論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	WHOLE GENOME SEQUENCING-BASED MUTATION ANALYSIS OF HUMAN CELLS IN CULTURE
著者(和文)	MILAIEnkhbaatar
Author(English)	Enkhbaatar Milai
出典(和文)	学位:博士(理学), 学位授与機関:東京科学大学, 報告番号:甲第252号, 授与年月日:2025年3月26日, 学位の種別:課程博士, 審査員:松本 義久,塚原 剛彦,片渕 竜也,鷹尾 康一朗,山村 雅幸
Citation(English)	Degree:Doctor (Science), Conferring organization: Institute of Science Tokyo, Report number:甲第252号, Conferred date:2025/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	要約
Type(English)	Outline

WHOLE GENOME SEQUENCING-BASED MUTATION ANALYSIS OF HUMAN CELLS IN CULTURE

by:

Enkhbaatar MILAI

Academic Supervisor:

Yoshihisa MATSUMOTO, Ph.D.

Doctoral Thesis

Submitted to:

Department of Transdisciplinary Science and Engineering

School of Environment and Society

Institute of Science Tokyo

Tokyo, Japan

In Partial Fulfillment of the Requiring for the Degree of:

DOCTOR OF SCIENCE

2025

TABLE OF CONTENTS

ACKNOWLEDGEMENT

I am deeply grateful to my supervisor, Prof. Yoshihisa Matsumoto, for his invaluable guidance and unwavering support throughout my graduate education. His expert mentorship and enthusiasm for radiobiology have been a constant source of inspiration, keeping me engaged and motivated during my research journey. I am also profoundly thankful for his personal generosity, which greatly enriched my experience at Science Tokyo and made my time there truly memorable.

I extend my heartfelt gratitude to Assist.Prof. Mikio Shimada for his consistent willingness to provide guidance, support, and encouragement over the years. His availability and expertise have been instrumental in shaping the direction of my research. I am equally thankful to all the former and current members of the Matsumoto Laboratory for their assistance, stimulating discussions, and the enjoyable moments we shared.

A special note of appreciation goes to Ms.Tomoko Miyake, whose kindness extended beyond research assistance to making my life in Japan much easier and more fulfilling. I am also grateful to Ms. Dulguun Tsumbuukhuu, who helping me embark on my bioinformatics journey, as well as to Mr. Isao Yoda from the Cobalt Irradiation Facility for his critical support in irradiating my samples.

I am deeply indebted to Dr. Takashi Sugihara in the Institute of Environmental Science and Prof. Toshiro Matsuda and Prof. Hirokuni Yamanishi from Kindai University for their indispensable cooperation in conducting the irradiation experiments. I am also grateful to Prof. Yoichi Gondo in Tokai University and Osaka University for fruitful discussion leading to the conceptualization this research. Your expertise and support were vital to my work, and I truly value your contributions.

I would also like to thank my former colleagues and professors at the Nuclear Research Center, National University of Mongolia, for shaping my academic foundation. A special thanks to Prof. Norov Nanzad, whose guidance introduced me to the world of research and set me on this path.

To my lifelong brothers—our thirty years of friendship and unwavering trust have been a source of strength and encouragement. Your belief in me has been a driving force, and I am profoundly grateful for your constant support. Our bond is truly timeless.

I express my deepest gratitude to my parents, siblings, and their families, whose unconditional love and encouragement have sustained me throughout this journey. To my sons, Temuulen Enkhbaatar, Khantenger Enkhbaatar, and Guntengis Enkhbaatar, thank you for bringing endless joy, energy, and motivation into my life. Your presence has been a constant reminder of what truly matters and has given me the strength to persevere through the challenges of this journey.

Finally, to my beloved wife, Otgonjargal Galbadrakh, I cannot thank you enough for your unwavering support and continuous encouragement throughout my research and life's challenges. Your belief in me has been the cornerstone of this accomplishment. This achievement would not have been possible without your love, sacrifices, and steadfast presence by my side. You are my greatest strength, and I dedicate this milestone to you.

Thank you all.

ABSTRACT

Radiation-induced mutagenesis is a critical area of study, shedding light on the intricate mechanisms of DNA damage and repair under varying conditions of ionizing radiation. My doctoral thesis explores somatic mutation dynamics in human cell systems using single-cell culturing methods and whole-genome sequencing (WGS) to unravel the distinct mutational profiles driven by different radiation types, dose rates, and exposure durations. The work spans three major studies, each elucidating unique aspects of radiation-induced somatic mutagenesis.

In the first study, the impact of chronic low-dose gamma radiation on NB1RGB fibroblasts was assessed. Employing single-cell culturing and high-throughput sequencing, this work revealed subtle increases in oxidative damage-related base substitutions, such as C>A transversions. Despite these increases, the mutational spectra remained consistent with control samples, underscoring the effectiveness of base excision repair (BER) in mitigating low-dose radiation damage. These findings highlight the resilience of non-proliferative cells under low-dose exposure, while emphasizing the importance of sensitive detection techniques for subtle genomic alterations.

The second study focused on the comparative effects of acute X-ray and nuclear reactor (NR) mixed radiation on TK6 human lymphoblastoid cells. X-rays predominantly induced oxidative base damage, leading to transitions such as T>C and A>G substitutions, whereas NR radiation generated broader mutational spectra, including diverse indel types and sizes. The mixed radiation from NR exposure was linked to more complex DNA lesions, challenging repair systems and activating error-prone mechanisms like non-homologous end joining (NHEJ). Mutational signature analysis revealed distinct profiles, with X-rays associated with oxidative stress signatures and NR radiation exhibiting signatures indicative of replication stress and clustered damage. This study underscores the critical role of radiation quality in shaping DNA damage and repair dynamics.

The third study explored the cumulative effects of chronic low-dose gamma radiation on clonal evolution in TK6 cells. Primary and secondary clonal populations were analyzed to track mutation accumulation and adaptive responses over extended exposure periods. A dose-dependent increase in mutation burdens was observed, with secondary colonies showing evidence of clonal selection, favoring the elimination of deleterious single nucleotide polymorphisms (SNPs) while retaining

indels driven by error-prone repair pathways. Interestingly, while chronic radiation amplified mutation frequencies, it did not fundamentally alter the qualitative nature of the mutational landscape. The results highlight the role of clonal selection and repair pathway dynamics in shaping genomic outcomes under chronic low-dose irradiation.

Collectively, this thesis advances our understanding of how radiation-induced somatic mutations arise, emphasizing the interplay between radiation quality, dose rate, DNA damage, and repair mechanisms. By integrating single-cell culturing methods, stringent bioinformatics pipelines, and WGS, the studies establish a robust framework for investigating somatic mutagenesis at high resolution. Future research could build upon these findings by incorporating analyses of structural variants, copy number variations, and higher-order genomic rearrangements to provide a more comprehensive view of radiation-induced genomic instability. These insights not only deepen our knowledge of radiation biology but also refine computational and experimental approaches for studying mutational processes in response to environmental stressors.

Submitted manuscripts included in this thesis

- **Enkhbaatar Milai,** Dulguun Tsumbuukhuu, Takashi Sugihara, Mikio Shimada, Yoichi Gondo, Yoshihisa Matsumoto. Whole genome sequencing-based analysis of low dose rate radiation-induced mutations in normal fibroblast NB1RGB in culture. Proceedings of ICRP 2023 Symposium. (Accepted for publication 19 August 2024).

International conferences

- **Enkhbaatar Milai,** Dulguun Tsumbuukhuu, Takashi Sugihara, Mikio Shimada, Yoichi Gondo, Yoshihisa Matsumoto. Whole genome sequencing-based analysis of spontaneous and low dose rate radiation-induced mutations in human fibroblast NB1RGB in culture. Biological Effects and Applications of Radiation. Tokyo Japan. 2024
- **Enkhbaatar Milai,** Dulguun Tsumbuukhuu, Toshiro Matsuda, Hirokuni Yamanishi, Mikio Shimada, Yoichi Gondo, Yoshihisa Matsumoto. Whole genome sequencing analysis of X-ray and nuclear reactor irradiation-induced mutations in human lymphocyte TK6 in culture. Biological Effects and Applications of Radiation. Tokyo Japan. 2024
- **Enkhbaatar Milai,** Dulguun Tsumbuukhuu, Takashi Sugihara, Mikio Shimada, Yoichi Gondo, Yoshihisa Matsumoto. Whole genome sequencing-based analysis of spontaneous and low dose rate radiation-induced mutations in human fibroblast NB1RGB in culture. 7th International Symposium on the System of Radiological Protection. Tokyo Japan. 2023

Domestic conferences

- **Enkhbaatar Milai,** Dulguun Tsumbuukhuu, Toshiro Matsuda, Hirokuni Yamanishi, Mikio Shimada, Yoichi Gondo, Yoshihisa Matsumoto. Whole genome sequencing analysis of naturally occurring and radiation-induced mutations in human lymphocyte TK6 in culture. Japanese Radiation Research Society 66th annual meeting. AS3-1 November 2023. Tokyo. Japan.
- Tsumbuukhuu Dulguun, **Enkhbaatar Milai,** Takashi Sugihara, Mikio Shimada, Yoichi Gondo, Yoshihisa Matsumoto. “Whole Genome Sequencing and Mutation Analysis of Human Cell Clones in Culture” 23rd Sugahara-Ohnishi Memorial Symposium for the Sensitization of Cancer Treatment, Oral Presentation-03. February, 2023. Kyoto, Japan

- **Enkhbaatar Milai,** Takashi Sugihara, Mikio Shimada, Yoichi Gondo and Yoshihisa Matsumoto. “Whole-Genome Sequencing Analysis on Irradiated Cells.” The 21th Young Researchers and Engineers Conference, P-8. January, 2023. Tokyo, Japan
- **Enkhbaatar Milai,** Takashi Sugihara, Mikio Shimada Yoichi Gondo and Yoshihisa Matsumoto. “Whole Genome Sequencing-based mutation analysis in human fibroblast NB1RGB.” The 65th Annual Meeting of JRRS. PS1-17. September, 2022, Osaka, Japan

LIST OF ABBREVIATIONS

APE1: Apurinic/apyrimidinic endonuclease1

APTX: Aprataxin

BER: Base Excision Repair

BRCA: Breast Cancer

BWA: Burrows-Wheeler Aligner

CFTR: Cystic Fibrosis Transmembrane Conductance Regulator

CNV: Copy Number Variants

COSMIC: Catalogue of Somatic Mutation in Cancer

CPD: Cyclobutene Pyrimidine Dimers

CSA/B: Cockayne Syndrome Protein A and B

ddNTP: Dideoxynucleotide

DDR: DNA Damage Response

DNA: Deoxyribonucleic Acid

DMEM: Dulbecco's Modified Eagle Medium

DSB: Double-Strand Breaks

ERBB2: ERB-B2 Receptor Tyrosine Kinase 2

FBS: Fetal Bovine Serum

HGP: Human Genome Project

HR: Homologous Recombination

H2AX: H2A histone family member X

INDEL: Insertion and Deletion

IR: Ionizing Radiation

NB1RGB:

NHEJ: Non-Homologous End Joining

HS: High Sensitivity

LET: Linear Energy Transfer

LIG: Ligase

LQ: Linear Quadratic

MMEJ: Microhomology-Mediated End Joining

MMR: Mismatch Repair

NER: Nucleotide Excision Repair

PARP1: Poly ADP-ribose polymerase1

PMP22: Peripheral Myelin Protein 22

PNKP: Polynucleotide kinase 3′-phosphatase

POL: Polymerase

RNA: Ribonucleic acid

RPMI: Roswell Park Memorial Institute

ROS: Reactive Oxygen Species

SF: Survival Fraction

SBS: Single Base Substitutions

SDSA: Synthesis-Dependent Strand Annealing

SNP: Single Nucleotide Polymorphism

SSB: Single-Strand Breaks

SV: Structural Variants

TDP1: Tyrosyl-DNA Phosphodiesterase 1

TFIIH: Transcription initiation Factor II H

TOP1: DNA topoisomerase1

UV: Ultraviolet

VAF: Variant Allele Frequency

WGS: Whole Genome Sequencing

XPC: Xeroderma Pigmentosum, complementation group C

XRCC: X-ray Repair Cross-Complementing protein

53BP1: Tumor suppressor p53-binding protein1

LIST OF FIGURES

Chapter I

Chapter III

LIST OF TABLES

Chapter III

CHAPTER I: GENERAL INTRODUCTION

1.1 RADIOBIOLOGY

1.1.1 Types of radiation and mechanisms of interaction

Ionizing Radiation

Ionizing radiation is a high-energy type of radiation release by X-Ray machines, radioactive substances, and cosmic rays that enter Earth's atmosphere. This type of radiation has enough energy to remove electrons from atoms, leading to ion formation and potentially serious biological effects. While low-dose exposure is relatively harmless, higher doses can increase cellular activity and, over time, raise the risk of developing health issues, such as cancer (Tulchinsky et al. 2023, Hall and Giaccia 2019).

Types of Ionizing Radiation

Ionizing radiation includes high energy electromagnetic waves (like X-rays and γ-rays) and high-energy particles (such as α-particles, β-particles, neutrons, and protons). The energy in these forms of radiation is much higher than that in visible or UV light, which means they can directkly damage biological material, especially DNA, leading to mutations, cellular death, or cancer.

- Alpha Particle (α-particles): Alpha particles are heavy, positively charged particles. They pose health risks mainly when inhaled or ingested, as they cannot penetrate the skin. Alpha emissions primarily come from natural sources like uranium, radium, and certain mineral deposits. However, when alpha-emitting materials are disturbed throuhg mining or other prcesses, they can contaminate the air and water, increasing exposure risks.
- Beta Particle (β-particles): Beta particles are lighter, negatively charged particles that can penetrate deeper into tissues compared to alpha particles. They are commonly used in medical procedures like thyroid treatments and certain diagnostic tests. Beta particle exposure can result from nuclear accidents or medical sources. Like alpha particles, inhalation and ingestion of beta emitters can cause cellular damage and increase cancer risks over time.

- Gamma Rays and X-rays (X-rays and γ-rays): Gamma rays and X-rays are highly penetrating waves with no mass or charge. They can travel through the human body, impacting organs and tissues deep within. Gamma rays come from radioactive materials and are used in various applications, including cancer treatment, sterilization of medical equipment, and food preservation. While most gamma ray exposure is external, ingestion or inhalation of radioactive isotopes can also lead to internal exposure.
- Neutron: Neutron radiation is a high-energy, uncharged particle which is neutral and can penetrate deeply into materials, including human tissue, making them particularly harmful. Neutrons are indirectly ionizing, which means they do not ionize atoms themselves. Instead, they transfer energy to atomic nuclei, causing them to become ionized or to emit other forms of radiation (like alpha or gamma particles) as secondary effects.

Non-Ionizing Radiation

In contrast, non-ionizing radiation (such as ultraviolet [UV] light, microwaves, and radio waves) lacks the energy required to ionize atoms. Instead, it interacts with biological material primarily by exciting electrons in atoms, which raises them to higher energy levels without removing them. This process generates relatively low-energy interactions, causing changes in temperature or vibration in molecules but generally not inducing direct DNA damage. However, specific forms of non-ionizing radiation, such as UV light, can still lead to biological effects, like DNA modifications, through prolonged or intense exposure (Omer et al. 2010).

1.1.2 Interaction of radiation with cells

Cells exposed to radiation undergo a series of interactions and changes at various stages, ultimately leading to cell death or abnormal behavior. The interaction of ionizing radiation with cells can be divided into four distinct stages: physical, pre-chemical, chemical, and biological (IAEA 2010).

1. Physical Stage: In this initial stage, the kinetic energy of radiation is transferred to atoms or molecules within the cell, resulting in excitation and ionization. This energy transfer

disrupts molecular stability, leading to the formation of charged particles (ions) and excited molecules.

2. Pre-Chemical Stage: The ionization process generates free radicals due to the absorption of the kinetic energy released during the physical interactions. Free radicals are highly reactive species that can damage cellular components.
3. Chemical Stage: The formation of free radicals leads to chemical reactions that produce abnormal biomolecules. These abnormal structures can disrupt normal cellular functions, resulting in injuries that affect all levels of biological organization, from individual cells to entire organisms.
4. Biological Stage: The effects of radiation at this stage can manifest in various ways, including cell death, mutation, or abnormal cellular behavior. If radiation damage leads to oncogenesis, the onset of overt cancer may be delayed for an extended period as the damaged cells undergo uncontrolled proliferation.

Radiation can be classified into two types: directly ionizing and indirectly ionizing. Charged particles (such as alpha and beta particles) are predominantly directly ionizing, while electromagnetic radiation (like X-rays and γ-rays) is typically indirectly ionizing. When radiation interacts with biological material, it may interact directly with critical cellular targets – a process known as direct action.

Indirect action: Radiation often interacts with water molecules, which comprise approximately 80% of cellular content. When radiation interacts with water, it can ionize the water molecules, producing hydroxyl radicals. This process can be represented as follows:

$$H_2O \rightarrow H_2O^+ + e^-$$

The ionized water molecule can then react with another water molecule, leading to the formation of a highly reactive hydroxyl radical ($\cdot OH$):

$$H_2O^+ + H_2O \rightarrow H_3O^+ + \cdot OH$$

The hydroxyl radical, characterized by having nine electrons with one unpaired, is a potent free radical capable of diffusing short distances within the cell to reach and damage critical targets. If

cellular damage occurs, biological effects may not be immediately evident; instead, they may manifest hours to days later as the damaged cells attempt to divide. In cases of oncogenic damage, the development of overt cancer may take years to become apparent.

DNA and Chromosome Damage Induced by Radiation

DNA and chromosomes are primary targets for radiation because of their critical roles in cellular function and heredity. Ionizing radiation, which transfers high energy to biological molecules, can lead to diverse forms of damage in DNA and chromosomes, disrupting genetic integrity and potentially leading to severe cellular outcomes. Beyond single-strand and double-strand breaks, which compromise DNA stability, radiation exposure can cause chromosomal breaks and aberrations - significant structural damage where chromosome arms are fragmented or mis-joined, resulting in abnormalities like dicentrics and acentric fragments. Furthermore, chromosome exchanges and rearrangements occur when fragments from multiple chromosomes rejoin incorrectly, leading to translocations or inversions that disrupt gene function and elevate cancer risk. Radiation can also induce base modifications and cross-linking in DNA, which interfere with replication and transcription, promoting genomic instability and cellular dysfunction.

1.1.3 Radiation dose, exposure duration and biological effects

Understanding how radiation impacts biological systems requires considering both the dose and duration of exposure, as these factors significantly influence cellular and tissue-level outcomes. Radiation exposure can occur significantly influence cellular and tissue-level outcomes. Radiation exposure can occur in two primary forms:

- Acute Exposure: which is refer to a high dose of radiation absorbed in a short time, often from a single event, such as an accidental exposure or medical imaging involving ionizing radiation. High dose, short term exposures can cause immediate and severe cellular damage, often leading to cell death (necrosis or apoptosis) or irreversible DNA damage.
- Chronic Exposure: Chronic exposure involves lower doses of radiation over an extended period, such as continuous occupational exposure or background environmental radiation. Chronic low-dose exposure may not cause immediate cell

death but can lead to gradual accumulation of DNA damage. For over time, this can result may occur increased mutational burden or increased cancer risk.

Each exposure type interacts with cellular mechanisms differently, affecting the likelihood of immediate versus cumulative biological damage.

Additionally, the relationship between dose level and biological responses often follows specific dose-response patterns, which are crucial for predicting radiation risks, including tissue damage and potential long-term effects.

- Deterministic Effects: These effects occur above a certain threshold dose. For example, skin burns, radiation sickness, and organ-specific damage require doses beyond a specific threshold to manifest. Below this threshold, no immediate biological effect is observed. The threshold dose varies based on exposure type (acute or chronic) and the sensitivity of the tissue. High radiation doses can trigger deterministic effects rapidly, while chronic exposures rarely exceed these thresholds but can still lead to cumulative cellular changes.
- Stochastic Effects: These effects, including cancer and genetic mutations, have no clear threshold and are probabilistic. The likelihood of occurrence increases with dose, but the severity does not; hence, any dose is thought to carry some risk, though the probability is lower at low doses.

In addition, the biological effects of radiation are often modeled in terms of linear and non-linear dose-response relationships. Assumes that even the smallest dose has some risk of causing cancer or mutations, with risk increasing linearly with dose (IAEA 2010, Ministry of the Environment. Booklet to Provide Basic Information Regarding Health Effects of Radiation. 4th Edition).

1.1.4 Summary of Radiobiology

In summary, ionizing radiation, characterized by its high energy, can cause significant biological effect by directly or indirectly damaging cellular structures, especially DNA. Types of ionizing radiation include alpha, beta, gamma rays, X-rays, and neutrons, each interacting uniquely with biological tissue. Ionizing radiation impacts range from immediate effects like cellular damage do

delayed consequences such as mutations or cancer, depending on exposure dose and duration. Acute exposure at high doses can lead to rapid tissue damage, while chronic, low-dose exposure may result in gradual DNA damage and heightened risk. Radiation effects are categorized as deterministic, with dose thresholds triggering effects like burns or radiation sickness, and stochastic, whit no clear threshold, where any dose could elevate risk. By understanding the mechanisms of radiation interaction and dose-response relationships, we can better anticipate and mitigate its biological impact.

1.2 LONG-TERM LOW-DOSE IRRADIATION

The impact of long-term exposure, especially at low doses, is critical are of study in radiation biology, with significant implications for public health occupational safety, and environmental protection. Research on prolonged radiation exposure has revealed unique mutation patterns, DNA damage responses, and repair pathway activation that differ from those seen in acute, high-dose radiation scenarios. This section summarizes key findings from studies that have examined the genetic effects of extended or low-dose radiation exposure, focusing on the accumulation of mutations and the specific genetic responses that develop over time.

1.2.1 Effect of prolonged radiation exposure on cellular genetics

Research on the effects of prolonged radiation exposure has highlighted several important aspects of cellular response to sustained radiation stress, including the gradual accumulation of DNA mutations, structural variants, and the activation of stress response pathways. Unlike acute radiation exposure, which typically induces extensive DNA double-strand breaks (DSBs) and large-scale structural rearrangements, prolonged radiation exposure at low-to-moderate levels often results in subtler forms of genomic instability that accumulate gradually.

Studies have shown that cells exposed to prolonged radiation accumulate point mutations, small insertions and deletions (indels), and structural variants over time, although at a lower frequency compared to acute exposure scenarios. For example, long-term studies on animal models exposed to chronic gamma radiation have documented a gradual increase in point mutations and small indels in tissues with high cell turnover rates, such as the gastrointestinal and hematopoietic systems. This gradual accumulation suggests that sustained low-level DNA damage can eventually

overwhelm repair pathways, leading to an observable mutational burden even in non-cancerous tissues (Uchimura et al. 2015).

Cells exposed to prolonged radiation tend to activate repair pathways, such as homologous recombination (HR) and base excision repair (BER), at higher baseline levels compared to unexposed cells. Research has shown that this adaptive response can sometimes increase cellular resilience to DNA damage (Jeggo et al. 2006). However, prolonged exposure can also lead to repair pathway inefficiencies, as continuous DNA damage taxes the cellular repair machinery. Over time, this stress may contribute to repair inaccuracies, leading to small indels and mutations that would otherwise be corrected under lower stress conditions.

1.2.2 Studies on continuous low-dose radiation

Continuous low-dose radiation, es experienced in certain occupations or environmental conditions, has distinct genetic impacts compared to both acute and prolonged high-dose exposure.

Studies examining continuous low-dose exposure, such as those conducted on nuclear industry workers or populations living near high-radiation areas, have revealed that while mutation rates are lower than those seen with acute exposure, they are still significantly elevated compared to non-exposed populations. This finding suggests a cumulative dose-response relationship, where even low doses of radiation, when sustained over long periods, can lead to notable genetic alterations. Some research has proposed that continuous low-dose exposure may “sensitize” cells, making them more susceptible to DNA damage and less efficient at repair, potentially resulting in a higher mutation rate over time.

One of the primary mechanisms by which low-dose radiation induces genetic changes is through chronic oxidative stress. Low doses of radiation continuously generate reactive oxygen species (ROS), which cause oxidative damage to DNA bases and lead to base substitutions (C>A transversions). This continuous oxidative stress can result in cumulative DNA damage that eventually overwhelms repair mechanisms, particularly in tissues with limited regeneration capacities, such as the brain and heart. The study found that workers exposed to radiation had an excess relative risk of 0.97 per Sievert (Sv) for cancers other than leukemia, with a 95% confidence interval of 0.14 to 1.97. Although smoking could be a confounding factor, it likely does not fully

explain the increased risk. For leukemia (excluding chronic lymphocytic leukemia), the excess relative risk was 1.93 per Sv, with a range from less than 0 to 8.47. Based on these risk estimates, radiation exposure may be responsible for 1-2% of cancer deaths among workers in this group (Cardis et al. 2005). Prolonged oxidative damage has also been associated with age-related diseases, raising concerns about the long-term health impacts of continuous low-dose radiation exposure.

Continuous low-dose radiation exposure has been linked to altered DNA repair dynamics, with cells frequently activating error-prone repair pathways such as non-homologous end joining (NHEJ). Studies have shown that these cells often exhibit increased repair errors, resulting in small indels and structural variations that accumulate over time. These genetic alterations are particularly concerning in tissues with a high cancer susceptibility, such as the thyroid, where low-dose radiation exposure has been associated with elevated cancer risk. Research indicates that individuals exposed to low doses of radiation over extended periods may exhibit unique mutational signatures, which can serve as biomarkers for assessing cancer risk associated with chronic radiation exposure (Asaithamby et al. 2009).

In conclude, studies on prolonged and continuous low-dose radiation exposure reveal a range of genetic effects, from gradual mutation accumulation to persistent oxidative stress and altered repair dynamics. These findings suggest that long-term exposure, even at low doses, has significant implications for genetic integrity, potentially increasing cancer risk and contributing to genomic instability. Further research is essential to refine radiation exposure guidelines and to develop biomarkers that can monitor genetic damage in populations exposed to chronic radiation.

1.3 SHORT-TERM LOW-DOSE IRRADIATION

1.3.1 Immediate genetic and cellular response to low-dose irradiation

Short-term low-dose irradiation is an area of study that focuses on the immediate and potentially reversible effects of short-term radiation exposure at low doses. While low-dose irradiation is less likely to cause extensive DNA double-strand breaks (DSBs), it dose produce single-strand breaks (SSBs) and base lesions, primarily through the generation of reactive oxygen species (ROS) (Sage et al. 2017).

Short-term low-dose radiation leads to a transient increase in ROS, causing oxidative damage to DNA bases. This oxidative stress primarily results in single-base lesions and mild strand breaks. Research has identified specific base substitutions, such as C>A and G>T transversions, as indicative of ROS-related DNA damage in irradiated cells. This immediate oxidative response is usually temporary and may not lead to permanent mutations if cellular repair mechanisms are efficient. However, in cells with compromised repair capacity, even these minor oxidative lesions could contribute to long-term genomic instability.

Cells exposed to short-term low-dose radiation activate DNA repair mechanisms, particularly base excision repair (BER) for oxidative base lesions and SSB repair pathways. Studies have found that low-dose irradiation triggers a rapid but transient activation of DNA repair proteins, such as XRCC1 and PARP1, which help resolve SSBs and base modifications. The cellular response is often proportional to the radiation dose, with low doses inducing minimal but measurable activation of repair pathways. This rapid repair response helps cells to recover from low-dose irradiation without accumulating significant DNA damage, suggesting that the impact of short-term exposure may be minimal in healthy cells with robust repair mechanisms (Lomax et al. 2013).

Interestingly, some studies have found that short-term low-dose irradiation may induce an adaptive response, where cells become temporarily more resistant to subsequent radiation exposure. This phenomenon, known as "radiation-induced hormesis," involves the activation of cellular stress responses that enhance DNA repair capacity and antioxidant defenses. For instance, low-dose irradiation has been shown to upregulate genes involved in detoxifying ROS and repairing damaged DNA, thereby reducing the likelihood of mutations in future exposures. However, this adaptive response varies between cell types and may not occur uniformly across different tissues.

1.3.2 Mutation induction and cancer risk in low-dose exposures

While short-term low-dose irradiation is generally considered to have minimal mutagenic effects, some studies suggest that it can still influence mutation rates, particularly in susceptible individuals or cell types. The risk of mutation and cancer from low-dose irradiation remains a topic of ongoing research and discussion.

Research has indicated that short-term low-dose irradiation can induce specific mutational signatures, though these are usually less pronounced than those seen in high-dose exposures. Studies have identified an increase in small deletions, single-base substitutions, and minor indels in cells exposed to low doses of radiation for brief periods. These mutational changes are often dose-dependent, with lower doses inducing fewer and more subtle mutations. Importantly, these mutational effects are usually temporary, as cells with effective DNA repair mechanisms can often resolve the majority of the damage without long-term consequences.

The link between short-term low-dose irradiation and cancer risk is still uncertain. Some epidemiological studies have suggested a possible increased risk of cancer with cumulative low-dose exposures, even from brief medical imaging procedures. However, the exact contribution of each low-dose exposure to cancer risk is difficult to quantify. For example, studies on medical workers exposed to frequent low-dose radiation from X-rays and CT scans have reported slightly elevated cancer rates over time, though these increases are generally small. Current research aims to establish dose thresholds below which the cancer risk is negligible, with findings indicating that single exposures at very low doses may not contribute significantly to long-term cancer risk in most individuals (Little et al. 2009).

Some cell types, such as hematopoietic and reproductive cells, are more sensitive to radiation-induced DNA damage, even at low doses. In these cell types, short-term low-dose radiation may have more pronounced effects, potentially leading to transient increases in mutation rates or epigenetic changes. Additionally, genetic differences in DNA repair capacity can make certain individuals more susceptible to the mutagenic effects of low-dose radiation, leading to a higher risk of radiation-induced mutations or even cancer. Identifying these susceptibility factors remains an important focus for personalized radiation safety guidelines.

In conclude, studies on short-term low-dose irradiation suggest that, while the genetic impact may be minimal compared to high-dose or prolonged exposure, it still induces measurable oxidative damage, activates DNA repair pathways, and may result in transient mutational signatures. The effects are generally temporary, with cells often recovering quickly due to efficient repair mechanisms. However, certain cell types and genetically predisposed individuals may experience heightened sensitivity, potentially leading to increased mutation rates or epigenetic changes.

Further research is needed to understand the long-term health implications of short-term low-dose irradiation, especially in cases of repeated exposure, such as medical imaging. Establishing safe exposure thresholds and identifying susceptibility factors will be essential for refining radiation safety standards and improving risk assessments for low-dose radiation exposure in both clinical and occupational settings.

1.4 BIOLOGY

1.4.1 Molecular biology

DNA

Molecular biology is the branch of science that seeks to understand the molecular basis of biological processes, with DNA (Deoxyribonucleic acid) as its fundamental element. DNA is the hereditary material in almost all living organisms, encoding the genetic information that directs growth, development, and cellular functions. This information is stored in the form of a code made up of four nucleotides: adenine (A), guanine (G), cytosine (C), and thymine (T). These nucleotides link together to form a DNA strand, with the nitrogenous bases extending inward from the backbone of the strand (Bruce et al., 2015). DNA bases pair with each other according to specific rules: adenine (A) pairs with thymine (T), and guanine (G) pairs with cytosine (C). Adenine and guanine are purines with a double-ring structure, while cytosine and thymine are pyrimidines with a single-ring structure. Together, these base pairs form units known as base pairs, as illustrated in Figure 1-1. These bases precise order, or sequence, determines the genetic information necessary for an organism's development and function. Each base is also bonded to a sugar molecule and a phosphate group, collectively forming a nucleotide. These nucleotides are arranged in two complementary strands that twist into a structure called a double helix. The strands are held together by weak hydrogen bonds between the paired bases—adenine with thymine, and guanine with cytosine. Although most DNA is located in the nucleus of the cell, small amounts can also be found in the mitochondria, the cell’s energy-producing organelle (Friedberg et al. 2006).

In the molecular mechanism, DNA serves not only as mechanisms for replication but also as the template for the synthesis of other essential molecular within the cell. This process known as gene expression, facilitates the production of two critical classes of polymers: RNAs (Ribonucleic acid)

and Proteins (Ministry of the Environment, Japan https://www.env.go.jp/en/chemi/rhm/basic-info/). Although RNA shares structural similarities with DNA, it differs by containing uracil (U) in place of thymine (T), while the other bases - adenine (A), cytosine (G), and guanine (G) – remain the same in both molecules. RNA, composed of the bases U, A, C and G, contrast with DNA, which is composed of T, A, C and G.

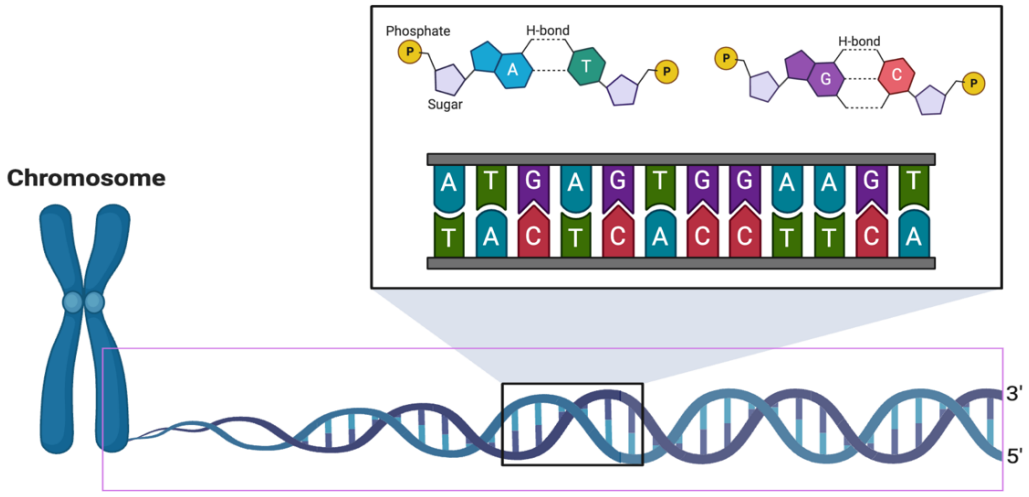

Figure 1.1 Structure of DNA base-pairs (created with Bioinder.com)

During transcription, RNA monomers are selected and polymerized based on the sequence of a template DNA strand, in a process similar to DNA replication. Within the DNA, specific sequences known as genes encode the instructions for synthesizing proteins, which are vital for the structure, function, and regulation of the body's tissues and organs. Like DNA and RNA, proteins are ling, unbranched polymer chains made up monomeric building blocks, which follow a standard arrangement across all living cells. These linear sequences in proteins store information, much like DNA and RNA, and are essential to cellular function. Protein molecules, due to their abundance and diversity, make up the majority of a cell's mass and are central to its operation (Ringo 2012)

Cell cycle

The primary function of the cell cycle is to accurately duplicate the vast amount of DNA contained within the chromosomes and ensure its proper segregation into two genetically identical daughter cells. The cell cycle is composed of four distinct phases G1, S, G2, and M. These phases

collectively define the two major stages of the cell cycle. In the G1phase, chromosomes are not yet replicated, and the cell has not committed to division. A critical checkpoint occurs at the G1-to—S phase transition to ensure proper conditions for DNA replication. Chromosome duplication takes place during the S phase (DNA synthesis), which lasts approximately 10-12 hours and constitutes about half of total cell cycle time in typical mammalian cell. Following DNA synthesis, the cell enters the G2 phase, during which it continues to grow and produce essential proteins. Finally, in the M phase, chromosome segregation and cell division (mitosis) occur, a process that takes around one hour [1] as illustrated in Figure 1-2.

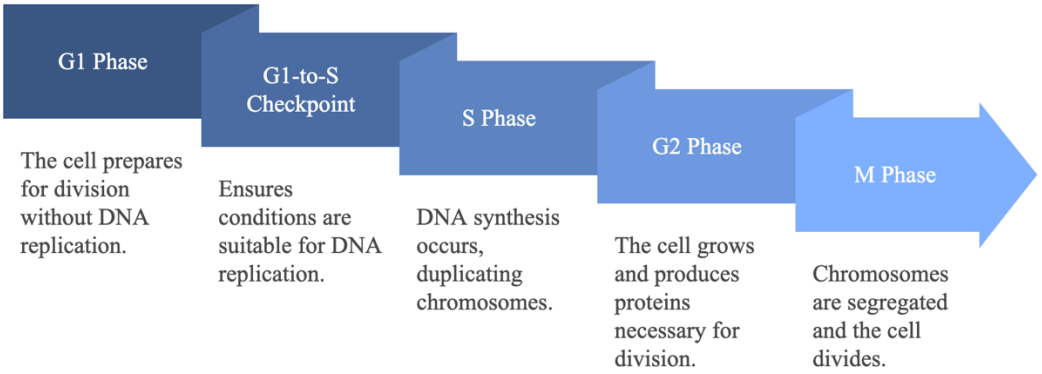

Figure 1.1 The process of eucaryotic cell (created with Bioinder.com)

DNA damage

DNA damage refers to any alteration in the DNA sequence or structure that disrupts it is normal function in processes like transcription, replication, or cell division. It is a critical factor in genetic diseases, aging and cancer development. While cells have evolved mechanisms to repair DNA damage, unrepaired or mis-repeated lesions can lead to mutations, genomic instability, or cell death as shown in Figure 1-3 (Patrick et al. 2023, Guo et al. 2023).

DNA damage can be broadly classified into two main classes based on its origin: endogenous and exogenous.

1. **Endogenous DNA damage**

a. Reactive Oxygen Species (ROS): By-products of cellular metabolism, ROS are highly reactive and can cause oxidative damage to DNA, leading to base modifications, strands breaks, and crosslinking (Guo et al., 2023).
b. Replication Errors: Mistakes during DNA replication, such as mismatched bases or slipped strand mispairing, can introduce mutations (Miguel et al. 2006).
c. Spontaneous Chemical Changes: Some damage arises naturally, such as the hydrolysis of DNA leading to the loss of bases or deamination (e.g., conversion of cytosine to uracil) (Nabel et al. 2012)

2. Exogenous Sources

a. Radiation: Ionizing radiation (like X-ray and gamma rays) and non-ionizing radiation (like UV light) can cause direct DNA damage. Ionizing radiation leads to double-strand breaks (DSBs) and single-strand breaks (SSBs) while UV light induces the formation of pyrimidine dimers (Borrego et al. 2015).
b. Chemical Agents: Exposure to mutagenic chemicals like alkylating agents, tobacco smoke, and certain industrial pollutants can induce DNA adducts or crosslinks, which is impair DNA function (Hang et al. 2010, Lin 2011).
c. Environmental Toxins: Exposure to environmental carcinogens, such as polycyclic aromatic hydrocarbons (PAHs) and aflatoxins, can modify DNA bases, leading to mutations (Moorthy et al. 2015).

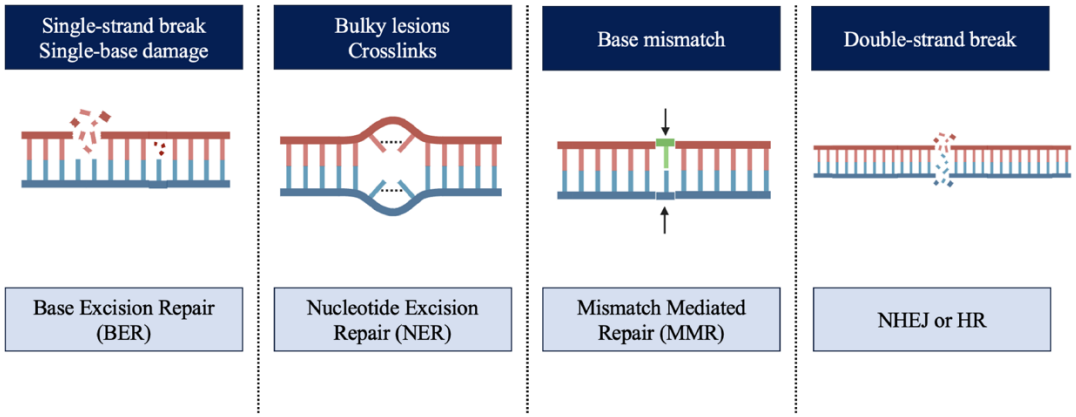

Figure 1.2 DNA damage and repair pathways (created with Bioinder.com)

DNA damage response (DDR): Following DNA damage, lesion-specific sensor proteins trigger the DNA damage response (DDR), a coordinated network of mechanisms responsible for detecting damage, signaling its presence, and facilitating repair (Harper et al. 2007). The recruitment of DDR factors occurs through a tightly regulated, spatiotemporal process, with these components assembling sequentially at the damage sites. Chromatin remodeling plays a crucial role in this process by enabling key post-translational modifications that facilitate the recruitment of specific DDR and repair proteins. Mutations in DDR components are associated with several cancer predisposition syndromes, emphasizing the importance of these pathways in maintaining genomic stability and preventing disease. Nevertheless, DNA repair pathways efficiently resolve most DNA lesions, mitigating the risk of mutations and ensuring the integrity of critical cellular processes such as replication and transcription. Without these repair mechanisms, persistent damage could lead to cellular senescence, apoptosis, or uncontrolled cell proliferation (Van Attikum et al. 2005).

DNA repair pathways

Cells experience more than a million DNA damage events daily due to normal metabolic activities, chemical exposure, and radiation (Tubbs et al. 2017). Among these, ionizing radiation poses a significant threat as it can induce various types of DNA lesions. Fortunately, cells have evolved robust DNA repair mechanisms to detect and correct these abnormalities, ensuring genomic stability. Of the different types of DNA damage, double-strand breaks (DSBs) are the most lethal to cells (Borrego et al. 2015).

DNA repair pathways mainly following categories, each responsible for addressing different forms of damage:

Base Excision Repair (BER): This pathway primarily addresses small-scale DNA lesions caused by oxidation, alkylation, deamination, and abasis sites. This pathway plays a vital role in maintaining genomic stability, especially during the G1 phase of the cell cycle. The BER process begins with chromatin remodeling around the lesion, allowing access to a family of DNA glycosylases that identify and remove damaged bases (Scharer et al. 2013).

BER can be proceed via short-patch or long-patch repair pathways, depending on the type of glycosylase activity and the lesion processed.

1. Short-Patch BER pathway:
    - The abasis site generated by monofunctional glycosylases is cleaved by APE1(apurinic/apyrimidinic endonuclease 1), at the $5'$ site, leaving a deoxyribose phosphate (dRP) group at the $3'$ end.
    - POLβ (DNA Polymerase β) performs both gap-tailoring (removing the dRP group) and single-nucleotide filling.
    - Ligation is completed by LIG1 or a LIG3-XRCC1 (Ligase - X-ray Repair Cross-Complementing protein) complex, depending on the cell context

2. Long-Patch BER pathway:
    - Bifunctional glycosylases create gaps that are processed by APE1, followed by strand-displacement synthesis by POLβ in non-dividing cells, or POLδ/ε in dividing cells.
    - A flap endonuclease (FEN1) removes the displaced strand before LIG1 seals the nick.

BER plays a crucial role in genome stability. Dysregulation or mutations in key BER proteins have been implicated in aging, neurodegenerative disease, and cancers.

Nucleotide Excision Repair (NER): Nucleotide excision repair (NER) is critical DNA repair mechanism that removes bulky DNA lesions caused by various environmental and endogenous factors. These lesions include cyclobutene pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4 PP) formed due to ultraviolet (UV) radiation, as well as bulky adducts resulting from exposure to chemical mutagens such as benzo[a]pyrene and certain chemotherapeutic agents (Scharer O.D et al., 2013). The accumulation of such DNA damage, if unrepaired, can lead to significant genomic instability, contributing to the development of disease like cancer.

NER operates through two primary pathways: Global Genome NER (GG-NER) and Transcription-Coupled NER (TC-NER).

1. Global Genome NER is responsible for scanning the entire genome for DNA damage, irrespective of whether the DNA is actively being transcribed. The initial recognition

of DNA lesions in GG-NER is mediated by the XPC (Xeroderma Pigmentosum, complementation group C) protein, which forms a complex with RAD23B and CETN2. This complex detects distortions in the DNA helix caused by bulky lesions and initiates the repair process. Once a lesion is recognized, the transcription initiation factor II H (TFIIH) complex is recruited to the site (Yokoi et al. 2000).

2. Transcription Coupled NER, on the other hand, specifically targets DNA lesions that stall RNA polymerase II during transcription. When RNA polymerase II encounters a lesion, it triggers the recruitment of TC-NER-specific proteins CSA (Cockayne Syndrome Protein A) and CSB (Cockayne Syndrome Protein B). These proteins facilitate the assembly of the TC-NER machinery, which overlaps with components of the GG-NER pathway to remove the obstruction and resume normal transcription. This pathway ensures that actively transcribed genes are promptly repaired, maintaining cellular function and preventing the accumulation of transcription-blocking lesions (Fousteri et al. 2006).

Nucleotide Excision Repair is indispensable for maintaining genomic integrity by effectively removing a wide range of bulky DNA lesions.

Mismatch Repair (MMR): MMR is an evolutionarily conserved, post replicative pathway that ensures high replication fidelity by correcting errors that arise during DNA synthesis, improving accuracy by at least 100-fold. MMR specifically targets base mismatches and insertion-deletion loops (IDLs) formed due to strand slippage in repetitive sequences. These errors, it left unrepaired, can lead to genome instability and microsatellite expansion, both of which are linked to various human disease (Kunkel et al. 2009).

In eukaryotic cells, mismatch recognition is primarily executed by two heterodimers: $MutS\alpha$ (MSH2/MSH6), which identifies single-base mismatches and small IDLs, and $MutS\beta$ (MSH2/MSH3), which detects larger IDLs (Kunkel et al. 2005). Recent studies have shown that MutL, another key component, can trap $MutS$ at the mismatch site, preventing premature disassembly of the repair complete. Once $MutS$ binds the mismatched region, the $MutS\alpha$

heterodimer (MLH1/PMS2) is recruited to stabilize the repair site and regulate the excision process (Nicolaides et al. 1994).

Single Strand Break Repair (SSB repair): SSBs are often generated by oxidative damage, abasic sites, or the erroneous activity of DNA topoisomerase1 (TOP1). Unresolved SSBs can cause replication collapse, transcription stalling, and activation of PARP1 (poly ADP-ribose polymerase1). PARP1 triggers the release of NAD+, ATP, and apoptosis-inducing factors, leading to cell death if not properly managed. Defects in SSBs repair are associated with genetic disorders such as spinocerebellar ataxia with axonal neuropathy1 and ataxia-oculomotor apraxia1, which are marked by genetic instability and increased cancer risk (El-Khamisy et al. 2005).

SSBs repair operates through three pathways based on the nature of the SSB:

1. Long Patch Repair Pathway: In this pathway, PARP1 identifies the SSBs and undergoes rapid cycle of poly (ADP) ribosylation before disengaging. The ends are processed by enzymes like APE1, PNKP (polynucleotide kinase $3'$ - phosphatase), and APTX (aprataxin), FEN1 removes damaged termini, and POLβ, POLδ, or POLε fill the gap. Finally, LIG1 seals the DNA with the assistance of PCNA and XRCC1.

2. Short Patch Repair Pathway: This pathway repairs smaller lesions, often arising during BER. APE1 processes the break, and the gap is filled exclusively by POLβ. The repair is completed by LIG3, assisted by XRCC1.
3. TOP1-SSBs Pathway: In this variant, the enzyme TDP1 (tyrosyl-DNA phosphodiesterase 1) removes trapped TOP1 from the DNA ends, facilitating proper repair.

Double Strand Break Repair (DSB repair): DSBs are the most dangerous form of DNA damages induced by various chemical and physical agents, such as radiation and certain chemicals. Unresolved DSBs can lead to severe genomic instability and are implicated in various human disorders and cancers (Pfeiffer et al. 2000). An initial response to DSBs: the cellular response to DSBs begins with chromatin modifications, which signal the presence of damage and trigger a cascade of events. Key early responders include the activation of ATM protein, targeted phosphorylation of histone H2AX, chromatin poly (ADP- ribosylation), and recruitment of

mediator proteins such as MDC1, 53BP1, and BRCA1 to the break site. Notably, 53BP1 and BRCA1 exhibit antagonistic roles in the repair process (Xie et al. 2007).

Cells have evolved two primary pathways to repair DSBs: Homologous Recombination (HR) and Non-Homologous End Joining (NHEJ).

1. Non-Homologous End Joining (NHEJ)

   The Non-Homologous End Joining (NHEJ) pathway is a critical mechanism for repairing DNA double-strand breaks (DSBs), a common and severe form of DNA damage induced by radiation exposure. In this pathway, the Ku heterodimer (Ku70/Ku80) rapidly binds to DSBs, preventing excessive DNA end resection and serving as a scaffold for recruiting other repair proteins. DNA-PKcs (DNA-dependent protein kinase catalytic subunit), highlighted in studies by Matsumoto et al. (2021, 2022), is a pivotal sensor in this process, structurally and functionally resembling Ataxia-Telangiectasia Mutated (ATM) kinase. DNA-PKcs interacts with the Ku heterodimer to stabilize DNA ends and promote downstream repair activities. Proteins such as Artemis, PNKP, and APLF process DNA ends to remove obstructive groups, allowing for gap filling by family X polymerases. The final ligation step is mediated by LIG4, assisted by XRCC4 and XLF, which restore DNA integrity. As discussed by Matsumoto et al. (2022), targeting DNA-PKcs has therapeutic potential in cancer treatment, underscoring its importance in both maintaining genomic stability and modulating the DNA damage response. This pathway's efficiency is particularly relevant to understanding mutation dynamics and DNA repair mechanisms in TK6 cells exposed to radiation, as explored in this study.
2. Homologous Recombination (HR)

   HR is characterized by its use of DNA strand invasion and template-directed repair synthesis, ensuring high-fidelity repair. The HR pathway includes several sub-pathways, notably the traditional DSB-induced HR, synthesis-dependent strand annealing (SDSA), and break-induced replication.

   The HR process is initiated by the MR complex (MRE11-RAD50-NBS1), which recognizes RM. Activated ATM phosphorylates H2AX, which serves as an anchor for MDC1 (Li et al. 2008, Altmeyer et al. 2013).

During the S/G2 phase, BRCA1 opposes the action of 53BP1 and initiates the ubiquitination of CtIP. The next step involves end resection, where nucleolytic degradation generates 3′ overhangs, committing the cell to the HR pathway. Initial resection is performed by the MRN complex and CtIP, followed by long-range resection mediated by EXO1 or BLM along with DNA.

Following end resection, RPA coats the 3′ overhangs, which are subsequently displaced by RAD51, forming a nucleoprotein filament. BRCA2 and PALB2 facilitate the filament formation, allowing it to invade a homologous DNA duplex and form a D-loop (Zhang et al. 2009).

Comparison of Homologous recombination and non-homologous end joining: In mammalian cells, homologous recombination (HR) is an error-free DNA repair pathway that requires a sister chromatid or homologous DNA sequence as a template for repair. However, HR is time-consuming due to the complexity of the mechanism, and it can only proceed efficiently during the S and G2 phases of the cell cycle, when a sister chromatid is available nearby. If homologous sequences are not present—such as during G1 phase—or if the DSB occurs in a region lacking accessible homologs, HR cannot proceed effectively.

In cases where the sister chromatid is present and physically close—such as during the S phase—HR becomes feasible. At this point, HR proteins promote strand invasion and template-based DNA synthesis to restore the broken DNA with high fidelity. However, when the cell is outside the S/G2 phases or when the sister chromatid is not accessible, the non-homologous end joining (NHEJ) pathway becomes the preferred option.

Unlike HR, NHEJ is faster and does not require a homologous template, making it the primary pathway for DSB repair during the G1 phase or under conditions where rapid repair is necessary to prevent genomic instability. However, NHEJ is more error-prone, as it directly ligates DNA ends, potentially leading to small insertions or deletions at the break site. This dynamic interplay between HR and NHEJ ensures that cells have options to maintain genomic integrity depending on the stage of the cell cycle and the nature of the damage.

1.4.2 Genomics Sequencing and its Mutations

Genomic Sequence Basics

The genome is organized into chromosomes, which are long DNA molecules that carry genetic information. In humans, there are 23 pairs of chromosomes (46 total), including one pair of sex chromosomes. Each chromosome is a compact structure made of DNA wrapped around proteins called histones, forming a complex called chromatin.

- Chromosomes ensures that DNA can fit within the cell nucleus and is properly segregated during cell division. Genes are segments of DNA on chromosomes that contain instructions for building proteins, which perform most cellular functions. The human genome contains approximately 20,000-25,00 genes.

- Genes are divided exons and introns. Exons are coding sequences, meaning they contain the instructions that are ultimately translated into proteins. Introns are non-coding sequences interspersed within genes. While they are transcribed into RNA, they are spliced out during RNA processing.

- Non-coding Regions: The genome also including vast regions of non-coding DNA, which do not encode proteins but play crucial roles in gene regulation, chromosome structure, and genome stability.

The DNA sequence is composed of four nucleotide bases- adenine (A), thymine (T), cytosine (C), and guanine (G). These bases are arranged in specific sequences, with sets of three bases (codons) in coding regions that correspond to amino acids, forming the genetic instructions for protein synthesis (Bruce et al. 2015).

Genomic Variants

Althrough the human genome is largely shared amond indivuals, small genetic variations make each person unique and can significantly influence health, physical traits, and environmenal repsonses. Among these variations, mutations or variants can arise in different forms, each with

distinct effects on the genome and potential impacts on gene function (National Institutes of Health (US) 2007).

1. Single Nucleotide Polymorphisms (SNPs)

   SNPs are most common type of genetic variation, representing changes at single nucleotide positions in the DNA sequence. These may occur in both coding and non-coding regions (The 1000 Genomic Project Consortium., 2010).

   - Functional Impact: While many SNPs have no effect, some may alter gene expression or protein structure, potentially influencing traits or disease risk. For example, SNPs is regulatory regions may increase or decrease a gene's expression in response to environmental factors, impacting susceptibility to conditions like asthma or hear disease.

2. Small Insertions and Deletions (Indels)

   Indels represent additions or losses of a few nucleotides (usually 1 to 50 base pairs) within the genome. These can have varied consequences depending on their location:

   - Insertions: In coding regions, insertions can disrupt the reading frame, resulting in a frameshift that may produce nonfunctional or truncated proteins. In non-coding regions, they may alter regulatory elements, impacting gene expression
   - Deletions: Similarly, deletions can lead to frameshift mutations in coding sequences, potentially silencing gene function. When occurring in regulatory areas, deletions may impact gene activity, possibly leading to abnormal gene expression patterns involved in disease like cancer.

3. Copy Number Variants (CNVs)

   CNVs are larger genomic alterations where segments of DNA (often containing one or more genes) are duplicated or deleted. These can have significant implications for gene dosage and overall genome stability:

   - Gene Dosage Effects: CNVs increase or decrease the number of gene copies, affecting the expression levels of encoded proteins. In some cases, increased dosage can drive overexpression, as seen in certain cancers. For instance, duplications of

the ERBB2 gene are linked to breast cancer due to heightened protein levels that promote cell proliferation.

- Disease Susceptibility: CNVs can also impact susceptibility to neurodevelopmental disorders like schizophrenia and autism. Deletions at chromosome 22q11.2, for example, lead to 22q11.2. deletion syndrome, affecting multiple systems in the body due to the loss of several genes.

4. Structural Variants (SVs)

   SVs encompass larger changes in the genome's architecture, including duplications, translocations, inversions, and large deletions (Feuk et al. 2017). These variants can disrupt gene function or regulatory networks:

   - Duplications: Extra copies of genes can lead to overexpression, affecting cellular processes. For instance, duplication of the PMP22 gene cause Charcot-Marie-Tooth disease, a disorder affecting the peripheral nerves.
   - Translocations: These involve segments of DNA moving from one chromosome to another, potentially disrupting genes at breakpoints or leading to aberrant gene regulation. A classic example is the Philadelphia chromosome in chronic myeloid leukemia, where a translocation between chromosomes 9 and 22 creates a fusion gene, driving cancer progression.
   - Inversions: Reversed segments within a chromosome may affect gene regulation if breakpoints occur near regularity regions, even if they do not alter protein sequences. This can lead to misexpression or loss of gene function, contributing to developmental disorders.
   - Large Deletions: Significant sections of DNA, including multiple genes, can be lost, often resulting on severe effects. For instance, large deletions in the DMD gene result in Duchenne muscular dystrophy by preventing the production of the functional dystrophin protein needed for muscle stability.

Each type of genetic variation can contribute uniquely to biological diversity, disease susceptibility, and evolutionary adaption. Together, these variants form a complex network of genetic influences that impact human health, resilience, and individuality

Germline and Somatic Mutations

Germline Mutation: Germline mutation occur in reproductive cells, such as sperm and eggs, making them inheritable. This type of mutation is present from conception and is integrated into the DNA of every cell in the offspring, affecting their overall genetic makeup and potentially increasing susceptibility to certain disease throughout their lifetime (Milholland et al. 2017).

Mutation pathways:

- Hereditary Cancer Genes: BRCA1 and BRCA2 mutations are well-known examples of germline mutations associated with increased cancer risk. These genes are involved in DNA repair processes, and mutations hinder this repair function, allowing DNA damage to accumulate over time.
- Inherited Genetic Disorders: Germline mutations can also lead to other hereditary diseases. For instance, mutations in the CFTR gene cause cystic fibrosis, a disorder characterized by the production of thick mucus in the lungs and digestive system. Each cell in an effected individual's body carries the mutation, resulting in systemic symptoms from birth.

Germline mutations affect individuals from birth and impact every cell in their body, resulting in lifelong health implications. These mutations can also affect entire families and populations due to hereditary transmission. Certain populations, such as Ashkenazi Jewish population, have a higher prevalence of BRCA1 and BRCA2 mutations, increasing the risk of hereditary cancer within these groups due to genetic drift and founder effects.

Somatic Mutation: Somatic mutation occurs in non-reproductive cells, meaning they are not passed on to offspring. These mutations accumulate throughout an individual's lifetime, often due to environmental factors or errors in cell division. somatic mutations typically affect tissues derived from the mutated cell, which restricts their impact to specific body parts (Milholland et al. 2017)

Mutation pathways:

- Environmental triggers: Ultraviolet (UV) radiation, for instance, induces the formation the formation of thymine dimers, which distort the DNA structure, leading to mutation that may contribute to skin cancer. Carcinogens in cigarette smoke, such as benzo[a]pyrene, bind directly do DNA in lung cells, increasing the mutation rate and associated cancer risk.
- Cellular pathways affect: Somatic mutations often occur in genes regulating cell growth and apoptosis. For example, mutations in the tumor suppressor gene TP53 disrupt cell cycle control, allowing damaged cells to escape apoptosis and continue dividing unchecked. Oncogenes like KRAS, when mutated, promote cell proliferation by continuously signaling growth, even without external growth signals.

Somatic mutations can lead to localized diseases, such as cancer, without affecting the individual's offspring. While somatic mutations are confined to specific cells or tissues, environmental exposure to mutagens can increase the likelihood of similar mutations across individuals as shown in Figure 1.4.

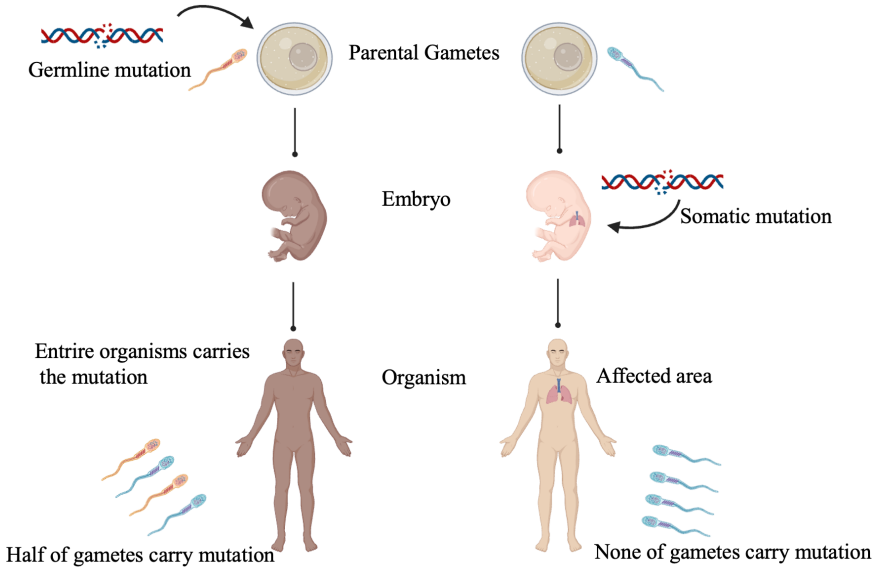

Figure 1.3 Differences between germline and somatic mutation (created with Bioinder.com)

Mutation Rate

Mutation rates represent the frequency at which changes occur in DNA sequences over time. These rates influenced by various factors, including DNA sequence composition, replication fidelity, and DNA repair efficiency. In general, regions of the genome with complex or repetitive sequences tend to have higher mutation rates, while highly conserved regions experience fewer mutations due to string selective pressures to maintain function.

Variable mutation rates across genomic regions;

- Repetitive DNA Sequences: Regions of DNA with repetitive sequences, such as microsatellites (short tandem repeats), are prone to slippage during DNA replication. This slippage increases the likelihood of insertions or deletions, leading to higher mutation rates in these areas. As a result, repetitive regions contribute significantly to genetic variability and can serve as markers for studying population diversity.
- Coding vs. Non-Coding Regions: Mutation rates differ between coding and non-coding regions. Non-coding regions, particularly introns and intergenic areas, can tolerate higher mutation rates since many mutations in these regions are less likely to disrupt essential functions. Conversely, coding regions (exons) experience lower mutation rates, as changes here may alter protein structure and function, leading to potentially harmful effects that natural selection tends to remove.
- Hotspot Regions: Certain regions within the genome, known as "mutation hotspots," experience mutations more frequently. These hotspots can occur in areas of high transcriptional activity where DNA is more exposed, or at sites where the DNA sequence is inherently unstable. Such regions can evolve more rapidly, contributing to localized diversity within populations

Higher mutation rates in non-coding and repetitive sequences contribute to genetic diversity without compromising essential functions. This diversity is crucial for evolution, enabling populations to adapt to environmental changes over time. On the other hand, lower mutation rates

in coding and highly conserved regions help preserve vital genetic functions, ensuring stability in fundamental biological processes (Satoh et al. 2023).

1.4.3 Summary biology

Molecular biology is rooted in understanding the essential role of DNA as the genetic blueprint guiding cellular functions, development, and heredity. The molecular structure of DNA, with its double-helix form and base-pairing rules, underpins this role by encoding the vast information needed for cellular processes and organismal growth. Through the tightly regulated cell cycle, cells duplicate their DNA, ensuring genetic continuity across generations, while complex repair pathways work to maintain genomic stability by addressing DNA damage from both endogenous and exogenous sources.

DNA damage and repair mechanisms, especially those addressing damage from oxidative stress, replication errors, and environmental mutagens, highlight the intricate cellular machinery evolved to protect genomic integrity. Key repair pathways like Base Excision Repair (BER), Nucleotide Excision Repair (NER), and Mismatch Repair (MMR) demonstrate specialized responses to varied types of DNA lesions, underscoring the importance of these mechanisms in preventing mutations, genetic diseases, and cancer. In particular, double-strand breaks, the most severe form of DNA damage, pose a unique challenge due to their potential to compromise cell viability if left unrepaired.

Altogether, the fidelity of DNA replication, the robustness of DNA repair systems, and the precision of the cell cycle underscore the cellular commitment to preserving genetic information. These mechanisms provide a foundation for understanding genetic diseases, aging processes, and cancer, and their study continues to drive innovations in research.

The study of genomic organization and mutations reveals the complexity and resilience of the human genome. Each level of genetic variation, from SNPs and indels to large structural changes, contributes to the individuality and adaptability of humans. While some mutations, particularly in coding regions, can have significant consequences, non-coding variations often provide a buffer that maintains essential functions while allowing genetic diversity.

Both germline and somatic mutations play distinct roles in health and disease. Germline mutations contribute to hereditary conditions and are passed across generations, while somatic mutations accumulate over an individual's life, frequently due to environmental factors, and can lead to localized diseases. Mutation rates vary across the genome, with repetitive and non-coding regions experiencing higher rates, promoting evolutionary flexibility, while conserved coding regions maintain essential cellular functions with lower mutation frequencies.

Together, the organization of the genome, the types of mutations, and the variable mutation rates across genomic regions demonstrate a balance between stability and adaptability. This balance allows for diversity among individuals and within populations while safeguarding critical biological processes essential for survival and reproduction.

1.5 WHOLE GENOME SEQUENCING

Whole Genome Sequencing (WGS) has revolutionized genomics by providing researchers with the capacity to examine the entire genetic blueprint of an organism at an unprecedented scale and resolution. Unlike traditional targeted sequencing methods that are confined to specific genes or genomic regions, WGS captures every nucleotide, offering a comprehensive view of both coding and non-coding regions (Goodwin et al. 2016). This approach enables a complete examination of an organism's genomic landscape, revealing insights into complex genetic variations, structural changes, and mutation patterns across the entire genome (Korbel et al. 2007, Zhang et al. 2009). The unbiased nature of WGS is instrumental in allowing scientists to identify genetic variants without presuppositions about mutation locations, making it particularly advantageous in studies of complex diseases, evolutionary genetics, and personalized medicine.

The extensive reach of WGS has been transformative in several fields. In cancer genomics, for instance, WGS reveals the full spectrum of germline or somatic mutations that drive cancer progression, including those in non-coding regions that may regulate gene expression (Pleasance et al. 2019, Alexandrov et al. 2013). In evolutionary biology, WGS enables researchers to trace lineage-specific variations, uncover adaptive mutations, and assess genetic diversity within populations. In personalized medicine, it provides insights into individual genetic predispositions, supporting risk assessment, disease prevention, and the development of personalized therapeutic strategies (Nielsen et al. 1977). Through these applications, WGS advances our understanding of genomic complexity and the intricate relationship between genetic variation and phenotypic diversity.

1.5.1 Early genomic sequencing efforts

In the early development of DNA sequencing, two main approaches achieved significant success: Sanger sequencing, introduced by Frederick Sanger in 1977, and Maxam-Gilbert sequencing, developed by Allan Maxam and Walter Gilbert in the same year. These pioneering methods revolutionized molecular biology by providing accurate techniques to decode DNA, despite their distinct approaches and inherent limitations.

Sanger Sequencing: The development of Sanger sequencing, introduced by Frederick Sanger in 1977, marked a groundbreaking advancement in molecular biology. Sanger's method, also known as chain-termination or dideoxy sequencing, became the first widely adopted DNA sequencing technique and transformed genetic research by enabling scientists to accurately sequence DNA for the first time. This method involves the selective incorporation of chain-terminating dideoxynucleotides (ddNTPs) during DNA replication. Each ddNTP stops the replication process at a specific base, resulting in DNA fragments of varying lengths that can be separated by size and read sequentially to determine the DNA sequence (Sanger et al., 1977).

The precision and reliability of Sanger sequencing established it as the standard for genetic research for over two decades. Its accuracy was fundamental to early genetic studies and allowed scientists to identify genes, map genomes, and develop an understanding of DNA's role in heredity and disease. Sanger sequencing became the cornerstone of molecular biology, widely used in research labs for sequencing small regions of DNA, individual genes, and even whole bacterial genomes (Ansorge et al. 2009). However, while revolutionary, Sanger sequencing was limited in terms of speed and scalability. Sequencing large genomes with Sanger's method required labor-intensive steps and was costly, time-consuming, and technically challenging.

These limitations became particularly evident as scientists began anticipate the possibility of sequencing the entire human genome. Despite its constraints, Sanger sequencing remained the method of choice due to its high accuracy, and it was the primary sequencing technology used in the Human Genome Project (Lander et al. 2001).

Maxam-Gilbert Sequencing: Developed by Allan Maxam and Walter Gilbert in 1977, Maxam-Gilbert sequencing was another pioneering technique in DNA sequencing history (Maxam and Gilbert 1977). This chemical-based method relied on the cleavage of DNA at specific bases using chemical reactions. DNA fragments of varying lengths were generated through these reactions and separated by electrophoresis to deduce the DNA sequence. The Maxam-Gilbert method was particularly useful for sequencing DNA fragments with complex secondary structures and provided a valuable alternative to Sanger sequencing in the early days of genomic research.

However, Maxam-Gilbert sequencing was less widely adopted due to its reliance on hazardous chemicals and the labor-intensive nature of the process. The technique required radioactive labeling of DNA and posed significant safety challenges, making it less practical for widespread use compared to Sanger's method. Despite these drawbacks, Maxam-Gilbert sequencing played a crucial role in advancing DNA sequencing methodologies and contributed to the foundational knowledge necessary for the development of modern sequencing technologies.

Human Genome Project: The Human Genome Project (HGP) launched in 1990 and completed in 2003 (Venter et al. 2001), was an ambitious international research attempt, aimed at mapping and sequencing the entire human genome. With the goal of identifying and cataloging the approximately three billion base pairs that constitute human DNA, the HGP represented a monumental milestone in biology and genetics. The project involved scientist from 20 research institutions worldwide including Japan (International Human Genome Sequencing Consortium., 2004). The HGP relied primarily on Sanger sequencing, and its successful completion in 2003 provided a reference human genome, transforming our understanding of genetic contributions to health and disease. The project's achievements were vast: it identified thousands of genes, revealed patterns of genetic variation across populations, and established a foundational database of human genetic information (Warburton et al. 2019).

However, the project underscored significant limitations in sequencing technology at the time. The cost of the HGP exceeded $3 billion, and it took 13 years to complete the sequencing process. The sheer time and expense involved revealed the need for more efficient, scalable sequencing methods, as sequencing large genomes with Sanger technology was neither practical nor sustainable on large scale. These limitations drove researchers and technology developers to innovate, leading to the emergence of Next- Generation Sequencing (NGS) technologies in the mid-2000s. NGS offered high-throughput sequencing capabilities at drastically reduced costs and timelines, paving the way for whole-genome sequencing to become accessible in both research and clinical contexts (Warburton et al. 2019).

1.5.2 Next-generation sequencing

The mid-2000s saw a pivotal shift in genome sequencing with the introduction of Next-Generation Sequencing (NGS), a technology that replaced the single-fragment sequencing approach of Sanger with a high-throughput model. NGS platforms enabled the simultaneous sequencing of millions of DNA fragments, reducing the time and cost needed to sequence entire genomes. This shift made Whole Genome Sequencing (WGS) feasible on a scale previously unattainable, allowing for large-scale genomic studies and diverse applications in research and clinical settings. Modern NGS platforms have evolved to include both short- and long-read sequencing technologies, each with unique capabilities. Short-read sequencing generates large numbers of highly accurate short DNA fragments, ideal for identifying single nucleotide polymorphisms (SNPs) and small insertions and deletions (indels). These reads are cost-effective and support high-throughput data production, making them suitable for large-scale studies. Long-read sequencing, in contrast, produces extended DNA sequences that span complex regions, improving the detection of structural variations (SVs) and copy number variations (CNVs). This ability to accurately read longer fragments has made long-read sequencing especially valuable in genome assembly, structural variant analysis, and studies of highly repetitive DNA (Goodwin et al. 2016).

The vast data produced by NGS requires advanced computational tools for alignment, variant calling, and data management. Modern alignment algorithms quickly map millions of reads to reference genomes, ensuring efficient and accurate positioning of each fragment (Li et al. 2009). Variant calling tools then identify SNPs, indels, and larger structural variants, using probabilistic models to distinguish true variations from sequencing errors. Additionally, improvements in data compression, cloud storage, and automated bioinformatics pipelines have made it possible to store, share, and analyze massive datasets, supporting collaborative and high-throughput studies (Schatz et al. 2010).

1.5.3 Advancements and future directions in WGS technologies

Recent Advances in Sequencing Technologies: Recent advancements in sequencing focus on third-generation, single-molecule technologies that capture DNA directly without amplification, improving both read length and accuracy, Technologies like long-read sequencing, exemplified by

platforms such as PacBio and Oxford Nanopore, generate sequences that can span thousands of base pairs, making it possible to detect SVs and conduct complete, de novo genome assembly. These capabilities address limitations in short-read sequencing, especially for complex regions with repetitive sequences or large structural rearrangements, thus enhancing applications in SV detection and reference genome creation (Schatz et al. 2009).

Single-cell sequencing and spatial genomics represent a major advancement in applying WGS at a cellular level. Single-cell sequencing enables analysis of individual cell genomes or transcriptomes, capturing genetic diversity within tissues-critical for understanding cellular heterogeneity in tumors or developmental processes (Wang et al. 2015). Spatial genomics further maps this genomic information to specific cellular locations within tissue samples, offering insights into cellular organization and gene expression dynamics within a physical context. This approach is particularly valuable in oncology for studying tumor microenvironments and in neuroscience for analyzing tissue architecture.

The immense datasets generated by modern sequencing require sophisticated computational approaches for effective analysis. Recent advancements in bioinformatics focus on enhancing the efficiency and accuracy of read alignment, variant calling, and data handling. Improved algorithms enable fast, memory-efficient alignment of both short- and long-read sequences, while variant calling tools use probabilistic models to accurately detect SNPs, indels, and larger structural changes. Additionally, the rise of cloud computing, automated pipelines, and data compression technologies has streamlined WGS data storage, processing, and collaboration, making large-scale sequencing projects more feasible.

Current and Future Directions: Efforts to reduce WGS costs continue, with innovations in sequencing chemistry, hardware, and analysis methods pushing towards making WGS more accessible and affordable. Precision improvements focus on reducing errors in sequencing reads, increasing accuracy in variant detection, and minimizing biases in data processing. These advancements are expected to drive WGS toward clinical ubiquity, supporting diagnostic precision and personalized medicine at a more accessible cost (Muir et al. 2016).

The future of WGS lies in its integration with other omics technologies-such as transcriptomics, proteomics, and epigenomics-forming a comprehensive, multi-dimensional view of biological systems. By combining WGS with transcriptome and protein data, researchers gain insights into how genetic variations impact cellular function. AI and machine learning further enhance this integration, providing tools to analyze complex interactions across omics layers, identify patterns, and generate predictive models. This multi-omics approach has broad applications, from identifying disease biomarkers to exploring genetic-environments interactions and advancing precision medicine.

1.6 APPLICATION IN MUTATIONS ANALYSIS

1.6.1 Somatic and germline mutation detection

Ionizing radiation (IR) differentially affects somatic and germline cells, leading to distinct mutational outcomes that are influenced by cellular repair mechanisms and environmental factors. Recent studies have leveraged WGS to systematically investigate these mutations, revealing unique patterns in both somatic and germline cells that reflect their respective biological roles and vulnerability to environmental damage (Kadhim et al. 2023).

Differences in mutation rates and spectra: Somatic and germline exhibit markedly different mutation rates in response to IR, with somatic cells generally accruing mutations at a higher rate due to their ongoing exposure to environmental stressors and limited repair fidelity. A study found that the somatic mutation rate in mouse and human cells is nearly two orders of magnitude higher than that in germline cells, underscoring the "protected" status of germline cells in maintaining genomic integrity for inheritance. The higher mutation rates in somatic cells align with findings of increased cancer risk due to accumulated mutations, especially in highly proliferative cell types (Milholland et al. 2017)

Radiation-induced mutation patterns in somatic cells: Somatic cells exposed to IR exhibit unique mutational signatures that are not typically observed in spontaneous mutations. Some result showed that mouse hematopoietic stem (HSCs) exposed to X-rays develop distinct base substitution patterns, such as C>A transversions, likely due to oxidative damage from reactive oxygen species (ROS) generated by IR (Matsuda et al. 2023). These ROS-induced mutations cluster within certain genomic regions, indicating that IR may cause localized DNA damage hotspots that somatic cells fail to repair adequately. Moreover, structural variants, such as large deletions, inversion, and complex rearrangements, are more frequency observed in irradiated somatic cells. These changes reflect the limitations of repair mechanisms, such as NHEJ, in somatic cells exposed to high levels of DNA damage, which increase the risk of tumorigenesis over time.

Recent work highlighted the complexity of IR-induced SV in somatic cells, documenting specific pattens like balanced inversion and chromothripsis (a complex chromosomal rearrangement) in

irradiated cells (Youk et al. 2024). These complex rearrangements are associated with catastrophic mutational event and can disrupt large genomic regions, thereby elevating cancer risk in affected tissues. This characteristics SV profile in irradiated somatic cells is distinct from the mutation spectra seen in germline cells, reinforcing the notion that IR exposure can lead to somatic mutations that drive disease processes, particularly malignancies.

Germline mutation patterns and mechanisms: In contrast to somatic cells, germline cells display a lower mutation rate and a different mutational response to IR, reflecting evolutionary pressure to preserve genomic integrity for future generations. Germline cells have robust DNA repair mechanisms that limit mutation accumulation, even under environmental stressors like radiation. When IR does induce germline mutations, these are often point mutations or small indels that occur less frequently than in somatic cells (Uchimura et al. 2015). The study showed that germline mutations can have cumulative effects over generations, suggesting that IR exposure may subtly alter the germline genome in ways that could impact long-term population health. Notably, germline mutations induced by IR are often enriched in non-coding regions or repetitive sequences, likely due to the distinct repair processes in germline cells. These subtle changes may contribute to heritable phenotypes without causing immediate harmful effects, a finding that has implications for evolutionary biology and genetic diversity in exposed populations.

Implications for radiation biology and long-term genetic stability: The differential effects of IR on somatic versus germline cells have significant implications for understanding genetic stability and mutation-drive diseases. Somatic mutations contribute to aging and cancer development, with IR-induced mutation signatures serving as potential biomarkers for radiation exposure in tissues. Conversely, germline mutations, though less frequent, introduce heritable genetic variations that can influence population health over generations. These findings underscore the need for distinct risk assessments and protective measures for somatic and germline exposure (Milholland et al. 2017). Somatic mutation signatures provide valuable tools for quantifying IR exposure and assessing cancer risk in exposed populations. In germline cells, preserving genomic integrity is crucial to minimizing transgenerational genetic risks associated with radiation exposure, particularly in areas with high environmental radiation levels or in occupations involving radiation.

1.6.2 SNPs and small indels

Single Nucleotide Polymorphisms (SNPs) and small insertions and deletions (indels) are fundamental forms of genetic variation that can have substantial impacts on gene expression, protein function, and genome stability. While SNPs involve a single base pair substitution, indels represent small additions or deletions of nucleotides within the DNA sequence. Radiation exposure is known to induce both types of mutations, contributing to genetic instability in somatic and germline cells.

Radiation-induced mutation patterns in SNPs and Indels: Studies have shown that radiation exposure to distinct mutational signatures, with specific base substitutions and small indels serving as markers of IR effects. For example, SNPs in irradiated mouse hematopoietic cells often present as C>A transversion, a substitution pattern indicative of oxidative stress caused by IR (Matsuda et al. 2023). Reactive oxygen species (ROS), which are generated during radiation exposure, cause oxidative damage to DNA bases, particularly guanine, leading to mutations that are both frequent and difficult to repair accurately.

Moreover, radiation-induced SNPs in human cells exposed to proton and alpha particle radiation (Delhomme et al. 2023). Their study demonstrated that these radiation types not only increase overall SNP frequency but also favor specific substitution patterns, such as G>T and C>A transversions, further supporting the role of oxidative stress in radiation mutagenesis. These unique base changes serve as mutational signatures that can be used to identify radiation exposure levels and assess the extent of DNA damage.

Radiation-induced indels and DNA repair mechanisms: Radiation-induced indels are primarily the result of error-prone repair processes that follow double-strand breaks (DSBs) caused by IR. Indels frequently occur when repair pathways, such as non-homologous end joining (NHEJ), attempt to re-ligate broken DNA ends without template guidance, leading to the loss or addition of nucleotides at the break site. A significant increase in small indels in human cell lines exposed to high-dose IR, particularly in regions where NHEJ was the dominant repair mechanism (Youk et al. 2023). This finding underscores the mutagenic potential of NHEJ, especially under

conditions of extensive DNA damage where the repair process may become overwhelmed or imprecise.

Indels are particularly impactful in coding regions where frameshift mutations can disrupt gene function (Montgomery et al. 2013). For example, radiation-induced indels in mutator mouse models often lead to frameshift mutation, resulting in loss of function in essential genes, which can have severe phenotypic consequences. These findings are significant for understanding radiation-induced pathologies, as indels in key regulatory or structural genes can drive cancer progression, cell death, or impaired tissue function in exposed individuals.

Detection and analysis of radiation-induced SNPs and Indels: The detection of radiation-specific SNP and indel patterns has been greatly facilitated by advances in WGS and bioinformatics tools. Modern WGS approaches enable high-resolution mapping of SNPs and indels across the genome, allowing researchers to detect even subtle increases in mutation rates and distinguish radiation-induced mutations from background mutation rates (Alexandrov et al. 2013). Bioinformatics tools for SNP and indel detection have advanced, allowing for precise variant calling and annotation even in high-damage contexts like radiation exposure. Algorithms designed to analyze mutation spectra can now distinguish between spontaneous mutations and those caused by environmental factors, offering insights into the mutation mechanisms activated by radiation. Studies have demonstrated that machine learning models trained on radiation-induced mutational data can enhance SNP and indel analysis by accurately classifying mutational patterns associated with radiation, potentially improving exposure assessments.

1.6.3 Structural variants and copy number variants

Structural Variants (SV) and Copy Number Variants (CNV) are substantial genomic alterations that affect gene dosage and regulatory regions, leading to significant biological and phenotypic consequences.

Radiation-induced SVs and mechanisms of DNA repair: Studies have shown that IR induced specific types of SVs, such as large deletions, duplications, and complex rearrangements, primarily through DNA double-strand breaks (DSBs) and subsequent repair pathways. Post-Chernobyl papillary thyroid carcinomas (PTCs) and observed that radiation exposure was associated with

increased rates clonal deletions and balanced SVs in tumor cells, largely repaired by NHEJ. This repair mechanism, while rapid, is error-prone, leading to SVs that drive oncogenic transformations. Their findings underscore that early DNA repair mechanisms, particularly NHEJ, are key contributors to the genomic instability observed in radiation-exposed populations, especially among those exposed at younger ages.

Specifically, proton and alpha particle exposure on various human cell lines have been investigated, which is reported that proton radiation increased the frequency of deletions, particularly in blood and lung cells line, further supporting the view that different radiation types produce distinct mutational footprints (Delhomme et al. 2023). This specificity has implications for assessing the risks associated with various radiation sources, including medical therapies and environmental exposures.

CNVs and Tumorigenesis in radiation-exposed cells: CNV, a subset of SVs, have been frequently linked to cancer progression, particularly through dosage alterations of oncogenes and tumor suppressor genes. Studies have demonstrated that radiation exposure can elevate CNV frequencies, contributing to tumor development (Yeager et al. 2021).

Radiation-induced CNVs have also been identified as contributors to increased risk in populations exposed to nuclear incidents. Contribution of CNVs in mutation-prone environments, mothing that radiation can exacerbate the mutation burden in tissues exposed to high levels of IR. CNVs affecting regions associated with cell-cycle regulation and DNA repair further emphasize the potential of radiation to drive oncogenic changes by altering gene dosage balance. This impact is significant in tissues like the thyroid, where exposure to radioactive iodine has been directly linked to increased cancer incidence, as observed in Chernobyl-exposed populations.

Detection of complex SVs with advanced WGS: The ability to detect and characterize SVs and CNVs has significantly improved with WGS, especially long-read sequencing platforms, which offer better resolution in complex or repetitive regions of the genome. WGS in irradiated mouse models and demonstrated how long-read technologies could capture complex rearrangements that are typically undetectable with short-read sequencing. Their study revealed high rates on non-repeat deletions and complex translocations in hematopoietic cells exposed to radiation,

illustrating how advanced WGS technologies can illuminate the full extent of radiation-induced genomic instability. Study about the mutational impacts of IR in human cells and identified SV patterns like chromothripsis- an extreme form of genome shattering and rearrangement in heavily irradiated samples. These finding indicate that high doses of IR can lead to catastrophic structural changes, which have been implicated in tumorigenesis. Such research underscores the necessity of using comprehensive sequencing approaches to understand the mutational landscapes induced by environmental mutagens and to identify potential biomarkers of radiation exposure (Youk et al. 2024).

In conclusion, the research demonstrates that SVs and CNVs induced by radiation exposure are closely linked to DNA damage response mechanisms, particularly NHEJ, which often leads to genomic rearrangements associated with tumorigenesis. Advanced WGS methodologies, such as long-read sequencing, have been instrumental in detecting these variants with high resolution, providing insights into the specific mutational footprints of radiation exposure. The cumulative evidence indicates that radiation-induced structural changes are somatically prevalent, with significant implications for cancer risk assessment, particularly in populations and professions exposed to high levels of IR.

1.6.4 The challenge and scope of WGS data analysis

Whole Genome Sequencing (WGS) has become an indispensable tool in genomics, offering comprehensive insights into genetic variation across entire genomes. However, the analysis of WGS data presents numerous challenges due to its complexity, size, and the need for accurate and scalable bioinformatics solutions. This section reviews the main challenges in WGS data analysis, alongside the scope of tools and methodologies developed to manage and interpret these vast datasets effectively (Muir et al. 2016).

- Data volume and Storage Requirements

  One of the most pressing challenges in WGS data analysis is managing the sheer volume of data produced. Sequencing a single human genome at high depth generates hundreds of gigabytes of raw data, including reads and quality scores. When scaled to large cohort studies, this data volume becomes difficult to store,

transfer, and process efficiently. Cloud-based platforms and data compression techniques have emerged to address these storage challenges. Solutions like the Genomics Data Commons and cloud computing services offer scalable storage and processing power, enabling researchers to access and analyze large datasets without the need for extensive local infrastructure.

- Computational Power and Processing Time

  The computational resources required for WGS analysis are considerable, particularly when aligning reads to a reference genome, detecting variants, and performing downstream analyses. Tasks such as variant calling, alignment, and structural variant detection involve computationally intensive algorithms, which can become bottlenecks in WGS workflows. High-performance computing (HPC) clusters, parallel processing, and GPU-based computing are commonly employed to reduce processing times. However, the accessibility of these resources can be limited, especially in smaller research institutions, presenting a barrier to large-scale WGS studies.
- Accuracy in variant calling and structural variant detection

  Accurate variant calling, particularly for rare mutations, small indels, and structural variants, is another significant challenge in WGS data analysis. Differences in sequencing technology, read length, and depth can influence variant detection accuracy. For example, short-read sequencing may miss complex structural variants (SVs) that are detectable with long-read technologies. Bioinformatics tools like GATK and Deep-Variant have been developed to improve the accuracy of variant calling across various types of genomic variations, from single nucleotide polymorphisms (SNPs) to large-scale rearrangements. However, achieving high sensitivity and specificity across all variant types remains a complex problem, requiring ongoing refinement of algorithms and machine learning models to address false positives and false negatives effectively (DePristo et al. 2011).
- Interpretation of variants and functional annotation

  Another substantial challenge lies in the interpretation of detected variants, especially in distinguishing pathogenic variants from benign polymorphisms. The

functional impact of many variants, particularly those in non-coding regions, remains unknown, creating a bottleneck in translating WGS data into meaningful biological insights. Variant annotation databases, such as dbSNP, and the 1000 Genomes Project, provide valuable references for variant interpretation. However, many rare and novel variants identified through WGS lack functional annotation, requiring experimental validation or in silico prediction models. The integration of machine learning tools for functional prediction has expanded the scope of WGS analysis by helping prioritize variants with potential clinical relevance, yet the interpretation of non-coding variants remains an area of active research (MacArthur et al. 2014).

The challenges of WGS data analysis highlight the need for continued advancements in computational resources, bioinformatics tools, and data interpretation frameworks. As the scope of WGS applications expands, so too does the demand for scalable, accurate, and secure data analysis methodologies.

1.7 SPECIFIC ON NB1RGB AND TK6 CELL LINES

NB1RGB and TK6 cell lines are widely used in radiation biology to study mutagenesis, DNA repair pathways, and dose-response relationships, particularly under low-dose and long-term radiation exposure. Their unique cellular characteristics make these models valuable for examining the impact of different radiation types—such as low-dose X-rays, gamma rays, and neutron radiation on DNA integrity, mutation accumulation, DNA variations and cellular aging.

The NB1RGB cell line, derived from human skin fibroblasts, is commonly used to study the effects of radiation on non-cancerous, normal cells, making it highly relevant for low-dose and long-term radiation exposure studies. Due to its fibroblast origin, NB1RGB is particularly valuable for exploring the effects of radiation on tissue types that typically have low proliferation rates. This attribute aligns well with studies examining cellular senescence and genomic stability under prolonged low-dose irradiation (Suzuki et al. 2003).

- DNA Repair Mechanisms in Low-Dose Radiation: Research on NB1RGB cells has shown strong activation of DNA repair pathways, such as non-homologous end

joining (NHEJ) and homologous recombination (HR), in response to low-dose gamma and X-ray exposure. While NHEJ is often employed as a quick but error-prone repair mechanism, HR provides a more accurate but slower response. Studies indicate that low-dose, long-term irradiation can cause cumulative DNA damage in NB1RGB cells, which may gradually impair HR and increase reliance on NHEJ, leading to an elevated risk of mutations and chromosomal aberrations.

- Cellular Senescence and Tissue Aging: One of the key features of NB1RGB cells under low-dose radiation exposure is their tendency to undergo cellular senescence when exposed to prolonged or cumulative DNA damage. This response is crucial for studying long-term radiation effects, as senescence is associated with tissue aging and degeneration. The ability of NB1RGB cells to enter a senescent state highlights their potential as a model for investigating the aging effects of low-dose, chronic radiation exposure, which is relevant in medical imaging fields and environmental radiation exposure contexts.

TK6 cell line, derived from human lymphoblasts, is particularly sensitive to DNA damage, making it an ideal model for studying genotoxicity, especially under low-dose radiation. As a suspension cell line that closely resembles hematopoietic cells, TK6 is highly responsive to both low- and high-linear energy transfer (LET) radiation types, including gamma rays and neutron-mixed radiation. This sensitivity provides insights into radiation effects on lymphocytes and other hematopoietic cells, which are crucial in understanding radiation risks in blood-forming tissues (Amundson et al. 2001).

- Mutation and DNA repair response: Studies using TK6 cells have documented a high frequency of point mutations, small insertions and deletions (indels), and structural variants following exposure to low-dose radiation. These cells possess proficient DNA mismatch repair (MMR) and homologous recombination pathways, which enable them to efficiently repair radiation-induced DNA lesions under low-dose conditions. However, due to their high radiosensitivity, TK6 cells often exhibit significant DNA damage accumulation when exposed to long-term low-dose radiation, which can lead to the formation of complex structural variants and

increased mutation frequencies. This response is valuable for assessing the risk of radiation-induced mutations in sensitive cell types like lymphocytes, especially under conditions that mimic chronic, low-dose exposure.

- Apoptotic and cytotoxic response: TK6 cells are prone to apoptosis and other forms of programmed cell death when subjected to severe DNA damage, even at relatively low doses of radiation. This makes them particularly useful for analyzing dose-response relationships and cytotoxicity thresholds in radiation studies. The apoptotic sensitivity of TK6 cells provides insights into how low-dose radiation can affect immune cell populations, highlighting potential cytotoxic and immunosuppressive effects of prolonged exposure in sensitive tissues.

Using both NB1RGB and TK6 cells allows researchers to examine a wide range of responses to low-doses and long-term radiation. NB1RGB cells offer a model for studying radiation effects on non-proliferative, normal tissue cells, providing insights into tissue aging, genomic stability, and senescence under prolonged radiation exposure. TK6 cells, on the other hand, provide a highly sensitive model for studying radiation-induced mutations and DNA repair processes in proliferative, hematopoietic cells, allowing for detailed analysis of mutation accumulation and cytotoxicity in response to low-dose and neutron radiation.

1.8 RESEARCH OBJECTIVE AND THEIR SIGNIFICANCE

For this study, I will evaluate three primary aims, integrating single-cell culturing methods with whole-genome sequencing (WGS) to investigate radiation-induced somatic mutations. Each aim corresponds to specific chapters, focusing on understanding the mechanisms of mutagenesis under various radiation conditions.

Aim 1, Chapter III:

Investigate the mutation profiles in NB1RGB cells exposed to chronic low-dose gamma radiation using single-cell culturing methods.

Objectives:

i. Establish single-cell culturing techniques to isolate and propagate clonal populations of NB1RGB cells for subsequent genomic analyses.
ii. Develop computational pipelines to identify single nucleotide polymorphisms (SNPs) and small insertions/deletions (indels) from WGS data in NB1RGB cells subjected to low-dose gamma radiation.
iii. Analyze mutation frequencies and spectra, focusing on identifying DNA damage signatures.
iv. Explore the involvement of specific DNA repair pathways, particularly base excision repair (BER), in mitigating low-dose gamma radiation-induced damage

Expected Outcomes:

This study aims to establish a robust single-cell culturing approach to facilitate high-resolution mutational analyses. It is anticipated that low-dose gamma radiation will induce subtle but detectable increases in mutation frequencies, particularly oxidative damage-related mutations. The findings are expected to emphasize the role of DNA repair mechanisms, such as BER, in maintaining genomic stability under low-level irradiation.

Aim 2, Chapter IV:

Characterize the mutational mechanisms induced by acute X-ray and nuclear reactor (NR) radiation in TK6 cells using clonal populations.

Objectives:

i. Utilize single-cell culturing methods to establish clonal populations of TK6 cells exposed to acute X-ray and NR radiation for mutational analyses.
ii. Develop a comparative framework to analyze mutational profiles, including SNPs, indels, and mutational signatures, across X-ray- and NR-exposed TK6 cells.
iii. Differentiate the mutational mechanisms of X-rays and NR radiation.
iv. Identify radiation-specific mutational signatures linked to distinct DNA damage and repair pathways, such as oxidative stress and replication-associated mechanisms.

Expected Outcomes:

By applying single-cell culturing to propagate clonal populations, this study expects to reveal distinct mutational patterns induced by X-rays and NR radiation. X-rays are anticipated to induce DNA damage, whereas NR radiation is expected to produce a broader spectrum of mutations due to mixed radiation quality. These findings aim to elucidate the unique mechanisms of somatic mutagenesis associated with different radiation types.

Aim 3, Chapter V:

Examine the cumulative effects of chronic low-dose gamma radiation on mutational dynamics and clonal evolution in TK6 cells.

Objectives:

i. Conduct single-cell culturing to track mutation accumulation in primary and secondary clonal populations of TK6 cells exposed to low-dose gamma radiation.
ii. Assess dose-dependent mutation burdens and shifts in mutation types, focusing on differences between primary and secondary colonies.

iii. Investigate the role of error-prone repair mechanisms, such as microhomology-mediated end joining (MMEJ), in shaping the mutational landscape under chronic radiation exposure.
iv. Analyze clonal selection dynamics to determine how chronic radiation exposure drives the evolution of specific mutational profiles and genomic adaptations.

Expected Outcomes:

This aim anticipates identifying dose-dependent increases in mutation frequencies, with secondary colonies showing evidence of clonal selection and an increased reliance on error-prone repair pathways. While chronic low-dose radiation is not expected to introduce new mutation types, it is likely to amplify existing mutational processes, offering insights into how clonal evolution adapts to sustained genomic stress.

CHAPTER II: GENERAL MATERIALS AND METHODOLOGY

2.1 INTRODUCTION

The objective of this chapter is to present the core methodologies and protocols that underpin the experiments and analyses described in the subsequent chapters. While each experimental chapter (Chapters III, IV, and V) incorporates unique elements tailored to specific cell lines, irradiation conditions, or analytical focuses, they all rely on a common methodological foundation. By consolidating the general materials and methods here, repetition is minimized, and a coherent framework is established for understanding the specialized procedures presented later.

2.2 OVERVIEW OF CELL LINES AND GENERAL CULTURE CONDITIONS

This research involved human cell lines maintained under standard culture conditions to ensure consistency, reproducibility, and comparability across multiple experimental setups. Although the details of cell line selection, their unique doubling times, and particular media supplements may vary, all experiments adhered to the following general guidelines:

The irradiation procedure was carefully designed to ensure uniform exposure across all samples using a 96-well plate setup, with wells aligned for homogeneous radiation. Consistent conditions were maintained throughout, and only wells with robust clonal growth post-irradiation were selected for further analysis.

Each well was verified to contain a single cell before irradiation using limiting dilution techniques, ensuring colonies originated from individual cells. Surviving single cells expanded into colonies, harvested, and processed for downstream analyses, such as DNA extraction and whole-genome sequencing. This method enabled precise tracking of radiation-induced effects while minimizing variability.

2.2.1 Cell line maintenance

All cell lines were grown at 37℃ in a humidified atmosphere containing 5% CO_2. Cells were maintained in suitable growth media, typically consisting of a basal medium (e.g., Dulbecco's Modified Eagle Medium [DMEM] or Roswell Park Memorial Institute [RPMI] medium) supplemented with 10% fetal bovine serum (FBS), penicillin (100 U/mL), and streptomycin (100 $\mu g/mL$). In some cases, non-essential amino acids on DMEM or horse serum on RPMI were added,

as described in the relevant experimental chapters. These supplements supported robust cell growth and maintained cellular homeostasis throughout the experiments.

2.2.2 Subculturing and passage conditions

Cells were sub-cultured at 70–80% confluence to maintain exponential growth. For adherent cells (e.g., fibroblasts), a brief incubation with 0.05% w/v trypsin-EDTA facilitated detachment, while suspension cell lines (e.g., TK6) were harvested by gentle centrifugation. Uniform seeding densities were employed to minimize variations in growth kinetics and avoid density-dependent effects. By following standard aseptic techniques and consistent culture conditions, baseline cellular health was preserved across all experiments.

(Note: Specific details regarding cell lines used in Chapters III, IV, and V—such as their sources, growth characteristics, and any modifications to standard media—are provided within their respective chapters.)

2.3 GENERAL IRRADIATION SETUP

All irradiation procedures were conducted using precisely controlled low-dose or moderate-dose radiation sources. A key objective was to maintain stable and accurate dose rates, ensuring that the observed effects could be attributed to the cumulative dose delivered.

In this study, low-dose radiation effects are defined using both dose rate and cumulative dose thresholds. Low-dose exposure was characterized by radiation delivered at rates below 100 mGy/day (Feng et al. 2017), specifically 1 mGy/day and 20 mGy/day, which align with standard definitions in radiobiology. Over 21 days, cumulative doses of 21 mGy and 420 mGy were examined, remaining below thresholds for acute cytotoxicity or genomic instability. Low-dose effects were distinguished by subtle molecular changes, such as oxidative stress-related point mutations, without significant genomic instability or apoptosis, in contrast to high-dose effects that induce overt DNA damage and cell death. This was supported by whole-genome sequencing, showing baseline mutation rates with minor changes in mutation spectra under low-dose conditions, indicating efficient repair mechanisms.

2.3.1 Irradiation sources and facilities

A ^{137}Cs or ^{60}Co source was employed for gamma irradiation, housed within a dedicated facility to ensure stable low-dose-rate exposures. The irradiations were performed within specialized incubators or holders designed to maintain physiological conditions (37℃, 5% CO_2, appropriate humidity) throughout the irradiation period.

2.3.2 Dosimetry and dose verification

The radiation dose rates were established through calibration with dosimeters, verified periodically to ensure consistent exposure levels. Low-dose conditions (e.g., on the order of mGy/day) were maintained for extended periods, while separate experiments may have involved higher, acute doses (e.g., ~1 Gy). Each experimental chapter provides exact dose parameters and schedules relevant to that study.

By keeping these core conditions consistent, the differences observed in downstream analyses could be reliably linked to the specific dose rates and exposure durations presented in Chapters III, IV, and V.

2.4 GENERAL DNA EXTRACTION AND QUALITY CONTROL

High-quality genomic DNA was essential for accurate whole-genome sequencing and variant calling**.**

2.4.1 DNA extraction

Total genomic DNA was isolated from cultured cells using a commercial DNA extraction kit (e.g., NucleoSpin DNA extraction kit). Cells were pelleted by gentle centrifugation, resuspended in lysis buffer, and processed according to the manufacturer's instructions. The resulting DNA was stored at -30°C until library preparation.

2.4.2 DNA quality and quantity assessment

Purity of DNA was evaluated using a NanoDrop Spectrophotometer, ensuring $\dot{A}$260/280 and $\dot{A}$260/230 ratios within acceptable ranges. DNA concentration was measured using a Qubit High Sensitivity (HS) assay, providing accurate quantification for downstream WGS library preparation. Only samples meeting stringent purity and concentration thresholds were advanced to the sequencing stage.

2.5 GENERAL WHOLE-GENOME SEQUENCING (WGS) PROTOCOL

Whole-genome sequencing formed the backbone of the mutation detection process.

2.5.1 Library preparation

DNA libraries were constructed using a TruSeq DNA Nano Library Kit (Illumina). Genomic DNA was shared to a target insert size of approximately 400–500 bp. Following end-repair, A-tailing, and adapter ligation, libraries were PCR-amplified under minimal cycles to avoid amplification biases. Library quality and fragment size distribution were assessed using Bioanalyzer systems.

2.5.2 Sequencing parameters

Sequencing was performed on an Illumina HiSeq platform, generating 151 bp paired-end reads at approximately 30X coverage. This coverage level was chosen to balance genomic completeness

and cost-effectiveness, enabling reliable variant detection across the human genome. Any deviations or enhancements to coverage depth or read length are reported in the relevant experimental chapters.

2.6 GENERAL BIOINFORMATICS PIPELINE

A standardized computational workflow ensured consistency in data processing, variant calling, and initial filtering steps.

2.6.1 Read alignment and preprocessing

Raw FASTQ files were subjected to quality control checks (FastQC) before alignment to the human reference genome (GRCh38/hg38) using an established aligner (e.g., iSAAC and BWA-MEM). Duplicate reads were marked with Picard tools (http://broadinstitute.github.io/picard/), and base quality score recalibration was conducted using GATK's BaseRecalibrator to improve variant calling accuracy (https://gatk.broadinstitute.org/hc/en-us).

2.6.2 Variant calling and initial filtering

Variants were identified using state-of-the-art calling tools (e.g., Strelka or GATK HaplotypeCaller). Resulting VCF files underwent preliminary filtering based on mapping quality, depth, and known polymorphisms (dbSNP databases). Generic quality thresholds, such as QD < 2.0 or FS > 60.0, served as baseline criteria to remove low-confidence variants. Additional, experiment-specific filtering steps and variant classifications are provided in each experimental chapter.

2.6.3 Mutation frequency and spectrum analysis

General scripts (Python) were employed to compute mutation frequencies, rates, and classify SNVs by mutation type (e.g., transitions vs. transversions). Indels were categorized by size and type. This consistent approach allowed for direct comparisons across different experimental conditions, with chapter-specific adjustments described in their respective methods sections.

2.6.4 Signature profiling

Broadly, mutational signatures were derived by comparing observed mutation patterns to known COSMIC reference signatures using tools like SigProfiler (Alexandrov et al., 2020). This general method provided insight into underlying mutational processes. The exact signatures of interest and any unique analytical steps are detailed in the relevant chapters.

2.7 GENERAL STATISTICAL ANALYSIS

All statistical analyses were performed using recognized statistical tools and packages (Python, and SciPy). Standard thresholds ($p < 0.05$) defined statistical significance unless otherwise specified. Confidence intervals and p-values were computed for comparisons of mutation rates, variant counts, and other key metrics. Experiment-specific statistical tests or modeling approaches are indicated in Chapters III, IV, and V, as required by the specific hypotheses tested in each study.

In conclude, chapter II has outlined the fundamental materials and methodologies that underpin the research presented in this thesis. By establishing a uniform set of practices for cell culture, irradiation, DNA extraction, sequencing, bioinformatics processing, and data analysis, Chapter II provides a solid framework for the specialized experimental designs and comparative analyses that follow. The methods described here serve as reference points, ensuring that chapters dedicated to individual experimental conditions can focus on their unique aspects without redundant methodological descriptions.

CHAPTER III: MUTATION ANALYSIS OF HUMAN FIBROBLAST NB1RGB CELLS EXPOSED TO LOW DOSE RATE GAMMA IRRADIATION BY WHOLE GENOME SEQUENCING

3.1 INTRODUCTION

Gamma radiation, a high-energy form of electromagnetic radiation, poses a unique challenge to genomic stability due to its potential to cause long-term damage to DNA. While acute exposure to gamma radiation has been extensively studied, the biological effects of prolonged low-dose exposure remain less understood. Gamma radiation induces various types of DNA damage, including base modifications, single-strand breaks, and double-strand breaks. If not repaired accurately, these lesions can give rise to mutations such as single nucleotide polymorphisms (SNPs) and small insertions or deletions (indels), which may accumulate over time and alter the genomic landscape of affected cells.

A key question in radiation biology is whether long-term exposure to low doses of gamma radiation leads to distinct mutational patterns. Understanding these patterns is crucial not only for assessing radiation-associated risks but also for gaining insights into the underlying mechanisms of DNA damage and repair. However, despite decades of research, the mutational effects of low-dose, long-term gamma radiation exposure on human cells remain poorly characterized. The lack of definitive studies in this area has hindered the development of robust models for radiation risk assessment.

Advances in whole-genome sequencing (WGS) have provided powerful tools to study mutational processes with unprecedented resolution. Unlike earlier approaches limited to small genomic regions, WGS enables the comprehensive analysis of the entire genome (~6Gb), facilitating the identification of SNPs, small indels, and mutational signatures. These signatures, which represent the unique patterns of base substitutions associated with specific DNA damage and repair processes, have emerged as a critical tool for understanding the biological impact of environmental and chemical exposures, including radiation.

The primary objective of this study is to explore the DNA variations of low-dose gamma irradiation on normal human fibroblast NB1RGB cells. Utilizing whole-genome sequencing (WGS), this investigation seeks to quantify and characterize the mutation spectra induced by continuous low-dose irradiation. Specifically, the study aims to:

- Determine Mutation Rates and Profiles: Evaluate the frequency and types of mutations, including single nucleotide polymorphisms (SNPs) and small insertions/deletions (Indels), occurring in irradiated versus unirradiated cells.
- Identify Mutational Signatures: Analyze the spectra of mutations, focusing on characteristic patterns such as oxidative damage-related mutations (e.g., C>A transversions) and their correlation with irradiation.
- Establish a WGS-Based Pipeline for Radiation-Induced Mutation Analysis: Develop and apply a systematic filtering pipeline to identify de novo mutations by excluding germline and low-quality variants, ensuring robust and reproducible results.

This study aims to advance our comprehension of radiation-induced mutagenesis at low doses, offering a foundation for improved radiation risk assessment and potential interventions to mitigate genomic instability.

3.2 METHODS

3.2.1 Cell culture and irradiation

Normal human skin fibroblast cells (NB1RGB), characterized by a population doubling time of 15 hours, were sourced from the RIKEN Bio-Resource Center (Tsukuba, Ibaraki, Japan). These cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Nacalai Tesque, Kyoto, Japan), enriched with: 10% v/v fetal bovine serum (Hyclone, GE Healthcare, Chicago, IL, USA), 100 U/mL penicillin and 100 μg/mL streptomycin (Nacalai Tesque), and 1% v/v MEM Non-Essential Amino Acids Solution (100x) (Nacalai Tesque). The cells were maintained at 37°C in a humidified atmosphere containing 5% CO_2. For subculturing, cells were dissociated from culture dishes using 0.05% w/v Trypsin and 0.2 mM EDTA. Care was taken to ensure uniform seeding densities across experimental conditions to avoid confounding effects due to cell confluence or density-dependent growth changes.

For gamma irradiation experiments, the cells were exposed to low-dose-rate gamma rays generated from a ^{137}Cs source at the Institute of Environmental Science. Two exposure conditions were applied: 1 mGy/day and 20 mGy/day for 21 days, respectively.

Control cells were cultured in identical conditions but without gamma irradiation. Cells were irradiated within a CO_2 incubator designed for continuous low-dose-rate irradiation studies to ensure consistency in exposure and minimize environmental variation.

The survival of cells post-irradiation was evaluated using a Colony formation assay, where cells were plated at low densities, irradiated, and allowed to form colonies. Up to the 5Gy doses used in colony formation assays are typical for assessing cellular radiosensitivity and survival curves under high-dose, short-duration exposures (Polgar et al. 2022). Colonies consisting of ≥ 50 cells were scored, and the surviving fraction was calculated relative to the unirradiated controls. Radiosensitivity was determined by assessing the reduction in colony formation following exposure to gamma irradiation. All irradiations for survival fraction analysis were performed at the Chiyoda Technol Cobalt-60 Irradiation Facility at the Tokyo Institute of Technology.

3.2.2 DNA extraction and whole-genome sequencing

Genomic DNA was extracted from the cultured cells and clonal populations using the NucleoSpin DNA extraction kit (Takara Bio). The extracted DNA was quantified using both:

1. NanoDrop Spectrophotometer to assess purity based on $\dot{A}260/280$ and $\dot{A}260/230$ ratios, and
2. Qubit High Sensitivity (HS) Assay to measure DNA concentration accurately

To ensure DNA integrity, samples were stored at -30℃ until downstream processing.

Whole-genome sequencing (WGS) was performed at Kurabo Co., Ltd. (Neyagawa, Osaka, Japan). DNA library preparation was carried out using the TruSeq DNA Nano Library Kit (Illumina), which enables efficient and high-accuracy sequencing. Libraries were fragmented to a median insert size of 400-500bp. Paired-end sequencing was conducted on the Illumina HiSeq platform, producing 151 bp reads with an average genome coverage of >30X.

3.2.3 Genome alignment and variants calling

Raw sequencing reads were processed using standard bioinformatics pipelines. Paired-end reads were mapped to the human genome reference GRCh38 (hg38) using the iSAAC aligner

(v04.18.11.09), a tool optimized for high-throughput sequencing data. Reads were aligned to autosomes, excluding sex chromosomes, to prevent confounding by differences in copy number and variant allele frequency.

Variants, including single nucleotide polymorphisms (SNPs) and small insertions/deletions (indels), were called using Strelka (v2.9.10). Rigorous quality filtering was applied in three stages:

1. Germline Filter: Variants detected in the bulk culture were excluded from the clonal variants to remove pre-existing mutations.
2. First Quality Filter: Variants were removed if:
   - They appeared in more than 100 or fewer than 10 sequencing reads.
   - They were located within 10 bp of another variant (potential sequencing error hotspot).
   - They had a mapping quality score <20
3. Second Quality Filter: Remaining variants were filtered to retain:

- A variant allele frequency (VAF) between 0.35 and 0.65, reflecting the expected heterozygous frequency of de novo mutations.
- Variants with no significant strand bias ($p > 0.05$, Fisher's exact test).
- Variants absent from public databases of known polymorphisms (dbSNP138 and dbSNP154).

These filtering steps ensured that the final dataset represented high-confidence de novo mutations.

3.2.4 Mutation signature analysis

To investigate mutational processes associated with gamma radiation, we performed mutational signature analysis using a framework developed by the Alexandrov laboratory (https://github.com/AlexandrovLab/SigProfilerAssignment). This analysis focuses on deconstructing base substitution patterns into single base substitution (SBS) signatures, which reflect underlying DNA damage and repair mechanisms.

1. Input Data Preparation:

- High-confidence single nucleotide polymorphisms (SNPs) identified post-filtering were classified into one of six possible mutation types: C>A, C>G, C>T, T>A, T>C, and T>G.
- Each mutation type was further contextualized based on the trinucleotide sequence (5' and 3' bases flanking the mutated site), resulting in 96 possible mutation categories.

2. Signature Decomposition:
   - The frequency of each trinucleotide mutation was calculated for each sample, generating a mutational spectrum.
   - Mutational spectra were compared against the COSMIC (Catalogue of somatic mutations in cancer) reference database of known SBS signature
   - Signature deconvolution was performed using a non-negative matrix factorization (NMF) approach, as implemented in the SigProfiler framework developed by Alexandrov et al.

3. Signature Attribution:
   - Contributions of specific SBS signatures were quantified for each sample.
   - Special attention was given to signatures associated with oxidative damage (e.g., SBS18, indicative of ROS-mediated damage) and spontaneous deamination (e.g., SBS5), both of which are relevant to radiation-induced mutagenesis.
   - Comparative analysis was performed to identify differences in signature contributions between irradiated and unirradiated samples.

4. Visualization and Interpretation:
   - Mutational spectra and signature contributions were visualized using bar plots and cosine similarity heatmaps to assess the alignment of sample spectra with reference SBS signatures.
   - The results were interpreted in the context of known biological processes associated with low-dose gamma radiation.

3.2.5 Statistical Analysis

All statistical analyses were performed using R or Python-based tools. Differences in mutation frequencies, mutational spectra, and signatures between groups (e.g., irradiated versus unirradiated samples) were assessed. P-value of <0.05 was considered statistically significant.

3.3 RESULTS

3.3.1 Radiation sensitivity on the NB1RGB cell line

Investigate the effects of ionizing radiation (IR) on cell viability and the associated molecular mechanisms, first I utilized NB1RGB, normal human skin fibroblast. This model offers complementary insights, with NB1RGB cells serving as representatives of non-proliferative, fibroblast-like cells commonly found in normal tissue. The unique properties enable a comprehensive analysis of chronic versus acute radiation effects on cell growth, survival, and mutational outcomes (Miyake et al. 2019, Shimada et al. 2019, Shimada et al. 2023).

Growth Dynamics and Viability

NB1RGB cells exhibited predictable growth dynamics under normal culture conditions, as shown in Figure 3.1. These cells are non-proliferative in nature, making them an ideal model for studying long-term radiation exposure and the accumulation of DNA damage over time. Growth assays indicated steady proliferation until day 6, with a plateau phase observed by day 8, suggesting density-dependent growth inhibition in the absence of external stressors.

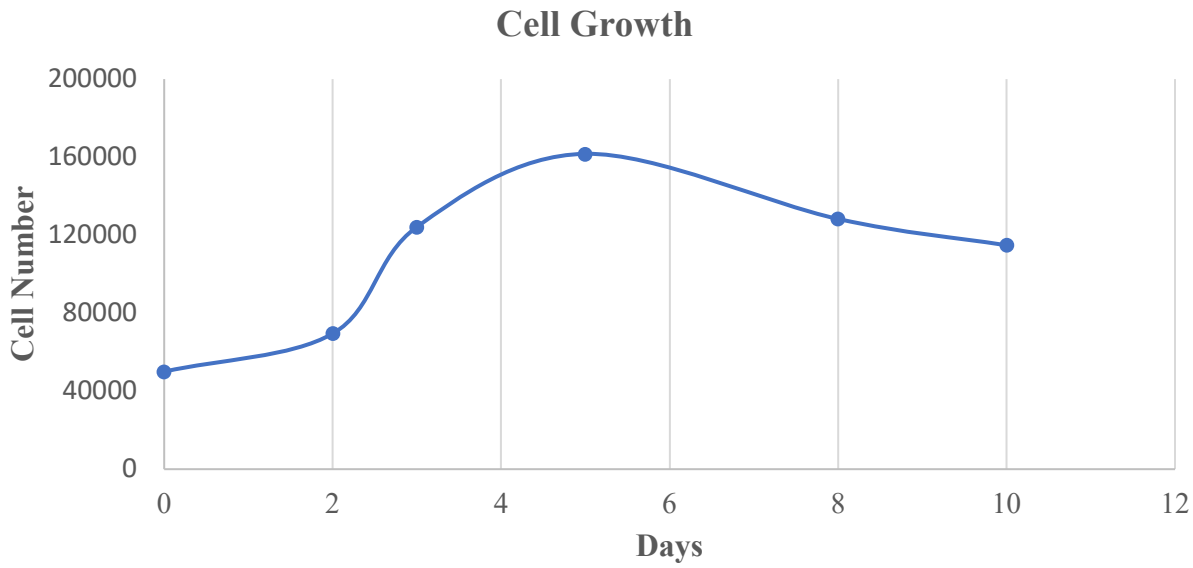

Figure 3.1 Cell growth in NB1RGB

Colony Formation and Radiosensitivity

To assess the radiosensitivity of NB1RGB cells, a colony formation assay was conducted following γ-irradiation. Under control conditions (0 Gy), the plating efficiency (PE) was determined to be approximately 50%, indicating a robust baseline ability to form colonies. Upon exposure to 2 Gy, the survival fraction (SF) decreased to 40%, reflecting moderate radiosensitivity. The data suggest that NB1RGB cells, while relatively resilient under normal conditions, experience significant viability reductions even at low-dose radiation exposure.

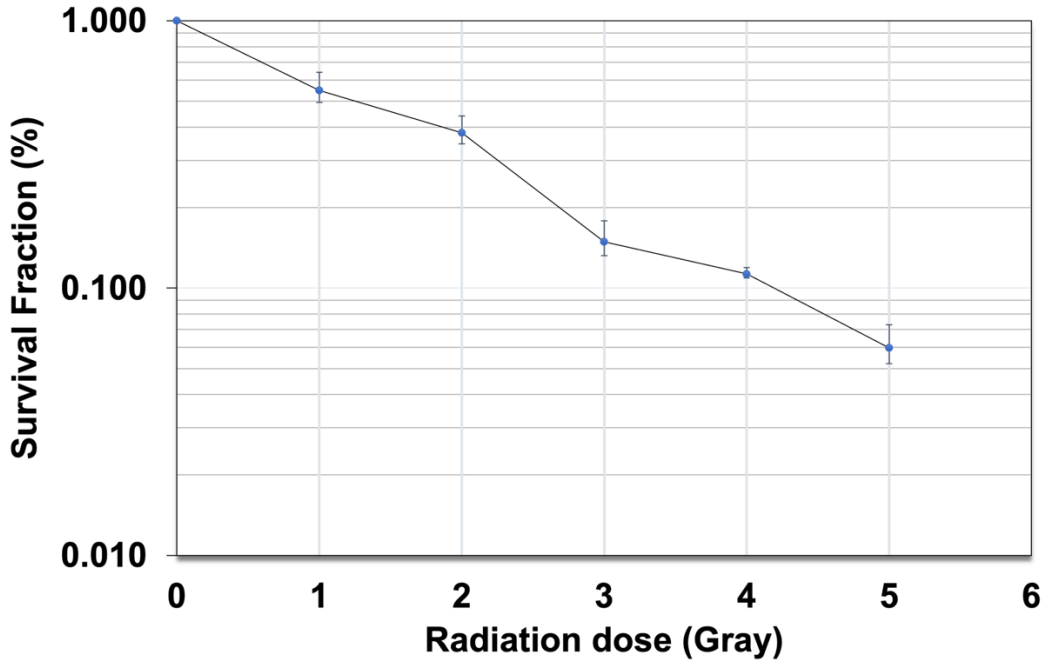

Figure 3.2 Radiation dose-survival curve of NB1RGB

The Linear-Quadratic (LQ) models is widely used to describe the relationship between cel survival fraction (SF) and ionizing radiation dose. It is particularly effective in capturing the effects of both low and high radiation doses on cellular viability, as it accounts for two key mechanisms of cell damage: linear (single-event damage) and quadratic (two-event damage).

The LQ model effectively describes how survival fraction decease with increasing doses, with the combined effects of single and double-event mechanisms becoming more apparent at higher doses. The graph in Figure 4.3 demonstrates two characteristic curves, one for closely ionizing radiation

and the other barely ionizing radiation. These differences reflect the distinct biological effects of different radiation types, emphasizing the importance of radiation quality when predicting cellular responses.

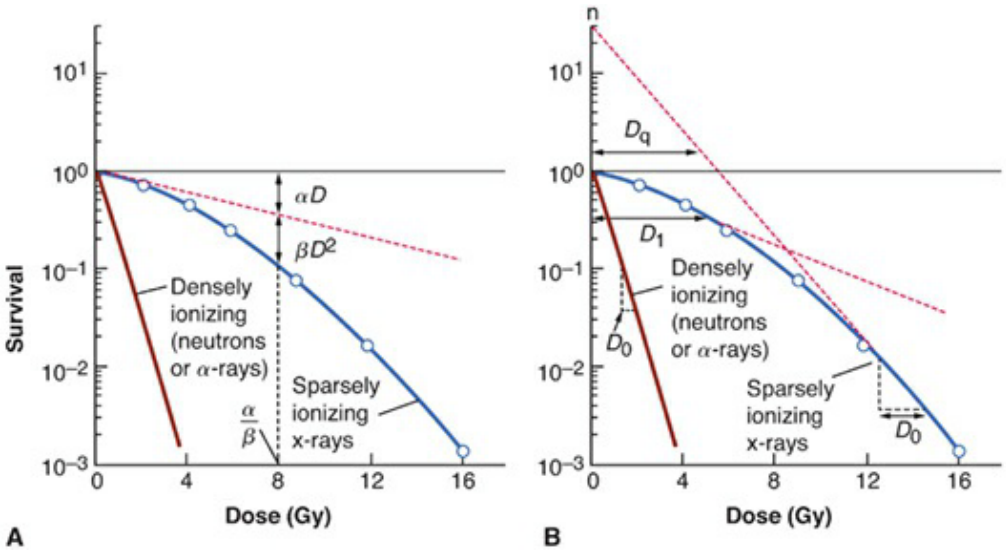

Figure 3.3 Shape of survival curve for cells exposed to radiation (Hall and Giaccia 2019)

Linear-Quadratic (LQ) model equation is

$$S(D) = e^{-\alpha D - \beta D^2}$$

- $S(D)$: The survival fraction of cells after radiation dose D
- α: The linear component, representing cell damage proportional to the radiation dose. This is primarily due to single-event DNA damage such as direct strand breaks.
- β: The quadratic component, representing cell damage proportional to the square of the dose (D^2). This term accounts for damage resulting from two-event interactions, such as the cumulative effects of multiple DNA breaks in close proximity.

The survival curve of NB1RGB cells was modeled using both the Linear-Quadratic (LQ) model and the Multitarget model to analyze the dose-response relationship following ionizing radiation

exposure (Figure 3.4). The LQ model provided a strong fit to the experimental survival fraction (SF) data, which were plotted on a logarithmic scale. At low doses (≤2 Gy), the SF decreased gradually, reflecting effective DNA repair mechanisms capable of resolving single-strand breaks (SSBs) and a proportion of double-strand breaks (DSBs). These repair mechanisms, including base excision repair (BER) and homologous recombination (HR), maintained high cell viability at these doses. However, at higher doses (>2 Gy), the SF dropped sharply, consistent with the saturation of repair pathways and the accumulation of irreparable DNA damage. The dose-dependent nature of this decline highlights the radiosensitivity of NB1RGB cells, particularly at doses where the repair capacity becomes overwhelmed.

Using the Multitarget model, a complementary analysis revealed the presence of a small shoulder region in the survival curve, representing a threshold dose below which sublethal damage was efficiently repaired without immediate cell death. This shoulder region, indicative of NB1RGB cells' ability to repair sublethal DNA damage, diminished as the dose increased, resulting in a steep decline in survival at higher doses. The comparison between the LQ and Multitarget models underscores the limited repair capacity of NB1RGB cells, particularly under sparsely ionizing radiation, where sublethal damage accumulates progressively. The multitarget model further highlights that, beyond the repair threshold, cellular mortality is driven by irreparable DSBs and chromosomal instability, consistent with the biological outcomes of exposure to doses exceeding 2 Gy.

α/β Ratio: The calculated α/β ratio for NB1RGB cells indicates a moderate contribution of both linear (single-event) and quadratic (cumulative) damage mechanisms to cell death. The linear component (α) reflects the effects of direct radiation-induced DNA strand breaks, while the quadratic component (β) captures damage caused by two-event interactions, such as clustered breaks in closely spaced genomic regions. At lower doses, the survival curve shows evidence of repairable damage, aligning with the region observed in sparsely ionizing radiation curves. However, at higher doses, the quadratic component becomes dominant, leading to a steep logarithmic decline in survival fractions as repair mechanisms fail to keep pace with the accumulation of DNA lesions. This dose-dependent trend reinforces the radiosensitivity of NB1RGB cells and their reliance on intact repair mechanisms to mitigate genomic instability.

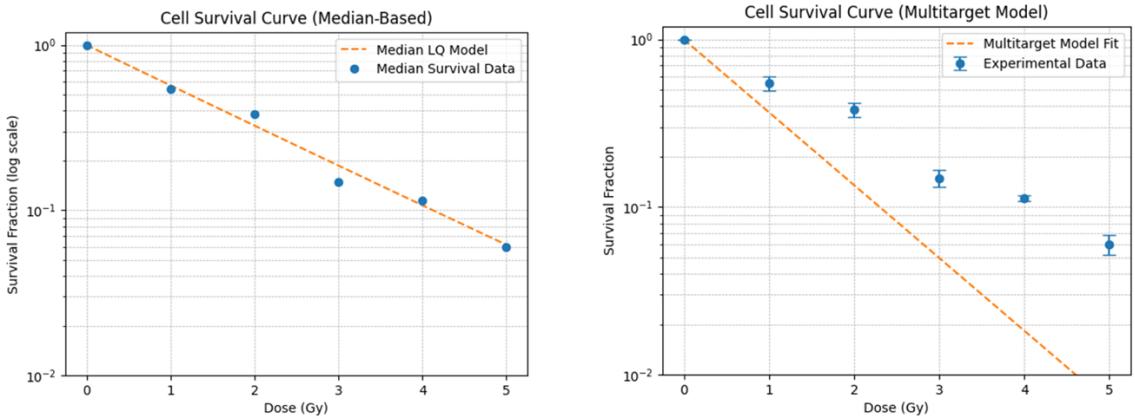

Figure 3.4 Calculation of shape of survival curve for cells exposed to radiation on NB1RGB

These results confirm the dose-dependent nature of radiation-induced cell death in NB1RGB cells, illustrating the interplay between repairable and cumulative damage and the saturation of cellular repair machinery at higher radiation doses.

3.3.2 Whole genome sequencing profile

The whole-genome sequencing (WGS) approach yielded an initial average of approximately 4.5 million single nucleotide variants (SNVs) and small insertions/deletions (indels) per sample. This result aligns with the estimated personal genetic variability in the human genome, which is approximately 0.1% of the total genome size (The 1000 Genomes Project Consortium, 2015). The extensive number of initial variants reflects both germline variation and sequencing artifacts, necessitating rigorous filtering to identify de novo mutations induced by low-dose gamma irradiation.

The filtering process consisted of three sequential steps:

1. Germline Filter: This step excluded variants present in the bulk culture, thereby removing pre-existing germline mutations shared across all clones. Applying this filter reduced the total number of variants by approximately 15,000-fold.
2. First Quality Filter: Additional filtering removed variants that were supported by too many (>100) or too few (<10) sequencing reads, variants located within 10 base pairs of others

(potential sequencing artifacts), and variants from reads with low mapping quality (<20). This step further reduced the number of calls by 2-fold.

3. Second Quality Filter: The final filter retained variants with a variant allele frequency (VAF) between 0.35 and 0.65, assuming heterozygosity for de novo mutations, and excluded those with significant strand bias ($p < 0.05$) or those registered in known single nucleotide polymorphism (SNP) databases (e.g., dbSNP138 and dbSNP154). This filter resulted in a 5-fold reduction in calls.

Following this pipeline, the average number of de novo substitutions and small indels detected per sample was 888 and 2,539, respectively (Table 3.1). Notably, the mutation rates were comparable between irradiated and unirradiated cultures, suggesting that the applied low-dose gamma irradiation (1 mGy/day or 20 mGy/day) did not significantly increase the frequency of de novo mutations under the studied conditions. The filtered variants were inspected using the Integrative Genomic Viewer (IGV) to confirm the mutation calls (Figure 3.5). IGV visualization of mutations observed in NT (Non-Treatment) and radiation-treated samples. The upper panel (black box) represents the IGV results for the NT sample, serving as a baseline. The lower panel shows IGV results for radiation-treated samples at 1 mGy/Day (a: small deletion, b: insertion) and 20 mGy/Day (c: substitution, d: deletion). Mutations are highlighted in the respective regions, with dose-dependent differences observed in mutation type and frequency.

Table 3.1 A descriptive summary of the variant calling pipeline

Sample	Initially Called SNVs	Germline Filter	1st Quality Filter	2nd Quality Filter	
				Substitutions	Indels
NT-1	4,556,200	31,718	15,356	797	2,474
NT-2	4,542,010	33,832	16,815	1,187	2,643
1 mGy/day_1	4,547,141	32,672	15,787	889	2,386
1 mGy/day_2	4,521,778	32,725	15,989	829	2,594
20 mGy/day_1	4,563,734	33,012	16,275	765	2,592
20 mGy/day_2	4,550,253	32,693	15,894	858	2,542

NT: Non-Treatment

To contextualize these results, the NB1RGB cells used in this study were cultured starting at a population doubling time of 15 hours. After two additional population doublings, the minimum number of cell divisions per clone was estimated at 17 hours, with an upper limit of approximately

50 hours, which corresponds to the general lifespan of human fibroblasts. Given the genome size of 2.88 × 10^9 base pairs (bp) for autosomes, the estimated mutation rate ranged between:

- 3.1×10^{-9} mutations per bp per division (lower limit), and
- 9.1×10^{-9} mutations per bp per division (upper limit)

These estimates are consistent with the somatic mutation rate of 2.66 × 10^{-9} mutations per bp per division previously reported by (Milholland et al., 2017), highlighting the validity of the WGS-based approach used in this study.

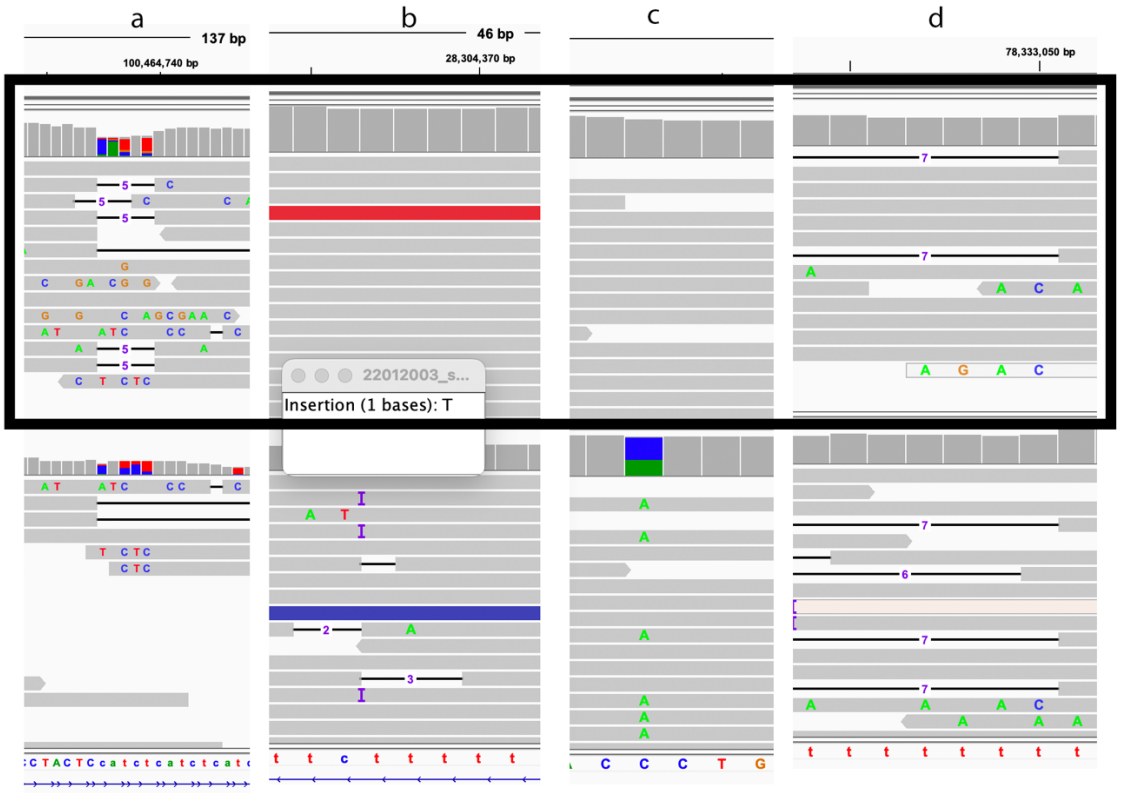

Figure 3 5 IGV visualization of mutations

The similarity in mutation rates and spectra between irradiated and unirradiated samples suggests that the cumulative radiation doses of 21 mGy (1 mGy/day) and 420 mGy (20 mGy/day) over 21 days may be insufficient to induce detectable differences in mutagenesis under the experimental conditions. Further analyses, such as focusing on specific mutational signatures or increasing

radiation exposure durations, may be required to uncover subtle effects of low-dose gamma irradiation.

3.3.3 Qualitative analysis of SNVs

The mutational spectra of single nucleotide variants (SNVs) were analyzed to compare the differences between irradiated and unirradiated clones. Across all samples, no discernible differences were observed in the mutation spectrum, indicating that low-dose gamma irradiation did not result in significant qualitative alterations in mutational patterns. In all clones, the C>A transversion, commonly associated with guanine oxidation, was the most frequent type of mutation. This was followed by T>C transitions, which are linked to pyrimidine dimer formation and other oxidative processes. Other types of substitutions, such as C>G, C>T, T>A, and T>G, were detected at lower frequencies, and their proportions were consistent across irradiated and unirradiated samples.

Figure 3.5 illustrates the mutation spectra for the different experimental conditions, highlighting the dominance of C>A transversions across all groups. Despite minor variations in the frequencies of specific substitution types, these variations were not statistically significant, further emphasizing the minimal qualitative impact of low-dose gamma irradiation under the studied conditions.

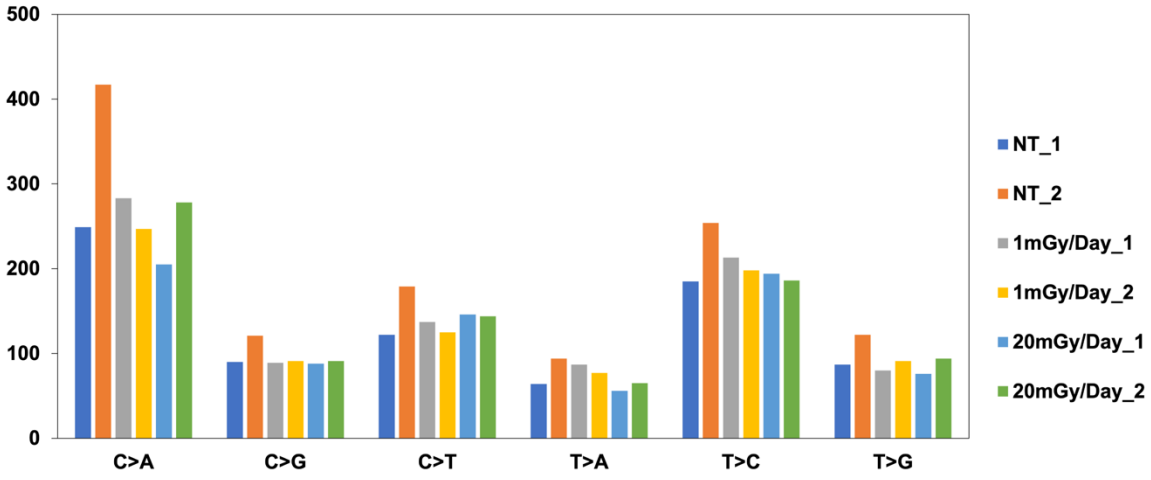

Figure 3.6 SNPs mutational signatures

To gain deeper insights into mutational processes, mutational signatures were analyzed based on established single base substitution (SBS) signatures. The results revealed the presence of SBS5, SBS36, and SBS40 across all samples, with varying proportions. SBS5 is a known clock-like signature associated with aging and environmental exposures, while SBS36 and SBS40 have been linked to oxidative damage and other mutational process.

Table 3.2 summarizes the relative contributions of each mutational signature for both unirradiated and irradiated samples. Although there were differences in the relative contents of SBS5, SBS36 and SBS49 among samples, no consistent pattern emerged between irradiated and unirradiated clones. For example:

- SBS5 dominated in some samples (e.g. NT-1 with 70.39%), while contribution was lower on others (e.g., 1 mGy/D-1 with 20.36%).
- SBS36 and SBS40 exhibited higher proportions in irradiated samples such as 1 mGy/D-2 and 20 mGy/D-1, but the variation was not statistically significant.

Table 3.2 Summary of mutational signature

Samples	SBS3 (%)	SBS5 (%)	SBS36 (%)	SBS40 (%)	Similarity
NT-1		70.39	29.61		0.92
NT-2		66.13	33.87		0.94
1 mGy/day_1		25.20	20.36	54.44	0.93
1 mGy/day_2		21.71	20.75	57.54	0.94
20 mGy/day_1	44.58	35.16	20.26		0.93
20 mGy/day_2		67.95	32.05		0.94

Despite these funding, the similarity scores across all samples (ranging from 0.92 to 0.94) indicate that the overall mutational signatures were highly comparable between conditions, further supporting the conclusion that low-dose rate gamma irradiation did not induce distinct qualitative changes in mutational process.

3.4 CHAPTER DISCUSSION

The findings from this study demonstrate that NB1RGB cells exhibit moderate radiosensitivity, as evidenced by the reduction in survival fraction (SF) with increasing gamma irradiation dose. The application of the Linear-Quadratic (LQ) model provided a nuanced understanding of dose-dependent cellular responses, revealing the interplay between single-event and cumulative damage mechanisms. The calculated α/β ratio underscores the balance between linear damage (e.g., direct strand breaks) and quadratic damage (e.g., cumulative or clustered breaks), aligning with established radiobiological models (Hall and Giaccia 2012). However, the sharp decline in SF at doses exceeding 2 Gy suggests a threshold at which repair mechanisms become overwhelmed, leading to the accumulation of irreparable double-strand breaks (DSBs) and eventual cell death. This finding is consistent with earlier studies reporting similar dose-dependent transitions in radiosensitivity in fibroblast-like cell models (Jeggo et al. 2011).

Interestingly, despite their non-proliferative nature, NB1RGB cells showed measurable radiosensitivity even at low-dose exposures, with significant reductions in viability at 1 mGy/day and 20 mGy/day. These findings suggest that low-dose gamma irradiation, while insufficient to induce acute cytotoxic effects, may still trigger subtle molecular and genomic changes. This underscores the importance of further examining low-dose exposure thresholds in non-proliferative human cells, particularly in the context of long-term environmental or occupational radiation exposure.

I used whole-genome sequencing (WGS) to analyze radiation-induced mutations at an unprecedented resolution comprehensively. The filtering pipeline effectively excluded germline variants, sequencing artifacts, and low-confidence calls, resulting in a high-confidence dataset of de novo mutations. The comparable mutation rates between irradiated and unirradiated samples highlight the limited impact of low-dose gamma irradiation (1 mGy/day or 20 mGy/day) on genomic stability under the studied conditions. This aligns with prior studies suggesting that doses below 100 mGy often fall below the threshold for detectable mutagenesis (Little et al. 2009).

The estimated mutation rates 3.1×10^{-9} to 9.1×10^{-9} (mutations per base pair per division) were consistent with baseline somatic mutation rates reported in human cells (Milholland et al.

2017). These rates reflect the efficacy of cellular repair mechanisms in mitigating the mutagenic potential of low-dose irradiation. However, the lack of significant differences in mutation rates between irradiated and control samples suggests that the cumulative doses (21 mGy and 420 mGy over 21 days) were insufficient to induce observable genomic instability. This finding raises important questions about the dose and duration thresholds required for low-dose radiation to produce measurable mutational effects.

Importantly, the study reaffirms the principle that low-dose gamma radiation, delivered at rates such as 1 mGy/day or 20 mGy/day, often falls below the thresholds required to induce measurable increases in genomic instability. This finding aligns with previous research indicating that low-dose radiation is generally well-tolerated by non-proliferative cells and tissues (Little et al. 2009). However, a low radiation dose did not increase variations in cells, but its opportunity to understand radiation's effects on human beings. Previous studies have similarly shown that low-dose exposures often fail to produce detectable mutagenesis in vitro, particularly in cell types with robust repair mechanisms (Hall and Giaccia 2012). Chronic irradiation over extended durations, or higher cumulative doses closer to occupational exposure limits, may reveal mutagenic effects that were not observable under the current experimental conditions.

Additionally, the study primarily focuses on fibroblast cells, which exhibit lower rates of replication-associated DNA damage compared to proliferative or cancer-prone cells. This raises the need to explore the effects of low-dose radiation in other cell types, including highly proliferative cells such as epithelial cells, hematopoietic stem cells, or lymphoblastoid cells. Proliferative cells, with their heightened vulnerability to replication stress and accumulated DNA damage, may respond differently to low-dose gamma radiation, resulting in distinct mutational patterns or rates (Little et al. 2009). Similarly, cancer-prone cells, characterized by inherent genomic instability and impaired repair mechanisms, may exhibit amplified sensitivity or unique mutational profiles under chronic low-dose exposure (Alexandrov et al. 2020).

Moreover, the study’s focus on gamma radiation limits its broader applicability to other types of ionizing radiation, such as alpha particles, beta particles, X-rays, or neutrons, which exhibit distinct linear energy transfer (LET) characteristics and biological effects. High-LET radiation, for instance, produces more clustered DNA damage, which is more challenging to repair and may lead

to greater genomic instability compared to the sparsely ionizing gamma rays used in this study (Ward et al. 1994). Expanding the scope of future research to include different radiation types would provide a more comprehensive understanding of how various radiation qualities influence DNA damage, repair mechanisms, and mutagenesis across different cell types and contexts.

Finally, the findings emphasize the importance of evaluating not only mutation rates and spectra but also broader biological outcomes such as epigenetic modifications, transcriptional changes, and clonal evolution in response to low-dose radiation. Similarly, the study of clonal dynamics in irradiated cell populations may provide insights into how radiation exposure influences cellular heterogeneity, selection pressures, and the emergence of adaptive or deleterious phenotypes over time (Nik-Zainal et al. 2016).

CHAPTER IV: MUTATION ANALYSIS OF HUMAN LYMPHOBLASTOID TK6 CELLS EXPOSED TO X-RAY AND NUCLEAR REACTOR RADIATION BY WHOLE GENOME SEQUENCING

4.1 ABSTRACT

Single cell-derived primary clones generated from human lymphoblast TK6 and the secondary clones were generated after exposure to X-ray and nuclear reactor radiation, which is a mixture of neutron and g-ray. Variants were called using GATK HaplotypeCaller, and somatic mutations, including single nucleotide variants (SNVs) and small insertions and deletions (Indels), were identified through stringent filtering criteria and variant allele frequency (VAF) thresholds to ensure accurate comparisons across samples.

Radiation exposure was associated with a dose-dependent increase in mutation, with distinct mutation spectra observed for X-ray and nuclear reactor irradiation. Base substitution patterns in SNVs suggested a significant role for reactive oxygen species (ROS) in radiation-induced mutagenesis. Single base substitution (SBS) signatures revealed radiation-specific mutation profiles, reflecting the differing biological impacts of X-ray and mixed radiation on TK6 cells.

This study provides valuable insights into radiation-induced mutagenesis, demonstrating the importance of variant filtering and VAF-based approaches for analyzing mutation dynamics. These findings contribute to the understanding of the distinct mutational mechanisms driven by low-dose radiation exposures.

CHAPTER V: MUTATION ANALYSIS OF HUMAN LYMPHOBLASTOID TK6 CELLS EXPOSED TO LOW DOSE RATE GAMMA RADIATION BY WHOLE GENOME SEQUENCING

5.1 ABSTRACT

This chapter investigates the mutational effects of low-dose-rate gamma radiation on human lymphoblastoid TK6 cells using whole-genome sequencing (WGS). TK6 cells were seeded into 96-well plates to establish single-cell-derived primary clones, which were cultured for 10 days under the same three conditions as described in Chapter 3: non-irradiation, 1mGy per day, and 20mGy per day. Following the culture period, primary clones were isolated, and genomic DNA was extracted from these clones, as well as from a non-clonal TK6 cell population.

WGS analysis was conducted to determine whether dose rate influences both the qualitative and quantitative characteristics of mutations. By comparing the mutation spectra and frequencies among the different exposure conditions, this study aimed to elucidate the potential dose-rate dependency of radiation-induced mutagenesis in TK6 cells.

CHAPTER VI: CONCLUSION AND PERSPECTIVES

6.1 Forward

The studies collectively provide a detailed examination of somatic mutations induced by different types and conditions of ionizing radiation. Through whole-genome sequencing (WGS) and robust bioinformatics pipelines, the results elucidate distinct mutational profiles driven by radiation type, dose rate, and exposure duration. These findings offer valuable insights into the mechanisms underlying radiation-induced somatic mutations, emphasizing the interplay between DNA damage, repair pathways, and clonal selection in shaping the genomic landscape of irradiated cells.

6.2 Radiation-induced mutational profiles in NB1RGB

The mutation analysis in NB1RGB cells exposed to chronic low-dose gamma radiation (Chapter III) revealed subtle but detectable increases in mutation frequencies, particularly oxidative stress-related base substitutions such as C>A transversions. These mutations, driven by reactive oxygen species (ROS), reflect the direct biochemical impact of gamma radiation on DNA bases. Despite these increases, the mutational spectra across irradiated and non-irradiated cells were remarkably similar, with no distinct signatures emerging in the context of low-dose exposure. This consistency highlights the robustness of DNA repair systems in mitigating the impact of low-level irradiation, particularly through pathways like base excision repair (BER) that efficiently handle oxidative damage.

The survival analysis, modeled using the Linear-Quadratic framework, further emphasized that while low-dose irradiation induced repairable DNA lesions, the threshold for significant genomic instability was not reached under the experimental conditions. The data suggest that NB1RGB fibroblasts, as non-proliferative cells, are relatively resistant to the accumulation of radiation-induced somatic mutations at low doses. This limited impact underscores the need for highly sensitive techniques and stringent filtering criteria to detect subtle mutational changes induced by chronic low-dose radiation

6.3 Broader implications for understanding somatic mutagenesis

In summary, this research establishes a novel approach for analyzing radiation-induced somatic mutations by integrating single-cell-derived clonal analysis with whole-genome sequencing

(WGS). By applying this methodology, successfully characterized mutation patterns under various radiation exposure conditions, providing new insights into the mutagenic effects of both low-dose-rate and acute radiation exposure.

6.4 Concluding perspectives and potential for deep analysis

The insights gained from these studies advance our understanding of how ionizing radiation induces somatic mutations, focusing primarily on small nucleotide changes such as SNPs and indels. However, a more comprehensive exploration of structural variants (SVs) and copy number variations (CNVs) could significantly enhance our understanding of radiation-induced genomic instability, offering a deeper view of the mutational landscape shaped by radiation exposure.

6.4.1 Deeper analysis of SNPs and Indels

While this study primarily examined SNPs and small indels, further analyses could leverage the following approaches to better characterize these mutations:

- Contextual Analysis of SNPs: Expanding the trinucleotide sequence context analysis to fully capture mutation biases induced by different radiation types. This could help identify whether specific DNA motifs are more susceptible to damage under certain radiation conditions.
- Indel Microhomology and Sequence Context: Investigating the prevalence of microhomology at indel breakpoints, especially in samples exposed to chronic low-dose gamma radiation, may clarify the role of error-prone repair pathways such as microhomology-mediated end joining (MMEJ).
- Mutation Clustering: Assessing whether SNPs and indels are spatially clustered within the genome or distributed randomly. This could reveal whether certain regions are more vulnerable to radiation-induced damage due to chromatin accessibility, replication timing, or DNA secondary structures.

6.4.2 Structural variants (SVs): A missing perspective

Structural variants, including deletions, duplications, inversions, and translocations, represent significant sources of genomic instability, particularly under high or chronic radiation exposure.

While this study focused on SNPs and indels, future work could incorporate SV analyses to capture larger genomic changes induced by radiation.

- Detection and Characterization of SVs: Using algorithms to detect SVs from WGS data could provide a comprehensive view of how radiation disrupts chromosomal architecture. For example, gamma radiation is known to generate double-strand breaks that may lead to complex SVs.
- Chromothripsis and Radiation: Radiation-induced chromothripsis, a phenomenon where chromosomes shatter and reassemble in a single catastrophic event, could be explored, especially in cells exposed to acute or high-dose radiation. SV analysis could uncover such patterns, which have implications for understanding large-scale genomic disruptions.
- Breakpoint Analysis: Mapping SV breakpoints at nucleotide resolution may help identify sequence motifs or genomic features that predispose regions to structural rearrangements under radiation stress.

6.4.3 Copy number variations (CNVs): A layer of complexity

Copy number variations, encompassing gains or losses of genomic regions, may represent a significant component of radiation-induced genomic alterations that were not addressed in this study. CNV analysis could provide insights into dosage-sensitive genes and broader genomic instability.

Estimating the overall burden of CNVs in irradiated versus non-irradiated samples could quantify the extent of genomic instability induced by radiation. CNVs often arise from replication stress or repair errors, particularly in cells exposed to mixed or chronic radiation.

6.4.4 Integrative analysis across all variant types

An integrative approach combining SNP, indel, SV, and CNV analyses would provide a comprehensive view of radiation-induced mutagenesis. For instance:

- SNPs and indels could be analyzed in the context of SV and CNV breakpoints to determine whether small-scale mutations cluster around larger genomic alterations.
- CNVs and SVs could be linked to mutational signatures to identify whether specific types of structural damage correlate with oxidative damage or replication stress signatures observed in SNP analyses.
- Clonal selection could be evaluated using all variant types to uncover the combined impact of small and large genomic changes on cell viability and adaptation under chronic irradiation.

6.4.5 Future directions and implications for radiation research

Incorporating these additional layers of analysis would significantly expand the scope of this work, providing a more complex understanding of how ionizing radiation drives somatic mutagenesis. While the current study emphasizes the quantitative increase in small mutations, exploring SVs and CNVs could uncover more profound genomic disruptions and provide critical insights into the interplay between radiation-induced DNA damage and repair processes. Integrating these analyses into future research would allow for a more complete characterization of the mutational landscape induced by radiation, advancing our understanding of the mechanisms and consequences of radiation exposure at the genomic level.

REFERENCES

Alexandrov L.B., Nik-Zainal S., Wedge D.C., Aparicio S.A.J.R., Behjati S., Biankin A.V., Bignell G.R., Bolli N., Borg A., Borresen-Dale, A.L. Signatures of mutational processes in human cancer. Nature, 500(7463), 415–421, 2013.

Altmeyer M., Lukas J. Guarding against collateral damage during chromatin transactions. Cell. 153(7), 1431-41, 2013.

Amundson S.A., Grace M.B., McLeland C.B., Epperly M.W., Yeager A., Zhan Q., Greenberger J.S., Fornace A.J. Human lymphocyte gene expression responses to ionizing radiation exposure. Rad Res. 155(2), 263-270, 2001.

Ansorge W.J. Next-generation DNA sequencing techniques. Nat Biotechnol. 25(4), 195-203. 2009.

Asaithamby A., Chen D.J. Mechanism of cluster DNA damage repair in response to ionizing radiation: implications in risk assessment. Mutat Res Genet Toxicol Environ Mutagen. 669(1-2), 18-26, 2009.

Behjati S., Gundem G., Wedge D.C., Roberts N.D., Tarpey P.S., Cooke S.L., Van Loo P., Kearney A., Davies H., Nik-Zainal S., et al. Mutational signatures of ionizing radiation in second malignancies. Nat Commun. 7, 12605, 2016.

Borrego S.G., Ortiz L.R., Rojas M.A. Ionizing radiation-induced DNA injury and damage detection in patients with breast cancer. Genet Mol Biol. 38(4), 420-432, 2015.

Bruce A., Alexander J., Julian L., David M., Martin R., Keith R., Peter W. Molecular biology of the cells. 6th Edition. Garland Science. 2015.

Cardis E., Vrijheid M., Blettner M., Gilbert E., Hakama M., Hill C., Howe G., Kaldor J., Muirhead C.R., Schubauer-Berigan M., et al. Risk of cancer after low doses of ionizing radiation: retrospective cohort study in 15 countries. BMJ. 331(7508), 77. 2005.

Chatterjee N., Walker G. C. Mechanisms of DNA damage, repair and mutagenesis. Environmental and Molecular Mutagenesis, 58(5), 235–263. 2017.

Delhomme T.M., Munteanu M., Buonanno M., Grilj V., Biayna J., Supek F. Proton and alpha radiation-induced mutational profiles in human cells. Sci Rep. 13, 9791, 2023.

DePristo M.A., Banks E., Poplin R., Garimella K.V., Maguire J.R., Hartl C., Philippakis A.A., del Angel G., Rivas M.A., Hanna M., et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 43(5), 491-498, 2011.

Dianov G.L., Hubscher U. Mammalian base excision repair: The forgotten archangel. Nucleic Acids Res. 40, 170-180, 2013.

Eid J., Fehr A., Gray J., Luong K., Lyle J., Otto G., Peluso P., Rank D., Baybayan P., Bettman B., et al. Real-time DNA sequencing from single polymerase molecules. Science. 323(5910), 133-138, 2009.

El-Khamisy S.F., Saifi G.M., Weinfeld M., Johansson F., Helleday T., Lupski J.R., Caldecott K.W. Defective DNA single-strand break repair in spinocerebellar ataxia with axonal neuropathy-1. Nature. 434(7029), 108-13, 2005.

Feuk L., Carson A., Scherer A. Structural variation in the human genome. Nat Rev Genet. 7(2), 85-97 2006.

Feng R.T., Weng K.L., Boo C.K. Low dose or low-dose-rate ionizing radiation induced bioeffects in animal models. J Rad Res. 58(2), 165-182, 2017.

Fousteri M., Vermeulen W., van Zeeland A.A., Mullenders L.H. Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. Mol Cell. 23, 471-482, 2006.

Friedberg E.C., Walker G.C., Wolfram S., Wood R.D., Schultz R.A., Ellenberger T. DNA repair and mutagenesis. 2nd Edition. ASM Press. 2006.

Goodwin S., McPherson J.D., McCombie W.R. Coming of age: Ten years of next-generation sequencing technologies. Nat Rev Gen. 17(6), 333-351. 2016.

Guo W., Xing Y., Luo X., Li F., Ren M., Liang Y. Reactive oxygen species: a crosslink between plant and human eukaryotic cell systems. Int J Mol Sci. 24(17), 13052, 2023.

Hang B. Formation and repair of tobacco carcinogen-derived bulky DNA adducts. J Nucleic Acids. 10.4061/709521, 2010.

Hall E.J., Giacca A.J. Radiobiology for the Radiologist 8th Edition. Wolters Kluwer. 2019.

Harper J.W., Elledge S.J. The DNA damage response: ten years after. Mol Cell. 28(5), 739-45. 2007.

Helleday T., Eshtad S., Nik-Zainal S. Mechanisms underlying mutational signatures in human cancers. Nat Rev Genet. 15(9), 585-598, 2014.

IAEA. Radiation Biology: A Handbook for Teachers and Students. Molecular and Cellular radiobiology. Training Course Series 42, 2010.

International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature. 431(7011), 931-945, 2004.

Jeggo P.A., Lobrich M. Radiation-induced DNA damage responses: their impact on human health and cancer risk. Radiat Environ Biophys. 45(2), 155-166, 2006.

Kadhim M., Salomaa S., Wright E., Hildebrandt G., Belyakov O.V., Prise K.M., Little M.P. Nontargeted effects of ionizing radiation–implications for low dose risk. Mutat Res Rev Mutat Res. 752(1), 84–98, 2013.

Korbel J.O., Urban A.E., Affourtit J.P. Paired-end mapping reveals extensive structural variation in the human genome. Science. 318(5849), 420-426, 2007.

Kumar K., Kumar S., Datta K., Fornace A.Jr., Suman S. High-LET-Radiation-induced persistent DNA damage response signaling and gastrointestinal cancer development. Curr Oncol. 30(6), 5497-5514, 2023.

Kunkel T.A. Evolving views of DNA replication (in)fidelity. Cold Spring Harb Symp Quant Biol. 74:91-01. 2009.

Kunkel T.A., Erie D.A. DNA mismatch repair. Annu Rev Biochem. 72, 681-710. 2005

Lander E.S., Linton L.M., Birren B., Nusbaum C., Zody M.C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., et al. Initial sequencing and analysis of the human genome. Nature. 409(6822), 860-921, 2001.

Lin H. S-Adenosylmethionine-dependent alkylation reactions: when are radical reactions used? Bio-org Chem. 39, 161-170, 2011.

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

Little M.P., Wakeford R., Tawn E.J., Bouffler S.D., Berrington de Gonzalez A. Risks associated with low doses and low dose rates of ionizing radiation: Why linearity may be (almost) the best we can do. Radiology. 251(1), 6-12, 2009.

Li X., Heyer W.D. Homologous recombination in DNA repair and DNA damage tolerance. Cell Res. 18, 99-113, 2008.

Lomax M.E., Folkes L.K., O'Neill P. Biological consequences of radiation-induced DNA damage: Relevance to radiotherapy. Clin Oncol. 25(10), 578-585, 2013.

Liu Q., Peng Q., Zhang B., Tan Y. X-ray cross-complementing family: the bridge linking DNA damage repair and cancer. J Transl Med, 21(602), 1–12, 2023.

MacArthur D.G., Manolio T.A., Dimmock D.P., Rehm H.L., Shendure J., Abecasis G.R., Adams D.R., Altman R.B., Antonarakis S.E., Ashley E.A., et al. Guidelines for investigating causality of sequence variants in human disease. Nature. 508(7497), 469-476, 2014.

Matsuda Y., Uchimura A., Satoh Y., Kato N., Toshishige M., Kajimura J., Hamasaki K., Yoshida K., Hayashi T., Noda A., Tanabe O. Spectra and characteristics of somatic mutations induced by ionizing radiation in hematopoietic stem cells. Proc Natl Acad Sci USA. 120(15), e2216550120, 2023

Matsumoto Y., Sharma M.K. DNA-dependent protein kinase in DNA damage response: Three decades and beyond. J Rad Cancer Res. 11, 123-134, 2020.

Matsumoto Y., Asa A.D.D.C., Modak C., Shimada M.. DNA-dependent protein kinase catalytic subunit: The sensor for DNA double-stranded breaks structurally and functionally related to Ataxia–Telangiectasia Mutated. Genes, 12, 1143, 2021.

Matsumoto Y.. Development and evolution of DNA-dependent protein kinase inhibitors toward cancer therapy. Int J Mol Sci. 23(8), 4264, 2022.

Maxam, A.M., Gilbert, W. A new method for sequencing DNA. Proc Natl Acad Sci USA. 74(2), 560-564, 1977.

Miguel G.D., Katarzyna B., Joseph M.K., Lars C.P., Thomas A.K. Structural analysis of strand misalignment during DNA synthesis by a human DNA polymerase. Cell. 124(2), 331-342, 2006.

Muir P., Li S., Lou S., Wang D., Spakowicz D.J., Salichos L., Zhang J., Weinstock G.M., Isaacs F., Rozowsky J., Gerstein M. The real cost of sequencing: scaling computation to keep pace with data generation. Genome Biol. 17(1), 53, 2016.

Ministry of the Environment. Booklet to Provide Basic Information Regarding Health Effects of Radiation. 4th Edition. https://www.env.go.jp/en/chemi/rhm/basic-info/.

Milholland B., Dong X., Zhang L., Hao Xi., Suh Y., Vijg J. Differences between germline and somatic mutation rates in humans and mice. Nat Commun 8, 15183. 2017.

Miyake T., Shimada M., Matsumoto Y., Okino A. DNA damage response after ionizing radiation exposure in skin keratinocytes derived from human-induced pluripotent stem cells. Int J Radiation Oncol Biol Phys, 105(1), 193-205. 2019.

Montgomery S.B., Goode D.L., Kvikstad E., Albers C.A., Zhang Z.D., Mu X.J., Ananda G., Howie B., Karczewski K.J., Smith K.S., et al. The origin, evolution, and functional impact of short insertion–deletion variants identified in 179 human genomes. Genome Res. 23(5), 749-761. 2013.

Moorthy B., Chu C. and Carlin D.J. Polycyclic aromatic hydrocarbons: from metabolism to lung cancer. Toxicol Sci. 145, 5-15, 2015.

Morton L.M., Karyadi D.M., Stewart C., Bogdanova T.I., Dawson E.T., Steinberg M.K., Dai J., Hartley S.W., Schonfeld S.J., Glover C.E., et al. Radiation-related genomic profile of papillary thyroid carcinoma after the Chernobyl accident. Science. 372(6543), eabg2538. 2021.

Nabel C.S., Manning S.A., Kohil R.M. The curious chemical biology of cytosine: deamination, methylation, and oxidation as modulators of genomic potential. ACS Chem Biol.7(1), 20-30, 2012.

National Institutes of Health (US); Biological Sciences Curriculum Study. NIH Curriculum Supplement Series [Internet]. Bethesda (MD): National Institutes of Health (US) Understanding Human Genetic Variation. https://www.ncbi.nlm.nih.gov/books. 2007.

Nielsen R., Paul J.S., Albrechtsen A., Song Y.S. Genotype and SNP calling from next-generation sequencing data in population genomics. Nat Rev Genet. 12(6), 443-451. 2011.

Nik-Zainal S., Davies H., Staaf, J., Ramakrishna M., Glodzik D., Zou X., Martincorena I., Alexandrov L. B., Martin S., Wedge D. C., Van Loo P., Ju Y. S., Smid M., Brinkman A. B., Butler, A., et al. The landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature, 534(7605), 47–54. 2016.

Nicolaides N.C., Papadopoulos N., Lui B., Wei Y.F., Carter K.C., Ruben S.M., Haseltine Q.A., Fleischmann R.D., Fraser C.M., et al. Mutations of two PS homologues in hereditary nonpolyposis colon cancer. Nature. 371(6492), 75-80. 1994 .

Omer H. Radiobiological effects and medical applications of non-ionizing radiation. Saudi J Biol Sci. 28(10), 5585-5592.

Patrick G.P., Chad T., Gordon B.M., Timothy A.Y. State of the art strategies for targeting the DNA damage response in cancer. Nat Rev Clin Oncol. 16. 81-104, 2019

Pfeiffer P., Goedecke W., Obe G. Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. Mutagenesis. 15(4), 289-02, 2000

Pleasance E.D., Stephens P.J., O'Meara S., McBride D.J., Meynert A., Jones D., Lin M., Beare D., Lau K.W., Greenman C., Varela I., Nik-Zainel S., et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. Nature. 463(7278), 184-90, 2010

Polgar S., Schofield P.N., Madas B.G. Datasets of in vitro clonogenic assays showing low dose hyper-radiosensitivity and induced radioresistance. Sci Data. 9(1), 555, 2022

Ramadan K., Shevelev I.V., Maga G., Hubscher U. De novo DNA synthesis by human DNA polymerase lambda, DNA polymerase mu and terminal dexoribonucleotidyl transferase. J Mol Biol. 339(2), 395-404, 2004.

Ringo J. Fundamental genetics. Cambridge University Press. 2012

Sage E., Shikazono N. Radiation-induced clustered DNA lesions: Repair and mutagenesis. Free Radic Biol Med. 107, 125-135. 2017.

Sanger F., Nicklen S., Coulson A.R. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA. 74(12), 5463-5467, 1977.

Satoh Y., Asakawa Ji., Nishimura M., et al. Characteristic of induced mutations in offspring derived from irradiated mouse spermatogonia and mature oocytes. Sci Rep. 10, 37, 2019.

Scharer O.D. Nucleotide excision repair in eukaryotes. Cold Spring Harb Perspect Biol. 5(10), 012609, 2013.

Schatz M.C., Langmead B., Salzberg S.L. Cloud computing and the DNA data race. Nat Biotech. 28(7), 691-693, 2010.

Shimada M, Tsukada K, Kagawa N, Matsumoto Y. Reprogramming and differentiation-dependent transcriptional alteration of DNA damage response and apoptosis genes in human induced pluripotent stem cells. J Radiat Res. 60(6), 719-728, 2019.

Shimada M, Tokumiya T, Miyake T, Tsukada K, Kanzaki N, Yanagihara H, Kobayashi J, Matsumoto Y. Implication of E3 ligase RAD18 in UV-induced mutagenesis in human induced pluripotent stem cells and neuronal progenitor cells. J Radiat Res. 64(2), 345-351, 2023.

Stephens P.J., Greenman C.D., Fu B., Yang F., Bignell G.R., Mudie L.J., Pleasance E.D., Lau K.W., Beare D., Stebbings L.A., et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell. 144(1), 27-40, 2011.

Suzuki K., Ojima M., Kodama S., Watanabe M. Radiation-induced DNA damage and delayed mutation in normal human skin fibroblasts. Int J Rad Biol. 79(4), 275-280, 2003.

The 1000 Genomic Project Consortium. A map of human genome variation from population-scale sequencing. Nature. 467, 1061-1073, 2010

Tubbs A., Nussenzweig A. Endogenous DNA damage as a source of genomic instability in cancer. Cancer Cell. 168(4), 644-656, 2017.

Tulchinsky T.H., Varavikova E.A., Cohen M.J. Chapter 9 Environmental and Occupational Health. New Public Health 4th Edition. Academic Press. 2023.

Uchimura A., Higuchi M., Minakuchi Y., Ohno M., Toyoda A., Fujiyama A., Miura I., Wakana S., Nishino J., Yagi T. Germline mutation rates and the long-term phenotypic effects of mutation accumulation in wild-type laboratory mice and mutator mice. Genome Res. 25(8), 1125-34. 2015.

Van Attikum H., Gasser S.M. The histone code DNA breaks: a guide to repair? Nat Rev Mol Bio. 6(10), 757-65, 2005.

Venter J.C., Adams M.D., Myers E.W., Li P.W., Mural R.J., Sutton G.G., Smith H.O., Yandell M., Evans C.A., Holt R.A., et al. The sequence of the human genome. Science. 291(5507), 1304-1351, 2001.

Wang Y., Navin N.E. Advances and applications of single-cell sequencing technologies. Mol Cell. 58(4), 598-609, 2015.

Warburton P.E., Sebra R.P. Long-Read DNA Sequencing: Recent Advances and Remaining Challenges. Annu Rev Genome Hum Genet. 20, 529-553, 2019.

Ward J. F. The complexity of DNA damage: relevance to biological consequences. Int J Radiat Biol. 66(5), 427–432, 1994.

Xie A., Hartlerode A., Stucki M., Odate M., Puget N., Kwok A., Nagaraju G., Yan C., Alt F.W., Chen J., Jackson S.P., Scully R. Distinct roles of chromatin-associated proteins MDC1 and 53BP1 in mammalian double-strand break repair. Mol Cell. 28(6), 1045-1057, 2007.

Yeager M., Machiela M.J., Nasrin S., Perez M., Setiawan V.W., Schumacher F.R., Wentzensen N., Han S.S., Litchfield K., Freedman N.D., et al. Radiation exposure and the landscape of copy number variations in cancer risk. Nat Genet. 53(12), 1689-1698, 2021.

Yokoi M., Masutani C., Maekawa T., Sugasawa K., Ohkuma Y., Hanaika F. The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA. J Biol Chem. 275, 9870-9875, 2000.

Youk J., Kim H., Lee S., Park S., Choi Y., Kim J., Seo J., Kang Y., Song J., Lim D. Complex structural variations, including balanced inversions and chromothripsis, induced by ionizing radiation in somatic cells. Genome Res. 34(2), 215-225, 2024.

Zhang F., Fan Q., Ren K., Andreassen PR. PALB2 functionally connects the breast cancer susceptibility proteins BRCA1 and BRCA2. Mol Cancer Res. 7(7), 1110-1118, 2009.

Zhang F., Gu W., Hurles M.E., Lupski J.R. Copy number variation in human health, disease, and evolution. Annu Rev Genome Hum Genet. 10, 451-481, 2009