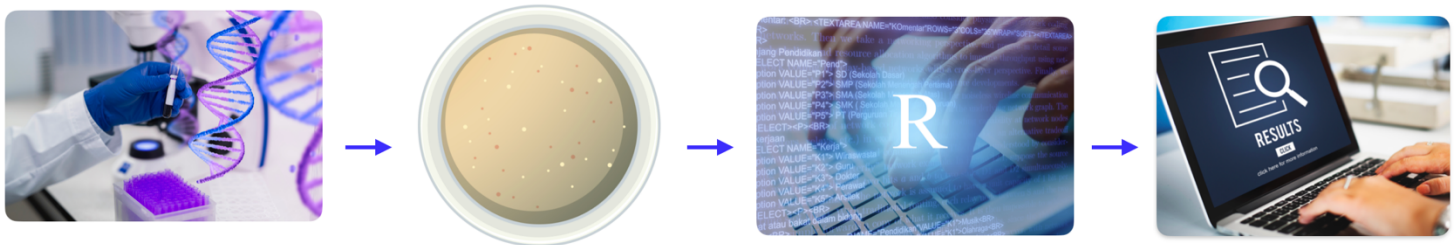
	Page Number
Chapter 1 Abstract	4
Chapter 2 Introduction	8
• H3k9me and Cancer Development	20
• Significance of This study	23
• Limitations of the Current Research	24
• Novel Method to Solve Current Research Limitations	
• Research purpose 1	28
• Research purpose 2	32
Chapter 3 Material and Methods	36
Chapter 4 Results	54
Chapter 5 Discussion	69
Chapter 6 Conclusion	76
Chapter 7 References	79
Copyright Remarks	87

Abstract

Genetic variation, driven by mutations, plays a crucial role in the evolution of species and the development of cancer. The mutation rate varies across the genome and can be influenced by chromatin organization and other factors. However, the relationship between chromatin organization and mutation rates has yet to be fully understood. One modification that affects chromatin organization is the methylation of histone H3 at lysine 9. This modification creates heterochromatin, which represses transcription in euchromatin and maintains genome stability, which is essential for the organism's survival. This research has two significant objectives. Firstly, it aims to investigate the impact of H3K9 methylation alone on mutation rates in fission yeast without

considering other histone markers. This groundbreaking research method uses fluctuation assays and tetracycline-activated genes to estimate mutation rates accurately and precisely on a single gene under two different conditions in a single experiment using only one isogenic clone. Secondly, this research seeks to solve the problem of statistically calculating mutation events and their 95% confidence intervals using the most accurate method, "Ma-Sandri-Sarkar Maximum Likelihood" (as mentioned in the graphical abstract below). The results suggest that the H3K9me markers can increase the phenotypic mutation rate of a single gene. Moreover, this study presents an R-code that uses the non-recursive version of the MSS likelihood to statistically analyze the genetic data efficiently and avoid software-related issues. The significance of the research goes beyond the results themselves. The genetic

assay and R-code utilized in this study have the potential to be used as powerful tools by future scientists to investigate the organization of chromatin and its influence on mutation rates. This could lead to significant breakthroughs in genetic mutation research, cancer, and evolutionary biology while paving the way for life-saving treatments.



Graphical Abstract: The Second Research Purpose. It is vital to provide an easy and user-friendly statistical analysis method for upcoming scientific researchers to analyze fluctuation assay results.

List of Abbreviations

(5-FOA)	5-fluoroorotic acid drug
***ura4*$^{+}$ gene**	The marker gene used in this study encodes for orotidine-5'-phosphate decarboxylase, an enzyme essential for the synthesis of uracil.
(tet)	Tetracycline drug
(-tet condition)	Indicating the absence of tetracycline in the cell culture and the presence of H3K9me on the *ura4*$^{+}$ gene
(+tet condition)	Indicating the presence of tetracycline in the cell cultures and absence of H3k9me on *ura4*$^{+}$ gene.

Introduction

1. The Growth and Survival of Human Beings

Genetic material is of great significance to all living organisms. It contains the genetic information that determines an organism's physical characteristics, such as its appearance, behavior, and susceptibility to certain diseases. Genetic material is passed down from parents to their offspring and is responsible for the diversity of life on Earth. Understanding the importance of genetic material can help us better comprehend the complexities of living organisms and the processes that occur within them [1]. As shown in Figure 1, the genetic material is stored in the nucleus of most eukaryotic cells. This genetic material is present in the form of linear and double-stranded DNA molecules, which are organized into structures known as

chromosomes. DNA is packed and condensed by the nuclear proteins to form chromatin. Based on the condensation level, chromatin can be categorized into heterochromatin and euchromatin. The genetic material performs three essential functions: the genotypic function, which involves replication; the phenotypic function, which affects gene expression; and the evolutionary function, which involves mutations that help the organism adapt to environmental changes and control its growth and development [1].

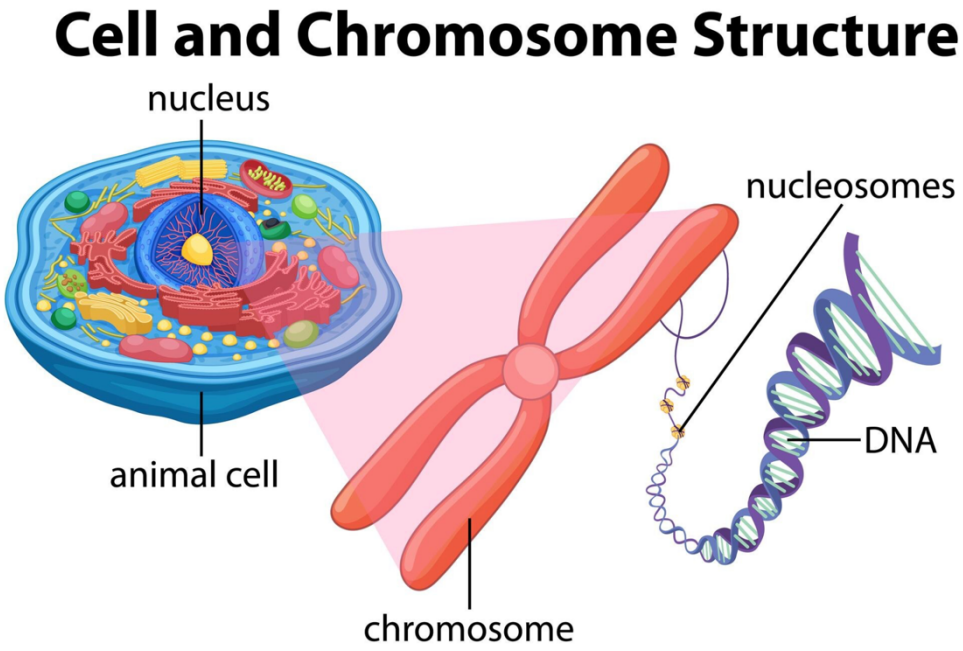


Figure 1: The Structure of Eukaryotic Cells [31]

2. The Daily Threats to Human Beings:

Mutation, which refers to the changes in genetic information, is one of the major obstacles that affect human development in daily life. Understanding the impact of mutation on human development is crucial. This is because these genetic changes, caused by metabolic by-products, DNA processing enzymes, errors during DNA synthesis, and other mutagens, can hinder human development [2] and induce spontaneous mutation. In addition, exposure to environmental factors, such as UV radiation and certain chemotherapeutic drugs like cisplatin, can lead to the formation of bulky lesions in DNA. These lesions can cause induced mutation and significantly distort the structure of the DNA double helix, making it difficult for the cell to replicate and transcribe genetic information. Compared to base damage, bulky lesions are more harmful because

they can strongly block DNA replication and transcription, causing damage to the overall health of the cell and cancer [2].

3. The Human Body`s Self-defense Mechanism:

To deal with mutations and maintain genome integrity, DNA damage response (DDR) pathways play a crucial role in repairing DNA [3] and fixing these issues (Figure 2). Also, this repair helps prevent cancer [4]

Figure 2: Imagine How Our Body Fix the Mutations Issues Every Day

There are several types of DNA-repairing mechanisms. The type of mutation in DNA determines the specific pathway involved. For example, base excision repair (BER) [5] is the crucial process that eliminates base lesions. This process relies heavily on the coordinated action of a series of enzymes, including DNA glycosylase, apurinic/apyrimidinic (AP) endonuclease, gap-filling DNA polymerase, and DNA ligase. The sequential coordination of these enzymes ensures that the repair product from one enzyme becomes the substrate for the next one, thus efficiently repairing the damage. This coordinated action is essential to prevent the accumulation of more harmful BER intermediates, such as AP sites or strand breaks. Without BER, these intermediates can accumulate and cause more severe damage. Therefore, BER is vital for

maintaining the integrity of the DNA molecule and, ultimately, the organism's overall health.

In addition, nucleotide excision repair (NER) is the most crucial pathway for repairing bulky lesions in our cells [6]. NER comprises over 30 repair proteins and is part of a larger group of repair mechanisms known as excision repair (ER). Together, BER and NER make up ER, acting as the first line of defense against a plethora of common DNA lesions, ensuring the maintenance of genome stability. While ER is an essential process, it can pose a challenge in chemotherapy efforts, acting as a double-edged sword. The significance of NER and its role in maintaining genomic health cannot be overstated - it is the cornerstone of DNA repair mechanisms.

4. The Chromatin Dynamics and Structure:

The nucleosome is the building block of chromatin, and its importance cannot be overstated. Composed of about 147 base pairs of DNA wrapped tightly around an octamer of histone proteins, it is the foundation upon which the structure and function of DNA are based (Figure 3). During nucleosome formation, the subnucleosomal structure called the tetrasome is formed by assembling two H3-H4 dimers on DNA, which are then incorporated with two H2A-H2B dimers to form the mature nucleosome. The histone linker protein H1 plays a crucial role in the compaction of chromatin into higher-order structures that make up chromosomes [7]. The formation of specific nucleosome arrays along the genome, resembling "beads on a string," confers different structural and functional chromatin states. This variability helps regulate gene expression, with the promoter

regions of actively transcribed genes having fewer nucleosomes and increasing nucleosome occupancy into coding regions. The binding of the negatively charged DNA phosphate backbone with the many positively charged (essential) amino acids of the nucleosome histones results in a strong DNA-histone association, which compacts chromatin into a tight heterochromatin structure [7].

Genes associated with heterochromatin are transcriptionally silent, and chromatin remodeling can alter the structure of chromatin. The nucleosome is a fundamental unit of chromatin responsible for forming specific nucleosome arrays along the genome, which regulate the expression of genes. The histone linker protein H1 helps in the compaction of chromatin into higher-order structures, which make up chromosomes. This tight chromatin structure inhibits cellular processes requiring direct interactions with the genome.

Alteration of chromatin structure, termed ‘‘chromatin remodeling,’’ and DNA accessibility are critical to gene expression, with implications for various biological processes. The complexity of this molecular mechanism demands regulation by a range of chromatin modifications. These include DNA methylation, non-coding RNAs (such as micro RNAs and long non-coding RNAs), polycomb-group proteins, and post-translational histone modifications. Their collective influence ensures genome integrity [7].

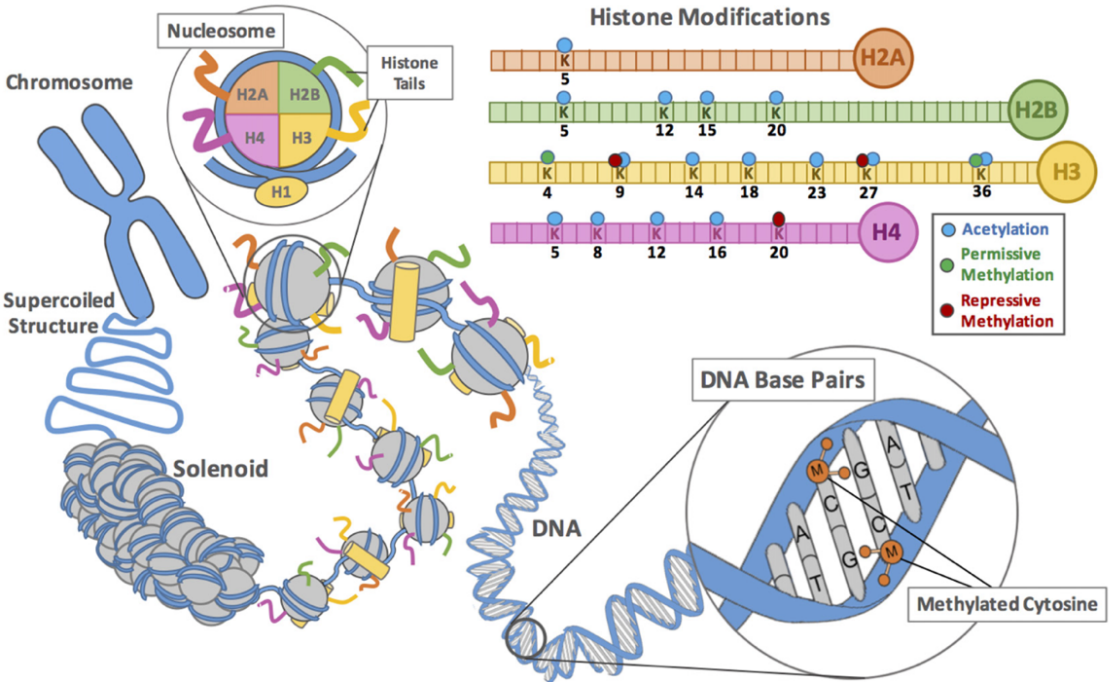


Figure 3: The Structure of Chromatin and Histone Proteins Organization [7]

5. The Correlation Between Chromatin Dynamics and DNA Repair Pathways:

Genetic stability is paramount in preventing mutations that can lead to grave consequences. Fortunately, Chromatin Dynamics and DNA Repair [8] Pathways are indispensable in maintaining this stability. However, when it comes to DNA repair, the chromatin state becomes a critical factor that determines the effectiveness of the repair process and the location of the mutations [9]. ER proteins must detect and remove damage across various chromatin structural domains and at different compaction levels. In other words, open chromatin or euchromatin regions are repaired more efficiently, while closed regions like heterochromatin are repaired less efficiently (Figure 4). This means that high mutation densities in cancer are often found in

closed chromatin and low mutation densities in open chromatin. Therefore, studying DNA repair in eukaryotes is essential as it helps us understand how repair factors handle the diverse chromatin landscapes to repair DNA damage and prevent mutations efficiently.

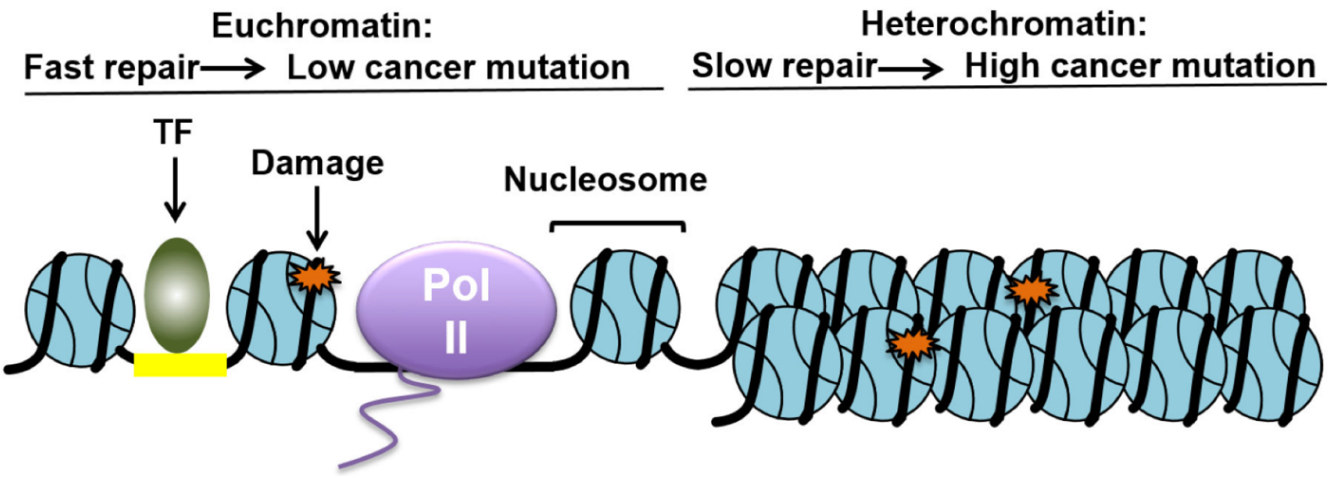


Figure 4: The Relationship Between Chromatin Structure and Repair Mutations in DNA [9]

6. Additional Factors Affect Chromatin and DNA Repair Pathways:

The relationship between DNA repair pathways and chromatin is intricate and influenced by crucial factors. These factors are essential to understanding genetic material protection and maintenance [8]. Histone post-translational modifications (PTMs), histone variants [10], chaperones [11], and chromatin remodeling complexes [11,12] are among these factors and have been identified as key players in the development and progression of diseases. Recent studies have revealed the existence of various PTMs, such as methylation, acetylation, phosphorylation, ADP-ribosylation, ubiquitylation, and SUMOylation. Their relationship with diseases has been explored, and

understanding their role in pathological processes is crucial in developing effective treatments for a wide range of diseases [13].

7. H3k9me and Its Role in Cancer Development

H3K9me is a vital post-translational modification (PTMs) that can be a powerful tool in creating transcriptional repression within euchromatin and generating heterochromatin [14] This process is a crucial component of eukaryotic cells, and it occurs in three distinct ways: mono-methylation (me1), di-methylation (me2), or tri-methylation (me3). With its ability to create these essential changes within cells, H3K9me is an indispensable tool that can transform how we understand and approach eukaryotic cell biology. These Methylation and demethylation processes regulate gene transcription, chromatin structure, and epigenetic state (Figure 5).

These processes are controlled by methyltransferases such as SUV39H and G9a, which are responsible for H3K9 methylation, and lysine demethylases (KDMs) that remove H3K9me, and H3K9 demethylases that reverse the function of H3K9 methylation [14]. Epigenetic regulators such as H3K9 demethylases have been found to positively impact repairing DNA double-strand breaks (DSBs), thus maintaining genomic integrity essential for healthy living.

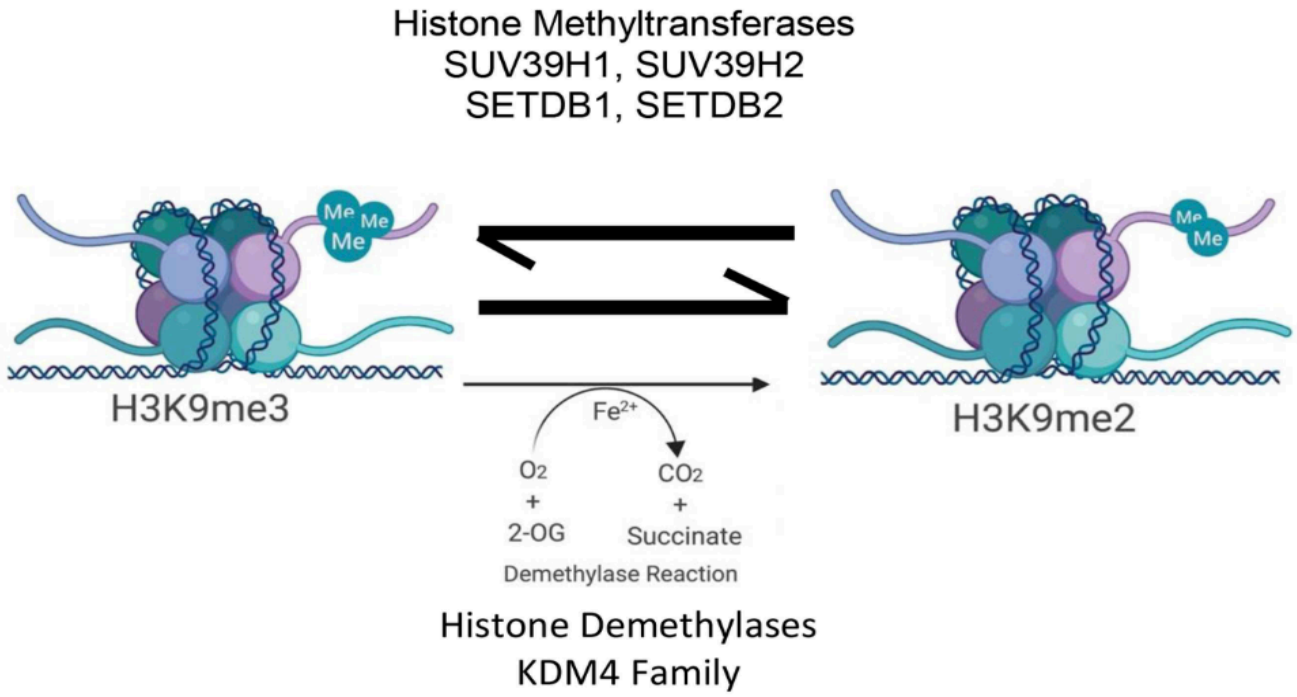


Figure 5: The regulation of H3K9me3 and its influence on transcription [14]

The dynamic nature of H3K9me3 emphasizes the involvement of multiple enzymes in its regulation. A precise group of readers, writers, and erasers of this histone mark establish and maintain the transcriptional landscape. However, even slight alterations in these enzymes can contribute to various diseases [14]

H3K9me3 is a vital histone modification that plays a significant role in developing various diseases, including hematopoietic malignancies and several types of cancer [15] [16]. Despite progress in developing inhibitors for H3K9me3 methyltransferases or demethylases, achieving selectivity remains an unmet challenge. The enzymes that regulate H3K9me3, such as the KDM4 family, have been found to exert a strong regulatory influence on oncogenic gene transcription.

Significance of this Study

Based on what is mentioned above, it's crucial to comprehensively understand H3K9me3's oncogenic function and epigenetic regulators that control its abundance in leukemia and other cancer cells. Through exploring new avenues and driving forward with this research, scientists can develop efficient therapeutic strategies to unlock the full potential of H3K9me3 as a therapeutic target. This will not only improve the lives of those affected by hematopoietic malignancies and other diseases where this histone modification is pathologically deregulated but also pave the way for developing groundbreaking treatments in the future [17].

8. The Limitations of the Current Research:

Understanding the intricate relationship between H3K9 methylation and the spontaneous mutation rates of genes has long been a focal point in genetic research. Despite its significance, this understanding has remained elusive over an extended period. Illustrated in Figure 6, numerous scholars have endeavored to explore the ramifications of H3K9me modification levels and heterochromatin on mutation rates in eukaryotes, as documented in references such as [18], [19], [20], [21], [22], among others. However, the clarity of their findings has often been compromised by questionable experimental methodologies and conflicting results. This could be attributed to the oversight evidence in many studies regarding the significant influence of genomic landscape variations across the same chromosome on mutation rates [23] [24]. Specifically, comparisons between mutation rates in euchromatin and heterochromatin domains have frequently been made without adequate control for variables such as distinct domains, genes, chromosomes, strains, or experimental conditions. Consequently, the

conclusions drawn from previous research remain partial, emphasizing the need for further investigations that can address these discrepancies within a unified experimental framework. Establishing such a framework is crucial to avoid biases arising from variable factors and to ensure a comprehensive understanding of the genetic phenomenon under study.

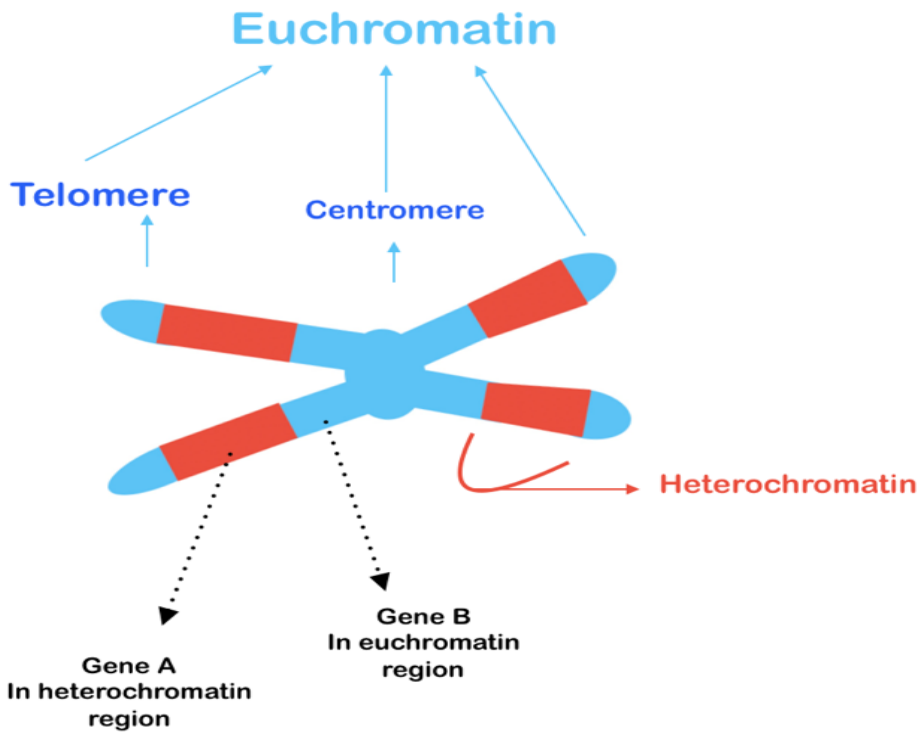
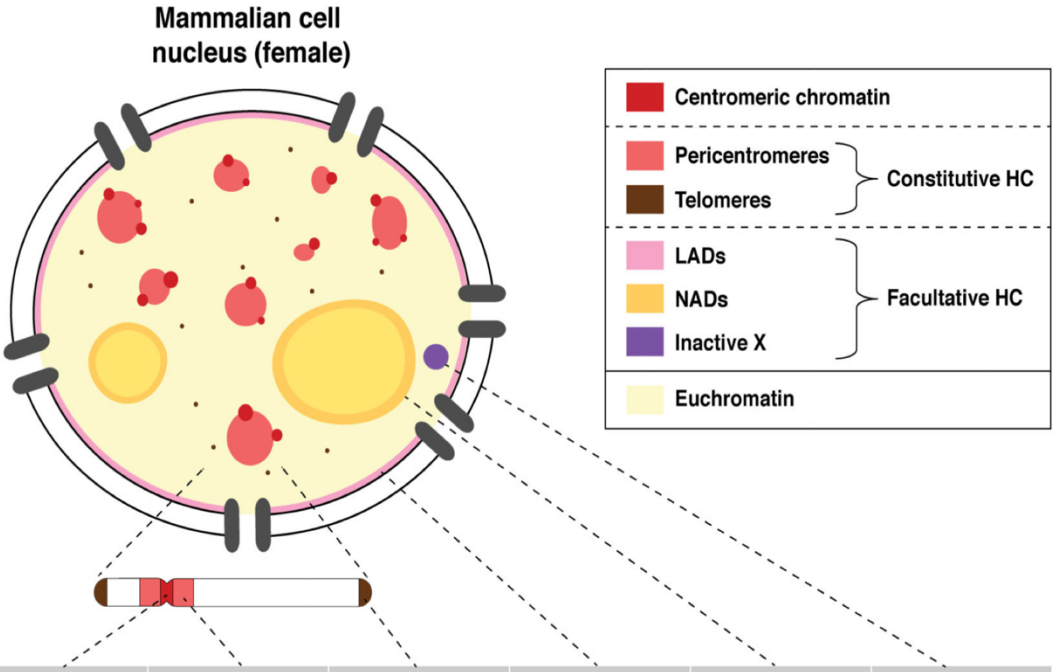


Figure 6: Most previous research compares two different genes with distinct DNA sequences, for example, A and B, into different locations to estimate the effect of heterochromatin on the mutation rate.

On the other hand, the second research dilemma is the sole existence of H3K9me. To gain a comprehensive understanding of H3K9 methylated heterochromatin's role in cancer development, it is imperative to concentrate solely on studying it in isolation at the same loci. Heterochromatin formation is related to various methylated histone modifications, such as H3K27me or H3K20me, that can coexist with H3K9me in the same domain, as shown in Figure 7 [25]. Investigating H3k9me together with other histone methylations like H3K27me, H3k20me, or different protein markers could potentially dilute or obscure the findings. Therefore, the research on H3K9 methylated heterochromatin alone must be prioritized to produce more accurate and conclusive results. By focusing solely on H3K9me, researchers can gain a deeper understanding of the role of its inhibitors in cancer therapy [14] and beyond. This knowledge can pave

the way for other drug developments [26] and contribute to a more effective cancer treatment.



	Centromere	Pericentromere	Telomere	LAD	NAD	Inactive X
Histone variants	CENP-A					MacroH2A
Histone modifications	H3K4me2 H3K36me2	H3K9me3 H3K27me1 H3K56me3 H4K20me3	H3K9me3 H4K20me3	H3K9me2 H3K27me3	H3K9me3 H3K27me3 H4K20me3	H3K27me3
Associated proteins / complexes	CENPs	HP1	Shelterin TRF1/2	Lamins		PRC2

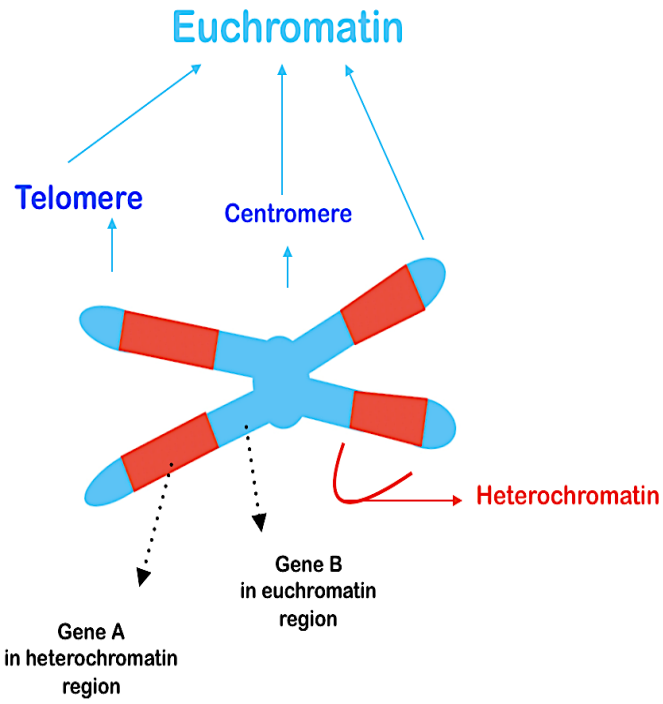
Figure 7: This diagram shows the histone variants, modifications, and associated proteins of constitutive and facultative heterochromatin domains.

Novel Method to Solve Current Research Limitations

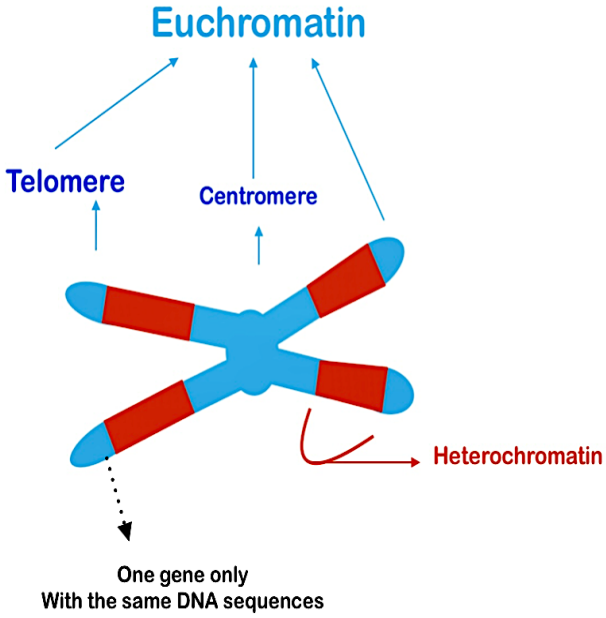
Research Purpose (1)

The experimental approach employed herein represents a significant contribution to scientific research. This pioneering study strategically bridges gaps in our understanding of genetic regulation, as demonstrated by a compelling comparison between previous research and our innovative experimental concept (Figure 8). This comprehensive research approach on the same gene involved switching between euchromatin and heterochromatin states using one isogenic clone in the same experiment. Furthermore, this study examines the influence of H3K9 methylation alone on the spontaneous mutational events per culture. The findings of this method provide reliable and precise data for the first time on the

impact of H3K9-methylated heterochromatin on spontaneous mutation rates, leaving no doubt about its remarkable significance. The results of this study represent a significant breakthrough in comprehending gene regulation by shedding light on this aspect of genetic regulation and setting a new standard for experimental design in this field.



Most previous research compares two different genes with distinct DNA sequences, such as Gene A and Gene B, placed in different locations to estimate the effect of heterochromatin on the mutation rate.



The Novelty of Our Study

Within the same experiment, a single gene can be switched between a euchromatic state and a heterochromatic state by adding tetracycline to the cell culture, allowing for the estimation of the difference in mutation rates between these states.

(-tet) Tetracycline conditions indicate the presence of H3K9me (heterochromatin)

(+tet) Tetracycline conditions represent the absence of H3K9me (euchromatin)

Figure 8: The uniqueness of this study. This diagram highlights the study's novelty, demonstrating how it stands out from previous research. This experimental design focuses solely on H3K9-methylated heterochromatin, unlike other studies examining the impact of various genetic and histone modifications on the mutation rate. Additionally, using an isogenic clone, this study confidently measures the impact of H3K9me on the mutation rate of the same gene in a single experiment. This approach distinguishes our study from previous research, which lacked the same level of precision and specificity.

To investigate these two scientific gaps, experiments were conducted to evaluate the impact of H3K9-methylated heterochromatin on the spontaneous mutation rate, using the fission yeast *Schizosaccharomyces pombe* as a model system [27]. *Schizosaccharomyces pombe*, similar to higher organisms, displays numerous chromatin modifications. [28]. Also, the *ura4*$^{+}$ gene was used as a reporter gene to evaluate gene silencing using a (5-FOA) drug [29]. Moreover, gene regulation by the tetracycline-activatable system or the inducible heterochromatin system at the *ura4*$^{+}$ gene was used.

This system was developed by [30], as mentioned in the method section and illustrated (Figure 10). This innovative system enabled to switch between activating and silencing the *ura4*$^{+}$ gene by adding (tet) into the cell culture to remove H3K9 me during the same experiment for a single clone using fluctuation assay, as shown in Figure 10.

This novel approach implemented in this study allowed for a confident investigation into the impact of H3K9-methylated heterochromatin on mutation rates of a single gene with identical DNA sequences while accurately determining the influence of H3K9me alone. The findings of this study and its highly accurate method have the potential to significantly impact various fields, including life science, evolutionary biology, and epigenetic therapy. This method has great potential to significantly contribute to drug development by providing scientists with a helpful research tool.

Research Purpose (2):

For several decades, scientists have been utilizing different methods, such as fluctuation assays, mutation-accumulation assays, and mutant accumulation [31,32] with whole-genome sequencing, to explore DNA mutations. However, these methods can be time-consuming and require a lot of effort and resources compared to fluctuation experiments [8]. That's why fluctuation assays have been an indispensable tool since 1943 [33] for accurately estimating the mutation rate. By using these powerful assays, scientists can efficiently unravel the mysteries of DNA mutations, saving time and resources while advancing our understanding of the genetic field. Because of the long-term significance of fluctuation assays, my

collaborator and I have two main objectives related to this research purpose.

First, this study offers valuable and user-friendly tools for future scientists to accurately analyze results from fluctuation assays and calculate mutation rates, as illustrated in Figure 9. To achieve this, the study provides R codes that use the Ma-Sandri-Sarkar maximum likelihood method to calculate the parameter (m). These codes are freely available and can be used with R software.

Second, this R-code version provides a reliable solution to calculate the 95% confidence intervals (CI) [34] using the MSS likelihood method. In some cases, the data generated by fluctuation tests may exceed the conventional values of (r). This creates a significant challenge when using R software's recursive version of this

likelihood. Therefore, this solution is critical to ensuring an accurate interpretation of the results [35].

Our main objective is to simplify the process for scientists to conduct statistical analyses. By doing so, we aim to ensure that the statistical analyses are more precise and dependable, which will help researchers make informed decisions and draw meaningful conclusions in the fields of population genetics of mutations and bioinformatics [36].

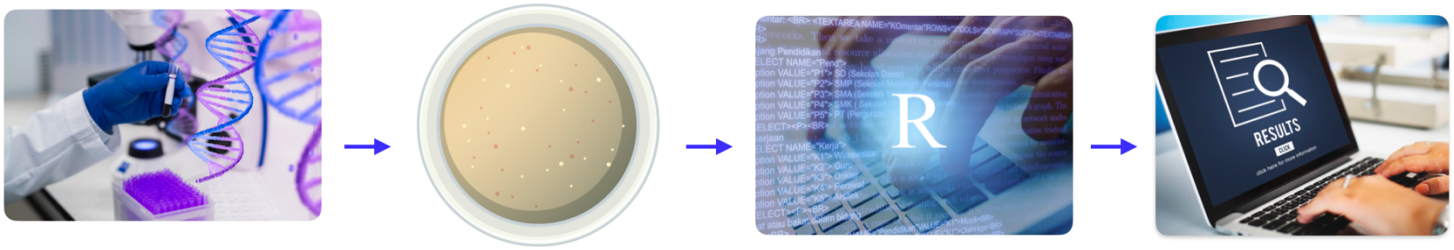


Figure 9. The Significance of This Study. It is crucial to provide upcoming scientific researchers with an easy, user-friendly method for statistically analyzing the results of fluctuation assays. These assays are typically conducted in laboratory settings, where the culture is plated on an agar plate with selective media. Subsequently, it becomes necessary to calculate the number of mutant colonies that appear on the surface in R-Software, as they are indicative of mutations. Our R-codes provide a time-efficient solution for data analysis, enabling researchers to obtain the parameter (m) and 95% intervals for the parameter (m) within a few minutes compared to other laborious techniques or estimators. This helps researchers gain insights about the mutation rate for genetic studies and drug development.

Materials and Methods

Materials and Methods for Research Purpose (1):

1. Strain and Media:

The fission yeast or *Schizosaccharomyces pombe* strain used in this study was obtained from Moazed Danesh's group and named SPY5101, as described in Ragunathan's study [30]. These researchers inserted *10XtetO* sites upstream of a 114bp fragment of the *ura4*$^{+}$ promoter at the endogenous locus of the *ura4*$^{+}$ gene. Furthermore, the GFP fusion protein was inserted downstream of the *ura4*$^{+}$gene. Under these conditions, TetR-Clr4-I can bind with *10XtetO*, leading to the establishment of ectopic heterochromatin and loss of *ura4*$^{+}$ gene function after silencing, as demonstrated in (Figure 10).

In contrast, the addition of (tet) is critical for the regulation of *ura4*$^{+}$ gene expression, resulting in the removal of ectopic heterochromatin and activation of the *ura4*+ gene, as illustrated in Figure 4B and Figure S3's red bar in Ragunathan's study [30]. As a result, the (*10XtetO- ura4 -GFP)* marker shows sensitivity to (5-FOA) drug in the presence of (tet).

By utilizing this system during these experiments, the lack of tetracycline or (- tet condition) implies silencing the *ura4*$^{+}$ gene due to the presence of H3K9-methylated heterochromatin. However, the presence of (tet) in the cell culture or (+ tet condition) means activation of the *ura4*$^{+}$ gene in the euchromatin region, which leads to cell death because the active *ura4*$^{+}$gene is sensitive to the (5-FOA) drug.

Under these 2 conditions during one experiment, mutant colonies might appear on solid media with (5-FOA drug + tet) only if there is a

mutation within the *ura4*$^{+}$ gene. This procedure allowed for estimating the phenotypic mutation rate in the absence and presence of (tet) at the same gene locus (*ura4*$^{+}$ gene). By utilizing this method, a meticulous examination of the correlation between H3K9-methylated heterochromatin and mutation rates of a particular gene can be conducted. Furthermore, this approach allows for the exclusive scrutiny of the impact of H3K9-methylated heterochromatin on the gene's mutation rate, resulting in highly accurate outcomes.

As for the media for cell growth, this strain, which contains (*10XtetO-ura4+-GFP*), was cultivated using a fully synthetic medium EMMG [37] during fluctuation assays. The drug concentrations were precisely

adjusted after several pilot experiments and a literature review. 5-FOA was set to (1 g/L), while (tet) was set to (10 µM/L).

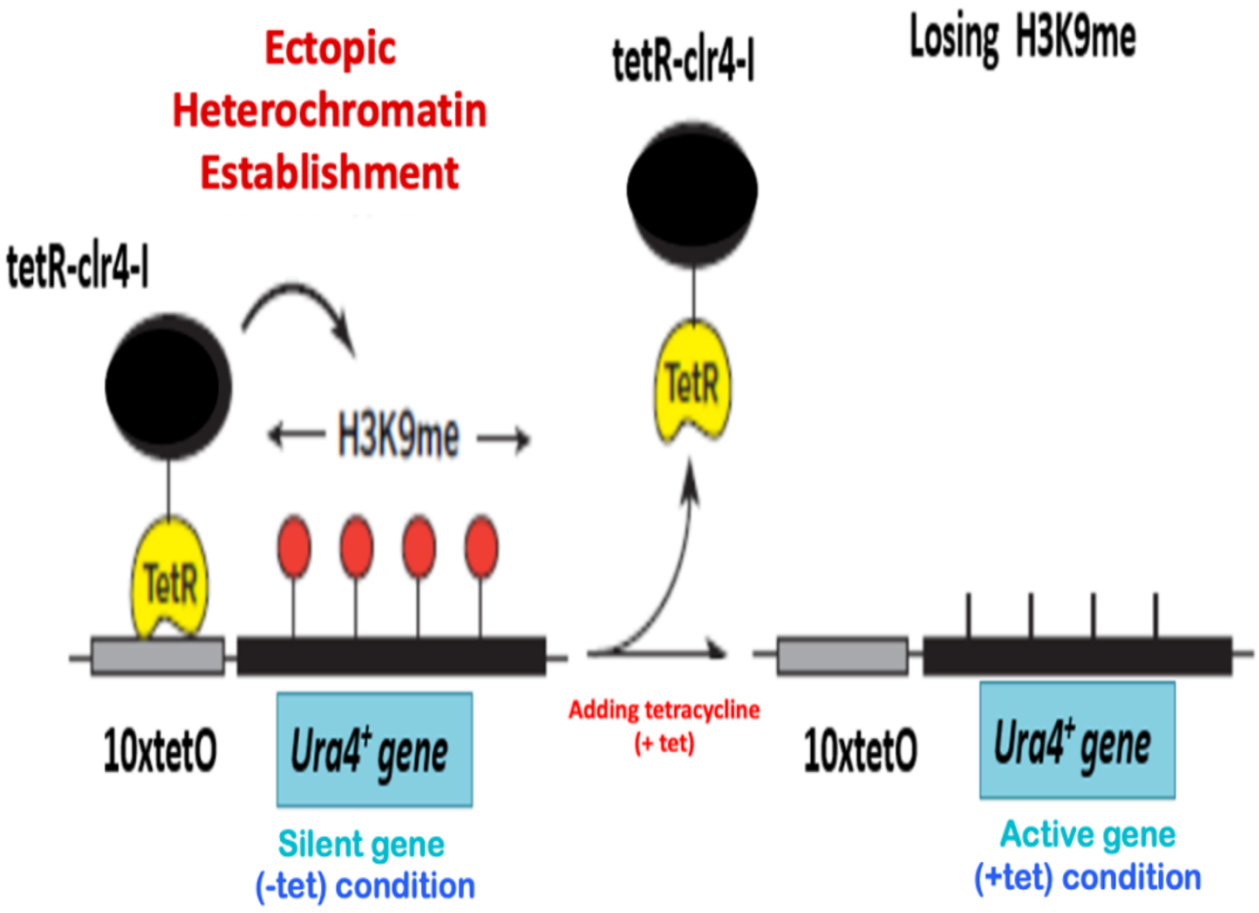


Figure 10: The diagrammatic representation presented herein elucidates the establishment of H3K9-methylated heterochromatin at the *ura4+* gene by the binding of TetR-Clr4-I with *10XtetO*, which is inserted upstream of the gene. Upon addition of tetracycline to the cell culture, TetR-Clr4-I's affinity *to 10XtetO* is lost, resulting in the dissipation of H3K9me and subsequent reactivation of the *ura4+* gene. This leads to the cells becoming sensitive to (5-FOA).

2. Silencing Assay:

To ensure a successful estimation of the mutation rate, it is paramount to validate the strain's sensitivity to both drugs (5-FOA + tet) in media. Taking this crucial step will guarantee accurate results and prevent errors in the subsequent analysis. Therefore, the strain with the modified reporter (*10XtetO- ura4*$^{+}$*-GFP*) was expertly spotted on plates containing yeast extract (YE), (5-FOA) (1 g/L), and (tet) (10 µM/L) after a precise four- or five-fold serial dilution. In the meantime, the present experiment utilized the wild-type *ura4*$^{+}$ gene as a positive control, while *ura4*$^{+}$Δ cells were implemented as the negative control. To determine the mutation rate with precision, it was sufficient to merely count the number of colonies or mutants (r) produced at the end of each fluctuation assay on media with (5-FOA + tet).

3. Developing Fluctuation Assay for Increased Precision:

Accurately estimating the mutation rate is essential to comprehending the mechanisms of evolution and disease. Selecting the appropriate method is vital to obtain an accurate estimation of the parameter (m), which denotes the number of mutation events per culture. Although several methods are available, fluctuation assays stand out as the most precise and straightforward approach to detecting the parameter (m). This method can ensure the accuracy of our findings and mitigate the shortcomings of other techniques, such as mutation accumulations and mutant accumulation assays. Therefore, employing fluctuation assays can increase the reliability of our results and provide us with a

better understanding of the underlying mechanisms of evolution and disease.

To obtain precise results, it is imperative to adopt a rigorous approach to experimental design and data analysis. The proposed enhancements in this study, which are grounded in the most recent [38] and original fluctuation assay protocols [33], have made it possible to obtain reliable data.

➢ Newly Developed Fluctuation Assay Steps:

These assays were done for 2 isogenic clones of the previously mentioned strain after doing single-clone isolation, as mentioned in Figure 11. Then, a single clone was grown in liquid EMMG media at 32°C for each assay until saturation. The initial inoculum (No) of approximately 200 cells was obtained by diluting the preculture into fresh media and splitting it into two different media types. The first media condition, which contains solely liquid EMMG media, is a crucial aspect of the experiment as it represents the absence of tetracycline, also known as the (-tet condition). The second media condition has EMMG with (tet) (10 µM/L), representing the presence of tetracycline (+tet condition).

With the use of 96-well plates, 10 µl of diluted cultures were dispensed for each experimental condition to enhance the overall precision of the

essay. This 96-well format allows adjustment of all parallel cultures to 10 microliters and uses foil covers to prevent evaporation and light exposure. After that, these cultures in 96-well plates were left to grow until they reached saturation point. 24 parallel cultures were taken afterward from each experimental condition to determine the average number of cells per culture after saturation. The remaining 72 parallel cultures were plated separately for each condition with and without (tet) on a 9 cm petri dish containing EMMG, (5-FOA) (1 g/L), and (tet) (10 μM/L). After being incubated at 32°C in the dark for 3-4 days, the number of mutants (r) was recorded for each parallel culture separately. All observed mutants were recorded except for jackpot cultures, which were excluded from calculating parameter (m). Concurrently, some of these colonies were picked up and replated on

a 9 cm petri dish containing EMMG, (5-FOA) (1 g/L), and (tet) (10 μM/L) to double-check their phenotypic resistance on these drugs.

Isolation of a single colony on agar plate

Pickup 1 isogenic clone to prepare preculture in a liquid EMMG media

Incubate the preculture overnight until it reaches saturation

Simultaneously dilute the previous preculture into two types of fresh media

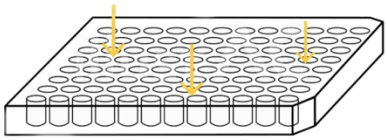
First Culture

Fresh media EMM with (10 µM/L) tetracycline (+ tet condition), indicating the absence of H3k9me and the presence of euchromatin

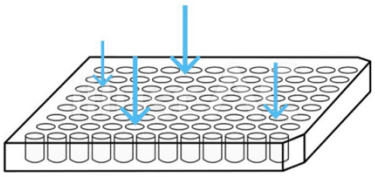
Secound Culture

Fresh media EMM without tetracycline (- tet condition), representing presence of H3k9me and the ectopic heterochromatin

Dispense 10 µL of this culture containing tetracycline into a 96-well plate



Dispense 10 µL of this culture into a 96-well plate



Incubate the 2 well plates at 32 degree Celsius until saturation is reached

Plating 72 wells or saturated cell culture.

Each well is transferred individually onto an agar plate, containing EMMG media, supplemented with 5-fluoroorotic acid or 5-FOA at (1 g/L), and anhydrotetracycline at (10 µM/L).

Plating 72 wells or saturated cell culture.

Each well is transferred individually onto an agar plate, containing EMMG media, supplemented with 5-fluoroorotic acid or 5-FOA at (1 g/L), and anhydrotetracycline at (10 µM/L).

Incubate all 9 cm Petri dishes from both conditions in the 32 degree Celsius incubator for 3 - 4 days

Calculate the number of mutant (r) or 5-FOA residence colonies in each agar plate individually

This represent the count of mutants for (+ tet condition)

Calculate the number of mutant (r) or 5-FOA residence colonies in each agar plate individually

This represent the count of mutants for (- tet condition)

Figure 11: Conducting a Fluctuation Assay for Each Clone. The figure presented here explains the detailed procedure for conducting a fluctuation assay for a single clone. Furthermore, it is important to mention that two different assays were carried out for two distinct clones to ensure the accuracy and reliability of the results.

4. Statistical Analysis for the Results of Fluctuation Assays Using MATLAB Software:

To calculate the expected number of mutation events per culture or parameter (m) in MATLAB, the number of mutants or colonies on each agar plate was counted for each parallel culture. According to equation 18 from [31], the phenotypic mutation rates (μ) were determined using (μ = parameter m/Nt), where (Nt) represents the average number of cells per culture before plating on the agar plates. By calculating the 95% confidence intervals for the parameter (m) using the MSS Maximum Likelihood or Ma–Sandri–Sarkar maximum likelihood to

ensure the most accurate results [32]. The MSS-maximum-likelihood [31] approach in R with the recursive likelihood function [39] was used. This calculation method has proven to be the most reliable over the years, making it the optimal choice.

This study led us to valuable data beyond the conventional values of (r), which posed a significant challenge when using the recursive version of the likelihood in R. Despite this, a non-recursive version of the likelihood function, along with non-parametric bootstrapping, was developed to implement the MSS Maximum Likelihood approach. This innovative approach produced excellent results and filled a gap in the literature, as no other non-recursive version of the likelihood appears to be available. For more details about calculating 95% confidence intervals, please read the methods section of research purpose (2) below. With the use of this technique for each of the 4 experimental

setups, the confidence intervals were calculated for the findings to ensure accuracy.

5. Computational Simulation Using MATLAB Software:

To ensure the accuracy of the fluctuation data, a comparison was made between the predicted cumulative frequency distribution of mutants and the experimental data for each experimental condition. A plot was generated using MATLAB® to represent the results visually. This methodology allowed for verifying the accuracy of the data and ensuring that our findings were based on reliable statistical distributions.

By harnessing the power of a MATLAB file from a comparable study [40], highly precise and accurate computational simulations were

performed in this study. Lang's groundbreaking study has conclusively demonstrated that the Luria-Delbrück distribution can be fully accounted for by the one-parameter model [33]. In addition, the probability distribution for the number of mutants per culture in the two-parameter model is the joint distribution of the Luria-Delbrück (parameter m) and the Poisson (parameter n = md) distributions. The mutation following the plating process (d) was accurately assessed using this two-parameter model. This Div. post-plating or value (d) represents the total number of cell divisions capable of producing mutants after plating the cells on the selective media with both (5-FOA drug + tet).

By fitting the results of fluctuation assays into reliable one-parameter and two-parameter distributions, the accuracy of a fluctuation assay simulation was improved to ensure reliable findings.

Materials and Methods for Research Purpose (2):

The study utilized R software to design functions that accurately calculate the parameter (m) using the Ma-Sandri-Sarkar maximum likelihood [31] method. The main goal, as mentioned in the introduction section, has been accomplished through the creation of these functions or R-code.

As part of our study, it is crucial to address the challenge of using the recursive version of the MSS likelihood in R software. To achieve this, a code that calculates 95% confidence intervals using a recursive equation was developed that efficiently computes the Luria-Delbruck distribution based on the Lea-Coulson generating function. Although the recursive equation is used to estimate the parameter m in the Ma-Sandri-Sarkar maximum likelihood approach, it can cause computational problems when implementing it in R due to stack

growth. To avoid this, a non-recursive version of the likelihood function and used non-parametric bootstrapping were employed to quantify parameter uncertainty. By using these techniques, it is possible to get results when using the MSS Maximum Likelihood approach and fills a gap in the literature, as no other non-recursive version of the likelihood appears to be available.

To determine how many cultures returned each distinct (r) value, a unique experimental (r) value vector and a count vector for each (r) value were leveraged. It is necessary to employ a powerful non-recursive function for Pr(m) and a negative log-likelihood function, written in terms of m, the two datasets, and the Pr(m) function. By minimizing this function via R's optimization, 95% confidence intervals were calculated using equations 29 and 30 [31]. Table 1 illustrates the metadata of the R code for reproducibility.

Additionally, a non-parametric bootstrapping technique was used to estimate confidence intervals. By resampling the experimental values 1000 times and calculating m using the MSS Maximum Likelihood approach for each bootstrap replicate, we generated confidence intervals by extracting the 2.5th and 97.5th percentiles from the 1000 bootstrap replicates.

Nr	Code metadata description	
C1	Current code version	v1
C2	Permanent link to code/repository used for this code version	https://github.com/OlaAbdalla/Fluctuation-Assay
C3	Permanent link to reproducible capsule	https://codeocean.com/capsule/8197897/tree/v1
C4	Legal code license	Apache License, 2.0
C5	Code versioning system used	none
C6	Software code languages, tools and services used	R

Table 1: The Code Metadata

Results

The Results for Research Purpose (1):

1. Tetracycline removes ectopic H3K9-methylated heterochromatin and increases cell sensitivity to the drug (5-FOA):

Before commencing the fluctuation assays, it is imperative to execute a silencing assay to confirm the successful removal of H3K9me from the *(10XtetO-ura4+-GFP)* gene by the tetracycline and to evaluate the cell sensitivity to the 5-FOA drug.

In Figure 12, the results show that (tet) effectively removed ectopic H3K9-methylated heterochromatin, consistent with the results of Ragunathan's study shown in Figure 4B [30].

As a consequence of its removal, the *(10XtetO-ura4+-GFP)* marker simultaneously exhibited sensitivity to the 5-FOA drug in the presence of (tet) in a manner analogous to that of wild-type cells containing the ura4+ gene (Figure 12). The mutation rate for each experimental condition was accurately determined by simply counting the colonies or mutants (r) generated at the end of each fluctuation assay, after plating the cells on the selective media with both (5-FOA drug and tet).

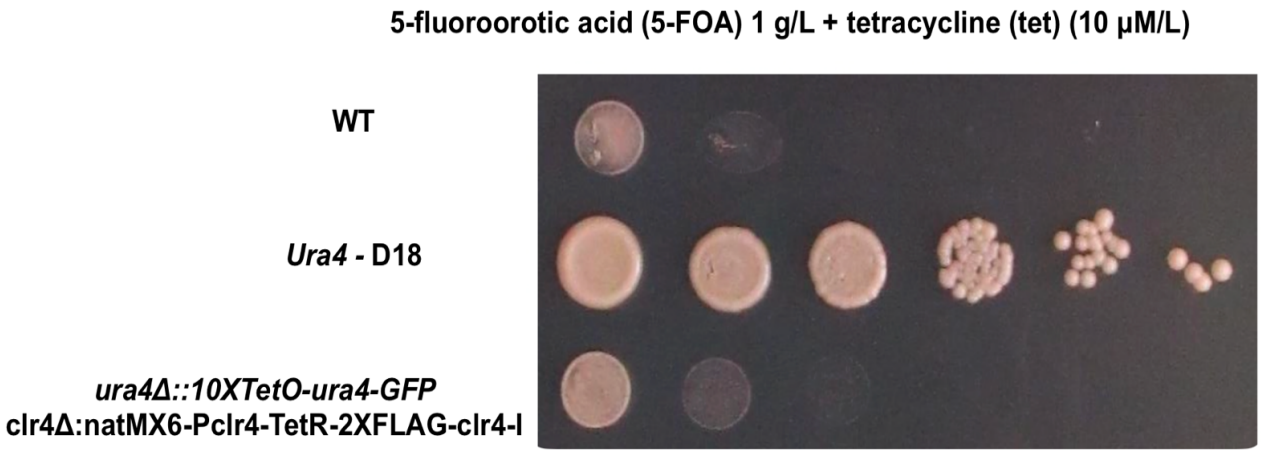


Figure 12: Removal of Ectopic H3K9-methylated Heterochromatin in the Presence of Tetracycline (tet). The effectiveness of (tet) in removing H3K9me was tested with specific media containing (5-FOA and tet). The purpose was to determine the successful reactivation of the (*10XtetO-ura4+-GFP*) gene and cell sensitivity. The sensitivity of the (*10XtetO-ura4+-GFP*) reporter, the only strain used in fluctuation assays, was tested using 5-fluoroorotic acid (5-FOA), as shown in the figure. In addition, positive and negative control cells were used, with wild-type *ura4* gene as the positive control and *ura4Δ* cells as the negative control.

2. Ectopic H3K9-methylated heterochromatin has Increased the Phenotypic Mutation Rate:

After conducting a fluctuation assay for each clone, it is essential to accurately determine the number of mutants or colonies on agar plates and estimate parameter (m). To ensure the accuracy of data, the

number for every parallel culture was individually calculated under two conditions: (-tet and +tet).

Then, the calculations of the phenotypic mutation rates were performed following the specific steps elaborated in the Materials and Methods section.

According to the statistical data presented in Table 2, it can be concluded that establishing H3k9me on the *ura4*$^{+}$ gene locus increases the number of mutation events per culture (parameter m). For clones 1 and 2, the values of the parameter (m) under (-tet) condition are approximately 1.4773 and 9.3216, respectively. Similarly, the mutation rates of clones 1 and 2 under the (-tet) condition are about 4.6310344×10^{-6} and 3.7436144×10^{-5}, subsequently. Based on all the previous values, the results suggest that the number of mutation events per culture (parameter m) and mutation rates are

higher than what is observed in clone 1 and clone 2 cell cultures when there are no H3K9me markers and tetracycline is present, as in the (+tet condition). Furthermore, we could calculate 95% confidence intervals of all parameters (m) separately, as detailed in Table 2.

By statistically combining the phenotypic mutation rates of Clone 1 (4.6310344×10^{-6}) and Clone 2 (3.7436144×10^{-5}) under (-tet conditions), we were able to determine a phenotypic mutation rate of (1.907426×10^{-5}) with a 95% confidence interval via bootstrap [41] [1.653×10^{-5}, 2.211×10^{-5}]. Similarly, when we statistically combined the phenotypic mutation rates of Clone 1 (2.3932659×10^{-6}) and Clone 2 (6.9356807×10^{-6}) under (+tet conditions), we obtained a phenotypic mutation rate of (4.755245×10^{-6}) with a 95% confidence interval via bootstrap [41] [3.794×10^{-6}, 5.960×10^{-6}].

Thus, these results showed that the presence of H3k9me on the *ura4*$^{+}$ gene locus under (-tet conditions) increased the phenotypic mutation rate (μ) in clones 1 and 2.

	Clone 1		Clone 2	
Experimental conditions	Presence of H3K9me (- tet condition)	Absence of H3K9me (+ tet condition)	Presence of H3K9me (- tet condition)	Absence of H3K9me (+ tet condition)
Parameter (m)	1.4773 (Figure 13 A)	0.7108 (Figure 13 B)	*9.3216* (Figure 13 C)	1.4773 (Figure 13 D)

Division post-plating (d)	0	0	0	0
95% confidence intervals (CI) for parameter (m) via bootstrap [42]	[0.995, 2.018]	[0.469, 1.028]	[7.942, 11.015]	[1.146, 1.884]
Phenotypic mutation rate (μ)	4.6310344×10^{-6}	2.3932659×10^{-6}	3.7436144×10^{-5}	6.9356807×10^{-6}

Table 2: The Presence of H3K9me on the *ura4*$^{+}$ gene Increases the Rate of Phenotypic Mutations. This study's results show that the presence of H3K9-methylated heterochromatin significantly raised the number of mutation events per culture (parameter m) of the same gene. These results were obtained by measuring the number of mutants (r) or 5-fluoroorotic acid resistance (5-FOA) colonies in the presence of tetracycline (tet) for two different clones of the same strain with the *(10XtetO-ura4*$^{+}$*-GFP)* marker. When calculating the mutation rate using the equation (μ= parameter m/Nt), our findings revealed a significant difference in phenotypic mutation rates between the presence and absence of H3K9me.

3. Luria-Delbrück Model Distributions Match Fluctuation Assay Results:

Upon completion of statistical analysis, the precision of our fluctuation results and parameter (m) in Table 3 were assessed by comparing mutants' predicted cumulative frequency distribution to the experimental data. To present these findings, a plot using MATLAB was created, allowing for demonstrating the results' accuracy.

As shown in Figure 13, the results obtained in this study are in perfect alignment with one-parameter models. This data is best described by the Luria-Delbrück distribution or the one-parameter model, suggesting that post-plating growth and mutation do not occur. This is supported by the (d = 0) for the two-parameter model. The agreement between the results and the model indicates that this study`s findings

are both accurate and reliable. Therefore, these results provide supporting the conclusion that H3K9me has increased the phenotypic mutation rate of the same gene.

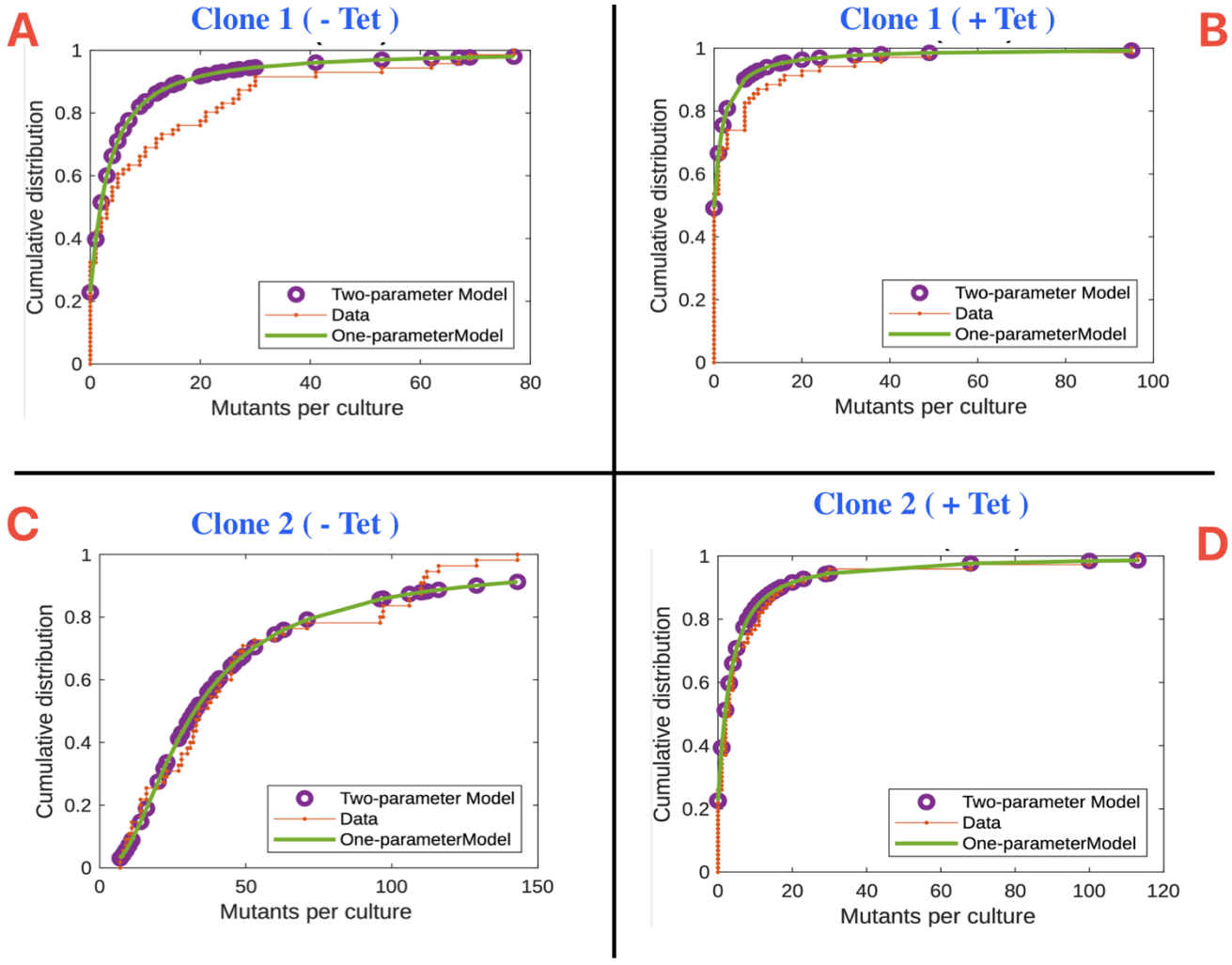


Figure 13: The Computational Analysis of the Fluctuation Assays Results. Our experiment using clone 1 and clone 2, both with and without tetracycline, yielded some significant results. The data we obtained is represented by a solid red curve called "data". The green solid lines represent the Luria-Delbrück (one-parameter model) distributions, which is a perfect fit for our accurate data. According to Luria-Delbrück's experimental principle, the alignment of experimental data with the Luria-Delbrück distribution is crucial in determining the accuracy of fluctuation assay results. Moreover, the purple circles indicated the two-parameter model for Luria-Delbrück and passion distributions, which also fit our data, with a Div. post-plating value (d = 0). Notably, in Figures 4A and 4B for clone 1, the parameter "m" value was 1.4773 in the absence of tetracycline and 0.7108 in its presence. Similarly, in Figures 4C and 4D for clone 2, the value of the parameter "m" was 9.3216 in the absence of tetracycline and 1.4773 in its presence.

The Results for Research Purpose (2):

This brief study utilized biological data obtained from fluctuation assays in a previous genetic study (Table 3 of the article [32]). The aim of using this biological data is to showcase the implementation of our R codes for calculating the parameter 'm' using the MSS maximum likelihood method. With the aid of the R code in this study, the parameter 'm' for experimental data containing multiple mutants was computed. Below is Table 3 in this resrach, providing a concise and easily understandable representation of the data.his code includes two functions, conditionalProb_r() and loglik():conditionalProb_r() takes as input *r* and *m*, and returns as output the probability of obtaining *r* mutations in a culture with expected number of mutations *m*, assuming the Luria-Delbruck distribution;

- loglik()takes as input: the expected number of mutations per culture *m*, a vector of unique experimental mutations *rVec*, and the corresponding vector of counts for each unique experimental r *rCard* and returns as output the log-likelihood assuming the Luria-Delbruck distribution.

Input	(r) Values or the number of mutants per culture	The data is in two forms in the code. One is just a list of the values obtained (called "dt"): 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 3 3 3 4 4 4 4 5 6 7 7 9 The second form is as a pair of vectors, one with unique values (called "rDt"): 0 1 2 3 4 5 6 7 9 and one with the number of times it occurred (called "cDt"): 11 17 12 3 4 1 1 2 1

		So there 11 values of 0 in dt, 17 of 1, etc… From (Table 3 of the article [32]).
Output	Parameter (m) based on MSS-Maximum-Likelihood Method	1.106733
	95% confidence intervals (CL) for parameter (m) via equatic 29 and 30 [31]	[0.827, 1.481]
	95% confidence intervals (CL) for the parameter (m) via bootstrap	[0.898, 1.358]

Table 3. By utilizing the most accurate method available, MSS-Maximum-Likelihood, and the R-codes developed in this study, we have obtained statistical analysis results that are both reliable and insightful. Generally, scientists begin by calculating the number of mutants on selective media for each parallel culture following fluctuation assays. They then utilize software to statistically calculate the number of mutational events per culture or parameter (m), as well as the 95% confidence intervals for parameter (m) like this example.

These findings demonstrate that a non-recursive version of the function can be employed to overcome the issue of recursive format. The R-code we used can calculate the parameter (m) and 95% confidence intervals using both formulas 29 and 30 [31] and Bootstrap. Thus, these results showcase the effectiveness of the R-code that utilizes the MSS maximum likelihood method in generating precise and dependable outcomes, making it a valuable tool for future researchers in the field of bioinformatics and beyond.

Discussion

The comprehension of the impact of H3K9 methylation on gene mutation rates is a paramount concern in scientific inquiry. However, some researchers have used unsuitable and questionable methods to investigate this phenomenon. Specifically, their analysis relied on data from two genes situated in different chromatin regions across the chromosome as visualized in Figure 6, or alternative methods. Such flawed approaches can lead to erroneous conclusions. It is imperative to refine the scientific approach and employ more precise methods to advance scientific research and drug development.

The attention should be focused on a meticulous investigation of the impact of H3K9me on the mutation rate of the same gene.

This study introduces a novel and advanced approach (Figure 8) that combines several practical components and crucial breakthroughs. The tetracycline-activatable system can be utilized with full assurance, enabling scientists to switch on and off the expression of a specific gene with certainty, as demonstrated in Figure 10.

A new methodology has been introduced in this study, which enabled the investigation of the impact of H3K9-methylated heterochromatin on the mutation rates of a gene with identical DNA sequences. The study also presents a clear and concise step-by-step process for performing the fluctuation assay, enhancing accuracy, and improving the results, as shown in Figure 11. Thus, this comprehensive approach can be utilized by future scientists in evolutionary biology and other fields to advance their research with greater accuracy, leading to more reliable outcomes.

Regarding the results of research purpose No.1, it is evident that clones 1 and 2 exhibit significant differences in parameter (m) in the presence or absence of H3k9me, as illustrated in Table 2. Additionally, the calculation of the phenotypic mutation rate using the formula (μ = parameter m / Nt) revealed a significant increase in the phenotypic mutation rate in the presence of H3k9me on the ura4+ gene.

Based on the values of (μ), the presence of H3k9me on the ura4+ gene in clones 1 and 2 elevated the phenotypic mutation rates of the same gene. This indicates that the objective of conducting quantitative analysis to investigate the research goal was successfully achieved.

However, the objective of conducting a semi-quantitative analysis has not yet been fulfilled. Specifically, the present study, despite its meticulous approach, failed to present the mutational spectra arising

from the phenotypic mutation rate due to the absence of DNA sequencing experiments for the mutants after the fluctuation assay.

Nevertheless, this study introduces a novel approach for investigating multiple epigenetic modifications on the same gene mutation rate, thereby offering a fresh perspective and conceptual framework to the field. The findings provide valuable insights into the complexities of genetic analysis and hold significant potential for the future of drug development for H3K9me-related diseases for H3K9me-related diseases [43] and evolutionary biology [44].

Regarding research purpose No.2, the Luria-Delbrück fluctuation assay has emerged as a pivotal experiment in the calculation of mutation rates, particularly in the fields of genetic and mutation research. Its reliability and accuracy have rendered it the preferred method for numerous researchers, providing a robust foundation for

further research. This study introduces an effective statistical method that employs R software to analyze genetic outcomes and the number of mutants obtained from fluctuation assays (Figure 9).

This R-code facilitates researchers in analyzing their data more efficiently. The present R-codes are distinct from other estimators [45], as they can calculate the parameter (m) and 95% confidence intervals without requiring the input of each mutant's number alongside the number of cells (Nt). This is a significant advantage as it saves valuable time that would otherwise be expended in inputting each mutant number with the corresponding number of cells.

The ability to calculate the parameter (m) and 95% confidence intervals without the need for additional input is a unique feature of the present R-codes. This feature is particularly useful for researchers who are working with large datasets and need to analyze their data

quickly and accurately. By eliminating the need for additional input, the present R-code can help researchers save time and increase their productivity.

Most importantly, the present methodology utilizing the R-codes is designed to be accessible to molecular biologists in the genetic field without requiring a sophisticated interface [46]. This is particularly advantageous as such interfaces often necessitate an understanding of advanced statistics to calculate parameters (m), which can be a challenge for those without a statistical background. This approach circumvents this issue, making it a more efficient and user-friendly option for researchers in the field.

This R-code is based on the Ma-Sandri-Sarkar Maximum Likelihood method, which provides reliable and accurate data interpretation. This method offers a promising alternative to current approaches and

provides researchers with a user-friendly, efficient, and dependable tool to enhance the fields of bioinformatics and genetics. Thus, these software codes are a valuable asset to researchers who seek to streamline their work and obtain accurate results in a timely manner.

Conclusion

This study has successfully addressed two significant scientific challenges through a precise and innovative experimental approach. It introduced a novel method to estimate the impact of H3K9-methylated heterochromatin on gene mutation rates by utilizing identical isogenic clones under different experimental conditions within a single experiment.

Moreover, the study suggests that H3K9me3 histone markers alone significantly increase the expected number of mutation events per culture (parameter m) and the phenotypic mutation rate.

The findings of this study are of great significance to the scientific community, as they provide a better understanding of the impact of H3K9-methylated heterochromatin on gene mutation rates. The

study's precise and innovative experimental approach serves as a model for future research in this field. By elucidating the mechanisms of H3K9me-associated mutations, this breakthrough empowers scientists to develop more effective cancer treatments, potentially saving millions of lives worldwide. This research represents a significant contribution to the field and has the potential to revolutionize cancer treatment

On the other hand, the Luria-Delbrück fluctuation assay is widely recognized as a reliable and effective method for detecting mutations across a diverse range of organisms, including bacteria [47] [48] and yeast [49] [50] [51]. This assay is a valuable tool for identifying low-frequency mutations and estimating mutation rates, providing insights into the impact of genetic factors on the evolution of organisms. The R code utilized in this study enables researchers to

perform statistical analyses of the fluctuation assay results and calculate 95% confidence intervals, making it an indispensable tool for mutation research.

References

[1] S. Minchin, J. Lodge, Understanding biochemistry: structure and function of nucleic acids, Essays Biochem. 63 (2019) 433–456. https://doi.org/10.1042/ebc20180038.

[2] DNA Repair and Mutagenesis, (2005). https://doi.org/10.1128/9781555816704.

[3] R. Huang, P.-K. Zhou, DNA damage repair: historical perspectives, mechanistic pathways and clinical translation for targeted cancer therapy, Signal Transduct. Target. Ther. 6 (2021) 254. https://doi.org/10.1038/s41392-021-00648-7.

[4] A. Ui, N. Chiba, A. Yasui, Relationship among DNA double-strand break (DSB), DSB repair, and transcription prevents genome instability and cancer, Cancer Sci. 111 (2020) 1443–1451. https://doi.org/10.1111/cas.14404.

[5] R. Prasad, W.A. Beard, V.K. Batra, Y. Liu, D.D. Shock, S.H. Wilson, A review of recent experiments on step-to-step "hand-off" of the DNA intermediates in mammalian base excision repair pathways, Mol. Biol. 45 (2011) 536. https://doi.org/10.1134/s0026893311040091.

[6] M. Kusakabe, Y. Onishi, H. Tada, F. Kurihara, K. Kusao, M. Furukawa, S. Iwai, M. Yokoi, W. Sakai, K. Sugasawa, Mechanism and regulation of DNA damage recognition in nucleotide excision repair, Genes Environ. 41 (2019) 2. https://doi.org/10.1186/s41021-019-0119-6.

[7] I.C.G. Weaver, A.C. Korgan, K. Lee, R.V. Wheeler, A.S. Hundert, D. Goguen, Stress and the Emerging Roles of Chromatin Remodeling in Signal Integration and Stable Transmission of Reversible Phenotypes,

Front. Behav. Neurosci. 11 (2017) 41. https://doi.org/10.3389/fnbeh.2017.00041.

[8] N. Nair, M. Shoaib, C.S. Sørensen, Chromatin Dynamics in Genome Stability: Roles in Suppressing Endogenous DNA Damage and Facilitating DNA Repair, Int. J. Mol. Sci. 18 (2017) 1486. https://doi.org/10.3390/ijms18071486.

[9] P. Mao, J.J. Wyrick, Organization of DNA damage, excision repair, and mutagenesis in chromatin: A genomic perspective, DNA Repair 81 (2019) 102645. https://doi.org/10.1016/j.dnarep.2019.102645.

[10] J. Ferrand, A. Plessier, S.E. Polo, Control of the chromatin response to DNA damage: Histone proteins pull the strings, Semin. Cell Dev. Biol. 113 (2021) 75–87. https://doi.org/10.1016/j.semcdb.2020.07.002.

[11] U. Chakraborty, Z.-J. Shen, J. Tyler, Chaperoning histones at the DNA repair dance, DNA Repair 108 (2021) 103240. https://doi.org/10.1016/j.dnarep.2021.103240.

[12] J. Stadler, H. Richly, Regulation of DNA Repair Mechanisms: How the Chromatin Environment Regulates the DNA Damage Response, Int. J. Mol. Sci. 18 (2017) 1715. https://doi.org/10.3390/ijms18081715.

[13] X. Wu, M. Xu, M. Geng, S. Chen, P.J. Little, S. Xu, J. Weng, Targeting protein modifications in metabolic diseases: molecular mechanisms and targeted therapies, Signal Transduct. Target. Ther. 8 (2023) 220. https://doi.org/10.1038/s41392-023-01439-y.

[14] L. Monaghan, M.E. Massett, R.P. Bunschoten, A. Hoose, P.-A. Pirvan, R.M.J. Liskamp, H.G. Jørgensen, X. Huang, The Emerging Role of H3K9me3 as a Potential Therapeutic Target in Acute Myeloid Leukemia, Front. Oncol. 9 (2019) 705. https://doi.org/10.3389/fonc.2019.00705.

[15] H.-Y. Jeon, A. Hussain, J. Qi, Role of H3K9 demethylases in DNA doublestrand break repair, J. Cancer Biol. 1 (2020) 10–15. https://doi.org/10.46439/cancerbiology.1.003.

[16] M. Zhou, J. Yan, Q. Chen, Y. Yang, Y. Li, Y. Ren, Z. Weng, X. Zhang, J. Guan, L. Tang, Z. Ren, Association of H3K9me3 with breast cancer prognosis by estrogen receptor status, Clin. Epigenetics 14 (2022) 135. https://doi.org/10.1186/s13148-022-01363-y.

[17] Y. Yang, M. Zhang, Y. Wang, The roles of histone modifications in tumorigenesis and associated inhibitors in cancer therapy, J. Natl. Cancer Cent. 2 (2022) 277–290. https://doi.org/10.1016/j.jncc.2022.09.002.

[18] M.V. de la Peña, P.A.M. Summanen, M. Liukkonen, I. Kronholm, Chromatin structure influences rate and spectrum of spontaneous mutations in Neurospora crassa, Genome Res. 33 (2023) 599–611. https://doi.org/10.1101/gr.276992.122.

[19] M. Habig, C. Lorrain, A. Feurtey, J. Komluski, E.H. Stukenbrock, Epigenetic modifications affect the rate of spontaneous mutations in a pathogenic fungus, Nat. Commun. 12 (2021) 5869. https://doi.org/10.1038/s41467-021-26108-y.

[20] P. Polak, R. Karlić, A. Koren, R. Thurman, R. Sandstrom, M.S. Lawrence, A. Reynolds, E. Rynes, K. Vlahoviček, J.A. Stamatoyannopoulos, S.R. Sunyaev, Cell-of-origin chromatin organization shapes the mutational landscape of cancer, Nature 518 (2015) 360–364. https://doi.org/10.1038/nature14221.

[21] J.G. Prendergast, H. Campbell, N. Gilbert, M.G. Dunlop, W.A. Bickmore, C.A. Semple, Chromatin structure and evolution in the human genome, BMC Evol. Biol. 7 (2007) 72. https://doi.org/10.1186/1471-2148-7-72.

[22] J. Xia, L. Han, Z. Zhao, Investigating the relationship of DNA methylation with mutation rate and allele frequency in the human genome, BMC Genom. 13 (2012) S7. https://doi.org/10.1186/1471-2164-13-s8-s7.

[23] K.D. Makova, R.C. Hardison, The effects of chromatin organization on variation in mutation rates in the genome, Nat. Rev. Genet. 16 (2015) 213–223. https://doi.org/10.1038/nrg3890.

[24] A. Gonzalez-Perez, R. Sabarinathan, N. Lopez-Bigas, Local Determinants of the Mutational Landscape of the Human Genome, Cell 177 (2019) 101–114. https://doi.org/10.1016/j.cell.2019.02.051.

[25] A. Fortuny, S.E. Polo, The response to DNA damage in heterochromatin domains, Chromosoma 127 (2018) 291–300. https://doi.org/10.1007/s00412-018-0669-6.

[26] J. Du, W. Liao, H. Wang, G. Hou, M. Liao, L. Xu, J. Huang, K. Yuan, X. Chen, Y. Zeng, MDIG-mediated H3K9me3 demethylation upregulates Myc by activating OTX2 and facilitates liver regeneration, Signal Transduct. Target. Ther. 8 (2023) 351. https://doi.org/10.1038/s41392-023-01575-5.

[27] C.S. Hoffman, V. Wood, P.A. Fantes, An Ancient Yeast for Young Geneticists: A Primer on the Schizosaccharomyces pombe Model System, Genetics 201 (2015) 403–423. https://doi.org/10.1534/genetics.115.181503.

[28] R. Nakamura, J. Nakayama, Regulation of the SUV39H Family Methyltransferases: Insights from Fission Yeast, Biomolecules 13 (2023) 593. https://doi.org/10.3390/biom13040593.

[29] H.P. Cam, S. Whitehall, Reporter Gene Silencing Assays in Fission Yeast, Cold Spring Harb. Protoc. 2016 (2016) pdb.prot091512. https://doi.org/10.1101/pdb.prot091512.

[30] K. Ragunathan, G. Jih, D. Moazed, Epigenetic inheritance uncoupled from sequence-specific recruitment, Science 348 (2015) 1258699. https://doi.org/10.1126/science.1258699.

[31] P.L. Foster, Methods for Determining Spontaneous Mutation Rates, Methods Enzym. 409 (2006) 195–213. https://doi.org/10.1016/s0076-6879(05)09012-9.

[32] W.A. Rosche, P.L. Foster, Determining Mutation Rates in Bacterial Populations, Methods 20 (2000) 4–17. https://doi.org/10.1006/meth.1999.0901.

[33] S.E. Luria, M. Delbrück, MUTATIONS OF BACTERIA FROM VIRUS SENSITIVITY TO VIRUS RESISTANCE, Genetics 28 (1943) 491–511. https://doi.org/10.1093/genetics/28.6.491.

[34] L. Hespanhol, C.S. Vallio, L.M. Costa, B.T. Saragiotto, Understanding and interpreting confidence and credible intervals around effect estimates, Braz. J. Phys. Ther. 23 (2019) 290–301. https://doi.org/10.1016/j.bjpt.2018.12.006.

[35] D. Kallogjeri, D.O. Francis, J.F. Piccirillo, Use of Confidence Intervals to Improve the Quality of Research Results Reporting and Interpretation, JAMA Otolaryngol.Head Neck Surg. 145 (2019) 491–493. https://doi.org/10.1001/jamaoto.2019.0793.

[36] J.S. Morris, V. Baladandayuthapani, Statistical contributions to bioinformatics: Design, modelling, structure learning and integration, Stat. Model. 17 (2017) 245–289. https://doi.org/10.1177/1471082x17698255.

[37] J. Petersen, P. Russell, Growth and the Environment of Schizosaccharomyces pombe, Cold Spring Harb. Protoc. 2016 (2016) pdb.top079764. https://doi.org/10.1101/pdb.top079764.

[38] G.I. Lang, Genome Instability, Methods and Protocols, Methods Mol. Biol. 1672 (2017) 21–31. https://doi.org/10.1007/978-1-4939-7306-4_3.

[39] R Core Team (2021) R A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. - References - Scientific Research Publishing, (n.d.). https://www.scirp.org/(S(czeh2tfqw2orz553k1w0r45))/reference/referencespapers.aspx?referenceid=3131254 (accessed October 14, 2023).

[40] G.I. Lang, A.W. Murray, Estimating the Per-Base-Pair Mutation Rate in the Yeast Saccharomyces cerevisiae, Genetics 178 (2008) 67–82. https://doi.org/10.1534/genetics.107.071506.

[41] O. Abdalla, C. Walker, K. Ishimori, R-code for calculating fluctuation assay results and 95% confidence intervals based on Ma-Sandri-Sarkar maximum likelihood, Softw. Impacts (2024) 100661. https://doi.org/10.1016/j.simpa.2024.100661.

[42] O. Abdalla, C. Walker, R-codes for Calculating Fluctuation Assay Results and 95% Confidence Intervals Based on Ma-Sandri-Sarkar Maximum Likelihood, (2023). https://doi.org/10.21203/rs.3.rs-3646152/v1.

[43] S.P. Rowbotham, F. Li, A.F.M. Dost, S.M. Louie, B.P. Marsh, P. Pessina, C.R. Anbarasu, C.F. Brainson, S.J. Tuminello, A. Lieberman, S. Ryeom, T.M. Schlaeger, B.J. Aronow, H. Watanabe, K.K. Wong, C.F. Kim, H3K9 methyltransferases and demethylases control lung tumor-propagating cells and lung cancer progression, Nat. Commun. 9 (2018) 4559. https://doi.org/10.1038/s41467-018-07077-1.

[44] M.I. Lind, F. Spagopoulou, Evolutionary consequences of epigenetic inheritance, Heredity 121 (2018) 205–209. https://doi.org/10.1038/s41437-018-0113-y.

[45] bz-rates - mutation rates calculator for fluctuation assays, (n.d.). https://www.lcqb.upmc.fr/bzrates (accessed November 26, 2023).

[46] dbRIP - FALCOR, (n.d.). https://lianglab.brocku.ca/FALCOR/FALCOR_old.html (accessed November 26, 2023).

[47] C.F. Pope, D.M. O'Sullivan, T.D. McHugh, S.H. Gillespie, A Practical Guide to Measuring Mutation Rates in Antibiotic Resistance, Antimicrob. Agents Chemother. 52 (2008) 1209–1214. https://doi.org/10.1128/aac.01152-07.

[48] B.A. Bridges, The fluctuation test, Arch. Toxicol. 46 (1980) 41–44. https://doi.org/10.1007/bf00361244.

[49] P. Jiang, A. Ollodart, M.J. Dunham, A Modified Fluctuation Assay with a CAN1 Reporter in Yeast, BIO-Protoc. 12 (2022). https://doi.org/10.21769/bioprotoc.4435.

[50] P. Jiang, A.R. Ollodart, V. Sudhesh, A.J. Herr, M.J. Dunham, K. Harris, A modified fluctuation assay reveals a natural mutator phenotype that drives mutation spectrum variation within Saccharomyces cerevisiae, ELife 10 (2021) e68285. https://doi.org/10.7554/elife.68285.

[51] H. Liu, J. Zhang, The rate and molecular spectrum of mutation are selectively maintained in yeast, Nat. Commun. 12 (2021) 4044. https://doi.org/10.1038/s41467-021-24364-6.

[52] Dissertation cover photo, (n.d.). https://www.freepik.com/ (accessed November 8, 2023).

[53] O. Abdalla, C. Walker, Estimation of mutation rate for fluctuation assay via MSS Maximum Likelihood | Code Ocean, (n.d.). https://codeocean.com/capsule/8197897/tree/v1 (accessed April 28, 2024).

[54] O. Abdalla, C. Walker, GitHub - OlaAbdalla/Fluctuation-Assay: R-code for Calculating Fluctuation Assay Results and 95% Confidence Intervals Based on Ma-Sandri-Sarkar Maximum Likelihood, (n.d.). https://github.com/OlaAbdalla/Fluctuation-Assay (accessed April 28, 2024).

Copyrights Remarks

First, in this research project, the author of this dissertation, Ola Abdalla, developed a novel method termed "fluctuation assays" (Figure 11). She employed the fission yeast, *Schizosaccharomyces pombe* strain SPY5101, which was obtained from Moazed Danesh's group, as documented in Ragunathan's study [30]. Ola Abdalla devised this method (Figure 11) after testing various experimental conditions and conducting pilot experiments based on the principles of Luria-Delbrück [33].

Second, Ola Abdalla personally financed part of this study, using her own resources to conduct statistical analysis, data analysis, and more, due to the lack of research tools available at the university to complete her research. **By investing her personal funds, she has acquired**

partial intellectual property rights to this research project alongside Hokkaido University's rights.

Third, Ola Abdalla acknowledges that the copyright for the R-code used in this study belongs to Professor Cameron Walker and the University of Auckland. This R-code [53] [54] was developed during a collaborative research effort between Ola Abdalla from Hokkaido University and Professor Cameron Walker from the University of Auckland. It is important to note that both **Hokkaido University and the University of Auckland have authorized these collaborative research activities to promote international cooperation**. Therefore, it is hoped that Hokkaido University will avoid any intellectual property and copyright infringement by duly recognizing and respecting Auckland University's contribution to this research project.

Fourth, Ola Abdalla used her private money, not the university's, to obtain copyrights or licenses for the cover photo, abstract photos, doctoral presentation photos, and some images in this dissertation.

Finally, Ola Abdalla thanks Dr. Kaushik Ragunathan for granting permission to modify and reuse a part of Figure 1-A from his study [30].